Computers, Environment and Urban Systems 79 (2020) 101404

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Revealing deep semantic commercial patterns: Insights from urban landscape depiction

Hanfa Xing[a,b], Yuan Meng[a,*]

[a] College of Geography and Environment, Shandong Normal University, Jinan, Shandong 250300, China
[b] School of Geography, South China Normal University, Guangzhou, Guangdong 510631, China

ARTICLE INFO

Keywords:
Urban landscape
Urban function
Commercial pattern
Structural equation model

ABSTRACT

Commercial functions play a significant role in economic growth and sustainable development. Influenced by socioeconomic changes which were investigated based on commercial behaviours using big geo-data, urban environment shows polycentric structures and further leads to more complex commercial landscapes. Thus, research considered environmental consequences of urban growth to better understand the commercialization process. However, only general commercial areas or certain commercial retail locations were investigated with the effects of urban environmental landscapes in existing studies. The decomposition of the commercial patterns such as accommodation and catering functions remains undiscussed. To fill this gap, this paper proposes a framework to decompose commercial patterns based on insight from different urban landscapes. First, we depict urban landscapes using spatial metrics, including the density of points of interest (POIs), percentage of landscape of buildings, landscape shape index of buildings and density of roads. Then, we examine the spatial distribution of the decomposed commercial patterns by delineating potential deep semantic topics. Finally, a structural equation model is utilised to determine the relations among urban landscapes and commercial patterns. The results show that four types of urban landscapes exhibit different spatial patterns, revealing various perspectives of commercial characteristics. Meanwhile, the decomposed commercial patterns, including those for catering and accommodation functions, display a heterogeneous distribution. These commercial patterns are directly affected by various configurations of buildings, POIs and roads. On this basis, suggestions are offered to improve commercial pattern development, including integrating urban landscape construction, organising the commercial pattern distribution and enhancing the harmony between the urban environment and commercial land uses.

1. Introduction

The rapid development of urbanisation has led to the emergence of and shifts in different urban functions, including residential, commercial, public and service uses (Anderson, 1976; Ward, Phinn, & Murray, 2000). Highlighting the increase in the consumer market, commercial functions play a significant role in economic development and improving the quality of urban residents (Berry & Garrison, 1958; Wang, Wang, Zhao, & Fu, 2018). Such functions serve various commercial categories, including accommodations, entertainment, catering, and businesses. Exploring the distribution patterns of certain commercial functions and investigating the corresponding trends are necessary for predicting future changes and supporting urban planning (Porta et al., 2009).

Commercial function patterns have long been discussed by authorities and scholars. Both socioeconomic drivers and environmental consequences of urban growth have been considered to better understand the commercialization process (Leichenko, 2001; Wentz, Stefanov, Gries, & Hope, 2006). Traditionally, delineating commercial functions in detail has been challenging due to heavy reliance on interviews, statistical yearbooks and field surveys (Murphy, 1954; Wolf, Guensler, & Bachman, 2001). For example, Yan, Zhou, and Leng (2000) considered housing prices to estimate the commercial areas in Guangzhou and investigated the traffic conditions for each road to determine the commercial area boundaries; they further applied a field survey approach to detect the central business districts (CBD) in Guangzhou. These conventional approaches are limited to data collection because socioeconomic and environmental data are expensive and time consuming to acquire, and it is difficult to improve their spatial and temporal resolutions (Xing, Meng, & Shi, 2018).

The rise of big geo-data has provided new opportunities to investigate the changes in commercial patterns from the perspective of

* Corresponding author.
E-mail addresses: xinghanfa@sdnu.edu.cn (H. Xing), mengyuan@stu.sdnu.edu.cn (Y. Meng).

https://doi.org/10.1016/j.compenvurbsys.2019.101404
Received 15 June 2019; Received in revised form 2 August 2019; Accepted 6 September 2019
Available online 12 September 2019
0198-9715/ © 2019 Elsevier Ltd. All rights reserved.

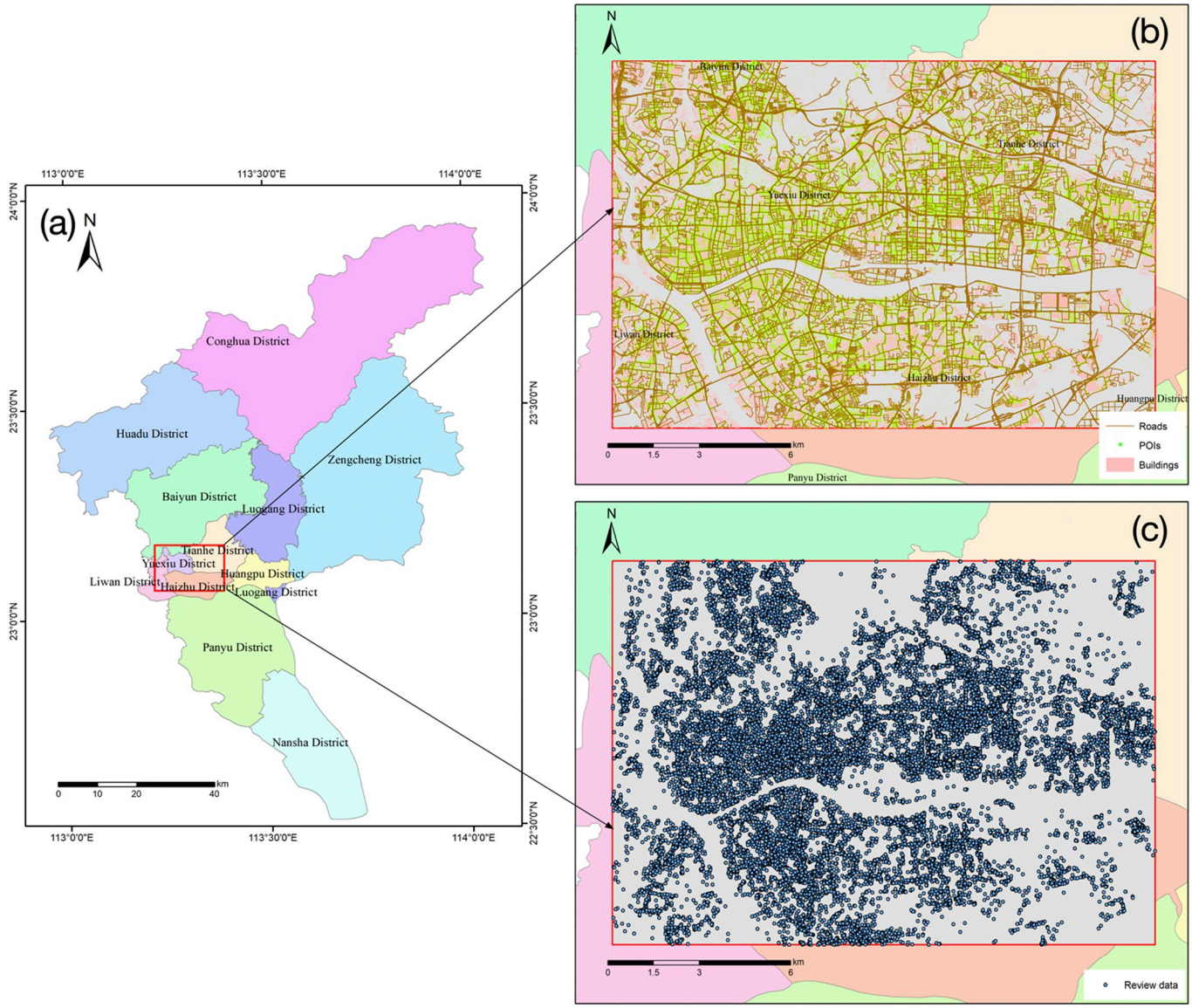


Fig. 1. Study area and data sources. (a) The location of the study area. (b) The spatial distributions of roads, POIs and buildings for depicting the urban landscape. (c) The spatial distribution of review data for revealing commercial patterns.

human activities (Liu et al., 2015). Through movement data, social media data and cell phone data, activity patterns can be automatically extracted corresponding to certain urban functions (Chen et al., 2017; Yuan, Zheng, & Xie, 2012). Lloyd and Cheshire (2017) took geo-tagged Twitter data into consideration to establish retail centres and define catchment areas, and the hypothesis that big geo-data trends corresponded to retail activities was confirmed. Apart from the retail patterns, Wang et al. (2018) utilised the customer count and satisfaction data from social networks to analyse the spatial distribution of urban facilities in Beijing, China, based on weighted network-constrained kernel density estimation. Hu, Bai, Wang, and Ai (2019) focused on the detection of commercial area boundaries by proposing an exploratory spatial analysis method using Weibo check-in data. On this basis, a time consequence extraction method for commercial areas was proposed to analyse the variational trends of commercial patterns. Siła-Nowicka and Fotheringham (2019) used GPS data to derive human activity flows from the perspective of shopping behaviours. Compared with traditional approaches, the utilisation of big geo-data has been considered as an alternative freely available resource that effectively captures commercial behaviours.

Influenced by commercial behaviours, more complex commercial landscapes have been generated with structures varying from monocentric to polycentric (Yu, Ai, & Shao, 2015), and this process was followed by the emergence of urban environmental-related big geo-data such as buildings and roads from navigation maps (Liu, Lan, & Xing, 2019). Accordingly, exploring the environmental consequences is inevitable for detecting the variations in commercial patterns. By measuring urban landscapes using spatial metrics, Vanderhaegen and Canters (2017) focused on the configuration of built-up and open space areas for distinguishing various urban functions, including commercial types. Considering the effects of road networks, Lin, Chen, and Liang (2018) investigated the relationship between commercial stores and street centrality by calculating different characteristics, including the closeness, betweenness and straightness. Similarly, Ma et al. (2019) analysed the spatial coupling effects between urban public transport and commercial complexes. In summary, existing studies have investigated the environmental factors to identify commercial patterns, explored the effects of different landscape elements for generating commercial functions, and discussed the challenges associated with developing commercial areas in the context of variable environmental

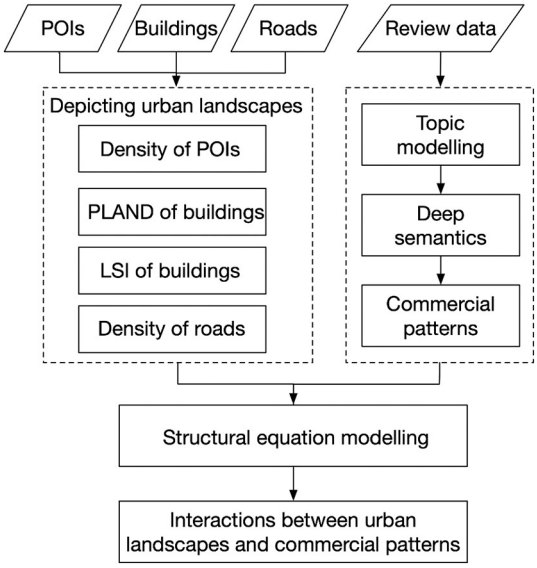


Fig. 2. The overall framework.

consequences. However, most of these studies solely analysed how environmental consequences affect general commercial areas covering all types of commercial functions or specific commercial retail locations. The effects of environmental factors on the decomposition of commercial patterns, such as those related to accommodation and catering functions, have not been investigated in detail.

In this work, we investigate deep semantic commercial patterns by depicting urban landscapes. Certain landscape characteristics that are related to commercial functions are quantified using spatial metrics. Different commercial functions are decomposed from the textual information in review data. Finally, a structural equation model (SEM) is utilised to investigate the potential correlations among urban landscape and commercial pattern features. The contributions of the proposed research can be summarised as follows: 1) Urban environment that corresponds to commercial landscape characteristics is depicted from the perspective of commercial functions; 2) The spatial distribution of certain commercial patterns, including catering and accommodation functions, is obtained by modelling topics from review data; 3) The relationships of urban landscapes and commercial patterns are investigated by establishing the SEM framework; 4) Several findings and suggestions are proposed for organising urban environments and commercial land uses to support city planning.

The remainder of this paper is organised as follows. Section 2 introduces the study area and data sources used in this study. The methodology and results are presented in Section 3 to illustrate how urban landscapes and commercial patterns are interconnected. Further discussion of the findings and suggestions based on this research are given in Section 4. Section 5 concludes the study and discusses future works.

2. Study area and data sources

The study area is located in the central downtown region of Guangzhou in Guangdong Province, with a total area of 213.6 km^2, as shown in Fig. 1(a). Guangzhou has become the political, economic and cultural centre of southern China. Encompassing Yuexiu District and partial areas of Tianhe District, Haizhu District and Liwan District, this area contains complex urban structures with various building, road and urban facility morphologies. Moreover, with economic growth and urban development, more commercial facilities have become available for public services and further shaped the different spatial distributions of commercial patterns.

To delineate the landscapes in the study area, we use vector datasets, including building, road and point of interest (POI) datasets, from AutoNavi. Displayed in Fig. 1(b), the building datasets describe the specific location and shape of each independent building from the top view. The multicategory POIs are used to indicate urban facilities for public services. To meet this goal, we excluded POIs described as natural places or with administrative names, such as those POIs categorised as mountains or villages. The road datasets containing main roads, secondary roads and pedestrian roads were used to organise and divide buildings into different regions and connect urban facilities with various distributions; thus, they represent the main structure of the city.

To represent commercial patterns from a deep semantic perspective and investigate the spatial distributions of and semantic variations in commercial patterns, review data were considered. Review data are widely available from many websites, such as Yelp and Dianping. These data provide the locations of commercial sites and relevant consumer comments. In this study, a review dataset with 784,300 records was obtained from AutoNavi. In particular, these review data are for commercial sites, including restaurants and hotels, as shown in Fig. 1(c).

It should be noted that the abovementioned data sources, including buildings, POIs, roads and review data, were all collected in the year 2018. Considering the rapid growth of the economy, this time consistency ensures the accuracy and certainty of this research.

3. Methodology and results

By depicting the spatial configuration of the urban landscape, this study proposes a new perspective for revealing semantic commercial patterns. The overall framework is schematically illustrated in Fig. 2. First, urban landscapes are depicted by quantifying vector datasets, including POIs, buildings and roads, using spatial metrics. The details of this process are given in Section 3.1. Second, commercial patterns are semantically decomposed from review data using topic modelling in Section 3.2. Third, interactions among urban landscapes and commercial patterns are analysed based on structural equation modelling, as discussed in Section 3.3.

3.1. Urban landscape depiction

The spatial configuration of the urban landscape is not identical from place to place. To measure this variation, the study area is divided into 200 m*200 m regular grids. By treating each grid as an independent landscape, we quantify the urban landscapes constituted by different POIs, buildings and roads by means of spatial metrics (Section 3.1.1) and then analyse the spatial distribution of urban landscapes among regular grids (Section 3.1.2).

3.1.1. Depicting urban landscapes using spatial metrics

With a focus on measuring the typical features of urban landscapes, we consider POIs, buildings and roads as independent landscape classes and propose class-level spatial metrics. In particular, four types of spatial metrics are considered in Table 1.

3.1.1.1. Density of POIs. We selected the density of POIs to highlight the spatial distributions of various urban facilities. The selection is based on the hypothesis that urban facilities, especially those that serve commercial functions, play an important role in constituting commercial patterns (Wang et al., 2018). Thus, the capability of providing urban facilities in each grid is determined. It should be noted that apart from the density of POIs, population density is also a key characteristic that depicts commercial patterns (Long & Huang, 2017). Since the aim of this research focuses on the perspective of urban landscapes, population density is not considered.

3.1.1.2. Percentage of landscape of buildings (PLAND of buildings). Buildings are one of the basic components for describing urban morphology, modelling urban areas and distinguishing urban

Table 1
Spatial metrics for depicting urban landscapes.

Spatial metrics	Abbreviation	Equation	Descriptions
Density of POIs	*DP*	$DP = \frac{N_i}{A_i}$	N_i refers to the number of POIs in each grid; A_i is the ith grid.
PLAND of buildings	*PLAND*	$PLAND = \frac{\sum_{j=1}^{n} b_{ij}}{A_i}$	b_{ij} refers to the area of the jth building in the ith grid; n refers to the total number of buildings in each grid.
LSI of buildings	*BLSI*	$BLSI = \frac{\sum_{j=1}^{n} l_{ij}}{2\sqrt{\pi A_i}}$	l_{ij} refers to the perimeter of the jth building in the ith grid.
Density of roads	*DR*	$DR = \frac{\sum_{k=1}^{n} r_{ik}}{A_i}$	r_{ik} refers to the length of the kth road in the ith grid.

functions. The variable PLAND of buildings is proposed based on the modification of the traditional spatial metric PLAND (McGarigal & Marks, 1995). With this new variable, we measure the percentage of buildings in each grid.

3.1.1.3. Landscape shape index of buildings (LSI of buildings). With regard to measuring the spatial heterogeneity of buildings, the LSI of buildings was proposed by Xing and Meng (2018), and this metric is utilised in this study. This metric was designed based on the conventional landscape shape index (Patton, 1975; Rutledge, 2003). Regardless of the other landscapes, the LSI of buildings focuses solely on buildings, with more complex building distributions having much larger values.

3.1.1.4. Density of roads. According to the research of Zhu, Wang, Wu, and Liu (2017), we assume that road systems provide a significant perspective for linking city construction areas and urban functions. Thus, to capture the spatial characteristics of roads, we measure the density of roads, namely, the ratio of the road length to the area of a regular grid cell.

3.1.2. Distribution of the urban landscape analysis

Fig. 3 displays the spatial distributions of four types of urban landscapes. The four types of urban landscapes exhibit different spatial patterns, with high values represented by red-filled grids and low values represented by blue-filled grids. Density of POIs is shown in Fig. 3(a). The high values are distributed in sparse clusters, and the average and low values occupy most of the area. From visual inspection, the use of urban facilities delineates the potential region of the commercial pattern, as shown in yellow- and red-filled grids, but detailed landscapes that correspond to some commercial patterns are missing. This landscape depiction is improved by adding metrics in Fig. 3(b)-(d). Compared with the density of POIs, PLAND of buildings exhibits dense clusters spread across the study area, indicating that buildings are occupied more than other open areas and facilities. Additionally, the LSI of buildings reflects the spatial heterogeneity of buildings. In particular, buildings with complex distribution patterns may not occupy larger areas in each grid and thus reveal more detailed characteristics of the internal landscape. Moreover, the density of roads displays a distinct spatial distribution with high-value grids spreading and connecting different landscapes.

3.2. Deep semantic commercial pattern decomposition

To decompose commercial patterns with specific urban functions, deep semantic topics can be effectively modelled using the latent Dirichlet allocation (LDA) topic model (Blei, Ng, & Jordan, 2003) according to review data. On this basis, deep semantic topics are associated with certain commercial functions, and the corresponding spatial distributions are further analysed.

3.2.1. Modelling deep semantic topics from review data

Treating the textual information in review data as a corpus $\boldsymbol{C}$ at the grid scale, documents $\boldsymbol{D}$ are established from the corpus $\boldsymbol{C}$ in each grid. Based on the LDA hypothesis, each grid includes one document d that contains N topics with different proportions, and each topic is associated with a set of words $\boldsymbol{w}$ from corpus $\boldsymbol{C}$ with a variety of distributions. The corresponding equation is shown as follows:

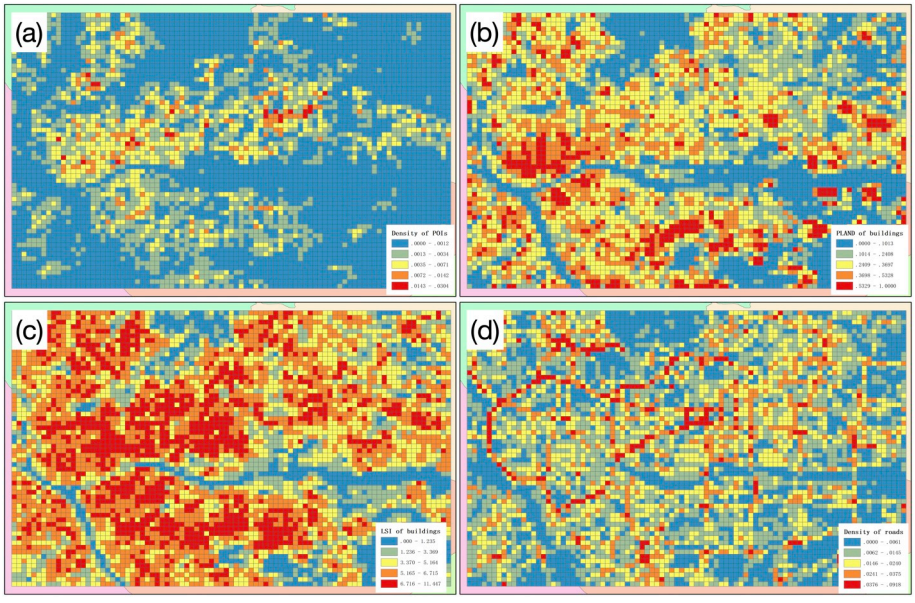


Fig. 3. Spatial distributions of four types of urban landscapes. (a) Density of POIs; (b) PLAND of buildings; (c) LSI of buildings; (d) Density of roads.

$$p(\theta, \boldsymbol{z}, \boldsymbol{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta) \quad (1)$$

where $\boldsymbol{z}$ represent the deep semantic topics, α and β refer to the smoothing parameters that are selected according to the parameter selection scheme proposed by Griffiths and Steyvers (2004), and $\boldsymbol{\theta}$ indicates the probability that document d contains the deep semantic topics.

In general, we utilise deep semantic topics to define commercial patterns. In the LDA model, $k = 2, \ldots\ldots, 50$ topics can be utilised. On this basis, the value of k is determined using expert knowledge, for which topics provide maximum information and minimal dimensionality. In addition, each topic represents specific functions, such as food, night-time entertainment and accommodations, related to commercial patterns based on words with high probabilities of occurrence.

3.2.2. Detecting deep semantic commercial patterns

Based on expert knowledge, we chose $k = 10$ as the topic number. Among the 10 topics, 6 topics were manually interpreted to represent specific functions. In particular, topics 2, 4, 5, and 7 reflect catering-related commercial patterns, and topics 3 and 9 are related to accommodations and commercial living conditions. Table 2 presents some of the word distributions for the selected topics. Words such as “Taste”, “Delicious”, “Buy”, and “Porridge” are displayed with higher probabilities for topics 2, 4, 5, and 7, whereas words such as “Hotel”, “Room”, “Convenience”, and “Clean” have higher probabilities of being used for topics 3 and 9. To detect commercial patterns with certain functions, we utilised topics 2, 4, 5, and 7 to infer catering-related commercial patterns and topics 3 and 9 to represent accommodation-related commercial patterns.

Based on the probabilities of topics in each grid, we display the spatial distribution of the selected topics in Fig. 4. Because the word probability for each topic is not the same, even for topics with similar commercial functions, the spatial distributions vary from topic to topic. Fig. 4(a)–(d) represent the spatial distribution of catering-related topics. From visual inspection, catering-related commercial patterns are usually concentrated as clusters, in which neighbourhood grids contain similar values. Thus, restaurant and commercial food sites have a tendency to be grouped to improve business. This finding suggests that these commercial sites usually exist as business districts and malls. In contrast, the accommodation-related commercial patterns displayed in Fig. 4(e)-(f) indicate that grids with high values are usually separately distributed from other grids. Thus, we can conclude that compared with catering-related commercial sites, accommodation-related commercial sites benefit more from having a sparse distribution than a clustered grouping pattern.

3.3. Revealing deep semantic commercial patterns from urban landscapes

Utilising 4 spatial metrics and 6 deep semantic topics as key variables, SEM is introduced in this study to investigate the relations between urban landscapes and commercial patterns. The basic approach and design of the SEM framework are detailed in Section 3.3.1. On this basis, the characteristics of the relations between urban landscapes and commercial patterns are addressed in Section 3.3.2.

3.3.1. Designing a structural equation model

Integrating path analysis and confirmatory factor analysis, the SEM is utilised to investigate the relationships among urban landscapes and commercial patterns. The SEM uses path analysis, confirmatory factor analysis and statistical testing methods to explore the relationships between different variables (Doloi, Iyer, & Sawhney, 2011; Hellier, Geursen, Carr, & Rickard, 2003). In general, observed variables and latent variables constitute the basic variables of the SEM. As shown in Table 3, the observed variables represent the variables that are directly measured, including the density of POIs, PLAND of buildings, LSI of buildings, and density of roads, as well as the calculated topics. Latent variables are indirectly measured variables, including the urban landscapes and commercial patterns. In other words, latent variables describe characteristics that cannot be directly quantified. On this basis, latent variables can be divided into exogenous latent variables and endogenous latent variables, with the former considered explanatory variables that impact other variables and the latter being affected by exogenous variables and other endogenous variables. In this study, urban landscapes are considered exogenous latent variables, and both catering-related and accommodation-related patterns are considered endogenous latent variables. Based on the proposed variables, the SEM comprises a measurement model and a structural model. Specifically, the measurement model is composed of observed variables and latent variables, and the structural model measures the correlation among latent variables.

The measurement model is constructed based on the following equations:

$$X = \bigwedge_x \xi + \delta \quad (2)$$

Table 2
Some of the word distributions for selected topics. High values indicate that a word has a high possibility of being used.

	Topic 2	Topic 3	Topic 4	Topic 5	Topic 7	Topic 9
Taste	0.01740102	0.00000090	0.00876305	0.02299370	0.03158083	0.00245172
Delicious	0.01587106	0.00053591	0.00266687	0.00835030	0.02094536	0.00000006
Buy	0.00917558	0.00021622	0.00284767	0.00024186	0.00074915	0.00079199
Porridge	0.00000008	0.00004528	0.00480664	0.00099374	0.00058960	0.00000006
Environment	0.00473854	0.01141356	0.00067809	0.02423066	0.01046995	0.01755557
Service	0.00763714	0.01689758	0.00000010	0.01744861	0.00956580	0.01004761
Dish	0.00000008	0.00000008	0.00060885	0.01072644	0.00772390	0.00000006
Waiter	0.00342254	0.00208833	0.00000010	0.00956452	0.00487500	0.00177723
Purchase	0.00886569	0.00000008	0.00008953	0.00954103	0.00646101	0.00078055
Hotel	0.00000008	0.03773161	0.00000010	0.00032599	0.00000005	0.00662094
Room	0.00000008	0.02658194	0.00000010	0.00086868	0.00000005	0.02785184
Convenience	0.00412978	0.02199288	0.00144167	0.00349949	0.00143002	0.00905875
Clean	0.00051205	0.01984792	0.00281786	0.00041164	0.00175716	0.01227284
Stay	0.00000688	0.01223209	0.00123587	0.00000690	0.00029156	0.01240653
Location	0.00433949	0.01188364	0.00232355	0.00219128	0.00309431	0.01326832
Healthy	0.00072516	0.01057447	0.00154458	0.00014408	0.00102662	0.01315630

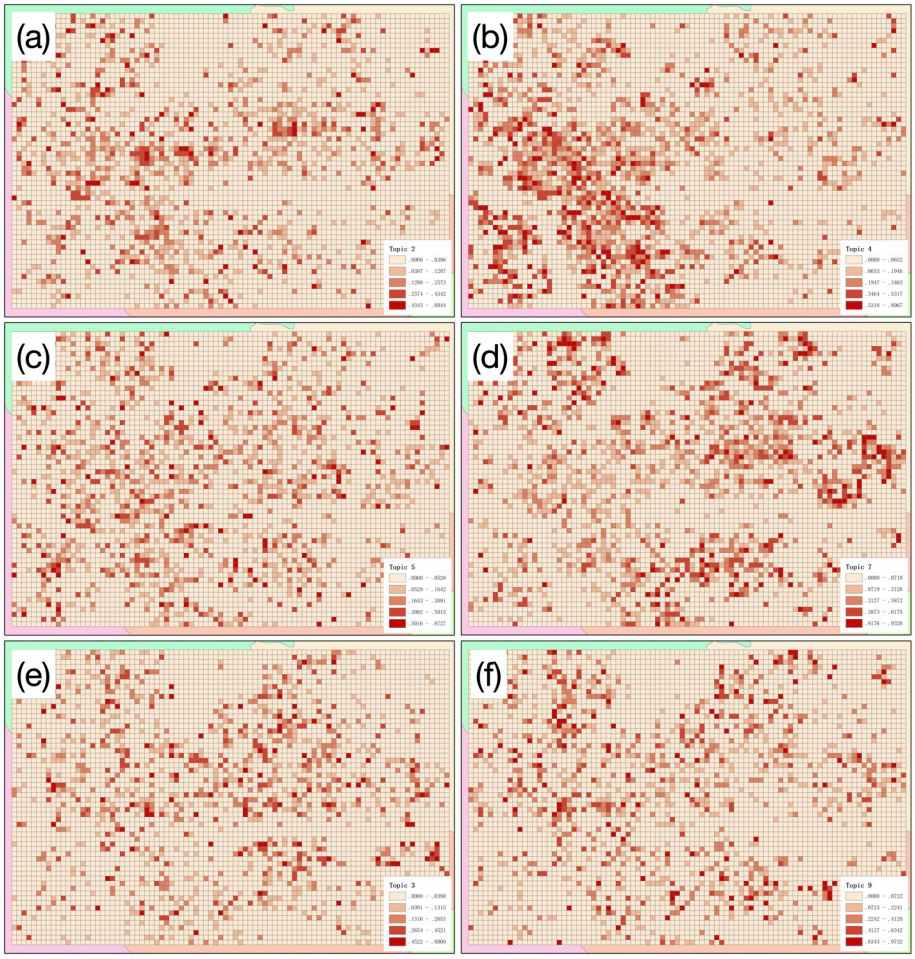


Fig. 4. Spatial distribution of topics for revealing commercial patterns. (a)-(d) Distributions for catering-related topics; (e)-(f) Distributions for accommodation-related topics. (a) Distribution of topic 2; (b) Distribution of topic 4; (c) Distribution of topic 5; (d) Distribution of topic 7; (e) Distribution of topic 3; (f) Distribution of topic 9.

Table 3
Variables in the structural equation model.

Latent variables	Observed variables
Urban landscapes	Density of POIs
	PLAND of buildings
	LSI of buildings
	Density of roads
Catering-related commercial patterns	Topic 2
	Topic 4
	Topic 5
	Topic 7
Accommodation-related commercial patterns	Topic 3
	Topic 9

$$Y = \bigwedge_y \eta + \varepsilon \tag{3}$$

where X and Y represent the observed variables, including the calculated spatial metrics and topics. ξ and η are the exogenous variables and endogenous variables, respectively, reflecting the urban landscapes and both catering-related and accommodation-related patterns in the proposed study. In addition, $\bigwedge_x$ and $\bigwedge_y$ are the factor loading matrices of X and Y, respectively. δ and ε are the measurement errors of X and Y.

Additionally, the structural model is formed as follows:

$$\eta = \mathrm{B}\eta + \Gamma + \zeta \tag{4}$$

where B and Γ refer to the endogenous and exogenous latent variable coefficient matrices, respectively, and ζ represents the equation error.

Because we aim to investigate the relations among urban landscapes and commercial patterns, the SEM is designed according to the modelling path in Fig. 5. Four spatial metrics, including the density of POIs, PLAND of buildings, LSI of buildings and density of roads, serve as the observed variables to directly assess urban landscapes. Similarly, topics 2, 4, 5, and 7 reflect catering-related commercial patterns, and topics 3 and 9 constitute accommodation-related commercial patterns. On this basis, we assume that urban landscapes have a direct impact on both catering and accommodation-related commercial patterns. In addition, catering-related and accommodation-related commercial patterns have impacts on each other. In Fig. 5, e1-e12 are error variables measured for each observed variable.

3.3.2. Relations among urban landscapes with deep semantic commercial patterns

Based on the above-designed SEM, a model that investigates the relations among urban landscapes and commercial patterns was constructed. To revise the constructed model by including the covariance of error variables, a check of the evaluation indices, including χ^2, df, RMSEA, GFI, AGFI, NFI, RFI, IFI, TLI, CFI, PGFI, PNFI and PCFI (Hooper, Coughlan, & Mullen, 2008), was proposed. According to the reference values and model values in Table 4, except for PGFI, which is close to the reference value, the other indices meet the requirements of model assessment. Therefore, the constructed model effectively fits the experimental data.

Fig. 6 displays the paths and parameters of the constructed model. Overall, the modelling results support the hypothesis in Fig. 5 from four

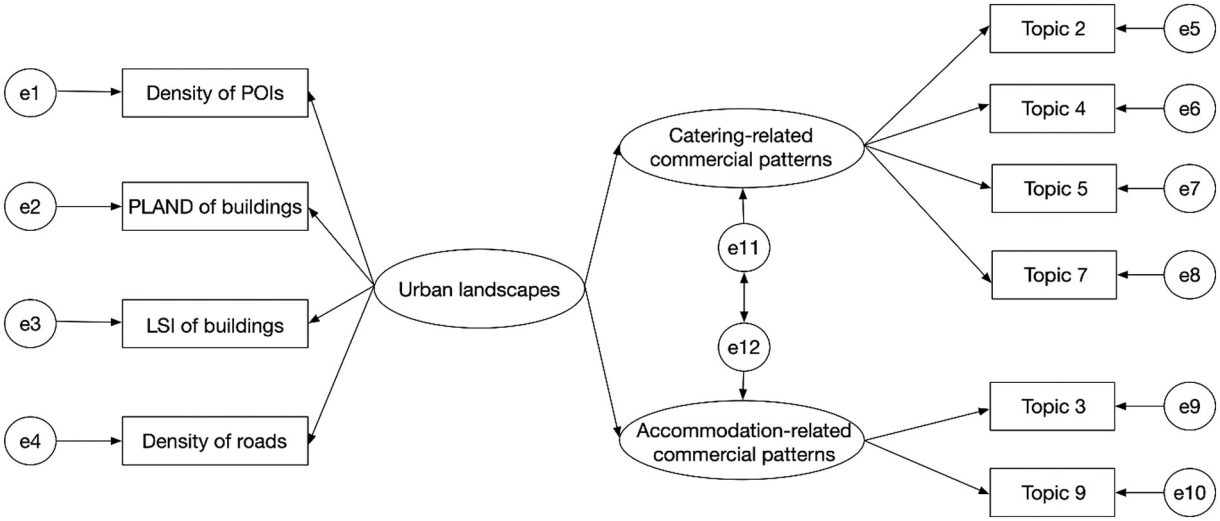


Fig. 5. Structural equation model design for urban landscapes and commercial patterns.

Table 4
Model fit indices for assessing the structural equation model.

Model fit indices	Reference value	Model value
Chi-square (χ^2)	–	51.341
Degree of freedom (df)	–	23
Root mean square error of approximation (RMSEA)	< 0.08	0.015
Goodness of fit index (GFI)	> 0.90	0.998
Adjusted goodness of fit index (AGFI)	> 0.90	0.995
Normed fit index (NFI)	> 0.90	0.994
Relative fit index (RFI)	> 0.90	0.998
Incremental fit index (IFI)	> 0.90	0.997
Tucker-Lewis index (TLI)	> 0.90	0.993
Comparative fit index (CFI)	> 0.90	0.997
Parsimony goodness of fit index (PGFI)	> 0.50	0.417
Parsimony normed fit index (PNFI)	> 0.50	0.508
Parsimony comparative fit index (PCFI)	> 0.50	0.509

perspectives. 1) The urban landscape has a direct impact on the density of POIs, PLAND of buildings, LSI of buildings and density of roads, with standardised estimates of 0.53, 0.63, 0.69, and 0.29, respectively. 2) Catering-related commercial patterns have a direct impact on topics 2, 4, 5 and 7, with standardised estimates of 0.28, 0.42, 0.19, and 0.29, respectively. 3) Accommodation-related commercial patterns have a direct impact on topics 3 and 9, with standardised estimates of 0.58 and 0.70, respectively. 4) Urban landscapes have a direct impact on catering-related and accommodation-related commercial patterns, with standardised estimates of 0.71 and 0.24, respectively.

Further illustrations of quantifying the interactions among urban landscapes and commercial patterns are shown in Table 5. With respect to each variable path, an unstandardised estimate, the standard error (S.E.), the critical ratio (C.R.), the significance (p) and a standardised estimate are calculated. Specifically, all values of C.R. are > 1.96, reaching the 0.05 significance level, and the corresponding p values reach the 0.001 significance level. This result confirms the validity of the paths and parameters generated from the theoretical assumptions. Moreover, we can conclude that urban landscapes have a much greater impact on catering-related commercial patterns and a lesser impact on accommodation-related commercial patterns with standardised estimates of 0.71 and 0.24. In fact, the organisation of urban facilities and buildings potentially influences the choice of catering locations, while accommodations are affected more by the purpose of travelling or business. For example, people tend to choose a restaurant near various urban facilities for the convenience of shopping and consider a hotel for the ability to travel to scenic locations. According to the variable paths,

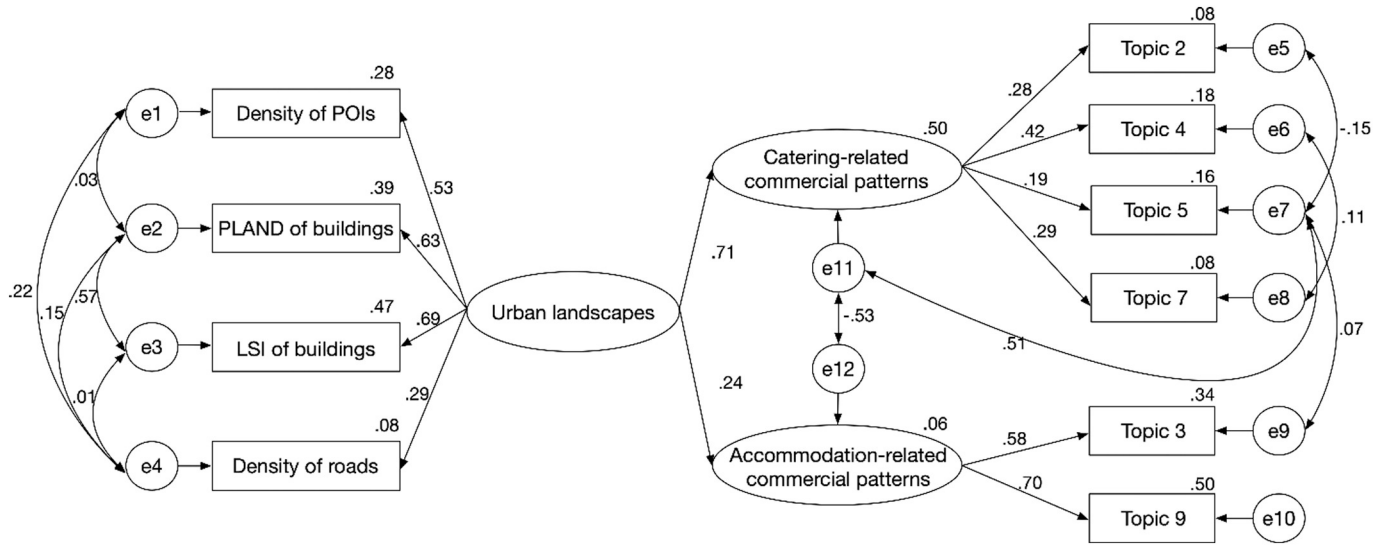


Fig. 6. SEM paths and parameters for assessing urban landscapes and commercial patterns.

Table 5
SEM paths and regression weights for revealing the relationships among urban landscapes and commercial patterns.

Path			Unstandardised estimate	S.E.	C.R.	*p*	Standardised estimate
LSI of buildings	⟵	Urban landscapes	1.000	–	–	–	0.686
Density of roads	⟵	Urban landscapes	0.425	0.047	9.037	***	0.291
PLAND of buildings	⟵	Urban landscapes	0.911	0.032	28.747	***	0.626
Density of POIs	⟵	Urban landscapes	0.776	0.048	16.179	***	0.533
Topic 2	⟵	Catering-related commercial patterns	0.670	0.062	10.842	***	0.284
Topic 4	⟵	Catering-related commercial patterns	1.000	–	–	–	0.424
Topic 5	⟵	Catering-related commercial patterns	0.452	0.056	8.130	***	0.192
Topic 7	⟵	Catering-related commercial patterns	0.673	0.045	14.924	***	0.285
Topic 3	⟵	Accommodation-related commercial patterns	0.829	0.083	10.015	***	0.584
Topic 9	⟵	Accommodation-related commercial patterns	1.000	–	–	–	0.704
Catering-related commercial patterns	⟵	Urban landscapes	0.436	0.030	14.525	***	0.705
Accommodations-related commercial patterns	⟵	Urban landscapes	0.242	0.028	8.560	***	0.236

*** Indicates significance at the 0.001 level.

Table 6
Covariance estimates.

			Estimate	S.E.	C.R.	*p*	Correlation
e11	⟷	e12	−0.110	0.011	−10.037	***	−0.534
e4	⟷	e2	−0.113	0.020	−5.522	***	−0.152
e3	⟷	e2	0.324	0.029	11.352	***	0.571
e2	⟷	e1	0.018	0.013	1.359	0.174	0.027
e7	⟷	e9	0.053	0.012	4.237	***	0.071
e5	⟷	e7	−0.133	0.019	−6.953	***	−0.152
e7	⟷	e11	0.139	0.017	8.333	***	0.505
e6	⟷	e8	0.098	0.018	5.292	***	0.112
e3	⟷	e4	0.007	0.022	0.337	0.736	0.010
e4	⟷	e1	0.175	0.020	8.941	***	0.216

*** Indicates significance at the 0.001 level.

the LSI of buildings, PLAND of buildings and density of POIs contribute the most to depicting urban landscapes, with standardised estimates of 0.686, 0.626 and 0.533. Thus, it is reasonable that urban landscapes show a greater impact on catering-related rather than accommodation-related commercial patterns.

Tables 6 and 7 display the covariance and variance estimates of the variables, respectively. In Table 6, most of the covariance estimates among variables are significant at the 0.001 level. In particular, the relationships among different variables are displayed through the correlation between each pair of error variables in the range of e1-e12 in Table 6. These error variables are measured for each observed variable, and thus, the correlations between catering-related and accommodation-related commercial patterns and between each pair of spatial metrics can be analysed. For the relationship between catering-related and accommodation-related commercial patterns, the correlation value of −0.53 indicates that the two commercial patterns show significant impacts on each other. However, the covariance between e2 and e3, as well as between e3 and e4, is less significant, with correlations of 0.027 and 0.010, respectively. Accordingly, we conclude that the relationships between the PLAND of buildings and the LSI of buildings, as well as that between the LSI of buildings and density of roads, remain unclear. In addition, the variance estimates of each variable display significance at the 0.001 level in Table 7.

Table 7
Variance estimates.

	Estimate	S.E.	C.R.	*p*
Urban landscapes	0.471	0.033	14.309	***
e11	0.090	0.021	4.259	***
e12	0.468	0.049	9.461	***
e3	0.529	0.030	17.449	***
e4	0.915	0.024	38.167	***
e2	0.609	0.033	18.332	***
e1	0.716	0.022	32.441	***
e5	0.919	0.020	45.886	***
e6	0.820	0.026	31.245	***
e7	0.838	0.025	33.894	***
e8	0.918	0.021	44.318	***
e9	0.659	0.036	18.299	***
e10	0.504	0.050	10.090	***

*** Indicates significance at the 0.001 level.

4. Discussion

This research statistically describes urban landscapes and commercial patterns and further demonstrates the relations among urban landscapes and commercial patterns, emphasising that the spatial characteristics of urban facilities, buildings and roads are strongly correlated with the commercial functions of catering and accommodation uses. The findings reveal that the presence of urban landscapes contributes to the commercial organisation to some extent. Moreover, we deduce that optimising the existing urban environment can guide commercial function allocation and support city planning. On this basis, several findings and suggestions are proposed in the following sections.

4.1. Integrating urban facilities, buildings and roads to support urban planning

The systematic integration of urban facilities, buildings and roads serves as a vital component that aids city planning. As shown in Fig. 6 and Table 6, error variables e1-e4 that directly measure the corresponding four spatial metrics are correlated. However, the covariance estimates reveal that among the relations between e1 and e2, e2 and e3, e3 and e4, and e1 and e4, the estimated values are 0.018, 0.324, 0.007, and 0.175, respectively. The interactions between both e2 and e3 and e1 and e4 are significant at the 0.001 level, and the *p* values for e1 and e2 and e3 and e4 are only 0.174 and 0.736, respectively. These statistical results indicate that the spatial arrangement of urban facilities, buildings and roads is far from balanced.

Four examples selected from the study area are displayed in Fig. 7. Influenced by the environmental consequences, such as the distributions of grasslands and water bodies, the configuration of buildings, roads and urban facilities vary from Fig. 7(a) to Fig. 7(d). High densities of buildings and urban facilities are displayed in Fig. 7(a), but buildings and urban facilities are sparsely distributed in Fig. 7(b). Compared with these configurations, Fig. 7(c) and Fig. 7(d) show a well-organised distribution among buildings, roads and urban facilities. Moving from Fig. 7(a)-(b) to Fig. 7(c)-(d), the trends towards a more organised spatial configuration suggests that city planners should take the urban landscape configuration into consideration as an alternative but effective approach for developing sustainable urban function uses.

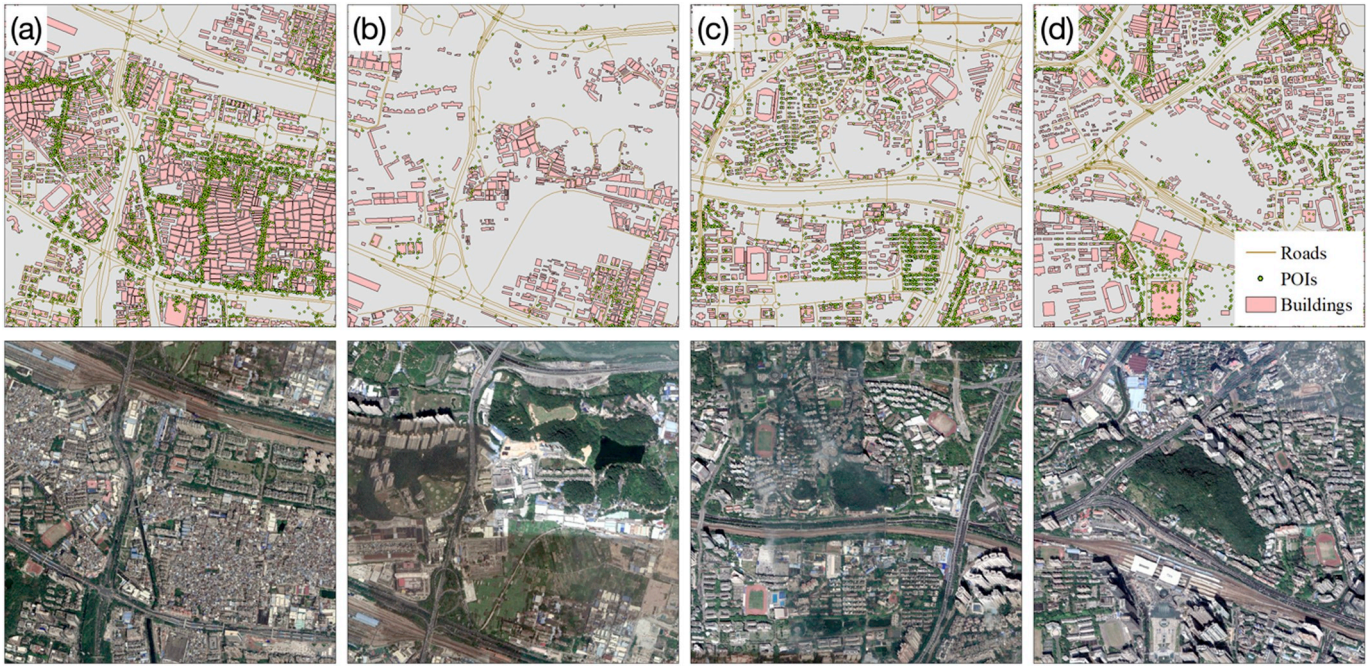


Fig. 7. Four examples of urban landscapes in the study area and the corresponding images.

4.2. Arranging catering and accommodation sites to organise commercial patterns

Although previous studies have examined the spatial distributions of different commercial sites (Rui, Huang, Lu, Wang, & Wang, 2016; Wu, Chen, & Zhao, 2014), the variations among different commercial patterns have not been discussed before. This research concentrated on the spatial distributions of catering and accommodation sites and depicted the related distinguishing characteristics. As shown in Fig. 4, catering sites are distributed with a higher density of clusters than are accommodation sites. We further highlight the high values of topic probability in each grid in Fig. 8. In particular, grids with probability values > 0.5 are selected to represent high probabilities that certain commercial functions exist. As a result, grids filled with blue reflect catering sites, and grids in red are considered accommodation sites. From visual inspection, both the catering and accommodation sites are spatially integrated. Specifically, some accommodation sites and catering sites are adjacent, indicating mixed commercial patterns in many commercial areas. Moreover, the catering sites are widely distributed compared with the accommodation sites. This finding is reasonable because catering businesses tend to serve more people to meet their daily needs, and accommodation businesses usually target tourists.

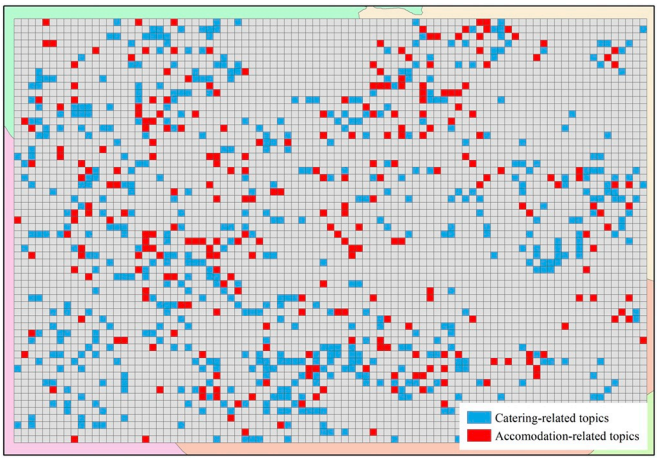


Fig. 8. Spatial distributions of catering-related topics and accommodation-related topics with probability values > 0.5.

However, biases exist in terms of representing commercial patterns using review data. Since most reviews are generated by people who have visited a certain commercial site, the derived semantic topics are highly affected by the variation of people's perceptions. For example, young people contribute the most to the reviews of commercial places, while the voices of older individuals are often ignored. Moreover, review numbers for different commercial places typically vary. In fact, the spatial distribution of the semantic topics in Fig. 8 presents the people perception-based commercial patterns. Although this may lead to the described biases in commercial patterns., it can provide a new perspective to depict commercial patterns based on the attractiveness of commercial places.

4.3. Enhancing the harmony between the urban environment and commercial land uses

The acceleration of urbanisation with the rapid expansion of and variation in the urban landscape has influenced the sustainable development of commerce. Considering the SEM paths and regression weights in Table 5, urban landscapes have a direct impact on both catering-related commercial patterns and accommodation-related commercial patterns, with 0.705 and 0.236 standardised path parameters, respectively. Accordingly, from a theoretical perspective, the development of urban commercial functions is not only directly influenced by socioeconomic factors such as employment and the GDP but also by the existing urban environment.

In China, governments have established a series of policies for developing megacity regions and city agglomerations. As a result, massive urbanised regions with commercial ties have been progressively generated. An example lies in the Guangdong-Hong Kong-Macao Greater Bay Area of southern China. The transport systems in this area serve as a key component for connecting independent cities and have reshaped the undeveloped land uses to form various commercial patterns to support sustainable economic growth (Hui, Li, Chen, & Lang, 2018). This example further supports the need to enhance the integrated construction of urban environments and commercial land uses.

5. Conclusions

In this study, we revealed the spatial distribution of commercial patterns from the perspective of the urban landscape. The conceptual framework was designed based on an *urban landscape depiction – commercial pattern decomposition – relationship* analysis workflow. The proposed approaches and results are as follows. 1) Four spatial metrics, including the density of POIs, PLAND of buildings, LSI of buildings and density of roads, were provided to quantify various urban landscape characteristics. 2) We utilised textual information in review data to investigate commercial topics. Topics related to catering and accommodation functions were selected to indicate commercial patterns. 3) On this basis, an SEM was designed to illustrate the relations among urban landscapes and commercial patterns. In particular, urban landscapes have a direct impact on catering- and accommodation-related commercial patterns.

Based on the proposed research and relevant findings, we proposed three suggestions for improving the commercial pattern development from an urban landscape perspective. First, the construction of urban facilities, buildings and roads should be integrated to support urban planning. Second, examples of catering and accommodation sites are provided to better organise commercial patterns. Third, the combined design of the urban environment and commercial land uses must be prioritised in the urbanisation process.

Although this study analysed the relations among urban landscapes and commercial patterns, rapid commercial changes are complex and can be affected by both landscape and socioeconomic factors (Xing & Meng, 2018). Future studies should focus on using socioeconomic observed variables to assess commercial patterns at fine scales.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 41971406; 41501420).

Declaration of Competing Interest

The authors declare no conflict of interest.

References

Anderson, J. R. (1976). *A land use and land cover classification system for use with remote sensor data. Vol. 964*. US Government Printing Office.
Berry, B. J., & Garrison, W. L. (1958). A note on central place theory and the range of a good. *Economic Geography, 34*, 304–311.
Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.
Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., ... Pei, F. (2017). Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method. *Landscape and Urban Planning, 160*, 48–60.
Doloi, H., Iyer, K., & Sawhney, A. (2011). Structural equation model for assessing impacts of contractor's performance on project success. *International Journal of Project Management, 29*, 687–695.
Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*, 5228–5235.
Hellier, P. K., Geursen, G. M., Carr, R. A., & Rickard, J. A. (2003). Customer repurchase intention: A general structural equation model. *European Journal of Marketing, 37*, 1762–1800.
Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Articles, 2*.
Hu, Q., Bai, G., Wang, S., & Ai, M. (2019). Extraction and monitoring approach of dynamic urban commercial area using check-in data from Weibo. *Sustainable Cities and Society, 45*, 508–521.
Hui, E. C., Li, X., Chen, T., & Lang, W. (2018). Deciphering the spatial structure of China's megacity region: A new bay area—The Guangdong-Hong Kong-Macao Greater Bay Area in the making. *Cities*.
Leichenko, R. M. (2001). Growth and change in US cities and suburbs. *Growth and Change, 32*, 326–354.
Lin, G., Chen, X., & Liang, Y. (2018). The location of retail stores and street centrality in Guangzhou, China. *Applied Geography, 100*, 12–20.
Liu, Y., Lan, Z., & Xing, H. (2019). Inferring spatial distribution patterns in web maps for land cover mapping. *CMES-computer modeling in engineering & sciences. Vol. 119. CMES-computer modeling in engineering sciences* (pp. 311–330).
Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., ... Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers, 105*, 512–530.
Lloyd, A., & Cheshire, J. (2017). Deriving retail Centre locations and catchments from geo-tagged twitter data. *Computers, Environment and Urban Systems, 61*, 108–118.
Long, Y., & Huang, C. C. (2019). Does block size matter? The impact of urban design on economic vitality for Chinese cities[J]. *Environment and Planning B: Urban Analytics and City Science, 46*(3), 406–422.
Ma, F., Ren, F., Yuen, K. F., Guo, Y., Zhao, C., & Guo, D. (2019). *The spatial coupling effect between urban public transport and commercial complexes: A network centrality perspective.* Sustainable cities and society101645.
McGarigal, K., & Marks, B. J. (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. *Gen. Tech. Rep. PNW-GTR-351*. US Department of Agriculture, Forest Service, Pacific Northwest Research Station.
Murphy, R. E., & Vance, J. E., Jr. (1954). Delimiting the CBD. *Economic Geography, 30*, 189–222.
Patton, D. R. (1975). A diversity index for quantifying habitat" edge". *Wildlife Society Bulletin (1973–2006), 3*, 171–173.
Porta, S., Strano, E., Iacoviello, V., Messora, R., Latora, V., Cardillo, A., ... Scellato, S. (2009). Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning. B, Planning & Design, 36*, 450–465.
Rui, Y., Huang, H., Lu, M., Wang, B., & Wang, J. (2016). A comparative analysis of the distributions of KFC and McDonald's outlets in China. *ISPRS International Journal of Geo-Information, 5*, 27.
Rutledge, D. T. (2003). *Landscape indices as measures of the effects of fragmentation: Can pattern reflect process?* .
Siła-Nowicka, K., & Fotheringham, A. S. (2019). Calibrating spatial interaction models from GPS tracking data: An example of retail behaviour. *Computers, Environment and Urban Systems, 74*, 136–150.
Vanderhaegen, S., & Canters, F. (2017). Mapping urban form and function at city block level using spatial metrics. *Landscape and Urban Planning, 167*, 399–409.
Wang, T., Wang, Y., Zhao, X., & Fu, X. (2018). Spatial distribution pattern of the customer count and satisfaction of commercial facilities based on social network review data in Beijing, China. *Computers, Environment and Urban Systems, 71*, 88–97.
Ward, D., Phinn, S. R., & Murray, A. T. (2000). Monitoring growth in rapidly urbanizing areas using remotely sensed data. *The Professional Geographer, 52*, 371–386.
Wentz, E. A., Stefanov, W. L., Gries, C., & Hope, D. (2006). Land use and land cover mapping from diverse data sources for an arid urban environments. *Computers, Environment and Urban Systems, 30*, 320–346.
Wolf, J., Guensler, R., & Bachman, W. (2001). Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data. *Transportation Research Record, 1768*, 125–134.
Wu, L., Chen, X., & Zhao, C. (2014). Taylor's law and the spatial distribution of urban facilities. *arXiv preprint arXiv:1407.5368*.
Xing, H., & Meng, Y. (2018). Integrating landscape metrics and socioeconomic features for urban functional region classification. *Computers, Environment and Urban Systems, 72*, 134–145.
Xing, H., Meng, Y., & Shi, Y. (2018). A dynamic human activity-driven model for mixed land use evaluation using social media data. *Transactions in GIS, 22*, 1130–1151.
Yan, X.-p., Zhou, C.-s., & Leng, Y. (2000). Functional features and spatial structure of CBDs in Guangzhou. *Acta Geographica Sinica-Chinese Edition, 55*, 485–493.
Yu, W., Ai, T., & Shao, S. (2015). The analysis and delimitation of central Business District using network kernel density estimation. *Journal of Transport Geography, 45*, 32–47.
Yuan, J., Zheng, Y., & Xie, X. (2012). Discovering regions of different functions in a city using human mobility and POIs. *Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 186–194). ACM.
Zhu, D., Wang, N., Wu, L., & Liu, Y. (2017). Street as a big geo-data assembly and analysis unit in urban studies: A case study using Beijing taxi data. *Applied Geography, 86*, 152–164.