Tourism Management Perspectives 39 (2021) 100860

Contents lists available at ScienceDirect

Tourism Management Perspectives

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

Comparing differences in the spatiotemporal patterns between resident tourists and non-resident tourists using hotel check-in registers

Yuquan Xu [a], Xiaobin Ran [a], Yuewen Liu [a,*], Wei Huang [a,b]

[a] School of Management, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an 710049, China
[b] College of Business, Southern University of Science and Technology, No. 1088 Xueyuan Avenue, Nanshan District, Shenzhen 518055, China

ARTICLE INFO

Keywords:
Big data
Spatiotemporal patterns
Resident tourists and non-resident tourists
Multiple city travel
Hotel check-in registers

ABSTRACT

Previous research studied the spatiotemporal patterns in different visitor segments but lacks evidence of the segmentation of resident tourists and non-resident tourists in multi-city travel. To fill this gap, this study conducts a big data study using hotel check-in registers. The exploratory data analysis visualizes the spatiotemporal patterns and the differences between resident tourists and non-resident tourists. Then, the spatiotemporal patterns are measured by the length of stay and the number of visited cities. The regression shows that both the length of stay and the number of visited cities of non-resident tourists are higher than those of resident tourists. Moreover, non-resident tourists reduce their length of stay and their number of visited cities more than resident tourists on three-day holidays, while they increase their number of visited cities less than resident tourists on seven-day holidays. This study has significant implications for understanding spatiotemporal patterns and visitors' segmentations.

1. Introduction

Understanding the spatiotemporal patterns of visitors is an important and classic issue in tourism research (Li, Xu, Tang, Wang, & Li, 2018; Zhao, Lu, Liu, Lin, & An, 2018). It is the foundation of tourism practices, such as tourism planning, tourism marketing, destination management and travel recommendations (Park, Xu, Jiang, Chen, & Huang, 2020; Salas-Olmedo, Moya-Gómez, García-Palomares, & Gutiérrez, 2018). Usually, spatiotemporal patterns vary with the visitor segmentations, such as age, gender and nationality (Batista e Silva et al., 2018; Esiyok, Kurtulmuşoğlu, & Özdemir, 2017; Lee, 2015; Losada, Alén, Nicolau, & Domínguez, 2017; Raun, Ahas, & Tiru, 2016; Yang, Lin, & Han, 2010). One of the important segmentations is resident versus non-resident tourist (Hoogendoorn & Hammett, 2020; Jeuring & Haartsen, 2016). Resident tourists are the tourists who travel inside their usual residential area while non-resident tourists are the tourists who travel outside their usual residential area. The range of the residential area could be a city, a province/state or a country in existing research (Hasnat & Hasan, 2018; Li, Zhou, & Wang, 2018). Since a tourism location attracts both resident tourists and non-resident tourists, it may cause some tourism conflicts, such as crowd traffic and hotel room shortages. To avoid the conflicts, resident tourists and non-resident tourists show different spatiotemporal patterns, such as changing their travel time and destinations. Thus, investigating the differences in the spatiotemporal patterns between resident tourists and non-resident tourists could help tourism managers and administrators design tourism products, develop local economies and optimize regional planning (Liu, Wang, & Ye, 2018; Su, Spierings, Dijst, & Tong, 2019).

Current studies on these spatiotemporal patterns are mainly concerned with temporal patterns and spatial patterns, such as travel time, spatial distribution and destination identification (Ferrante, Lo Magno, & De Cantis, 2018; Li, Zhou, & Wang, 2018; Mou et al., 2020; Salas-Olmedo et al., 2018; Tang & Li, 2016). Some of these studies discussed the differences between resident visitors and non-resident visitors (Hoogendoorn & Hammett, 2020; Singh & Krakover, 2015). In detail, they found that the spatial patterns of non-resident visitors are more concentrated in central urban areas whereas resident visitors visit suburban areas. The temporal patterns of non-resident visitors vary considerably whereas those of resident visitors are relatively stable (Khan, Wan, & Yu, 2020; Liu et al., 2018; Su et al., 2019; Wu et al., 2020). All of these studies mainly focused on the differences within a city (single city travel). However, multiple city travel is a common type of travel and a main part of the tourism market (Önder, 2017). Far less is known about the difference in the spatiotemporal patterns between

* Corresponding author.
E-mail addresses: xuyuquan@stu.xjtu.edu.cn (Y. Xu), liuyuewen@mail.xjtu.edu.cn (Y. Liu).

https://doi.org/10.1016/j.tmp.2021.100860
Received 30 September 2020; Received in revised form 22 June 2021; Accepted 19 July 2021
Available online 28 July 2021

resident tourists and non-resident tourists in multiple city travel. First, the findings on single city travel in the existing literature cannot be extended to multiple city travel due to the differences between these two types of travel. Specifically, the increasing costs of intercity travel, such as the costs of transportation and accommodations, would lead to different temporal patterns including the length of stay and the daily and monthly trends of visitor arrivals (Santos, Ramos, & Rey-Maquieira, 2014). The spatial patterns in single city travel are impacted by urban facilities and urban functional zones, but the spatial patterns in multiple city travel may be impacted by the cities' location, climate and other factors. Second, the data in the existing studies are mainly traditional survey data and big data such as online photo, GPS and mobile roaming data (Batista et al., 2020; Li, Xu, et al., 2018; Xue & Zhang, 2020). Most of these studies need to perform tourist identification and demographic information inference. Thus, it is difficult to distinguish resident tourists and non-resident tourists. Hotel check-in registers could solve such issues, but limited extant literature explores the spatiotemporal patterns of tourists using hotel check-in registers. Third, a few studies discussed the spatiotemporal patterns in multiple city travel, but they mostly focused on non-resident tourists' patterns while paid less attention to resident tourists' patterns (Hasnat & Hasan, 2018; Sinclair, Mayer, Woltering, & Ghermandi, 2020; Su, Wan, Hu, & Cai, 2016). Overall, few studies focused on the spatiotemporal patterns of resident and non-resident tourists, especially in multiple city travel.

Given these gaps in the current literature, the objective of this study is to distinguish the main differences in the spatiotemporal travel patterns between resident tourists and non-resident tourists via a big data study. Unlike the existing research, this study uses hotel check-in registers to observe tourists' spatiotemporal patterns. Hotel check-in registers are generated and stored in hotel information systems and have broad coverage among all tourist groups and hotels in different cities, which could fill the abovementioned gaps in the existing studies. Specifically, collaborating with an IT company providing information systems for hotels in Yunnan, China, this study uses a sample that contains more than 6.29 million hotel check-in records of 670 thousand individuals (including approximately 170 thousand resident tourists and 500 thousand non-resident tourists) from January 1, 2015 to December 31, 2018, covering more than 45 thousand hotels distributed in 16 cities in Yunnan Province. This study first summarizes the temporal patterns by analyzing the monthly and daily trends of check-in records and the spatial patterns by describing the spatial distribution and the seasonal tendency of spatial patterns. Two indicators of spatiotemporal patterns are adopted, which are the Length of Stay (LS) that reflects people's temporal patterns and the Number of Visited Cities (NVC) that reflects people's spatial patterns. Then we run regressions to study the correlations between tourist demographic variables and the two spatiotemporal pattern indicators. Last, this study further analyzes the differences between resident tourists and non-resident tourists during public holidays in the regressions. The results show that (1) both the LS and the NVC of non-resident tourists are higher than those of resident tourists. Specifically, the LS of non-resident tourists is 30.70% higher than that of resident tourists, and the NVC of non-resident tourists is 24.91% higher than that of resident tourists; (2) non-resident tourists reduce their LS and NVC more than resident tourists on three-day holidays, while they increase their NVC less than resident tourists on seven-day holidays.

The contributions of this study are as follows. First, to the best of our knowledge, this study is the first attempt using a large dataset of hotel check-in registers as the primary data source to study the spatiotemporal patterns of tourists, and thus provides a completely new perspective of travel study. Compared to traditional data in tourism research, hotel check-in registers could avoid some data preprocessing issues such as tourist identification and demographic information inference (Reif & Schmücker, 2020). Second, using demographic information, this study provides evidence of the spatiotemporal patterns of resident tourists and non-resident tourists. In the existing literature, scholars pay less attention to the spatiotemporal patterns of resident tourists while increasingly more resident tourists travel their residential areas (Hoogendoorn & Hammett, 2020). Finally, this study provides evidence of the spatiotemporal patterns of multiple city travel using a large hotel check-in register dataset from tens of thousands of hotels. The findings in this study are meaningful for stakeholders in the tourism industry to explore the value of hotel check-in registers and to understand the segmentation of tourists and the spatiotemporal patterns in multiple city travel. The related fields, including tourism planning, tourism marketing and tourism destination management, could benefit from the findings of this study.

The remainder of this paper is structured as follows. Section 2 reviews the existing literature on spatiotemporal patterns. Section 3 describes the data and the method. Section 4 presents the results; and Section 5 summarizes the main conclusions, discussions and future research directions.

2. Literature review

2.1. The spatiotemporal patterns of visitors

Regarding temporal patterns, many studies have mainly discussed the seasonality of tourism, the arrival time, the number of arrivals and the length of stay (Duro, 2016; Duro & Turrión-Prats, 2019). For example, Connell, Page, and Meyer (2015) found that the role of attraction-based events could counter the effects of the decrease in seasonal demand, and important events could maintain tourists' interest through the off-peak season. Turrión-Prats and Duro (2018) found that tourism income are inversely related to seasonality, while the appreciation of the exchange rate increases the seasonality in Spain. Liu, Zhang, Zhang, Sun, and Qiu (2019) found that the number of arrivals that occur 2–4 days later could be accurately predicted using web search data. Among the factors, the length of stay is a critical factor for understanding temporal patterns and is widely discussed in the tourism literature (Alén, Nicolau, Losada, & Domínguez, 2014; Santos et al., 2014; Wang, Fong, Law, & Fang, 2017). Thrane and Farstad (2011) found a positive but diminishing relationship between the length of stay and tourism expenditures. Zhao et al. (2018) found that the length of stay varies with the arrival time on public holidays.

Regarding spatial patterns, existing studies focus on the topics of the identification of popular areas and travel routes, the seasonal tendency of spatial patterns and other items (Guedes & Jiménez, 2015; Sun, Fan, Helbich, & Zipf, 2013; Zhong, Sun, & Law, 2019). For instance, Zhong et al. (2019) identified the top attractions and the most popular travel routes in Tibet using mobile positioning data. Sun et al. (2013) studied the seasonal tendency of the spatial pattern of tourist accommodations in Vienna and found that the distribution of hotspots is most dispersed in spring and most concentrated in autumn. Fernández-Morales, Cisneros-Martínez, and McCabe (2016) found that the international demand for UK destinations was dominated by tourists from European countries. Specifically, the closer to the destinations the countries of origin were, the less seasonally concentrated the tourists. Duro and Turrión-Prats (2019) studied the seasonality of European countries and found that a country's location affects its seasonal tendency of spatial patterns: the higher the latitude is, the greater the seasonality. They suggested that geographical location and the income of the major markets of origin are significant predictor variables for the number of arrivals. Ferrante et al. (2018) demonstrated a strong connection between seasonal patterns and spatial distribution.

2.2. The spatiotemporal patterns in different population segmentations

Many studies have discussed the spatiotemporal patterns of visitors among various segmentations, such as demographic characteristics and travel party size (Chen & Shoemaker, 2014; Esiyok et al., 2017; Losada et al., 2017; Rodríguez, Martínez-Roget, & González-Murias, 2018; Santos et al., 2014; Thrane, 2016). For example, Carr (2002) found that tourists between 16 and 24 years old preferred hedonic travel such as beach-oriented travel. Barros and Machado (2010) found that the length of stay is largely explained by the sociodemographic profile of the tourist and moderated by the perceived characteristics of the destination. Alegre and Pou (2006) found that the over 60 years old group stays more days than younger age groups (under 30 years old), and hotel prices and daily costs are negatively related to the length of stay. Martínez-Garcia and Raya (2008) found that travel time, country of origin, age, occupation, the type of destination and travel service prices (such as accommodation prices) are related to the length of stay. Travelers decrease their length of stay when their income is low, their time is short, and the related travel service prices are high. Zhao et al. (2018) used mobile positioning data to present different tourist movement patterns across different travel party sizes in terms of spatial and temporal characteristics. Kellner and Egger (2016) found that tourists' movement patterns vary among nationalities in Salzburg City, Austria. In detail, Asian visitors tend to spend more time at Mozart's residence and in the river promenade, while American tourists spend most of their time at the Hohensalzburg Fortress.

One of the important segmentations is the segmentation of resident tourists and non-resident tourists. On the one hand, non-resident tourists spend on food, transportation and accommodation in tourism locations (Hwang, Lee, & Park, 2012; Marrocu, Paci, & Zara, 2015), and thus can benefit the local economy (Antonakakis, Dragouni, & Filis, 2015; Mayer, Müller, Woltering, Arnegger, & Job, 2010; Ridderstaat, Croes, & Nijkamp, 2014; Wan & Li, 2013). For example, Seetanah (2011) found that tourism significantly contributes to the growth of island economy regardless of whether the context is developing or developed countries. On the other hand, the influx of non-residents has great impacts such as crowd traffic and environmental issues on the lives of residents (Al Haija, 2011; Shen, Li, Luo, & Chau, 2017; Sheng & Tsui, 2010; Tsaur, Yen, & Teng, 2018). For example, Wan and Li (2013) found that tourism industry booms could lead to negative economic consequences, such as a serious shortage of human resources and great pressure on the supply of available land. Therefore, understanding the segmentation of resident tourists and non-resident tourists is beneficial for tourism management. A few studies have discussed the spatiotemporal patterns of resident tourists and non-resident tourists. For example, Ferrante et al. (2018) studied the similarities and differences of seasonal patterns between resident groups and non-resident groups. Hasnat and Hasan (2018) identified popular cities/areas for non-resident tourists and compared the difference in the travel time between resident tourists and non-resident tourists. In summary, existing studies related to resident tourists and non-resident tourists are insufficient, and there is little evidence of the spatiotemporal patterns in multiple city travel.

2.3. The spatiotemporal patterns at different geographical levels

The analyses of spatiotemporal patterns are usually discussed at some geographical levels (related to the place of residence), including the country level, province level and city level. At the country level, foreign tourists show different spatiotemporal preferences. For instance, foreign tourists in China are more interested in China's cultural attractions than natural attractions (Yang et al., 2010). East Asian and Oceanian visitors exhibit more preferences towards the eastern coastal areas of China while European and North American visitors show more interest in the western and northern areas of China (Su et al., 2016). In Estonia, Germans are the most summer-oriented visitors among all the tourists, while Russians are the most equally distributed visitors throughout the year (Raun et al., 2016). Batista e Silva et al. (2018) found that countries in coastal areas and islands are popular year-round and that their peaks are significant in summer months, compared to other European countries. The countries in the Alpine region display high tourist densities in both summer and winter but are comparatively less dense in spring and fall. At the province level, comparing to non-resident tourists, resident tourists may travel earlier in holidays, and travel nearer their residential areas. For example, Hasnat and Hasan (2018) used location-based data from Twitter to analyze the travel behavior of tourists. Regarding spatial patterns, they showed Florida's most popular tourism attractions and some emerging tourism attractions for tourists. They also found that tourists in Florida start their travel after 12 a.m. while residents start from 11 a.m. Marrocu and Paci (2013) found that the number of arrivals in a province shows a spatial dependency induced by neighboring provinces, including travel from and to neighboring provinces. Specifically, long distance and high population density have negative impacts on the number of arrivals, while income, accessibility and attractions are crucial determinants of the number of arrivals. At the city level, the spatiotemporal patterns of tourist are influenced by urban form and urban functional zones. For instance, Li, Zhou, and Wang (2018) found that tourists' destinations are clustered around the city center while residents' destinations distribute around the cities and some attractions along the main transportation corridors.

The existing research has studied spatiotemporal patterns within a single destination such as a city, province or country while few studies have paid attention to multiple city travel. Notably, multiple city travel, as a type of multiple-destination travel, is a common travel mode and an important part of the tourism market (Parroco, Vaccina, De Cantis, & Ferrante, 2012).

3. Data and method

3.1. Hotel check-in registers

When guests check in at a hotel, the reception staff usually logs the guest ID card information into the hotel information system. When visitors check in at different hotels during a trip, visitors' travel trajectories can be extracted. Indeed, existing studies have used various data sources, including geotagged check-in platforms such as Twitter, Weibo and Flickr (Barchiesi, Moat, Alis, Bishop, & Preis, 2015; Li, Zhou, & Wang, 2018), mobile phones (Zhao et al., 2018), zenith images (Donaire, Galí, & Gulisova, 2020) and GPS devices (De Cantis, Ferrante, Kahani, & Shoval, 2016; East, Osborne, Kemp, & Woodfine, 2017), to track visitors' movements. However, to the best of our knowledge, there is little literature that studies visitors' movements using hotel check-in registers, partially because of the difficulty of collecting all check-in records distributed in various hotels. Compared to the above-mentioned data, hotel check-in registers have the following features: (1) clear tourist identification, (2) accurate demographic information, and (3) exact dates and locations of accommodations.

In this study, the dataset is hotel check-in registers collected from a hotel information systems company in Yunnan Province. The guests' IDs are encrypted to protect guests' privacy. The company provides us a sample, which is completely random sampling based on tourist IDs to ensure the completed records of each tourist. After removing 33 inconsistent records (less than 0.01% of all records), we obtained a final dataset, which contains 670 thousand individuals' completed records from January 1, 2015 to December 31, 2018, a total of 6.29 million records.

Table 1 presents examples of hotel check-in registers. The first row shows that guest "e8c6b4" is male, and his age is between 30 and 39 (to protect privacy, we do not know the precise age). He is a non-resident tourist (according to his ID card information). He checked three times on "October 1, 2018", "October 3, 2018" and "December 21, 2018". The corresponding hotels belong to cities "1201", "1207" and "1201".

Fig. 1 depicts the demographic information of the sample. As shown in Fig. 1, the proportion of visitors is mainly distributed in young adult groups, such as the (20, 29) and (30, 39) age groups. Moreover, the proportions of resident tourists in the (0, 19) and (20, 29) age groups are 5% and 15% higher than those of non-resident tourists, respectively, while the proportions of resident tourists in the (30, 39), (40, 49), (50, 59) and (60, +) age groups are 1% to 7% less than those of non-resident

Table 1
Examples of hotel check-in registers.

ID	Gender	Age group	Type	Date	City code
e8c68f	Male	(30, 39)	Non-resident tourist	October 1, 2018	1201
e8c68f	Male	(30, 39)	Non-resident tourist	October 3, 2018	1207
e8c68f	Male	(30, 39)	Non-resident tourist	December 21, 2018	1201
69b8c4	Female	(20, 29)	Non-resident tourist	May 3, 2018	1207
69b8c4	Female	(20, 29)	Non-resident tourist	May 7, 2018	1129
69b8c4	Female	(20, 29)	Non-resident tourist	December 17, 2018	2107
1a14s1	Female	(40, 49)	Resident tourist	June 10, 2018	4501
1a14s1	Female	(40, 49)	Resident tourist	October 16, 2018	6629
1a14s1	Female	(40, 49)	Resident tourist	October 27, 2018	6607

tourists. Simply put, the average age of non-resident tourists is older than that of resident tourists. We find that the proportion of males is greater than the proportion of females. More specifically, the male-female ratio of non-resident tourists (approximately 1.88) is 0.40 larger than that of resident tourists (approximately 1.48).

3.2. Study area and method

Yunnan Province, located in the southwestern area of China, is one of the main contributors to China's tourism market and one of the most ethnically diverse provinces of China (National Bureau of Statistics of China, 2019). It is famous for colorful natural landscapes and cultural heritage sites such as the Three Pagodas, the Chongsheng Temple, Shangri-La and the Stone Forest. The rich natural landscapes and national culture bring enormous tourism revenues for Yunnan. In 2018, the tourism earnings from domestic tourism in China were approximately 7.26 trillion dollars while the domestic tourism revenues in Yunnan Province were 1.34 trillion dollars (approximately 16.96% of the total tourism revenues of the country) (Yunnan Statistical Bureau, 2019). As shown in Fig. 2, Yunnan Province consists of 16 cities, including Kunming, Qujing, Dali, Lijiang, etc. The term "city" discussed in this paper refers to the entire administrative area of each city. In order to show an intuitive picture, we color the cities based on their tourism revenues according to the Yearbook (Yunnan Statistical Bureau, 2019). From Fig. 2, we find that Kunming, Lijiang, Dali, Xishuangbanna and Honghe are major cities contributing to tourism revenues.

We also examine the relationship between tourism revenues and Gross Domestic Product (GDP) of each city. In Fig. 3, each dot represents a city, the line is a fitting line and the shadow is the 95% confidence interval. The fitting line shows that the tourism revenues and GDP are positively correlated. The cities above the line have relatively high tourism revenues, such as Lijiang, Xishuangbanna and Dali, while the cities below the line have relatively high GDP, such as Qujing and Yuxi. Some cities have both high tourism revenues and GDP, such as Kunming.

To process the hotel check-in registers, we use SQL Server for data storage and R for visualization and regression analyses. First, we show the differences between resident tourists and non-resident tourists in their temporal patterns and spatial patterns through descriptive visualization analysis. Then, we calculate two variables to measure the spatiotemporal patterns, namely, the LS reflecting the temporal pattern and the NVC reflecting the spatial pattern. The LS and the NVC are calculated using individual check-in records during a trip. Following the work of Raun et al. (2016), we define a trip as the sequence of check-in records where the interval of continuous check-in records is within 14 days.

We calculate the summary statistics of the dependent variables in Table 2. As shown in Table 2, regarding the average LS, non-resident tourists (3.060) stay 0.767 more days than resident tourists (2.293). Similarly, we find that non-resident tourists (1.595) travel to 0.480 more cities than resident tourists (1.115). Overall, non-resident tourists stay longer and travel to more cities than resident tourists.

4. Results

4.1. Temporal patterns from the perspective of monthly and daily trends

We analyze the temporal patterns from the monthly trends and daily trends of check-in records. Regarding the monthly trends, we find that the trends in Yunnan Province are seasonal. The highest proportion of check-in records usually appears in summer months such as August, and the lowest proportion appears in winter months such as February (Fig. 4a). There are several extremely low points, such as December 2016, March 2017 and April 2017. The reason for the declines in March 2017 and April 2017 is the negative impact of a vicious tourism attack (Li & Yang, 2017). The reason for the decline in December 2016 could be the data quality since there were no obvious events that impact tourism

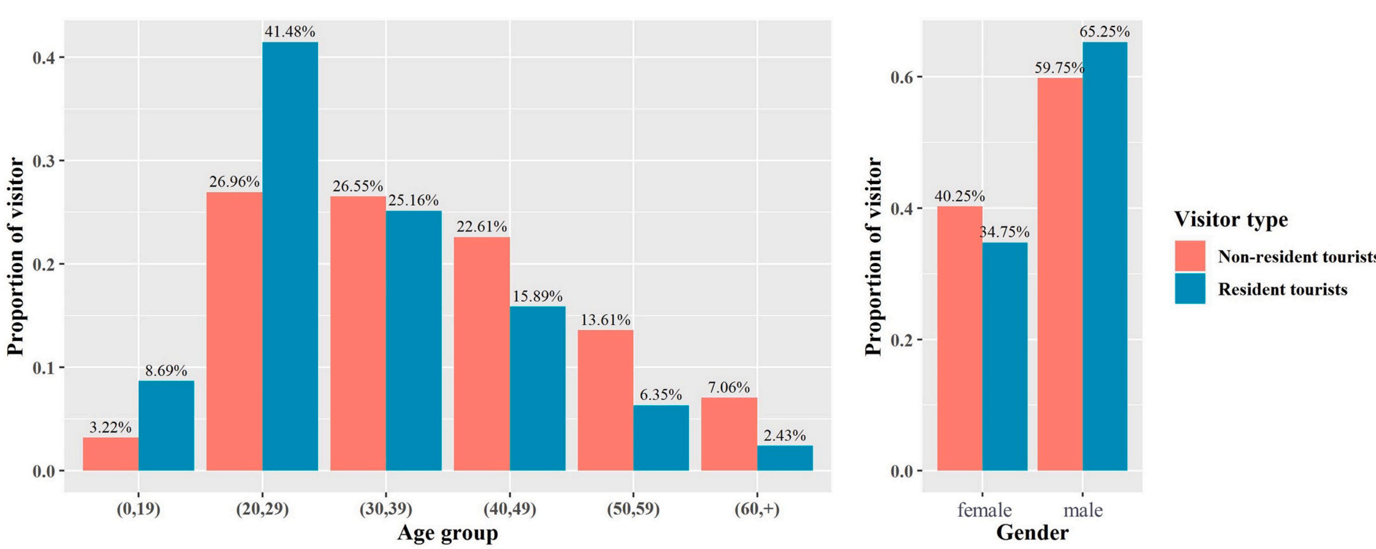

Fig. 1. The demographic information of the sample.

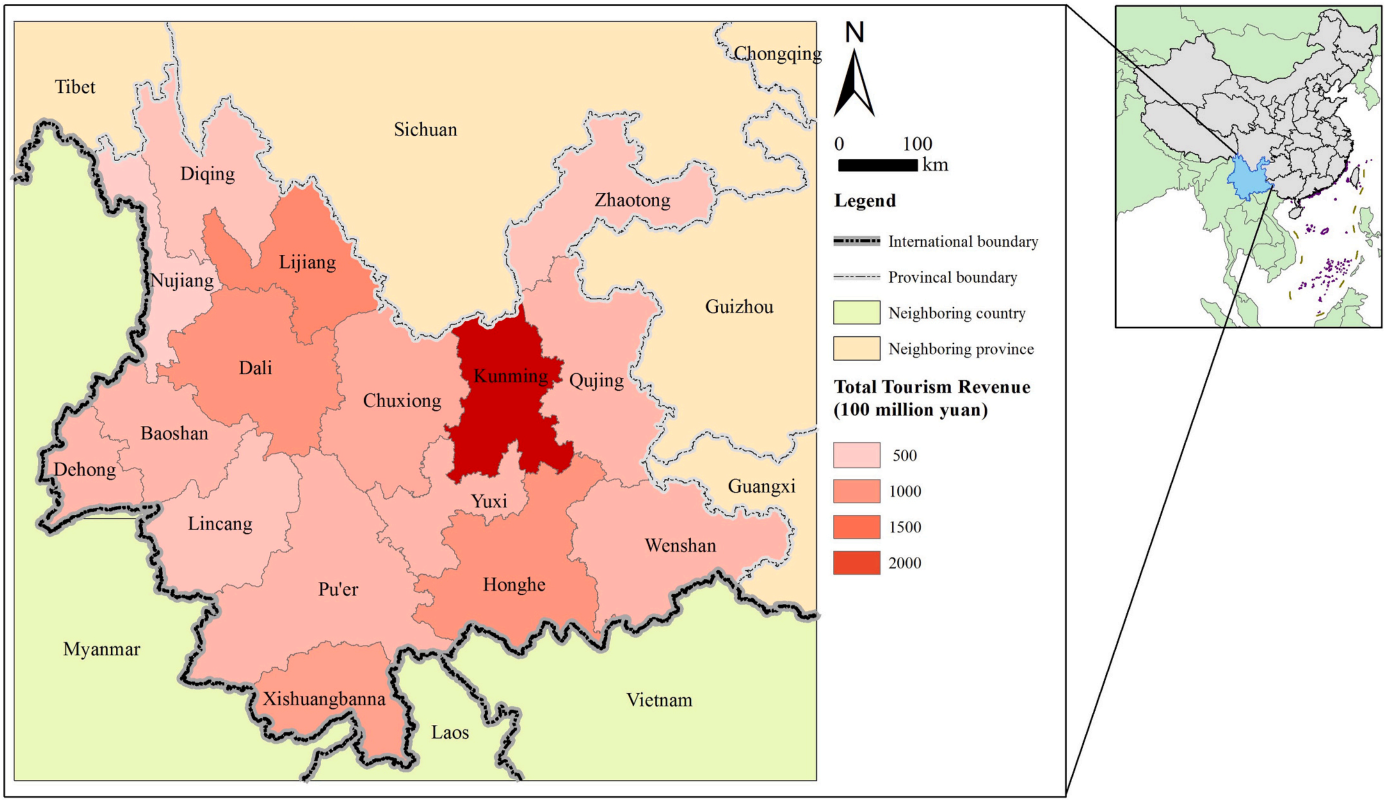

Fig. 2. The locations and tourism revenues of cities in Yunnan Province.

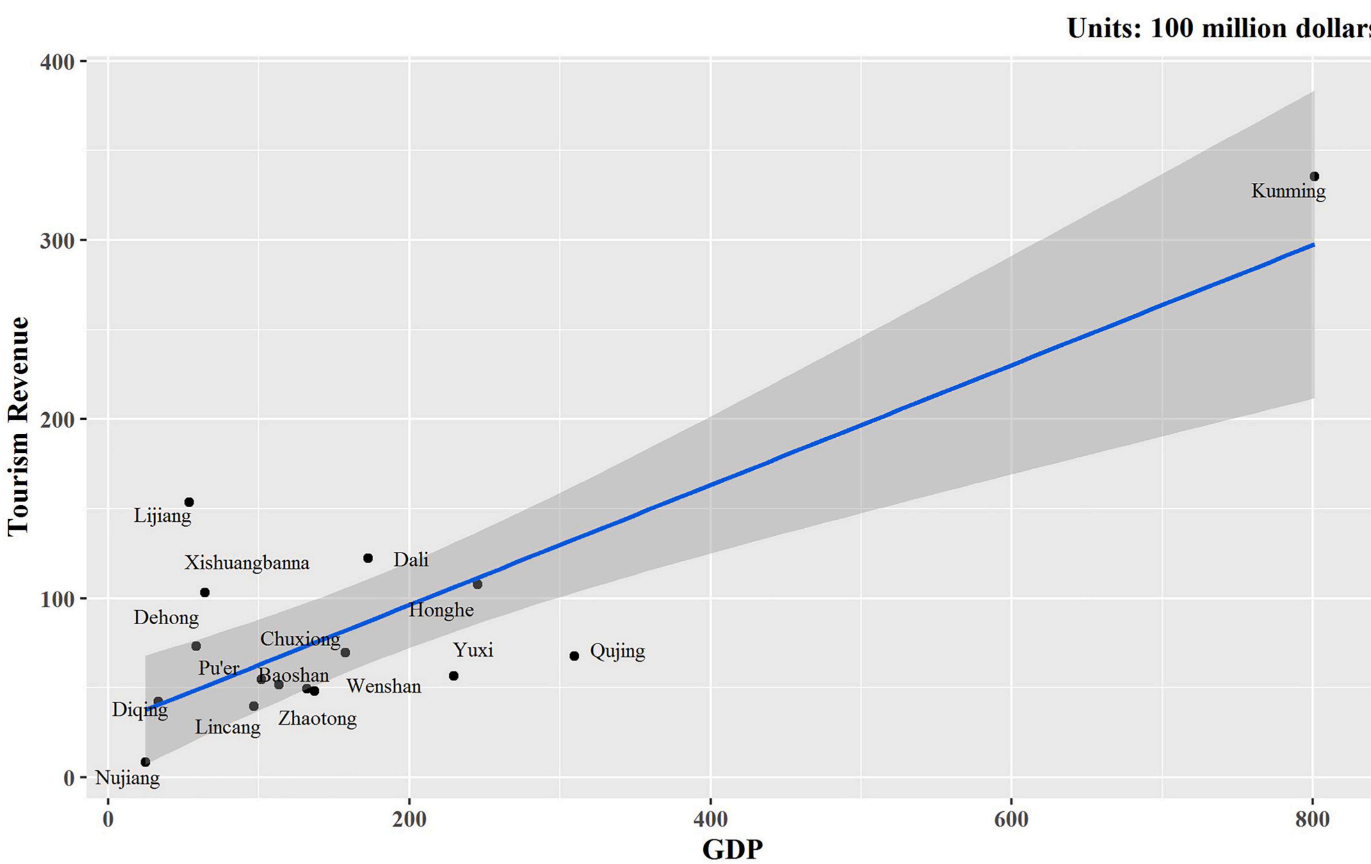

Fig. 3. The GDP and tourism revenues of the 16 cities in Yunnan Province.

Table 2
The summary statistics of the LS and NVC of tourists.

	Resident tourists		Non-resident tourists	
Statistic	LS	NVC	LS	NVC
Mean	2.293	1.115	3.060	1.595
St. Dev.	2.577	0.379	2.657	0.931
Min	1	1	1	1
Max	14	9	14	10

in December 2016, and we will show the reason in the latter analysis of the daily trends.

Fig. 4b depicts the differences between resident tourists and non-resident tourists. We find that the monthly fluctuations in the check-in records of non-resident tourists are greater than those of resident tourists. The standard deviation of the check-in records of non-resident tourists is 1.54% while that of resident tourists is 0.90%. The figure also shows that the portion of resident tourists do not have a notable lowest month (excluding anomalies), while the portion of non-resident tourists have a notable lowest month (February).

The seasonal fluctuations of the monthly trends may be caused by various reasons, such as the climate, public holidays and special events. To explore the possible reasons, we analyze the temporal patterns from the perspective of daily trends. We find that the daily trends of check-in records fluctuate around a horizontal level (approximately 0.25%) (Fig. 5a). To highlight the impacts of holidays and other events, we color the date types including workdays, weekends, and public holidays (three-day holiday and seven-day holiday). The orange areas of the background show the peak of summer holidays (July and August). We also find spikes for public holidays. Besides, there are small spikes on weekends, and these spikes are smaller than the spikes on public holidays. The highest percentage of check-in records is approximately 0.47%, which is 1.7 times the mean value (0.27%). The proportion of check-in records on several days is near zero. The reason could be that network failure and system failure cause missing records and lead to the proportion of check-in records to be close to zero on several days. This is also the reason of the abnormal points in the monthly trends (Fig. 4) during December 2016, March 2017 and April 2017.

Fig. 5b and Fig. 5c depict the daily trends of resident tourists and non-resident tourists, respectively. We find that non-resident tourists show more significant fluctuations than resident tourists (the standard deviation of the check-in records of resident tourists is 0.0608% while that of non-resident tourists is 0.0670%). The spikes of the check-in records of non-resident tourists occur during seven-day holidays and summer while the spikes of the check-in records of residents occur during seven-day holidays, three-day holidays and weekends.

4.2. Spatial pattern from the perspective of the distribution

We visualize the spatial pattern in Fig. 6 based on the proportion of check-in records of each city in Yunnan Province. Each city is colored based on its quartile of the proportion of check-in records, which results in four levels with a red gradient. As shown in Fig. 6a, the most visited cities have either relatively high tourism revenues or relatively high GDP, such as Kunming (32.37%), Lijiang (8.21%), Dali (8.19%) and Qujing (7.87%). For resident tourists, the most visited cities have relatively high GDP, such as Kunming (30.92%) and Qujing (10.11%). For non-resident tourists, the most visited cities have relatively high tourism revenues, such as Kunming (34.58%), Lijiang (14.44%) and Dali (10.29%). The difference of the spatial pattern shows that resident tourists are more likely to travel in the cities with relatively high GDP, while non-resident tourists are more likely to travel in the cities with relatively high tourism revenues (i.e., tourism cities).

Considering that spatial patterns may change with time (Ferrante et al., 2018), to obtain deep insights into the spatial patterns, we illustrate the seasonal tendency of the spatial patterns (Fig. 7) following the work of Sun et al. (2013). Fig. 7 depicts the change in the check-in records in different cities, where the horizontal axis represents the month, and the vertical axis represents the city. First, we find that the seasonal tendency in each city is different (Fig. 7a). For non-resident tourists (Fig. 7c), some cities such as Lijiang and Diqing, reach their peak season during the summer months (July to August), and reach their off season during the winter months (December to January). In contrast, some other cities such as Xishuangbanna and Baoshan, reach their peak season in winter months, and reach their off season in summer months. Resident tourists have contrary seasonal tendencies in some cities (Fig. 7b). For example, the peak months of Xishuangbanna and Baoshan are in summer months such as August. Second, the change in the spatial distribution of non-resident tourists is more regular than that of resident tourists. We measure the seasonal tendency of the spatial patterns through entropy. The average entropy of the proportion of check-in records of non-resident tourists is 2.46 while that of resident tourists is 2.48. This means that the seasonal tendency of the spatial patterns of non-resident tourists is different from that of resident tourists.

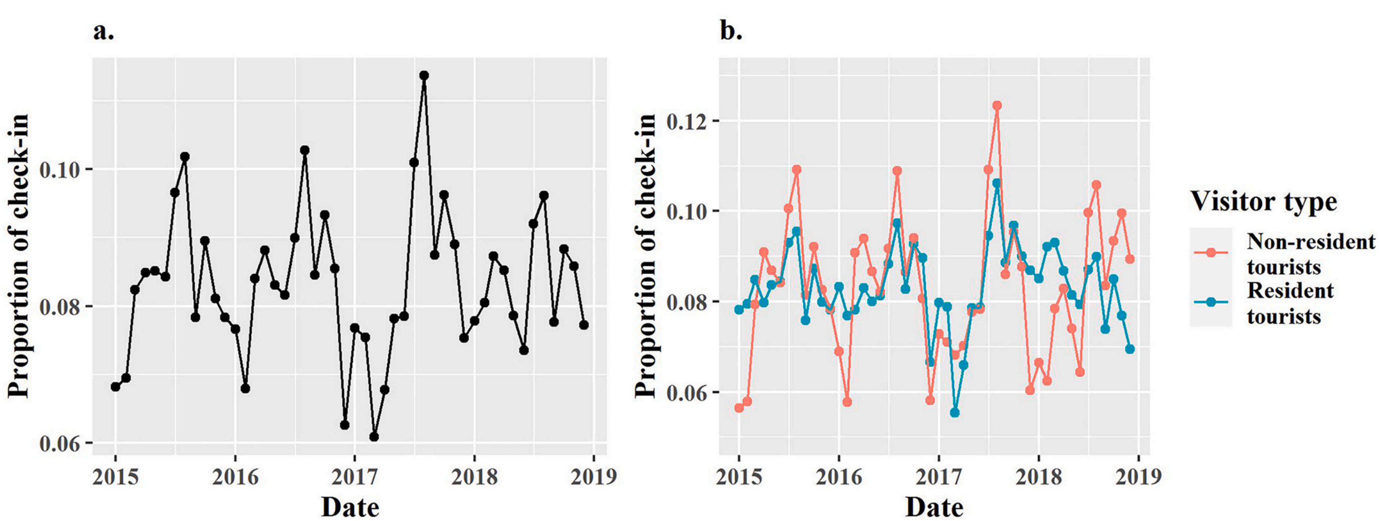

Fig. 4. Monthly trends of the check-in records in Yunnan Province. Panel a shows the trend of all visitors, and Panel b shows the trends of resident tourists and non-resident tourists.

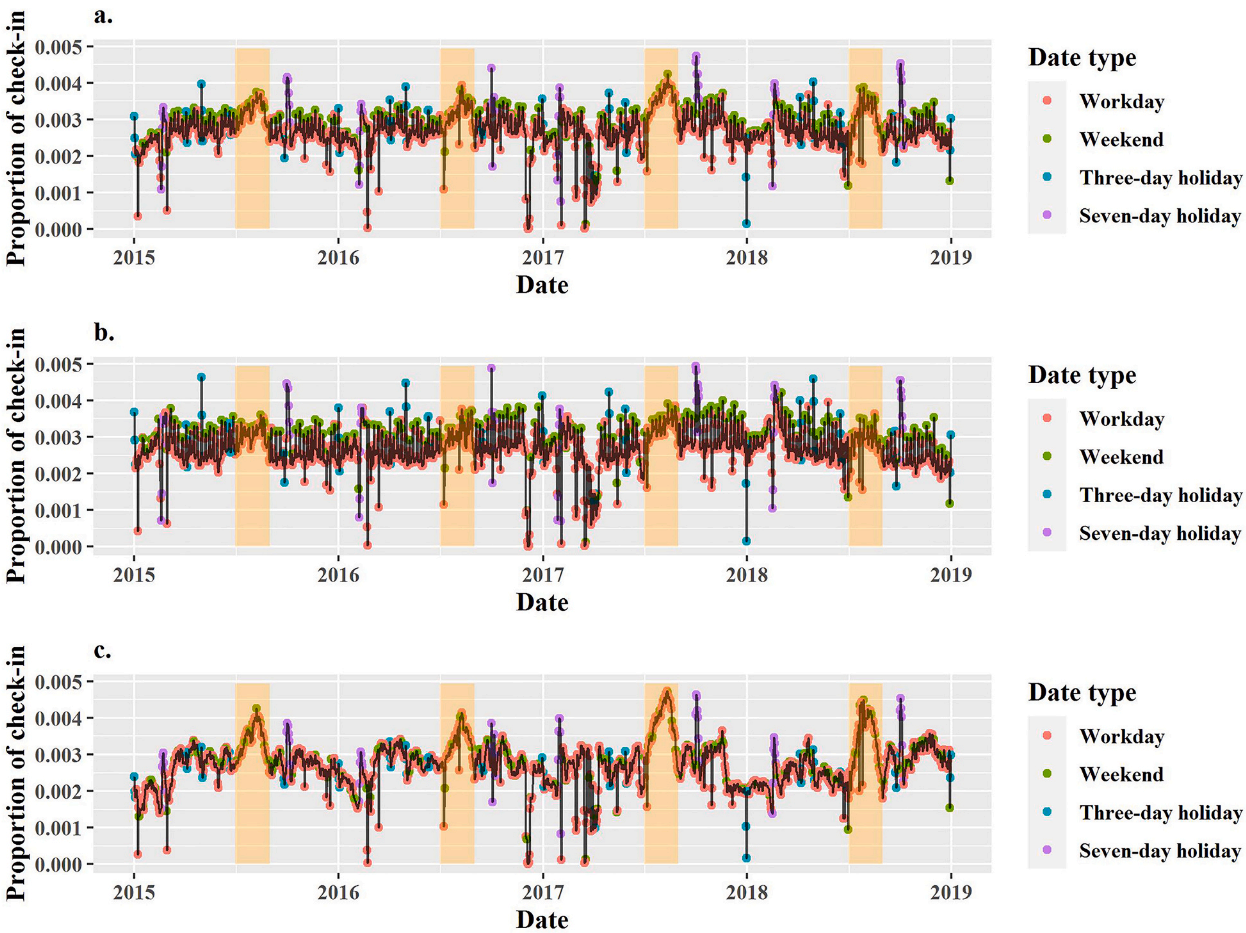

Fig. 5. Daily trends of check-in records in Yunnan Province. Panel a shows the trends of all visitors, Panel b shows the trends of resident tourists, and Panel c shows the trends of non-resident tourists.

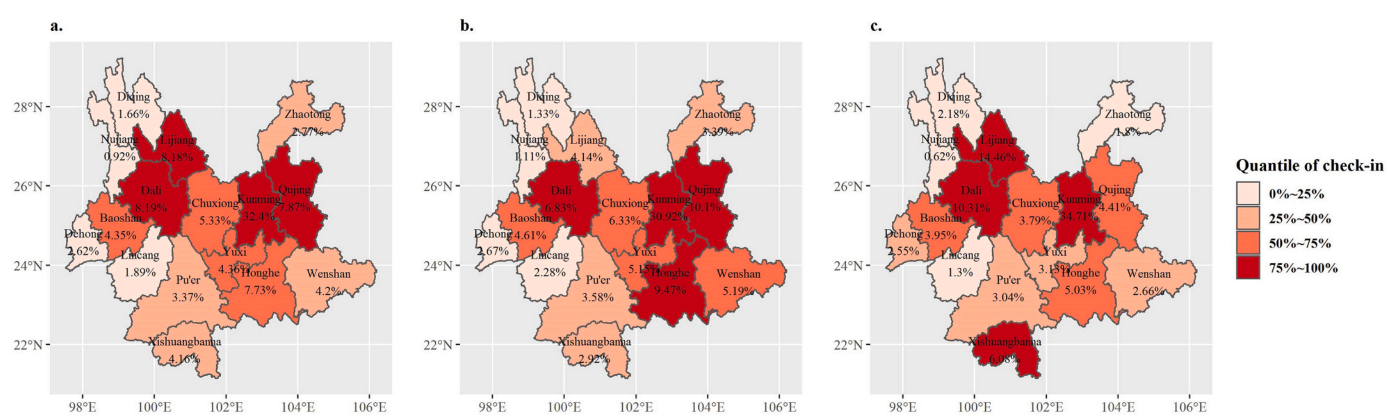

Fig. 6. The spatial distribution of the check-in record quantiles in Yunnan cities. Panel a shows the distribution of all visitors, Panel b shows the distribution of non-resident tourists, and Panel c shows is the distribution of non-resident tourists.

4.3. Regression analyses

To quantify the difference in the spatiotemporal patterns between resident tourists and non-resident tourists, we conduct regression analyses using the log-level model (Models 1 and 2 in Table 3). The results show that the LS of non-resident tourists is 30.70% higher than that of resident tourists, and the NVC of non-resident tourists is 24.91% higher than that of resident tourists. The coefficients of year show that the LS and NVC decrease each year. Regarding age groups, young (age between 20 and 29) and elderly people (age above 50) have a relatively high LS and NVC while middle-aged people (age between 30 and 49) have a relatively low LS and NVC, which is similar to a U-shaped curve. The coefficient of males shows that the LS of males is 2.35% less than that of females, and the NVC of males is 6.74% less than that of females. The coefficients of months show that compared to the reference month (January), the LS and NVC increase more during peak seasons (the LSs in July, August and September increase 7.43%, 7.20%, and 6.05%, respectively; while the NVCs in these months are 4.55%, 5.06%, and

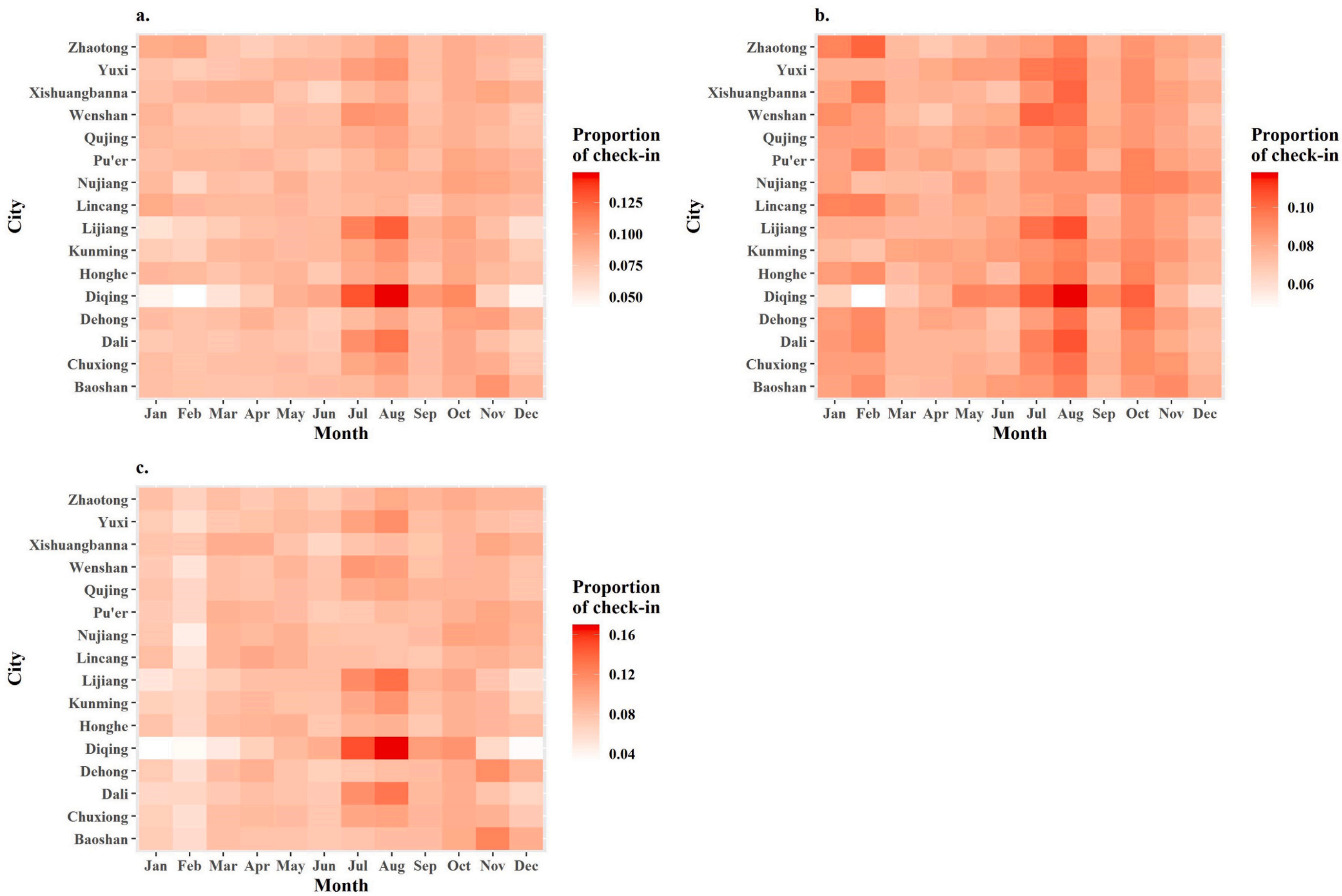

Fig. 7. The proportion of check-in records in each city in different months. Panel a shows the trend of all visitors, Panel b shows the trend of resident tourists, and Panel c shows the trend of non-resident tourists.

2.93% higher than the reference month, respectively).

As holiday travel is a considerable portion of the tourism market, we include interaction terms of holiday types and visitor types to further explore the difference in the spatiotemporal patterns between resident tourists and non-resident tourists during public holidays (Models 3 and 4 in Table 3). In Model 3, the coefficients of three-day holiday and seven-day holiday are −0.0754 and 0.0004, respectively, which indicate that tourists decrease their LS during short holidays but insignificantly increase their LS during long holidays compared to normal days. Furthermore, the coefficients of the interaction terms show that holidays have a stronger impact on non-resident tourists than resident tourists. Specifically, non-resident tourists decrease their LS 3.41% more in a three-day holiday than resident tourists. Similar results are shown in Model 4 for the NVC. We find that a 1.29% decrease and a 3.79% increase in the NVC occur during three-day holidays and seven-day holidays, respectively. Compared to resident tourists, non-resident tourists have an extra 3.73% decrease and 0.58% decrease during three-day holidays and seven-day holidays, respectively. Taken together, both resident tourists and non-resident tourists tend to decrease their LS and NVC during three-day holidays, and this drop is stronger for non-resident tourists than for resident tourists. In contrast, on seven-day holidays, both resident tourists and non-resident tourists are likely to increase their NVC but maintain their LS, and this effect is weaker for non-resident tourists.

5. Conclusion and discussion

In this study, we empirically study the spatiotemporal patterns of multiple city travel with a sample of hotel check-in registers, which contains 6.29 million records from January 1, 2015 to December 31, 2018. The massive hotel check-in register data could help us find general and solid spatiotemporal patterns. In detail, we mainly use exploratory data analysis to analyze the differences in the temporal patterns (the monthly trends and daily trends) and spatial patterns (the spatial distribution and seasonal tendency of the spatial patterns) between resident tourists and non-resident tourists. Then, we use regressions to quantify the differences between resident tourists and non-resident tourists through two dependent variables, namely, the length of stay (representing the temporal pattern) and the number of visited cities (representing the spatial pattern). From the various analyses of the spatiotemporal patterns expounded in the results section, we can conclude that resident tourists and non-resident tourists show different spatiotemporal patterns in multiple city travel, especially during holidays.

The temporal pattern analyses show that there are significant seasonal fluctuations in the monthly trends and daily trends. In our dataset, the peak season appears in the summer months, and the off season appears in the winter months. The reasons for seasonal fluctuations could be classified into two broad groups (Duro & Turrión-Prats, 2019). The first group of reasons are the related variables associated with climatic conditions and other environmental factors. The second group of reasons are institutional factors, such as specific holidays. In this paper, first, the temperature is cool for most cities in Yunnan during summer months, and summer months are the best time for many natural landscapes and cultural heritages in Yunnan (Allendorf, Brandt, & Yang, 2014; Yang, 2011). Second, some travel behaviors could cause seasonal fluctuations, such as seasonal retirement migrants (Wu, Hannam, & Xu, 2018), seasonal lifestyle tourism (Salazar & Zhang, 2013) and students'

Table 3
Regression on LS and NVC.

	Dependent variable			
	ln(LS)	ln(NVC)	ln(LS)	ln(NVC)
	(1)	(2)	(3)	(4)
Year 2016	−0.0115***	−0.0010	−0.0115***	−0.0011
Year 2017	−0.0351***	−0.0265***	−0.0351***	−0.0265***
Year 2018	−0.0475***	−0.0356***	−0.0475***	−0.0356***
Age group (20, 29)	0.0251***	0.0179***	0.0253***	0.0180***
Age group (30, 39)	0.0041*	0.0173***	0.0042*	0.0174***
Age group (40, 49)	−0.0014	0.0361***	−0.0013	0.0362***
Age group (50, 59)	0.0544***	0.1007***	0.0545***	0.1008***
Age group (60, +)	0.1338***	0.1768***	0.1338***	0.1767***
Male	−0.0235***	−0.0674***	−0.0235***	−0.0674***
February	−0.0066**	0.0144***	−0.0065**	0.0144***
March	0.0178***	0.0159***	0.0178***	0.0159***
April	0.0507***	0.0250***	0.0507***	0.0249***
May	0.0384***	0.0159***	0.0384***	0.0158***
June	0.0407***	0.0152***	0.0407***	0.0152***
July	0.0743***	0.0455***	0.0743***	0.0454***
August	0.0720***	0.0506***	0.0720***	0.0505***
September	0.0605***	0.0293***	0.0606***	0.0293***
October	0.0424***	0.0159***	0.0424***	0.0158***
November	0.0326***	0.0138***	0.0327***	0.0138***
December	0.0090***	−0.0040***	0.0090***	−0.0040***
Non-resident tourist	0.3070***	0.2491***	0.3080***	0.2510***
Three-day holiday	−0.0869***	−0.0254***	−0.0754***	−0.0129***
Seven-day holiday	0.0034	0.0359***	0.0004	0.0379***
Non-resident tourist: Three-day holiday			−0.0341***	−0.0373***
Non-resident tourist: Seven-day holiday			0.0085	−0.0058**
Constant	0.4488***	0.0834***	0.4483***	0.0827***
Observations	2,733,424	2,733,424	2,733,424	2,733,424
Adjusted R^2	0.0408	0.1444	0.0409	0.1445
Note:	*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$			

summer vacations in July and August (Bicikova, 2014; Thrane, 2016). We also find that the seasonal fluctuations of non-resident tourists are greater than those of resident tourists. The reason may be that traveling in the local province is more convenient than traveling in other provinces (in terms of distance), and thus resident tourists do not need to travel intensively at a certain time such as holidays. This phenomenon is similar to the findings of Segota and Mihalic (2018). They found that on the Slovenian coast, the contribution of international tourists to seasonal fluctuations was larger than that of domestic tourists. Our findings provide solid big data evidence that non-resident tourists bring more seasonal fluctuations than resident tourists from the perspective of multiple city travel within a province.

In addition to seasonality, we also find public holidays could cause spikes in check-in records. The spikes of seven-day holidays are larger than those of three-day holidays. First, people have extra time to travel on public holidays. Second, the governments have established policies to support travel and promote tourism development, such as Travel Golden Week (Wu, Xue, Morrison, & Leung, 2012) and free charge for highways on holidays (Hao, Yuxi, & Chunxiao, 2014; Zhao, Cui, Zhang, & Luo, 2019). Regarding the differences between resident tourists and non-resident tourists, we find that the spikes during weekends and three-day holidays are mainly contributed by resident tourists, and the spikes during seven-day holidays are mainly contributed by non-resident tourists. The reason for this phenomenon may be that short holidays are suitable for short distance travel. Therefore, three-day holidays and weekends are suited for resident tourists, while seven-day holidays are suited for non-resident tourists. Furthermore, seven-day holidays also provide relaxed time for residents to travel to other provinces. With these findings, stakeholders of the hospitality industry could better respond to fluctuations in tourism demand from the resident tourists and non-resident tourists.

The analyses of the spatial patterns show that the most visited cities have either high tourism revenue or high GDP. After visualizing the spatial distributions of resident tourists and non-resident tourists, we find that the check-in records of non-resident tourists are mainly in the cities with relatively high tourism revenues, while those of resident tourists are mainly in the cities with relatively high GDP. The difference in spatial patterns between resident tourists and non-resident tourists suggests that the motivations behind their trips could be different (Xue & Zhang, 2020). According to Xue and Zhang (2020)'s findings, resident tourists may travel more for business or shopping in high GDP cities, while non-resident tourists may travel more for leisure, sightseeing or different culture experience in tourism cities. Furthermore, the spatial patterns are not static and change along with the months. One possible reason is that Yunnan has various geographical environments, such as cold weather in high mountain areas and high temperatures in areas with low latitudes or low altitudes. This finding is similar to the seasonal tendency of the spatial patterns in Europe (Batista e Silva et al., 2018). Regarding the difference between resident tourists and non-resident tourists, one possible reason is that non-resident tourists face more restrictions such as time. Therefore, their travel is more concentrated in periods with long holidays. As a result, the seasonal tendency of the spatial patterns of non-resident tourists may be relatively fixed and regular.

Using regression analyses, we quantify the correlations between the segmentation variables and the spatiotemporal pattern indicators LS and NVC, as well as the differences between resident tourists and non-resident tourists on the LS and NVC. We find that the LS and NVC decline each year. The reason could be that the development of transportation decreases the time costs. We also find that elderly tourists have a larger LS and NVC than young visitors. Compared to young people, elderly people have more money and leisure time, and thus elderly people can travel to more destinations and stay longer. This finding is consistent with the booming market of senior visitors reported in the literature of Martínez-Garcia and Raya (2008) and Alén et al. (2014). Compared with females, males have a lower LS and NVC. A possible reason is that females spend more money and time in tourism markets (Bernini & Fang, 2020; Lin, Qin, Li, & Wu, 2020; Yang, Wu, & Lu, 2019). This gender difference is consistent with the findings of Santos et al. (2014). Regarding the differences in months, the LS and NVC in summer months are higher than those in other months because summer months are school summer holidays. On public holidays, we find that both the LS and NVC decline during three-day holidays while the NVC increases during seven-day holidays compared to normal days. We also explore the difference between resident tourists and non-resident tourists on public holidays. On three-day holidays, non-resident tourists decrease their LS and NVC more than resident tourists, while non-resident tourists increase their NVC less than resident tourists on seven-day holidays. The main reason for this difference may be that the length of the holiday limits the travel time. Given the time limitation, according to utility maximization, tourists try their best to travel as many places as possible to improve their travel experience. Non-resident tourists need more time to arrive in Yunnan than resident tourists. Thus, non-resident tourists have less time in Yunnan and travel to more cities to maximize their travel experiences.

Understanding multi-destination travel is an important issue for tourism management. The existing research extensively explores the spatiotemporal patterns of the travel within a city or a country but lacks evidence of multiple city travel. Our research uses hotel check-in registers to complement the existing literature from the viewpoint of multiple city travel within a province. Furthermore, our research provides new insights for tourism market segmentation, which is the segmentation of resident tourists and non-resident tourists. Our research also expounds the difference in the spatiotemporal patterns between resident tourists and non-resident tourists and their differences during public holidays. Through the findings of the differences in spatiotemporal patterns, some practical implications could be beneficial for tourism management. For example, the spatiotemporal patterns of multiple city travel could help some travel agencies design travel packages considering suitable spatiotemporal arrangements, such as developing seasonal travel routes and potential travel routes. Governments could invest in efficient transportation infrastructure, such as intercity railways, shuttle buses, and special travel lines; and enact tourism policy, such as establishing cross-border tourism economic zones, to accelerate the development of regional tourism integration. The difference in the segmentation of resident tourists and non-resident tourists could improve the performance of segment marketing and develop personalized travel products for segmentation. For instance, travel agencies could sell travel packages with different travel routes to different tourists. Hotels could offer some value-added services (such as tickets booking services) to both resident tourists and non-resident tourists. Hotel stakeholders could better understand the main segmentations of their guests when they make decisions of hotel locations in different cities. The difference in public holidays could enable tourism managers to improve tourism marketing, such as holiday product design and product and service pricing. Policy makers could effectively enact transportation policies to promote tourism development, such as free-charge highways and peak avoidance travel during public holidays.

Hotel guests have to register in most countries when they check in at a hotel. Thus, the tourism spatiotemporal patterns in other countries and regions could also be studied using hotel check-in registers. For example, Fyfe, Holdsworth, and Weaver (2009) used hotel check-in registers from commercial hotels in three small places in central Pennsylvania in America during the late 19th century to analyze visitation patterns. However, this was a small sample and may not be suitable for today's tourism situation. To the best of our knowledge, our study is the first attempt to study the spatiotemporal patterns of visitors using large-scale hotel check-in registers. Traditionally, hotel check-in registers are handwritten. With the development of information systems, hotel check-in registers have been digitized, which makes it easier to collect and analyze large-scale hotel check-in registers. Besides, handwritten registers are likely to encounter invalid records and missing records while digitized registers could also avoid such issues and improve data quality. There are several important advantages of hotel check-in registers. First, hotel check-in registers contain the demographic information of tourists. Unlike mobile phone positioning data, GPS data and geotagged social media data, analyzing hotel check-in registers does not need to identify tourists and estimate the demographic information of tourists, such as their ages and genders. Second, continuous check-in records could track multiple city travel trajectories. The traditional survey method can obtain visitors' demographic information, travel-related information, etc. in detail, but it is difficult to verify visitors' actual travel behaviors and analyze a large-scale population sample. In summary, hotel check-in registers are a new attractive data source that could contain demographic information on tourists and track multiple city travel.

There are some limitations and future research directions of this study. First, the partition of resident tourists and non-resident tourists is based on ID card information, but ID cards do not update residents' address in a timely manner. This may lead to some bias because of the misclassification of residents and non-residents. For example, someone may change their place of residence from a city in Yunnan to another city out of Yunnan, but their ID card may not timely reflect the change. Future research could avoid the misclassification of resident tourists and non-resident tourists if more detailed demographic information can be accessed. Second, the hotel check-in registers in our study are at the city level. Some cities may have only one attraction while some other cities have many attractions. Thus, the spatial preferences of visitors may be different in these cities. If the specific location of each hotel could be accessed, then the analysis could be deeper into the inner city, and the spatiotemporal patterns could be better analyzed. Third, the length of stay may be underestimated since hotel check-in records do not contain the exact time that tourists leave Yunnan. Future research could benefit from multisource data fusion such as transportation data. Finally, our dataset contains no information about the travel motivations of the tourists (e.g., leisure or business). Since the travel motivation influences travel behaviors, future research could improve the current work by combining hotel check-in registers datasets and survey datasets.

Credit author statement

Yuquan Xu: conceptualization, literature review, data analysis, writing - original draft.

Xiaobin Ran: writing - review & editing.

Yuewen Liu: conceptualization, supervision, data collection, writing - review & editing.

Wei Huang: supervision, writing - review & editing.

Declaration of interest

None.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant 71301128, 62041206, 71731009 and 72061127002), and is partially supported by the Government Fund of Guangdong Province (Grant No. 2018WZDXM020), Shenzhen Key Research Base in Arts & Social Sciences, and the Intelligent Management & Innovation Research Center, SusTech.

References

Al Haija, A. A. (2011). Jordan: Tourism and conflict with local communities. *Habitat International, 35*(1), 93–100. https://doi.org/10.1016/j.habitatint.2010.04.002.

Alegre, J., & Pou, L. (2006). The length of stay in the demand for tourism. *Tourism Management, 27*(6), 1343–1355. https://doi.org/10.1016/j.tourman.2005.06.012.

Alén, E., Nicolau, J. L., Losada, N., & Domínguez, T. (2014). Determinant factors of senior tourists' length of stay. *Annals of Tourism Research, 49*, 19–32. https://doi.org/10.1016/j.annals.2014.08.002.

Allendorf, T. D., Brandt, J. S., & Yang, J. M. (2014). Local perceptions of Tibetan village sacred forests in Northwest Yunnan. *Biological Conservation, 169*, 303–310. https://doi.org/10.1016/j.biocon.2013.12.001.

Antonakakis, N., Dragouni, M., & Filis, G. (2015). How strong is the linkage between tourism and economic growth in Europe? *Economic Modelling, 44*, 142–155. https://doi.org/10.1016/j.econmod.2014.10.018.

Barchiesi, D., Moat, H. S., Alis, C., Bishop, S., & Preis, T. (2015). Quantifying international travel flows using Flickr. *PLoS One, 10*(7), Article e0128470. https://doi.org/10.1371/journal.pone.0128470.

Barros, C. P., & Machado, L. P. (2010). The length of stay in tourism. *Annals of Tourism Research, 37*(3), 692–706. https://doi.org/10.1016/j.annals.2009.12.005.

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management, 68*, 101–115. https://doi.org/10.1016/j.tourman.2018.02.020.

Batista, E. S. F., Freire, S., Schiavina, M., Rosina, K., Marin-Herrera, M. A., Ziemba, L., … Lavalle, C. (2020). Uncovering temporal changes in Europe's population density patterns using a data fusion approach. *Nature Communications, 11*(1), 4631. https://doi.org/10.1038/s41467-020-18344-5.

Bernini, C., & Fang, R. (2020). Living standard and Chinese tourism participation. *International Journal of Tourism Research*, 1–14. https://doi.org/10.1002/jtr.2406.

Bicikova, K. (2014). Understanding student travel behavior: A segmentation analysis of British university students. *Journal of Travel & Tourism Marketing, 31*(7), 854–867. https://doi.org/10.1080/10548408.2014.890154.

Carr, N. (2002). Defining young tourists visiting beach-oriented resorts: A Behavioural analysis. *Anatolia, 13*(1), 49–61. https://doi.org/10.1080/13032917.2002.9687014.

Chen, S. C., & Shoemaker, S. (2014). Age and cohort effects: The American senior tourism market. *Annals of Tourism Research, 48*, 58–75. https://doi.org/10.1016/j.annals.2014.05.007.

Connell, J., Page, S. J., & Meyer, D. (2015). Visitor attractions and events: Responding to seasonality. *Tourism Management, 46*, 283–298. https://doi.org/10.1016/j.tourman.2014.06.013.

De Cantis, S., Ferrante, M., Kahani, A., & Shoval, N. (2016). Cruise passengers' behavior at the destination: Investigation using GPS technology. *Tourism Management, 52*, 133–150. https://doi.org/10.1016/j.tourman.2015.06.018.

Donaire, J. A., Galí, N., & Gulisova, B. (2020). Tracking visitors in crowded spaces using zenith images: Drones and time-lapse. *Tourism Management Perspectives*, 35. https://doi.org/10.1016/j.tmp.2020.100680.

Duro, J. A. (2016). Seasonality of hotel demand in the main Spanish provinces: Measurements and decomposition exercises. *Tourism Management, 52*, 52–63. https://doi.org/10.1016/j.tourman.2015.06.013.

Duro, J. A., & Turrión-Prats, J. (2019). Tourism seasonality worldwide. *Tourism Management Perspectives, 31*, 38–53. https://doi.org/10.1016/j.tmp.2019.03.010.
East, D., Osborne, P., Kemp, S., & Woodfine, T. (2017). Combining GPS & survey data improves understanding of visitor behaviour. *Tourism Management, 61*, 307–320. https://doi.org/10.1016/j.tourman.2017.02.021.
Esiyok, B., Kurtulmuşoğlu, F. B., & Özdemir, A. (2017). Heterogeneity in the determinants of length of stay across middle age and senior age groups in thermal tourism. *Journal of Travel & Tourism Marketing, 35*(4), 531–540. https://doi.org/10.1080/10548408.2017.1374906.
Fernández-Morales, A., Cisneros-Martínez, J. D., & McCabe, S. (2016). Seasonal concentration of tourism demand: Decomposition analysis and marketing implications. *Tourism Management, 56*, 172–190. https://doi.org/10.1016/j.tourman.2016.04.004.
Ferrante, M., Lo Magno, G. L., & De Cantis, S. (2018). Measuring tourism seasonality across European countries. *Tourism Management, 68*, 220–235. https://doi.org/10.1016/j.tourman.2018.03.015.
Fyfe, D. A., Holdsworth, D. W., & Weaver, C. (2009). Historical GIS and visualization. *Social Science Computer Review, 27*(3), 348–362. https://doi.org/10.1177/0894439308329762.
Guedes, A. S., & Jiménez, M. I. M. (2015). Spatial patterns of cultural tourism in Portugal. *Tourism Management Perspectives, 16*, 107–115. https://doi.org/10.1016/j.tmp.2015.07.010.
Hao, C., Yuxi, W., & Chunxiao, L. (2014). Highway toll charges guide holiday tourism supply and demand. *Tourism Tribune, 29*(1), 83–91.
Hasnat, M. M., & Hasan, S. (2018). Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. *Transportation Research Part C-emerging Technologies, 96*, 38–54. https://doi.org/10.1016/j.trc.2018.09.006.
Hoogendoorn, G., & Hammett, D. (2020). Resident tourists and the local 'other'. *Tourism Geographies*, 1–19. https://doi.org/10.1080/14616688.2020.1713882.
Hwang, J., Lee, K., & Park, J. (2012). A comparison of dining preference between resident and non-resident groups. *Journal of Foodservice Business Research, 15*(2), 156–175. https://doi.org/10.1080/15378020.2012.677392.
Jeuring, J. H. G., & Haartsen, T. (2016). The challenge of proximity: The (un) attractiveness of near-home tourism destinations. *Tourism Geographies, 19*(1), 118–141. https://doi.org/10.1080/14616688.2016.1175024.
Kellner, L., & Egger, R. (2016). Tracking tourist spatial-temporal behavior in urban places, A methodological overview and GPS case study. In A. Inversini, & R. Schegg (Eds.), *Information and communication technologies in tourism 2016*. Cham.
Khan, N. U., Wan, W., & Yu, S. (2020). Spatiotemporal analysis of tourists and residents in Shanghai based on location-based social network's data from Weibo. *ISPRS International Journal of Geo-Information, 9*(2). https://doi.org/10.3390/ijgi9020070.
Lee, C.-F. (2015). Tourist satisfaction with forest recreation experience: A segment-based approach. *Anatolia, 26*(4), 535–548. https://doi.org/10.1080/13032917.2014.997768.
Li, D., Zhou, X., & Wang, M. (2018). Analyzing and visualizing the spatial interactions between tourists and locals: A Flickr study in ten US cities. *Cities, 74*, 249–258. https://doi.org/10.1016/j.cities.2017.12.012.
Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management, 68*, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009.
Li, X., & Yang, W. (2017). Six jailed for vicious attack on tourist in Yunnan. http://www.chinadaily.com.cn/kindle/2017-08/18/content_30781070.htm.
Lin, V. S., Qin, Y., Li, G., & Wu, J. (2020). Determinants of Chinese households' tourism consumption: Evidence from China family panel studies. *International Journal of Tourism Research*, 1–13. https://doi.org/10.1002/jtr.2425.
Liu, P., Zhang, H., Zhang, J., Sun, Y., & Qiu, M. (2019). Spatial-temporal response patterns of tourist flow under impulse pre-trip information search: From online to arrival. *Tourism Management, 73*, 105–114. https://doi.org/10.1016/j.tourman.2019.01.021.
Liu, Q., Wang, Z., & Ye, X. (2018). Comparing mobility patterns between residents and visitors using geo-tagged social media data. *Transactions in GIS, 22*(6), 1372–1389. https://doi.org/10.1111/tgis.12478.
Losada, N., Alén, E., Nicolau, J. L., & Domínguez, T. (2017). Senior tourists' accommodation choices. *International Journal of Hospitality Management, 66*, 24–34. https://doi.org/10.1016/j.ijhm.2017.06.014.
Marrocu, E., & Paci, R. (2013). Different tourists to different destinations. Evidence from spatial interaction models. *Tourism Management, 39*, 71–83. https://doi.org/10.1016/j.tourman.2012.10.009.
Marrocu, E., Paci, R., & Zara, A. (2015). Micro-economic determinants of tourist expenditure: A quantile regression approach. *Tourism Management, 50*, 13–30. https://doi.org/10.1016/j.tourman.2015.01.006.
Martínez-Garcia, E., & Raya, J. M. (2008). Length of stay for low-cost tourism. *Tourism Management, 29*(6), 1064–1075. https://doi.org/10.1016/j.tourman.2008.02.011.
Mayer, M., Müller, M., Woltering, M., Arnegger, J., & Job, H. (2010). The economic impact of tourism in six German national parks. *Landscape and Urban Planning, 97*(2), 73–82. https://doi.org/10.1016/j.landurbplan.2010.04.013.
Mou, N., Yuan, R., Yang, T., Zhang, H., Tang, J., & Makkonen, T. (2020). Exploring spatio-temporal changes of city inbound tourism flow: The case of Shanghai, China. *Tourism Management, 76*. https://doi.org/10.1016/j.tourman.2019.103955.
National Bureau of Statistics of China. (2019). Number of oversea visitor arrivals by region. http://www.stats.gov.cn/tjsj/ndsj/2019/indexeh.htm.
Önder, I. (2017). Classifying multi-destination trips in Austria with big data. *Tourism Management Perspectives, 21*, 54–58. https://doi.org/10.1016/j.tmp.2016.11.002.
Park, S., Xu, Y., Jiang, L., Chen, Z., & Huang, S. (2020). Spatial structures of tourism destinations: A trajectory data mining approach leveraging mobile big data. *Annals of Tourism Research, 84*. https://doi.org/10.1016/j.annals.2020.102973.
Parroco, A. M., Vaccina, F., De Cantis, S., & Ferrante, M. (2012). Multi-destination trips and tourism statistics: Empirical evidences in Sicily. *Economics: The Open-Access, Open-Assessment E-Journal, 6*(2012–44). https://doi.org/10.5018/economics-ejournal.ja.2012-44.
Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management, 57*, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006.
Reif, J., & Schmücker, D. (2020). Exploring new ways of visitor tracking using big data sources: Opportunities and limits of passive mobile data for tourism. *Journal of Destination Marketing & Management, 18*. https://doi.org/10.1016/j.jdmm.2020.100481.
Ridderstaat, J., Croes, R., & Nijkamp, P. (2014). Tourism and long-run economic growth in Aruba. *International Journal of Tourism Research, 16*(5), 472–487. https://doi.org/10.1002/jtr.1941.
Rodríguez, X. A., Martínez-Roget, F., & González-Murias, P. (2018). Length of stay: Evidence from Santiago de Compostela. *Annals of Tourism Research, 68*, 9–19. https://doi.org/10.1016/j.annals.2017.11.001.
Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2018). Tourists' digital footprint in cities: Comparing big data sources. *Tourism Management, 66*, 13–25. https://doi.org/10.1016/j.tourman.2017.11.001.
Salazar, N. B., & Zhang, Y. (2013). Seasonal lifestyle tourism: The case of Chinese elites. *Annals of Tourism Research, 43*, 81–99. https://doi.org/10.1016/j.annals.2013.04.002.
Santos, G. E. D. O., Ramos, V., & Rey-Maquieira, J. (2014). Length of stay at multiple destinations of tourism trips in Brazil. *Journal of Travel Research, 54*(6), 788–800. https://doi.org/10.1177/0047287514532370.
Seetanah, B. (2011). Assessing the dynamic economic impact of tourism for island economies. *Annals of Tourism Research, 38*(1), 291–308. https://doi.org/10.1016/j.annals.2010.08.009.
Segota, T., & Mihalic, T. (2018). Elicitation of tourist accommodation demand for counter-seasonal responses: Evidence from the Slovenian Coast. *Journal of Destination Marketing & Management, 9*, 258–266. https://doi.org/10.1016/j.jdmm.2018.02.002.
Shen, H., Li, X., Luo, J. M., & Chau, K. Y. (2017). One country, two strata: Implications of social and cultural conflicts of Chinese outbound tourism to Hong Kong. *Journal of Destination Marketing & Management, 6*(3), 166–169. https://doi.org/10.1016/j.jdmm.2016.10.003.
Sheng, L., & Tsui, Y. (2010). Foreign investment in tourism: The case of Macao as a small tourism economy. *Tourism Geographies, 12*(2), 173–191. https://doi.org/10.1080/14616681003725219.
Sinclair, M., Mayer, M., Woltering, M., & Ghermandi, A. (2020). Using social media to estimate visitor provenance and patterns of recreation in Germany's national parks. *Journal of Environmental Management, 263*. https://doi.org/10.1016/j.jenvman.2020.110418.
Singh, S., & Krakover, S. (2015). Tourist experience at home – Israeli domestic tourism. *Tourism Management, 46*, 59–61. https://doi.org/10.1016/j.tourman.2014.06.007.
Su, S., Wan, C., Hu, Y., & Cai, Z. (2016). Characterizing geographical preferences of international tourists and the local influential factors in China using geo-tagged photos on social media. *Applied Geography, 73*, 26–37. https://doi.org/10.1016/j.apgeog.2016.06.001.
Su, X., Spierings, B., Dijst, M., & Tong, Z. (2019). Analysing trends in the spatio-temporal behaviour patterns of mainland Chinese tourists and residents in Hong Kong based on Weibo data. *Current Issues in Tourism*, 1–17. https://doi.org/10.1080/13683500.2019.1645096.
Sun, Y., Fan, H., Helbich, M., & Zipf, A. (2013). Analyzing human activities through volunteered geographic information: Using Flickr to analyze spatial and temporal pattern of tourist accommodation. In J. M. Krisp (Ed.), *Progress in location-based services* (pp. 57–69). Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-642-34203-5_4.
Tang, J., & Li, J. (2016). Spatial network of urban tourist flow in Xi'an based on microblog big data. *Journal of China Tourism Research, 12*(1), 5–23. https://doi.org/10.1080/19388160.2016.1165780.
Thrane, C. (2016). Students' summer tourism: Determinants of length of stay (LOS). *Tourism Management, 54*, 178–184. https://doi.org/10.1016/j.tourman.2015.11.010.
Thrane, C., & Farstad, E. (2011). Domestic tourism expenditures: The non-linear effects of length of stay and travel party size. *Tourism Management, 32*(1), 46–52. https://doi.org/10.1016/j.tourman.2009.11.002.
Tsaur, S.-H., Yen, C.-H., & Teng, H.-Y. (2018). Tourist–resident conflict: A scale development and empirical study. *Journal of Destination Marketing & Management, 10*, 152–163. https://doi.org/10.1016/j.jdmm.2018.09.002.
Turrión-Prats, J., & Duro, J. A. (2018). Tourist seasonality and the role of markets. *Journal of Destination Marketing & Management, 8*, 23–31. https://doi.org/10.1016/j.jdmm.2016.11.004.
Wan, Y. K. P., & Li, X. (2013). Sustainability of tourism development in Macao, China. *International Journal of Tourism Research, 15*(1), 52–65. https://doi.org/10.1002/jtr.873.
Wang, L., Fong, D. K. C., Law, R., & Fang, B. (2017). Length of stay: Its determinants and outcomes. *Journal of Travel Research, 57*(4), 472–482. https://doi.org/10.1177/0047287517700315.
Wu, B., Xue, L., Morrison, A. M., & Leung, X. Y. (2012). Frame analysis on golden week policy reform in China. *Annals of Tourism Research, 39*(2), 842–862. https://doi.org/10.1016/j.annals.2011.10.002.
Wu, Y., Wang, L., Fan, L., Yang, M., Zhang, Y., & Feng, Y. (2020). Comparison of the spatiotemporal mobility patterns among typical subgroups of the actual population

with mobile phone data: A case study of Beijing. *Cities, 100*. https://doi.org/10.1016/j.cities.2020.102670.
Wu, Y.-F., Hannam, K., & Xu, H.-G. (2018). Reconceptualising home in seasonal Chinese tourism mobilities. *Annals of Tourism Research, 73*, 71–80. https://doi.org/10.1016/j.annals.2018.09.005.
Xue, L., & Zhang, Y. (2020). The effect of distance on tourist behavior: A study based on social media data. *Annals of Tourism Research*, 82. https://doi.org/10.1016/j.annals.2020.102916.
Yang, C.-H., Lin, H.-L., & Han, C.-C. (2010). Analysis of international tourist arrivals in China: The role of world heritage sites. *Tourism Management, 31*(6), 827–837. https://doi.org/10.1016/j.tourman.2009.08.008.
Yang, C.-W., Wu, C.-L., & Lu, J.-L. (2019). Exploring the interdependency and determinants of tourism participation, expenditure, and duration: An analysis of Taiwanese citizens traveling abroad. *Tourism Economics*. https://doi.org/10.1177/1354816619896656.
Yang, L. (2011). Ethnic tourism and cultural representation. *Annals of Tourism Research, 38*(2), 561–585. https://doi.org/10.1016/j.annals.2010.10.009.
Yunnan Statistical Bureau. (2019). Basic statistics of tourism development. http://stats.yn.gov.cn/tjsj/tjnj/201912/t20191202_908222.html.
Zhao, J., Cui, J., Zhang, Y., & Luo, T. (2019). In W. Wang, K. Bengler, & X. Jiang (Eds.), *Impact of holiday-free policy on traffic volume of freeway: An investigation in Xi'an.* Singapore: Green Intelligent Transportation Systems.
Zhao, X., Lu, X., Liu, Y., Lin, J., & An, J. (2018). Tourist movement patterns understanding from the perspective of travel party size using mobile tracking data: A case study of Xi'an, China. *Tourism Management, 69*, 368–383. https://doi.org/10.1016/j.tourman.2018.06.026.
Zhong, L., Sun, S., & Law, R. (2019). Movement patterns of tourists. *Tourism Management, 75*, 318–322. https://doi.org/10.1016/j.tourman.2019.05.015.

Yuquan Xu is a Ph.D. student in Management Science and Engineering at the School of Management, Xi'an Jiaotong University from 2016. His main research interests are travel behavior research, big data analysis and machine learning.

Xiaobin Ran is currently a Ph.D. student in Management Science and Engineering at the school of Management, Xi'an Jiaotong University from 2015. His research interests are social network analysis, data-driven individual behavior research, with special attention to the interaction between individuals, and related research on the path and results of individual behavior diffusion.

Yuewen Liu is currently a Professor with Xi'an Jiaotong University, Xi'an, China. He received the Ph.D. degree of Information Systems from City University of Hong Kong (2010), and a Ph.D. degree in Management Science and Engineering from the University of Science and Technology of China (2010). His main research fields include user behavior research, innovation diffusion research, electronic commerce, social network etc. He published more than 20 research papers in international journals, including in top-tier journals such as Information Systems Research, Annals of Statistics, IEEE Transactions on Cybernetics, Decision Support Systems, etc.

Wei Huang is a professor at the Management School of Xi'an Jiaotong University. He was a senior visiting fellow at Queensland University of Technology, Australia. He was a fellow of Harvard University and a visiting scholar at University of Georgia, USA. He has worked as a full-time faculty in top-tier research universities in Australia, Singapore, Hong Kong and USA. He published more than 60 research papers in international journals, including in top-tier journals such as MIS Quarterly, Journal of MIS, Communications of ACM, ACM Transactions, IEEE transactions, European Journal of Information Systems, etc.