Current Issues in Tourism

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rcit20

Agglomeration and/or differentiation at regional scale? Geographic spatial thinking of hotel distribution – a case study of Guangdong, China

Lei Fang , Yujing Xie , Shenjun Yao & Ting Liu

CURRENT ISSUES IN TOURISM
https://doi.org/10.1080/13683500.2020.1792852

Agglomeration and/or differentiation at regional scale? Geographic spatial thinking of hotel distribution – a case study of Guangdong, China

Lei Fang[a,b,c], Yujing Xie[b], Shenjun Yao[d] and Ting Liu[a]

[a]Institute of Remote Sensing and Earth Sciences, Hangzhou Normal University, Hangzhou, People's Republic of China; [b]Department of Environmental Science and Engineering, Fudan University, Shanghai, China; [c]Shenzhen Institute of Research and Innovation, The University of Hong Kong, Shenzhen, People's Republic of China; [d]School of Geographical Science, East China Normal University, Shanghai, People's Republic of China

ABSTRACT

This study is among the first attempts to understand hotel location choice by developing a spatial framework based on spatial autocorrelation to investigate hotel pattern. By taking Guangdong Province, China as the study context, this research analyzes the differentiation and agglomeration effects of star-rated hotels in terms of absolute number and density at the county level with Local Moran's *I*. The spatially diverse relationships between 13 factors and the number of star-rated hotels were quantified by regions classified by the Local Moran's *I*. Results indicate that (1) counties with positive agglomeration were located south of the Pearl River Delta; (2) factors influencing hotel location choice varied across regions; and (3) attractions and urbanization do not always affect hotel location choices.

ARTICLE HISTORY
Received 26 February 2019
Accepted 26 June 2020

KEYWORDS
Hotel; spatial analysis; China; Guangdong; Moran's I; GIS

1. Introduction

As part of the service industry, hotels rely heavily on effective location strategies to successfully attract guests and win the competition. One research direction in hotel location is the agglomeration effect, which refers to the benefits that hotels can receive from clustering. Under an agglomeration framework, a hotel located in a cluster has advantages, such as more accessible resources and lower searching costs for consumers (Canina et al., 2005). The theory is supported by microeconomics studies suggesting that the knowledge spillover theory of entrepreneurship holds for regions and industries (e.g. Audretsch & Lehmann, 2005). However, the benefits that hotels gain from agglomeration effects can be heterogeneous (Chung & Kalnins, 2001). Hotels that pursue a high level of differentiation may contribute to the positive agglomeration effects of other hotels (Canina et al., 2005).

From a geographical perspective, the nature of spatial agglomeration effects is a problem of 'space' which is characterized by a set of geographical coordinates and spatial interactions. Thus, despite the plethora of research on hotel clustering, little progress has been made to confront this crucial issue. Previous studies (e.g. Barros, 2005; Urtasun & Gutiérrez, 2006b) considered hotels' unbalanced distribution in space as a phenomenon of agglomeration heterogeneity and ignored the spatial variation law concerning the changes of spatial distances shown by this distribution phenomenon. Previous studies on agglomeration also focused closely on the clustering tendency of individual hotels. They treated clusters independently without considering the relationship between clusters

CONTACT Ting Liu tingliu21@gmail.com
Supplemental data for this article can be accessed at https://doi.org/10.1080/13683500.2020.1792852

and their neighbourhoods. Hence, these studies failed to fully reflect the agglomeration or differentiation effects of their target study areas. This deficiency is known as 'the dirty little secret of hotel agglomeration effect research' (Short et al., 1996). Most previous studies on hotel clusters focused on non-spatial factors, such as the decision making, strategy development and management of investors and managers, and the operation, size and brand of hotels. One factor that has not been widely considered is whether a hotel shows a clustering tendency in space. Without adequate research efforts in this issue, the works of applied geographers and other professionals toward the policy implications in addressing the agglomeration effects of hotels could be inappropriate.

The current research has the following objectives: (1) to develop a framework to analyze the differentiation and agglomeration effects of hotels based on spatial autocorrelation, (2) to analyze the spatial patterns of star-rated hotels and four- or five- star – rated (hereafter referred to as 4-5-star-rated) hotels in Guangdong Province in 2013 by using the proposed framework, and (3) to explore the potential factors contributing to the spatial patterns of star-rated hotels and 4-5-star-rated hotels. With counties of different sizes, this study analyzes the hotels in terms of absolute number and density.

This paper is organized as follows: Section 2 describes the related studies in the areas of hotel location, differentiation/agglomeration effects, and the geographic spatial method. Section 3 illustrates the fundamental research design, including the framework developed under geographic spatial thinking, the dataset, and the local Moran's *I* model. Section 4 presents the data analysis and results in detail, including the spatial patterns of star rated and 4-5-star-rated hotels and the factors that may influence the spatial distribution. Correlation analysis between hotel pattern and its potential factors is described in this section. Based on the results of the local Moran's *I*, we divided the entire region into differentiation/agglomeration zones and analyzed the potential factors for each zone. The outcomes were compared with those with no division. Section 5 concludes this research and presents directions for future research.

2. Literature review

Hotels in urban areas are usually located in clearly defined clusters, which indicates that their distribution is not random. A number of prior empirical studies have examined the determinants of hotel location choice, and identified factors such as the convenience of transportation and parking (e.g. Li et al., 2015; Tsaur & Tzeng, 1996), accessibility to tourism attractions (e.g. Yang et al., 2012), surrounding public service infrastructure and economic environment (e.g. Yang et al., 2014), hotel characteristics (e.g. Yang et al., 2014), and agglomeration effects (e.g. Freedman & Kosová, 2012; Luo & Yang, 2016; Yang et al., 2012). A major incentive for hotels to locate close to each other is the ensuing increase in hotel efficiency because they are able to benefit from the positive spillover effects from their neighbours (Barros, 2005). This strategy is especially obvious in Chinese cities because of the importance of agglomeration economies brought about by the large size of these cities (Egan et al., 2006). Canina et al. (2005) ventilated some reasons behind clusters from both the demand and production perspectives. From a consumer's perspective, agglomeration cuts down the cost of searching. From a producer's perspective, it enables cluster members to gain easier access to major suppliers, exclusive services, or relationships. However, some hotels may still not enjoy the benefits of agglomeration. Chung and Kalnins (2001) found through their study of the lodging industry in Texas that the benefits of agglomeration to hotels are heterogeneous. Similar hotels are likely to benefit less from agglomeration than hotels dissimilar to nearby competitors.

In previous empirical studies, regression analysis was generally used to identify agglomeration areas and establish the effects of hotel location according to different research aims. Baum and Haveman (1997) examined the effects of localized competition on failure rates in the Manhattan hotel industry from 1898 to 1990. In their exponential regression model, the independent variables were size, geographic location (relative position of hotels), price, and population density. Urtasun and Gutiérrez (2006a) investigated geographic location, size, price, and services to ascertain how the distribution of existing hotels influences the positioning of new competitors. They identified the relative values of these four factors, combined the data in four simultaneous equations, and compared the results using the OLS method. Yang et al. (2012) investigated the potential impacted factors that contribute to hotel location choice by dividing hotels into five categories according to their distance to the city centre and using a logit model with location characteristics in Beijing. They found the factors such as years after opening, star rating, ownership, service diversification, agglomeration effect, accessibility to tourist sites, road accessibility, public service infrastructure, and metro accessibility were important in determining location choice of hotels. However, these studies did not address the agglomeration effect of hotels from the perspective of space. From a geographical perspective, the nature of spatial agglomeration effects is a problem of 'space' which is characterized not only by a set of geographical coordinates but also by spatial interactions. Most previous studies on agglomeration always focused on the clustering tendency of individual hotels. They treated clusters independently without considering the relationships between clusters and their neighbourhoods and thus could not fully reflect the agglomeration or differentiation effects of their study areas.

In the literature, there were a few researchers who realized the 'space' problem and conducted research on agglomeration effects by utilizing spatial methods. Since the early 1990s, mapping and GIS have been used to help to describe hotel location distribution and the possible factors that shape hotel patterns (e.g. Bloomfield, 1996; Bull & Church, 1994; Dökmeci & Balta, 1999; Ferreira & Boshoff, 2013; Prayag et al., 2012; Rogerson, 2013a, 2013b). Other more complex spatial statistical tools have also been developed in recent years aimed at understanding hotel distribution and associated agglomeration effects. Wall et al. (1985) employed point pattern analysis tools to study the spatial distribution of accommodations in Toronto, Canada. They found the agglomeration patterns of distribution and revealed the changes in the mean centre, dispersion, and orientation of the distribution over periods. Another study by Broadway (1993) computed the geographic centres of hotel distribution over different periods in Montreal, Canada. The results indicated that the centre moderately shifted over time. As for other point pattern analysis tools, and Yang and Fik (2011) employed the K-function to understand hotel agglomeration areas at multiple distances, and Sund (2006) investigated the inequity of hotel distribution using the Lorenz curve. Spatial regression was also used by Helmers (2010) to identify the spatial agglomeration pattern of new hotel entrants. Luo and Yang (2013) distinguished the star-rated hotel hotspots in Chinese cities and explained four factors shaping this pattern. Among these studies, only a few acknowledged the clustering tendency of hotels (e.g. Luo & Yang, 2013; Yang & Fik, 2011). Luo and Yang (2013) did consider spatial relationships, but they focused closely on the detection of hotspots at the regional level.

3. Methodology

3.1. Study area and data

China, as a transforming society, has witnessed a rapid development of the hotel industry in recent decades. In 2013, the number of star-rated hotels reached 11,687, 3,100 of which were awarded three-star rating. This research chose Guangdong Province as the study area because it ranked first during this period in terms of the number of star-rated hotels (917 in total) among provincial spatial units. It is located in the southern part of China, with an area of 179.7 thousand square km.

The address information of star-rated hotels was collected from Guangdong Tourism Statistical Yearbook 2013, whereby the geographical information (longitude and latitude) of these hotels was obtained with Baidu Geocoder API (http://lbsyun.baidu.com/index.php?title=webapi/guide/webservice-geocoding). The Baidu Geocoder API can provide the geographical coordinates of each named hotel. In a GIS environment, these hotels were plotted onto a map as a set of points. Looking into the locations of these hotels (see Figure 1), one may easily observe their heterogeneous distribution. To answer questions such as 'Are the hotels located in random places?' or

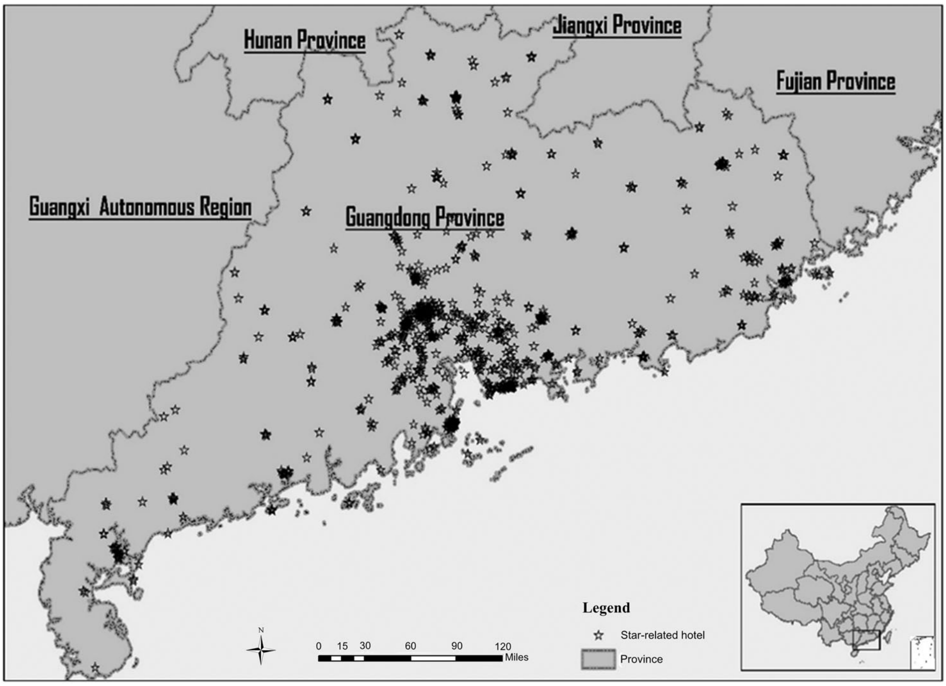

Figure 1. Spatial distribution of hotels in Guangdong Province in 2013(CIT remains strictly neutral with respect to jurisdictional claims on disputed territories in published maps, and the naming conventions used in the maps included in the figure).

'Do the hotels tend to cluster together to attract more customers?', the spatial patterns of hotels should be appropriately explored because the spatial location is one of the most important factors of hotel establishment and management (Canina et al., 2005; Chung & Kalnins, 2001; Urtasun & Gutiérrez, 2006b).

In an aggregate analysis, the spatial scale is crucial. While dividing an entire region into a set of regular zones allow us to hold a constant size of the analytical spatial units, obtaining the socio-economic characteristics of these zones is difficult because the attributes are always recorded based on administrative units or census tracts with varied size. For the spatial analysis of a tourism phenomenon, one would expect to explain the possible reasons. Hence, administrative spatial units or census tracts may be appropriate for analyzing hotel patterns. From the perspective of policy implication for the local tourism government, this study analyzed the hotel pattern of Guangdong Province at the county level. As of the study year, Guangdong Province comprises 121 counties. In analyzing the potential factors that influence hotel distribution patterns, independent variables were chosen through surveys or reviews of previous studies. The descriptive statistics on hotel and 13 potential factors are presented in Table 1. The names of attractions were collected from Ctrip (https://www.ctrip.com/), which is the top online travel agency in China. The attractions' locational information was derived from Baidu Geocoder API with their names. With geographical coordinates, the attractions were then plotted onto the map. Demographic and socio-economic data were collected from Guangdong Statistical Yearbook 2013. The variables describing land use characteristics were derived from a land-use dataset of 1-km spatial resolution, which is provided by the Resource and Environment Data Could Platform (http://www.resdc.cn/data.aspx?dataid=184). Figure 1 delineates spatial distributions of variables. It can be observed that they were not randomly distributed. For instance, the hotels and population were likely to cluster in the Pearl River Delta (PRD) region. In dealing with this issue, a spatial analytical framework is desirable for analyzing the hotels (Figure 2).

Table 1. Descriptive statistics of variables.

Variables	Description	*N*	Minimum	Maximum	Sum	Mean	S.D.
Dependent variable							
StarHotel	Number of star-rated hotels	121	0	90	1100	9.091	13.635
HighStar	Number of 4-5-star-rated hotels	121	0	46	304	2.512	5.172
HotelDen	Number of star-rated hotels per km^2	121	0.000	1.205	4.284	0.035	0.130
HStarDen	Number of 4-5-star-rated hotels per km^2	121	0.000	0.427	1.288	0.011	0.043
Demographic & Socio-economic characteristic							
Pop	Population size	121	90515	8220207	102971990	851008.182	932819.323
UrbanPop	Urban population size	121	32200	7271322	68242108	563984.364	875826.380
UrbanRatio	Proportion of urban population	121	0.199	1.000	71.097	0.588	0.265
GDP	Gross domestic product (RMB ¥10,000)	121	171193	42464527	475139592	3926773.49	6223853.93
PGDP	Per capita GDP (RMB ¥)	121	6735.000	394866.390	5236968.74	43280.733	48811.9
UrbanPInc	Urban per capita income (RMB ¥)	121	14192.0	75985.0	3711090.92	30670.173	11144.653
Attractions & Transportation accessibility							
Attr	Number of attractions	121	0	27	643	5.314	5.329
RoadDen	Length of the road network per km^2 (km/km^2)	121	0.035	12.605	160.899	1.330	2.169
Land use characteristics							
pGreen	Proportion of green land area	121	0.000	0.897	62.959	0.520	0.273
pConstruction	Proportion of construction area	121	0.002	0.936	18.772	0.155	0.202
pFarmland	Porportion of farmland area	121	0.000	0.613	30.611	0.253	0.136
pUnused	Proportion of unused land area	121	0.000	0.044	0.163	0.001	0.005
pWater	Porportion of water area	121	0.000	0.516	8.495	0.07	0.09

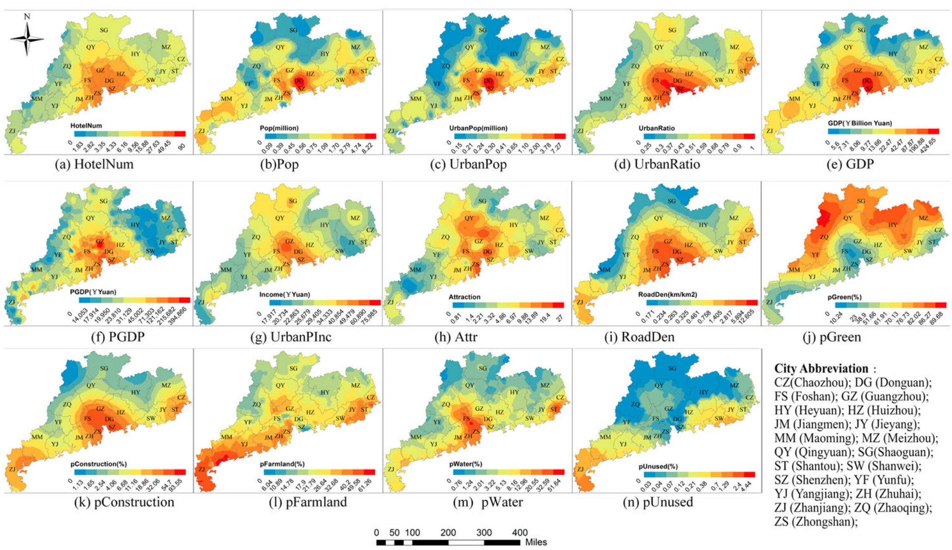

Figure 2. Illustration of variables.

3.2. Spatial analytical framework

With traditional approaches, the spatial patterns of hotels were summarized as agglomeration ('H') or differentiation ('L'). By focusing on the number of hotels and spatial interactions, this study represented the spatial patterns from a unary manner to a binary manner. A binary representation can be positive agglomeration ('HH'), negative agglomeration ('LL'), positive-negative agglomeration ('HL'), and negative-positive agglomeration ('LH'). 'HL' and 'LH' can be regarded as differentiation effects. The fifth type is insignificant agglomeration, which means that hotels may be randomly distributed. 'HH' (or 'LL') suggests that hotels in the target and neighbouring areas are densely (or sparsely) distributed with small spatial variation. 'HL' (or 'LH') indicates that hotel distribution in the target area reflects a clustering (or dispersion) tendency while hotel patterns in neighbouring areas suggest dispersion (or clustering).

Previous research usually established statistical models to describe the overall data relationships in entire study areas. If the relationships are consistent across study areas, a global statistical method can be used to model the relationships well. When the relationships differ by regions of the study area; however, the regression model describes a global average of the mix of relationships and cannot represent all relationships. One solution that has been widely used in previous studies is to divide the study area into several parts according to specific rules (Yang et al., 2012). A major limitation of such is the heavy reliance on division rules. In the current research, we proposed a new framework (see Figure 3) to measure the varying relationships between hotel location and corresponding influential factors when the hotel distribution shows a spatial autocorrelation. The proposed framework includes the following steps: (1) spatial autocorrelation test to identify the spatial autocorrelation of hotels and potential factors, (2) spatial division of the region according to the results of spatial autocorrelation, and (3) statistical modelling for analyzing hotels in different types of regions.

A well-known spatial statistic for measuring spatial autocorrelation is Moran's I (Moran, 1948, 1950), which refers to the degree at which the value of a variable at a location covaries with the values of that variable at nearby locations. By decomposing the global Moran's I, Anselin (1995) further developed the local Moran's I statistic to detect clustering tendencies at the local level. The current research employed the local Moran's I statistic to implement the analytical framework.

For county i, the local Moran's I index is given by Anselin (1995) as

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_{j=1, j \neq i}^{n} W_{ij}(x_j - \bar{x}) \quad (1)$$

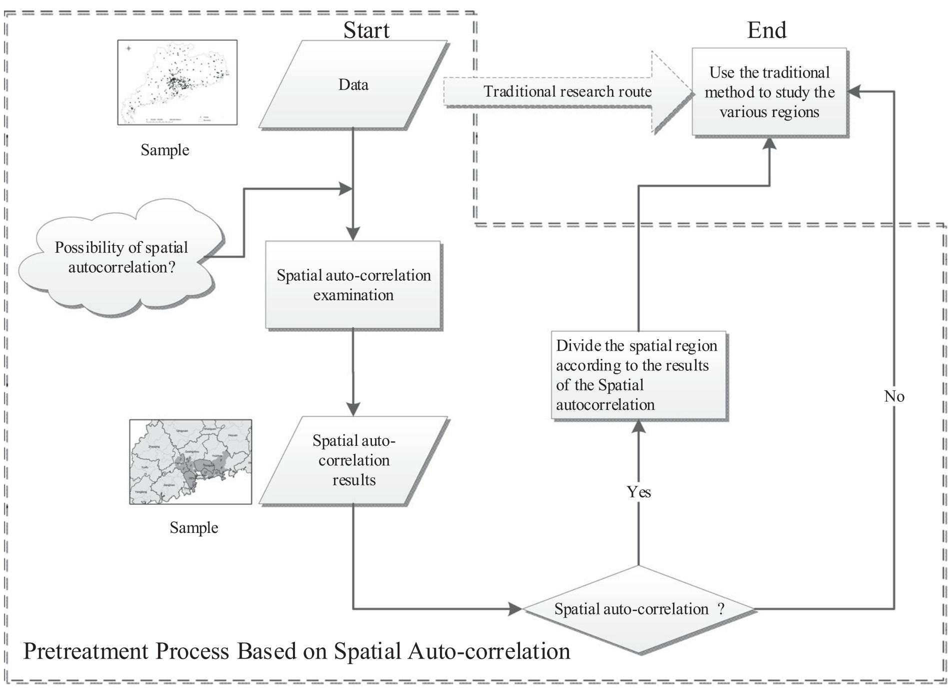

Figure 3. Illustration of proposed study framework.

where n is the number of counties, x_i is the value of attribute X (e.g. number of star-rated hotels) at county i, $\bar{x}$ is the mean value of X, and W_{ij} is a matrix indicating the spatial weight between counties i and j. S^2 is the variance of X calculated by

$$S^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n} \tag{2}$$

A positive value for ***I*** indicates that a county and its neighbouring counties have a similar high or a low number of X values. A negative value for *I* indicates that a county and its neighbouring counties have significantly different X values. We defined a county and its neighbouring counties that form a statistically significant cluster of high values as HH; a county and its neighbouring counties that constitute a statistically significant cluster of low values as LL; a county with a high value and its neighbours with low values as HL; and a county with a low value and its neighbours with high values as LH.

To explore the possible reasons that could explain the spatial patterns of star-rated hotels, we analyzed the potential explanatory factors using Moran's *I*. ArcGIS software 10.2 was used for the analysis. Weight matrix *W* was defined by the 'contiguity' method, which only includes neighbouring polygons for computation.

4. Results and discussion

4.1. Spatial patterns of star-rated hotels

Figure 4 depicts the spatial pattern of star-rated hotels. It can be observed that 16 out of the 121 counties are significantly autocorrelated. They are labelled with numbers in Figure 4b and details are shown in Table 2. All the counties are situated in the south of the PRD, which is one of the most developed regions of China.

Star-rated hotels are highly clustered (HH in Table 2) in 15 counties of 7 cities. These hotels are located in urban areas, thereby indicating positive agglomeration effects. Zhongshan and Dongguan were two small cities where the development of hotels partly depended on their unique tourism resources. Zhongshan was known as the hometown of Sun Yat-sen, and Dongguan was well-known for its porn industry. Nonetheless, the hotels in these two cities might be affected by their neighbourhoods. Zhongshan could be influenced by the Shunde District of Foshan in the north, Panyu District of Guangzhou City in the northeast, and Xiangzhou District of Zhuhai City in the south (see Table 2 and Figure 4). These areas are located in the urban core of the PRD region, which has undergone rapid economic development in recent decades. The agglomeration of hotels in Zhongshan had no significant relationship with Jiangmen City in the west, where the economy was less developed than that in other cities. The pattern constituted a north-south corridor. Dongguan was significantly and positively associated with the spatial agglomeration of star-rated hotels in Panyu District of Guangzhou in the west, Baoan District of Shenzhen in the south, and two districts of Huizhou in the east, thereby constituting an east-west corridor. No significant neighbourhoods were situated in the north, probably because these areas were not in the urban core of the city and had limited tourism resources.

In the HH regions, the hotel industry in a county (or a district) could be influenced by its neighbourhood. For instance, as the Chinese government cracked down on the porn industry in Dongguan, a significant number of hotels might not survive. The effect might reach neighbouring counties such as the Baoan District of Shenzhen.

Nansha District, which is far from the downtown area of Guangzhou City, is famous for its beautiful natural scenery. However, only one five-star hotel could be found in the district. The high clustering tendency detected in its neighbourhoods might be partly explained by the convenient traffic from neighbouring districts. The 'LH' relationship indicates that the tourism resources of the Nansha District might have been utilized by its neighbours.

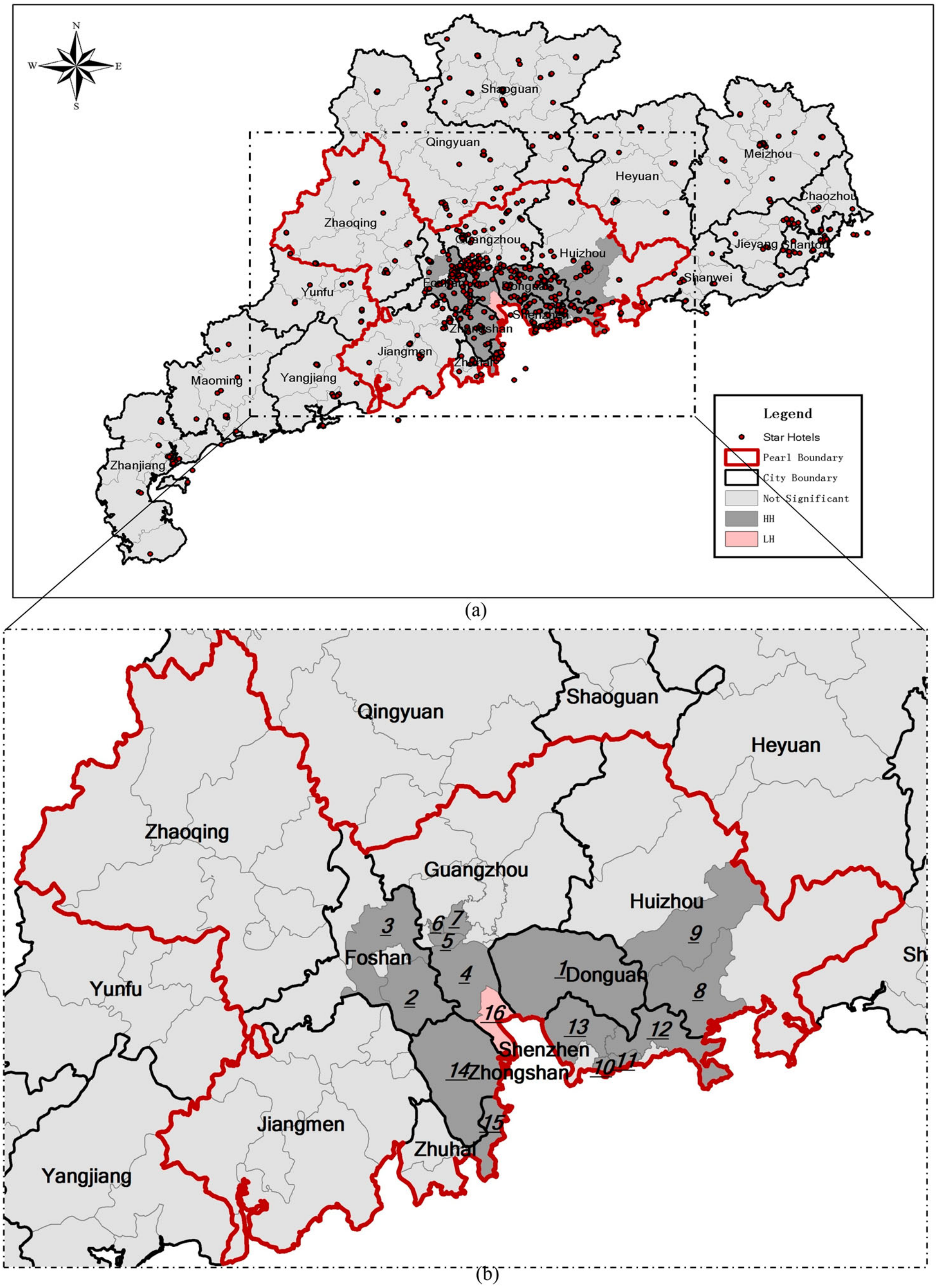

Figure 4. Spatial pattern of star-rated hotels in Guangdong Province.

In terms of the density of star-rated hotels, only 6 HH counties were detected. The two positive agglomeration clusters were in Guangzhou and Shenzhen. In comparison with the pattern for the number of star-rated hotels, five districts were detected as significant in terms of both absolute number and density (highlighted in grey in Table 3). The clusters were in the urban core of the two most important cities (that is, Guangzhou and Shenzhen) in Guangdong Province (see Figure 5). The CBDs of these two cities are in these two clusters. Highly integrated services, such as transport systems, might significantly improve accessibility. Also, the Shenzhen cluster might be positively

Table 2. Significant counties with star-rated hotels in Guangdong Province.

No.	Moran's *I* Index	*Z* Score	*P*-value	Type	Hotel Count	District	City
1	2.137	7.509	<0.001	HH	90	Dongguan	Dongguan
2	1.263	3.088	0.002	HH	27	Shunde	Foshan
3	0.626	2.220	0.026	HH	31	Nanhai	Foshan
4	1.266	4.218	<0.001	HH	25	Panyu	Guangzhou
5	1.098	2.688	0.007	HH	20	Haizhu	Guangzhou
6	3.120	6.774	<0.001	HH	48	Yuexiu	Guangzhou
7	2.299	5.604	<0.001	HH	56	Tianhe	Guangzhou
8	1.381	3.008	0.003	HH	19	Huiyang	Huizhou
9	1.533	3.744	<0.001	HH	27	Huicheng	Huizhou
10	2.168	4.065	<0.001	HH	29	Futian	Shenzhen
11	2.324	5.049	<0.001	HH	55	Luohu	Shenzhen
12	1.097	2.685	0.007	HH	16	Longgang	Shenzhen
13	2.463	6.004	<0.001	HH	24	Baoan	Shenzhen
14	1.002	3.140	0.002	HH	30	Zhongshan	Zhongshan
15	3.298	5.026	<0.001	HH	68	Xiangzhou	Zhuhai
16	−1.497	−2.781	0.005	LH	2	Nansha	Guangzhou

Table 3. Significant counties for star-rated hotel density in Guangdong Province.

No.	Moran's *I* Index	*Z* Score	*P*-Value	Type	Hotel Density	District	City
1	23.640	14.588	<0.001	HH	0.35	Tianhe	Guangzhou
2	34.053	23.465	<0.001	HH	1.20	Yuexiu	Guangzhou
3	13.648	8.433	<0.001	HH	0.19	Haizhu	Guangzhou
4	3.547	2.211	0.027	HH	0.08	Liwan	Guangzhou
5	10.599	8.445	<0.001	HH	0.35	Futian	Shenzhen
6	10.468	7.229	<0.001	HH	0.59	Luohu	Shenzhen

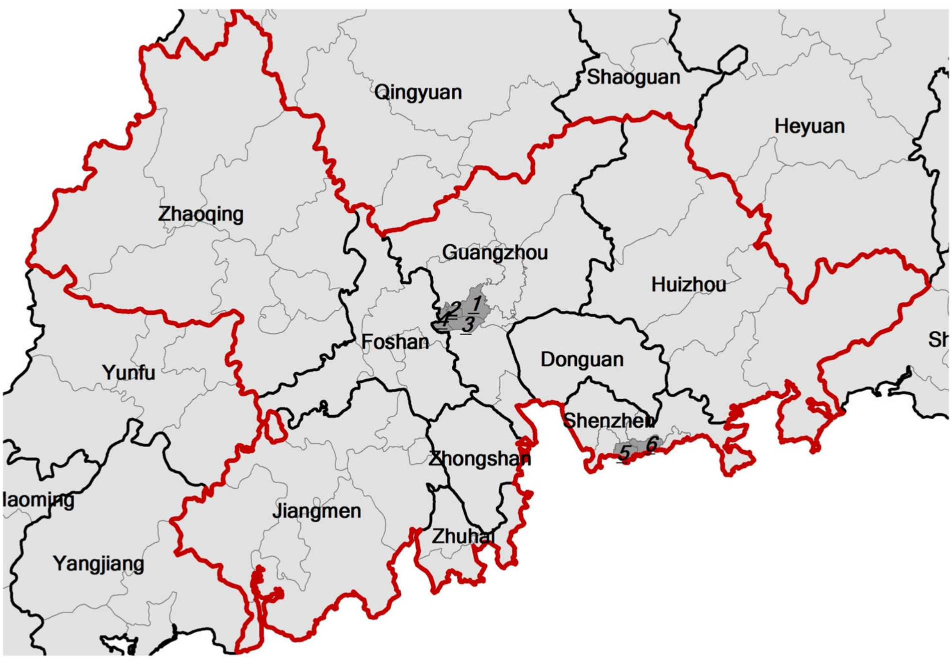

Figure 5. Spatial pattern of star-rated hotel density in the Pearl River Delta.

influenced by tourists travelling to Hong Kong because Luohu and Futian checkpoints are situated in these two districts. The metro system of Hong Kong allows tourists to travel from the checkpoints to major shopping areas, including Tsim Sha Tsui and Mong Kok, within one hour. Although Nanshan District of Shenzhen has one checkpoint, the transport system that connects the checkpoint and the downtown area of Hong Kong is not as convenient as that in Luohu and Futian. This discrepancy can partly explain why Nanshan District does not have many star-rated hotels. The 'core' area could be ranked first if policymakers are willing to restrict the development of the hotel industry. For researchers, these areas should be worth studying if they focus on agglomeration effects.

By comparing Tables 2 and 3, it can be observed that 10 HH districts (or counties) identified in terms of the number of star-rated hotels were detected as insignificant in the pattern of star rated hotel density. However, looking at Figures 4 and 5, one may find that most of the hotels in these ten districts were located surrounding the two clusters. A probable reason could be the saturation of the hotel industry in the two regions, where the limited space resulted in 'hotel dispersal' to the neighbourhood.

4.2. Spatial patterns of 4-5-star-rated hotels

The spatial pattern identified for the 4-5-star-rated hotels (see Figure 6 and Table 4) was slightly different from that identified for the star-rated hotels (see Figures 4 and 6). The positive agglomeration area was in the urban area of the southern part of the PRD. This location was the most developed region in the PRD and covered 12 administrative units, including the Nanshan District of Shenzhen (not shown in Table 2). The exception suggests that in Shenzhen, investors were likely to establish 4-5-star-rated hotels in Nanshan District, probably because of the proximity to the central business district (CBD) of Shenzhen. The north-south corridor in the pattern of star-rated hotels did not exist in 4-5-star-rated hotels, whereas the east-west corridor covered similar counties (or districts).

In the examination of the density of 4-5-star-rated hotels, two clusters were located in Guangzhou and Shenzhen, respectively (see Figure 7 and Table 5). The spatial pattern was generally the same as that for star-rated hotels (see Figures 5 and 7). The result indicates that the star-rated hotels and 4-5-star-rated hotels had the same hotel dispersal core. Although Liwan District in Guangzhou was a significant 'HH' district in the pattern of star-rated hotel density (see Tables 3 and 5), investors were less likely to establish 4-5-star-rated hotels. The reason might be that Liwan was not the urban core of the city. Another possible reason could be that the dispersal direction of 4-5-star-rated hotels (west-east corridor) differed from that of star-rated hotels (north-south corridor and west-east corridor).

4.3. Spatial patterns of potential factors

The potential factors were also analyzed by the Local Moran's I (see Figure 8). Taking the urban population size (UrbanPop) as the example, positive (HH) agglomeration effects were observed (see Figure 8a). The positive agglomeration zones were located in the south of the PRD where hotels were also distributed in a positive agglomeration pattern. By comparing Figure 8 and Figure 4 (or Figure 6), the hotels were highly clustered in the positive agglomeration region where the urban population size was relatively high. As spatial autocorrelation is an inherent part of spatial data, it is worth examining the influence of these potential factors in a 'spatial' approach.

4.4. Linear regression modelling

For star-rated hotels, there were three types of counties (or districts), namely HH, HL, and Random. As only one county belonged to the kind of HL, we divided all counties into two categories, HH and others. Table 6 shows the results when using a linear regression model to estimate the coefficients of the variables shown in Table 1 after normalization. Based on correlation analysis between

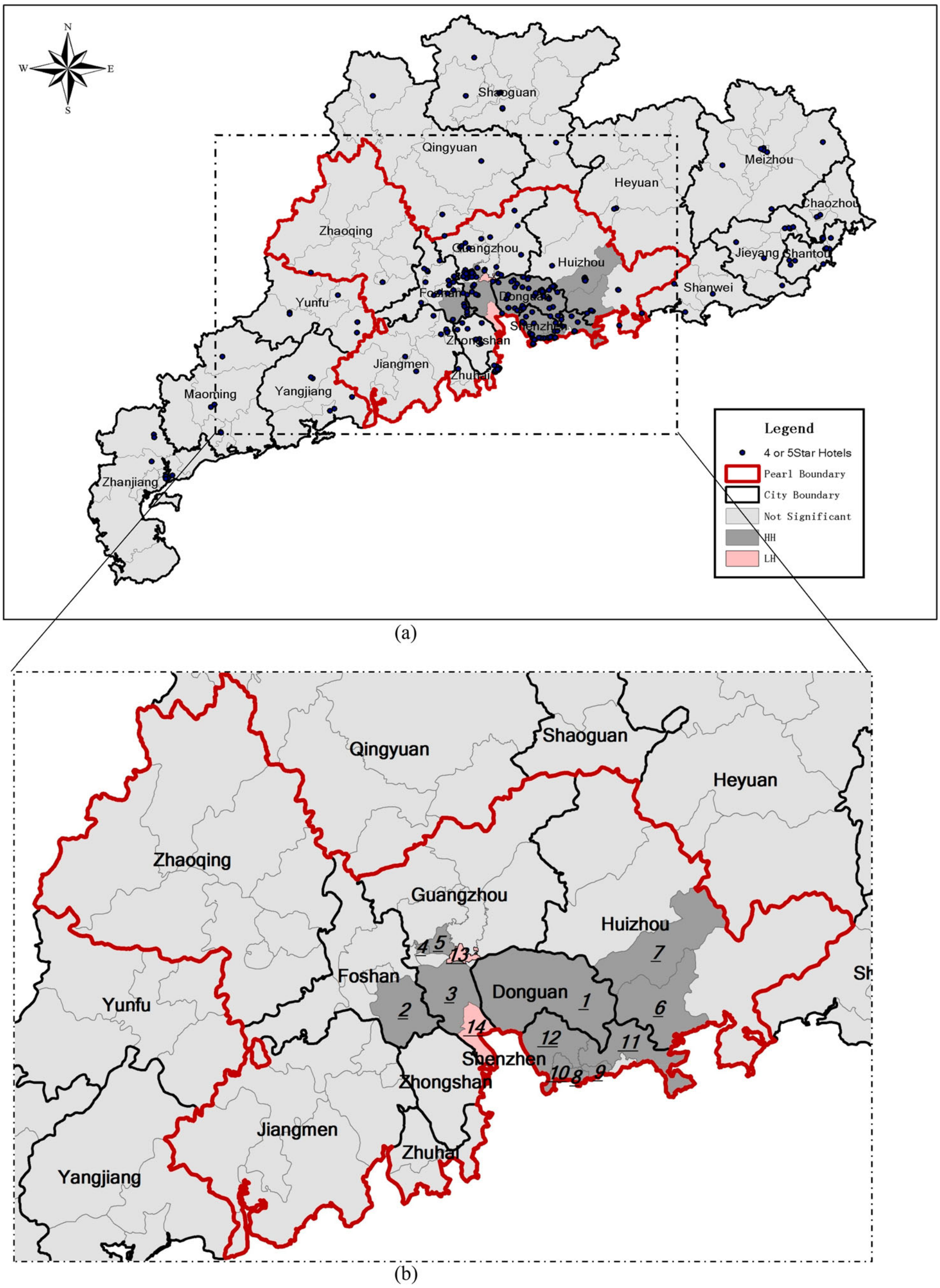

Figure 6. Spatial pattern of 4-5-star-rated hotels in Guangdong Province.

independent variables, some variables were eliminated to deal with the multicollinearity problem. For instance, *Pop, GDP,* and *UrbanPop* were significantly collinear; and *UrbanRatio, pConstruction, PGDP* and *UrbanPInc* were highly correlated. On average, hotel distribution was globally and significantly associated with the number of urban population (*UrbanPop*), the number of attractions (*Attr*), and the road network density (*RoadDen*) when geographic thinking (in this case, agglomeration effect) was ignored. The variable *Attr* was statistically significant, in accordance with previous research (e.g. Arbel & Pizam, 1977; Li et al., 2015; Shoval, 2006). A county with more attractions and

Table 4. Significant counties for 4-5-star-rated hotels in Guangdong Province.

No.	Moran's *I* Index	*Z* Score	*P*-value	Type	Hotel Count	District	City
1	27.434	10.998	<0.001	HH	46	Dongguan	Dongguan
2	4.898	2.765	0.005	HH	11	Shunde	Foshan
3	7.482	3.177	0.001	HH	6	Panyu	Guangzhou
4	7.555	4.739	<0.001	HH	17	Yuexiu	Guangzhou
5	8.361	4.703	<0.001	HH	17	Tianhe	Guangzhou
6	4.738	2.980	0.003	HH	5	Huiyang	Huizhou
7	4.039	2.284	0.022	HH	5	Huicheng	Huizhou
8	5.858	4.234	<0.001	HH	11	Futian	Shenzhen
9	5.081	3.194	0.001	HH	9	Luohu	Shenzhen
10	2.680	2.372	0.018	HH	7	Nanshan	Shenzhen
11	8.010	4.506	<0.001	HH	6	Longgang	Shenzhen
12	18.595	10.431	<0.001	HH	10	Baoan	Shenzhen
13	−6.191	−3.144	0.002	LH	0	Huangpu	Guangzhou
14	−2.966	−2.117	0.034	LH	1	Nansha	Guangzhou

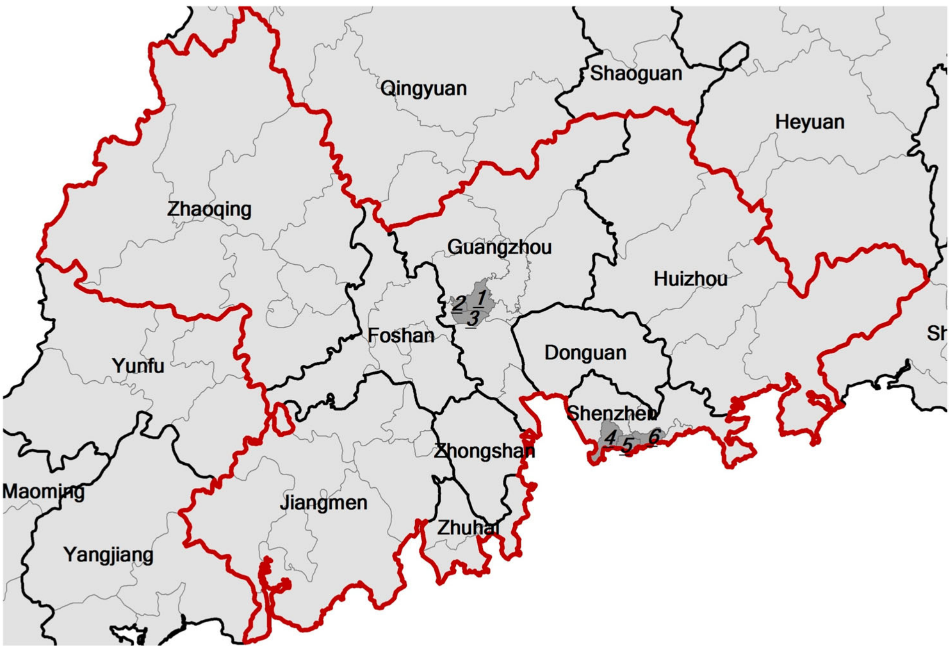

Figure 7. Spatial pattern of 4-5-star-rated hotel density in the Pearl River Delta.

Table 5. Significant counties for 4-5-star-rated hotel density in Guangdong Province.

No.	Moran's *I* Index	*Z* Score	*P*-value	Type	Hotel Density	District	City
1	21.579	15.128	<0.001	HH	0.11	Tianhe	Guangzhou
2	23.307	18.297	<0.001	HH	0.43	Yuexiu	Guangzhou
3	6.989	4.919	<0.001	HH	0.04	Haizhu	Guangzhou
4	2.022	2.269	0.023	HH	0.04	Nanshan	Shenzhen
5	7.191	7.090	<0.001	HH	0.13	Futian	Shenzhen
6	7.377	5.809	<0.001	HH	0.10	Luohu	Shenzhen

transportation facilities was more likely to have a more significant number of star-rated hotels, which provided services for leisure and sightseeing tourists (Yang et al., 2012). Hotels tend to be located in counties with a larger urban population, due probably to the better services urban area could

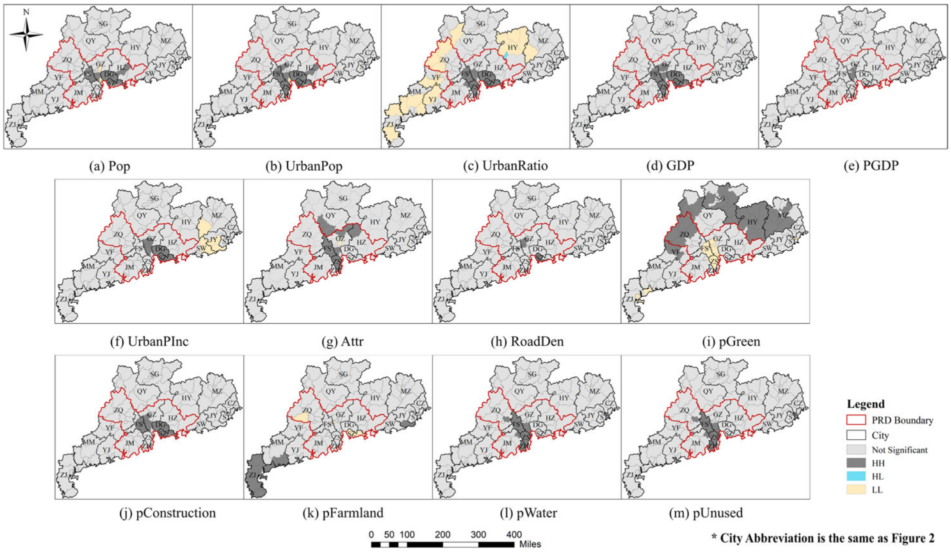

Figure 8. Spatial pattern of independent variables in Guangdong Province.

Table 6. Estimation results of linear regression in the whole study area[a].

	Star-rated hotel		4-5-star-rated hotel	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
(Constant)	−0.333	0.178	−0.074	0.111
UrbanPop	0.630	0.086***	0.669	0.054***
UrbanRatio	0.019	0.044	−0.022	0.028
Attr	0.176	0.047***	0.088	0.029**
RoadDen	0.299	0.089**	0.179	0.056**
pGreen	0.105	0.063	0.069	0.040
pFarmland	0.059	0.069	0.034	0.043
pUnused	0.068	0.077	0.081	0.048
UrbanPInc	0.115	0.074	0.072	0.046

[a]Significant at 0.1 level; *Significant at 0.05 level; **Significant at 0.01 level; ***Significant at 0.001 level.

provide. Another possible reason might be that urban areas could attract more business travellers who need accommodation.

Table 7 shows the linear regression results in the agglomeration area. In the HH area, the independent variables *UrbanPop*, *RoadDen*, and *Attr* are not statistically significant in the HH type areas (agglomeration areas) while the *UrbanPInc* are significant (see Table 7). The agglomeration zones were located in areas with large urban population size and hight road density. The result reflects that public services such as traffic accessibility might not be an important factor for hotels in locations

Table 7. Estimation results of linear regression in agglomeration area[a].

	Star-rated hotel		4-5-star-rated hotel	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
(Constant)	−0.515	0.416	−0.099	0.305
UrbanPop	0.159	0.333	0.472	0.244
Attr	0.562	0.314	0.369	0.230
pGreen	0.268	0.208	0.12	0.152
pUnused	0.195	0.277	0.001	0.203
UrbanPInc	0.682	0.286*	0.405	0.209

[a]Significant at 0.1 level; *Significant at 0.05 level; **Significant at 0.01 level; ***Significant at 0.001 level.

Table 8. Estimation results of linear regression in other areas[a].

	Star-rated hotel		4-5-star-rated hotel	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
(Constant)	0.105	0.221	−1.075	0.292***
UrbanPop	0.413	0.1***	0.395	0.133**
UrbanRatio	0.093	0.067	0.180	0.089*
Attr	0.259	0.076**	0.201	0.101*
RoadDen	−0.144	0.139	0.545	0.184**
pGreen	0.024	0.097	0.39	0.128**
pFarmland	−0.027	0.105	0.361	0.139*
pUnused	0.191	0.105	0.381	0.139**
UrbanPInc	0.104	0.07	0.149	0.093

[a]Significant at 0.1 level; *Significant at 0.05 level; **Significant at 0.01 level; ***Significant at 0.001 level.

with good services (e. g. convenient transportation) in urban areas. Hotels were likely to be located in more developed areas where people had a higher socio-economic status.

Table 8 shows the linear regression results in other areas. The variables *UrbanPop* and *Attr* are positively and significantly associated with the number of star-rated hotels, in accordance with the global model. For the 4-5-star-rated hotels, the coefficients of all variables expect *UrbanPInc* were positive and significant. Counties with more green, farm, or unused lands attracted more high-rated hotels. It could be attributed to the natural tourism resources in these regions. For instance, in other areas, there were a significant number of mountainous counties with springs. The unique tourism resource attracted many 4-5-star-rated hotels.

4.5. Discussion

When analyzing an entity with spatial characteristics, one cannot neglect the effects of spatial autocorrelation. The effects of spatial autocorrelation, particularly spatial agglomeration (e.g. Canina et al., 2005; Chung & Kalnins, 2001; Urtasun & Gutiérrez, 2006b), can always be detected among hotels. This practice violates the underlying assumption of many traditional (non-spatial) methods that observations are independent. In addition, surrounding environmental factors, as depicted in Section 4.3, were also autocorrelated, resulting in considerably complicated situations. The failure to account for spatial effects in statistical models may lead to inappropriate conclusions. For instance, Urtasun and Gutiérrez (2006b) argued that the reason behind the longstanding existence of economy hotels in the city centre of Madrid is that the availability of misused and obsolete office blocks and the existence of various subsidies and a favourable planning regime encouraged the conversion of existing building stocks into hotels of this type. However, this situation raised further questions, such as 'Were economy hotels possibly clustered just because they would like to benefit from the agglomeration effects?', 'Will more hotels be established when the number of obsolete office blocks is increased?'. In the case of Guangdong, 'Do hotel clusters in Guangzhou and Shenzhen emerge due to high traffic accessibility?' One may not get the correct answer if a statistical method is employed without filtering the 'spatial attributes' of data. While some researchers neglected spatial autocorrelation, others treated autocorrelation as a bad incident that should be removed from the data through methods such as resampling. However, spatial autocorrelation is evidence of important underlying spatial processes, that is, being part of the data. Removal of spatial context may result in only one side of the story. In the current study, the hotel distribution, as well as important variables, show significant agglomeration in Guangdong Province. Thus, we proposed a framework based on spatial autocorrelation for analyzing hotels at an aggregate analysis by dividing the entire study area into different parts and quantifying the influences of potential factors, respectively.

We established three models, and the results reflect the influences of explanatory variables varied across the study area. Taking the number of attractions as an example, it was statistically significant in the global and the 'other' models, but insignificant in the agglomeration model. This finding was consistent with research by Fang et al. (2019), who proposed a 'poached egg' model for delineating the spatial heterogeneity in influential factors. Our results reflect that their theory cannot only hold at a city but can also be applied to a hotel study at a regional scale.

5. Conclusion, implications, and way forward

This study contributes to the current theoretical research by analyzing hotel phenomena with local spatial thinking at a macro scale. Factors such as the urban population, road density, salary, or attractions do not always affect hotel location choices. They usually vary across regions. In this research, in the hotel agglomeration area, most factors were not significantly associated with the number of star-rated hotels. It reflected that hotels clustered in positive agglomeration areas could because they would like to receive benefits from agglomeration effects. Hence, one can choose a location for establishing a new hotel in one of these agglomeration regions by taking advantage of the agglomeration effects. The hotel industry in these regions would not decline in the short term if the development condition were stable. The development of hotels in a positive agglomeration zone can influence the development of hotels in neighbouring zones as well. This research also pointed out that geographical thinking is desirable for a better understanding of the influence of potential factors. For example, the comparison results in this research indicated that if the study area was not divided according to spatial autocorrelation, the results of linear regression indicated that attractions play a major role in determining the location of hotels in Guangdong Province. However, in the agglomeration area, the impact of attraction count was not statistically significant.

In this study, irregular cells (administrative units) were used for aggregating hotels. Hence, the investigation of hotel density is important in controlling area variations of counties (or districts). The results demonstrate that indicators could provide useful information and help understand hotel patterns. When comparing star-rated hotel count and density patterns, we proposed a 'hotel dispersal' conjecture, which was inferred from the observed pattern in 2013. If time-series data are available, the process of dispersal can be examined. The results may then shed light on the development of the hotel industry in transforming societies.

This research developed a new framework to study hotel locations in relation to the autocorrelation problem. The research framework can be extended to other hotel studies at different scales. In particular, investigations of spatial autocorrelations should be conducted as the first step in this framework because of the importance of geographical thinking. This framework can be applied to the prediction of the future development of hotel locations or mining potential impact factors affecting hotel location and distribution.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work described in this paper was fully supported by grants from the National Key Research and Development Program of China (grant number 2016YFC0502700), the National Natural Science Foundation of China, China (Project No. 41301423) and Zhejiang Provincial Natural Science Foundation of China under grant number LQ17D010005.

References

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Arbel, A., & Pizam, A. (1977). Some determinants of urban hotel location: The tourists' inclinations. *Journal of Travel Research*, *15*(3), 18–22. https://doi.org/10.1177/004728757701500305

Audretsch, D. B., & Lehmann, E. E. (2005). Does the knowledge spillover theory of entrepreneurship hold for regions? *Research Policy*, *34*(8), 1191–1202. https://doi.org/10.1016/j.respol.2005.03.012

Barros, C. P. (2005). Evaluating the efficiency of a small hotel chain with a Malmquist productivity index. *International Journal of Tourism Research*, *7*(3), 173–184. https://doi.org/10.1002/jtr.529

Baum, J. A., & Haveman, H. A. (1997). Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898–1990. *Administrative Science Quarterly*, *42*, 304–338. https://doi.org/10.2307/2393922

Bloomfield, G. T. (1996). Lodging at the interchange in London, Ontario. *The Canadian Geographer*, *40*(2), 173–180. https://doi.org/10.1111/j.1541-0064.1996.tb00444.x

Broadway, M. J. (1993). Montreal's changing tourist landscape. *Canadian Journal of Urban Research*, *2*(1), 30–48.

Bull, P. J., & Church, A. (1994). The geography of employment change in the hotel and catering industry of Great Britain in the 1980s: A subregional perspective. *Regional Studies*, *28*(1), 13–25. https://doi.org/10.1080/00343409412331348036

Canina, L., Enz, C. A., & Harrison, J. S. (2005). Agglomeration effects and strategic orientations: Evidence from the U.S. Lodging industry. *Academy of Management Journal*, *48*(4), 565–581. https://doi.org/10.5465/amj.2005.17843938

Chung, W., & Kalnins, A. (2001). Agglomeration effects and performance: A test of the Texas lodging industry. *Strategic Management Journal*, *22*(10), 969–988. https://doi.org/10.1002/smj.178

Dökmeci, V., & Balta, N. (1999). The evolution and distribution of hotels in Istanbul. *European Planning Studies*, *7*(1), 99–109. https://doi.org/10.1080/09654319908720504

Egan, D. J., Chen, W., & Zhang, Y. (2006). The intra-urban location of hotels in the Chinese cities of Beijing, Shanghai & Shenzhen. *China Tourism Research*, *2*(4), 516–530.

Fang, L., Li, H., & Li, M. (2019). Does hotel location tell a true story? Evidence from geographically weighted regression analysis of hotels in Hong Kong. *Tourism Management*, *72*, 78–91. https://doi.org/10.1016/j.tourman.2018.11.010

Ferreira, S. L., & Boshoff, A. (2013). Post-2010 FIFA Soccer World Cup: oversupply and location of luxury hotel rooms in Cape Town. *Current Issues in Tourism*, Epub ahead of print 19 March 2013.

Freedman, M. L., & Kosová, R. (2012). Agglomeration, product heterogeneity and firm entry. *Journal of Economic Geography*, *12*(3), 601–626. https://doi.org/10.1093/jeg/lbr022

Helmers, C. G. (2010). *The importance of ownership form and spatial location in the lodging industry: An empirical analysis of the hotel sector in Texas*. [Unpublished Ph.D Dissertation]. Purdue University, West Lafayette, Indiana.

Li, M., Fang, L., Huang, X., & Goh, C. (2015). A spatial–temporal analysis of hotels in urban tourism destination. *International Journal of Hospitality Management*, *45*, 34–43. https://doi.org/10.1016/j.ijhm.2014.11.005

Luo, H., & Yang, Y. (2013). Spatial pattern of hotel distribution in China. *Tourism and Hospitality Research*, *13*(1), 3–15. https://doi.org/10.1177/1467358413508579

Luo, H., & Yang, Y. (2016). Intra-metropolitan location choice of star-rated and non-rated budget hotels: The role of agglomeration economies. *International Journal of Hospitality Management*, *59*, 72–83. https://doi.org/10.1016/j.ijhm.2016.09.007

Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society, Series B*, *10*, 243–251. https://doi.org/10.1111/j.2517-6161.1948.tb00012.x

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1–2), 17–23. https://doi.org/10.1093/biomet/37.1-2.17

Prayag, G., Landré, M., & Ryan, C. (2012). Restaurant location in Hamilton, New Zealand: Clustering patterns from 1996 to 2008. *International Journal of Contemporary Hospitality Management*, *24*(3), 430–450. https://doi.org/10.1108/09596111211217897

Rogerson, J. M. (2013a). Reconfiguring South Africa's hotel industry 1990–2010: Structure, segmentation, and spatial transformation. *Applied Geography*, *2013*(36), 59–68. https://doi.org/10.1016/j.apgeog.2012.06.004

Rogerson, J. M. (2013b). The economic geography of South Africa's hotel industry 1990–2010. *Urban Forum*, *24*(3), 425–446. https://doi.org/10.1007/s12132-012-9186-5

Short, J. R., Kim, Y., Kuus, M., & Wells, H. (1996). The dirty little secret of world cities research: Data problems in comparative analysis. *International Journal of Urban and Regional Research*, *20*(4), 697–717. https://doi.org/10.1111/j.1468-2427.1996.tb00343.x

Shoval, N. (2006). The Geography of hotels in cities: An empirical validation of a Forgotten model. *Tourism Geographies*, *8*(1), 56–75. https://doi.org/10.1080/14616680500392499

Sund, K. J. (2006). The geographical concentration of hotels in Switzerland and the industry life cycle. *Tourism and Hospitality Planning & Development*, *3*(1), 1–18. https://doi.org/10.1080/14790530600638788

Tsaur, S.-H., & Tzeng, G.-H. (1996). Multiattribute decision making analysis for customer preference of tourist hotels. *Journal of Travel & Tourism Marketing*, *4*(4), 55–69. https://doi.org/10.1300/J073v04n04_04

Urtasun, A., & Gutiérrez, I. (2006a). Hotel location in tourism cities: Madrid 1936–1998. *Annals of Tourism Research*, *33*(2), 382–402. https://doi.org/10.1016/j.annals.2005.12.008

Urtasun, A., & Gutiérrez, I. (2006b). Tourism agglomeration and its impact on social welfare: An empirical approach to the Spanish case. *Tourism Management*, *27*(5), 901–912. https://doi.org/10.1016/j.tourman.2005.05.004

Wall, G., Dudycha, D., & Hutchinson, J. (1985). Point pattern analyses of accommodation in Toronto. *Annals of Tourism Research*, *12*(4), 603–618. https://doi.org/10.1016/0160-7383(85)90080-5

Yang, Y., & Fik, T. (2011). *Agglomeration effects and hotel location: Empirical analysis from major China cities*. 2011 AAG annual meeting, 15 April 2011, Seattle, Washington.

Yang, Y., Luo, H., & Law, R. (2014). Theoretical, empirical, and operational models in hotel location research. *International Journal of Hospitality Management*, *36*, 209–220. https://doi.org/10.1016/j.ijhm.2013.09.004

Yang, Y., Wong, K. K., & Wang, T. (2012). How do hotels choose their location? Evidence from hotels in Beijing. *International Journal of Hospitality Management*, *31*(3), 675–685. https://doi.org/10.1016/j.ijhm.2011.09.003