World Scientific Research Journal Volume 8 Issue 5, 2022
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

The Spatial Distribution Characteristics and Influencing Factors of Housing Prices in Qingdao

Feiyao Gao

School of Information Engineering, China University of Geosciences (Beijing), Beijing, 100000, China

Abstract

Housing price has always been a hot issue of people's livelihood. This study collected the relevant data of Fangtianxia, Lianjia, AutoNavi Map, Baidu Map and other websites to obtain the average price data of second-hand housing transactions of various residential buildings in Qingdao, selected as a sample point, and used The four interpolation methods of ordinary kriging (OK), simple kriging (SK), universal kriging (UK) and disjunctive kriging (DK) have obtained the continuous distribution layer of housing prices in Qingdao, using cross To verify the method, the optimal interpolation model is selected as ordinary Kriging second-order interpolation, and the spatial distribution characteristics of housing prices in Qingdao are analyzed through the optimal interpolation results. The research shows that the distribution of second-hand housing prices in Qingdao has obvious spatial gradient characteristics. The above performance is that it gradually decreases from the center of Shinan and Shibei districts to the peripheral districts, and explores the correlation between the spatial distribution characteristics of housing prices and the distribution of transportation, hospitals, education, and entertainment facilities. The conclusion is that the density of facilities is related to housing prices. It is not absolute, but it shows regularity to a certain extent.

Keywords

Housing Price; Kriging Interpolation; Cross-Validation; Spatial Distribution Characteristics.

1. INTRODUCTION

1.1. Question and Research Significance

Spatial analysis is one of the important parts of GIS, and interpolation is a more commonly used method in spatial analysis, covering a wide range. Combined with real life, due to the limitation of work cost and the difficulty of measurement work, it is impossible to measure the value of every location in the study area. Therefore, reasonable selection of sampling points can be considered, through the measured values of the sampling points, after analyzing the data and using an appropriate mathematical model, all positions in the study area can be predicted to form a surface of measured values to achieve the ultimate goal.

Selecting an adaptive interpolation method under different regional conditions has strong feasibility and practicability to summarize and summarize the spatial distribution characteristics of the selected research area.

1.1.1 Statement of Question

Housing has always been a rigid demand for residents, and urban housing prices have always been a hot topic of social concern. In recent years, the soaring housing prices in cities in my country have led to the emergence of many urban social problems. Among them, second-hand

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

houses account for a high proportion of urban housing, and most second-hand houses are located in superior geographical locations. The research on the factors affecting urban housing prices seems to be very important. important[1]. However, the composition of the urban system is complex and diverse, and the spatial distribution of the housing prices of second-hand houses is also different. Due to the influence of multiple factors, it is necessary to combine data and methods to explore the spatial distribution and influencing factors.

1.1.2 Significance of Research

Since the 21st century, with the comprehensive promotion of China's housing marketization, urban housing prices have become a common topic of concern in different industries and disciplines[2]. With the rising housing prices, some new features and problems have emerged in the urban spatial structure. Combined with GIS, the spatial distribution and influencing factors of second-hand housing prices are studied through the interpolation method in spatial analysis, which is of great significance to real estate development and solving the rigid needs of residents.

2. OVERVIEW OF THE STUDY AREA AND DATA ACQUISITION

2.1. The Overview of the Study Area

According to the research, Qingdao ranks first in the ranking of the most livable cities in China. Qingdao is located at the southern end of the Shandong Peninsula (35°35'-37°09'N, 119°30'-121°00'E). Qingdao City is located in the southeast of Shandong Peninsula, bordering the Yellow Sea in the east and south, adjacent to Yantai City in the northeast, connected with Weifang City in the west, and bordered by Rizhao City in the southwest. The area is 758.16 square kilometers, with a total area of 11,293 square kilometers.

2.2. Data Source and Preprocessing

The house price data used in this research comes from Fangtianxia[3] and Lianjia[4]. Python (crawler technology) is used to crawl the relevant information of a total of 2,621 communities in 10 districts in Qingdao, including the area where the communities are located, specific addresses and coordinates, as well as July 2020. The average unit area price of second-hand housing transactions from June to June 2021, and the specific spatial distribution of the community is shown in Figure 1.

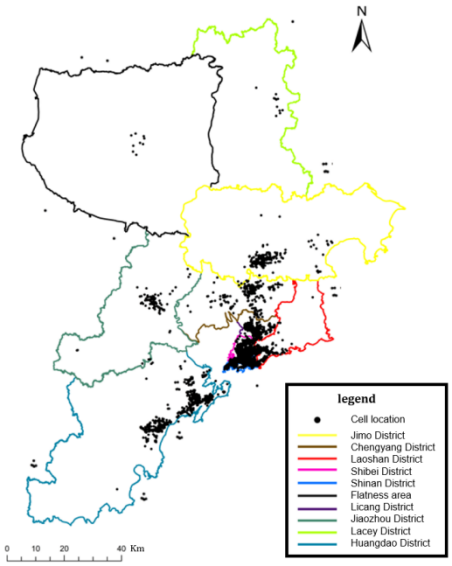

Figure 1. Spatial distribution of selected communities in Qingdao

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

2.3. Data Exploratory Analysis

After obtaining the data of all sample points, preliminary exploratory analysis of the data is required to test the distribution of the data and observe the distribution trend of the data. Use the data exploration and analysis tools provided by the ArcMap 10.8 Geostatistical Wizard to examine the analysis data.

2.3.1 Test of Data Distribution

Comparing the observatory plot and the normal QQ plot of the data after the Log change, it is found that the changed data satisfies the normal distribution.[5]

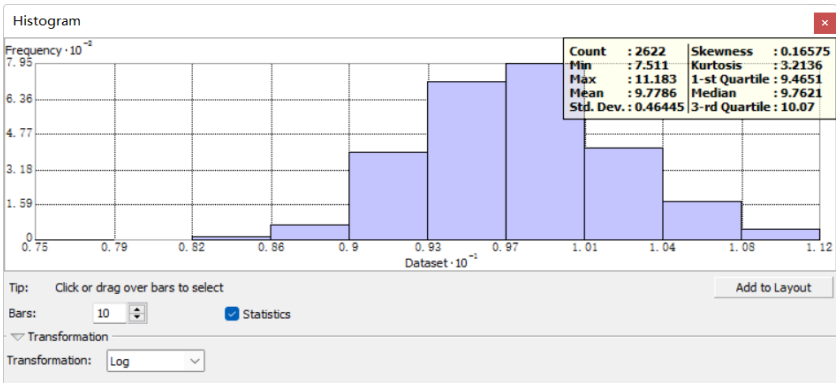

Figure 2. Histogram of house price data (after log change)

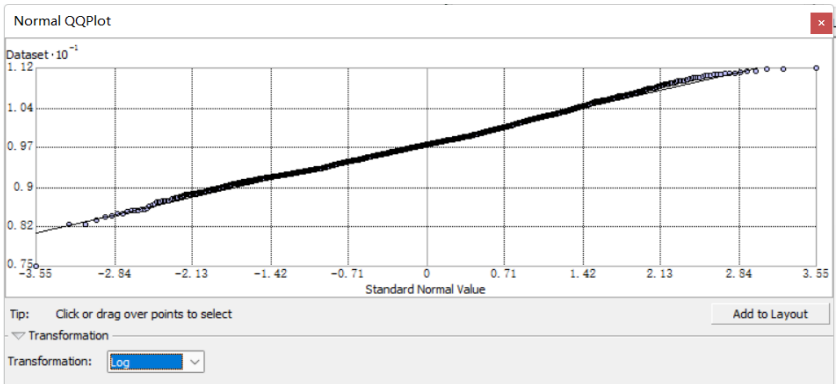

Figure 3. Normal QQ plot of house price data (after log change)

2.3.2 Trend Analysis of Data

It can be seen that when rotating these points, the trend always presents an inverted "U" shape, so it is considered that the data exhibits a second-order distribution in X and a slight second-order distribution in the Y direction.

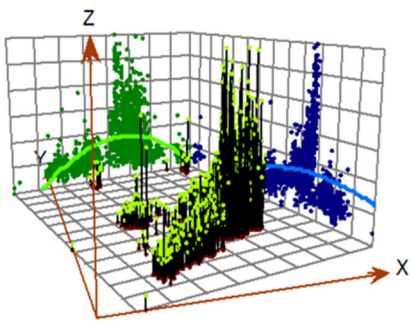

Figure 4. Community house price data trend chart

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

3. RESEARCH METHOD

3.1. Method of Interpolation

The interpolation model is essentially a mathematical smoothing method. In order to take into account the law of the spatial evolution of housing prices, it is necessary to select the most appropriate spatial interpolation method. By analyzing the characteristics of the data, from the perspective of overall interpolation, the analysis of the trend gradient characteristics is carried out. The degree of approximation is not high, and it is difficult to make accurate contour lines. So choose 4 different kriging interpolation methods to get the result.

Kriging interpolation, also known as spatial local estimation or spatial local interpolation, is a linear and unbiased optimal estimation of the values of regional variables in a limited area based on the theory of variograms and structural analysis. a method.

3.2. Accuracy Check

This paper uses cross-validation to determine the effect of interpolation, that is, using all data to estimate the trend and autocorrelation model, it removes one data location at a time, and then uses data from other locations to predict the data associated with it, and the predicted value is compared with The mean error and root mean square error were obtained by comparing the measured values[6].

In this study, four indicators including root mean square error (RMSE), standard mean error (MSE), standard root mean square error (RMSSE) and average standard error (ASE) were selected as verification indicators. The overall verification and verification criteria are as follows: ①The closer the standard root mean square error is to 1, the better; ②the smaller the root mean square error, the better, and the closer the standard mean error is to the root mean square error, the better (approximately equivalent to |RMS-ASE| The smaller the value, the better); ③ The closer the standard mean error is to 0, the better.

The above indicators and relevant verification criteria are selected to verify and compare the horizontal and vertical accuracy of the interpolation results, and the best interpolation method is obtained.

4. COMPARISON OF DIFFERENT INTERPOLATION METHODS

4.1. Determine the Parameters

In order to obtain the best interpolation model, four Kriging interpolation methods and three variograms are selected for interpolation comparison, and the default step size is set to 12, and the calculation is performed without considering anisotropy.

4.2. Analysis of Ordinary Kriging Interpolation Results

Using the simple ordinary kriging model for interpolation, and observing Table 2, it can be observed that the effect of the second-order trend surface is significantly better than that of the first-order trend surface. And when the function is a Gaussian function and a spherical function, the cross-validation index is better, and the validation criteria are met. All horizontal comparisons, in the ordinary kriging interpolation method, when the function is a Gaussian function and a spherical function, the effect is the best.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

Table 1. Precision Analysis of Ordinary Kriging Results

Method	Interpolation model	MSE	$RMSSE$	$\lvert RMS - ASE \rvert$
Ordinary Kerkin Interpolation	Gaussian first order	0.0250	0.9020	765.21
	Gaussian second order	0.0096	0.9920	135.00
	Spherical first order	0.0420	0.8233	943.02
	Spherical second order	0.0060	0.9945	143.18
	Exponential first order	0.0247	0.9036	728.96
	Exponential second order	0.0127	0.9572	454.85

4.3. Analysis of Simple Kriging Interpolation Results

The simple kriging model is used for interpolation. Observe Table 3. It can be observed that the effect of the same second-order trend surface is significantly better than that of the first-order trend surface. When the function is an exponential function, the cross-validation index is better, it can meet the verification criteria, and the horizontal comparison is the best when the function is an exponential function in the simple kriging interpolation method.

Table 2. Precision Analysis of Simple Kriging Results

Method	Interpolation model	MSE	$RMSSE$	$\lvert RMS - ASE \rvert$
Simple Kerkin Interpolation	Gaussian first order	0.0250	0.9020	765.26
	Gaussian second order	0.0355	0.9531	439.71
	Spherical first order	0.0690	0.7737	1170.58
	Spherical second order	0.0318	0.9567	437.261
	Exponential first order	0.0658	0.7867	1166.28
	Exponential second order	0.0214	0.9831	289.46

4.4. Analysis of Disjunctive Kriging Interpolation Results

The disjunctive kriging model is used for interpolation. Observe Table 4. It can be observed that the effect of the second-order trend surface is significantly better than that of the first-order trend surface. Among them, when the function is an exponential function, the cross-validation index is better, and it can meet the verification criteria. Horizontal comparison, the disjunctive kriging interpolation method has the best effect when the function is an exponential function.

Table 3. Precision Analysis of Disjunctive Kriging Results

Method	Interpolation model	MSE	$RMSSE$	$\lvert RMS - ASE \rvert$
Disjunctive Kerkin Interpolation	Gaussian first order	0.0524	0.7277	1625.22
	Gaussian second order	0.0127	0.9302	1242.77
	Spherical first order	0.0448	0.7585	1947.34
	Spherical second order	0.0005	0.9616	1402.50
	Exponential first order	0.0367	0.7900	1880.88
	Exponential second order	0.0065	0.9905	1228.03

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

4.5. Analysis of Universal Kriging Interpolation Results

Observing the accuracy index record table of universal kriging interpolation results (Table 5), it is found that the interpolation method has a large error. When the sample data is housing prices, the housing prices are distributed in a "clustered" shape in some areas, so this method is not applicable. Choose to discard this method.

Table 4. Precision Analysis of Universal Kriging Results

Method	Interpolation model	MSE	$RMSSE$	$\lvert RMS - ASE \rvert$
Universal Kerkin Interpolation	Gaussian first order	0.0700	0.7730	764.05
	Gaussian second order	0.0355	0.9531	439.71
	Spherical first order	0.0700	0.7737	770.58
	Spherical second order	0.0318	0.9567	434.23
	Exponential first order	0.0658	0.7867	669.27
	Exponential second order	0.0214	0.9831	289.39

4.6. Optimal Interpolation Model Selection

Through the horizontal analysis of the above four different interpolation methods, the preliminary interpolation effect is obtained as follows: ordinary kriging > simple kriging > disjunctive kriging > universal kriging[6].

Secondary comparisons were performed to test the accuracy longitudinally. By comparing the ordinary kriging Gaussian second-order, ordinary kriging spherical second-order, simple kriging index second-order, and disjunctive kriging index second-order four methods, compare the accuracy test index RMSSE, MSE, and sort the accuracy It is: ordinary kriging Gaussian second order > ordinary kriging sphere second order > disjunctive kriging index second order > simple kriging index second order[7]. The error of the ordinary Kriging Gaussian second order is small in all aspects, which is more in line with the actual situation, so it is used as the best interpolation model for this study.

5. APPLICATION OF OPTIMAL INTERPOLATION RESULTS

5.1. Spatial Distribution Characteristics of Housing Prices in Qingdao

After obtaining the optimal interpolation result model, the spatial distribution characteristics of second-hand housing prices in Qingdao are further analyzed and summarized through the interpolation results[8].

By observing the housing price trend in Qingdao from December 2020 to December 2021 (Figure 5), combined with the best interpolation results (Figure 6) showing the distribution of second-hand housing prices in Qingdao, the spatial characteristics can be summarized as: The housing prices in the administrative districts vary greatly and are hierarchical. In space, they show obvious spatial agglomeration and spatial gradient characteristics. Generally, they gradually decrease from the center of Huangpu District to the peripheral districts. The overall distribution of second-hand housing prices in Qingdao is decreasing from the center to the periphery. The areas with the highest house prices are distributed in the southeast direction in the city center, and there are a few peripheral high-value areas in the suburbs. There is a strong correlation between the spatial distribution of housing prices and the circle structure of urban space.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

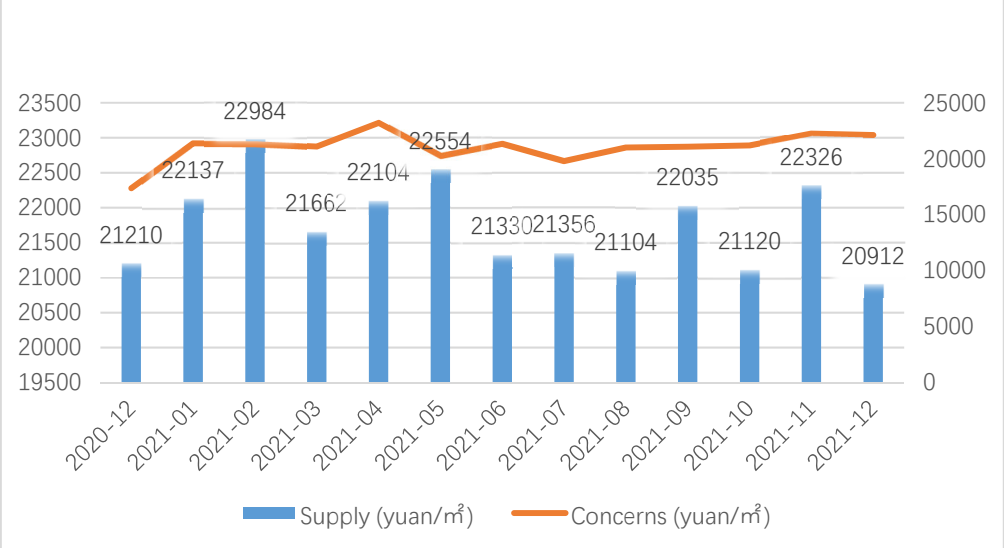

Figure 5. Qingdao residential house price details in the past year

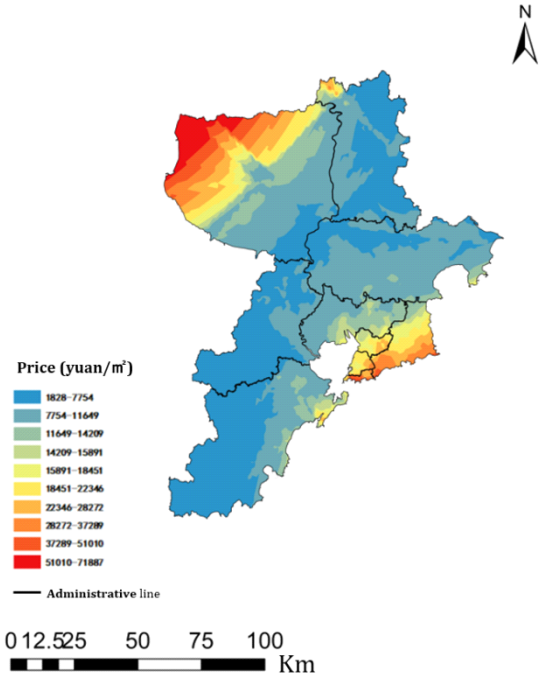

Figure 6. Ordinary Kriging Gaussian Second-Order Interpolation Results

For further analysis, it is embodied in the following three points:

(1) The price span is large, and there is no linear relationship between the number of houses and the price.

The highest value of second-hand housing prices in the central urban area of Nanning is nearly 10 times the lowest value, and the growth trend of the number of second-hand houses at different price levels is also different. After analysis, for example, in the "Huangqi Fugui Lily Garden" community in Jiaozhou City, when the data was crawled, the number of houses was large, and the housing prices were still low. The minimum value was 1849 yuan/㎡.

(2) The price gradually decreases outward from a certain point, and the regional changes are obvious.

The second-hand housing in the central urban area of Qingdao has formed from Shinan District and Shibei District as high-priced agglomeration centers, and Shinan District, Qingdao,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

a more prosperous area, both in supporting and facilities are relatively mature, so the housing price is stable and at a high price.

The emerging city of Pingdu has also developed rapidly in recent years. With the policies issued by the government departments, shopping malls and entertainment facilities have continued to enter, making its housing prices also high. Most people choose to buy second-hand houses in this area.

(3) Some remote areas still have high housing prices.

5.2. Factors Affecting the Spatial Distribution of Housing Prices in Qingdao

Based on the results of interpolation analysis, the spatial characteristics of housing price distribution in Qingdao are preliminarily summarized. In order to further verify whether the spatial characteristics are accurate, we will continue to explore the factors affecting their distribution through the above results: transportation, education, hospitals, and entertainment facilities [9].

By capturing the POI data of Baidu and AutoNavi, the POI data of the distribution of bus stops and expressways, education (schools), hospitals and entertainment facilities in Qingdao are obtained. And put all kinds of facilities on the layer after house price interpolation for correlation analysis.

5.2.1 Transportation

Superimpose the house price distribution map of second-hand housing in Qingdao with the map of expressways and bus stations in Qingdao to observe the correlation between the spatial distribution of housing prices and subway lines. There is a strong correlation between the spatial distribution of housing prices in Qingdao and bus stations, and the housing prices in the surrounding areas of the bus stations are significantly higher than those in areas without bus stops. In the peripheral high-value areas of housing prices, such as Pingdu and the Laixi area of Jimo, there are expressways passing through, which also shows to a certain extent that the convenient transportation can promote the rise of surrounding housing prices.

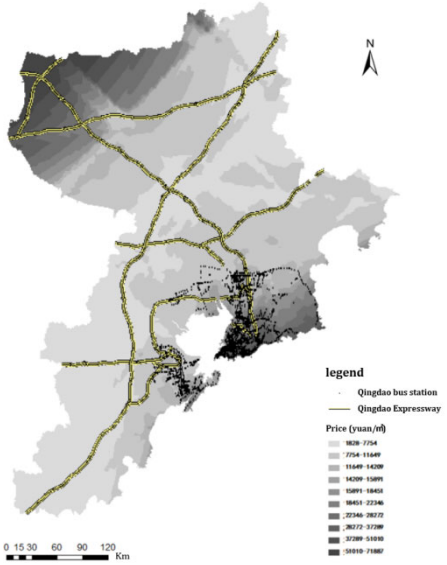

Figure 7. Interpolation results of transportation POI and Qingdao housing prices

5.2.2 Education

After superimposing the housing price distribution map and the point layer of education school distribution, it is analyzed that the spatial distribution of housing prices has a certain correlation with the spatial distribution of primary and secondary schools and kindergartens.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

Due to the high residential density in the main urban area, the distribution density of educational schools is also high. In the suburbs, there is no obvious correlation between the distribution of education and the spatial distribution of housing prices.

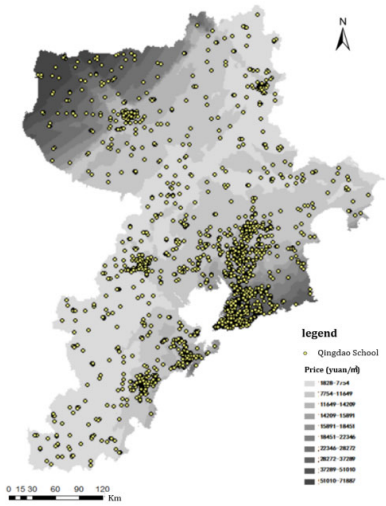

Figure 8. Interpolation results of education POI and Qingdao housing prices

5.2.3 Hospitals

The housing price distribution map is superimposed with the point layer of medical facilities distribution, and it is analyzed that the spatial distribution of housing prices has a certain correlation with the spatial distribution of hospitals and pharmacies. Due to the high residential density in the main urban area, the distribution density of hospitals is also high. In the suburbs, most residents in the suburban areas with high housing prices live in the suburbs for a quieter and more natural living environment and more spacious and comfortable living space, and they believe that such residents have the ability to obtain education, medical services and other services in the city center. Therefore, this group of people does not have high requirements for the facilities around the house, so the distribution of facilities in the suburbs has little impact on the spatial distribution of housing prices.

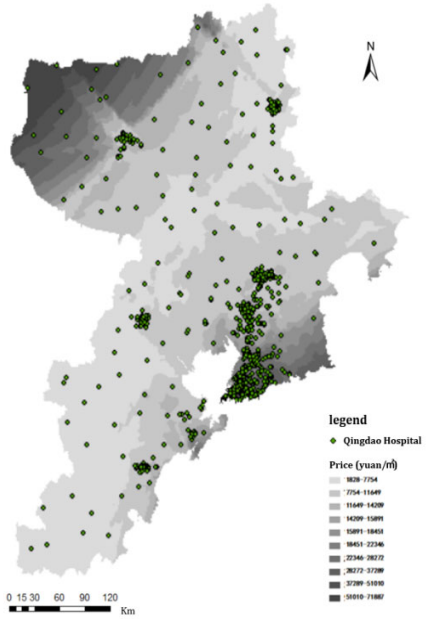

Figure 9. Interpolation results of hospital POI and Qingdao housing prices

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

5.2.4 Recreational Facilities.

The house price distribution map is superimposed with the point layer of the distribution of entertainment facilities (supermarkets, shopping malls, restaurants), and it is analyzed that the spatial distribution of house prices has a certain correlation with the distribution of the point group. It was concluded that the more densely populated the place, the higher the house price.

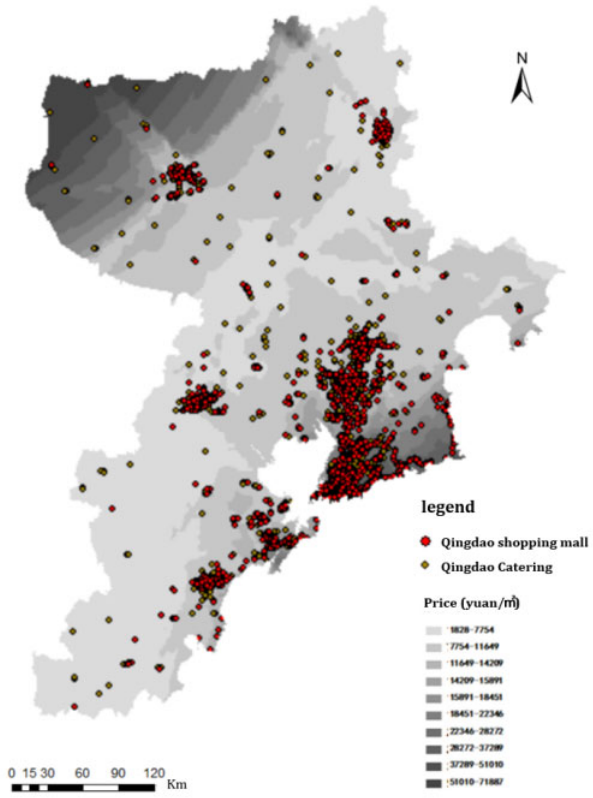

Figure 10. Interpolation results of recreational facilities POI and Qingdao housing prices

6. DISCUSSION AND SUMMARY

By collecting relevant data from Fangtianxia, Lianjia, AutoNavi Map, Baidu Map and other websites, this study obtained the average price of second-hand housing transactions in various residential quarters in Qingdao. After obtaining most of the POI data of the whole city, the spatial distribution characteristics of housing prices of second-hand residential houses in Qingdao are further analyzed, and the possible reasons for this situation and possible influencing factors are explored. It is verified that the optimal choice of the interpolation method is correct through various analyses.

ArcGIS is used to analyze the price difference of second-hand housing in different administrative districts in Qingdao, and then Kriging interpolation method is used to generate the continuous distribution layer of housing prices in Qingdao, and the specific characteristics of the spatial distribution of housing prices in Qingdao are obtained. Finally, the housing price distribution layer and the facility distribution layer are superimposed to analyze the correlation between the spatial distribution of housing prices and the spatial distribution of facilities, and to explore the possible factors that affect the spatial distribution of housing prices.

In the process of using different kriging interpolation methods to generate the continuous distribution layer of house prices in Qingdao, four different kriging interpolation methods were selected and divided into 24 sub-categories. Four interpolation methods of ordinary kriging,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

simple kriging, disjunctive kriging and universal kriging are used for the housing prices of Qingdao. Four indicators of root mean square error (RMSE), standard mean error (MSE), standard root mean square error (RMSSE) and average standard error (ASE) were selected as verification indicators. However, the research object of this research is the housing price of urban housing. Since the sampling points of housing prices are not as uniform as the traditional sampling points, a "cluster-like" distribution will occur in the case of high probability. "The smaller the root mean square error, the better, and the closer the standard mean error is to the root mean square error, the better" is changed to be approximately equivalent to "The smaller the value of |RMS-ASE|, the better." The final result obtained is OK>SK>DK>UK.

There are various types of spatial interpolation methods, and different research methods are used in the same area, resulting in different results. The research shows that the housing prices in the various administrative districts of Qingdao are quite different and hierarchical, and show obvious spatial agglomeration and spatial gradient characteristics in space. The overall performance is that it gradually decreases from the center of Shinan and Shibei to the peripheral districts. The areas with the highest housing prices are distributed in the southeast direction in the city center, and there are still a few peripheral high-value areas in the suburbs. There is a strong correlation between the spatial distribution of housing prices and the circle structure of urban space.

The housing prices in the various administrative districts of Qingdao are quite different and hierarchical. In terms of space, they show obvious spatial agglomeration and spatial gradient characteristics, which generally show a gradual decrease from the center of Shinan District to the peripheral districts. Therefore, the spatial distribution of housing prices has a certain correlation with the spatial distribution of educational, medical, entertainment and other facilities, but the correlation between facility density and housing prices is not strong.

The shortcomings of this study: ① This study only considers the price of second-hand housing, and the interpolation results are different from the real situation. ② This study can only crawl recent housing price data, and the results are not perfect in the time dimension.

The strength of this study is that it is relevant to real life, because housing prices have always been a concern, so this research has far-reaching significance. Different regions adopt different methods to obtain the spatial distribution of regional housing prices. This research method is universal, and based on this conclusion, it is expected that further in-depth research will obtain more accurate results.

REFERENCES

[1] S. Li, X. Ye, J. Lee, et al. Spatiotemporal Analysis of Housing Prices in China: a Big Data Perspective, Applied Spatial Analysis & Policy, Vol. (2017) No.10, p.421-433.

[2] Z.S. Yang, Z.L. Su, Y. Ding, et al. The External Effect of Jiaoyufication upon Urban Space, Human Geography, Vol.33 (2018) No.4, p.60-67. (In Chinese)

[3] Information on: https://www1.fang.com

[4] Information on: https://www.lianjia.com

[5] J.Y. Li, C.K. Li, Z.H. Yin: ArcGIS Based Kriging Interpolation Method and Its Application, Bulletin of Surveying and Mapping, Vol. (2013) No.9, p.87-97. (In Chinese)

[6] D.Zhang, B.Z. Jiang, G.H. Liu, et al. Applicability of Spatial Interpolation Methods to Predict Total Phosphorus in the Typical Irrigated Areas of the Sanjiang Plain, Chinese Journal of Eco-Agriculture, Vol.29 (2021) No.8, p.1405-1416. (In Chinese)

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202205_8(5).0058

[7] B.X. Zhao, L. Wang, D.Y. Cheng: Comparison of Spatial Interpolation Method for Meteorological Data and Distribution Characteristic in Anhui Province, Research of Soil and Water Conservation, Vol.24 (2017) No.3, p.141-145. (In Chinese)

[8] G.S. Bhunia, P.T. Shit, R. Maiti: Comparison of GIS-based Interpolation Methods for Spatial Distribution of Soil Organic Carbon (SOC), Journal of the Saudi Society of Agricultural Sciences, Vol.17(2018) No.2, p.114-126.

[9] D. Stadelmann: Which factors capitalize into house prices? A Bayesian averaging approach, Journal of Housing Economics, Vol.19 (2010) No.3, p,180-204.