DEVELOPING A 4D PROPERTY VALUATION MODEL BASED ON GEOSPATIAL DATA AT CITY SCALE (XI'AN, CHINA)

RUIJIE LI

February, 2019

SUPERVISORS:

Dr. M. N. Koeva

Dr. M. Kuffer

Dr. X.Li

ADVISOR:

Mr. MSc. Kwabena Asiama

DEVELOPING A 4D PROPERTY VALUATION MODEL BASED ON GEOSPATIAL DATA AT CITY SCALE (XI’AN, CHINA)

RUIJIE LI

Enschede, The Netherlands, February,2019

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Land Administration

SUPERVISORS:
Dr. M. N. Koeva
Dr. M. Kuffer
Dr. X.Li
ADVISOR:
Mr. MSc. Kwabena Asiama

THESIS ASSESSMENT BOARD:
Prof. Dr. Ir. J.A. Zevenbergen (Chair)
Dr. Maik Netzband (External Examiner, Name Institute)

Abstract

With the rapid urbanization in China, the residential property market has grown substantially over the last decade and has obtained high attention from the whole society. Xi'an as the biggest city in the north-west of China has a rapidly developing real estate market. After 2016, the price of properties in Xi'an started to increase faster than before. However, the spatial patterns of property prices in a complex urban environment, dominated by high-rise developments, is not well understood. This research based on geospatial and temporal data developed a 4D hedonic property price regression model to detect how property prices distribution and indicators change across the eight year.

Based on the indicators influencing property prices in Xi'an, selected from literature review, interview and questionnaires, spatial analysis and multiple linear regression model are used to develop a hedonic price model. Firstly, this research introduced the state-of-the-art on hedonic property price models relevant to the context of cities in China. According to previous researches, the property price is influenced by several indicators which can be categorized into location characteristics, environmental characteristics, and physical characteristics. Then appropriate indicators for Xi'an through interviews and questionnaires are collected. Spatial analysis is used to quantify the spatial indicators, followed by a multiple linear regression analysis to determine the most important indicators of property prices in 2010, 2014 and 2018. The result shows the change of the most indicators of local property prices. It can be found that local people give an increasing importance to environmental characteristics, comparing 2018 with previous years. However, the building height (an 3D indicator) is consistently throughout the years one of the most important indicators and plays a positive role in property prices.

The vizualizaion of results illustrates that the prices increased a lot after 2014, and the distribution of property prices also changed a lot during 2010 and 2018. Besides, the result is also visualized through CityEngine, which is a helpful tool to visualize the 3D indicators and be easily understood by potential property buyers and non-professional and technical officials in the government. It can help property buyers to understand the property market and provide suggestions when they not sure where to select for buying a property. For examples, it can provide suggestions for the buyers to select a cost-effective property with better physical, location and environmental characteristics and lower price.

Keywords: hedonic model, property valuation, geospatial and temporal data, CityEngine

ACKNOWLEDGEMENTS

For the one and a half year studying in ITC, I have obtained lots of help from many people and improved myself. It is exciting to say that after six months of effort, the master thesis is finished. On the occasion of the completion of this thesis, I would like to express my sincere gratitude and best wishes to all those who have helped me.

Firstly, I would like to express my gratitude to ITC and all staff in ITC for giving me the opportunity to study and complete my MSc thesis. Thanks to ITC for improving my science theory of geo-information, remote sensing and land administration. Besides, studying in ITC is my first-time study abroad, thanks to all staff in ITC who make me feel at home and I am so appreciated to meet many friends, teachers and enrich my life here.

I want to expand my sincere thanks to my supervisors, Dr. M. N. Koeva, Dr. M. Kuffer and Dr. X. Li for their patient guanidine, encouragement, support and constructive suggestions for my MSc research thesis. I would also like to express my appreciation to Mr. Kwabena Asiama, a Ph.D. candidate in LA, for giving me lots of suggestions and help to improve my thesis writing. Besides, thanks for Jong Wang, a Ph.D. candidate in AES, for proving helpful data for my research. Without their help, the research cannot be completed.

During the field work of this thesis, several in-depth interviews and a focus group discussion had been conducted. Gratefully acknowledgments are expressed to the interviewer respondents for their costing time to support my thesis research.

I own my gratitude to my friends, Ms. Jingxuan Zhang, Ms. Yue Ying, Ms. Shen Nie, Ms. Mengge Tian, Ms. Na Chen, and Ms. Li Liu, etc. They supported and encouraged me all the time and gave me many good suggestions for the research thesis. I really enjoy the time we spent together for cooking, traveling, chatting and learning.

Finally, special thanks to my parents and grandparents for their understanding, spiritual support and encouragement during my thesis research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1. Background and justification

Rapid urbanization and the growing housing price are one of the most pressing issues in urban China and has received common attention from scholars (Chao Wu et al., 2016). The change of the property value is a social issue that influenced and impacted the Chinese society in many ways (Li, Ye, Lee, Gong, & Qin, 2017). For example, the very high property price is unaffordable for some poor families (Jie Chen, Hao, & Stephens, 2010); on the other hand, the fluctuation of house prices will affect macroeconomic stability, and economic stability may affect social welfare (L. Chen & Fan, 2014). Therefore, the accurate estimation of property market values is important, as it can provide reliable support for city managers to make policies and property purchasers take their decisions (Wu, Deng, & Liu, 2014).

The process of property market value estimation is called property valuation (Kahr & Thomsett, 2005). During this process, many characteristics, which influence the price of a property, need to be considered. For instance, researchers found that the property prices are influenced by the property itself, location (Wyatt, 2013), accessibility to railway station (Debrezion, Pels, & Rietveld, 2007), waterside location (Garrod & Willis, 1994), noise from airport (Nguy, Sun, & Zheng, 2014), etc. These characteristics can be categorized into three types and researchers commonly evaluate the property price using them: physical characteristics, location, and environment (Jim & Chen, 2006; Kahr & Thomsett, 2005). In addition, the vertical dimension (building height) impacts the property value as well, because it is relevant to the time a person needs to reach from the building entrance to the property. Moreover, other factors as the view, orientation, size and position are of importance (Wong, Chau, Yau, & Cheung, 2011). In general, there is a spatial differentiation of real estate values in urban areas (F. Wu (2002).

Location, and environment characteristics are always influenced by the urban expansion (J. Xiao et al., 2006), in both horizontal and vertical direction (S. Liu, Fan, Wen, Liang, & Wu, 2014). The study by Qin, Fang, Wang, Li, and Wang (2015), for Yangzhou, a city in South China, has shown how three-dimensional urban expansion and buildings patterns changed. The time dimension is called the fourth dimension (Resch, Hillen, Reimer, & Spitzer, 2013). However, these researches are mainly related to 3D and 4D analysis on urban expansion not to property values. Property prices are dynamic and influenced by the change in factors, e.g., property prices are dynamic over time influenced by urban expansion. Chen, Guo and Wu (2011) examined empirically the factors of urban house prices using time-series and cross-sectional data for 29 cities and they found that property values were relative to the urbanization level. Besides, Li et al. (2017) found a clear relationship between property values, location and horizontally urban growth. It is clear that property prices are related to urban spatial expansion. However, their analysis was based only on 2D data.

To detect how indicators influence property prices, the hedonic pricing model is commonly used in this research domain (Abidoye & Chan, 2017). The basic theory of hedonic pricing models allows evaluating the benefits and services provided by the real estate (Andersson, Shyr, & Fu, 2010; Hamilton & Morgan, 2010). This model has been used for property valuation for a long time (Hamilton & Morgan, 2010). It is important for using hedonic model to decide the form of the function (Lee, Park, & Kim, 2003). Commonly used functions are: linear, quadratic, log-log and semi-log (Grislain-Letrémy & Katossky, 2014).

1.2. Research problem

In real life, property purchasers are willing to pay more for the best floor level, best views and a wonderful environment (Jim & Chen, 2009; Wong et al., 2011). In other words, property prices are related to location, accessibility and environment (Debrezion et al., 2007; Wyatt, 2013). Meanwhile, property prices are related to urban expansion (Li et al., 2017). However, most researches only focus on the relationship between property price and one of these factors. Though some researchers have studied more holistic sets of factors, most of them only focus on the two dimensions (Isikdag, Horhammer, Zlatanova, Kathmann, & Van Oosterom, 2015; Shin, Washington, & Choi, 2007). In other words, there is a lack of combining horizontal factors, vertical factors, the dynamics of urban growth (temporal factor) and property values in studies of property valuation. Consequently, this study is combining horizontal factors, vertical factors and time-series data, called 4D data (Resch et al., 2013) to analyse how urban growth and building heights impacts the property price for property valuation by using a hedonic model.

1.3. Conceptual framework

For this study, the concepts of property valuation, spatial-temporal (4D) data, hedonic property price model, and the characteristics of the model are the basic researched elements. The characteristics or indicators for property valuation include location, accessibility and environment, which cause the spatial differentiation of real estate value. Meanwhile, these factors are the basis for detecting the dynamic of property prices with urban growth. In that case, these factors are extracted for both spatial and temporal analysis.

The conceptual framework is shown in Figure 1 :

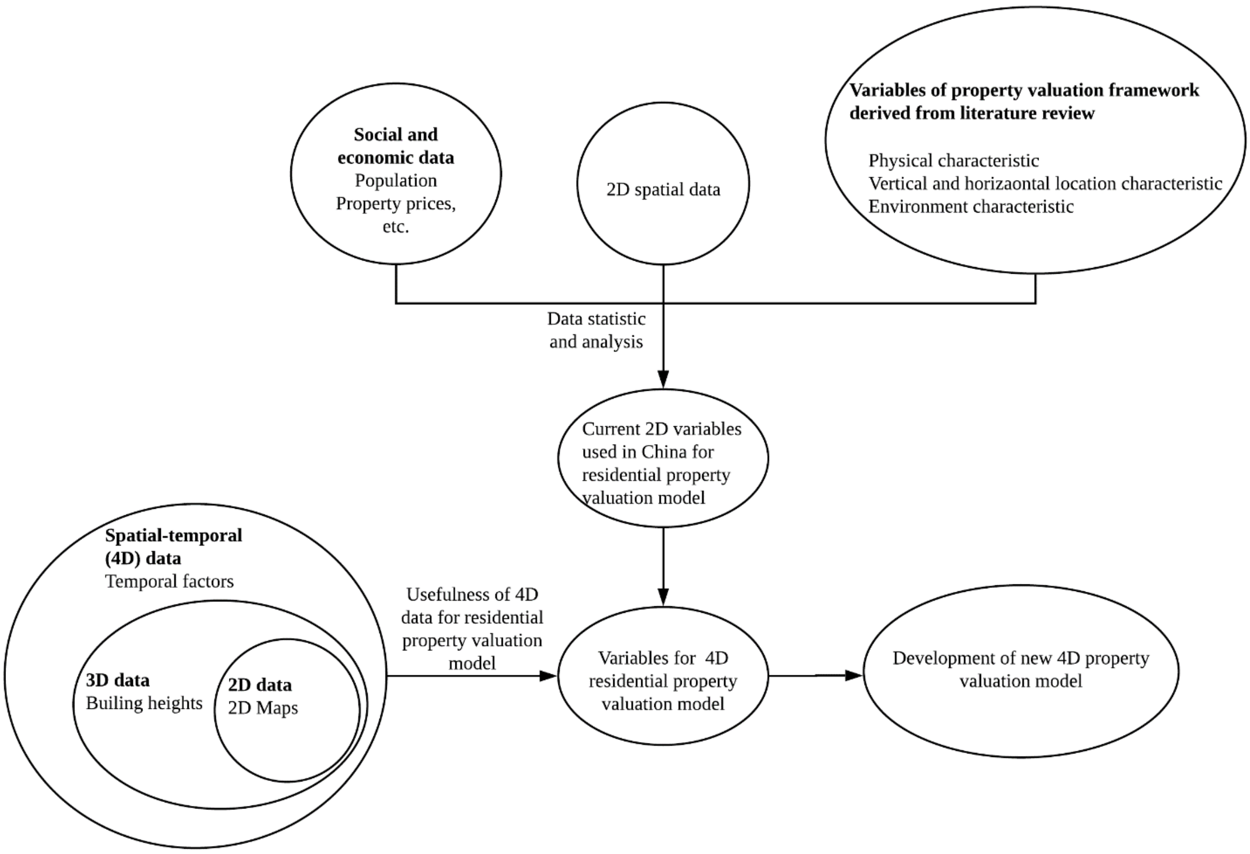


Figure 1 Conceptual framework

1.4. Research objectives

1.4.1. Main objective

To develop a 4D spatial hedonic property valuation model, which includes the temporal urban development and vertical factors besides 2D spatial factors.

1.4.2. Sub-objectives

1. To review the hedonic property pricing models in China.
2. To develop a 4D Property valuation model based on Remote Sensing Data
3. To evaluate the model

1.5. Research questions

1.5.1. To review the hedonic property pricing model in China.

- What is the state-of-the-art on hedonic property price model relevant to the context of cities in China?
- What are the benefits including 4D data for price valuation in property pricing model?

1.5.2. To develop a 4D Property valuation model

- What are the spatial patterns of the urban area of Xi'an over different times?
- Which current indicators are used for property valuation in Xi'an?
- What indicators would be beneficial to add in the 4D property valuation model?

1.5.3. To evaluate the model

- How to evaluate the model?
- Is the model acceptable and can be generalization?

1.6. Structure of the thesis

The first chapter provides a general overview of the research, including background and justification, research problem, research objectives and questions. Chapter two includes the literature review. Chapter three explains the methodology. Chapter four shows the result of the research and discussion. Chapter five and six is the discussion and the conclusion respectively.

2. Literature review

The hedonic model method is used to analyse a property price in a micro-view, which is different from traditional methods based on supply and demand, construction cost, etc. In hedonic models, the influence of property characteristics would be analysed and the price is determined by the satisfaction and preference of the buyer for the various characteristics of the house. In that case, the property valuation model can be developed through qualitative analysis to know the influence of various characteristics on house prices.

2.1. Hedonic model

Colwell and Dilmore (1999) believed that Haas was one of the first users of the hedonic model. Haas (1922) analyses farmland prices by a simple hedonic model based on the distance to the city center and the size of the city as two important characters. Waugh (1928) used multiple regression method, commonly used for a hedonic model, to analyse the relationship between vegetable price and quality factors. In 1939, the term "hedonic" was firstly coined by Andrew Court (Goodman, 1998). Court (1939) studied the effects of different characteristics of automobiles on their prices and tried to build a price index for the automobile industry. This research inspired other researchers, and this method was adopted into other consumer goods (Wen, Jia, & Guo, 2005). In the 1960s and 1970s, there was a great development of the hedonic model. Lancaster (1966) took the customer choices as the starting point of his research and suggested that the heterogeneity of goods is the root of consumers' demand for goods. Ridker and Henning (1967) used the hedonic model to analyse the market of residential real estate and found that there is a positive correlation between residential prices and air quality. Rosen (1974) constructed a property hedonic price model based on the theory of supply and demand equilibrium and hedonic model. After his study, scholars around the world have been adopting the hedonic model technique for property valuation (Abidoye & Chan, 2017). Since 1990s, some researchers have introduced other methods such as artificial neural networks, fuzzy logic, etc. into the hedonic model to enhance the interpretation of the hedonic model (Abidoye & Chan, 2017; Pagourtzi, Assimakopoulos, Hatzichristos, & French, 2003).

Hedonic pricing model decomposes the property into component characteristics and estimates the value of differences in property characteristics (Seo, Golub, & Kuby, 2014). These characteristics usually include structure characteristics, accessibility characteristics and neighbourhood characteristics (Wilhelmsson, 2002). Accessibility characteristics and neighbourhood characteristics contain the location factors, which is difficult to obtain in the traditional hedonic price model (Huang, Chen, Xu, & Zhou, 2017). In that case, many researchers started to study the spatial hedonic price model. Dubin (1998) used geostatistical methods (Kriging) to estimate the covariance structure of the model. Can (1990) used a spatial model to incorporate the spatial neighbourhood as a variable to capture its effects on property prices. With the development of GIS, GIS has been integrated with hedonic models and support improving the accuracy of property valuation (Huang et al., 2017). Researchers combined not only 2D data but also 3D data and temporal data respectively with the hedonic model. Table 1 is a summary of the work done recently.

Table 1 A summary of recent years studies

Authors	Study focus	Dimension	Study area
Kong, Yin, and Nakagoshi (2007)	The impact of urban green space on property prices	2D	Jinan City, China
Sander and Polasky (2009)	The impact of views and open space on residential property prices	2D	Ramsey County, Minnesota, USA
Andersson et al. (2010)	The impact of high-speed rail accessibility on residential property prices	2D	Southern Taiwan, China
Feng and Li (2011)	The significant variables of the property prices	2D	Qu Jiang New District, Xi'an, China
Wong et al. (2011)	The impact of vertical dimension on property prices	3D	Hong Kong
Y. Xiao, Webster, and Orford (2016)	The effects of changes in urban street configuration on property prices	Temporal data and 2D	Nanjing, China
Grislain-Letrémy and Katossky (2014)	The impact of hazardous industrial facilities on housing prices	2D	Bordeaux, Dunkirk, and Rouen, France
Noor, Asmawi, and Abdullah, (2015)	The influence of green space on housing prices	2D	Subang Jaya, Malaysia
Abidoye and Chan (2017)	Review of hedonic pricing model applications in property price valuation	2D	Nigeria
Yang, Zhou, Shyr, and Huo (2018)	The impact of bus accessibility on property prices	2D	Xiamen Island, China
Montero, Mínguez, and Fernández-Avilés (2018)	Comparison of different hedonic house pricing models for housing price prediction	3D	Madrid, Spain

2.1.1. The forms of the hedonic model formula

In general research of the hedonic model, commonly used function forms are divided into three types--linear, logarithmic and semi-logarithmic.

1. Linear function:

$$\mathrm{P} = b_0 + \sum a_i k_i + \varepsilon$$

This is the most common form, which is a linear regression of the characteristic variables. This function assumes that the characteristic variable and the price are linear (Zhang, 2014).

2. Logarithmic function:

$$\ln\mathrm{P} = b_0 + \sum a_i ln k_i + \varepsilon$$

In practice, sometimes the independent variables and dependent variables are not linear, so the form of logarithm can more accurately express their relationship, which is better for regression (Zhang, 2014).

3. Semi-logarithmic function:

$$\ln\mathrm{P} = b_0 + \sum a_i k_i + \varepsilon$$

Sometimes, this function can be more suit. Besides, the coefficient k represents the proportional relationship between the indicator (a_i) and the final price (P) (Zhang, 2014).

$$a_i = P_{k_i}/P$$

2.1.2. Comparison with traditional property valuation methods

1. Cost method

The property price is calculated based on the cost of the developer in a real estate project. The calculation method is shown following (Pagourtzi et al., 2003):

The sum of the property values of the project = replacement cost or reproduction cost – Depreciation+ land value

However, the cost method does not suit the situation that the land prices are volatile, and the objects are too old. In this research, the land price of Xi'an changed a lot during these years. Thus, the method is not suitable.

2. Comparable method

This method is used to estimate the price of the measured property based on the recent transaction price of similar house prices (Pagourtzi et al., 2003). This method needs a premise that the contrast object must be similar to the measured property. However, the residential property is so special that it is always unique and each property has its own attribute characteristics. However, the price of residential property in Xi'an increase rapid, the price of surrounding housing maybe not have the reference value.

3. Income capitalisation method

This method is based on the future benefits of the property to calculate the current value. It is more suitable for countries with stable housing prices and the results using this method in China are often unsatisfactory (Zhang, 2014). Due to inflation, housing prices are growing too fast, and there are too many uncertainties in the property market. This method is not very suitable for this study.

2.2. Common property value characteristics

As mentioned above, the hedonic pricing model is based on the characteristics of the property. There are many choices for characteristics, such as geographical location, greening, illuminance, floor level, transportation, etc. Butler (1982) divided the indicators that influence the price of the house into physical characteristics, the geographical location characteristics, and the surrounding environmental characteristics.

2.2.1. Physical characteristics

Scholars found that the price of residential property has a strong correlation with the physical characteristics, also named structural attributes, of the building itself. Ball (1973) argues that if one house with more desirable physical characteristics would have a higher price. Physical characteristic relates to the room numbers, the floor area, building height, age of the building, condition of the sunlight in the room, etc. Scholars has also done researches based on one of or part of these indicators. Fletcher, Gallimore, and Mangan (2000) prove that the number of rooms is positively correlated with the price of the house. Some studies suggest that the floor level is positively related to the property sale price (Carroll, Clauretie, & Jensen, 1996; Rodriguez & Sirmans, 1994). In addition, studies also discover that the building age is another important indicators negatively relating to the property value (Clark & Herrin, 2000; Rodriguez & Sirmans, 1994; Straszheim, 1975), which is because that the longer age of the building means the more cost for building maintenance (Salleh, Yakin, Ismail, & Talib, 2016). Kain and Quigley (1970) discover that the property structural and housing quality have a significant impact on residential prices. Glaeser, Gyourko, and Saks(2005) indicate that property prices tended to rise with the building height and they explain the reason of this phenomime is that taller building meant better sign views and better building qualities.

2.2.2. Location characteristics

The fixed location characteristics are quantified with respect to the whole urban area and pertain to some form of accessibility measure (Chau & Chin, 2002; Follain & Jimenez, 1985). The effects of location characteristics on property values have been noted by lots of studies. Locational characteristics include socioeconomic class, racial composition, aesthetic attributes, pollution levels, and proximity to local amenities (Chau & Chin, 2002; Dubin & Sung, 1990).

According to a research by Jie Chen and Hao (2008), the distance to city center significantly influences property prices, but the price gradient varies in different directions outwards from the city center, which is because the city is not completely homogeneous across the city (Adair, McGreal, Smyth, Cooper, & Ryley, 2000). At the same time, they find that the convenience of transportation around the house has an impact on the price of residential properties. For example, So, Tse, and Ganesan (1997) analyse the relationship between the transportation accessibility and property price based on the property market of Hong Kong, and their research proves that the property price is positively related to transportation convenience. Researchers have also found that the price of residential real estates around some landscapes is related to the convenience of reaching these landscapes (Benson, Hansen, Schwartz, Jr., & Smersh, 1998; Sander & Polasky, 2009). To be more exact, scholars indicate that the property buyers prefer to living near a good view, such as lakes and green space, and they are willing to pay more for these good views (Benson et al., 1998; Darling, 1973; Gillard, 1981; Sander & Polasky, 2009).

2.2.3. Surrounding environmental characteristics

Surrounding environmental characteristics are also understood as neighborhood characteristics, which refer to the socio-economic indicators, local government or municipal services and externalities (Chau & Chin, 2002). Clark and Herrin's (2000) research indicates that residential prices will decrease due to increased homicides, because of people's pursuit of residential safety. While according to Clark and Herrin (2000), existing educational institutions around the residential property has a greater impact on the property price thaneither crime or environmental quality, especially for those who have children. As for hospitals, Huh and Kwak (1997) indicated that hospitals exhibit a significant negative effect on property prices. While Peng and Chiang (2015) argued that hospitals would only be highly evaluated in a 'close-but-not-too-close' geographic location. In addition, the availability of surrounding facilities, such as entertainment or sports facilities, a landscaped garden and a clubhouse, have an positive effect on the property price (Chau & Chin, 2002; X. Feng & Humphreys, 2008; Mok, Chan, Kong, & Department, 1995; Tse & Love, 2000). Both proximities to the shopping center and the size of the shopping center are related to the surrounding housing prices (Des Rosiers, Lagana, Thériault, & Beaudoin, 1996; Sirpal, 1994). Besides, scholars also found that the noise from traffic would also influence the property price (Palmquist, 1992).

2.3. Indicators influencing the residential property price in China

As mentioned above, the property price is affected by physical characteristics, environmental characteristics, and location characteristics. According to the scholars, the physical characteristics include the building height, the building age, and the property area, etc.; the environment characteristics refers to distance to the parks, greening rate, etc.; the location characteristics include the distance to CBD, the distance to metro station, the number of nearby schools, the distance to hospital, etc. (A. Chen, 2006; Hai-zhen Wen, Zhang, & Peng, 2010; Haizhen Wen, Zhang, & Zhang, 2014; Yang et al., 2018). The indicators were collected as shown in the Table 2.

Table 2 Indicators influencing the property price in China

Authors	Type of characteristics	Indicators
Jim and Chen(2009)	Physical characteristics	Property age
Wong et al. (2011)		Building height
Jie Chen & Hao, 2008)	Location characteristics	Distance to city centre
Changshan Wu and Sharma (2012)		Distance to the city boundary
L. Wang and Sun (2014)		Accessibility to bus stops
Tian (2006)		Accessibility to nearby metro stations
Z. Chen and Haynes (2015)		Accessibility to train stations
Haizhen Wen et al. (2005)	Surrounding environment characteristics	Accessibility to shopping centres

Song (2012)	Accessibility to hospital
Kong et al. (2007)	Industrial pollution
Jim and Chen (2009)	Parks and landscape
Haizhen Wen et al. (2014)	Education facilities

2.4. Summary

This chapter introduces the development of the theory of hedonic property price model and compare the hedonic mode with traditional valuation methods. Besides, this chapter reviews the state-of-the-art on hedonic property price model indicators selecting based on physical characteristics, location characteristics and surrounding environment characteristics, and the most common indicators used in Chinese research are summarised.

3. Methodology

3.1. Study area

Xi'an is the capital of Shaanxi province and the economic center in Northwest China. The built-up area increased from 395 sq.km in 2010 to more than 566 sq.km in 2017 (Bureau of Statistics of Xi'an, 2017). The urban residential property prices per square meter increased a lot during the years of 2010 and 2017, from 4341yuan to 8531 yuan (NBSC, 2018). In 2010, the value-added by real estate industry was 186.42 billion yuan (Xi'an Statistical Bureau, 2018). In 2014 the value-added by real estate industry was 326.10 billion yuan. In 2017 the value-added by real estate industry was 550.35 billion yuan. The data of the value added by real estate in different years are shown as Table 3.

Table 3 The value added by the real estate industry in different years

(100 million yuan)

Year	Gross Domestic Product	Real Estate	Investment by Enterprises for Real Estate Development
2010	3242.86	186.42	670.41
2011	3869.84	235.20	836.05
2012	4394.47	260.53	1,003.85
2013	4924.97	292.41	1,226.28
2014	5492.64	326.10	1,321.91
2015	5801.20	398.34	1,304.60
2016	6282.65	459.20	1,337.35
2017	7471.89	550.35	1,505.03

The rapidly developing property market and urban area of Xi'an draws lots of attention from local people and local government (Y. Wang, Bai, & Wang, 2015). Face to the rapid increasing property price and changes of property price distribution across the city, buyers need suggestions to help them to make decisions about the location selection and prices. On the other hands, when local government make policies, the important indicators of property prices need to be known. So, an analysis of how urban expansion relates to property prices, and the development of a hedonic property price model for Xi'an is needed. Thus, Xi'an city is selected as the study area. The location of Xi'an city is shown as Figure 2.

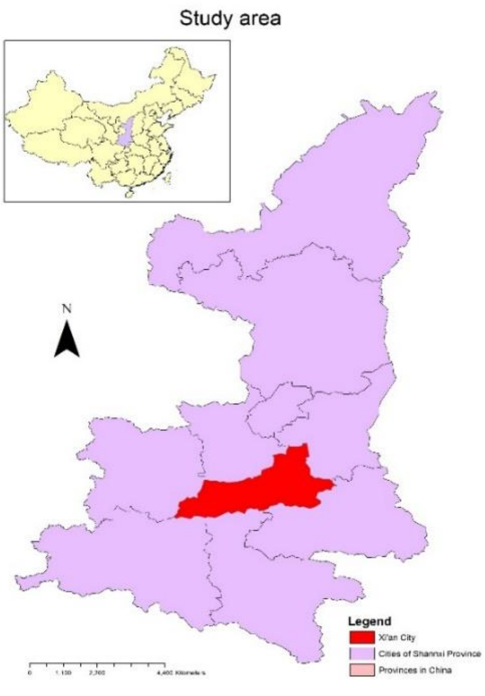


Figure 2 The location of Xi'an city

The urban area of Xi'an is centered on Tang Chang'an City, with the basic contour of the Ring Expressway, east to the Wei River, west to the Ring Expressway, south to Chang'an (Luo River), and north to the Wei River.

The spatial pattern is the administrative, commercial, and tourism center within the Ming City Wall; the military and light industrial areas centred on the Textile Town in the east of the city; the tourist resort centred on the Qujiang New District in the southeast; the research and cultural district centred on Yanta District and Chang'an District in the south; the high-tech industrial zone centred on the Xi'an Hi-tech Industries Development Zone in the southwest; and the industrial storage area centred on the Third Bridge in the west; Han Chang'an City in the northwest; processing and equipment industrial area centred on the northern economic and technological development zone; tourist resort centred around the Chan River, Ba River, and Wei Rivers in the northeast. The study area is shown as Figure 3.

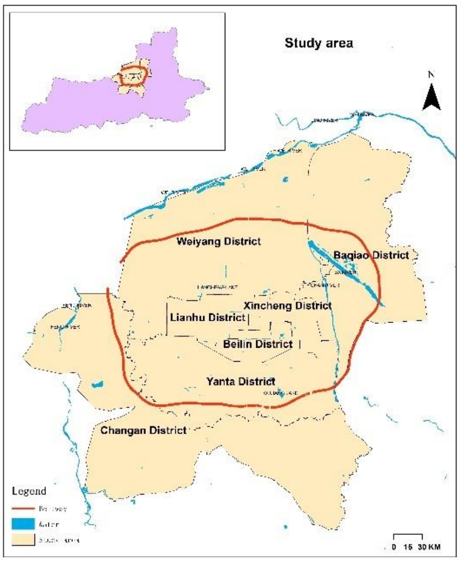


Figure 3 The study area

The urbanization process in Xi'an is with the northward shift of the city center, the construction of metros, the expansion of Qujiang District, the second venture of Gaoxin New District, the development of Daming Palace zone, the renovation of the old constructions in Daxing District, and the development of Textile Town. The expansion of the city promotes the rapid development of the property market and makes the development of various zones of Xi'an in a more balanced way. However, the north area and the east area of Xi'an have benefited the most from urbanization.

According to the status of commercial housing development in various districts of Xi'an, this thesis mainly takes Xi'an urban area, which is the area interior the beltway of Xi'an city, as the study area. At the same time, due to the development of real estate, the study area also extends south to Guodu Town and Weiqu Town in Chang'an District.

Introduction of Xi'an property market

After the cancellation of the welfare housing system in 1988, properties came into the real estate market and it began to be concerned by local residents. After 2008, a large number of commercial houses were built for sale. During this period, the housing prices in other cities in China grew rapidly, but the housing prices in Xi'an were relatively stable. Xi'an is one of the lowest-priced first-tier cities in China. Since 2016, Xi'an residential prices have started to grow rapidly, and house price growth has surged in the second half of 2017. People in Xi'an set off a boom in buying houses. The price of the residential property in different years is shown as Table 4.

Table 4 The average price of property in different years

Year	Property price (yuan/sq.m)	Area of residential properties for sale (10000 sq.m)
2010	4341.00	1523.24
2011	5829.79	1663.97
2012	6224.03	1379.13
2013	6435.00	1496.34
2014	6105.00	1514.15
2015	6221.00	1583.53
2016	6385.00	1866.50
2017	8166.00	2105.94

Since the rapid increase of the property price in Xi'an, the local price administration bureau and the department of property management has published several policies to control the price, especially the price of new properties.

1. To control the price of new residential properties, before the real estate company sell these new properties, the company should report the price of each apartment to Xi'an Municipal Price Bureau and get a permit from the Bureau (Xi'an Price Administration Bureau, 2017). This process is used to make sure the price of the new property is similar to the price of the surrounding buildings. So, if the company reports a very high property price to the Price Bureau, the company cannot get the permit and these residential properties are not allowed to be sold until both the bureau and government have an agreement of the price.
2. In the past, some real estate companies promised that they would provide opportunities for the children of the buyers to study in a good school to attract more buyers. These promises are not fared to the other households' children and stimulate increasing price of the properties near good schools. So, now the real estate development company is not allowed to contain the promise of attending school for the children of the purchaser (Xi'an Education Administration Bureau, Xi'an Housing Administration Bureau, & Xi'an Administration for Industry and Commerce, 2018).
3. Because of the development of Xi'an, more and more people move to this city. So sometimes the residential properties are in short supply, which may cause buyers have to pay more to compete for a residential property. To solve this problem, the local government issues that if the number of intention-purchasing houses is more than the number of saleable housing, it shall be publicly sold by means of notarization and lottery, and the notarization agency will supervise the whole process, and it is strictly forbidden to reserve the house internally or set the restrictive conditions such as the full-priority preferential selection (Xi'an Housing Administration Bureau, 2018).

3.2. Overall approach and methods

In this research both qualitative method and quantitative method were used. The qualitative method was mainly used in data collection and pre-process to select indicators to match the practice situation in Xi'an real estate market. And the quantitative method was used in spatial analysis and linear regression analysis parts.

3.3. Workflow

The research is separated into three phases: pre-fieldwork, field work and post-fieldwork. The aim of the first phase is having an overview of the research with research objectives and questions. Besides, during this phase, it needs to prepare for fieldwork and to do initial data collection as well. The second phase is to collect primary data through fieldwork. The main methods to collect are interview and questionnaires. The third phase aims at data statistic and analysis. Based on that, the 4D property valuation model will be developed. Therefore, after developing the model, it has been evaluated. Figure 4 illustrates the research design.

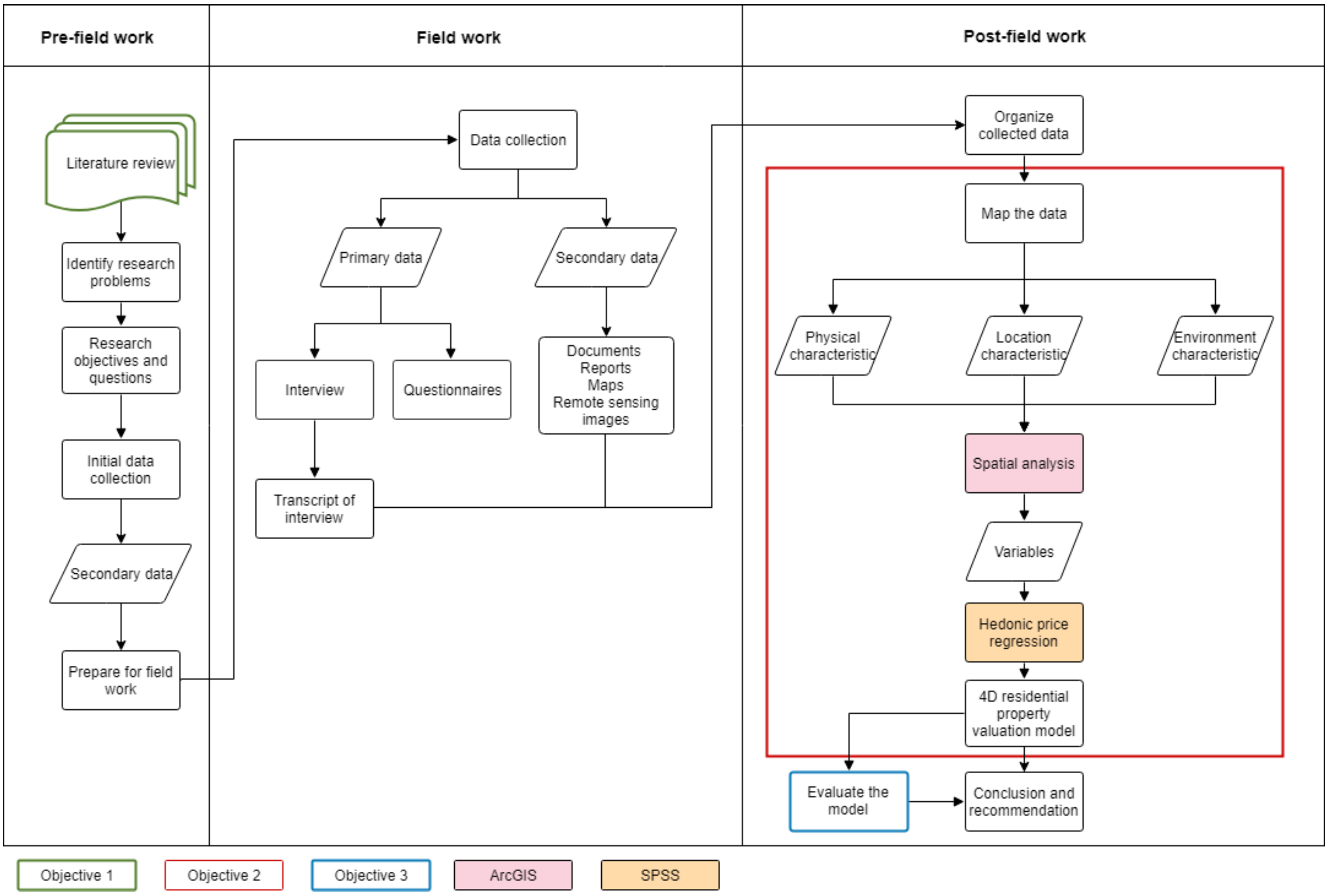


Figure 4: Workflow

3.4. Data and Methods

3.4.1. Data requirement and collection methods

This research requires specialised information regarding urban properties location, their prices and characteristics; temporal urban boundaries; etc.

Therefore, both qualitative and quantitative data are needed. Literature review; documents and records; expert interview; questionnaires are appropriate data collection methods for this study to deploy an integrated and multi-discipline set of information in addition to efficient feedbacks on the process (Y. Liu, Gupta, Springer, & Wagener, 2008).

The literature review is used to review the state-of-the-art property pricing models, how researchers use the model, which characteristics and functions are used to build the model. Besides, according to the interview with property owners to find the limitation of current property valuation models.

Collecting documents and records is the main way to collect secondary data, including geo-data (remote sensing images, maps, shapefiles, etc.), data about property prices, official text and statistical data. However, secondary data are not enough for this research.

Therefore, as a supplement, expert interview and questionnaires are used to collect primary data. Semi-structured expert interviews were used to collect data about the area grows fast in the vertical dimension, as these areas are difficult to detect automatically from remote sensing images. Experts mainly include urban planners. Questionnaires and interviews were used to collect data about local people's preference for properties, which cannot be acquired from other methods. The respondents are mainly citizens who have lived in Xi'an for several years. Table 5 is the table of data requirements.

Table 5 Data requirement

Data requirement	Date	Source
Remote sensing images	2010, 2014,2017	Aster (15m), GaoFen-1(16m)
Building height information for whole buildings in Xi'an	2017	Other researchers (A group of GIS specialists from Harbin Institute of Technology, Wuhan University and Shenzhen University)
Streets and subway lines	2018	Open street map
POI	2018	Baidu map (Website), GaoDe map
Property price	2010,2014,2018	Real estate agency (websites), and samples from websites of "Chinese house price"
Socio-economic factors		Fieldwork (interviews, questionnaire)

3.4.2. Data collection

In collecting the data, this research uses several methods, which are interviews, focus groups, questionnaires and secondary information collecting.

3.4.2.1. Primary data collection

In the literature review section, the relevant indicators influencing the property prices used in related studies have been formulated. In order to get the data related to the local context, these indicators collected from literature review were confirmed to local people and local experts. The primary data were collected through interviews, focus groups and questionnaires to make sure which indicators could be used to build the model, which can suit the local context. Respondents of the interviews include the officials working for planning, surveying and mapping, property valuation, landscape designing and property sale from different companies or institute in Xi'an. The respondents of two focus groups and questionnaires are residential property owners or the one living in Xi'an.

Interviews with officials specializing in planning and surveying: the interviews were conducted with the officials of Xi'an Survey and Mapping Institute, Xi'an Qujiang New District Management Committee and Xi'an Huadi Surveying and Mapping Technology Company who engaged in land planning, surveying and property valuation.

The aim of interview was to get an overview of indicators used for residential property valuation based on their working experience. In that case, people engaged in different positions were asked different questions based on their working tasks (see appendix1). For instance, the questions for planner were at a more macroscopic level, such as which indicators would influence property price and when they make planning if they would consider these indicators, etc. The questions for residential salesmen were at a more microcosmic level, such as what kinds of the residential property was most popular with property purchasers, what the characteristics of these properties are, and how the preferences of the property buyers changed during 2010 and 2018, etc. In addition, both of them were asked about their opinion of using 3D technology in their working field.

Interviews with staff in different real estate companies: the interviews were conducted with 4 property sale managers from Ziwei Real Estate Company, Jinhui Real Estate Company and Shanghai Industrial Urban Development Company (Xi'an) respectively, and two landscape designers from Jinhui Real Estate Company and Gemdale Real Estate Company respectively.

According to the aim of the interviews, the questions for the sale managers are mainly about what kind of property sells best in 2010, 2014 and 2018 respectively and what kind of common attribute they own. Besides, these sale managers were also asked what changes of characteristics of the best-selling property over the eight years are. In addition, the questions for landscape designers are mainly about what kind of landscape is more popular among the property buyers, and the changes in this popular landscape over the ten years.

Questionnaires: The questionnaire (see appendix2) aimed at detecting the local people's attitude to different indicators of property prices, so the respondents of the questionnaire are local people who own properties in Xi'an or living in Xi'an for several years. During the field work, the questionnaires were sent through paper and the internet link and 261 questionnaires have been collected.

The central part of the questionnaire contained questions about: respondent's attitude of different indicators of one residential property. These indicators were collected based on a literature review and local experts' interviews. Respondents were asked to select the importance level from five levels (very important, important, general, not important, not at all important) for each indicator according to their performance. Besides, the main questions, there are some questions about their personal information (such as age, gender, etc.), their prediction of the residential property price in Xi'an in the next one year, and their satisfaction level of their living environment. The answers to these questions were helpful for the analysis in the discussion section.

3.4.2.2. Secondary data collection

Residential property price data collection: collecting residential property price data is crucial in building the property valuation model. The price data of second-hand residential properties are collected because as mentioned in chapter two, the new property price is controlled by the local government too much, so the second-hand property price is more like a market price. The second residential property transaction data in 2018 are collected from a Chinese famous real estate transaction agency, Anjuke[1]. Anjuke is a real estate transaction website platform, providing the sales information and basic data of the property on selling, including the transaction price, the building height, the coordinate value, etc. The historical transaction data are collected from the sample data on the Website of "Chinese house prices"[2]. The website is hosted by China Real Estate Association, based on the real estate big data of Xitai.

Spatial data collection: In order to collect data on location characteristics and environmental characteristics, POI data are collected from Gaode Map (a digital Chinese map) and Open Street data. POI data includes the distribution of hospitals, schools, parks, bus stops, metro stations, residential property, etc. Building boundaries and building height data in 2017 are collected from the researchers in the group of GIS specialists from Harbin Institute of Technology, Wuhan University and Shenzhen. Build-up boundary data are collected from Resource and Environment Data Cloud Platform[3]. Aster 2010, Gaofen 2014 and 2017 are collected for land use analysis. All of these data use WGS 84 coordinates. The table of spatial data collection is shown as Table 6.

[1] https://xa.anjuke.com/

[2] http://www.creprice.cn/database/trans.html

[3] http://www.resdc.cn/

Table 6 Spatial data collection

Relevant characteristics	Type of spatial data	Data description
physical characteristics	Polygon	Polygons of buildings with the number of floors
	Point	Transaction properties
Location characteristics	Point	Banks, bus stations, metro stations, CBD, air quality monitoring stations
	Line	Build-up boundary, roads
Surrounding environment characteristics	Point	Schools, hospitals, parks, entertainments, factories, shopping malls, supermarkets
	Polygon	Water
	Satellite image	Aster 2010, Gaofen 2014 and 2017

Other no-spatial data collection: Some macroeconomic data, such as inflation rate, can be used to the adjust property price in different years (Capozza, Hendershott, Mack, & Mayer, 2002). Population and economic level are also two main indicators cause property price dynamic (Junhua Chen et al., 2011). Air quality value data are collected from PM2.5.in[1], a website providing air quality monitoring data collecting from the Ministry of Environmental Protection.

3.4.3. Methods for data analysis

Data analysis phase includes both qualitative and quantitative methods. Firstly, categorize and transform qualitative data collecting through interviews. In this process, the indicators repeated by the respondents most times are found out, and these indicators are set in questionnaires to detect how local people like them based on statistical analysis.

After collecting questionnaires, commonly, the Cronbach's Alpha is used to analyse the reliability of the result first. The larger the Cronbach's α coefficient value means the stronger internal consistency or the higher reliability (Cortina, 1993; DeVellis, 2003), as shown in Table 7. Then basic statistical measures were presented by the visualization and descriptive by charts and texts and the most concerned indicator of local people are discovered.

Table 7 Cronbach's Alpha and internal consistency

Cronbach's Alpha	Internal consistency
$0.9 \leq \alpha$	Excellent
$0.8 \leq \alpha < 0.9$	Good
$0.7 \leq \alpha < 0.8$	Acceptable
$0.6 \leq \alpha < 0.7$	Questionable
$0.5 \leq \alpha < 0.6$	Poor
$\alpha < 0.5$	Unacceptable

After selecting indicators, spatial analysis is used to quantify the selected indicators and linear regression analysis is used to detect how important of each indicator and the coefficient of the most important indicators to build the valuation model.

[1] http://pm25.in/

3.4.3.1. Spatial analysis

Spatial analysis will be used to calculate the characteristics. For instance, the accessibility to the CBDs map can be obtained by calculating the time needed to take a car or walk from the property to CBDs. The time is equal to multiply the speed of roads or walking by the distance between the property and the CBD.

Visualisation of the dependent variable

The second-hand property transaction price per square meter (Chinese yuan per square meter) of the property is defined as the dependent variable. The average property price of each Xiaoqu[1] is calculated as the dependent variable, because the average price is more valuable to detect how property price distributes across the whole city. On the other hands, some of the transaction data collected from websites may be not true, because some sellers are put a very low price on the websites to attract the buyers' attention, so using the average price of each Xiaoqu can avoid these errors. The location of the residential property is digitized in ArcGIS. The map of the property is shown as Figure 5.

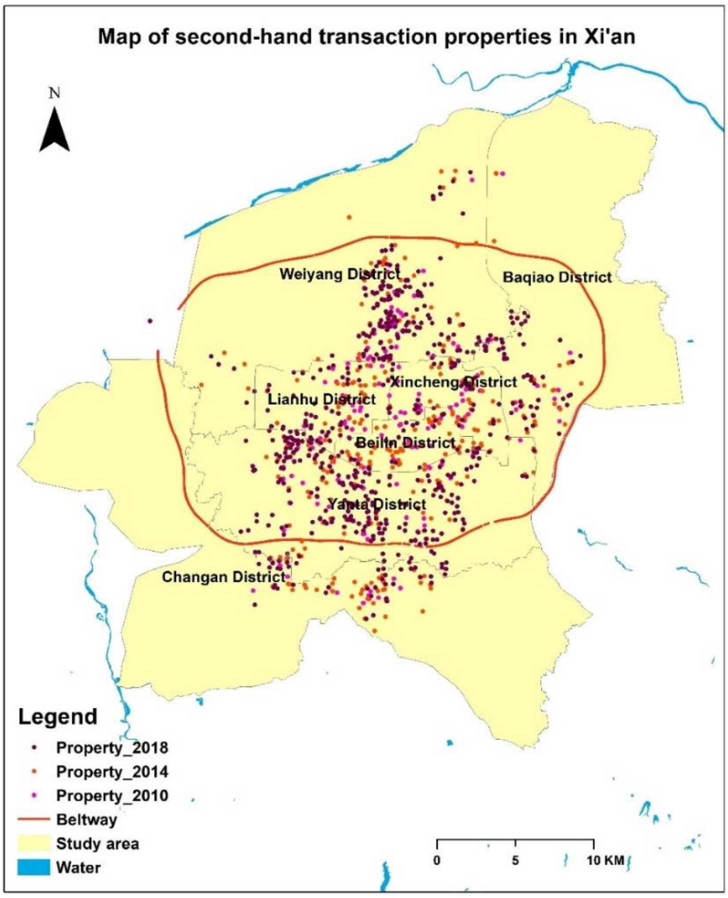


Figure 5 Maps of send-hand transaction properties in Xi'an

Measuring accessibility

According to the indicators selected above, the accessibility of metros, bus stations, hospital, parks, and schools, etc is needed to be calculated. Commonly, accessibility is only considered about the proximity of the

[1] Xiaoqu is a type of neighbourhood, which is also named micro-district. Xiaoqu, similar to the Russia neighbourhood unit, refers to large residential buildings with relatively independent living environment in a certain area of the city and it is equipped with complete sets of living service facilities, such as commercial facilities, hospitals and schools, etc. (Bray, 2005; Wallenwein, 2013).

property to the destinations. (Tribby & Zandbergen, 2012). In this thesis, the accessibility of metro stations, college and shopping mall should be considered as mentioned above. Calculating Euclidean distance is used to measure the proximity of this destination. For examples, the accessibility map of metros is shown as Figure 6. The other Euclidean distance maps are used the same way to be made. There are 3 metro lines has been constructed until 2018. The metro line No.2 penetrates the city north and south, while the other two lines penetrate the city west and east.

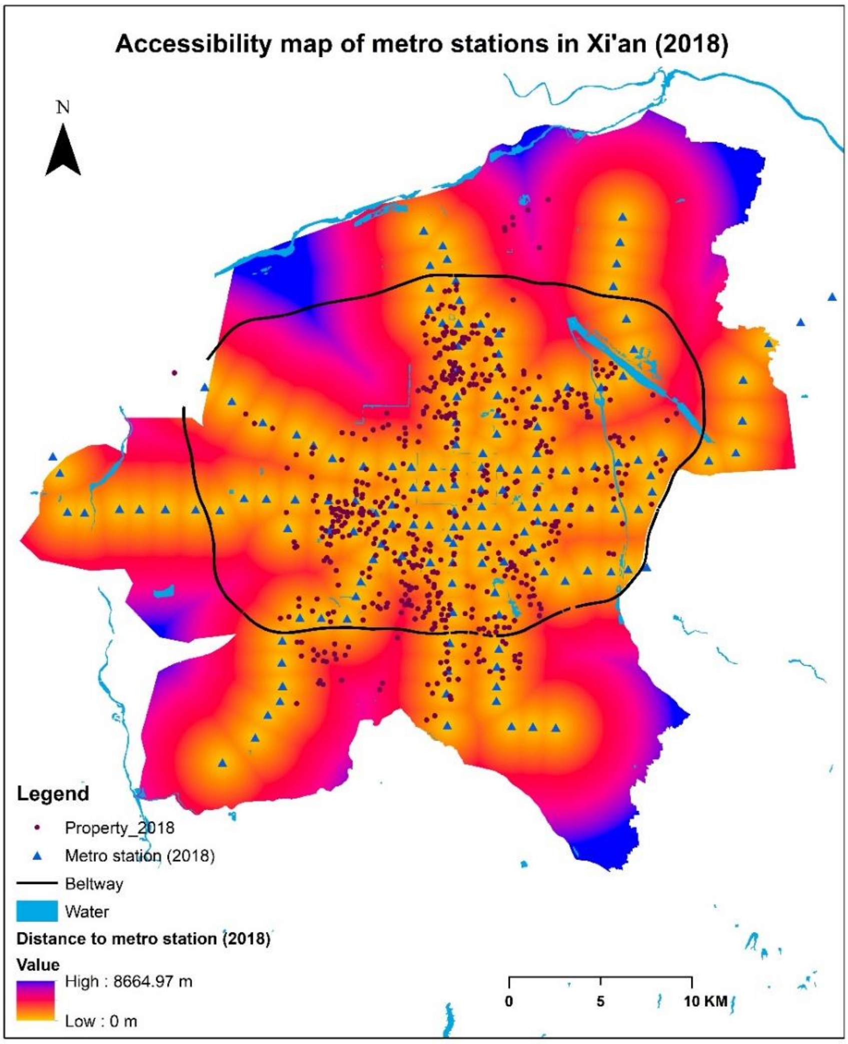


Figure 6 Accessibility map of metro stations in 2018

In addition, calculating the number of destinations (density) is another way to measure accessibility, especially when the range of distance of the property to each destination is similar (Cheng & Agrawal, 2010). In Xi'an city, the distance of each property to the nearest bus station is always less than 1.5 kilometers, while the density of bus station is unevenly distributed entire Xi'an city. In that case, the accessibility to the bus station is calculated by the density of bus stations near the property in this thesis. The density map of bus stations is shown as Figure 7. The figure11 illustrates that the density of bus stations is much higher in the south of the city and lower in the west of the city.

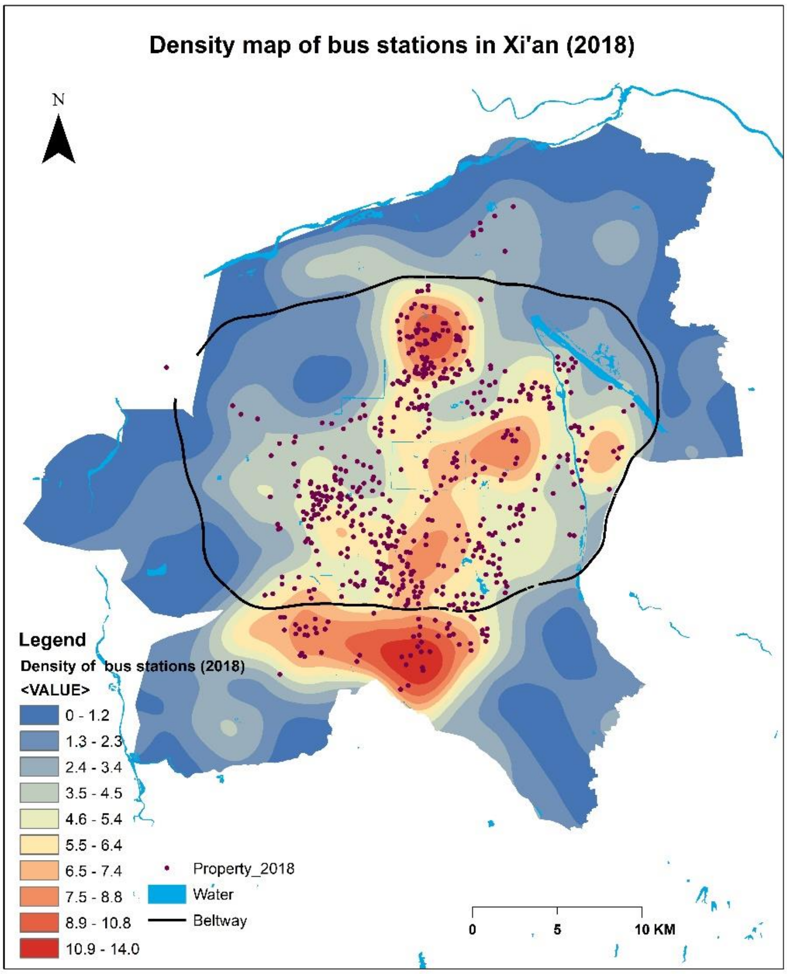


Figure 7 Density map of bus stations in Xi'an in 2018

Air quality index (AQI)

According to the interview and questionnaires, local people in Xi'an pay much attention to the surrounding environment of the property when they buy a residential property. Air pollution as a problem of Xi'an also attracts attention of local people. There are 11 air quality monitoring stations in the study area. These stations monitor the value of PM2.5, PM10, SO2, NO2, etc. of each station per hour, and provide the value of air quality index (AQI). AQI is an index value to report the daily air quality, and the higher value of AQI means more air pollution(AirNow, 2016). There are six levels of the value of AQI (AirNow, 2016), as shown in Table 8.

Table 8 AQI levels of heal concern

AQI levels of heal concern	AQI value range
Good	0-50
Moderate	51-100
Unhealthy for sensitive groups	101-150
Unhealthy	151-200
Very Unhealthy	201-300
Hazardous	301-500

In this thesis, the average AQI value of each station has been calculated and the interpolation map of AQI in the study area is shown as Figure 8. The Inverse distance weighting (IDW) method has been used to calculate the interpolation map of AQI value, because this method is commonly used in the environmentl science to calculate the distribution of air pollute from fixed monitoring stations (L. Wu et al., 2017). From the interpolation map, it can be found that the air quality in the south of the city is better than the air quality in the north area. And the area in the east of the city has better air quality than the west area of the city.

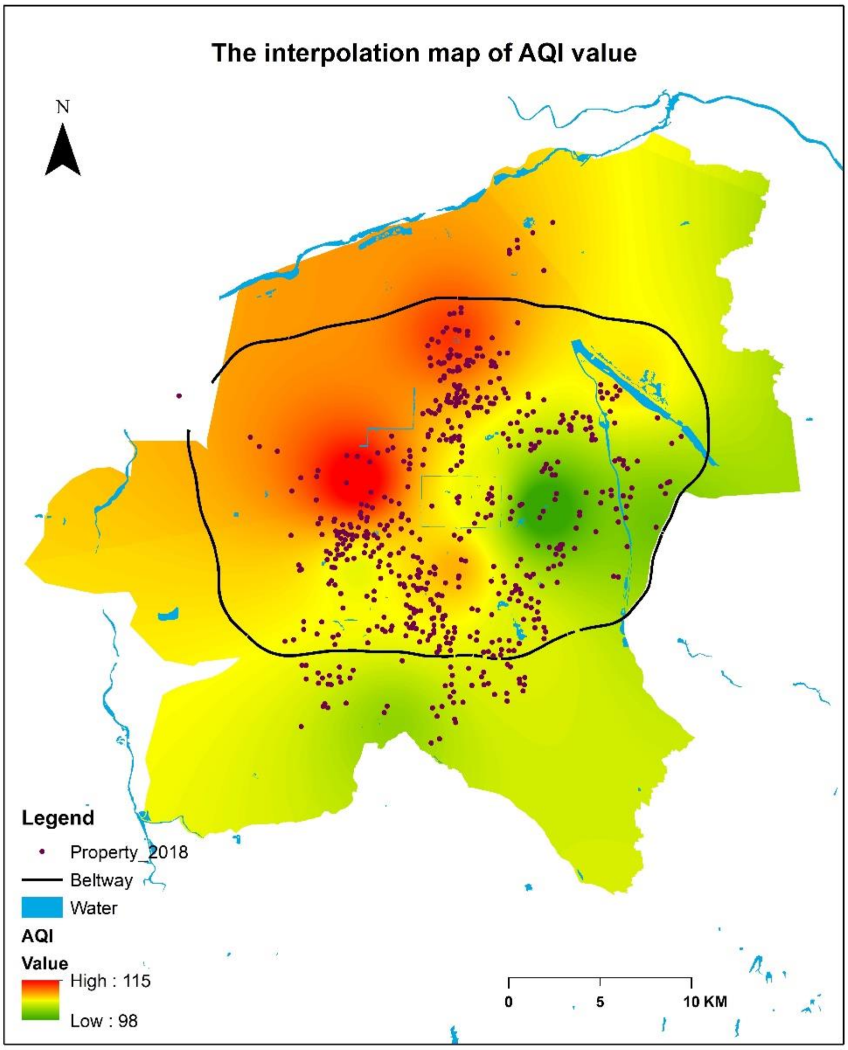


Figure 8 The interpolation map of AQI value

NDVI

Normalized difference vegetation index (NDVI) is used to detecting vegetation coverage through remote sensing images and "it is calculated from the visible and near-infrared light reflected by vegetation". (Weier & Herring, 2000). The formula of calculating NDVI value is:

$$\text{NDVI} = \frac{NIR - VIS}{NIR + VIS}$$

The result is given to each pixel. The range of the result of NDVI is from -1 to 1. the more the result close to 1, the more photosynthetic active vegetation cover (Weier & Hserring, 2000). The map of NDVI value in Xi'an in 2017 is shown as Figure 9.

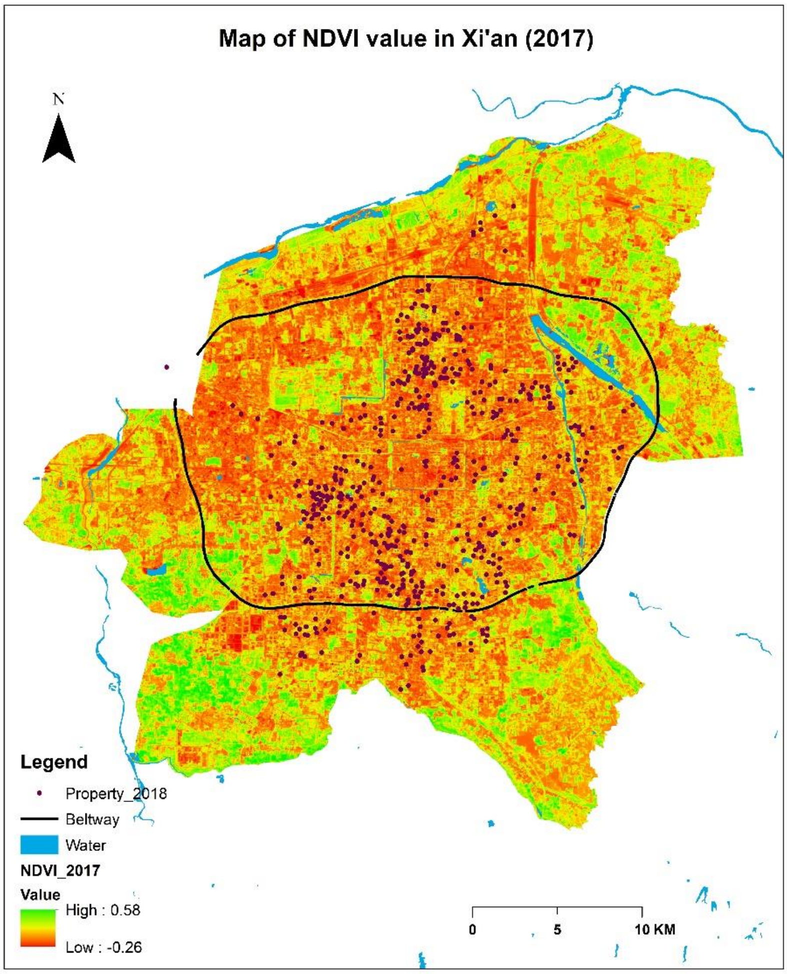


Figure 9 Map of NDVI value in Xi'an in 2017

Building height

The second-hand property transaction data in 2010, 2014 and 2018 collected from the website, and the geometric data of building polygons in 2017 collected from the group of GIS specialists from Harbin Institute of Technology include the number of the floors of the building where the property locates. As generally the height of each floor is 3 meters, in that case, the building height of a building (meter) is equal to the number of floors of the building multiplied by 3.

3.4.3.2. Multiple linear regression model

Regression is used to identify the variables of property valuation. The variables are identified from geo-data, which includes both spatial urban patterns and spatial property prices. Meanwhile, it should be sure that the independent variables are linearly related to the dependent variable and estimate the coefficients of variables through multiple regression analysis in SPSS (Seo et al., 2014). Then variables are grouped and filtered by the correlation matrix, and only relevant variables are used for model building (Hill, 2013).

In this thesis, according to the theory of hedonic model to calculate the relationship between indicators and property value, multiple linear regression model is suitable. In the multivariate linear regression model, the function form is used to establish the regression relationship between the property price (as the independent variable) and the price indicators (as the dependent variables). The sample data are used to statistically infer the parameters of the equation. If the regression coefficient passes the significance test, it indicates that the indicators have an impact on housing price and the goodness of fit of the regression model illustrates the fitting effect of the model. The multivariate linear regression model is easy to interpret because of its simple operation, and it is convenient for epitaxial inference to be widely used in real estate evaluation. In this research, the indicators are linear relevant to the dependent variables. The specific function form of the model is as follows:

$$\mathrm{P} = b_0 + \sum a_i k_i + \varepsilon$$

P is the price of property; b_0 is the constant term; a_i is the regression coefficient; k_i is the independent indicators, and ε is the residual terms of the model (Bremer, M.).

3.4.3.3. K-fold cross validation

When building multiple models of the same practical problem, it is often necessary to evaluate the quality of the model. Cross Validation, sometimes called Rotation Estimation, is a way to assess how the results of statistical analysis will generalize to an independent data set (Geisser, 2017; Kohavi, 1995). In other words, this method can be used to estimate the skill of the model on new data (Brownlee, 2018). This technique uses all available examples as training and test examples (Stone, 1974).

K-fold cross validation is one of the common cross validation methods (Anguita, Ghelardoni, Ghio, Oneto, & Ridella, 2012). The initial sample is separated into K subsamples. A single subsample is retained as test data, and the other K-1 subsamples are used for training the regression model. This validation process then repeats K times, and each of the K subsamples is used once as test data to validate the model. And the K-times results are averaged as a final estimate result (Brownlee, 2018; Ca & Fr, 2004). The advantage of K-fold cross validation is that all samples are used for both training and validation, and each test subsample is only used once (Barrow & Crone, 2013). Although K is an unfixed parameter, 10-fold cross validation is a common method (McLachlan, Do, & Ambroise, 2004).

3.5. Summary

In this chapter, the study area, research methods of data collection and analysis are distributed, which is the basis of building a valuation model. Selecting the appropriate indicators for later data analysis is important, so based on literature review, interviews and questionnaires are used to collect the local situation of property price. The main parts of data analysis in this research is spatial analysis to qualify the indicators and regression analysis. The result of regression analysis is described in the next chapter.

4. Result

4.1. Result of questionnaires and interviews

4.1.1. Result of questionnaires

After deleting the questionnaire with an answer time of fewer than 3 minutes or the respondent's IP address not in Xi'an, the number of valid questionnaires is 221. The result has passed the reliability analysis and validity analysis. Cronbach's Alpha is 0.842, which means the reliability is high (Bolarinwa, 2015). The result of the questionnaires is shown as Figure 10.

Through the results of the questionnaire, the most important indicator is "nearby factories without pollution", and over 66% of respondents agree it is very important. The second important indicator is the distance to the metro station. 33.4% and 45.75% of respondents think it is very important and important respectively. However, the distance to the railway station and the distance to the cultural heritages are the least important indicators. Only 7.08% and 3.07% people consider the two indicators in the very important level respectively. In that case, the two indicators are not used in building the property valuation model.

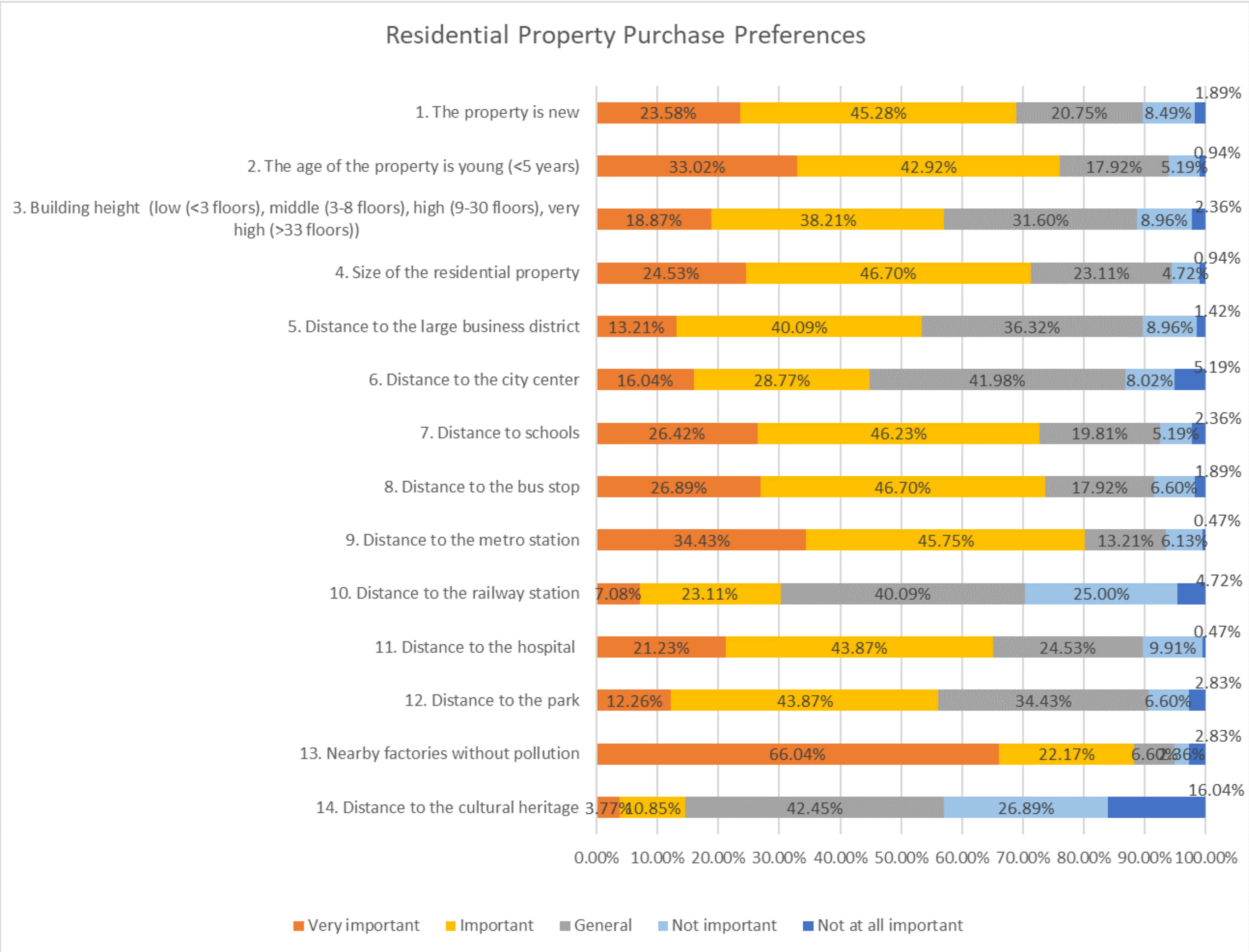


Figure 10 The result of questionnaires

4.1.2. Result of interviews

The aim of the interview was to collect the indicators of residential property valuation based on their working experience.

Through the interview with employees of different real estate companies, we found that the pricing strategies of different real estate companies are similar. The most popular method of setting the average price of a neighbourhood is the market approach. In other words, the company will set its average price according to other surrounding properties' prices. Mostly, education facilities, transportation accessibility (to bus stops and/or metro stations), the hospital, shopping and environment (less pollution and better view) are the most important indicators of property prices. In that case, the neighbourhood with wonderful infrastructure facilities and located nearby the beautiful scenery with little noise, such as the lake, the park would be more attractive. Therefore, its price would be also higher. Besides, the property in the neighbourhood with a low plot ratio always more expensive than the others. To set the price of each unit, firstly they will select the best one with best location (almost in the center of the neighbourhood and in the middle floor level) because of the balance of accessibility, noise and the view. Besides, the apartment with the orientation to both north and south is always more expensive than others, because of the better daylighting. Then other properties' prices would be lower than the best one.

When talking about how urbanization influences the residential property price in Xi'an, most of the respondents agree that urbanization stimulates the property price increase. In the past, the administrative district with the highest average residential property price was Gaoxin New District (in the west of Yanta District), but now it is Qujiang New District (in the east of Yanta District). The price of property located outside the Third Ring Road has been more than twice the price in the past. However, the price of property in the south of the Xi'an city is always higher than the price in other parts, because of more job opportunities, colleges and better views.

Since 2017, the local Price Bureau has started to control the price of residential properties to stabilize the real estate market. In that case, before the property is sold, the company need to get the permission form Price Bureau and the local Department of Housing Management, which means the property's design meets the requirements of the urban planning and the price of the property is not higher than the price given by the Price Bureau, which is calculated based on the market approach and inflation rate. The results of the interviews are shown as word clouds in Figure 11.

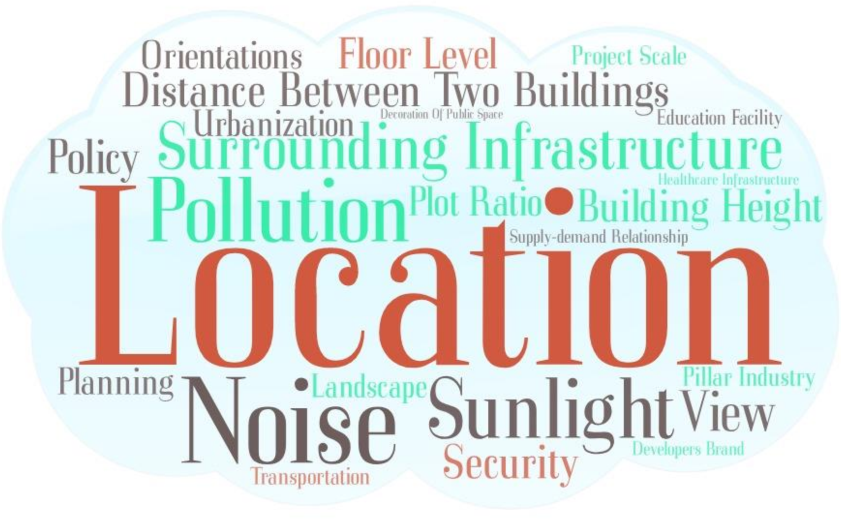


Figure 11 Words cloud of interview-working in real estate companies

In the interviews with officials specializing in planning and surveying, the officials mentioned that comprehensive city planning and balance between urbanization and residential price most times. In that case, the most important planning is the master city planning which will set aims for each district and the aim for the whole city. Then the detailed planning should comply with the requirements of the master plan. And in detail planning, the planner needs to make sure that each neighbourhood should have enough surrounding infrastructure facilities, such as primary school, high school, parks and sports facilities. However, they cannot ensure these facilities have the same quality, in that case, the property near the school with high quality would always be more expensive than others. The results of the interviews are shown as word clouds in Figure 12.

Figure 12 Words cloud of interview-planner and surveyor

4.2. Indicators selection and collinearity test

Because of the dependent variable is the average price of each Xiaoqu (a type of neighbourhood in China), the indicators of physical of each property are not used in this research, except the building height. According to the result of questionnaires, interviews and literature review, the indicators selected to be used are shown as Table 9.

Table 9 Property valuation indicators

Physical characteristics		Building height
Location characteristics	Transportation	Distance to metro
		Bus stations' density
		Distance to high way
	Location	Distance to CBDs
		Distance to the build-up area boundary
Surrounding environment characteristics	Education facilities	Kid gardens' density
		Primary schools' density
		High schools' density
		Distance to college
	Shopping facilities	Distance to shopping mall
		Supermarkets' density
	Health facilities	Distance to hospital
	Environment and landscape	Distance to park
		Factories' density
		Distance to water
		Air quality index (AQI)
		NDVI

4.3. Result of 2018 property value regression model

This process is intended to test whether these indicators determine the residential property price in Xi'an significantly. In that case, multiple linear regression model is a potentially useful tool to achieve the aim.

4.3.1. Descriptive statistic

A descriptive statistic is a brief summarization of the data set, which usually includes the mean, the median and mode, etc. (Investopedia, 2018).

The completeness of the data set was confirmed by the descriptive statistics shown in Figure 13and Table 10. The results indicate that the property prices are high variance. As shown in Figure 13, the prices are in accordance with the normal distribution.

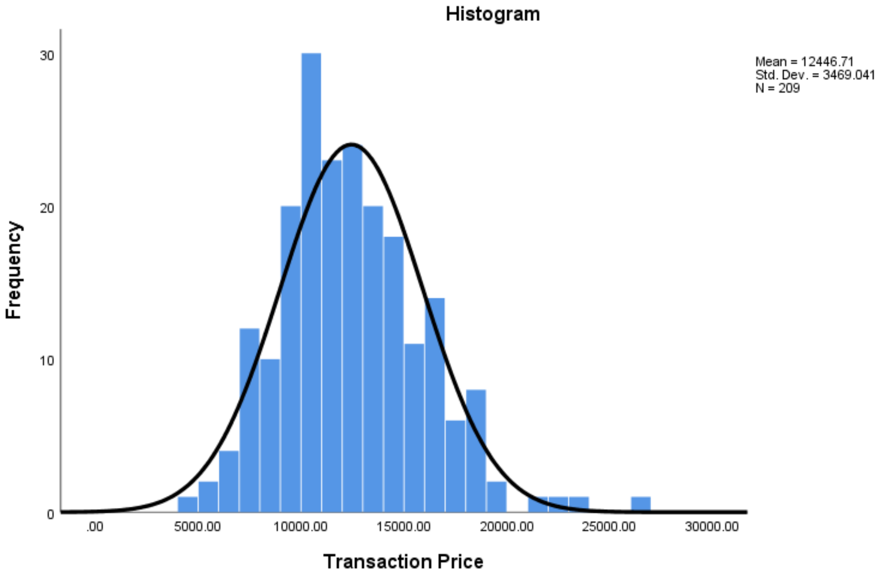


Figure 13 Histogram of dependent variable distribution

Table 10 shows the result of checking mean and standard deviations of each variables. The mean of a binary indicator indicates the percentage of properties possessing that characteristic.

Table 10 Descriptive statistics for the variables

	Mean	Std. Deviation	N
Price	12446.71	3469.04	209
Building height	62.97	29.18	209
Density of Key primary school	0.08	0.11	209
Density of Key high school	0.08	0.07	209
Density of park	0.61	0.41	209
NDVI	0.02	0.07	209
AQI	106.87	2.43	209
Density of Factory	0.76	0.54	209
Density of Hospital	0.67	0.66	209
Density of Bus station	6.04	1.73	209
Distance to metro station	898.20	659.60	209
Distance to high way	1076.40	844.09	209
Distance to CBD	2656.01	1911.43	209
Distance to college	1391.53	854.27	209
Density of kid garden	3.92	1.47	209
Density of high school	0.50	0.38	209
Distance to Eucdist shop mall	517.66	376.13	209
Density of Entertainment	13.64	9.86	209

4.3.2. Module summary

Step wise regression method was used to do the multiple linear regression analysis and select the signature indicators to build the price model, shown as Table 11.

Table 11 Table of model summary

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.367[a]	.134	.130	3235.21847	.134	32.153	1	207	.000	
2	.462[b]	.213	.205	3092.14454	.079	20.599	1	206	.000	
3	.521[c]	.272	.261	2982.46550	.058	16.430	1	205	.000	
4	.579[d]	.336	.323	2855.28290	.064	19.669	1	204	.000	
5	.592[e]	.351	.335	2828.85605	.015	4.829	1	203	.029	
6	.603[f]	.364	.345	2807.76887	.013	4.061	1	202	.045	
7	.622[g]	.387	.366	2762.05267	.024	7.742	1	201	.006	
8	.637[h]	.406	.382	2726.85836	.018	6.222	1	200	.013	
9	.647[i]	.419	.392	2704.03523	.013	4.390	1	199	.037	
10	.656[j]	.431	.402	2682.76156	.012	4.169	1	198	.043	
11	.668[k]	.446	.415	2652.61514	.016	5.526	1	197	.020	
12	.664[l]	.441	.412	2659.34773	-.006	2.006	1	197	.158	2.115

Predictors: (Constant), building height
Predictors: (Constant), building height, NDVI
Predictors: (Constant), building height, NDVI, density of park
Predictors: (Constant), building height, NDVI, density of park, density of hospital
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI, density of entertainment
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI, density of entertainment, distance to shopping mall
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI, density of entertainment, distance to shopping mall, distance to college
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI, density of entertainment, distance to shopping mall, distance to college, factory
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI, density of entertainment, distance to shopping mall, distance to college, factory, Distance to metro station
Predictors: (Constant), building height, NDVI, density of park, density of hospital, AQI, density of entertainment, distance to shopping mall, distance to college factory, distance to metro station, density of key high school
Predictors: (Constant), building height, NDVI, density of park, density of hospital, density of entertainment, distance to shopping mall, distance to college, factory, distance to metro station, density of key high school
Dependent Variable: average_price

In the best result, R2 is 0.446 and adjusted R2 is 0.415, which can illustrate that the indicators account for 41.5 percent of the property value variations. According to the regression model summary, these indicators, building height, NDVI, density of park, density of hospital, density of entertainment, distance to to shop mall, distance to college, density of factory, distance to metro station, AQI, and density of key high school, are significant to the property transaction prices. Use these indicators as independent variables and second-hand property transaction prices as the dependent variable to do linear multiple regression analysis, and the coefficients of the model are shown as Table 12.

Table 12 Regression model coefficients

	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	27097.425	8571.873		3.161	.002		
building height	39.224	6.545	.330	5.993	.000	.928	1.078
density of key high school	-9548.446	4061.870	-.185	-2.351	.020	.452	2.211
density of hospital	-1236.913	359.644	-.235	-3.439	.001	.600	1.666
density of park	3203.286	589.020	.376	5.438	.000	.589	1.699
NDVI	8899.388	2842.775	.188	3.131	.002	.778	1.286
AQI	-115.033	81.211	-.080	-1.416	.158	.871	1.149
density of factory	-1027.127	377.768	-.161	-2.719	.007	.802	1.247
distance to shop mall	-2.094	.586	-.227	-3.571	.000	.695	1.438
distance to metro station	-.910	.326	-.173	-2.789	.006	.730	1.370
distance to college	-.832	.264	-.205	-3.151	.002	.665	1.505
density of entertainment	-116.107	26.350	-.330	-4.406	.000	.501	1.996

The most important indicator is the density of park, whose coefficient is positive as expected, because local people like to buy the residential property with wonderful environment and less pollution. The second important indicator is building height, whose coefficient is positive, which means that local people prefer properties in higher buildings. It may because most new buildings are tall-buildings in Xi'an and the new property is always more expensive than the old property. The third important indicator is the density of entertainment places, whose coefficient is negative. The next two important indicators are density of hospital and distance to shop mall and both of them have negative influence on property prices.

The final model can be written as:

Price= 27097.425+(building height *39.224) -(density of key high school *9548.446) – (density of hospital * 1236.913) +(density of park *3203.286) +(NDVI*8899.388)- (AQI*115.033) - (density of factory *1027.127) – (distance to shop mall*2.094)- (distance to metro station*0.91) – (distance to college*0.832) – (density of entertainment*116.107)

4.3.3. Evaluating model

The four assumptions of linear regression are: no multicollinearity, homoscedasticity, independent errors and normally-distributed errors (Field, Miles, & Field, 2012).

No multicollinearity

All the indicators used in the final model are significant and have passed the t-test as shown in figure 22. Besides, the tolerance value of each indicator is more than 0.2 (Ziegel & Menard, 1996) and the VIF value of each indicator is less than 10 means that the model is no multicollinearity (Hair, Sarstedt, Ringle, & Mena, 2012).

Homoscedasticity/ Independent Errors

The size of the error term should be same across values of an independent variable (Field et al., 2012) and the residuals are uncorrelated with the dependent variable (Ai, 2005). Otherwise the situation is known as heteroscedasticity. As shown in Figure 14 the data obey the assumption.

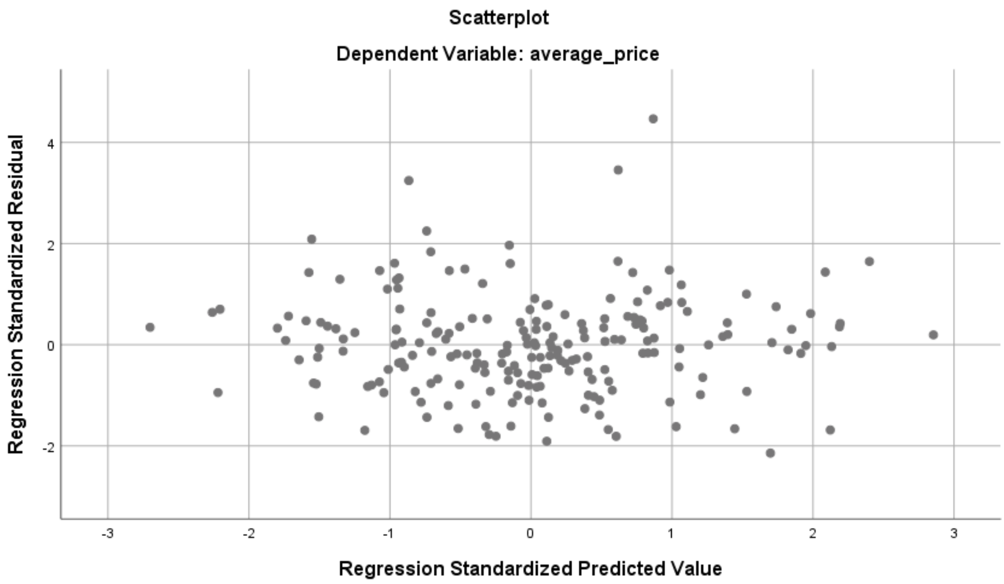


Figure 14 Regression standardized predicted value

Normally-distributed Errors

It checks whether the regression residuals are normally distributed with a mean of zero (Ziegel & Menard, 1996). As shown in Figure 15 the regression standardized residual is a normal distribution.

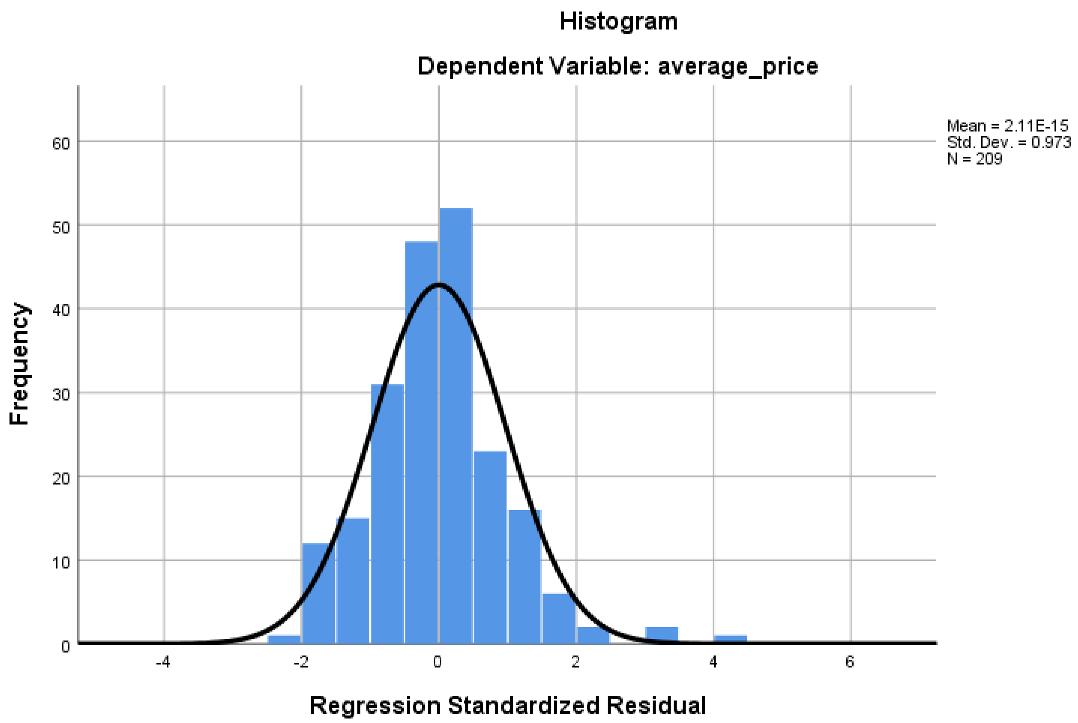


Figure 15 Regression standardized residual

4.3.4. Generalize the model to the study area.

Using the model builder in ArcGIS to generalize the model to the entire study area. The result is shown as Figure 16. The price of the property variation from 504.43 to 27262.10 and the mean value is 9728.14 (Chinese RMB yuan per sq. meter.), shown with different colours in the map. The shapes with red colour mean higher price than the shapes with blue colour. The cheapest area is in the north-east of the city and the most expensive area is mainly distributed in the south-east of the city and part of them are distributed in east of the city between Chan River and Ba River. The red cross symbols represent hospitals that is an important indicator. As for another important indicator, density of entertainment places is shown through grayscale in the map. The darker of the pixel, the higher the density is. In addition, the green point in the map represents parks.

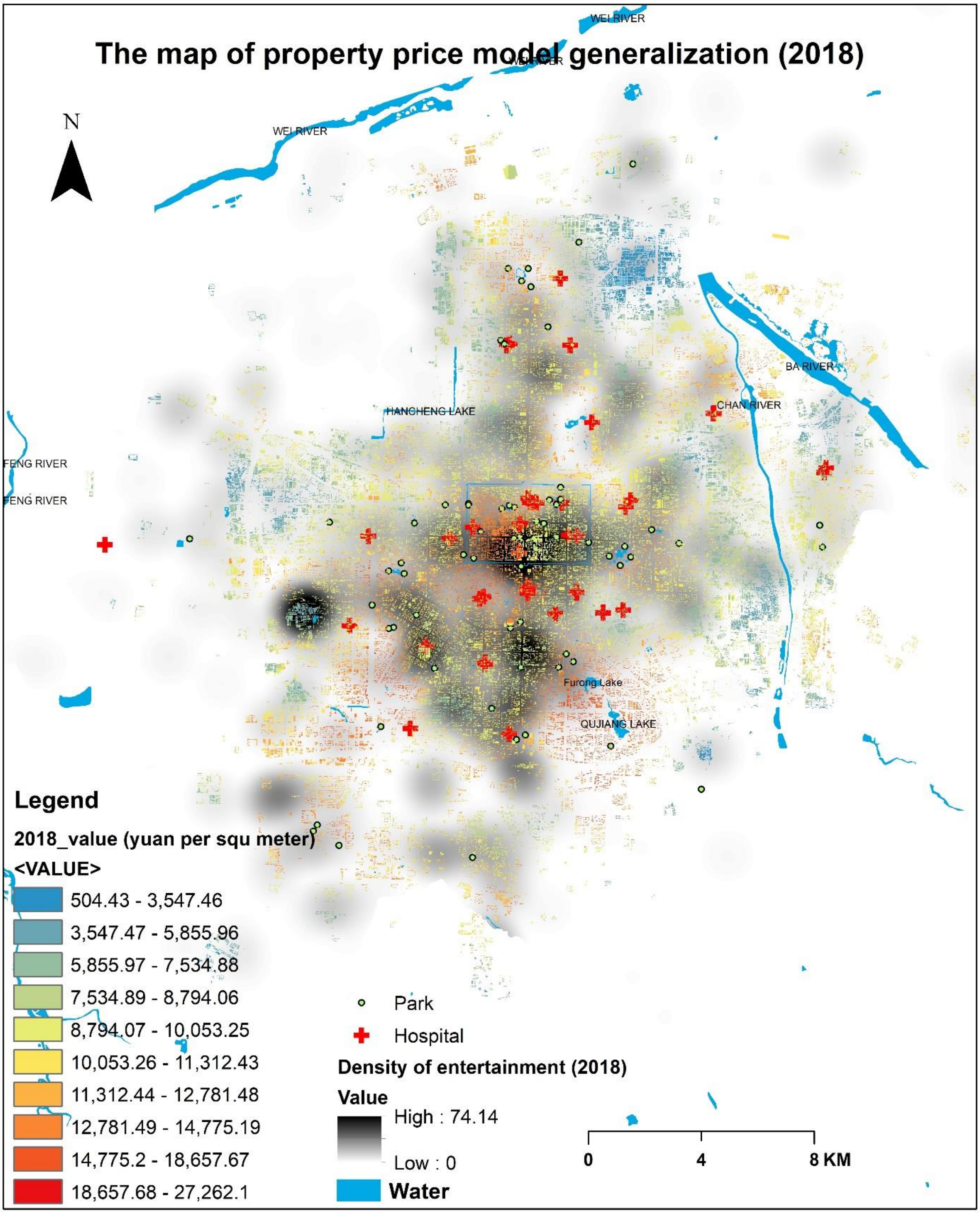


Figure 16 The map of property price model generalization (2018)

4.4. Result of 2014 property value regression model

4.4.1. Descriptive statistic

The completeness of the data set was confirmed by the descriptive statistics shown in Figure 17 and Table 13. The results indicate that the property prices are high variance. As shown in Figure 17, the prices are in accordance with the normal distribution.

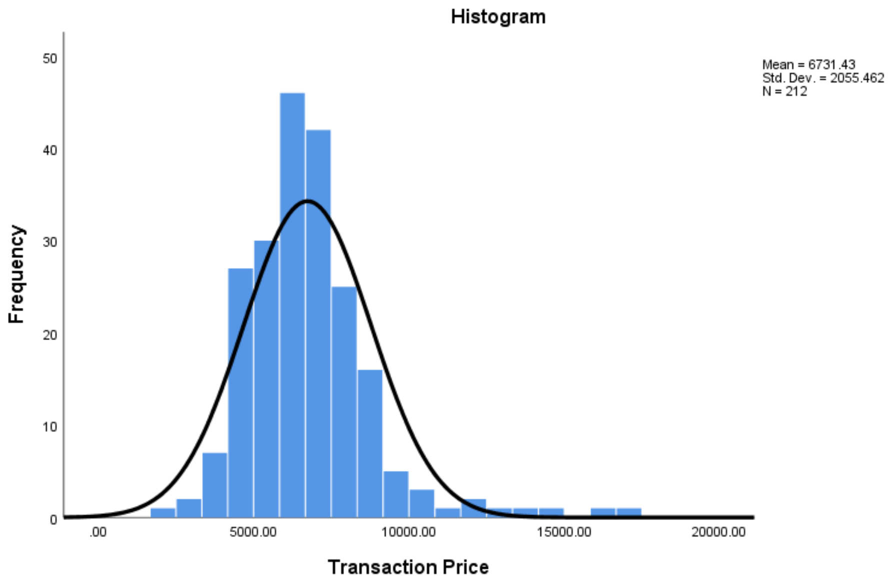


Figure 17 Histogram of dependent variable distribution

Table 13 shows the result of checking mean and stand deviations of each variable. The mean of a binary indicator indicates the percentage of properties possessing that characteristic.

Table 13 Descriptive statistics for the variables

	Mean	Std. Deviation	N
Price	6731.43	2055.46	212
Building_Height	62.89	29.14	212
Density of bus station	6.09	2.20	212
AQI	105.85	2.97	212
Distance to hospital	2155.96	1605.12	212
Distance to metro station	2530.04	1926.95	212
Density of park	0.47	0.48	212
Distance to college	1164.40	769.14	212
Distance to high way	1226.21	963.36	212
Distance to shopping mall	689.41	687.51	212
Density of high school	0.69	0.34	212
Density of primary school	0.69	0.34	212
Density of kid garden	1.56	0.79	212
Distance to build-up boundary	5348.69	2312.71	212
Density of factory	1.09	1.18	212
Distance to park	1094.76	630.58	212
NDVI	-0.02	0.07	212

4.4.2. Module summary

Stepwise regression method was used to do the multiple linear regression analysis and select the signature indicators to build the price model, shown as Table 14.

Table 14 Table of model summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
					Change Statistics					
1	.363a	.132	.128	1919.47883	.132	31.955	1	210	.000	
2	.449b	.202	.194	1844.99006	.070	18.299	1	209	.000	
3	.508c	.258	.247	1783.68859	.056	15.613	1	208	.000	
4	.528d	.278	.264	1762.94897	.021	5.923	1	207	.016	
5	.544e	.296	.279	1744.95521	.018	5.291	1	206	.022	
6	.561f	.314	.294	1726.93344	.018	5.322	1	205	.022	1.848

a. Predictors: (Constant), Distance to hospital
b. Predictors: (Constant), Distance to hospital, Building_Height
c. Predictors: (Constant), Distance to hospital, Building_Height, Distance to metro station
d. Predictors: (Constant), Distance to hospital, Building_Height, Distance to metro station, density of park
e. Predictors: (Constant), Distance to hospital, Building_Height, Distance to metro station, density of park, density of bus station
f. Predictors: (Constant), Distance to hospital, Building_Height, Distance to metro station, density of park, density of bus station, AQI
g. Dependent Variable: Price

The best result, R^2 is 0.314 and adjusted R^2 is 0.294, which can illustrate that the indicators account for 29.4 percent of the property value variations. The value of R^2 is less than the other two years is because there were fluctuations of property transaction prices in 2014 and the change of the transaction price is not due to the indicators used in the research, which is explained in the chapter of discussion. According to the regression model summary, these indicators, distance to hospital, building height, distance to metro, density of park, density of bus station and AQI, are significant to the property transaction prices. Use these indicators as independent variables and second-hand property transaction prices as the dependent variable to do linear multiple regression analysis, and the coefficients of the model are shown as Table 15.

Table 15 Regression model coefficients

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
(Constant)	-5455.728	4400.095		-1.240	.216		
Building_Height	19.226	4.227	.273	4.549	.000	.931	1.074
density of bus station	151.082	56.449	.162	2.676	.008	.918	1.089
AQI	93.635	40.588	.135	2.307	.022	.970	1.031
distance to hospital	-.482	.093	-.377	-5.205	.000	.639	1.565
distance to metro station	.346	.071	.324	4.840	.000	.746	1.340
density of park	659.056	279.391	.153	2.359	.019	.794	1.259

The most important indicator is distance to hospital, whose coefficient is negative. The second important indicator is distance to metro station, whose coefficient is negative as expected. It may because, in 2014 there are only two metro lines in Xi'an, so there are not many metro stations in Xi'an. In that case, local people pay much attention on the proximity to metro station when buying a property. The third important indicator is building height, whose coefficient is positive. The next two important indicators are the density of bus station and density of park and both of them have positive influence on property prices.

The final model can be written as:

$Price_{2014}$= 5455.728+(building height *19.226) +(density of bus station*151.082) + (AQI * 93.635) - (distance to hospital *0.482) +(distance to metro station * 0.346)- (density of park *659.056)

4.4.3. Evaluating model

No multicollinearity

All the indicators used in the final model are significant and have passed the t-test as shown in figure 29. Besides, the tolerance value of each indicator is more than 0.2 (Ziegel & Menard, 1996) and the VIF value of each indicator is less than 10 means that the model is no multicollinearity (Hair et al., 2012).

Homoscedasticity/ Independent Errors

The size of the error term should be same across values of an independent variable (Field et al., 2012) and the residuals are uncorrelated with the dependent variable (Ai, 2005). Otherwise the situation is known as heteroscedasticity. As shown in Figure 18 the data obey the assumption.

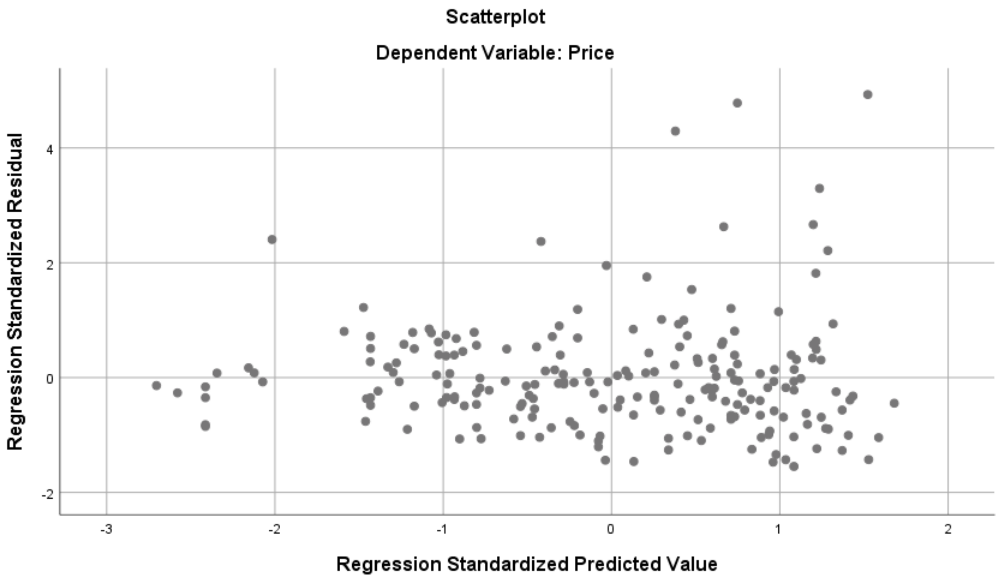


Figure 18 Regression standardized predicted value

Normally-distributed Errors

It checks whether the regression residuals are normally distributed with a mean of zero (Ziegel & Menard, 1996). As shown in Figure 19 the regression standardized residual is a normal distribution.

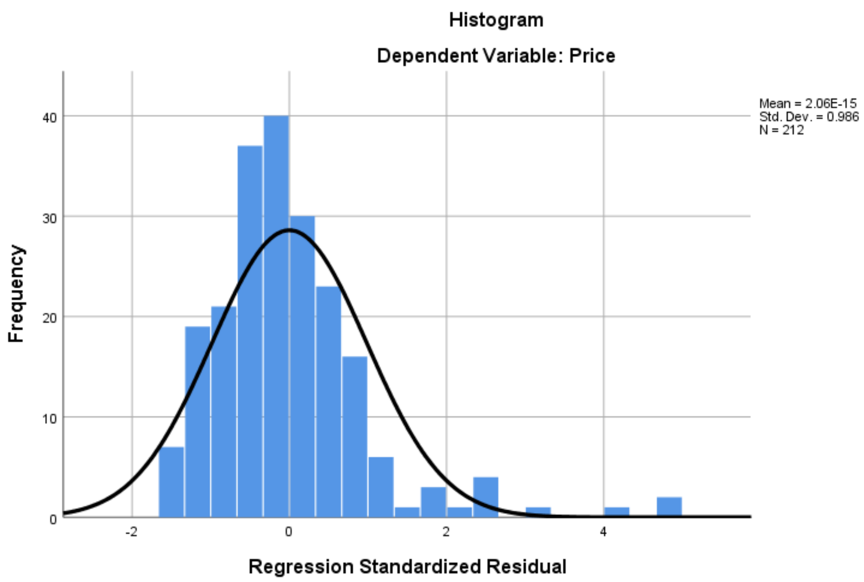


Figure 19 Regression standardized residual

4.4.4. Generalize the model to the study area.

Use the model builder in ArcGIS to generalize the model to the entire study area and adjust the value based on the inflation value. The result is shown as Figure 20.

The price of the property variation from 822.56 to 11735.20 and the mean value is 5914.62 (Chinese RMB yuan per sq. meter), shown with different colours in the map. The shapes with red colour mean higher price than the shapes with blue colour. The cheapest areas are distributed in the east and west part of the city. The most expansive part is distributed in the south west of the city. The red cross symbols represent hospitals that is an important indicator. As for another important indicator, density of bus stations, is shown through grayscale in the map. The darker of the pixel, the higher density is. In addition, the green point in the map represents park and the dark triangle represents metro stations in the map.

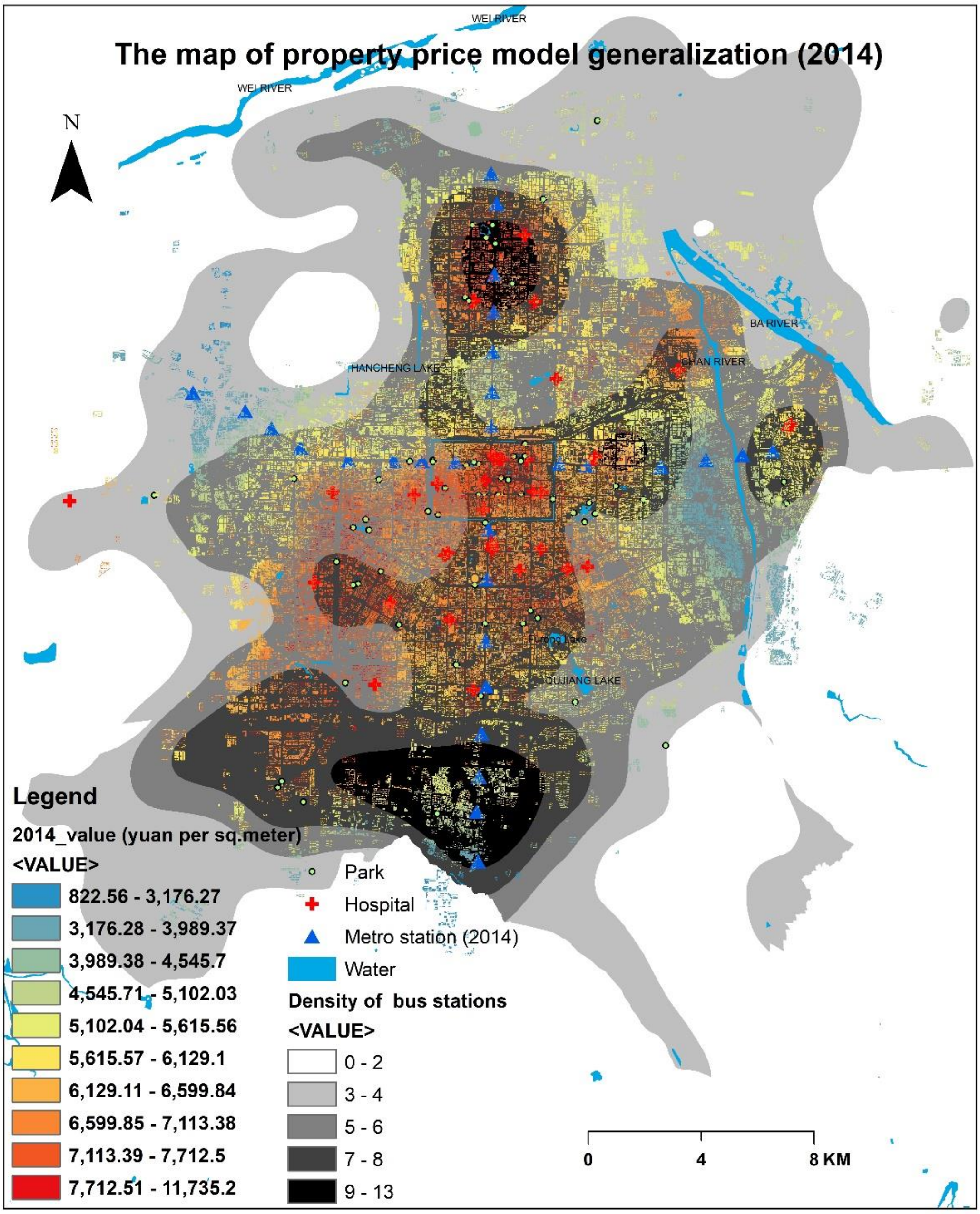


Figure 20 The map of property price model generalization (2014)

4.5. Result of 2010 property value regression model

4.5.1. Descriptive statistic

The completeness of the data set was confirmed by the descriptive statistics shown in Figure 21 Histogram of dependent variable distribution and Table 16. The results indicate that the property prices are high variance. As shown in Figure 21, the prices are in accordance with the normal distribution.

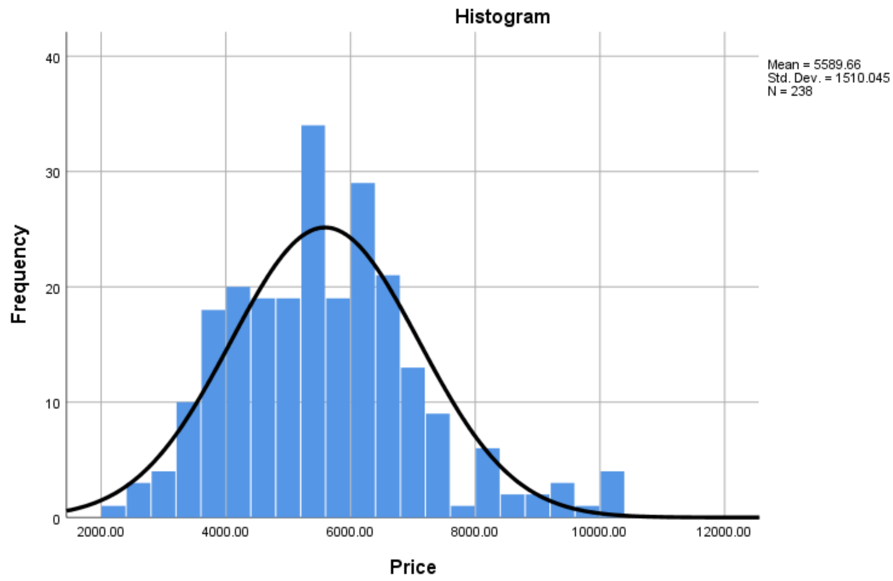


Figure 21 Histogram of dependent variable distribution

Table 16 shows the result of checking the mean and standard deviations of each variable. The mean of a binary indicator indicates the percentage of properties possessing that characteristic. In 2010, no metro line was open, so the indicators do not include distance to metro stations.

Table 16 Descriptive statistics for the variables

	Mean	Std. Deviation	N
Price	5589.66	1510.04	238
Building_height	49.65	28.66	238
Density of primary school	2.26	1.83	238
Density of factory	0.96	0.78	238
Distance to park	1699.89	1488.21	238
distance to CBD	5391.85	3521.07	238
Distance to hospital	1858.91	1614.17	238
Distance to shopping mall	632.71	504.41	238
AQI	106.37	2.63	238
NDVI	0.48	0.02	238
Density of high school	1.92	1.91	238
Density of kid garden	1.17	0.79	238
Distance to build-up boundary	3365.77	2019.67	238
Distance to high way	1417.16	1008.70	238
Distance to city center	6657.31	3502.96	238
Distance to college	1153.14	771.02	238
Density of bus station	11.66	5.40	238
Density of park	0.30	0.39	238

4.5.2. Module summary

Stepwise regression method was used to do the multiple linear regression analysis and select the signature indicators to build the price model, shown as Table 17.

Table 17 Table of model summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.423a	.179	.176	1370.909637213770300	.179	51.548	1	236	.000	
2	.547b	.299	.293	1269.280311055851600	.120	40.305	1	235	.000	
3	.607c	.369	.361	1207.319515972372800	.069	25.740	1	234	.000	
4	.630d	.397	.387	1182.551468821267500	.028	10.905	1	233	.001	
5	.639e	.409	.396	1173.643401693564800	.012	4.550	1	232	.034	
6	.648f	.420	.405	1165.259563150523500	.011	4.350	1	231	.038	
7	.656g	.430	.413	1156.904680022952000	.011	4.348	1	230	.038	1.536

a. Predictors: (Constant), Building_height
b. Predictors: (Constant), Building_height, Distance to CBD
c. Predictors: (Constant), Building_height, Distance to CBD, Density of factory
d. Predictors: (Constant), Building_height, Distance to CBD, Density of factory, Distance to shopping mall
e. Predictors: (Constant), Building_height, Distance to CBD, Density of factory, Distance to shopping mall, Distance to park
f. Predictors: (Constant), Building_height, Distance to CBD, Density of factory, Distance to shopping mall, Distance to park, Distance to hospital
g. Predictors: (Constant), Building_height, Distance to CBD, Density of factory, Distance to shopping mall, Distance to park, Distance to hospital, density of primary school
h. Dependent Variable: Price

In the best result, R^2 is 0.430 and adjusted R^2 is 0.413, which can illustrate that the indicators account for 41.3 percent of the property value variations. According to the regression model summary, these indicators, building height, distance to CBD, density of factory, distance to shopping mall, distance to park, distance to hospital and density of primary school, are significant to the property transaction prices. Use these indicators as independent variables and second-hand property transaction prices as the dependent variable to do linear multiple regression analysis, and the coefficients of the model are shown as Table 18.

Table 18 Regression model coefficients

Model	Unstandardized Coefficients: B	Unstandardized Coefficients: Std. Error	Standardized Coefficients: Beta	t	Sig.	Collinearity Statistics: Tolerance	Collinearity Statistics: VIF
(Constant)	6761.133	371.539		18.198	.000		
density of primary school	-118.372	56.765	-.143	-2.085	.038	.524	1.910
density of factory	-465.117	105.737	-.241	-4.399	.000	.826	1.211
distance to park	-.171	.068	-.168	-2.515	.013	.552	1.810
distance to CBD	-.157	.035	-.367	-4.535	.000	.379	2.641
distance to hospital	-.162	.065	-.174	-2.498	.013	.513	1.949
distance to shopping mall	.610	.183	.204	3.340	.001	.666	1.502
Building_height	12.020	2.868	.228	4.191	.000	.836	1.196

The most important indicator is distance to CBD, whose coefficient is negative. The second important indicator is density of factory, whose coefficient is negative as expected. The third important indicator is building height, whose coefficient is positive. The next two important indicators are the distance to shopping mall and the distance to the hospital.

The final model can be written as:

$Price_{2010}$= 6761.133 - (density of primary school * 118.372) –(density of factory *465.117) –(distance to park *0.171) –(distance to CBD *0.157) –(distance to hospital *0.162) –(distance to shopping mall*0.610)+(building height *12.020)

4.5.3. Evaluating model

No multicollinearity

All the indicators used in the final model are significant and have passed the t-test as shown in figure 36. Besides, the tolerance value of each indicator is more than 0.2 (Ziegel & Menard, 1996) and the VIF value of each indicator is less than 10 means that the model is no multicollinearity (Hair et al., 2012).

Homoscedasticity/ Independent Errors

The size of the error term should be same across values of an independent variable (Field et al., 2012) and the residuals are uncorrelated with the dependent variable (Ai, 2005); otherwise the situation is known as heteroscedasticity. As shown in Figure 22 the data obey the assumption.

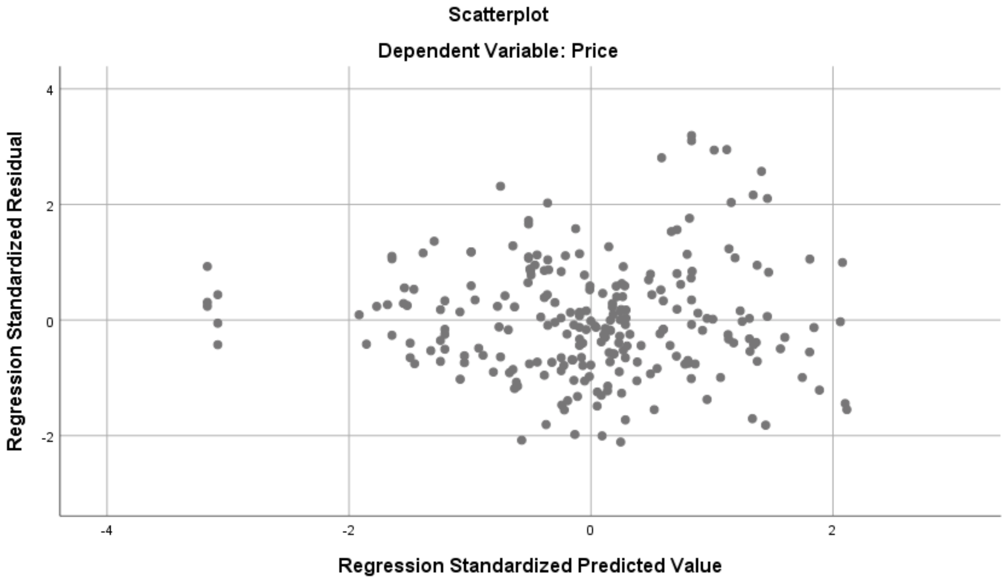


Figure 22 Regression standardized predicted value

Normally-distributed Errors

It checks whether the regression residuals are normally distributed with a mean of zero (Ziegel & Menard, 1996). As shown in Figure 23 the regression standardized residual is a normal distribution.

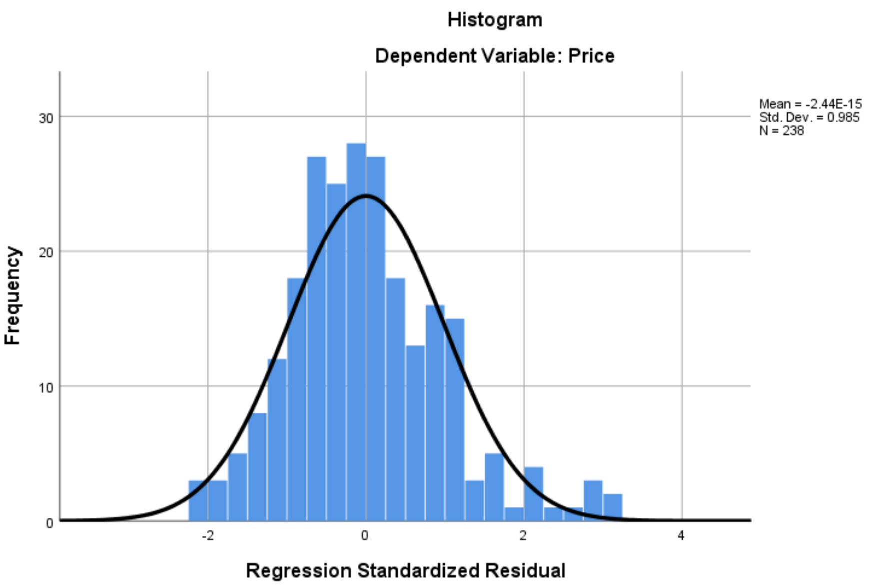


Figure 23 Regression standardized residual

4.5.4. Generalize the model to the study area.

Use the model builder in ArcGIS to generalize the model to the entire study area and adjust the value based on the inflation value. The result is shown as Figure 24.

The price of the property variation from about 2175.41 to 12220.60 and the mean value is 6201.95 (Chinese RMB yuan per sq. meter), shown with different colours in the map. The shapes with red colour mean higher price than the shapes with blue colour. Obviously, the price in the south of the city is higher than the other parts. The red cross symbols represent hospitals that is an important indicator. As for another important indicator, density of factory, is shown through grayscale in the map. The darker of the pixel, the higher density is. In addition, the green point in the map represents park and the dark triangle represents CBDs in the map.

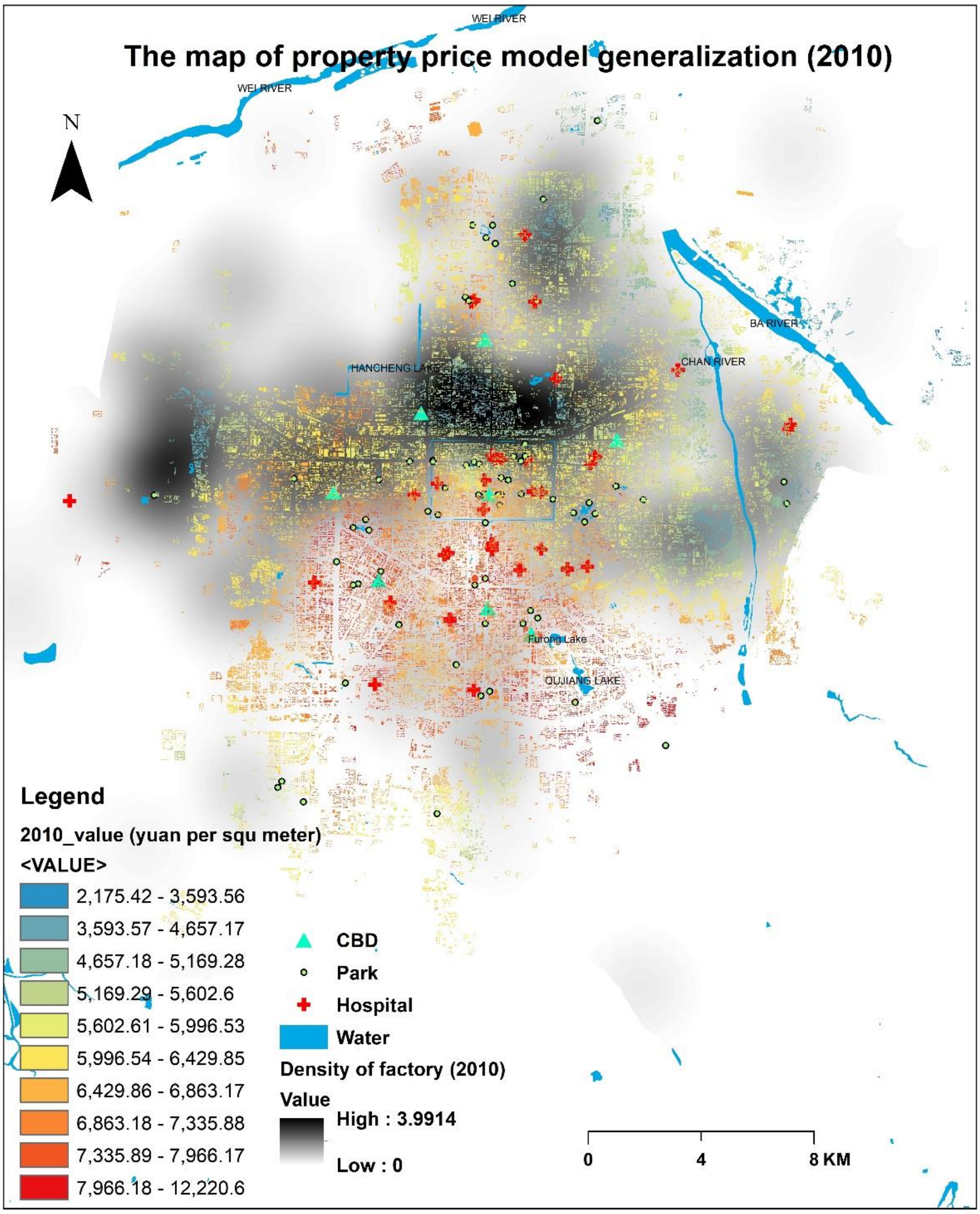


Figure 24 The map of property price model generalization (2010)

4.6. K-fold cross validation

In this research using ten-fold cross validation to check if the regression result is acceptable. Divide the data set into ten parts, and take 9 of them as training data and 1 part as test data in turn. Each test will result in a corresponding correct rate (or error rate). The average of the correct rate (or error rate) of the results of 10 times is used as an estimate of the accuracy of the model. The K-fold cross validation of each regression model in different year is shown as Table 19.

Table 19 Result of K-fold cross validation

Year	Result of k-fold cross validation
2018	2613.824
2014	1648.095
2010	1133.347

. As mentioned above, in this thesis, to fit the city scale, the average transaction price is in a Xiaoqu level and in the same Xiaoqu in Xi'an, the price variation range can be over 3000 RMB yuan, so the value of the K-fold cross validation is acceptable.

4.7. Visualization the result

Based on the generalization of property price model, the result is also visualized through CityEngine. Different colors show different ranges of price; the darker red represents the higher price, shown as the following Figure 25, Figure 26 and Figure 27. The results are shared in ArcGIS Online [1].

[1] https://www.arcgis.com/home/item.html?id=09e25f9df5d5443698c23b870652448c

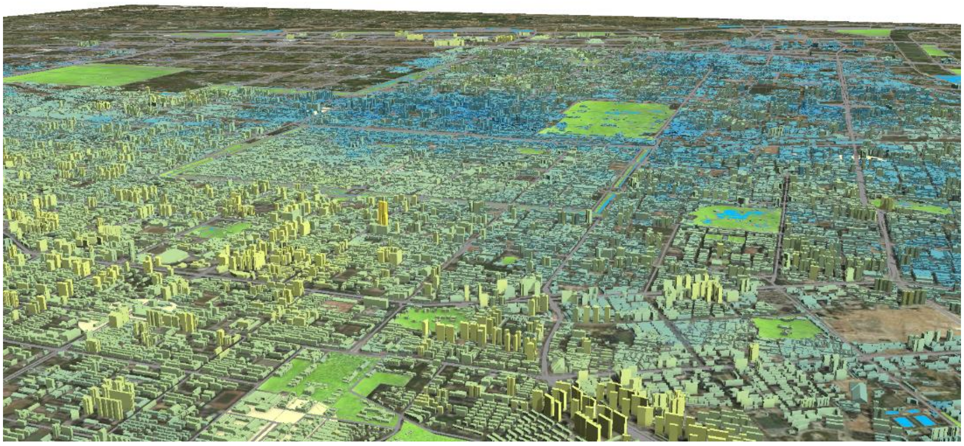
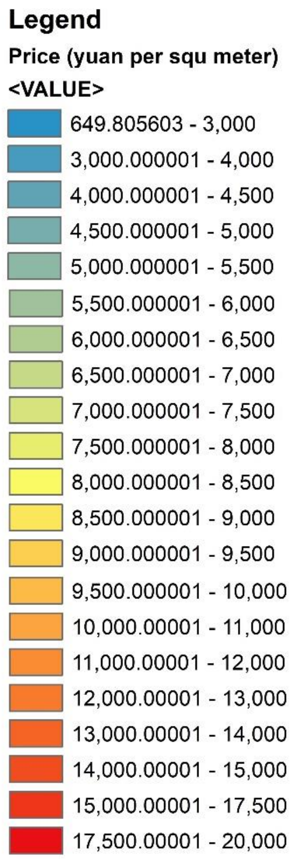


Figure 25 Result visualization (2010)



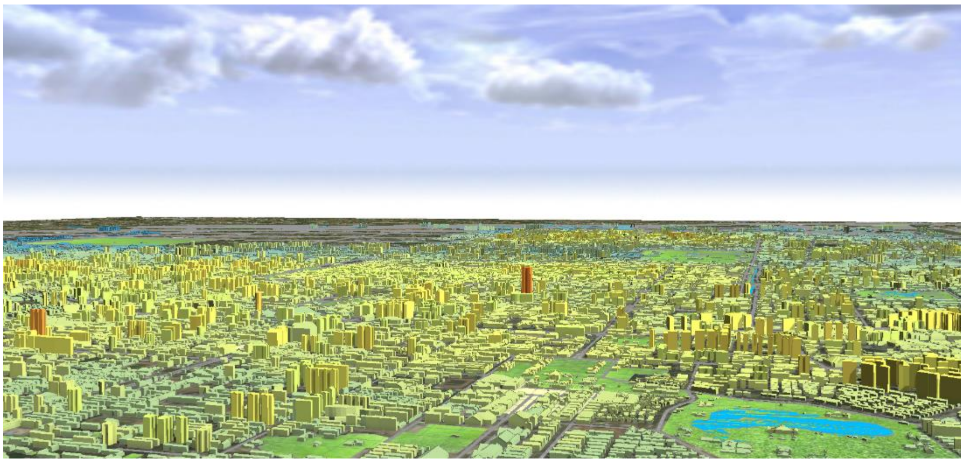


Figure 26 Result visualization (2014)

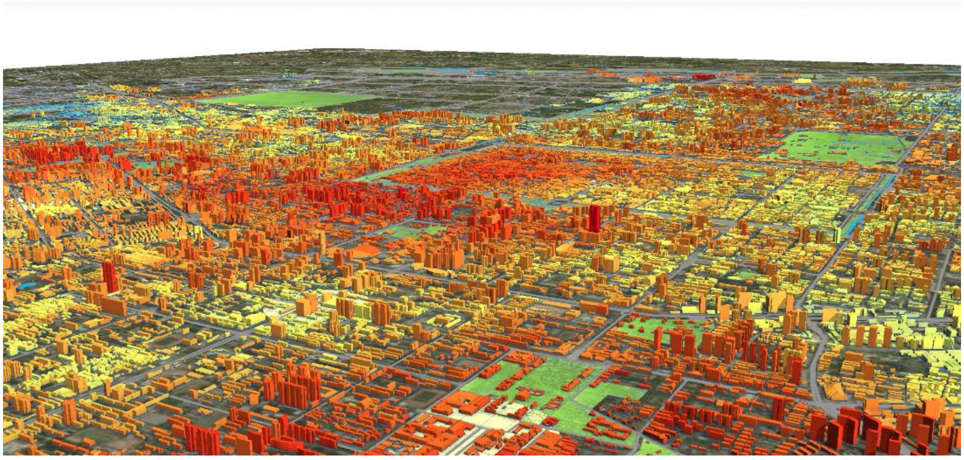


Figure 27 Result visualization (2018)

4.8. Summary

In this chapter, the result of questionnaires and interviews are described to help select the appropriate indicators for the next regression analysis. The second parts are the results of the regression analysis in 2010, 2014 and 2018. It illustrates the change of important indicators of property prices from 2010 to 2018. And the formulas of model for 2010,2014 and 2018 are the following:

$Price_{2010}$= 6761.133 - (density of primary school * 118.372) – (density of factory *465.117) – (distance to park *0.171) – (distance to CBD *0.157) – (distance to hospital *0.162) – (distance to shopping mall*0.610) +(building height *12.020)

$Price_{2014}$= 5455.728+(building height *19.226) +(density of bus station*151.082) + (AQI * 93.635) - (distance to hospital *0.482) +(distance to metro station * 0.346)- (density of park *659.056)

$Price_{2018}$= 27097.425+(building height *39.224) -(density of key high school *9548.446) – (density of hospital * 1236.913) +(density of park *3203.286) +(NDVI*8899.388)- (AQI*115.033) - (density of factory *1027.127) – (distance to shop mall*2.094)- (distance to metro station*0.91) – (distance to college*0.832) – (density of entertainment*116.107)

Visualizing the model in CityEngine is the third parts, which is helpful to show the relationship of 4D indicators and property price.

5. Discussion

This chapter is used to discuss the result findings and limitation of this research.

5.1. Indicators selecting and methods

Indicators of property prices can be categorized into three types: physical characteristics, location characteristics, and environment characteristics (Jim & Chen, 2006; Kahr & Thomsett, 2005). Through literature review, interviews and questionnaires, these indicators have been selected: the distance to metro station, bus stations' density, distance to high way (Adair et al., 2000), distance to CBDs, distance to the build-up area boundary, kid gardens' density, primary schools' density, high schools' density, distance to college(Clark & Herrin, 2000; Haurin & Brasington, 1996; Haizhen Wen et al., 2014), distance to shopping mall, supermarkets' density, distance to hospital (Des Rosiers et al., 1996; Peng & Chiang, 2015), distance to park, factories' density, distance to water, Air quality index (AQI) value and NDVI value (Park et al., 2017; Sander & Polasky, 2009; Smith & Huang, 1995).

The indicators are quantified by spatial analysis, such as the accessibility of indicators, the air quality, and the vegetation covering level, etc. In this research, for the same indicator, in different years, sometimes distance of the points is used and sometimes the density of the points is used, which is because with the development of the city, the number of the same type of the points are more. For example, the number of hospitals in 2018 is much larger than the other two years. So, the distance to hospital in 2018 is not significant anymore, while the density of hospitals works better in the regression model.

5.2. Explanation of results

5.2.1. Discussion about the value of R^2 of the regression result

All the value of R^2 of the regression analysis is acceptable, but the value of R^2 is not very high, shown as Table 20. It is because the price is influenced not only the indicators used in the research and the average price of each Xiaoqu calculated by the limited property transaction price. As mentioned, for each property, the price is influenced by the physical characteristics, such as the floor level, area, number of the rooms, building age, etc. For some Xiaoqu, the number of transaction property is small or even one transaction property, causing the average price is not true average price and still influenced by the physical characteristics.

Besides, the R^2 of the 2014 regression model is less than the other years. It is because the housing market was unstable and the property price fluctuated across the year of 2014. However, in this research, the indicators are calculated for the whole year and cannot show the variation over the year. On the other hand, the fluctuated price also influenced by other indicators, such as the purchase intention of the whole society in Xi'an and issuing the relevant policies, which are not considered in the research. In addition, the limitation of the data set may be another reason.

Table 20 R square of the regression analysis

Year	R^2	Adjusted R^2
2018	0.446	0.415
2014	0.314	0.294
2010	0.430	0.413

5.2.2. Discussion about the most important indicators

Through stepwise regression analysis, the main indicators of property price for each year in Xi'an have been selected. The number of most important indicators in different years are different due to the change of buyer's concern on the indicators when buying a residential property over the years.

According to the chapter of the result, it can be found that the building height and accessibility to the hospital are always the most important indicators in 2010, 2014, 2018. Building height is always a positive indicator. But the attitude of local people for the accessibility of hospital changed across the years. In 2010 and 2014, the less distance to the hospital means higher property prices. However, in 2018, the density of the hospital becomes a negative indicator. It is may because that the hospitals are accessible to most properties in 2018 and higher density of hospitals may bring more noise, such as the noise of ambulance.

The indicators about the environment are important in the three years. The most important indicators about the environment in 2010 are the distance to park and density of factories; in 2014 are AQI and density of park. In 2018 the most important indicators about the environment are more than the other years, including AQI, NDVI, the density of factory, the density of park. It illustrates that local people focus more on the environmental characteristics. It may because up to 2018, more properties are bought for improving the level of living and local people pay more attention to health, which is also mentioned by the respondents of the interviews.

The convenience of shopping is another main indicator across the years, but comparing with the situation in 2018, local people in 2010 focus more on the distance to CBD. It may because with the development of the city, more and more shop malls distribute across the city in 2018 and online shopping in the whole China develops fast, causing the convenience of shopping.

Building height is always important across the eight years, it may because in this period, most new properties are in high buildings. Besides, as mentioned, the properties in high height, always means a better viewer, better building structure quality, more facilities in the building, etc. (Jim & Chen, 2009; Wong et al., 2011; Yu, Han, & Chai, 2007).

5.2.3. Discussion about the result generalization

It is obvious that the whole price of the property increases a lot from 2010 to 2018, and the price of the properties in the south of the city is always higher than most prices of properties in the north of the city. And the highest property price zone transfers from south-west to south-east over the 8 years. In 2010, the most expensive properties were in the west of Yanta District, one of the reasons is at that time Gaoxin District developed rapidly and could provide most high-tech jobs. In addition, in the north of the city, the property price was also high, which is because the government of Xi'an had mover to this area to promote the development of Weiyang District. In 2014, parts of property prices are less than prices in 2010. The property with the highest price distributed in south-west and west of the city. In 2018, the property prices are much higher than the other two years. More properties with high price distributed in the south-east area of the city, which may because of the rapid development of the south-east part of Yanta District. Through the interview with local urban planners, it is found that the aims of the west area in Yanta and east area in Yanta are different. The east area in Yanta District focuses on travel cultural and historical cultural, while the west part in Yanta focuses on high-tech. Besides, the east part of Yanta District has better natural scenery and better air quality.

However, comparing with the price map in 2010, it can be found that the price of the properties between Chan River and Ba river increases a lot, and even properties in this area become one of the most expansive parts in 2018. Part of the reason is that closing to Chan River and Ba River, the environment of this area is one of the best parts across the city. Especially after the construction Xi'an World Expo Site, the properties there are more attractive. Besides, the land in the south of the city has been nearly used out until 2018 and the land in the east city is a new development zone, so most properties in the east area are younger than the properties in the south part. Therefore, the properties in the east part of the city become more and more popular. Inverse distance weighted (IDW) interpolation method is used to show how property prices distribute in different years. As shown in Figure 28.

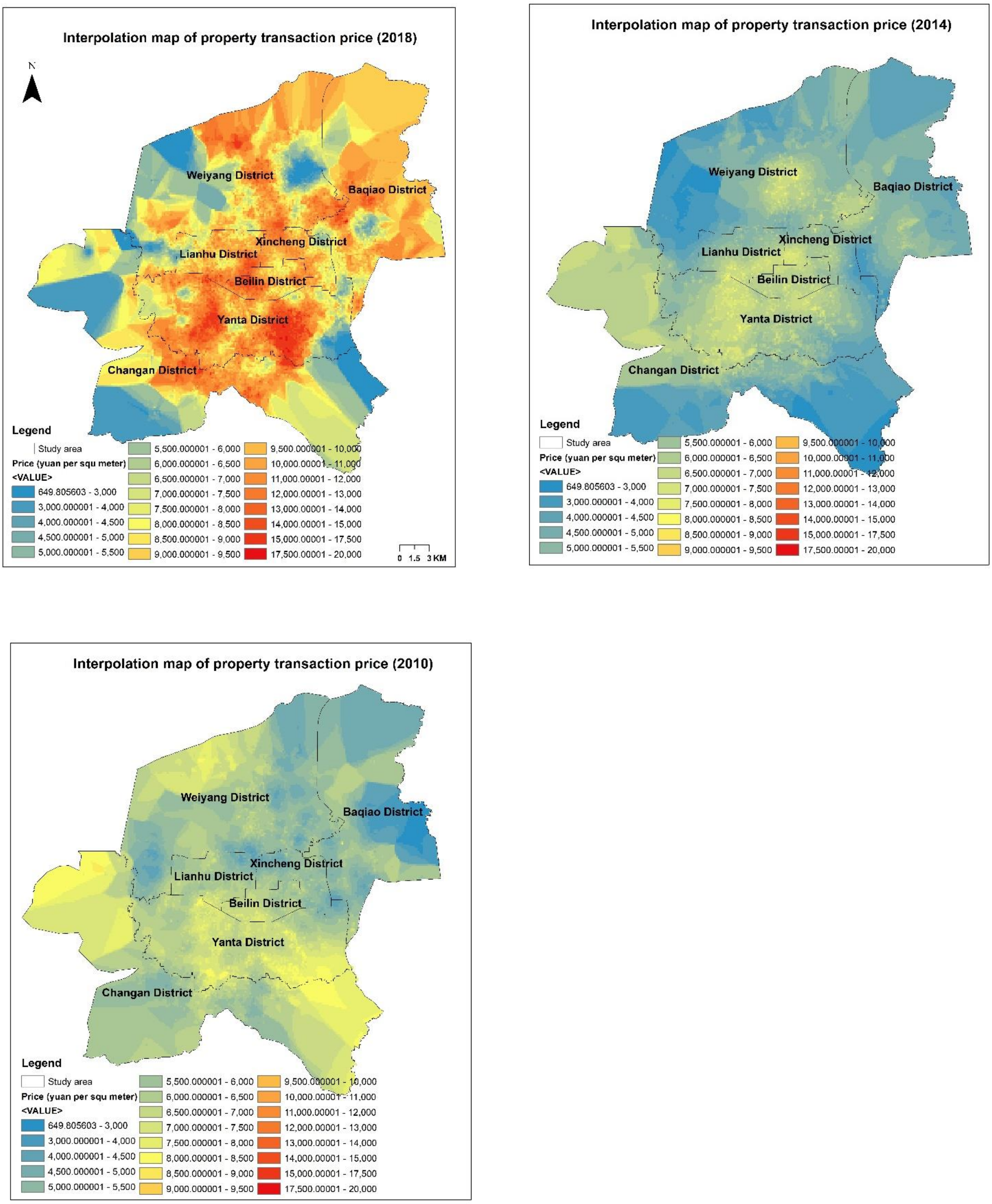


Figure 28 Interpolation map of the property transaction price

5.3. Limitation

Because of the limited time and data set, there are several limitations existing in this research.

1. Through the results from questionnaire and interviews, different people in different age range focus on different characteristics of the property. For instance, old people prefer to living near the hospital and park, and they prefer living on the lower floor level. For young people, they would focus more on the transportation of the property; the people who have or will have a kid prefer to live near the school. In this research, these people are not separated due to the time limitation and data set limitation. However, it would be better to add these factors in the model.

2. In this research, only one 3D indicator, building height, is used to develop the model. In the actual property market, there are many other 3D indicators would influence the property price, such as floor level, the sight of the apartment, the amount of light of an apartment per day, etc.

3. Lack of the polygons data of buildings in 2010 and 2014, the result of the regression model is generalized based on the building polygons in 2018. With the urbanization of Xi'an, some building with low height in 2010 has been reconstructed to high buildings, which would influence the generalization result of 2010 and 2014.

6. Conclusion

6.1. Reflection on research objective

As mentioned in the first chapter, with the quick urbanization in China, the residential property market has grown substantially over the decade. The residential property price has got high attention from the whole society. A good understanding of property prices, is important for housing security, property market development, urban planning and urban management. This can be achieved by using property price models. This can also help property buyer to better understand the property market and provide suggestions when they buying a residential property.

This research focused on analysing the important indicators of Xi'an city's second-hand residential property transaction prices in 2010, 2014 and 2018 respectively.

The research objectives are the following:

1. To review the hedonic property pricing models in China.
2. To develop a 4D Property valuation model based on Remote Sensing Data
3. To evaluate the model

6.2. Main conclusion

Selecting Xi'an, China as study area, this research used GIS to describe and analysis the spatial distribution of property prices and spatial characteristics across the city. This research also used regression analysis to detect the effects of each indicators on second hand property transaction prices. The main conclusions are the following:

1. In the past eight years, the spatial structure of residential land prices in Xi'an city has undergone changes. The development of the city is the internal driving force for the changes in the spatial distribution of urban property prices. For example, the price of properties in the main development areas planned by the city government will be higher or increase faster than other areas. Because the development of the area has promoted the improvement of infrastructure and public facilities, the ecological and living environment. It improves the property location characteristics, leading to the changes of spatial distribution pattern of urban property prices.
2. The price of properties in Xi'an increases a lot from 2010 to 2018. However, the average price in 2014 dose not change too much comparing with the price in 2010, and the price increase rapidly during 2014 and 2018, especially after 2016.
3. The spatial differences have been found in both property prices and spatial indicators, such as accessibility to public transportation facilities, education, hospitals and parks. It is found that high price properties are near the sufficient facilities, high transportation accessibility, and good environment (e.g. high air quality, few factories, high density of parks).
4. It can be found that the indicators influence the property prices in Xi'an has several changes over the eight years. Local people focus more on the environment characteristics, such as the density of factories, AQI, distance of parks, NDVI, etc. And with the urbanization of Xi'an, the living service facilities are improved a lot than the statues in 2010, so the local people do not care much about the proximity to CBDs. However, several indicators show their consistent importance throughout the years, i.e. density of factory and distance to park or density of park.
5. The 3D indicator, building height, is always the most important indicators of property price. In other words, property buyers would like to pay more for the property in a higher building.

6. As mentioned above, the research illustrated how property prices and most important indicators of property prices distributes across the eight years, which is helpful for property buyers to understand how real estate market develops. Besides, the visualization of result can provide suggestions for the buyers to select a cost-effective property with better physical, location and environmental characteristics and lower price.

6.3. Recommendations

6.3.1. Recommendations for future study

In the future study, some recommendations are as follows:

1. It would be better to add more temporal indicators to detect the regulation of property prices change and develop a more accurate 4D property valuation model, such as the GDP, population, income level of citizens of the city.
2. The object could be more detailed, such as an apartment instead of the average of price of each Xiaoqu. In that case, more 3D indicators could be used in research, including view, floor level, daylighting amount, etc.
3. Besides the second-hand transaction property price, new property transaction price could also be one research object. And comparing the mode of new properties and second-hand transaction properties could help to find out the differences between real estate company price policy and property price in the actual real estate market.
4. More scales could be considered, to find out how property price distributed in a bigger region or in a smaller region, such as the distribution and indicators in the whole country or in a district.
5. Based on more necessary indicators mentioned above, the model can be used to predict property price of the city and provide suggestions for local government to make more reasonable property transaction and urban planning policies.

6.3.2. Recommendations for practice

For potential property buyers, as mentioned above, the environment in the east of the city is good, the facilities are sufficient and the price is not the highest. Therefore, buying a property near Chan River and Ba River could be a good choice. Besides, the west part of the city is cheaper than the other parts, so if the buyers do not have much money, a property in the west of the city near Hancheng Lake could be a good choice. For the buyers who want to live with those who have a high income, buying properties in Yanta District, near the Qujiang Lake or Furong Lake is still one of the best choices.

For the government, the price of properties in the west of the city increased slower than the other parts across the years. To balance the real estate market, it would pay more attention to the construction of living facilities and improve the environment in this zone, such as moving the factory to an area farther away from the city and increasing the number of parks in this area.

The real estate development company can combine the hedonic price model and comparison method to set the sold price of a new property project. Nowadays, these companies in China commonly use a comparison method and cost method as the approach to set sold price. In that case, they may ignore the property attributes that are focused by the property buyers, which may cause the sold price setting unreasonable.

LIST OF REFERENCES

Abidoye, R. B., & Chan, A. P. C. (2017). Critical review of hedonic pricing model application in property price appraisal: A case of Nigeria. *International Journal of Sustainable Built Environment*, *6*(1), 250–259. https://doi.org/10.1016/J.IJSBE.2017.02.007

Adair, A., McGreal, S., Smyth, A., Cooper, J., & Ryley, T. (2000). House Prices and Accessibility: The Testing of Relationships within the Belfast Urban Area. *Housing Studies*, *15*(5), 699–716. https://doi.org/10.1080/02673030050134565

Ai, P. (2005). Residential land value modelling, 84.

AirNow. (2016). Air Quality Index (AQI) Basics. Retrieved February 12, 2019, from https://airnow.gov/index.cfm?action=aqibasics.aqi

Andersson, D. E., Shyr, O. F., & Fu, J. (2010). Does high-speed rail accessibility influence residential property prices? Hedonic estimates from southern Taiwan. *Journal of Transport Geography*, *18*(1), 166–174. https://doi.org/10.1016/J.JTRANGEO.2008.10.012

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., & Ridella, S. (2012). The ' K ' in K-fold Cross Validation. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. https://doi.org/10.1007/s00787-016-0866-x

Ball, M. J. (1973). Recent Empirical Work on the Determinants of Relative House Prices. *Urban Studies*, *10*(2), 213–233. https://doi.org/10.1080/00420987320080311

Barrow, D. K., & Crone, S. F. (2013). Crogging (cross-validation aggregation) for forecasting — A novel algorithm of neural network ensembles on time series subsamples. In *The 2013 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8). IEEE. https://doi.org/10.1109/IJCNN.2013.6706740

Benson, E. D., Hansen, J. L., Schwartz, Jr., A. L., & Smersh, G. T. (1998). Pricing Residential Amenities: The Value of a View. *The Journal of Real Estate Finance and Economics*, *16*(1), 55–73. https://doi.org/10.1023/A:1007785315925

Bolarinwa. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195. https://doi.org/10.4103/1117-1936.173959

Bray, D. (2005). *Social space and governance in urban China : the danwei system from origins to reform*. Stanford University Press. Retrieved from https://www.sup.org/books/title/?id=6335

Bremer, M. (n.d.). *Multiple Linear Regression*. Retrieved from http://mezeylab.cb.bscb.cornell.edu/labmembers/documents/supplement 5 - multiple regression.pdf

Brownlee, J. (2018). A Gentle Introduction to k-fold Cross-Validation. Retrieved February 6, 2019, from https://machinelearningmastery.com/k-fold-cross-validation/

Bureau of Statistics of Xi'an. (2017). 2017 Xi'an Statistical Yearbook. Retrieved August 13, 2018, from http://tjj.xa.gov.cn/ptl/def/def/index_923_6528_ci_trid_4417083.html

Butler, R. V. (1982). The Specification of Hedonic Indexes for Urban Housing. *Land Economics*, *58*(1), 96. https://doi.org/10.2307/3146079

Ca, B. U., & Fr, Y. G. (2004). *No Unbiased Estimator of the Variance of K-Fold Cross-Validation Yoshua Bengio Yves Grandvalet*. *Journal of Machine Learning Research* (Vol. 5). Retrieved from http://www.jmlr.org/papers/volume5/grandvalet04a/grandvalet04a.pdf

Can, A. (1990). The Measurement of Neighborhood Dynamics in Urban House Prices. *Economic Geography*, *66*(3), 254. https://doi.org/10.2307/143400

Capozza, D. R., Hendershott, P. H., Mack, C., & Mayer, C. J. (2002). *NBER WORKING PAPER SERIES DETERMINANTS OF REAL HOUSE PRICE DYNAMICS*. Retrieved from http://www.nber.org/papers/w9262

Carroll, T. M., Clauretie, T. M., & Jensen, J. (1996). Living next to godliness: Residential property values and churches. *The Journal of Real Estate Finance and Economics*, *12*(3), 319–330. https://doi.org/10.1007/BF00127540

Chau, K. W., & Chin, T. L. (2002, June 12). A Critical Review of Literature on the Hedonic Price Model. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2073594

Chen, A. (2006). Housing Project Hedonic Price Model Based on Principal Components Analysis. *Journal of Chongqing University(Natural Science Edition)*, (06), 144–148. Retrieved from http://en.cnki.com.cn/Article_en/CJFDTOTAL-FIVE200606036.htm

Chen, J., Guo, F., & Wu, Y. (2011). One decade of urban housing reform in China: Urban housing price dynamics and the role of migration and urbanization, 1995–2005. *Habitat International*, *35*(1), 1–8. https://doi.org/10.1016/J.HABITATINT.2010.02.003

Chen, J., & Hao, Q. (2008). The Impacts of Distance to CBD on Housing Prices in Shanghai: A Hedonic Analysis. *Journal of Chinese Economic and Business Studies*, *6*(3), 291–302. https://doi.org/10.1080/14765280802283584

Chen, J., Hao, Q., & Stephens, M. (2010). Assessing Housing Affordability in Post-reform China: A Case Study of Shanghai. *Housing Studies*, *25*(6), 877–901. https://doi.org/10.1080/02673037.2010.511153

Chen, L., & Fan, H. (2014). Housing price fluctuation, monetary policy and the social welfare loss in china (in Chinese). *Chinese Journal of Management Science*, *5*.

Chen, Z., & Haynes, K. E. (2015). Impact of high speed rail on housing values: an observation from the Beijing–Shanghai line. *Journal of Transport Geography*, *43*, 91–100. https://doi.org/10.1016/J.JTRANGEO.2015.01.012

Cheng, C., & Agrawal, A. (2010). TTSAT: A New Approach to Mapping Transit Accessibility. *Journal of Public Transportation*, *13*(1), 55–72. https://doi.org/10.5038/2375-0901.13.1.4

Clark, D. E., & Herrin, W. E. (2000). The impact of public school attributes on home sale prices in California. *Growth and Change*. https://doi.org/10.1111/0017-4815.00134

Colwell, P. F., & Dilmore, G. (1999). Who Was First? An Examination of an Early Hedonic Study. *Land Economics*, *75*(4), 620. https://doi.org/10.2307/3147070

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Court, A. T. (1939). Hedonic price indexes with automotive examples. In *The dynamics of automobile demand*.

Darling, A. H. (1973). Measuring Benefits Generated by Urban Water Parks. *Land Economics*, *49*(1), 22. https://doi.org/10.2307/3145326

Debrezion, G., Pels, E., & Rietveld, P. (2007). The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics*, *35*(2), 161–180. https://doi.org/10.1007/s11146-007-9032-z

Des Rosiers, F., Lagana, A., Thériault, M., & Beaudoin, M. (1996). Shopping centres and house values: an empirical investigation. *Journal of Property Valuation and Investment*, *14*(4), 41–62. https://doi.org/10.1108/14635789610153461

DeVellis, R. F. (2003). *Scale development: Theory and applications*. *International Journal of Industrial Ergonomics*.

Dubin, R. A. (1998). Spatial Autocorrelation: A Primer. *Journal of Housing Economics*, *7*(4), 304–327. https://doi.org/10.1006/JHEC.1998.0236

Dubin, R. A., & Sung, C.-H. (1990). Specification of hedonic regressions: Non-nested tests on measures of neighborhood quality. *Journal of Urban Economics*, *27*(1), 97–110. https://doi.org/10.1016/0094-1190(90)90027-K

Feng, L., & Li, F. L. (2011). Study of commercial residential character price in Xi'an Qu Jiang New District. In *2011 International Conference on Business Management and Electronic Information* (pp. 199–202). IEEE. https://doi.org/10.1109/ICBMEI.2011.5920428

Feng, X., & Humphreys, B. R. (2008). *Assessing the Economic Impact of Sports Facilities on Residential Property Values: A Spatial Hedonic Approach*. Retrieved from http://college.holycross.edu/RePEc/spe/FengHumphreys_PropertyValues.pdf

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using IBM SPSS Statistics*. *Statistics*. https://doi.org/10.1111/insr.12011_21

Fletcher, M., Gallimore, P., & Mangan, J. (2000). Heteroscedasticity in hedonic house price models. *Journal of Property Research*, *17*(2), 93–108. https://doi.org/10.1080/095999100367930

Follain, J. R., & Jimenez, E. (1985). Estimating the demand for housing characteristics: A survey and critique. *Regional Science and Urban Economics*, *15*(1), 77–107. https://doi.org/10.1016/0166-0462(85)90033-X

Garrod, G., & Willis, K. (1994). An economic estimate of the effect of a waterside location on property values. *Environmental & Resource Economics*, *4*(2), 209–217. https://doi.org/10.1007/BF00692204

Geisser, S. (2017). *Predictive Inference*. Routledge. https://doi.org/10.1201/9780203742310

Gillard, Q. (1981). THE EFFECT OF ENVIRONMENTAL AMENITIES ON HOUSE VALUES: THE EXAMPLE OF A VIEW LOT. *The Professional Geographer*, *33*(2), 216–220. https://doi.org/10.1111/j.0033-0124.1981.00216.x

Glaeser, E. L., Gyourko, J., & Saks, R. (2005). Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices. *The Journal of Law and Economics*, *48*(2), 331–369. https://doi.org/10.1086/429979

Goodman, A. C. (1998). Andrew Court and the Invention of Hedonic Price Analysis. *Journal of Urban Economics*, *44*(2), 291–298. https://doi.org/10.1006/JUEC.1997.2071

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. https://doi.org/10.1016/J.REGSCIURBECO.2014.09.002

Haas, G. C. (1922). A Statistical Analysis of Farm Sales in Blue Earth County, Minnesota, as a Basis for Farm Land Appraisal. Retrieved from https://conservancy.umn.edu/handle/11299/179691

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414–433. https://doi.org/10.1007/s11747-011-0261-6

Hamilton, S. E., & Morgan, A. (2010). Integrating lidar, GIS and hedonic price modeling to measure amenity values in urban beach residential property markets. *Computers, Environment and Urban Systems*, *34*(2), 133–141. https://doi.org/10.1016/J.COMPENVURBSYS.2009.10.007

Haurin, D. R., & Brasington, D. (1996). School Quality and Real House Prices: Inter- and Intrametropolitan Effects. *Journal of Housing Economics*, *5*(4), 351–368. https://doi.org/10.1006/JHEC.1996.0018

Hill, R. J. (2013). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*, *27*(5), 879–914. https://doi.org/10.1111/j.1467-6419.2012.00731.x

Huang, Z., Chen, R., Xu, D., & Zhou, W. (2017). Spatial and hedonic analysis of housing prices in Shanghai. *Habitat International*, *67*, 69–78. https://doi.org/10.1016/J.HABITATINT.2017.07.002

Huh, S., & Kwak, S.-J. (1997). The Choice of Functional Form and Variables in the Hedonic Price Model in Seoul. *Urban Studies*, *34*(7), 989–998. https://doi.org/10.1080/0042098975691

Investopedia. (2018). Descriptive Statistics. Retrieved January 22, 2019, from https://www.investopedia.com/terms/d/descriptive_statistics.asp#axzz2DxCoTnMM

Isikdag, U., Horhammer, M., Zlatanova, S., Kathmann, R., & Van Oosterom, P. (2015). Utilizing 3D Building and 3D Cadastre Geometries for Better Valuation of Existing Real Estate. *FIG Working Week 2015: From the Wisdom of the Ages to the Challenges of the Modern World*, (May 2015), 1–18.

Jim, C. Y., & Chen, W. Y. (2006). Impacts of urban environmental elements on residential housing prices in Guangzhou (China). *Landscape and Urban Planning*, *78*(4), 422–434. https://doi.org/10.1016/J.LANDURBPLAN.2005.12.003

Jim, C. Y., & Chen, W. Y. (2009). Value of scenic views: Hedonic assessment of private housing in Hong Kong. *Landscape and Urban Planning*, *91*(4), 226–234. https://doi.org/10.1016/J.LANDURBPLAN.2009.01.009

Kahr, J., & Thomsett, M. (2005). *Real estate market valuation and analysis*. Hoboken, New Jersey.: John Wiley & Sons, Incorporated.

Kain, J. F., & Quigley, J. M. (1970). Measuring the Value of Housing Quality. *Journal of the American Statistical Association*, *65*(330), 532–548. https://doi.org/10.1080/01621459.1970.10481102

Kohavi, R. (1995). *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. Retrieved from http://robotics.stanford.edu/~ronnyk

Kong, F., Yin, H., & Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning*, *79*(3–4), 240–252. https://doi.org/10.1016/J.LANDURBPLAN.2006.02.013

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Source: The Journal of Political Economy*, *74*(2), 132–157. Retrieved from https://core.ac.uk/download/pdf/6519151.pdf

Lee, H. S., Park, K., & Kim, S. Y. (2003). Estimation of information value on the Internet: application of hedonic price model. *Electronic Commerce Research and Applications*, *2*(1), 73–80. https://doi.org/10.1016/S1567-4223(03)00002-4

Li, S., Ye, X., Lee, J., Gong, J., & Qin, C. (2017). Spatiotemporal Analysis of Housing Prices in China: A Big Data Perspective. *Applied Spatial Analysis and Policy*, *10*(3), 421–433. https://doi.org/10.1007/s12061-016-9185-3

Liu, S., Fan, X., Wen, Q., Liang, W., & Wu, Y. (2014). Simulated impacts of 3D urban morphology on urban transportation in megacities: case study in Beijing. *International Journal of Digital Earth*, *7*(6), 470–491. https://doi.org/10.1080/17538947.2012.740079

Liu, Y., Gupta, H., Springer, E., & Wagener, T. (2008). Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. *Environmental Modelling and Software*, *23*(7), 846–858. https://doi.org/10.1016/j.envsoft.2007.10.007

McLachlan, G. J., Do, K.-A., & Ambroise, C. (2004). *Analyzing Microarray Gene Expression Data*. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/047172842X

Mok, H. M. K., Chan, P. P. K., Kong, H., & Department, H. (1995). *A Hedonic Price Model for Private Properties in Hong Kong YIU-SUN CHO. Journal of Real Estate Finance and Economics* (Vol. 10). Kluwer Academic Publishers. Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF01099610.pdf

Montero, J.-M., Mínguez, R., & Fernández-Avilés, G. (2018). Housing price prediction: parametric versus semi-parametric spatial hedonic models. *Journal of Geographical Systems*, *20*(1), 27–55. https://doi.org/10.1007/s10109-017-0257-y

NBSC. (2017). National Data. Retrieved August 13, 2018, from http://data.stats.gov.cn/english/adv.htm?m=advquery&cn=E0105

Nguy, A., Sun, C., & Zheng, S. (2014). Airport Noise and Residential Property Values: Evidence from Beijing. In *Proceedings of the 17th International Symposium on Advancement of Construction Management and Real Estate* (pp. 473–481). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35548-6_49

Noor, N. M., Asmawi, M. Z., & Abdullah, A. (2015). Sustainable Urban Regeneration: GIS and Hedonic Pricing Method in Determining the Value of Green Space in Housing Area. *Procedia - Social and Behavioral Sciences*, *170*, 669–679. https://doi.org/10.1016/j.sbspro.2015.01.069

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: a review of valuation methods. *Journal of Property Investment & Finance*, *21*(4), 383–401. https://doi.org/10.1108/14635780310483656

Palmquist, R. B. (1992). Valuing localized externalities. *Journal of Urban Economics*, *31*(1), 59–68. https://doi.org/10.1016/0094-1190(92)90032-G

Park, J., Lee, D., Park, C., Kim, H., Jung, T., Kim, S., … Kim, S. (2017). Park Accessibility Impacts Housing Prices in Seoul. *Sustainability*, *9*(2), 185. https://doi.org/10.3390/su9020185

Peng, T.-C., & Chiang, Y.-H. (2015). The non-linearity of hospitals' proximity on property prices: experiences from Taipei, Taiwan. *Journal of Property Research*, *32*(4), 341–361. https://doi.org/10.1080/09599916.2015.1089923

Qin, J., Fang, C., Wang, Y., Li, G., & Wang, S. (2015). Evaluation of three-dimensional urban expansion: A case study of Yangzhou City, Jiangsu Province, China. *Chinese Geographical Science*, *25*(2), 224–236. https://doi.org/10.1007/s11769-014-0728-8

Resch, B., Hillen, F., Reimer, A., & Spitzer, W. (2013). Towards 4D Cartography – Four-dimensional Dynamic Maps for Understanding Spatio-temporal Correlations in Lightning Events. *The Cartographic Journal*, *50*(3), 266–275. https://doi.org/10.1179/1743277413Y.0000000062

Ridker, R. G., & Henning, J. A. (1967). The Determinants of Residential Property Values with Special Reference to Air Pollution. *The Review of Economics and Statistics*, *49*(2), 246. https://doi.org/10.2307/1928231

Rodriguez, M., & Sirmans, C. (1994). Quantifying the Value of a View in Single Family Housing Markets. *Appraisal Journal*.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Salleh, N. A., Yakin, M. K. A. M., Ismail, K., & Talib, Y. (2016). Prelimanary Investigation on The Factors That Influencing The Maintenance Cost of Apartment. In *The 4th International Building Control Conference 2016 (IBCC 2016)* (Vol. 00046). https://doi.org/10.1051/matecconf/20166600046

Sander, H. A., & Polasky, S. (2009). The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. *Land Use Policy*, *26*(3), 837–845. https://doi.org/10.1016/J.LANDUSEPOL.2008.10.009

Seo, K., Golub, A., & Kuby, M. (2014). Combined impacts of highways and light rail transit on residential property values: a spatial hedonic price model for Phoenix, Arizona. *Journal of Transport Geography*, *41*, 53–62. https://doi.org/10.1016/J.JTRANGEO.2014.08.003

Shin, K., Washington, S., & Choi, K. (2007). Effects of Transportation Accessibility on Residential Property Values. *Transportation Research Record: Journal of the Transportation Research Board*, *1994*(1), 66–73. https://doi.org/10.3141/1994-09

Sirpal, R. (1994). Empirical Modeling of the Relative Impacts of Various Sizes of Shopping Centers on the Values of Surrounding Residential Properties. *Journal of Real Estate Research*, *9*(4), 487–506. Retrieved from https://ideas.repec.org/a/jre/issued/v9n41994p487-506.html

Smith, V. K., & Huang, J.-C. (1995). Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. *Journal of Political Economy*, *103*(1), 209–227. https://doi.org/10.1086/261981

So, H. M., Tse, R. Y. C., & Ganesan, S. (1997). Estimating the influence of transport on house prices: evidence from Hong Kong. *Journal of Property Valuation and Investment*, *15*(1), 40–47. https://doi.org/10.1108/14635789710163793

Song, Y. (2012). Urban villages and housing values in China. *Regional Science and Urban Economics*, *42*(3), 495–505. https://doi.org/10.1016/J.REGSCIURBECO.2011.06.003

Stone, M. (1974). *Cross-Validatory Choice and Assessment of Statistical Predictions. Source: Journal of the Royal Statistical Society. Series B (Methodological)* (Vol. 36). Retrieved from https://www.jstor.org/stable/pdf/2984809.pdf

Straszheim, M. R. (1975). An Econometric Analysis of the Urban Housing Market. Retrieved from https://econpapers.repec.org/bookchap/nbrnberbk/stra75-1.htm

Tian, L. (2006). Impacts of Transport Projects on Residential Property Values in China: Evidence from Two Projects in Guangzhou. *Journal of Property Research*, *23*(4), 347–365. https://doi.org/10.1080/09599910601095365

Tribby, C. P., & Zandbergen, P. A. (2012). High-resolution spatio-temporal modeling of public transit accessibility. *Applied Geography*, *34*, 345–355. https://doi.org/10.1016/J.APGEOG.2011.12.008

Tse, R. Y. C., & Love, P. E. D. (2000). Measuring residential property values in Hong Kong. *Property Management*, *18*(5), 366–374. https://doi.org/10.1108/02637470010360669

Wallenwein, F. (2013). *The Housing Model xiaoqu 小区: the Expression of an Increasing Polarization of the Urban Population in Chinese Cities ? Heidelberg University*. Heidelberg University. Retrieved from http://archiv.ub.uni-heidelberg.de/volltextserver/17143/1/The Housing Model xiaoqu_Wallenwein.pdf

Wang, L., & Sun, T. (2014). Capitalization of legal title: Evidence from small-property-rights houses in Beijing. *Habitat International*, *44*, 306–313. https://doi.org/10.1016/J.HABITATINT.2014.07.013

Wang, Y., Bai, P., & Wang, E. (2015). An appraisal of Xi'an real estate developers core competence. *The Strategies of China's Firms*, 109–122. https://doi.org/10.1016/B978-0-08-100274-2.00007-8

Waugh, F. V. (1928). Quality Factors Influencing Vegetable Prices. *Journal of Farm Economics*, *10*(2), 185. https://doi.org/10.2307/1230278

Weier, J., & Herring, D. (2000). Measuring Vegetation (NDVI & EVI). Retrieved from https://earthobservatory.nasa.gov/features/MeasuringVegetation

Wen, H., Jia, S., & Guo, X. (2005). Hedonic price analysis of urban housing: An empirical research on Hangzhou, China. *Journal of Zhejiang University-SCIENCE A*, *6*(8), 907–914. https://doi.org/10.1631/jzus.2005.A0907

Wen, H., Zhang, L., & Peng, L. (2010). Spatial Variation of Housing Prices in Hangzhou City: Two-Dimensional Analysis Based on Hedonic Prices. *China Land Science*, *24*(2), 51–56. https://doi.org/10.3969/j.issn.1001-8158.2010.02.009

Wen, H., Zhang, Y., & Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in Hangzhou, China. *Habitat International*, *42*, 155–163. https://doi.org/10.1016/J.HABITATINT.2013.12.004

Wilhelmsson, M. (2002). Spatial Models in Real Estate Economics. *Housing, Theory and Society*, *19*(2), 92–101. https://doi.org/10.1080/140360902760385646

Wong, S. K., Chau, K. W., Yau, Y., & Cheung, A. K. C. (2011). Property price gradients: the vertical dimension. *Journal of Housing and the Built Environment*, *26*(1), 33–45. https://doi.org/10.1007/s10901-010-9203-8

Wu, C., & Sharma, R. (2012). Housing submarket classification: The role of spatial contiguity. *Applied Geography*, *32*(2), 746–756. https://doi.org/10.1016/j.apgeog.2011.08.011

Wu, C., Ye, X., Ren, F., Wan, Y., Ning, P., & Du, Q. (2016). Spatial and Social Media Data Analytics of Housing Prices in Shenzhen, China. *PLOS ONE*, *11*(10), e0164553. https://doi.org/10.1371/journal.pone.0164553

Wu, F. (2002). Sociospatial differentiation in urban China : evidence from Shanghai ' s real estate markets. *Environment & Planning A*, *34*, 1591–1615. https://doi.org/10.1068/a34196

Wu, J., Deng, Y., & Liu, H. (2014). House Price Index Construction in the Nascent Housing Market: The Case of China. *The Journal of Real Estate Finance and Economics*, *48*(3), 522–545. https://doi.org/10.1007/s11146-013-9416-1

Wu, L., Jin, L., Shi, T., Zhang, B., Zhou, Y., Zhou, T., … Liu, Y. (2017). Association between ambient particulate matter exposure and semen quality in Wuhan, China. *Environment International*, *98*, 219–228. https://doi.org/10.1016/J.ENVINT.2016.11.013

Wyatt, P. (2013). *Property Valuation*. Chichester, West Sussex, United Kingdom: John Wiley & Sons, Incorporated.

Xi'an Education Administration Bureau, Xi'an Housing Administration Bureau, & Xi'an Administration for Industry and Commerce. (2018). Notice on Prohibition of Linking Housing Sales to School Districts and Schools_Xi'an

Housing Security and Housing Administration. Retrieved November 15, 2018, from http://fgj.xa.gov.cn/zw/Detail.aspx?id=10173

Xi'an Housing Administration Bureau. (2018). Notice on Further Strengthening the Management of Commercial Housing Sales_Xi'an Housing Security and Housing Administration. Retrieved November 15, 2018, from http://fgj.xa.gov.cn/zw/Detail.aspx?firsttypeid=4&id=9554

Xi'an Price Administration Bureau. (2017). Notice of Xi'an Price Bureau on Relevant Issues Concerning the Declaration of Commodity Housing Price. Retrieved November 15, 2018, from http://wjj.xa.gov.cn/ptl/def/def/index_1285_3892_ci_trid_2626196.html

Xi'an Statistical Bureau. (2018). *Xi'an Statistical Yearbook-2018*. (M. Zheng & X. Weng, Eds.). Xi'an: China Statistics Press.

Xiao, J., Shen, Y., Ge, J., Tateishi, R., Tang, C., Liang, Y., & Huang, Z. (2006). Evaluating urban expansion and land use change in Shijiazhuang, China, by using GIS and remote sensing. *Landscape and Urban Planning*, *75*(1–2), 69–80. https://doi.org/10.1016/J.LANDURBPLAN.2004.12.005

Xiao, Y., Webster, C., & Orford, S. (2016). Identifying house price effects of changes in urban street configuration: An empirical study in Nanjing, China. *Urban Studies*, *53*(1), 112–131. https://doi.org/10.1177/0042098014560500

Yang, L., Zhou, J., Shyr, O. F., & Huo, (Derek) Da. (2018). Does bus accessibility affect property prices? *Cities*. https://doi.org/10.1016/J.CITIES.2018.07.005

Yu, S. M., Han, S. S., & Chai, C. H. (2007). Modeling the value of view in high-rise apartments: A 3D GIS approach. *Environment and Planning B: Planning and Design*, *34*(1), 139–153. https://doi.org/10.1068/b32116

Zhang, S. (2014). *Research on the price of subdistricts in Shanghai based on hedonic price model (in Chinese)*. Fudan University.

Ziegel, E. R., & Menard, S. (1996). Applied Logistic Regression Analysis. *Technometrics*. https://doi.org/10.2307/1270433

APPENDIX

Questionnaires

Information Sheet

This questionnaire is part of the research "Developing a 4D Property Valuation Model Based on Remote Sensing Data at City Scale (Xi'an, China)" carried out by Ruijie Li majored in Land Administration of University of Twente.

The main goal of this research is to develop a 4D property valuation Model at city scale. This questionnaire will help to detect the main indicators of residential property price in Xi'an.

You are invited to fill in this questionnaire on a voluntary basis. It will take you 10 minutes to fill it.

No personal information will be collected and the data used during this research will be maintained confidential and anonymous.

Contact information for questions

For any information about this research, please contact Ruijie Li: r.li-2@student.utwente.nl

Please rate the importance of the following indicators of residential property price according to your own experience.

Indicators	Very important	Important	General	Not important	Not at all important
1. The property is new					
2. The age of the property					
3. The building height (low (<3 floors), middle (3-8 floors), high (9-30 floors), very high (>33 floors))					
4. Size of the residential property					
5. Distance to the large business district					
6. Distance to the city center					
7. Distance to schools					
8. Distance to the bus stop					
9. Distance to the metro station					
10. Distance to the railway station					
11. Distance to the hospital					
12. Distance to the park					
13. Nearby factories without pollution					
14. Distance to the cultural heritage					
15. Other indicators					

Other indicators: ______________________________

1. Your gender is?
A. Male
B. Female
2. Your age is?
A. Under 23 years old
B. Between 23 and 30 years old
C. Between 31 and 40 years old
D. Between 41 and 50 years old
E. Between 50 and 60 years old
F. Over 60 years old
3. Where are you living now? (put point on the map)

4. What is your current form of residence？
A. Self-purchasing
B. Renting a room (including student accommodation)
C. Living with my parents
D. Staff quarters provided by the workplace
5. Are you satisfied with the condition of the house you are living in now?
A. Very satisfied
B. Satisfied
C. Generally satisfied
D. Less satisfied
E. Very dissatisfied
Why? Or why not?

6. In the past three years, have you bought a residential property in Xi'an?
A. Yes
B. No
7. How do you expect residential prices of Xi’an to change in the coming year?
A. It will increase by 20%
B. It will increase by 15%
C. It will increase by 10%
D. It will increase by 5%
E. It will be basically unchanged
F. It will decrease by 5%
G. It will decrease by 10%
H. It will decrease by 15%
I. It will decrease by 20%
8. The housing price for per square meter that your family can accept is

A. Under 10,000 yuan
B. 10,001-15,000 yuan
C. 15,001-20,000 yuan
D. 20,001-30,000 yuan
E. 25,001-30,000 yuan
F. Over 30,000 yuan

9. What is your future residential property purchase plan?

A. Buying a house within one year
B. Buying a house within one to three years
C. Buying a house within three to five years
D. Buying a house after five years
E. There is currently no plan to buy a house

10. What is the purpose for you to buy a residential property in the future?

A. Self-occupation (first residential property purchase or self-occupied for marriage)
B. Residential property for offspring's wedding
C. Expanding living space and improving living conditions
D. Education (school district room)
E. Pensioned
F. Rental
G. Investment (preservation, appreciation)

11. If you will purchase a residential property in Xi'an, which three administrative districts will you select? And please range them

A. Yanta district
B. Weiyang district
C. Changan district
D. Beilin district
E. Lianhu district
F. Xincheng district
G. Baqiao district
H. Others

12. Are you satisfied the information you can get from the real estate company about how the property variance? ()

A. Very satisfied
B. Satisfied
C. Generally satisfied
D. Less satisfied
E. Very dissatisfied

Questions for interviews

Questions for planners and surveyors:

1. What is the past and future of Xi'an's urban development process in the urban planning?
2. When the planning department is planning for Xi'an, which factors are specifically considered for the planning of construction land? How to make sure the number of new construction land?
3. The policy of Xi'an has attracted a large number of young people to settle, but at the same time it caused a sudden increase in housing demand, resulting in a sharp increase in housing prices. What is the main reason behind the implementation of this policy? Why not slow down the introduction of population to avoid this contradiction?
4. Xi'an's housing prices have risen in recent years. In addition to the population growth, what other factors do you think have stimulated this phenomenon as well?
5. How do real estate prices and planning schemes affect each other?
6. In the face of a rapid population increasing, the speed of urban expansion in Xi'an will also be affected. How does the plan intend to keep the construction of infrastructure and supporting resources keeping pace with population growth?
7. What is the future development of each districts of Xi'an? How is the allocation of resources for these districts carried out?
8. In addition to the three-dimensional development of the city, is there any consideration of the development of the city in the planning process?
9. What is the application of 3D technology in planning? What are the possible future application fields?
10. What factors do you think cause the residential property price increase readily?
11. What factors driving the rise of land price in Xi'an?
12. How does the Price Bureau determine the sales price of commercial housing to stability the property market?

Questions for staff working in real estate companies:

1. Which type of the residential property is most popular based on your experience? What are the common attribute characteristics of them?
2. In your experience, what indicators do real estate buyers consider when buying a home? What indicators have been considered in the past few years (around 2010) and are there any differences between past and current?
3. What are the main steps/procedures when pricing real estate? What method is used?
4. What impact factors are considered when pricing? What is the weight of the different factors?
5.
6. Do you think that the super-high residential buildings in the same area will be more popular because of the openness of the height of the building?
7. Will you show your customers a 3D model when you are selling the property? What is the customer's evaluation of this display method?
8. Do you think that developing online 3D model display can better improve performance?
9. What do you think is the impact of the expansion of Xi'an city on residential prices?

Questions for the teacher in Chang'an University:

1. Which indicators stimulate the price of property in Xi'an increase rapidly?
2. Has the 3D information been used in land valuation in Xi'an?
3. How do you like the expectation of 3D information used in land valuation? Why?

4. Is it necessary to use 3D data in land valuation?
5. Is there any difficulty in data acquisition or administration when using 3D data or building a 3D system?

Questions for the focus group:

1. What do you think of the phenomenon that property price in Xi'an increase rapidly?
2. Please list the the three most important qualities of the tall building (the house itself + the external environment).
3. What do you least like about the characteristics of a residential property?
4. As time goes by, is there any change of your preferences of residential property?
5. What indicators do you think cause the different prices for different floors? What are the factors about height that you consider about in the process of buying a house? (example, floor, orientation, landscape, view)
6. What type of landscape view do you like best/ like least?
7. (Show pictures of the south-facing balconies on different floors), How much are you willing to pay for the buildings of different heights?
8. (Show pictures of different landscapes such as green spaces, streets, buildings, etc.) How much are you willing to pay for these landscapes?
9. Have you experienced 3D technology during your purchase a property, such as VR viewing? Do you think the existing sand table, model room, and display area are enough for you to understand the whole house? Do you need a 3D model?
10. What indicators do you think stimulate the property price of Xi'an rapid increase? What are your expectations for the future trend of Xi'an housing prices?
11. Have you ever encountered a problem that did not match the developer description after the purchase of the residential property? (Examples, supporting facilities are not perfect, unreasonable design of the apartment)
12. High-rise buildings, small high-rise buildings, bungalows, villas, which one do you prefer? why?
13. Does Xi'an's planning have an impact on your choice of purchase a property? If so, what kind of planning will affect? (Such as subway, medical, education) If you choose a property in the future, which administrative district do you prefer? why?

Rules used in CityEngine

Rules to show property values and building height

```
version "2018.1"

attr height=10

Pink="#FF69B4"
Texture="assets/Facades/building.jpg"
attr price=10000
attr RoofStyle="RoofGable"
Lot -->

  extrude(height)
  //s('3,'3,'3)

    comp(f) { side :  Set_Color Facade.|top : roof }

roof-->
    roofGable(30,0.5,1)
   Set_Color


 Set_Color-->
 case price <3000 : color ("#2892C7")
 case price >=3000 && price <4000: color("#479BBF")
 case price >=4000 && price <4500: color("#60A9B5")
 case price >=4500 && price <5000:color("#60A9B5")
 case price >=5000 && price <5500:color("#8CB8A4")
 case price >=5500 && price <6000:color("#A0C29B")
 case price >=6000 && price <6500:color("#AECC91")
 case price >=6500 && price <7000:color("#C6D986")
 case price >=7000 && price <7500:color("#D7E37D")
 case price >=7500 && price <8000:color("#D5ED6F")
 case price >=8000 && price <8500:color("#FAFA64")
 case price >=8500 && price <9000:color("#FCE75B")
 case price >=9000 && price <9500:color("#FCCF51")
 case price >=9500 && price <10000:color("#FCBA47")
 case price >=10000 && price <11000:color("#FCA43F")
 case price >=11000 && price <12000:color("#FA8D34")
 case price >=12000 && price <13000:color("#F77A2D")
 case price >=13000 && price <14000:color("#F56325")
 case price >=14000 && price <15000:color("#F24D1F")
 case price >=15000 && price <17500:color("#ED361A")
 else : color("#E81014")
```