Spatial non-stationarity and heterogeneity of metropolitan housing prices: the case of Guangzhou, China

Shaopei Chen, Ming Fang* and Dachang Zhuang

Department of Land Resource Management, Guangdong University of Finance and Economics, Guangzhou, China

*Corresponding author’s e-mail: fangming@gdufe.edu.cn

Abstract. Taking a case study of the city of Guangzhou, this paper applies sophisticated approaches, including spatial autocorrelation analysis models, traditional regression models and Geographically Weighted Regressing (GWR) model to analyse and discuss the spatial non-stationarity and heterogeneity characteristics of housing prices, in terms of the data of the average annual housing prices of each sub-district/town from 2009 to 2015. In this study, four types of influence factors, i.e., location condition (distance to central business district (CBD)), traffic access level, commercial service facility and public service facility are selected. As results, the spatial-temporal evolution characteristic of the housing prices in Guangzhou is closely related to the layout and development of the urban areas. Furthermore, among the selected factors, location condition is the most important factors affecting the housing prices, followed by commercial service facility, traffic access level and public service facility. Specifically, based on the GWR model analysis, the spatial patterns of the regression coefficients of location condition and the public service facility illustrate “east strong and west weak” of the housing prices in Guangzhou. By contraries, “west strong and east weak” of housing prices according to the traffic access level and commercial service facility. The overall pattern of spatial non-stationarity and heterogeneity of housing prices presents the circle structure characteristics with the centre of the CBD in Guangzhou.

1. Introduction

The real estate industry has gradually become the primary pillar industry of the national economic development in China since 1990s. Particularly, in the past 20 years, the rapid development of the real estate industry has caused the housing price in the cities of China to rise dramatically. In this scenario, the price changes of commercial housing are increasingly concerned by the relevant research organizations, scholars, and interest groups. Currently, it has become one of the most controversial issues in the social and economic development in China. Therefore, the study focuses on the critical issues, which involving the changes and spatial differences of housing prices and influencing factors, in the domains of economic geography and urban sustainable development in China (Yao et al., 2014).

Due to the commodity attributes of the house, since 1970s, the Heonic Model (Rosen, 1974) has been widely used in the study of housing prices (Setvension, 2004; Sirmans, 2005; Raymond 2002; Michaels and Smith, 1990; Le, 2015; Ma and Li, 2003). As the definitions by the Hedonic Model, the commodity prices should be composed of several different factors. Since the number of factors and the combination of the price of each commodity are different, diverse commodities should have various prices. Therefore, the Hedonic Model is usually used to decompose the influencing factors of housing prices and apply the regression model to quantify the influence of factors, thus for revealing the
diverse characteristics of housing prices (Rosen, 1974). Nevertheless, the spatial characteristic of the dwelling house makes housing price different from the general commodity price, mainly in the case of the spatial non-stationarity and heterogeneity of the housing price. However, the Hedonic Model is built based on the traditional regression model because of lacking consideration of the spatial characteristics of housing prices. This causes the spatial non-stationarity and heterogeneity of housing prices to be difficultly identified. Therefore, in recent years, several spatial econometric models such as spatial interpolation, spatial autocorrelation analysis, exploratory spatial data analysis (ESDA) and geographically weighted regressing (GWR) have been widely used to study the spatial characteristics and influencing factors of housing prices (Lu et al., 2014; Smerth and Smith, 2000; Lee et al., 2016; Grespo et al., 2007; Helbich et al., 2014). Among these models, the local spatial autocorrelation analysis model has obvious advantages in discussing the spatial characteristics of housing prices, furthermore, the GWR model expands the traditional regression model and introduces the local location factors that accurately reflect the spatial non-stationarity and heterogeneity of housing prices. Currently, the GWR model is widely applied to deal with the issues on quantitative detection of the spatial heterogeneity of housing prices and the influencing factors under the specific spatial relationships (Brunsdon et al., 1998). Also, the GWR has been an effective way to achieve spatial non-stationary analysis of housing prices (Tu and Xia, 2008).

Taking a case of the city of Guangzhou, China, this paper applies the sophisticated approaches, including spatial autocorrelation analysis models, traditional regression models and GWR model to analyse and discuss the spatial non-stationarity and heterogeneity characteristics of housing prices, supported by the data collected by 141 sub-districts/towns in the city of Guangzhou from 2009 to 2015. In this study, four types of influential factors, i.e., location condition (distance to central business district, traffic access level, commercial service facility and public service facility are selected.

2. Methodology

2.1 Regression Model

The Hedonic Model is built by the regression model, which aims to provide regression analysis of the commodity prices by quantifying different attributes of commodity into multiple explanatory variables (independent variables). The regression model can be generally presented as following:

\[ Y \approx f(X, \beta) \]  

(1)

In equation (1), \( Y \) is a dependent variable and \( X \) is independent variables, \( \beta \) is denoted as the unknown parameters, which may represent a scalar or a vector.

2.2 GWR Model

In the past decades, the geographically weighted regression (GWR) model has been used widely in the field of spatial methodology research. The GWR model is given to study the mutual spatial impacts and relationships between two or more variables with geographical differences based on the principle of linear regression analysis. Therefore, the GWR is the expansion of ordinary line regression model (OLR) (Sun and Xu, 2016).

The changes of housing prices are affected by complex and diverse factors. In order to ensure quantitative analysis under the context of GWR model, this paper focus on four factors, including location condition (\( \Delta \)), traffic access level (\( T \)), commercial service facility (\( C \)) and the construction level of public service facility (\( P \)). Each factor involves with one or more indicators (see table 1).

| Table 1. Factors and specific indicators. |
|-----------------------------------------|
| Factors | Indicators                          |
| \( D \) | distance to centre business district \( (d_1) \) |
| \( T \) | bus stop density \( (t_1) \) |
|         | number of metro station \( (t_2) \) |
|         | road density \( (t_3) \) |
| \( C \) | commercial building density \( (c_1) \) |
financial institution density \((c_2)\)
retail store/supermarket density \((c_4)\)

\[
P = \begin{array}{c}
\text{educational institution density (}p_1) \\
\text{medical institution density (}p_2) \\
\text{government agency density (}p_3) \\
\text{park/square density (}p_4) \\
\end{array}
\]

As the sub-districts/towns are identified as the research units, the average annual housing prices of each sub-district/town from 2009 to 2015 is collected. Furthermore, for a sub-district/town, the arithmetic average processing is performed on the factors with multiple indicators, including traffic access level, commercial service facility and public service facility.

Firstly, the ratio of a specific indicator value (such as road density or educational institute density) of sub-district/town \(i\) to the average indicator value of all sub-districts/towns is calculated, which is defined as the indicator degree \((A)\):

\[
A_i = \frac{x_i}{\sum_{i=1}^{n} x_i}/n
\]

As shown in equation (2), \(x_i\) is a specific indicator value of sub-district/town \(i\), \(n\) is the number of all sub-districts/towns.

Then, for each factor, the average indicator degree can be attained, which is identified as factor intensity indicator \((K)\). For example, the bus stop density \((t_1)\), number of metro station \((t_2)\) and road density \((t_3)\) are the indicators of the factor of traffic access level \((T)\) of sub-district/town \(i\), \(K\) can be calculated as following:

\[
K_i^T = \frac{A_i^{t_1} + A_i^{t_2} + A_i^{t_3}}{3}
\]

According to equation (3), other factor intensity indicators can be attained respectively, i.e., \(K_i^C\), \(K_i^D\) and \(K_i^P\).

### 3. Results and Analysis

#### 3.1 Spatial Autocorrelation Analysis

**3.1.1 Global Autocorrelation Analysis.** According to the use of global spatial autocorrelation analysis tool (Global Moran I) based on first order polygon contiguity provided by ArcGIS, the results are listed in table 2. As shown in Table 2, the average annual housing prices of each sub-district/town (total 141) in Guangzhou from 2009 to 2015 shows positive spatial autocorrelation characteristics, furthermore, illustrates significant spatial clustering statistics \((Z>1.96, P<0.001)\). This implies that when the housing prices in a sub-district/town is high, the housing prices of the surrounding sub-districts/towns are also high, and vice versa.

| Index | 2009     | 2010     | 2011     | 2012     | 2013     | 2014     | 2015     |
|-------|----------|----------|----------|----------|----------|----------|----------|
| Moran I | 0.195145 | 0.365298 | 0.268577 | 0.288585 | 0.272338 | 0.346694 | 0.549534 |
| Expectation | -0.009901 | -0.009901 | -0.009901 | -0.009901 | -0.007143 | -0.007143 | -0.007143 |
| Variance | 0.004628  | 0.004643  | 0.004591  | 0.004643  | 0.002473  | 0.002483  | 0.002465  |
| Z      | 3.014119  | 5.506398  | 4.109836  | 4.380566  | 5.620302  | 7.101614  | 11.211547 |
| P      | 0.002577  | 0.000000  | 0.000040  | 0.000012  | 0.000000  | 0.000000  | 0.000000  |
3.1.2 **Local Autocorrelation Analysis.** In order to explore whether the characteristic of spatial clustering statistics of the housing prices is a high-value aggregation or a low-value aggregation, and more accurately detect the spatial distribution pattern of housing prices in Guangzhou, the tool (Anselin Local Moran I) of local spatial autocorrelation analysis provided by ArcGIS is used. Figure 1 illustrates the spatial aggregation pattern of local indicators of spatial association (LISA) of the housing prices in Guangzhou in 2015. The housing prices aggregation pattern can be divided into four types, i.e., high-high (HH), low-low (LL), high-low (HL) and low-high (LH) for identifying the spatial relationship between a sub-district/town and its surrounding sub-district/town. In figure 1, HH aggregation is found into Yuexiu district (the political, historical and cultural centre of Guangzhou), Centre Business District (the economic centre of Guangzhou) and other key urban development areas, such as the Pearl River Coastal zone and the Baiyun new town. Moreover, HH areas are surrounded by LH areas. However, HL areas cannot be found, this implies that the high housing prices in the area have a certain spatial spillover effect. The LL areas are in the north and south Guangzhou. However, in the east Guangzhou, there is no significant housing prices aggregation, which reflecting that the housing prices in east Guangzhou are homogeneous.

![Figure 1. spatial aggregation distribution of local indicators of spatial association (LISA) of the housing prices in Guangzhou](image)

3.2 **Regression Model Analysis**

3.2.1 **Linear Regression Model Analysis.** In order to investigate the overall (average) impact on housing prices changes derived from the four factors of location condition \(D\), traffic access level \(T\), commercial service facility \(C\) and the public service facility \(P\), this paper applies a linear regression model in the environment of Microsoft Excel. In the linear regression model, the factor intensity indicators (i.e., \(K^D, K^T, K^P, K^C\)) are identified as the explanatory variables (independent variables). The results according to the data in 2015 are listed in table 3.

| Coefficients | Standard Error | t-Stat | P-value |
|--------------|----------------|--------|---------|
| Intercept    | 30388.85677    | 1543.939504 | 19.68267325 | 1.25352E-41 |

Table 3. Results of regression model analysis in 2015.
As shown in Table 3, the goodness of fit ($R^2$) of regression model is 0.60, and Significance $F < 0.000000001$ implies that the confidence is over 99.9999%. According to the Coefficients listed in Table 3, it can be found that the greatest impact on the housing prices is the factor of location condition, followed by commercial services intensity, traffic access level and the construction level of public service facilities. Specifically, when shorten the distance from the Central Business District by 1 kilometre, it causes the housing prices to rise to 636 yuan (about 100 dollars) per square meter. And the factor indicator value of commercial services facility increases by 1 unit, causing the housing prices to rise to 346 yuan per square meter. It can be found that the impact on housing prices changes from the factor of public service facility is weak. The reason is in Chinese traditional concepts, the residential area close to hospitals and other medical institution are not favourable conditions, and even have a negative impact on housing prices. The more detail can be seen in Table 4.

### Table 4. Results of regression model analysis of construction level of public service facilities in 2015.

| Coefficients | Standard Error | t-Stat | P-value |
|--------------|----------------|--------|---------|
| Intercept    | 20015.2         | 1125   | 17.791* | 0.00000 |
| $p_1$        | 744.486         | 371.051| 2.27074*| 0.04475 |
| $p_2$        | 157.37          | 130.695| 1.2041* | 0.23066 |
| $p_3$        | 115.329         | 130.695| 1.76768*| 0.07937 |
| $p_4$        | 2162.49         | 890.242| 2.42911*| 0.01645 |

### 3.2.2 GWR Model Analysis. The geographically weighted regression model (GWR Model) in the environment of ArcGIS is used to explore the spatial heterogeneity and non-stationarity of the housing prices in Guangzhou. Furthermore, the Akaike information criterion (AIC) is applied to optimize the bandwidth. In GWR Model, the 141 sub-districts/towns in Guangzhou are identified as research units, and their average housing prices in 2015 are defined as the dependent variables, and the factor intensity indicators (i.e., $K_D^i$, $K_T^i$, $K_P^i$, $K_B^i$) are independent variables. The results show that $R^2$ is 0.63, which is higher than the goodness of fit of the traditional linear regression model ($R^2 = 0.6$). According to the analysis of the global spatial autocorrection of the residuals in each research unit based on GWR Model, Moran’s I index is 0.03, Z score is 0.5, and P-value is 0.4. It implies that the residuals have the characteristics of spatial random distribution.

In term of the results derived from the GWR Model analysis, the impacts on the spatial heterogeneity and spatial non-stationarity of the housing prices in Guangzhou from the four factors are illustrated in Figure 2-5, respectively.

Figure 2 illustrates that the location condition has the significant impact on the housing prices in east and south Guangzhou. Although in the west Guangzhou, the Yuexiu, Liwan and Haizhu districts are close to the CBD, the factor of location condition show a weak impact on the housing price. As these three districts are the developed area in the city of Guangzhou, their housing prices are affected greatly by their own geographic conditions and Economic and human environments. In figure 3, under the influence of the factor of traffic access level, the housing prices are effected significantly in the west area (i.e., the Yuexiu, Liwan and Haizhu districts), and weakly in the east area. This implies that the traffic access level in the Yuexiu, Liwan and Haizhu districts are relatively higher than the...
other districts in the city of Guangzhou. Figure 4 indicates that the influence on housing prices of the commercial service facility radiates outward from the centre of the developed urban area of the the Yuexiu, Liwan and Haizhu districts. As shown in figure 5, on the whole, the public service facility has no significant impact on housing prices.

4. Conclusions
This study takes a case study of the city of Guangzhou, and applies the global and local spatial autocorrelation model to analyse the spatial cluster of the housing prices. Furthermore, the GWR model is used to investigate and discuss the influencing factors of the housing price changes, and then explain the price spatial heterogeneity and non-stationary characteristics caused by these factors. In this paper, the selection and processing of the quantitative indicators are still in the exploratory stage, and the existing research literatures have no perfect processing basis and solutions. Therefore, the goodness of fit affecting the regression models is not ideal, and the results may be biased. The future research work of this paper needs to implement more in-depth analysis and study on the selection and quantification of indicators, and combine comparative analysis and refinement with more cities in China to improve the current research.
Acknowledgments

This study was supported by the Guangzhou Science and Technology Project (No.201607010264).

References

[1] Moore R and Lopes J 1999 Paper templates. In TEMPLATE’06, 1st Int. Conf. on Template Proc. SCITEPRESS.
[2] Smith J 1998 The book, The publishing company. London, 2nd edition.
[3] Yao L, Gu G, and Wang J K 2014 The spatial effect of building new housing in Zhengzhou City-based on the spatial econometric model. Chinese J. of Economic Geog., 34(1): 69-74.
[4] Rosen S 1974 Hedonic prices and implicit markets: Product differentiation in pure competition. J. of Political Economy 82: 34-55.
[5] Stevenson S 2004 New empirical evidence on heteroscedasticity in hedonic housing model. J. of Housing Economics 13: 136-153.
[6] Sirmans G S, Macpherson D A and Zietz E N 2005 The composition of hedonic pricing models. J. of Real Estate Literature 13: 1-44.
[7] Raymond Y C T 2002 Estimating neighborhood effects in house prices: Towards a new hedonic model approach. Urban Economics 39: 1165-1180.
[8] Michaels R and Smith V 1990 Market segmentation and valuing amenities with hedonic models: The case of hazardous waste sites. J. of Urban Economics 28: 223-242.
[9] Le J M 2015 Influence factors analysis of housing price in Nanchang based on Hedonic Model. Chinese J. of East China Jiaotong University 32(5):128-135.
[10] Ma S X and Li A 2003 House price and its determinations in Beijing based on Hedonic Model. Chinese J. of Civil Engineering, 36(9): 59-64.
[11] Lu B B, Charlton M, Harris P and Fotheringham A S 2014 Geographically weighted regression with a non-Euclidean distance metric: a case study hedonic house price data. Int. J. of Geographical Information Science 28(4):660-681.
[12] Smerth G T and Smith M T 2000 Accessibility changes and urban house price appreciation: a constrained optimization approach to determining distance effects J. of Housing Economics 9:187-196.
[13] Lee G, Cho D and Kim K 2016 The modifiable areal unit problem in hedonic house-price models. Urban Geog. 37(2):223-245.
[14] Crespo R, Fotheringham S and Charlton M 2007 Application of Geographically Weighted Regression to a 19-year set of house price data in London to calibrate local hedonic price models. In GeoComputation 2007, 9th Int. Conf. on Geocomputation, National University of Ireland, Maynooth, Ireland.
[15] Helbich M, Brunauer W, Vaz E and Nijkamp P 2014 Spatial heterogeneity in hedonic house price models: the case of Austria. Urban Studies 51(2):390–411.
[16] Brunsdon C, Fotheringham S and Charlton M 1998 Geographically weighted regression. Journal of the Royal Statistical Society: Series D The Statistician 47(3):431–443.
[17] Tu J and Xia Z G 2008 Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. Science of the Total Environment 407:358-378.
[18] Sun K and Xu Z M 2016 The impacts of human driving factors on grey water footprint in China using a GWR Model. Chinese J.1 of Geographical Research 35(1):37-48.
[19] Tian J, Xiong F Q, Cheng X P, Wang R and Fang H Q 2016 Road density partition and its application in evaluation of road selection. Chinese J.1 of Geomatics and Info. Sci. of Wuhan University 41(9):1225-1231.
[20] Yang X, Webster C and Orford S 2014 Identifying house price effects of changes in urban street configuration: an empirical study in Nanjing, China. Urban Studies, 53:112-131.
[21] Cliff A D and Ord J K 1973 Spatial Autocorrelation. Pion. London: Pion.
[22] Anselin L 1995 The Local Indicators of Spatial Association-LISA. Geog. Analysis, 27:93-115.