ARTICLE

Multilevel effects of urban form and urban functional zones on housing prices: evidence from open-source big data

Guan Huang[1] · Si Qiao[1] · Anthony Gar-On Yeh[1]

Abstract
While the relationship between local housing prices and the urban form and distribution of urban functional zones in a single city is well-discussed, the conclusion is usually sensitive to a particular city context, and cross-city comparative study is limited. This study attempts to examine the influences of urban form and urban functional zone distribution on housing prices within and between cities after controlling the city-wide socio-economic and demographic differences. Based on multiple open-source big data, such as points-of-interest (POI) and historical housing transaction data, the hierarchical linear model is utilized to compare the housing market of 10 extra-large cities in China. Results indicate that the urban form and the urban functional zone distribution significantly influence housing prices after the socio-economic and demographic differences are controlled. For inter-city comparison, an urban form with high compactness, low centrality, low polycentricity, high density, and low dissimilarity in housing development is related to lower city-level housing prices. For intra-city, proximity to work centers, high-quality hospitals, and schools shows positive associations to housing prices.

1 Introduction

Research on housing prices based on the hedonic model has widely clarified that the properties of houses (square footage, number of bedrooms, house orientation, building age, and so on), the location, and the neighborhood characteristics undoubtedly affect the prices (Goodman, 1978; Rosen, 1974). Typically, these empirical studies state themselves as unique to the experience of a given city and attribute the uniqueness to the

✉ Si Qiao
siqiao@connect.hku.hk

Guan Huang
guanhuang@connect.hku.hk

Anthony Gar-On Yeh
hdxugoy@hku.hk

[1] Department of Urban Planning and Design, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China

Springer

distinctive urban spatial structure formed by the interaction between the city's physical characteristics and the inhabitants' social needs (Lefebvre, 2003; Lefebvre & Nicholson-Smith, 1991; Lefebvre et al., 1996). However, different case studies in different cities can lead to differences in empirical conclusions. For example, the bid rent model based on monocentric urban form (Alonso, 1964) may have a poor fitting performance for polycentric cities (Wen & Tao, 2015). Likewise, the same kind of urban functional zones in different cities may also affect housing prices differently, such as the impact of hospitals on surrounding housing prices. If hospitals are considered as a type of valuable amenity, then it could increase the commercial values of houses in Shanghai, China Tam et al. ()Tam et al.(2019, 2019), while the negative externalities of hospitals probably depreciate the surrounding housing prices in Beijing, China (Qiao et al., 2021). Meanwhile, no significant impact has been found in Greater Kuala Lumpur, Malaysia (Teck-Hong, 2012).

Regarding the uncertainty or even the conflicting conclusions in various cities, the comparative research is imperative to explore how the differences are caused by the city's uniqueness. Limited studies discussed the housing price difference at the city-level market is primarily driven by the city's economic fundamentals and social demographic structure (Mirkatouli et al., 2018; Waltert & Schläpfer, 2010; Wang et al., 2017; Yi & Huang, 2014). But, as Lynch and Rodwin (1958) Lynch et al. (1958) stated, a well-designed community, a good built environment, and adequate service facilities would not automatically produce the perfect settlements in a city. Houses and their prices are also affected by the city's organization and arrangement from a larger environment. In this vein, we raise a question that, *would the urban form and the urban functional zone distribution affect the housing prices within and between cities?* If yes, then *to what extent do the urban form and the urban functional zones affect the spatial distribution of housing prices?* The empirical evidence could tell whether the effects of urban form and urban functional zones on housing prices were underestimated in previous studies. Also, the findings could support the toolsets to formulate the urban housing policies related to planning human-friendly and affordable settlements when constructing a new town or developing the suburbs under fast urbanization in China and other regions.

Therefore, we construct three hypotheses for assessing the effects of urban forms and urban functional zones on housing prices. First, urban forms and urban functional zones are significantly different across cities. Second, housing prices vary between cities. Finally, the housing price variation is affected by differentiated urban forms and urban functional zones. This study selects 10 representative extra-large cities in China as comparative cases, namely, Shanghai, Beijing, Shenzhen, Guangzhou, Chengdu, Hangzhou, Wuhan, Nanjing, Changsha, and Xi'an (sorted by GDP in 2020). Housing estate price data are calculated from historical transactions to estimate the average price per square meter of a housing estate. Indicators of urban functional zones are identified through the points-of-interest (POI) data obtained from Gaode Map, which is the biggest map and navigation service provider in China, through the contextual semantics algorithm of the natural language process. The relevant indicators in urban form are derived from the integration of road network, building footprint, land use data, parcel data, and administrative boundaries. To estimate the exact effects of intra-city and inter-city influencing factors, we apply the hierarchical linear model (HLM), which supports the fitting of multilevel groups of observations.

The rest of this paper is organized as follows. Section 2 introduces the measurements of independent and dependent variables from previous studies and their findings on these influencing factors on housing prices. Section 3 exhibits 10 extra-large cities' layouts and descriptive characteristics of urban forms and urban functional zones. Section 4 elaborates

the steps of hypothesis validation. Section 5 presents the findings and contributions of this study.

2 Urban form and urban functional zones

The understanding and measurements of urban forms and urban functional zones are complex and mixed. And the selection of appropriate measure methods largely depends on the spatial scale of the case studies (Lynch & Rodwin, 1958). At the city scale, the discussion on urban form includes the city's physical characteristics, such as the size, the shape, the monocentric/polycentric urban structure, the spatial configuration of the urban area (market town, central business district, or suburbs), densities, and the land use patterns (Dempsey et al., 2010). At the same level, the urban functional zone is related to the comprehensive land use for different purposes, such as residential, work, social, commuting, recreational, and administrational activities (Long et al., 2015; Tu et al., 2018). Then, the effects of urban functional zones on housing prices are measured by the distance to the transportation infrastructure (metro/bus stations), the supply of public services, and the proximity to green space, parks, hospitals, and other valuable amenities. In other words, the distribution of urban functional zones represents the relations between the residents' social needs and the urban spatial forms, and the form of cities is conversely shaped by the needs of residents or societies. Therefore, the indicators of urban forms and urban functional zones must be separately considered and dealt with (see Lynch & Rodwin, 1958 pp. 203).

2.1 POI-based urban functional zone within cities

With the rapid urbanization and increased human activities and their interaction, different urban land parcels have gradually formed diverse divisions of urban functional zones (Lefebvre, 2003). As a proxy, urban functional zones represent a rich mixture of places to accommodate human activities, such as living, working, shopping, dining, and recreations, which are vital for sensing the daily living of citizens and capturing the evolving nature of urban functions (Crooks et al., 2015).

Considering the complexity of urban functional zones derived from human activities and interactions (Batty, 2008), the image feature detection method is the traditional and widely used way of extracting and analyzing land use patterns and urban functional zones and relies on remote sensing images with a high spatial resolution (Cao et al., 2020; Wen et al., 2015; Zhang & Du, 2015; Zhang et al., 2017; Zhong et al., 2015). However, although this method specializes in detecting and reflecting the natural properties and features of ground objects, such as land use types, it fails to capture detailed human-involved functional information and socio-economic characteristics (Liu et al., 2015). These limitations are generally overcome by the emergence of big geographic data, such as POIs (Hu et al., 2020; Jiang et al., 2015; Liu & Long, 2016; Liu et al., 2020; Qian et al., 2021; Yao et al., 2017; Yi et al., 2019; Zhai et al., 2019; Zhang et al., 2017, 2019), social media (Du et al., 2020; Gao et al., 2017; Liu et al., 2015; Qian et al., 2021; Tu et al., 2017; Zhang et al., 2019; Zhi et al., 2016; Zhou & Zhang, 2016), mobile phone and floating car trajectory data (Gao et al., 2019; Hu et al., 2021; Qi et al., 2011; Tu et al., 2017; J. Zhang et al., 2021a, 2021b; Zhong et al., 2014), due to the strength of big geographic data to accurately and large-scale capture the human activities. Numerous studies have exemplified the usability and reliability of methods integrated with such data sources in urban functional zones

inference, identification, and detection, despite the requirement of meticulous data processing (Li et al., 2016; Niu & Silva, 2020). The methods, data source, and their pros and cons for urban functional zones identification are summarized in Table 1.

In recent studies, POI data is the most widely used data source and has four main advantages for urban functional zones identification at a fine-grained scale. First, POI data is highly accessible and available from online map service providers. Second, POI data is highly reliable due to the continuous and multi-source maintenance by commercial map suppliers and volunteer geographical information (Batty et al., 2012). Third, POI data has a fine-grained spatial resolution and a continuous temporal sequence. Lastly, the approximate global coverage of POI data supports large-scale research and global comparative studies. For instance, Jiang et al. (2015) utilized POIs obtained from Yahoo Online and employment data from census data at the aggregate level to identify the land use types at the city block level. Liu and Long (2016), proposed an automatic detection method for urban functional zones by using open resource data, POI data, and road networks from OpenStreetMap.

However, two challenges in using POIs remain (Zhai et al., 2019). First, as the description of POIs has several layers, synonymy and ambiguity are common when directly using these layers' information. As shown in Appendix Table A1, the bottom layer of food service can be a component in residential or commercial urban functional zones, but the combined semantic signature may be different (such as *accommodation service–hotel–food and beverage* vs. *commercial housing–industrial park–food and beverage*). Second, because POI description can only represent the information in the location of the point, redundancy is commonly seen in large plots, like hospitals, parks, airports, or universities. For instance, in addition to one POI with a combined description of *tourist attraction–scenic spot–park*, the spatial area of a park contains many other ancillary facilities (e.g., ticket office, catering, shopping, sports facilities, and toilets). If only the spatial relationships between these POIs in a given area are considered when inferring the function of the place, redundant POIs would affect the result.

Therefore, we followed the automatic method proposed by Liu and Long (2016) and proposed a hybrid approach to urban functional zones detection to take full advantage of the POI data. The process involves two steps (Fig. 1):

●First, use POI types, which represent single land use, to assign the land use type for urban functional zones.
●Second, use the multi-layer combined descriptions of the POIs reclassified by Sense2Vec[1] to infer the remaining zones with mixed land use.

2.2 Urban form indicators at the city scale

Urban form represents the physical form of a city and comprises activities (Lynch & Rodwin, 1958). The definition and measurements of urban form in the literature are very diverse, mainly relying on the scale of the research (Batty, 2008; Batty & Longley, 1994). Huang et al. (2007) Huang et al. (2007) utilized satellite images of 77 metropolitan areas in Asia, the United States, Europe, Latin America, and Australia to quantify the five dimensions (i.e., compactness, centrality, complexity, porosity, and density) of urban form from the physical characteristics of the urban landscape. Clifton et al. (2008) Clifton et al. (2008) divided the quantitative methods of analyzing urban forms into five categories, namely, landscape ecology,

---

[1] Sense2Vec is a commonly used natural language processing (NLP) machine learning algorithm that transfers a sentence to vector, allowing for further similarity comparison and clustering processing.

Table 1 Five mainstreams for urban functional zones identification

Category	Primary data source	Pros	Cons	Sample papers
(1) Remote sensing (RS)	RS images	Reflect the natural properties of ground objects with high resolution	Difficult to detect building characteristics and socio-economic activities	Cao et al., (2020); Wen et al., (2015); Zhang and Du, (2015); Zhang et al., (2017); Zhong et al., (2015)
(2) Geodemography	Socio economic statistics; social surveys	Could describe the social demographic structure of small geographic regions	Constrained computational ability; the willingness to engage surveys; collection of data is costly; labor intensive and resource consuming	Adnan et al., (2010); Ashby and Longley, (2005); Xing and Meng, (2018)
(3) POI	POIs	Provide information reflecting the place of human activity; high data accessibility and low cost; maintained by the commercial map supplier with high reliability	Hard to reflect areal unit attributes due to lack of area-of-interests information and regional boundaries; complex semantic expression associated with human activities	Hu et al., (2020); Jiang et al., (2015); Liu et al., (2020); Liu and Long, (2016); Qian et al., (2021); Yao et al., (2017); Yi et al., (2019); Zhai et al., (2019); Zhang et al., (2017); Zhang et al., (2019)
(4) Trajectory	Trajectories of floating cars or mobile phone	Could observe and explore the people's move and interact on both spatial and temporal dimension	Lack of data accessibility; large data throughput requiring high computer processing performance	Gao et al., (2019); Hu et al., (2021); Qi et al., (2011); Tu et al., (2017); J. Zhang et al., (2021a, 2021b); Zhong et al., (2014)
(5) Social media	Volunteered geographic information on social media platform such as Weibo and Twitter check ins	Provide affluent information on human activities and socio-economic information	Difficult to distinguish areas of low human activity, Privacy concerns	Du et al., (2020); Gao et al., (2017); Liu et al., (2015); Qian et al., (2021); Tu et al., (2017); Zhang et al., (2019); Zhi et al., (2016); Zhou and Zhang, (2016)
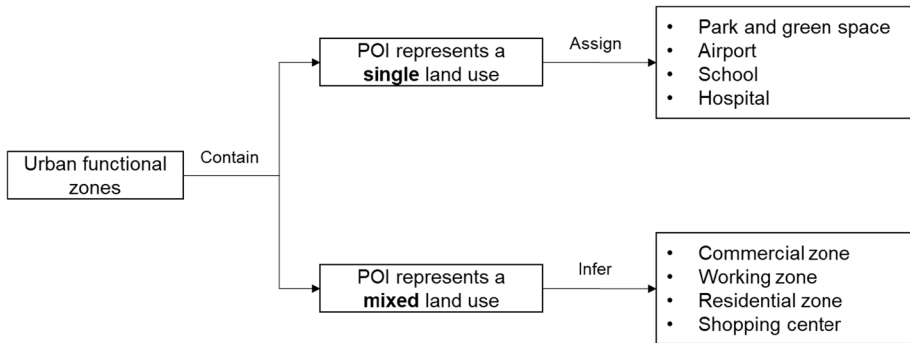
**Fig. 1** Flow diagram of detecting urban functional zones

economic structure, transportation, community design, and urban design, and sorted them by research scales from regional, metropolitan, sub-metropolitan, neighborhood to the smallest building block. At the city scale, the measures include the urban built-up size, density, and diversity, the urban structure, and the polycentricity. Dempsey et al. (2008) Dempsey et al. (2008) classified urban forms of different scales into five broad and interrelated elements that constitute a specific city and are encompassed by density, transport, housing, land use, and layout. The indicators of urban form at the city scale reflect the ratio of population, households, or housing units to the area of the entire city, including the city's population density, the household density, the dwellings density, and the city area. Echenique et al. (2012) Echenique et al. (2012) proposed that urban forms (compaction, dispersed, and planned expansion) in English city regions do not have obvious advantages for sustainable development. Conversely, population growth and changes in residents' lifestyles dominate the impacts on natural environment and resources. In comparative studies of global cities, openness and proximity are often used to represent the fragmentation or compactness of urban form (Dong et al., 2019; Schneider & Woodcock, 2008; Xu et al., 2020).

The frequently used urban form indicators for extra-large cities should be capable of indicating the outcomes that have great significance at the city or metropolitan scale and can be controlled and describe the different effects caused by the rearranged patterns at this scale (Lynch & Rodwin, 1958). In this study, we applied this criterion and utilized indicators to reflect the five dimensions of urban form, including compactness, centrality, polycentricity, dissimilarity, and density.

Compactness (Fig. 2a). Compactness does not only measure the shape of an individual developed block but also the fragmentation of the overall urban landscape (Li and Yeh, 2004; Angel et al., 2017, 2020). The estimation is based on the average comparison between the perimeters of each developed block and a circumscribed circle of the same area. The CI is calculated as follows:

$$Compactness = \frac{\sum P_i/p_i}{N^2} = \frac{\sum 2\pi\sqrt{S_i/\pi}/p_i}{N^2}$$

where S_i and p_i are the respective area and perimeter of developed block i, P_i is the perimeter of the circumscribed circle with area S_i, and N is the total number of blocks. Given two cities with the same amount of development, when city A is more compacted than city B, then the development of B is more evenly distributed than that of A. With a low degree of compactness, the sprawl-like development model is spreading in this city.
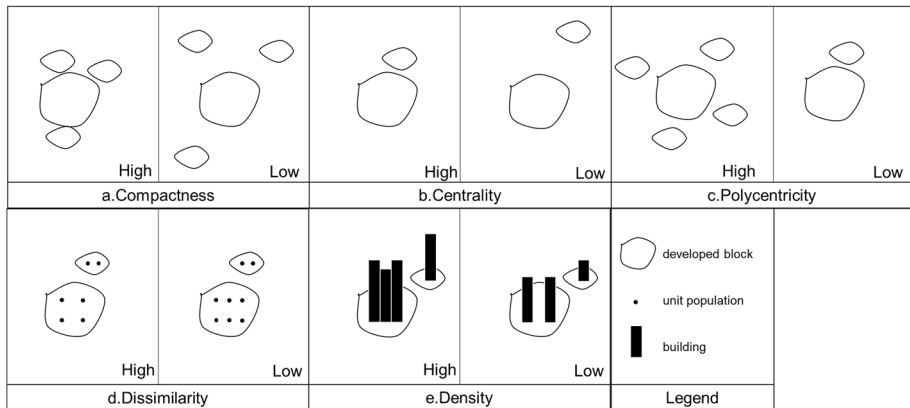
Fig. 2 Schematic of assessing spatial urban form

Centrality (Fig. 2b). Centrality is the degree of proximity of a city's developed areas to its urban central business district (CBD) (Galster et al., 2001). Centrality is a measurement of the areal size of a city. One of the most common forms of urban development is the loss of centrality, which results in increased travel distance and time as land development is far from the CBD. The centrality index increases with the decrease of the radius between the developed areas and the CBD. Conversely, the centrality index is low when a city has a great spread. The calculation formula of centrality is as follows:

$$Centrality = \frac{\sum D_i / N - 1}{R} = \frac{\sum D_i / N - 1}{\sqrt{S/\pi}}$$

where D_i is the distance of the centroid of developed block i to the centroid of the city center, N is the total number of developed blocks, R is the radius of a circle with area S, and S is the total built-up area of the city. Therefore, centrality is also sensitive to the shape of the city, that is, whether it is elongated or circular. The narrower the shape of the city is, the greater the centrality index, and vice versa.

Polycentricity (Fig. 2c). Polycentricity reflects the degree to which urban areas are characterized by a polycentric (rather than monocentric) development model. If a city's CBD is the only dense development zone, then this city has a monocentric structure and its polycentricity index equals one. If land development is scattered in several highly developed areas and each area contains an activity cluster, which accounts for a large part of the total number of such activities in the city, then this area is defined as the sub-center of the city. A polycentric city has a high polycentricity index, which equals to the number of sub-centers of the city.

Dissimilarity of housing development (Fig. 2d). The dissimilarity index refers to the unevenness of housing development. The housing development with a high dissimilarity index indicates the houses gathered at some developed places. While a low dissimilarity index represents houses evenly developed in a city (Schwarz, 2010; Tsai, 2005).

$$Dissimilarity = \frac{\sum_{i=1}^{N} |X_i - Y_i|}{2}$$

where X_i is the proportion of land area, Y_i is the proportion of housing development indicator of block i, and N is the total number of blocks in a city. A dissimilarity near 1 indicates

large spatial differences of housing development between blocks in a city, while a value near 0 indicates the even distribution of housing development. In this study, we utilized residential population as the housing development indicator.

Density (Fig. 2e). Density is generally accepted as a dimension of urban form. Usually, density is calculated by comparing the footprint of buildings to the extracted built-up area. Density indicates the number of residents per square kilometer of developable area in the city. However, building density cannot distinguish between the high-density forms of high- and low-rise buildings. With the enrichment and availability of 3D information, the building area ratio (BAR) has been defined to reflect the dense 3D urban form (Long et al., 2019). The BAR is expressed as follows:

$$BAR = \frac{\sum_n A_i}{S} = \frac{\sum_n P_i \times F_i}{S}$$

where $\sum A_i$ is the total indoor area of all the buildings within the city. S is the area of the urban built-up region. P_i is the footprint of individual building i, and F_i is the floor number of building i. n is the total number of buildings. A high BAR indicates a dense urban form and vice versa.

2.3 Effects of urban forms and urban functional zones on housing prices

Existing studies have extensively demonstrated the influence of house location and accessibility on the housing prices in a specific city on the basis of hedonic price models. Usually, scholars investigate the positive impact of the distribution of urban functional zones on housing prices with the point-like benefits facilities, such as educational facilities (Wen et al., 2014, 2019), sports facilities (Feng & Humphreys, 2012), parks (Wu et al., 2017), shared bikes (Qiao et al., 2021), transit-oriented-development facilities (Duncan, 2011), and comprehensive public facilities (Yuan et al., 2020). These empirical findings suggest that residents are willing to pay for good amenities. Despite the amenity effects of public facilities that bring premium to houses, the negative externalities of public facilities can also reduce the housing prices. Obnoxious public facilities are those necessary for society but not welcomed by nearby local residents because they have a negative impact on the sanitation environment of neighboring areas. Obnoxious facilities, such as hazardous industrial facilities (Grislain-Letrémy & Katossky, 2014), hazardous chemical facilities (Lee et al., 2008), resettlement houses (M. Zhang et al., 2021a, 2021b), and toxic waste sites (Kohlhase, 1991), have been shown to have negative impacts on housing prices.

However, past studies on the evaluation of the relationship between housing prices and facilities were usually based on the point-like distribution of a single type of facility, measuring the proximity to specific facilities and then exploring the impact of facility-based accessibility on housing prices. One of the shortcomings of the big data measurement of point service facilities is the multi-semantics of the data, as stated in Sect. 2.1. Meanwhile, the urban functional zone is a new representation approach to reflecting the complex and comprehensive human activities in space, taking advantage of the single point-like facilities measurement. The distribution of urban functional zones, such as business zones, working, shopping zones, parks and green spaces, hospital areas, and school catchments, presents the urban structure within a city. Thus, this study involved urban functional zones as a novel indicator to replace the single point-like facilities indicators.

In terms of the urban form's impacts on housing prices, existing studies proposed a hypothesis that localized urban forms have impacts on the housing prices within the local

area. For example, localized high density reduces the surrounding housing prices (Fesselmeyer & Seah, 2018), and increasing the compactness contributes to increased land values and may cause increased housing prices, especially in newly developed urban areas that must accommodate migrant populations (Hamidi & Ewing, 2015). However, with the increase in compactness, transportation costs have dropped faster than the housing costs have increased, resulting in a net decrease in household costs. Whereas Wassmer and Baass (2006) argued that no evidence that centralization would raise the housing prices in the urban area has been found. In view that a single indicator, such as compactness, density, or concentration, cannot fully show the comprehensive physical environment of the city (Galster et al., 2001), and the lack of comprehensive urban form indicators in the past comparative studies limited the understanding to the impact of urban form on urban housing prices. This study attempted to comprehensively measure the urban form of 10 large cities through the five dimensions described in Sect. 2.2 to bridge this research gap.

In general, the improved urban functional zones and urban forms indicators of the two sub-market levels, namely, intra- and inter-cities, help answer the research question on whether the layout of the urban physical environment can regulate the housing market. This study can help urban planners and policymakers have a better understanding of the impact of the urban physical environment on housing prices.

3 Comparative cases

3.1 Extra-large cities

This study measured the distribution of urban functional zones from the intra-city scale and the difference in urban form from the inter-city scale. These cross-level indicators were applied to estimate the impacts of urban forms and urban functional zones on housing prices between cities. We selected Shanghai, Beijing, Shenzhen, Guangzhou, Chengdu, Hangzhou, Wuhan, Nanjing, Changsha, and Xi'an (sorted by GDP), which are 10 major extra-large cities in China, as our research cases (see Fig. 3). Beijing is the capital city of China and the political and cultural center. Shanghai is China's financial center and an international metropolis. They are the representatives of extra-large cities in the northern and eastern regions of China respectively (Fig. 3a, d). Shenzhen represents the economically developed special administrative region along the southern coast (Fig. 3j). Shenzhen is also regarded as a pioneer city in China's reform policy, attracting numerous immigrants. The other cities are provincial capitals, representing the most thriving housing markets locally. These ten extra-large cities illustrate the diversity of urban forms and socio-economic developments in China.

3.2 Data description

This study involved four datasets. One is the road network data of 2018 obtained from Tianditu, which is an official map supplier certificated by the Land and Survey Department of China (https://map.tianditu.gov.cn). The road network data was used to divide the urban area into small zones for urban functional zone detection. Another primary dataset is the 2018 POI data crawled via the Gaode Map Service (http://map.gaode.com), which is the most widely used map and navigation service provider in China. We obtained 3,603,178 POI records of multiple categories for our study areas. As the category level is upgraded to three layers, a more detailed description is provided by combining three levels of separate Chinese phrases, which have the most integrated
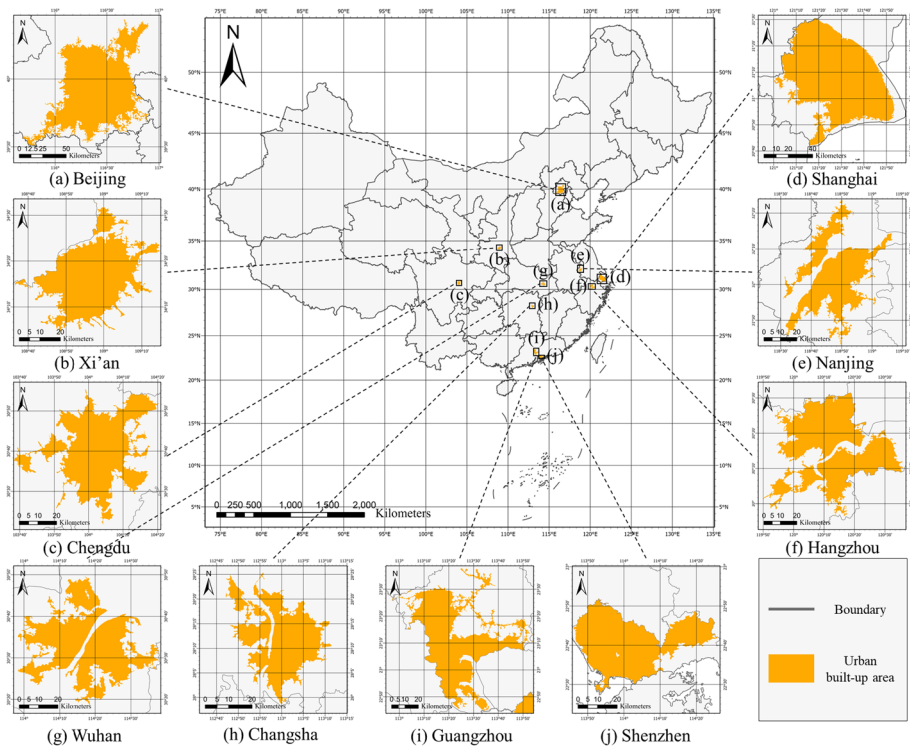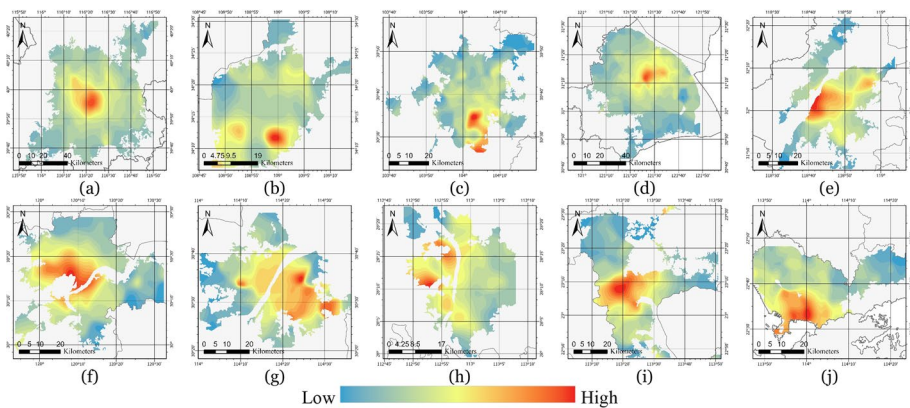
Fig. 3 Built-up area of the ten extra-large cities. (**a**) Beijing, (**b**) Xi'an, (**c**) Chengdu, (**d**) Shanghai, (**e**) Nanjing, (**f**) Hangzhou, (**g**) Wuhan, (**h**) Changsha, (**i**) Guangzhou, (**j**) Shenzhen

information. In our POI dataset, the top-level category has 13 tags, and the bottom-level category has 2,168 tags (Appendix Table A1). Third, the average housing estate price per-square-meter (HEPrice) was calculated from historical transaction records from Nov. 2020 to Oct. 2021. One year's data could smooth out the seasonal effects of cyclical fluctuation. Housing estate is referred to as *Xiaoqu* in Chinese. *Xiaoqu* is a complex of one or tens of residential buildings built by the same property developer or the municipal government. These buildings typically have the same housing qualities and shared facilities, such as gardens, swimming pools, courts, and fitness equipment. Thus, the price of housing estates enabled this study to overcome the impact of the housing structure attributes, thereafter focusing on the impact of urban forms and urban functional zones. The historical housing estate price is the "asking price", which was obtained from *Lianjia*, *Anjuke*, and *Fangtianxia*, which are three national real estate platforms in China, and the housing estates were counted only when their housing transactions exceeded three during the past year. The descriptive statistics of the average per-square-meter housing estate price in extra-large cities are shown in Table 2. Figure 4 illustrates the spatial distribution of the standardized housing prices for each city. The red color represents areas with high local housing prices, and the blue color represents areas with low housing prices. Finally, 47,316 housing estates in the 10 extra-large cities have validated average prices after data cross-validation from three platforms. Lastly, the social-economic data of the cities were obtained from China's National Statistical Yearbook 2020 (China, 2020).

Table 2 Descriptive statistics of average per-square-meter housing estate price in the ten extra-large cities

City name	Min	Max	Mean	Std. Dev	Counts
Shanghai	10,629	170,319	6,5061.69	31,560.97	9,454
Beijing	17,937	172,414	6,5477.49	34,375.78	5,966
Shenzhen	24,166	131,000	6,1640.42	23,162.43	3,432
Guangzhou	7,879	93,759	3,5901.15	20,316.75	4,396
Chengdu	5,089	30,032	12,035.49	45,65.732	8,026
Hangzhou	9,915	100,799	39,198.38	19,098.93	3,310
Wuhan	5,918	32,654	16,397.68	55,36.613	4,098
Nanjing	9,418	83,836	34,011.27	15,025.99	3,749
Changsha	5,356	24,582	10,011.98	39,25.467	2,019
Xi'an	6,093	40,084	14,086.01	61,03.178	2,956

[#]: Price unit: yuan/m^2, CNY. The currency exchange rate is 100 CNY to 15.65 USD for 22nd Oct 2021

Fig. 4 Interpolated housing prices. (**a**) Beijing, (**b**) Xi'an, (**c**) Chengdu, (**d**) Shanghai, (**e**) Nanjing, (**f**) Hangzhou, (**g**) Wuhan, (**h**) Changsha, (**i**) Guangzhou, (**j**) Shenzhen

3.3 Potential determinants

Several potential influencing factors possibly affect the HEPrice even after internalizing the housing structure attributes' effects. According to previous literature reviews, we classified three categories of indicators to evaluate the cross-scale potential determinants of housing estate price (see Table 3).

Category 1 is the individual category, including the structure and location attributes of specific housing estates. Estate building year has a significant impact on housing price due to the bank loan policy in China. A high down payment is associated with an older house, a shorter load period, and a lower available loan amount. The increasing burden of down payment with the increased building year would intensely affect buyers' choice and thus be reflected in the pricing strategy of sellers. Furthermore, the location attributes have a wide impact on the capability of accessing urban service opportunities and access to urban services may increase house values, as evidenced by higher prices for homes close to transit in previous studies (Duncan, 2011; Shen & Karimi, 2017).

Table 3 Variables and statistics based on three categories of spatial scales

Category	Sub-category	Variables [unit]	Mean	Std. dev	Level
1. Housing estate indicators	Estate structure attribute	Building year [year]	1999.67	11.487	Level 1
	Estate location attribute	Proximity to metro station [kilometer]	1.57	2.008	Level 1
		Proximity to bus station [kilometer]	0.27	0.136	Level 1
2. Intra-city indicators	Estate-based urban functions attribute	Proximity to working zone [kilometer]	0.53	0.386	Level 1
		Proximity to shopping center [kilometer]	1.14	0.715	Level 1
		Proximity to commercial business zone [kilometer]	2.74	2.276	Level 1
		Proximity to park and green space [kilometer]	1.57	1.788	Level 1
		Proximity to 3A hospital [kilometer]	5.00	5.462	Level 1
		Proximity to primary and middle school [kilometer]	0.65	0.426	Level 1
3. Inter-city indicators	City-based urban form attribute	Compactness [index]	0.13	.06	Level 2
		Centrality [index]	0.71	.116	Level 2
		Polycentricity [index]	3.97	1.683	Level 2
		Building area ratio [ratio]	1.47	.242	Level 2
		Dissimilarity [index]	0.43	0.08	Level 2
	City-based socio-economic attribute	Gini coefficient of population [continuous]	0.51	.085	Level 2
		Moran index of education [continuous]	0.20	.076	Level 2
		GDP per capita [1 million Yuan/population]	0.18	.022	Level 2

Category 2 attempts to explore the effects of the distribution of urban functional zones in intra-city, which is an extension of the impact of housing estates on the neighborhood attributes. The proximity to urban functional zones represents the time cost and expense burden of reaching potential destinations to obtain urban services, such as medical service, educational resources, and physical exercise amenities. The proximity indicators of urban functional zones, including working zones, shopping centers, commercial business zones, parks and green spaces, 3A hospitals, and primary and middle schools, are selected as the level 1 indicators (Qiao et al., 2021; Tam et al., 2019; Wang et al., 2017; Wen et al., 2017; Zhang et al., 2023).

Category 3 covers the city-level attributes of the socio-economic and physical urban form. The physical urban form indicators include the compactness, centrality, polycentricity, dissimilarity, and density. The details are presented in Sect. 2. The socio-economic factors include the Gini coefficient of population, the Moran index of education, and the GDP per capita, (Tsai, 2005; Villar & Raya, 2015; Zhang et al., 2016). The Gini coefficient of population represents the unequal population distribution in the metropolitan area. A high Gini coefficient approaching 1 indicates a high unevenness of population. The Moran index of education ranges from -1 to 1, illustrating the clustering pattern for the low-educated population (educated in primary school or below). If the value approaches 0, then the pattern is randomly scattered, while an absolute value near 1 represents a clustered pattern. This indicator shows whether low education, as a proxy of poverty, is agglomerated. The GDP per capita indicates a city's level of economic development. As the urban form indicator of building density is collinear with the population density, we omitted it. The final variables and statistics are shown in Table 3. The variance inflation factor (VIF) of the final variables is provided in Appendix Table A3. The VIFs of majority of variables are less than 5 and all VIFs are less than 10 indicating there is no significant multicollinearity between the final variables.

3.4 Hierarchical regression model

The traditional ordinary least squares (OLS) hedonic model has been criticized for ignoring the existence of housing market segmentation (Goodman & Thibodeau, 1998, 2003). In short, a house occupies a unique location in space, and the geographic attributes tied to its location cannot be replicated. The demand for houses in certain locations may be very inelastic, such as houses near prestigious schools (Wen et al., 2017). The irreplaceability of different housing locations shows that houses are essentially nested in spatial submarkets with different pricing premiums (Orford, 2000). Therefore, to solve the fixed effects of intra-city and inter-city submarkets, the hierarchical linear model (HLM) is introduced. In the first level, the individual housing estate attributes of proximity to urban functional zones are measured (see Table 3, "Level" column). This level reflects the housing values nested in intra-city functions. In the second level, the inter-city attributes of the urban physical forms and the city-based socio-economics of the habitants are assessed. The formulas of the two-level HLM used in this study are as follows (Goodman & Thibodeau, 2003):

$$Level_1 : P_{ij} = \alpha_{0j} + \alpha_{1j}(x_{1ij}) + \alpha_{2j}(x_{2ij}) + \cdots + \alpha_{pj}(x_{pij}) + e_{ij}, e_{ij} \in N(0, \sigma^2),$$

where P_{ij} is the logarithm-transformed HEPrice of estate i within city j; α_{pj} ($p=0, 1, 2, …, p$) is an intercept and regression coefficient. x_{1ij} is the level 1 variable that represent individual housing estate attributes. e_{ij} is a random error.
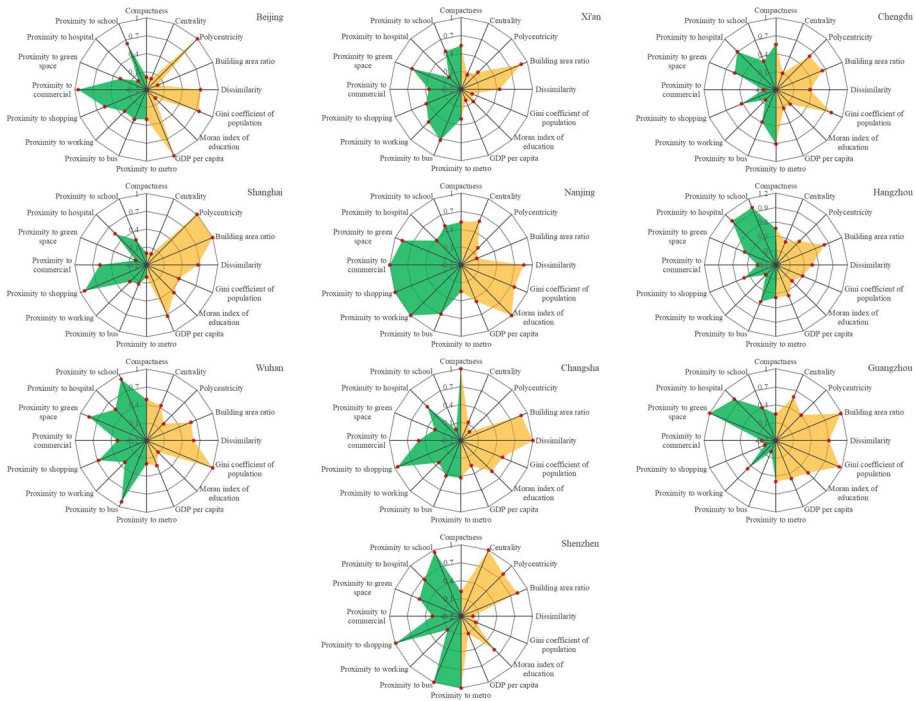
Fig. 5 Visualizing indicators of urban forms and urban functional zones between extra-large cities

$$Level_2 : \alpha_{pj} = \beta_{p0} + \beta_{p1j}(y_{1j}) + \beta_{p2}(y_{2j}) + \cdots + \beta_{pq}(y_{qj}) + \varepsilon_j$$

where β_{pq} ($q = 0, 1, 2, ..., q$) is the level 2 intercept and regression coefficient, and y_{qj} is the level 2 variable related to urban form indicators. ε_j is the random error at level 2.

4 Results

4.1 Disparities of urban forms and urban functional zones in ten extra-large cities

The visualized profiles of urban forms and urban functional zones illustrate the disparities between the 10 extra-large cities in Fig. 5 (see details in Appendix Table A2). From the different dimensions of measurements, clear disparities of urban forms and urban functional zones can be witnessed. For instance, Shanghai and Beijing both represent similar short distance to urban functions which benefits from their urban form of low compactness, low centrality and their six subcenters of working zones (Fig. 5a, d). However, the urban dwelling density shows that Shanghai developed as a high-density sprawl and Beijing maintained a relatively low-density. Shenzhen, which is also a polycentric developed city with five subcenters, has the highest centrality index among the 10 cities but the lowest dissimilarity index. This result can be attributed to the limitation of the developable land resource. Shenzhen had to spread along the narrow and long administrative boundary (Fig. 3j), presenting an urban form of high-density and a polycentric sprawl. Similarly,

Table 4 Analysis of variance of average per-square-meter housing estate price (HEPrice) within and between extra-large cities

Source	SS	df	MS	F	Prob > F
Between cities	2.39e + 13	9	2.65e + 12	5571.51	0.0000
Within cities	2.25e + 13	47,306	4.76e + 08		
Total	4.64e + 13	47,315			
Bartlett's test for equal variances:	chi2(9) = 4.2e + 04		Prob > chi2 = 0.000		

Wuhan and Changsha's urban built-up area are divided into parts due to the segmentation by lakes and rivers and thus formed a resource-limited sprawl pattern with a monocentric urban form (Fig. 3g, h). However, Changsha has a relatively high dissimilarity of housing development, causing this city to have a high-dissimilarity sprawl pattern compared to Wuhan, which becomes a low-centrality sprawl city. These two cities also belong to the monocentric urban form. Furthermore, the monocentric structure increases the proximity to urban functions compared to the polycentric city. Chengdu has the same compactness and centrality as Wuhan and Changsha but more subcenters. Thus, Chengdu has a shorter distance to urban functional zones with a typical polycentric sprawl with high-density and low-centrality characteristics (Fig. 3c).

Through the differences in these indicators, we can validate the first research hypothesis of this study, that is, significant differences in urban forms and functional zone distribution exist between these study cities. The disparities are formed under the influence of a series of complex factors as a result of land resources, urban planning, urban economy, and population structure.

4.2 Variation of housing prices in ten extra-large cities

Table 4 presents the analysis of variance of the HEPrice within and between the 10 cities through one-way ANOVA. The analysis compares the means of the HEPrice between the study cities and determines whether any of the housing price means are statistically, significantly different from each other. To specifically verify whether the majority of the cities exhibit a difference in HEPrice, we further performed post-hoc pairwise comparisons, that is, Fisher–Hayter pairwise comparisons, which are fit for the unequal cell sizes between groups. This comparison could help determine which specific cities differed from each other. The result shows that most of the cities have significant differences in HEPrice (Appendix Table A4), validating our second hypothesis, that is, housing prices vary between cities.

4.3 Effects of urban forms and urban functional zones on housing price

To evaluate the effects of urban forms and urban functional zones on the housing prices between cities, we constructed three different models based on the HLM as shown in Table 5. Model 1 only includes the building year as a baseline model, and Model 2 includes all level-1 variables. Model 3 controls both level-1 and level-2 indicators. The OLS model is a comparative model applied to the OLS method. Fitness indicators AIC and BIC show that Model 2 has a higher fit performance than baseline Model 1, and Model 3 has the highest fit performance among the models, indicating that the HLM with all variables is appropriate for this study.

Table 5 Effects of urban forms and urban functional zones on housing estate prices at multi levels

Price (log)	Model 1		Model 2		Model 3		OLS model	
	Coef	Robust Std.Err	Coef	Robust Std.Err	Coef	Robust Std.Err	Coef	Robust Std.Err
Building year [level 1]	−0.007***	0.002	−0.001	0.003	−0.001	0.003	0.000**	0.000
Proximity to metro station [level 1]			−0.008	0.007	−0.008	0.007	−0.010***	0.002
Proximity to bus station [level 1]			0.105	0.070	0.105	0.070	0.114***	0.015
Proximity to working zone [level 1]			−0.092***	0.033	−0.092***	0.033	−0.091***	0.005
Proximity to shopping center [level 1]			0.061***	0.009	0.061***	0.009	0.061***	0.003
Proximity to commercial business zone [level 1]			−0.012	0.009	−0.012	0.009	−0.016***	0.001
Proximity to park and green space [level 1]			0.013	0.011	0.013	0.011	0.016***	0.001
Proximity to 3A hospital [level 1]			−0.044***	0.003	−0.044***	0.003	−0.042***	0.001
Proximity to primary and middle school [level 1]			−0.084**	0.032	−0.084***	0.032	−0.088***	0.005
Compactness [level 2]					−10.996***	0.327	−10.822***	0.054
Centrality [level 2]					1.287***	0.225	1.436***	0.037
Polycentricity [level 2]					0.180***	0.054	0.189***	0.007
Building density [level 2]					−0.232***	0.065	−0.216***	0.009
Dissimilarity [level 2]					1.458***	0.425	1.682***	0.057
Gini index of population [level 2]					−1.375***	0.227	−1.457***	0.037
Moran index of education [level 2]					−0.718*	0.393	−0.835***	0.043
GDP per capita [level 2]					10.905***	1.042	10.894***	0.143
Level 1 Constant	24.585	5.998	11.849	5.633	10.901	5.494	10.193	0.420
Level 2 Constant	0.432	0.102	0.472	0.112	0.002	0.000	N.A	N.A
Level 2 Residual	0.199	0.024	0.131	0.012	0.131	0.012	N.A	N.A
Residual intra−class correlation	0.683		0.781		0.016		N.A	
Obs	47,316		47,316		47,316		47,316	
Groups	10		10		10		N.A	
AIC	58,176.63		38,539.57		38,502.13		38,939.39	
BIC	58,211.69		38,644.75		38,577.42		39,097.15	

*** $p < .01$, **$p < .05$, *$p < .1$

In the comparison of Models 2 and 3 in Table 5, the residential intra-class correlation reveals that the inter-city differences explain 76.5% of the variance of HEPrice. This answers exactly how much of the variance of HEPrice is contributed by the city's uniqueness. Model 3 also validates our third hypothesis that housing price variation is partially produced by differentiated urban forms and functions. Specifically, the HEPrice would decrease significantly (-10.996^{***}) with the increase of compactness, while centrality is positively associated with the housing price (1.287^{***}). Polycentricity stimulates the city-level housing price (0.180^{***}) because more working subcenters could attract migrants, thereby increasing the demands of residential settlements (Wen & Tao, 2015). The dissimilarity index is positively associated with the city's HEPrice (1.458^{***}), indicating that the uneven land development of the city is associated with high housing prices on the overall level. This result supplements the evidence of the impact of polycentricity on housing prices from another dimension. Balanced development was achieved through the allocation of state investment, which often drives the construction of new satellite cities/towns (Liu et al., 2018). The density index negatively affects the HEPrice (-0.232^{***}), suggesting that a high dwelling density is related to a low city-level housing price. This result is inconsistent with the previous intra-city evidence of housing premiums induced by local density (Kulish et al., 2012). High-density areas at the local level are often urban CBDs where the affluent are attracted to live nearby, thereby increasing the housing prices. Conversely, for city-level density comparison, high urban density typically means a high level of house supply at the overall level. When controlling other variables, such as the GDP and the population structure, a large amount of supply would depreciate the values of individual houses if the demand is fixed.

In terms of urban functional zones, Model 3 shows that only urban functions of working zones (-0.092^{***}), 3A hospitals (-0.044^{***}), and primary and middle schools (-0.084^{***}) have premiums on housing prices after controlling the city's differences. The findings are consistent with most of the previous empirical evidence that working opportunities, convenient commuting, and valuable amenities, like hospitals and schools, have capitalization effects on housing prices (Wang et al., 2017; Wen et al., 2017). Conversely, proximity to shopping centers likely inhibits the surrounding housing values (0.061^{***}). This result is consistent with that of a previous study where large-scale and high-end shopping complexes have a significant negative impact on the housing prices in core urban areas (Zhang et al., 2020).

5 Conclusions

Although previous studies have affirmed the contribution of a city's uniqueness to the housing market in that city, this study supplements the empirical evidence of the influence of urban forms and urban functional zones on the differentiated performance of the housing market on multiple scales (within and between cities). Our findings indicate that the difference in housing prices across cities is not only attributable to the city's uniqueness in the social and economic environment (Mirkatouli et al., 2018; Waltert & Schläpfer, 2010; Wang et al., 2017; Yi & Huang, 2014) but also to the urban form, which is led by urban growth strategies.

For the different urban forms in the five dimensions, our results show that land use development policies for low compactness, high centrality, high polycentricity, low density, and high dissimilarity in housing development may significantly increase the average housing price of the city's housing market. This phenomenon indicates that paying attention to the housing unaffordability within the cities due to the influences of house location should also regulate the housing market on the city level through appropriate urban planning and design. For example, when considering developing a new urban sub-center, constructing

high-density communities could provide more houses for migrants and reduce the housing prices (Fesselmeyer & Seah, 2018). Our research proposes to examine the impacts of the compactness, centrality, polycentricity, and dissimilarity in housing development and the density on housing prices in the evaluation of urban development projects.

For the distribution of urban functional zones within the cities, housing prices are affected by the spatial distribution of urban functions is well studied in previous studies (Qiao et al., 2021; Tam et al., 2019; Wang et al., 2017; Wen et al., 2017). But our finding provides a unified perspective for such single-city housing price studies (Shen & Karimi, 2017). After controlling the city as an influencing factor of housing price differences, working centers, high-quality hospitals, and schools can increase the housing prices of neighboring areas.

This research has its limitations. The spatial form of a city is always more or less related to its level of economic development. The world's most economically developed global cities (e.g., Chicago, Tokyo, Hong Kong, New York, Shanghai) are high-density cities with high housing prices. We cannot rashly infer that if a city is actively designed and built into a certain type of form, it will automatically push up or decrease local housing prices. The focus of this research is to answer whether the design and planning of the city's physical facilities and urban functional zones' layouts can certainly affect the housing prices if the social and economic conditions are equivalent to other cities. The findings emphasize that the moderating role of urban design and planning in social inequality caused by the housing unaffordability is feasible and our conclusions highlight a series of urban design indicators as a toolbox for planners and policymakers.

Appendix

See Tables 6, 7, 8, 9.

Table 6 POI categorizations of Gaode Map used in this study

Top type (13 categories, all)	Middle type (119 categories, examples)	Bottom type (2168 categories, examples)
Accommodation service	Hotel, inn	Economical hotel, motel, hostel
Commercial housing	Residential building, industrial park	Mansion, apartment, creative industry center
Daily life service	Ticket office, express	Train ticket office, barber shop, express
Educational and cultural service	School, museum	University, high school, science museum
Financial and insurance service	Bank, insurance	Bank of China, China Pacific Insurance Company, ATM
Firms and enterprises	Company, factory	Telecom company, medical company, construction company
Food and beverage	Chinese food, fast food	Sichuan food, fast food, sea food
Government agencies and social organizations	Social organization, government organization	Charity, court, city hall
Health care service	Hospital, clinic	General hospital, clinic, drug store
Shopping service	Supermarket, mall	Wal-Mart, book store, clothing store
Sports and recreation	Sports, recreation	Fitness center, cinema, playground
Tourist attraction	Scenic spot, park and square	Park, temple, botanical garden
Transportation facility	Airport, train station	Airport site, train station, subway station

Table 7 Disparities of urban forms and urban functional zones among megacities

Variables/Cities	Shanghai	Beijing	Shenzhen	Guangzhou	Chengdu	Hangzhou	Wuhan	Nanjing	Changsha	Xi'an
Proximity to working zone [kilometer]	0.402	0.440	0.476	0.545	0.614	0.380	0.881	0.538	0.541	0.670
Proximity to shopping center [kilometer]	1.094	1.272	1.251	1.260	0.963	1.092	1.277	1.173	1.136	1.098
Proximity to commercial business zone [kilometer]	1.951	2.685	3.453	3.262	2.045	2.172	4.532	2.700	4.385	2.836
Proximity to park and green space [kilometer]	1.484	1.500	0.575	1.031	2.245	1.253	2.023	1.979	1.032	1.731
Proximity to 3A hospital [kilometer]	5.442	5.263	4.759	4.987	5.721	6.634	4.070	4.764	2.545	2.899
Proximity to primary and middle school [kilometer]	0.620	0.731	0.608	0.563	0.633	0.738	0.652	0.724	0.674	0.649
Compactness [index]	0.062	0.064	0.110	0.117	0.187	0.165	0.172	0.178	0.286	0.182
Centrality [index]	0.619	0.619	1.005	0.847	0.656	0.706	0.786	0.848	0.668	0.642
Polycentricity [number]	6	6	5	3	4	3	2	2	1	2
Building area ratio [ratio]	1.710	1.019	1.587	1.695	1.465	1.570	1.439	1.074	1.631	1.635
Dissimilarity [index]	0.469	0.482	0.232	0.478	0.364	0.396	0.442	0.536	0.589	0.390

Note: the highest three values are highlighted as red for each variable; the lower values are in blue

Table 8 Result of variance inflation factor (VIF)

Variables	VIF
Building year	1.23
Proximity to metro station	1.73
Proximity to bus station	1.14
Proximity to working zone	1.23
Proximity to shopping center	1.11
Proximity to commercial business zone	1.72
Proximity to park and green space	1.62
Proximity to 3A hospital	2.22
Proximity to primary and middle school	1.43
Compactness	4.08
Centrality	5.61
Polycentricity	4.25
Building density	1.85
Dissimilarity	6.83
Gini index of population	4.33
Moran index of education	3.77
GDP per capita	3.32

Table 9 Fisher-Hayter pairwise comparisons of HEPrice for ten study cities

City vs. City	City means#	Mean diff	FH-test	
Beijing vs. Guangzhou	65.5	35.9	29.57	96.3627*
Beijing vs. Shanghai	65.5	65.1	0.42	1.6286
Beijing vs. Nanjing	65.5	34.0	31.47	97.7773*
Beijing vs. Wuhan	65.5	16.4	49.08	156.6610*
Beijing vs. Chengdu	65.5	12.0	53.44	202.4649*
Beijing vs. Xi'an	65.5	14.1	51.39	147.9686*
Beijing vs. Hangzhou	65.5	39.2	26.28	78.5239*
Beijing vs. Changsha	65.5	10.0	55.47	139.5118*
Beijing vs. Shenzhen	65.5	61.6	3.84	11.5009*
Guangzhou vs. Shanghai	35.9	65.1	29.16	103.4482*
Guangzhou vs. Nanjing	35.9	34.0	1.89	5.5054*
Guangzhou vs. Wuhan	35.9	16.4	19.50	58.1683*
Guangzhou vs. Chengdu	35.9	12.0	23.87	82.3705*
Guangzhou vs. Xi'an	35.9	14.1	21.82	59.3953*
Guangzhou vs. Hangzhou	35.9	39.2	3.30	9.2788*
Guangzhou vs. Changsha	35.9	10.0	25.89	62.3638*
Guangzhou vs. Shenzhen	35.9	61.6	25.74	72.6323*
Shanghai vs. Nanjing	65.1	34.0	31.05	104.1867*
Shanghai vs. Wuhan	65.1	16.4	48.66	168.5061*
Shanghai vs. Chengdu	65.1	12.0	53.03	226.2522*
Shanghai vs. Xi'an	65.1	14.1	50.98	156.6581*
Shanghai vs. Hangzhou	65.1	39.2	25.86	82.9332*
Shanghai vs. Changsha	65.1	10.0	55.05	145.4147*
Shanghai vs. Shenzhen	65.1	61.6	3.42	11.0098*
Nanjing vs. Wuhan	34.0	16.4	17.61	50.4725*
Nanjing vs. Chengdu	34.0	12.0	21.98	71.9427*
Nanjing vs. Xi'an	34.0	14.1	19.93	52.4602*
Nanjing vs. Hangzhou	34.0	39.2	5.19	14.0845*
Nanjing vs. Changsha	34.0	10.0	24.00	56.3024*
Nanjing vs. Shenzhen	34.0	61.6	27.63	75.2126*
Wuhan vs. Chengdu	16.4	12.0	4.36	14.7141*
Wuhan vs. Xi'an	16.4	14.1	2.31	6.2039*
Wuhan vs. Hangzhou	16.4	39.2	22.80	63.1849*
Wuhan vs. Changsha	16.4	10.0	6.39	15.2093*
Wuhan vs. Shenzhen	16.4	61.6	45.24	125.7092*
Chengdu vs. Xi'an	12.0	14.1	2.05	6.1722*
Chengdu vs. Hangzhou	12.0	39.2	27.16	85.1580*
Chengdu vs. Changsha	12.0	10.0	2.02	5.2634*
Chengdu vs. Shenzhen	12.0	61.6	49.60	156.0455*
Xi'an vs. Hangzhou	14.1	39.2	25.11	64.2650*
Xi'an vs. Changsha	14.1	10.0	4.07	9.1383*
Xi'an vs. Shenzhen	14.1	61.6	47.55	121.9721*
Hangzhou vs. Changsha	39.2	10.0	29.19	66.9353*
Hangzhou vs. Shenzhen	39.2	61.6	22.44	59.2678*
Changsha vs. Shenzhen	10.0	61.6	51.63	118.6187*

Table 9 (continued)

Studentized range critical value (0.05, 9, 47,306) equals to 4.3865448

the unit of price is 1,000 CNY

* Means the difference is significant between two cities

References

Adnan, M., Longley, P. A., Singleton, A. D., & Brunsdon, C. (2010). Towards real-time geodemographics: Clustering algorithm performance for large multidimensional spatial databases. *Transactions in GIS, 14*(3), 283–297.

Alonso, W. (1964). *Location and Land Use*. Harvard University Press.

Ashby, D. I., & Longley, P. A. (2005). Geocomputation, geodemographics and resource allocation for local policing. *Transactions in GIS, 9*(1), 53–72.

Batty, M. (2008). The size, scale, and shape of cities. *Science, 319*(5864), 769–771.

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., & Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics, 214*(1), 481–518.

Batty, M., & Longley, P. A. (1994). *Fractal Cities: A Geometry of Form and Function*. Academic press.

Cao, R., Tu, W., Yang, C., Li, Q., Liu, J., Zhu, J., Zhang, Q., Li, Q., & Qiu, G. (2020). Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS Journal of Photogrammetry and Remote Sensing, 163*, 82–97.

China, N. B. o. S. o. (2020). *China Statistical Yearbook 2020*. China Statistics Press

Clifton, K., Ewing, R., Knaap, G. J., & Song, Y. (2008). Quantitative analysis of urban form: A multidisciplinary review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 1*(1), 17–45. https://doi.org/10.1080/17549170801903496

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., Karagiorgou, S., Efentakis, A., & Lamprianidis, G. (2015). Crowdsourcing urban form and function. *International Journal of Geographical Information Science, 29*(5), 720–741.

Dempsey, N., Brown, C., Raman, S., Porta, S., Jenks, M., Jones, C., & Bramley, G. (2008). *Elements of Urban Form*. (pp. 21–51). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8647-2_2

Dempsey, N., Brown, C., Raman, S., Porta, S., Jenks, M., Jones, C., & Bramley, G. (2010). Elements of urban form. In *Dimensions of the Sustainable City* (pp. 21–51). Springer.

Dong, T., Jiao, L., Xu, G., Yang, L., & Liu, J. (2019). Towards sustainability? Analyzing changing urban form patterns in the United States, Europe, and China. *Science of the Total Environment, 671*, 632–643.

Du, S., Du, S., Liu, B., Zhang, X., & Zheng, Z. (2020). Large-scale urban functional zone mapping by integrating remote sensing images and open social data. *Giscience & Remote Sensing, 57*(3), 411–430.

Duncan, M. (2011). The impact of transit-oriented development on housing prices in San Diego. *CA. Urban Studies, 48*(1), 101–127.

Echenique, M. H., Hargreaves, A. J., Mitchell, G., & Namdeo, A. (2012). Growing cities sustainably: Does urban form really matter? *Journal of the American Planning Association, 78*(2), 121–137.

Feng, X., & Humphreys, B. R. (2012). The impact of professional sports facilities on housing values: Evidence from census block group data. *City, Culture and Society, 3*(3), 189–200.

Fesselmeyer, E., & Seah, K. Y. S. (2018). The effect of localized density on housing prices in Singapore. *Regional Science and Urban Economics, 68*, 304–315.

Galster, G., Hanson, R., Ratcliffe, M. R., Wolman, H., Coleman, S., & Freihage, J. (2001). Wrestling sprawl to the ground: Defining and measuring an elusive concept. *Housing Policy Debate, 12*(4), 681–717.

Gao, Q., Fu, J., Yu, Y., & Tang, X. (2019). Identification of urban regions' functions in Chengdu, China, based on vehicle trajectory data. *PLoS ONE, 14*(4), e0215656.

Gao, S., Janowicz, K., & Couclelis, H. (2017). Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS, 21*(3), 446–467.

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics, 5*(4), 471–484.

Goodman, A. C., & Thibodeau, T. G. (1998). Housing market segmentation. *Journal of Housing Economics, 7*(2), 121–143.

Goodman, A. C., & Thibodeau, T. G. (2003). Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics, 12*(3), 181–201.

Grislain-Letrémy, C., & Katossky, A. (2014). The impact of hazardous industrial facilities on housing prices: A comparison of parametric and semiparametric hedonic price models. *Regional Science and Urban Economics, 49*, 93–107.

Hamidi, S., & Ewing, R. (2015). Is sprawl affordable for Americans? Exploring the association between housing and transportation affordability and urban sprawl. *Transportation Research Record, 2500*(1), 75–79.

Hu, S., Gao, S., Wu, L., Xu, Y., Zhang, Z., Cui, H., & Gong, X. (2021). Urban function classification at road segment level using taxi trajectory data: A graph convolutional neural network approach. *Computers, Environment and Urban Systems, 87*, 101619.

Hu, S., He, Z., Wu, L., Yin, L., Xu, Y., & Cui, H. (2020). A framework for extracting urban functional regions based on multiprototype word embeddings using points-of-interest data. *Computers, Environment and Urban Systems, 80*, 101442.

Huang, J., Lu, X. X., & Sellers, J. M. (2007). A global comparative analysis of urban form: Applying spatial metrics and remote sensing. *Landscape and Urban Planning, 82*(4), 184–197.

Jiang, S., Alves, A., Rodrigues, F., Ferreira, J., Jr., & Pereira, F. C. (2015). Mining point-of-interest data from social networks for urban land use classification and disaggregation. *Computers, Environment and Urban Systems, 53*, 36–46.

Kohlhase, J. E. (1991). The impact of toxic waste sites on housing values. *Journal of Urban Economics, 30*(1), 1–26.

Kulish, M., Richards, A., & Gillitzer, C. (2012). Urban structure and housing prices: Some evidence from Australian cities. *Economic Record, 88*(282), 303–322.

Lee, S.-W., Taylor, P. D., & Hong, S.-K. (2008). Moderating effect of forest cover on the effect of proximity to chemical facilities on property values. *Landscape and Urban Planning, 86*(2), 171–176.

Lefebvre, H. (2003). *The Urban Revolution*. U of Minnesota Press.

Lefebvre, H., Kofman, E., & Lebas, E. (1996). *Writings on Cities* (Vol. 63). Blackwell Oxford.

Lefebvre, H., & Nicholson-Smith, D. (1991). *The Production of Space*. Oxford Blackwell.

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., & Stein, A. (2016). Geospatial big data handling theory and methods: A review and research challenges. *ISPRS Journal of Photogrammetry and Remote Sensing, 115*, 119–133.

Liu, K., Yin, L., Lu, F., & Mou, N. (2020). Visualizing and exploring POI configurations of urban regions on POI-type semantic space. *Cities, 99*, 102610.

Liu, X., Derudder, B., & Wang, M. (2018). Polycentric urban development in China: A multi-scale analysis. *Environment and Planning B: Urban Analytics and City Science, 45*(5), 953–972.

Liu, X., & Long, Y. (2016). Automated identification and characterization of parcels with OpenStreetMap and points of interest. *Environment and Planning B: Planning and Design, 43*(2), 341–360.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., & Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers, 105*(3), 512–530.

Long, Y., Li, P., & Hou, J. (2019). Three-dimensional urban form at street block level for major cities in China. *Shanghai Urban Planning Review, 3*(3), 10–15.

Long, Y., Shen, Z., Long, Y., & Shen, Z. (2015). Discovering functional zones using bus smart card data and points of interest in Beijing. *Geospatial Analysis to Support Urban Planning in Beijing, 45*, 193–217.

Lynch, K., & Rodwin, L. (1958). A theory of urban form. *Journal of the American Institute of Planners, 24*(4), 201–214. https://doi.org/10.1080/01944365808978281

Mirkatouli, J., Samadi, R., & Hosseini, A. (2018). Evaluating and analysis of socio-economic variables on land and housing prices in Mashhad, Iran. *Sustainable Cities and Society, 41*, 695–705.

Niu, H., & Silva, E. A. (2020). Crowdsourced data mining for urban activity: Review of data sources, applications, and methods. *Journal of Urban Planning and Development, 146*(2), 04020007.

Qi, G., Li, X., Li, S., Pan, G., Wang, Z., & Zhang, D. (2011). Measuring social functions of city regions from large-scale taxi behaviors. *2011 IEEE International Conference on Pervasive Computing and Communications Workshops* (PERCOM Workshops),

Qian, J., Liu, Z., Du, Y., Liang, F., Yi, J., Ma, T., & Pei, T. (2021). Quantify city-level dynamic functions across China using social media and POIs data. *Computers, Environment and Urban Systems, 85*, 101552.

Qiao, S., Gar-On Yeh, A., & Zhang, M. (2021). Capitalisation of accessibility to dockless bike sharing in housing rentals: Evidence from Beijing. *Transportation Research Part D: Transport and Environment, 90*, 102640. https://doi.org/10.1016/j.trd.2020.102640

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*(1), 34–55.

Schneider, A., & Woodcock, C. E. (2008). Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Studies, 45*(3), 659–692. https://doi.org/10.1177/0042098007087340

Schwarz, N. (2010). Urban form revisited—Selecting indicators for characterising European cities. *Landscape and Urban Planning, 96*(1), 29–47.

Shen, Y., & Karimi, K. (2017). The economic value of streets: Mix-scale spatio-functional interaction and housing price patterns. *Applied Geography, 79*, 187–202.

Tam, V. W., Fung, I. W., Wang, J., & Ma, M. (2019). Effects of locations, structures and neighbourhoods to housing price: An empirical study in Shanghai, China. *International Journal of Construction Management, 45*, 1–20.

Teck-Hong, T. (2012). Housing satisfaction in medium-and high-cost housing: The case of Greater Kuala Lumpur. *Malaysia. Habitat International, 36*(1), 108–116.

Tsai, Y.-H. (2005). Quantifying urban form: Compactness versus' sprawl'. *Urban Studies, 42*(1), 141–161.

Tu, W., Cao, J., Yue, Y., Shaw, S.-L., Zhou, M., Wang, Z., Chang, X., Xu, Y., & Li, Q. (2017). Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. *International Journal of Geographical Information Science, 31*(12), 2331–2358.

Tu, W., Hu, Z., Li, L., Cao, J., Jiang, J., Li, Q., & Li, Q. (2018). Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sensing, 10*(1), 141.

Villar, J. G., & Raya, J. M. (2015). Use of a Gini index to examine housing price heterogeneity: A quantile approach. *Journal of Housing Economics, 29*, 59–71.

Waltert, F., & Schläpfer, F. (2010). Landscape amenities and local development: A review of migration, regional economic and hedonic pricing studies. *Ecological Economics, 70*(2), 141–152.

Wang, Y., Wang, S., Li, G., Zhang, H., Jin, L., Su, Y., & Wu, K. (2017). Identifying the determinants of housing prices in China using spatial regression and the geographical detector technique. *Applied Geography, 79*, 26–36.

Wassmer, R. W., & Baass, M. C. (2006). Does a more centralized urban form raise housing prices? *Journal of Policy Analysis and Management, 25*(2), 439–462. https://doi.org/10.1002/pam.20180

Wen, D., Huang, X., Zhang, L., & Benediktsson, J. A. (2015). A novel automatic change detection method for urban high-resolution remotely sensed imagery based on multiindex scene representation. *IEEE Transactions on Geoscience and Remote Sensing, 54*(1), 609–625.

Wen, H., & Tao, Y. (2015). Polycentric urban structure and housing price in the transitional China: Evidence from Hangzhou. *Habitat International, 46*, 138–146.

Wen, H., Xiao, Y., & Hui, E. C. (2019). Quantile effect of educational facilities on housing price: Do homebuyers of higher-priced housing pay more for educational resources? *Cities, 90*, 100–112.

Wen, H., Xiao, Y., & Zhang, L. (2017). School district, education quality, and housing price: Evidence from a natural experiment in Hangzhou, China. *Cities, 66*, 72–80.

Wen, H., Zhang, Y., & Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in Hangzhou, China. *Habitat International, 42*, 155–163.

Wu, C., Ye, X., Du, Q., & Luo, P. (2017). Spatial effects of accessibility to parks on housing prices in Shenzhen, China. *Habitat International, 63*, 45–54.

Xing, H., & Meng, Y. (2018). Integrating landscape metrics and socioeconomic features for urban functional region classification. *Computers, Environment and Urban Systems, 72*, 134–145.

Xu, G., Zhou, Z., Jiao, L., & Zhao, R. (2020). Compact urban form and expansion pattern slow down the decline in urban densities: A global perspective. *Land Use Policy, 94*, 104563.

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., & Mai, K. (2017). Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *International Journal of Geographical Information Science, 31*(4), 825–848.

Yi, C., & Huang, Y. (2014). Housing consumption and housing inequality in Chinese cities during the first decade of the twenty-first century. *Housing Studies, 29*(2), 291–311. https://doi.org/10.1080/02673037.2014.851179

Yi, D., Yang, J., Liu, J., Liu, Y., & Zhang, J. (2019). Quantitative identification of urban functions with fishers' exact test and POI data applied in classifying urban districts: A case study within the sixth ring road in Beijing. *ISPRS International Journal of Geo-Information, 8*(12), 555.

Yuan, F., Wei, Y. D., & Wu, J. (2020). Amenity effects of urban facilities on housing prices in China: Accessibility, scarcity, and urban spaces. *Cities, 96*, 102433.

Zhai, W., Bai, X., Shi, Y., Han, Y., Peng, Z.-R., & Gu, C. (2019). Beyond Word2vec: An approach for urban functional region extraction and identification by combining Place2vec and POIs. *Computers, Environment and Urban Systems, 74*, 1–12.

Zhang, C., Jia, S., & Yang, R. (2016). Housing affordability and housing vacancy in China: The role of income inequality. *Journal of Housing Economics, 33*, 4–14.

Zhang, J., Li, X., Yao, Y., Hong, Y., He, J., Jiang, Z., & Sun, J. (2021a). The Traj2Vec model to quantify residents' spatial trajectories and estimate the proportions of urban land-use types. *International Journal of Geographical Information Science, 35*(1), 193–211.

Zhang, L., Zhou, J., & Hui, E.C.-M. (2020). Which types of shopping malls affect housing prices? From the perspective of spatial accessibility. *Habitat International, 96*, 102118.

Zhang, M., Luo, Z., Qiao, S., & Yeh, A.G.-O. (2023). Financialization, platform economy and urban rental housing: Evidence from Chengdu. *China. Applied Geography, 156*, 102993.

Zhang, M., Qiao, S., & Yeh, A.G.-O. (2021b). Blemish of place: Territorial stigmatization and the depreciation of displaced villagers' resettlement houses in Chengdu, China. *Cities, 117*, 103330.

Zhang, X., & Du, S. (2015). A Linear Dirichlet Mixture Model for decomposing scenes: Application to analyzing urban functional zonings. *Remote Sensing of Environment, 169*, 37–49.

Zhang, X., Du, S., & Wang, Q. (2017). Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. *ISPRS Journal of Photogrammetry and Remote Sensing, 132*, 170–184.

Zhang, Y., Li, Q., Tu, W., Mai, K., Yao, Y., & Chen, Y. (2019). Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Computers, Environment and Urban Systems, 78*, 101374.

Zhi, Y., Li, H., Wang, D., Deng, M., Wang, S., Gao, J., Duan, Z., & Liu, Y. (2016). Latent spatio-temporal activity structures: A new approach to inferring intra-urban functional regions via social media check-in data. *Geo-Spatial Information Science, 19*(2), 94–105.

Zhong, C., Huang, X., Arisona, S. M., Schmitt, G., & Batty, M. (2014). Inferring building functions from a probabilistic model using public transportation data. *Computers, Environment and Urban Systems, 48*, 124–137.

Zhong, C., Zhu, Q., & Zhang, L. (2015). Scene classification based on the multifeature fusion probabilistic topic model for high spatial resolution remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing, 53*(11), 6207–6222.

Zhou, X., & Zhang, L. (2016). Crowdsourcing functions of the living city from Twitter and Foursquare data. *Cartography and Geographic Information Science, 43*(5), 393–404.