The compactness of spatial structure in Chinese cities: measurement, clustering patterns and influencing factors

Fangqi Zhao\textsuperscript{a,b}, Lina Tang\textsuperscript{a}, Quanyi Qiu\textsuperscript{a} and Gang Wu\textsuperscript{a,c}

\textsuperscript{a}Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China; \textsuperscript{b}State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China

\textbf{ABSTRACT}

Rapid urbanization in China has led to an excessive urban expansion of built-up areas, which makes quantitative research on compact city important. We adopted density and the degree of mixed land use to measure the compactness of 160 Chinese cities. Spatial autocorrelation analysis was performed to identify spatial clustering patterns, and the relationships between compactness and five variables were explored through regression models. The result shows that in nearly half of the cases, the calculated values of two indices are less than the average. The high or low values of density and the degree of mixed land use tend to be spatially clustered. The hot spot regions of density and the degree of mixed land use lie mainly in the south of China, while the north present as cold spots or the insignificant regions. Urban compactness can be affected by multifaceted factors and the relationships between compactness and five variables are not consistent throughout the areas of analysis. The GWR model can identify this phenomenon and provides a better fit than the OLS model. This study proposed a new approach to measure the compactness, and the results of GWR analysis can conducive to appropriate policy-making based on different local conditions.

\textbf{Introduction}

Urban sprawl is usually defined as the unplanned growth of urban space that is characterized by low-density, scattered or leapfrog development with single use land and poor accessibility (Ewing 1997; Galster et al. 2001; Gordon and Wong 1985; Hamidi and Ewing 2014). Urban sprawl has been identified as one of the most critical issues in the urbanization process for it has resulted in a series of adverse effects on urban sustainability. These effects include encroachment on agricultural land, excessive car use, waste of energy and resources, and lack of social interactions (Angel et al. 2011; D’Amour et al. 2017; Deng et al. 2015; Ewing 2008). Since urban sprawl become an urgent problem, the compact city, a mode of sustainable urban development, has drawn attention from researchers and city planners. With an aim of eliminating the urban sprawl, alleviating pollution and promoting social justice, the compact city emphasizes high density, mixed land use construction and convenient public transportation (Burton 2002; Dantzig and Saaty 1973; Neuman 2005). Since the 1990s, some European countries or regions have implemented compact development strategies in order to solve the dilemma of urban decay (Burton 2000; Faludi 2004), and they achieved success in controlling the urban sprawl as well as improving the quality of city life.

So far, the definition of compact city is still lacking. High density and mixed land use are the most mentioned terms when trying to describe a compact city (Burton, Jenks, and Williams 2003; Jenks and Burgess 2000; Newman 1992). Urban density describes the carrying capacity of urban land so that high density means carrying more urban activities on less land. It is believed that high-density construction can directly reduce the occupation of cultivated land and save land resources (Jenks and Burgess 2000). Moreover, researchers also stated that high-density construction is associated with higher urban infrastructure utilization efficiency, less energy consumption and more urban vitality and creativity (Angel et al. 2018; Burton 2002; Glaeser and Kahn 2008; Güneralp et al. 2017; Cervero 2001). The concept of mixed land use states that urban activities of different types that are assembled within a specific range of area can functionally complement each other (Jacobs 1993). The benefits of mixed land use are more relevant to the needs of urban residents, which include vivifying communities, improving the accessibility of services and facilities, reducing private vehicle use and increasing residents’ physical activity (Grant 2002; Yan, Merlin, and Rodriguez 2013).

Since the 1970s, China’s economy has accelerated. The rural population flocked to the cities and the growth of the built-up areas in China has kept...
expanding at an unprecedented rate (Wei and Ye 2014). However, the land use presents as urban sprawl because it has been widely found in Chinese cities that the built-up areas are expanding at a much faster rate than the urban population growth (Gao et al. 2016; Seto et al. 2011; Xu et al. 2019; Angel et al. 2010). It was worth noting that the underlying mechanism of China’s urban sprawl is somewhat different from developed nations such as the United States (Wang, Shi et al. 2019). Unlike the urban sprawl in the U.S, which is closely related to urbanization and automobile dependence (Batty, Xie, and Sun 1999; Brueckner 2000; Ewing 1997), these factors are not obvious in China. One of the primary driving factors of China’s urban sprawl is the development of region, and the demands from migrants for urban space (Tan et al. 2005). About half of the built-up areas’ growth is at the expense of arable land in China, which poses challenges to food security (Bai, Shi, and Liu 2014; Deng et al. 2006). In addition, with the increased size of the built-up areas, the use of private cars has indirectly increased, which has led to an increase in greenhouse gas emissions and a large consumption of energy sources (Bart 2010; Hankey and Marshall 2010).

Although China seems to be much more “compact” than the developed countries, compact development is of great significance for Chinese cities to conserve the cultivated lands reduce air pollution and promote low-carbon development (Zhao et al. 2011; Li et al. 2018; Wang et al. 2019).

Urbanization on such a large scale has made China’s urban sprawl an important issue on academic research (Wei and Ye 2014), which has also triggered scholarly exploration on compact development in China. The measurement of compactness and the efficiency of compact city are what scholars are most concerned with. In previous studies, some of the scholars used a series of indices for compactness based on remote sensing data to identify the shape features of the regions (Liu, Song, and Song 2014; Song et al. 2017; Zhao Song, and Shi 2011). While others carried out their research using statistical data and evaluated compactness from a socioeconomic perspective (Chen, Jia, and Lau 2008; Fang, Qi, and Song 2008; Ma and Jin 2011). However, among the numerous studies, little research has measured the compact city from the internal spatial structure dimension, especially on a nationwide scale. We suspect the reason for this is the absence of enough easily obtainable data which can reflect the internal spatial structure of a city. The interpretation of high-resolution remote sensing data is cumbersome, while the statistical data ignores the spatial details (Long et al. 2018). In order to explore the applicability of compact development in China, scholars have studied a lot on the effects of the compact urban form (Chen, Jia, and Lau 2008; Shi, Yang, and Gao 2016; Zhao et al. 2014). However, what makes compactness different in different places has been rarely quantitatively discussed, although this kind of discussion is of great importance. Since China has a vast geographic area, the environments, locations, industries and policies of various kinds have led to the spatial heterogeneity of the urban spatial structure. Therefore, the compact development strategies in different regions of China need to be adapted to local conditions.

Against this background, this paper aims to propose a quantitative approach to measure compactness in the dimension of urban spatial structure, and then to explore the spatial pattern of the compactness across Chinese cities as well as its influencing factors. The first crucial task of this study is to determine how to measure urban compactness, as we try not only to identify the spatial structure in a fine-grained way, but also hope to understand the overall pattern of Chinese cities. The advancement of big data mining technology and the wide applications of open geodata (POI, Weibo check-in data, etc.) make this measurement possible, and therefore we developed indices using a point of interest (POI) dataset from electronic maps used for navigation purposes. The advantages of this kind of datasets include high spatial resolution, almost worldwide coverage and open access (Long and Liu 2013), all of which meets our criteria. The other critical issue is to investigate how the influencing factors affect compactness across the study area. In our case, the relationships between compactness and the independent variables may inevitably vary over space, which is referred to as spatial nonstationarity (Brunsdon, Fotheringham, and Charlton 1996). Therefore, a GWR model will be more appropriate to examine the varying relationships than the commonly used OLS model, because it takes local samples for parameter estimation and produces spatially altering coefficients, which will solve the problem of spatial nonstationarity (Tu and Xia 2008).

In this study, according to the spatial structure characteristics of compact city, density and the degree of mixed land use were adopted as the indices to measure the compactness in 160 Chinese cities. Then, we identified the spatial relevance and disparities at global and local scales with spatial autocorrelation analysis. Furthermore, we used the OLS and GWR models to explore the relationships between compactness and the influencing factors. Our paper is organized as follows. This section gives the background of our study. Section 2 introduces the study area and data sources and describes the analytical methods in detail. Section 3 presents the results of the compactness measurement, spatial autocorrelation analyses and regressions analyses. We discussed the compact development in China in Section 4, and some limitations and future work are also presented in this section. Finally, we summarized our work in Section 5.
Materials and methods

Study area

In this study, we chose 160 cities in China as the research objects: four municipalities, five state-plan cities, 24 provincial capital cities and 127 ordinary prefecture-level cities (Figure 1). Cities in Xinjiang, Tibet, Taiwan, Hong Kong, and Macau were excluded from this study due to the limitation of data availability and the distances between the cities. The 160 cities have differences in social, economic, cultural and climate conditions with the average urban population ranging from 0.15 to 21.27 million. Therefore, the exploration results of these cities can basically represent the overall pattern of the whole country.

The built-up area of a city is concentrated with urban population and facilities, which is often considered to be the best spatial scale for studying urban spatial structure (Kenworthy and Hu 2002). Unfortunately, although we can determine the extent of the built-up areas from the map, the built-up areas are not the recognized administrative urban socio-economic statistical units in China. The municipal district (including urban areas and suburbs) is one of China’s administrative divisions under the jurisdiction of municipalities or prefecture-level cities, which is also a basic statistical spatial unit. In this study, we focused on the spatial structure of the built-up areas within the municipal district, while considering the availability of socio-economic data, our study uses statistical data at the whole municipal district level.

Data sources

1. POI dataset, road networks and built-up area boundaries in 2015

The POI dataset, road networks and built-up area boundaries used in this study were acquired from Gaode Map Services (https://www.amap.com/). Gaode is one of the best-known corporations in China that provide navigation and location services. POIs in this database are divided into twenty types, each of which contains several subtypes. In this study, we reclassified these POIs into the following five types: residences, enterprises, public administration and services sites, commercial sites and entertainments sites (Table 1). These represent the five fundamental functions a city provides for urban inhabitants. Duplicate records or POIs of rare public cognition (such as public lavatory, telephone booths and ATMs) were removed.

Figure 1. Distribution of 160 Chinese cities.
mixed land use of the whole study area. The formula for the degree of mixed land use can be expressed as:

$$ R = \frac{1}{n} \sum_{i=1}^{n} S_i $$

(2)

where $R$ denotes the degree of mixed land use, with the value ranging from 1 to 5. $S_i$ is the number of POI types in the $i$th grid, and $n$ is the number of grids (containing POI ≥ 1).

**Spatial autocorrelation analyses**

Spatial autocorrelation is the fundamental property of the spatial units. When the value of a spatial variable is dependent on the value of its adjacent variables, spatial autocorrelation exists (Getis 2010).

(1) Global Moran’s I

Global Moran’s $I$ is one of the most commonly used tests for global spatial autocorrelation, and it identifies whether a pattern is clustered, dispersed, or random, and detects the strength of the spatial effects. Global Moran’s $I$ can be expressed as (Cliff and Ord 1973):

$$ I = \frac{1}{S} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x})(x_j - \bar{x}) $$

(3)

$$ S = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} $$

(4)

where $w_{ij}$ is the spatial weights between feature $i$ and $j$, $x_i$ and $x_j$ are the values of $x$ on the location $i$ and $j$, $\bar{x}$ is the average of $x$, $n$ is the number of features, and $S$ denotes the aggregate of the spatial weights. The value of $I$ ranges from −1 to 1; a positive $I$ value suggests that the features are spatially clustered, while a negative $I$ value indicates that the features tend to be dispersed. Features are randomly distributed when $I = 0$.

(1) Getis-Ord $G_i^c$

As a global statistic, Global Moran’s $I$ only measures the overall pattern of data, and cannot determine the specific location of the clustering. The local spatial autocorrelation statistics focus on smaller regions and assess the spatial autocorrelation related to a particular spatial unit. Getis-Ord $G_i^c$ is one of the local spatial autocorrelation statistics which identifies the spatial clusters of high values (hot spot) or low values (cold

**Table 1. POI types and the primary subtypes they contain.**

| Type                          | Primary subtypes                      |
|-------------------------------|---------------------------------------|
| Residences                    | Residence communities, cottages, dormitories, serviced apartments |
| Enterprises                   | Corporations, office buildings, industrial parks |
| Public administration and service sites | Hospitals, schools, government agencies, social organizations, public security bureaus |
| Commercial sites              | Restaurants, shopping malls, convenience stores, banks, hotels |
| Entertainment sites           | Parks, scenic spots, zoos, open squares |

(2) Social and economic data

The social and economic data involved in the GWR analysis were derived from China City Statistical Yearbook 2016, the content of which covers the primary socio-economic statistical data of cities at all levels in China for 2015.

**Measuring the compactness of urban spatial structure**

As mentioned above, our goal is to reveal the internal spatial structure of cities across the whole country. Therefore, the indices should be concise, valid, and easy to replicate. Here, we propose brief methods for calculating density and the degree of mixed land use based on the POI data. POIs represent real geographical entities, and their categories follow the current land use classification. Although the POI data are different from the conventional land use data, they can reflect the land use patterns (Yue et al. 2017).

(1) Density

High density means that the limited urban space can accommodate more urban activities. Density for each city can be calculated by:

$$ D = \frac{N}{A} $$

(1)

where $D$ is the density, $N$ is the total number of POIs within the study area, and $A$ is the patch size of the study area (km$^2$).

(2) The degree of mixed land use

In this study, we introduced the concept of species richness in ecology to represent the degree of mixed land use, and the index was calculated based on a fishnet method. We generated a fishnet with cell size of $1 \times 1$ km and placed it on the top of the study area. Then, we calculated the richness of POIs in each grid, and used the average of the richness of POIs as the degree of
spot) with statistical significance, the formula of the $G'_i$ statistic can be expressed as (Getis and Ord 2010):

$$G'_i = \frac{\sum_{j=1}^{n} w_{ij} x_j - \bar{X} \sum_{j=1}^{n} w_{ij}}{S \sqrt{\frac{1}{n} \sum_{j=1}^{n} x_j^2 - \frac{1}{n} \left( \sum_{j=1}^{n} x_j \right)^2}} \tag{5}$$

where $w_{ij}$ is the spatial weights between feature $i$ and $j$, $x_j$ is the observed value of feature $j$, $n$ is the number of features, and:

$$\bar{X} = \frac{1}{n} \sum_{j=1}^{n} x_j \tag{6}$$

$$S = \sqrt{\frac{1}{n} \sum_{j=1}^{n} x_j^2 - \left( \bar{X} \right)^2} \tag{7}$$

The $z$-score is the normalized $G'_i$ based on its sample mean and variance.

**Geographically weighted regression model**

The OLS is the most commonly used conventional regression model, and it can be expressed as:

$$y = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \varepsilon \tag{8}$$

where $y$ is the dependent variable, $\beta_0$ is the intercept, $\beta_i$ is the coefficient for the independent variable $x_i$, and $\varepsilon$ is the random error.

The geographic weighted regression (GWR) can be considered an extension of the conventional linear regression because it incorporates the spatial properties of the data into the model. GWR can be expressed as (Fotheringham, Brunsdon, and Charlton 2003):

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{9}$$

where $y_i$ denotes the dependent variable at location $i$, in this case the value of urban density or land use diversity for the $i$th city, $\beta_0(u_i, v_i)$ is the intercept coefficient at location $i$, $x_{ik}$ indicates the value of the $k$th independent variable at location $i$, $\beta_k(u_i, v_i)$ is the $k$th local regression coefficient for the independent variable $x_{ik}$, $\varepsilon_i$ is the random error, and $\varepsilon_i \sim N(0, \sigma^2)$. In this study, an adaptive bisquare kernel function was adopted to choose an appropriate spatial weight, and the optimal bandwidth was obtained by the corrected Akaike information criterion (AICC) method.

The compactness of the urban spatial structure is influenced by various factors, such as terrain, economics, built-up environments, transportation, and the sociodemographic condition. We performed a stepwise regression on several primitive variables in order to prevent multicollinearity. Finally, the five independent variables (road network density, the proportion of residences, the tertiary industry to GDP, investment in fixed assets per square kilometer and population size) were used in the regression analysis (Table 2). Among them, road network density and the proportion of residences represent the built-up environments factor, and the tertiary industry to GDP, investment in fixed assets per square kilometer and population size represent the industrial structure factor, economy devotion factor and the socio-demographic condition factor, respectively.

**Results**

**Measurement results of density and the degree of mixed land use**

The measurement results showed that the values of $D$ ranged from 8.51 (Liupanshui, Guizhou) to 405.77 (Daqing, Heilongjiang), with a mean value of 107.02. The values of $R$ ranged from 2.28 (Zhongwei, Ningxia) to 4.21 (Nanjing, Jiangsu), with a mean value of 3.348. The mean value of $D$ and $R$ in high administrative level cities (municipalities, state-plan cities, and capital cities) are 137.77 and 3.515, respectively, which are higher than that in the ordinary prefecture-level cities (99.44 and 3.309). We performed a nonparametric test on $D$ and $R$ for the two groups of cities, and the results showed that the differences are significant (Figure 2(a,b)).

Figure 3 shows the scatter plot of $R$-$D$, and the coordinate system was divided into four quadrants with the mean value of $D$ and $R$. Thus, the 160 cities were classified into four categories that are as follows: HD-HR (the first

| Table 2. Description of independent variables. |
|-----------------------------------------------|
| **Factor system** | **Independent variable** | **Abbreviation** | **Definition** |
| Built-up environments | Road network density | RD | Length of road network per square kilometer |
| | The proportion of residence | PRES | NRES = \# of POIs of residence type / NRES denotes the count of grids containing POIs of residence type, \( n \) is the count of grids (containing POIs=1) |
| Industrial structure | The tertiary industry to GDP | TI | Output value of tertiary industry / Total GDP |
| Economy devotion | Investment in fixed assets per square kilometer | FIX | Investment in fixed assets / City area |
| Sociodemographic condition | Population size | POP | Average urban population of 2015 |

TI, and FIX were calculated based on the statistical data of municipal districts due to the difficulty in obtaining data from built-up areas.
quadrant, high density – high richness), LD-HR (the second quadrant, low density – high richness), LD-LR (the third quadrant, low density – low richness), and HD-LR (the fourth quadrant, high density-low richness).

A total of 58 cities belong to the HD-HR category, including Beijing, Shanghai, Shenzhen and Nanjing. Tianjin, Taiyuan, Shijiazhuang and 18 other cities belong to the LD-HR category. A total of 79 cities belong to the LD-LR category including Harbin, Shenyang, Hohhot and Ningbo. Only Anqing and Zhangye belong to the HD-LR category.

We also found that the relationship between D and R can be described by a logarithmic function. R increases rapidly with D during the initial stage, and as D continues to increase, the slope of the curve gets smaller, and the growth gradually enters the stationary phase. The logarithmic relationship between D and R indicates that urban density can be regarded as the foundation of the mixed land use, which also further quantitatively proves that density is the basis of urban vitality.

Spatial clustering pattern

We performed Global Moran’s I tests using the tool in ArcGIS 10.3. This tool also computed a z-score and p-value to evaluate the significance of the index. The results of the tests showed that there were significant positive spatial autocorrelations for D and R in 160 cities (Moran’s I > 0, z > 0, p < 0.05) and that the distribution of D and R tend to be spatially clustered (Table 3).

Then, the hot spot analysis tool in ArcGIS 10.3 was performed to calculate the $G^*_c$ statistics. Figure 4 shows the cold-hot spot pattern of D and R, and the results were divided into seven categories based on the z-scores.

Table 3. Results of Global Moran’s I test for urban density and land-use diversity.

| Index | Global Moran’s I | z-score | p-value |
|-------|-----------------|---------|---------|
| D     | 0.40            | 8.88    | 0.00    |
| R     | 0.36            | 8.03    | 0.00    |
The cold-hot spot pattern of D shares some similarities with the pattern of R. A remarkable difference was found between the northern and southern areas in China. The hot spot regions of D and R are distributed mainly in provinces such as Fujian, Guangdong, Guangxi, Yunnan, Guizhou and Sichuan. The cold spot regions of D are mainly located in and around Henan and Jilin while the cold spot regions of R are located in Ningxia Jilin as well as its vicinity. Other regions did not show significant clustering of high or low values.

**Estimation results and the spatial distribution of GWR coefficients**

Regression analysis was performed in GWR4 to explore how the five factors affected D and R. Table 4 gives the parameter summary of the OLS and GWR models. In the regression model of D, the $R^2$ and the adjusted $R^2$ increased from 0.693 to 0.847 and 0.681 to 0.801, respectively, with the AICc declined from 1671.71 to 1618.99. For the regression model of R, the $R^2$ and the adjusted $R^2$ improved from 0.85 to 0.89, and 0.84 to 0.87 respectively, with the AICc declined from $-95.25$ to $-115.88$. The higher values of $R^2$ and adjusted $R^2$ indicate the stronger explanatory power of a regression model, and a model with lower AICc denotes a better fit with the observed data. Therefore, the GWR models of D and R showed better performance than their corresponding OLS models. According to the parameter estimates of the OLS, FIX had the most potent impact on D among the five variables, followed by RD, PRES, POP, and TI, while PRES exhibited the strongest effect on R, followed by RD, POP, FIX and TI. From the results of GWR-D and GWR-R, it can be seen that the effects of the five

![Figure 4. Results of hot spot analysis for (a) Density and (b) The degree of mixed land use.](image)

| Parameters | GWR-D | GWR-R |
|------------|-------|-------|
| Intercept  | 107.377 | 3.351 |
| RD         | 27.980 | 1.010 |
| PRES       | 25.703 | 0.324 |
| TI         | 8.536 | 0.017 |
| FIX        | 29.153 | 0.053 |
| POP        | −14.794 | 0.059 |
| $R^2$      | 0.693 | 0.053 |
| Adjusted $R^2$ | 0.681 | 0.850 |
| AICc       | 1671.712 | $-95.256$ |

Table 4. Parameter summary of the regression models.
variables are not always significant across the study area. For the GWR-D model, the percentage of cities with significance ranged from 26.88% to 77.50%, and for GWR-R model, the range is 1.88% to 100.00%. The built-up environment factor (RD, PRES) affects the compactness of the greatest number of cities while the industrial structure factor affects the fewest.

Figures 5–9 illustrate the patterns of the varying local coefficients over space as well as the significantly correlated regions (significant coefficients based on t-test at $\alpha = 0.05$). For better visualization, we displayed the results on an entire city scale; meanwhile, the values of the coefficients were divided into five groups using the Jenks natural breaks classification system.

(1) Road network density

The primary function of urban roads is transportation, and urban roads also closely related to the daily life of the residents. A dense road network divides the land into smaller patches, so it can help avoid the functional isolation by big blocks; it also encourages the urban development of different functions, and it also provides more opportunities for social communications (Jacobs 1993). The signs of the coefficients of RD are positive throughout the study area in the two models (Figure 5), which means cities with higher road network densities also tend to have higher D and R. In the GWR-D model, the significantly correlated region was found in the western, southern and eastern coastal areas, where the road network density itself is also high. The southwest was the region most affected by RD. In the GWR-R model, 92.50% of the cities were significantly correlated with RD, and the exception was inner Mongolia and its vicinity. The impact of RD decreased from south to north. Road network density in the north was lower than that in the south, and the differences between cities were also insignificant, which led to a lower sensitivity to the change of RD in the northern regions.

(2) The proportion of residence

The proportion of residential land in built-up areas in China was approximately 30% in 2015, which is lower than that of developed countries such as the United States and Japan. Overemphasis on industrial production and contempt for city life has created this pattern of land use structure in China. An increase in the proportion of residential land may have contributions to the enrichment of urban functions because it can lead to better urban facilities. The positive effect of increasing residential land on urban compactness has also been confirmed in the case of Malaysia (Abdullahi, Pradhan, and Mojaddadi 2017). In the GWR-D model, the significantly correlated region of Pres accounts for 77.50% of the study area (Figure 6). PRES shows a positive relationship with D in most of the study areas, with the exception of the Sichuan-Chongqing urban agglomeration. It is worth noting that the Sichuan-Chongqing urban agglomeration is precisely the region with high urban density and high proportion of residence, indicating that in these regions, a continuing increase in the proportion of residential land would cause urban sprawl rather than an increase of urban density. The southeastern coastal area is the most sensitive to the variation of Pres. In the GWR-R model, PRES has a significant positive effect on R over
the whole study area. In contrary to the RD, the coefficients of Pres show an apparent increasing trend from south to north.

(3) The tertiary industry to GDP

The tertiary industry in China is mainly based on service industries, public administration, and other knowledge-intensive industries. When compared with the secondary industry such as the energy industry, steel production, and conventional manufacturing, the tertiary industry is more environmentally friendly, occupies a smaller amount of land, and attracts a greater labor force. These characteristics should help form a more compact urban spatial structure. However, it can be seen from the Figure 7 that the influence of TI has no obvious association with D and R. The significant correlation area only accounts for a small fraction of the study area especially in the GWR-R model. The coefficients of TI decrease from south to north in the two models; coefficients are positive in the south and gradually become negative in inner Mongolia. The reason for this pattern we speculated is that China’s economy is mainly driven

Figure 6. Spatial distribution of regression coefficient of PRES.

Figure 7. Spatial distribution of regression coefficient of TI.
by the secondary industry, and the funds for urban construction are basically supported by the secondary industry. Although TI is positively correlated with D and R in a large number of cities, their effects are indirect. In resource-based cities, the effects of TI are even negative.

(4) Investment in fixed assets per square kilometer

In China, the high values of FIX aggregate mainly in two regions; the southwest and the eastern coastal area. The cities in the southwest usually have a smaller built-up area, so the investment in fixed assets per unit area seems to be more efficient; however, in the eastern coastal area, is due to the larger total investment in fixed assets. In the GWR-D model, FIX is significantly correlated with D in 63.13% of the cities, with the exception of the cities in the north China plain and the northeast. The coefficients in the significantly correlated region form a structure of circle layer of decreasing progressively from west to southeast, which indicate that the southwest is more sensitive to a change in FIX than the eastern coastal area. In the GWR-R model, the significantly correlated region is
located in the south of China. A total of 95% of the coefficients are positive; the signs of the coefficients are negative only in Ningxia.

(5) Population size

The population size is used as a standard for the city size in China. In the GWR-D model, the signs of the coefficients varying across regions (Figure 9). POP is significantly negatively correlated with D in the western area, indicating that a larger city tends to have a lower D. In the middle and coastal areas of China, POP played a positive role, although the impact is not significant. The effect of POP becomes negative again in the northeastern, while the impact is insignificant. In the GWR-R model, the coefficients exhibit spatial variation as well. POP has the most significant impact on R in southwestern China, and the effect is displayed by a hierarchical structural decrease in the direction of the northeast. A positive effect was found around the Beijing-Tianjin-Hebei region, while in northeastern China, the coefficients become negative again.

Discussion

Compact development in Chinese cities

In previous studies, the population density has often been adopted as the indicator of compactness, although it is more inclined to measure the degree of urban “crowdedness” rather than “compactness” (Bardhan, Kurisu, and Hanaki 2015). Since we used POI data representing geographic entities in this study, D and R measure urban compactness from the perspective of both the physical environment and socioeconomic activities, and this also led to different conclusions from studies based on population density.

In our study, 79 cities were found belong to the LD-LR category, and this was the largest number of cities among the four types. Thirteen of the cities with high administrative levels, including economically developed cities are in the east such as Ningbo and Qingdao. These results indicated that land use in quite a number of Chinese cities has not been highly efficient, and there is room for Chinese cities to become more compact. We have found that the hot-cold spot patterns of D and R are not consistent with the population density in China. The high values of population density agglomerate mainly in the eastern coastal areas of China. The Yangtze River Delta and the Pearl River Delta are the most significantly clustered regions. While in our study, the high values of D and R were clustered in Guangdong, Fujian, and most of the southwest part of China.

Since the concept of the compact city was first put forward, it has been subjected to academic dispute. Although numerous studies have confirmed the contribution of the compact city on environmental and social development, some scholars worry that an overcompact city might lead to more severe problems and therefore cast doubt on its applicability in developing countries (Gordon and Richardson 1997; Neuman 2005; Williams 2004). Nevertheless, the fact in China, according to our study, is that the land use in some cities is still inefficient. As more urbanization will occur in China, there is a need for cities to make room for the steady stream of immigrants (Angel et al. 2005; Seto et al. 2011). The viewpoint of this study is that in the context of inevitable urbanization, Chinese cities should increase their carrying capacity of immigrants and provide diversity services to residents by implementing a compact development strategy rather than expanding the boundaries of cities.

Limitations and future work

There are several limitations that exist in our current study. First, high density and mixed land use are the main spatial structure characteristics of the compact city, but they are not the only two features. It is not sufficient to equate “high density” and “mixed land use” with the idea of compact city. As the compactness indices were built based on the principles of briefness and applicability, our measurements may not comprehensively reflect the features of compact city. More open data sources are needed to support the examination of other features such as the construction of the public transport system and livability. Second, in the GWR analysis, some other important factors that may have a significant effect on compactness were not included. One reason is the lack of data. The collection of statistical data in some small cities of China is still incomplete which made some independent variables inapplicable. On the other hand, some social factors such as cultural customs, architectural features and local policies are difficult to quantify. Finally, this study did not answer the question that scholars are most concerned about which is what is the most appropriate compactness for sustainable urban development. In future studies, multi-sourced data should be adopted to complement the measurement of compactness and the regression analysis, and the performance of urban spatial structure should be analyzed for reasonable compactness from the economic, environment and social equity dimensions. Additionally, the specific organizations of urban spatial structure, such as the land mixed mode, are worth exploring further.

Conclusion

In this study, we built two indices based on the spatial structure characteristics of the compact city and chose POI as the data source to evaluate the compactness of
160 Chinese cities. When compared with the commonly used demographic data, the POI can reflect the space utilization pattern from both the physical and socioeconomic perspectives. Spatial autocorrelation analysis was performed to identify the clustering pattern, and we also explored the relationship between compactness and five influencing factors with a GWR model.

From the measuring density and the degree of mixed land use, we found that the values of D and R among the 160 cities were quite different. Most of the cities are of the “LD-LR” type, which means that the values of the two indicators are both less than the mean values. Significant positive spatial dependence and heterogeneity were found in the distributions of D and R. The hot spot regions of D and R are located in the southern part of China, such as Fujian, Guangdong, Guangxi, Yunnan, Guizhou and Sichuan, while the northern part of China is presented as the cold spots or insignificant clustering areas. Regression analysis revealed how the five influencing factors synthetically affected the compactness in Chinese cities from the dimensions of the built-up environment, economy devotion, industrial structure and sociodemographic conditions. The explanatory powers of the five variables are different in the regression models with D and R. We also found that the relationship between compactness and the five influencing factors were not constant over space, and the GWR model can identify the spatial nonstationarity and showed better performance than the corresponding OLS model. In the GWR model, the signs, sizes, and significance of local estimated parameters showed spatial heterogeneity, due to the difference in economics, industries and social conditions between the cities. In general, among the five variables, the built-up environment factors (RD and Pres) exhibited the broadest range of significant effective regions.

In the context of rapid urban growth, quantitative studies on the compact city for China are necessary. Our study provides a new method to quantitatively evaluate the compactness of urban spatial structure, which can be used in combination with other open data in the future for a comprehensive analysis. Meanwhile, as an effective technique for explaining the spatially varying relationship, the GWR model may be conducive to appropriate policy-making based on different local conditions.

Disclosure Statement
The authors declare no conflicts of interest.

Funding
This study was supported by the National Natural Science Foundation of China [Project No.: 41471137].

References
Abdullahi, S., B. Pradhan, and H. Mojaddadi. 2017. “City Compactness: Assessing the Influence of the Growth of Residential Land Use.” *Journal of Urban Technology* 25 (1): 21–46. doi:10.1080/10630732.2017.1390299.

Angel, S., J. Parent, D. L. Civco, A. Blei, and D. Potere. 2011. “The Dimensions of Global Urban Expansion: Estimates and Projections for All Countries, 2000–2050.” *Progress in Planning* 75 (2): 53–107. doi:10.1016/j.progress.2011.04.001.

Angel, S., J. Parent, D. L. Civco, and A. M. Blei. 2010. The Persistent Decline in Urban Densities: Global and Historical Evidence of ‘Sprawl’. Lincoln Institute of Land Policy.

Angel, S., S. A. Franco, Y. Liu, and A. M. Blei. 2018. “The Shape Compactness of Urban Footprints.” *Progress in Planning*. doi:10.1016/j.progress.2018.12.001.

Angel, S., S. C. Sheppard, D. L. Civco, R. With, A. Buckley, L. Chabaeva, A. Gitlin, J. Kraley, M. Parent, and Perlin. 2005. “The Dynamics of Global Urban Expansion.” The World Bank.

Bai, X., P. Shi, and Y. Liu. 2014. “Realizing China’s Urban Dream.” *Nature* 509 (750): 158–160. doi:10.1038/509158a.

Bardhan, R., K. Kurisu, and K. Hanaki. 2015. “Does Compact Urban Forms Relate to Good Quality of Life in High Density Cities of India? Case of Kolkata.” *Cities* 48: 55–65. doi:10.1016/j.cities.2015.06.005.

Bart, J. L. 2010. “Urban Sprawl and Climate Change: A Statistical Exploration of Cause and Effect, with Policy Options for the EU.” *Land Use Policy* 27 (2): 283–292. doi:10.1016/j.landusepol.2009.03.003.

Batty, M., Y. Xie, and Z. Sun. 1999. *Dynamics of Urban Sprawl*. London, UK: Centre for Advanced Spatial Analysis (UCL).

Brueckner, J. K. 2000. “Urban Sprawl: Diagnosis and Remedies.” *International Regional Science Review* 23 (2): 160–171. doi:10.1177/016001700761012710.

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. 1996. “Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity.” *Geographical Analysis* 28 (4): 281–298. doi:10.1111/j.1538-4632.1996.tb00936.x.

Burton, E. 2000. “The Compact City: Just or Just Compact? A Preliminary Analysis.” *Urban Studies* 37: 1969–2001. doi:10.1080/00420980050162184.

Burton, E. 2002. “Measuring Urban Compactness in UK Towns and Cities.” *Environment Planning B: Planning & Design* 29 (2): 219–250. doi:10.1068/b2713.

Burton, E., M. Jenks, and K. Williams. 2003. *The Compact City: A Sustainable Urban Form*. London, UK: Routledge.

Cervero, R. 2001. “Efficient Urbanisation: Economic Performance and the Shape of the Metropolis.” *Urban Studies* 38 (10): 1651–1671. doi:10.1080/00420980120084804.

Chen, H., B. Jia, and S. S. Y. Lau. 2008. “Sustainable Urban Form for Chinese Compact Cities: Challenges of a Rapid Urbanized Economy.” *Habitat International* 32 (1): 28–40. doi:10.1016/j.habitatint.2007.06.005.

Cliff, A. D., and J. K. Ord. 1973. *Spatial Autocorrelation*. London, UK: Pion.

D’Amour, C. B., F. Reitsma, G. Baiocchi, S. Barthel, B. Güneralp, K. Erb, H. Haberl, F. Creutzig, and K. C. Seto. 2017. “Future Urban Land Expansion and Implications for Global Croplands.” *Proceedings of the National Academy of Sciences* 114 (34): 8939–8944. doi:10.1073/pnas.1606036114.

Dantzig, G. B., and T. L. Saaty. 1973. *Compact City: A Plan for a Liveable Urban Environment*. San Francisco, USA: W.H. Freeman.

Deng, X., J. Huang, S. Rozelle, and E. Uchida. 2006. “Cultivated Land Conversion and Potential Agricultural Productivity in
Williams, K. 2004. Can Urban Intensification Contribute to Sustainable Cities? An International Perspective. City Matters: Official Electronic Journal of Urbanicity.

Xu, G., L. Jiao, M. Yuan, T. Dong, B. Zhang, and C. Du. 2019. “How Does Urban Population Density Decline over Time? An Exponential Model for Chinese Cities with International Comparisons.” Landscape and Urban Planning 183: 59–67. doi:10.1016/j.landurbplan.2018.11.005.

Yan, S., L. Merlin, and D. Rodriguez. 2013. “Comparing Measures of Urban Land Use Mix.” Computers Environment & Urban Systems 42 (7): 1–13. doi:10.1016/j.compenvurbsys.2013.08.001.

Yue, Y., Y. Zhuang, A. G. O. Yeh, J. Xie, C. Ma, and Q. Li. 2017. “Measurements of POI-based Mixed Use and Their Relationships with Neighbourhood Vibrancy.” International Journal of Geographical Information Science 31 (4): 658–675. doi:10.1080/13658816.2016.1220561.

Zhao, J., L. Xiao, L. Tang, L. Shi, X. Su, H. Wang, Y. Song, and G. Shao. 2014. “Effects of Spatial Form on Urban Commute for Major Cities in China.” International Journal of Sustainable Development & World Ecology 21 (4): 361–368. doi:10.1080/13504509.2014.922132.

Zhao, J., Y. Song, and L. Shi. 2011. “Study on the Compactness Assessment Model of Urban Spatial Form.” Acta Ecologica Sinica 31 (21): 6338–6343.

Zhao, J. Z., Y. Song, L. N. Tang, L. Y. Shi, and G. F. Shao. 2011. “China’s Cities Need to Grow in a More Compact Way.” Environmental Science & Technology 45 (20): 8607–8608. doi:10.1021/es203138c.