

GIS-based Analysis of Distribution Characteristics and Driving Factors of House Prices in Nanjing's Central District

Xiaolin Gong *

School of Public Administration, Central China Normal University, Wuhan, China

* Corresponding author: Xiaolin Gong (Email: athenagxl@163.com)

Abstract: This study aims to investigate the spatial distribution characteristics of house prices and their driving factors in the central area of Nanjing by using Geographic Information System (GIS) technology, combined with kernel density analysis, Moran index and geodetic detector method. The results show that: (1) The high values of house prices in central Nanjing have significant spatial aggregation characteristics, mainly concentrated in Qinhuai District, Gulou District and its border areas, and the house prices have significant spatial autocorrelation; (2) The results of the geo-detector show that the degree of traffic access and settlement characteristics have the greatest influence on house prices, while the influence of the construction of ancillary facilities is relatively weaker; (3) There are significant interactions between different factors, especially between transport accessibility and the construction of ancillary facilities, which play a crucial role in influencing house prices.

Keywords: House price, Spatial differentiation, Geo-detector, Main city of Nanjing.

1. Introduction

As urbanization continues to accelerate, the issue of housing prices has become one of the hot issues of social concern. The issue of housing prices involves thousands of families, living with a home is one of the most basic people's livelihoods, the 'CPC Central Committee on the formulation of the national economic and social development of the 14th Five-Year Plan and the 2035 Perspective Goals of the proposal' clearly pointed out that we should promote the healthy development of housing consumption. Secondly, the stability and healthy development of real estate market prices is crucial to the national economy. Based on data published by the National Bureau of Statistics, residential real estate prices rose by 6.8% in 2023 compared to 2022, and although the rate of increase was reduced, the growth trend remained unchanged, and residential real estate prices in the first-tier cities rose by a large margin; the data also showed that real estate accounted for 5.8% of the entire China's GDP in 2023, and although this The data also suggests that real estate will account for 5.8% of China's GDP in 2023, and although this percentage has declined in recent years, it still has a substantial effect on the overall economy, suggesting that price volatility will not only affect the real estate sector, but also the financial markets and the economy as a whole.

As an important first-tier city in eastern China, the trend of house prices in Nanjing's central city not only affects the quality of life of its citizens, but also reflects the overall level of urban development. This study is based on geographic information system (GIS) technology to conduct an in-depth analysis of the spatial distribution characteristics of housing prices and their driving factors in the central city of Nanjing, which is of great significance in the following aspects: firstly, it can help us to understand the differences in housing prices and their trends in different regions, analyse the causes of the differences, optimise the allocation of resources, and enhance the sense of well-being of residents' lives; secondly, it can explore in depth the various factors that affect housing prices

and provide scientific basis for the government to formulate policies and developers to make business decisions; thirdly, through the study of the distribution of housing prices and its driving factors, it can help the government to better manage the real estate market, formulate reasonable land use and property regulation policies, and enhance the government's governance ability.

Through reading the existing literature, it is found that most of the studies on real estate prices focus on socio-economic aspects, mostly discussing indicators such as population flow, land supply, economic development level, changes in government policies, and financial conditions [1-3]; most analyses of the current status of research on real estate prices start from the global, national, regional, or city-wide macro-level [2,4-7], and the characteristics of spatial differentiation of house prices within cities are analysed quantitatively. In general, existing literature studies on specific range micro-regions in specific cities are not common, and the studies mainly reflect the pattern of spatial differentiation of house prices from an economic perspective, lacking the public participation attribute of the main body of house buying.

In terms of research methodology, there have been few studies on real estate prices over the years, and their methods mostly use quantitative models to analyse the relationship between house prices and their influencing factors [8-10], mostly based on the Ordinary Least Squares (OLS) with the assumption of spatial mean, and the Geographically Weighted Regression (GWR) model based on spatial heterogeneity; however, the former does not take enough account of the spatial heterogeneity of house prices, and the latter takes account of the spatial heterogeneity but ignores the spatial influence scale difference of different variables, in this regard, some scholars also optimize the above problems and use MGWR model to conduct research [11]; in recent years, machine learning technology has been widely used in real estate research due to its unique advantage in processing massive data, and it has become a popular research method

[12-15]; based on this, many scholars at home and abroad have used machine learning methods to study factors such as traffic accessibility [16, 17], urban environment [18], and infrastructure support [19, 20], and to analyse their driving effects on housing prices. Analysing and summarising the above literature, it is found that geo-detector is less frequently used in the field of real estate price research.

2. Data Sources and Research Methodology

2.1. Data Sources

The research data of this paper mainly includes 3 parts: First, based on the second-hand housing source information of Shell Housing (www.ke.com), a property information website in Nanjing, the total supply got 2974 housing source information, and other attribute data of the housing source were also obtained, including coordinates, streets, floors, etc.; second, through the Gaode open platform and the data crawling plug-in, the relevant point-of-interest data within the research scope was collected and filtered; Third, the population raster dataset with 100m resolution from China's seventh population census (2020) shared by Prof Chen Yuehong's team on the figshare platform.

2.2. Research Methods

2.2.1. Analysis of spatial distribution characteristics

In this study, Kernel Density Analysis (KDA) is used to estimate the spatial density distribution of house prices in central Nanjing. By expanding each house price data point into a kernel function, a smooth density map is generated to show the spatial distribution of house prices. This approach can help identify hotspots in the distribution of second-hand houses for sale, assess the spatial distribution characteristics of house prices, as well as discover potential spatial patterns.

The model equation is:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

Where $\hat{f}(x)$ is the kernel density estimate at x ;

2.2.2. Spatial autocorrelation analysis

Spatial autocorrelation analysis primarily involves global and local spatial autocorrelation, which is quantified using Moran's I. The model equations are as follows:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} Z_i Z_j}{S_0 \sum_{i=1}^n Z_i^2}$$

Where I is the Moran index, n is the total number of elements, W_{ij} is the spatial weight between elements i and j , and S_0 is the set of all spatial weights.

2.2.3. Geographical Detector

In this paper, we will refer to the Geographical Detector (GD) model of Wang Jinfeng [21] et al. to explore the spatial relationship and the mechanism of action between the distribution of house prices and its driving factors in the central city of Nanjing. The driving factors include accessibility, construction of supporting facilities and characteristics of the settlement. By dividing the study area into sub-areas, the geoprobes will measure the extent to which

various factors explain the spatial distribution of house prices. The geodetector model includes a factor detector, an interaction detector, a risk detector, and an ecological detector:

Factor detector: detects the impact of a single factor (e.g., housing conditions, traffic conditions, infrastructure support, etc.) on the spatial distribution of house prices, measured by the q-value, with the following model formula:

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} = 1 - \frac{SSW}{SST}$$

$$SSW = \sum_{h=1}^L N_h \sigma_h^2, \quad SST = N \sigma^2$$

Interaction detector: To detect the effect of the interaction between different factors on the house price distribution, the evaluation is done by firstly calculating the q-values of the two factors X_1 and X_2 on Y : $q(X_1)$ and $q(X_2)$, respectively, and calculating the q-value of their interaction (the new polygonal distribution formed by the tangent of the two layers of the superimposed variables X_1 and X_2): $q(X_1 \cap X_2)$. Comparing $q(X_1)$, $q(X_2)$ with $q(X_1 \cap X_2)$.

Risk Detector: Identifies areas of high or low risk for higher or lower house prices and is tested using the t-statistic, modelled as follows:

$$t_{\bar{y}_{h=1} - \bar{y}_{h=2}} = \frac{\bar{Y}_{h=1} - \bar{Y}_{h=2}}{\left[\frac{Var(\bar{Y}_{h=1})}{n_{h=1}} + \frac{Var(\bar{Y}_{h=2})}{n_{h=2}} \right]^{1/2}}$$

Ecological detector: It is employed to determine whether there is a significant difference between the impacts of the two factors X_1 and X_2 on the spatial distribution of attribute Y , which is measured by the F-statistic, and the model formula is as follows:

$$F = \frac{N_{X_1}(N_{X_2} - 1)SSW_{X_1}}{N_{X_2}(N_{X_1} - 1)SSW_{X_2}}$$

$$SSW_{X_1} = \sum_{h=1}^{L1} N_h \sigma_h^2, \quad SSW_{X_2} = \sum_{h=1}^{L2} N_h \sigma_h^2$$

Through the geo-detector model, the spatial distribution characteristics and their interrelationships affecting the house prices in the central area of Nanjing are analysed in depth, offering a scientific foundation for the formulation of targeted house price regulation and urban planning strategies.

3. Characteristics of Spatial Differentiation of House Prices in The Main Urban Area of Nanjing City

Firstly, the Arcgis kernel density analysis tool is used to derive the spatial differentiation characteristics of second-hand houses for sale in the main urban area of Nanjing. According to the results of the analysis, it is found that the concentration areas of second-hand houses for sale in Nanjing are predominantly situated in the south-western part of the main city, namely Qinhuai District, Gulou District and their junction locations.

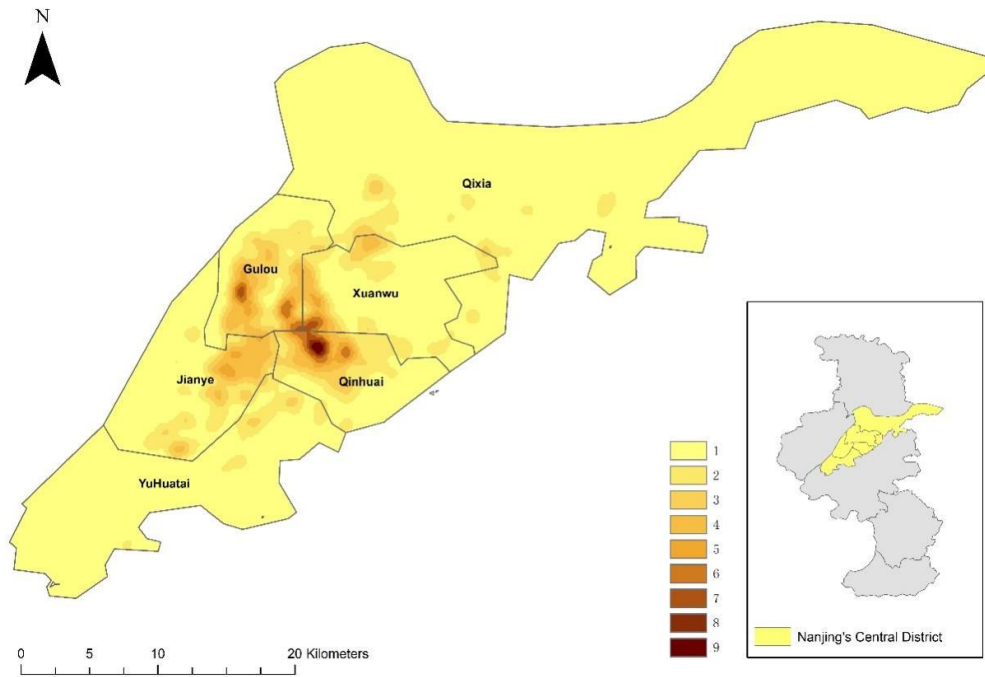


Figure 1. Kernel Density Analysis of House Prices in Nanjing's Central District

In order to study the clustering characteristics of house prices within the main urban area of Nanjing, this study intends to use Moran scatter plot and LISA clustering map to start the analysis. The spatial autocorrelation is analysed by using GEODA global Moran index, and the resultant I-value is 0.4669, which reveals that the house prices in the main urban area of Nanjing city show the clustering characteristics.

Subsequently, the local Moran index is used to analyse the local spatial autocorrelation type, and the clusters are first classified into four types: high - high, low - low, high - low, and low - high, and it is found that the high - high clusters of the house prices are mainly distributed in Gulou and Jianye districts as well as Xuanwu district, which are primarily located in the area bounded by Inner Ring South Road - Yingtian Avenue to the north, Inner Ring East Line to the west, and Beijing West Road to the south. This area features well-known commercial districts such as Xinjiekou, with comprehensive supporting facilities, convenient transportation, and is also the administrative center of Nanjing. Most of the low-high agglomerations are distributed in a ring shape along the Xuanwu Lake Park, which has the unique advantage of a better ecological environment. The low-high agglomerations are often found in the surroundings of the high-high agglomerations, which may be affected by factors such as the ageing of the houses themselves and the surrounding congested living environment, leading to lower house prices. Low-low aggregations are mainly located in Yuhuatai District, the northern part of Xuanwu District, and Qinhuai District, which are generally less developed and are mostly surrounded by industrial and commercial land.

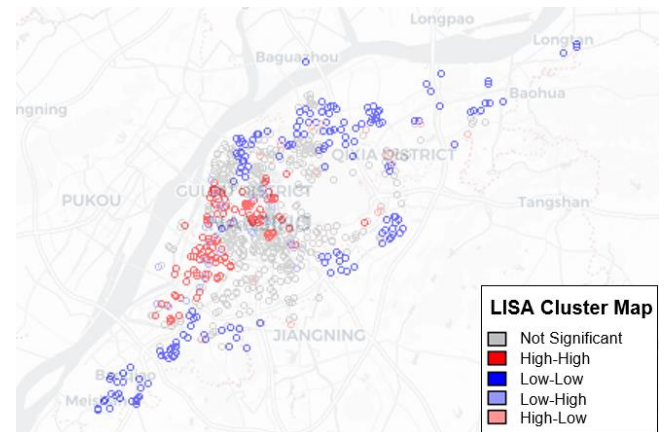


Figure 2. LISA Cluster Map

4. Factor Detection for Imaging Characteristics of House Price Divergence

In order to explore the driving factors of spatial variation of house prices in Nanjing, this study uses geodetectors to calculate the influence factor data, and obtains the results of factor detection, interaction detection, and ecological detection.

In this paper, the driving factors affecting the spatial pattern of house price divergence are classified into three categories, the first one is the degree of transport accessibility measured by the distance from metro stations, bus stops, and neighboring roads. The second category is the construction of ancillary facilities measured by the distance from neighbouring hospitals, educational institutions and shopping places. The third category is the characteristics of the neighbourhood measured by the population density of the street where it is located, the night-time lighting index (NLI) and the fractional vegetation cover (FVC).

Table 1. Impact factors

| | Drivers | Factors |
|----|---------------------------------------|-----------------------------------|
| X1 | Accessibility | Subway station |
| X2 | | Bus stops |
| X3 | | Road |
| X4 | Construction of supporting facilities | Shopping facilities |
| X5 | | Medical institutions |
| X6 | | Educational institutions |
| X7 | Characteristics of the settlement | Population density |
| X8 | | Fractional vegetation cover (FVC) |
| X9 | | Night light index (NLI) |

Based on the results of the factor probes, it is found that metro stations, bus stops, shopping places, hospitals, population, fractional vegetation cover, and nighttime lighting index are significant at the 5 per cent confidence level, with subway stations having a probability value (q) greater than 0.1 (0.113393), which has the greatest explanatory power for spatial divergence in house prices, followed by hospitals with a probability value close to 0.1 (0.092122), and shopping places and educational institutions have the weakest explanatory power for house prices. Overall, averaging the impact factors among the three categories of drivers reveals that the construction of neighbourhood amenities has the least impact on house price divergence, while the degree of transport access and settlement characteristics have a similar degree of impact on house price divergence, both being much higher than that of the impact brought by amenities.

Table 2. The results of factor detector

| Factors | p | q | Factors | p | q |
|---------|-------|--------|---------|-------|--------|
| X1 | 0.113 | 0.000* | X6 | 0.014 | 0.054 |
| X2 | 0.047 | 0.000* | X7 | 0.063 | 0.000* |
| X3 | 0.007 | 0.229 | X8 | 0.032 | 0.000* |
| X4 | 0.009 | 0.029* | X9 | 0.058 | 0.000* |
| X5 | 0.092 | 0.000* | | | |

* Indicates significant at 5% confidence level

Table 3. The results of ecological detector

| | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ | X ₆ | X ₇ | X ₈ | X ₉ |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| X ₁ | | | | | | | | | |
| X ₂ | N | | | | | | | | |
| X ₃ | Y | N | | | | | | | |
| X ₄ | Y | N | N | | | | | | |
| X ₅ | N | N | Y | Y | | | | | |
| X ₆ | Y | N | N | N | Y | | | | |
| X ₇ | N | N | N | N | N | N | | | |
| X ₈ | Y | N | N | N | N | N | N | | |
| X ₉ | N | N | N | N | N | N | N | N | |

Using an ecological detector to compare whether each factor has a significantly different impact on house prices, it is found that the difference in the impact of factors within

each type of driver on house prices is not large, while there is a certain difference in the force of the factor between the degree of transport access and the construction of ancillary facilities, but overall, the impact of each factor on house prices is not significant.

The interaction detector was utilized to identify the interactions among the drivers, and the results showed that the interaction force between any two of the nine selected influencing factors was greater than the influence of a single factor when it was in action. Among the drivers, X1∩X2, X4∩X5, X7∩X9 have the largest interaction power, 0.2185, 0.1538, 0.1053 respectively; among the different types of drivers, X1∩X5, X1∩X7, X5∩X7 have the largest interaction power, 0.2555, 0.1788, 0.1454 respectively. The maximum interaction force between accessibility and the construction of supporting facilities is greater than the maximum interaction force between the factors within itself. The factor interaction types include two-factor enhancement and non-linear enhancement, among which X1∩X9, X5∩X7, X5∩X9, X7∩X9, X8∩X9 show two-factor enhancement, and the rest show non-linear enhancement, which shows that the spatial pattern of house prices is the result of the joint action of different influencing factors.

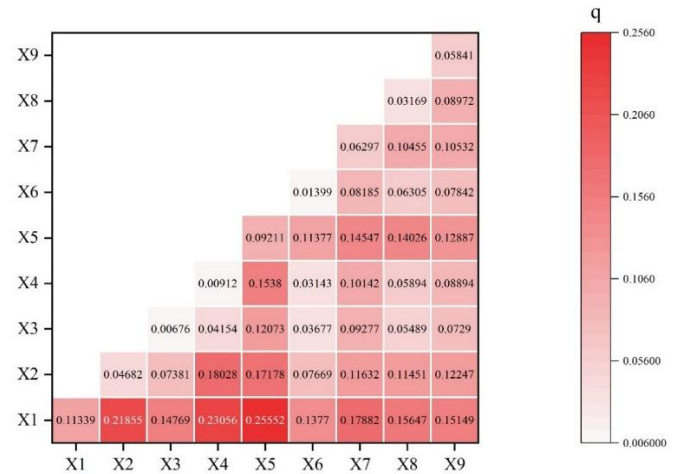


Figure 3. The results of interaction detector

5. Conclusion

This study analyses the spatial distribution pattern of house prices, as well as the global and local autocorrelation measures of house prices, using kernel density analysis and the Moran index, revealing the overall and local spatial correlation, aggregation and differentiation characteristics of house prices; and uses geo-detectors to identify spatial differentiation of house prices as well as the driving factors behind them. The main conclusions are as follows:

Firstly, the house prices in central Nanjing present obvious spatial aggregation characteristics, with the main aggregation areas located in Qinhuai District, Gulou District and their junction positions in the southwestern part of the main city; the global Moran's I index is 0.4669, this indicates that house prices exhibit significant spatial autocorrelation, with high-high clusters forming in Gulou District and Jianye District, and low-low clusters emerging in Yuhua District, northern Xuanwu District, and Qinhuai District.

Secondly, the results of the factor probes show that transport accessibility and settlement characteristics have a more substantial effect on house prices, and in particular the proximity of metro stations and bus stops has the strongest

explanatory power for the spatial pattern of house price divergence. On the contrary, the impact of the construction of supporting facilities on house prices is relatively weak.

Thirdly, the results of the interaction detector show that the interaction of different factors is obvious, and the interaction force between any two factors is greater than the influence of a single factor, especially there is a significant interaction between the degree of transport access and the construction of ancillary facilities; in addition, there is a complex relationship between the characteristics of the distribution of house prices and a variety of spatial factors, which indicates that the spatial distribution of house prices is the outcome of the combined influence of various factors.

References

- [1] Li Kai, Fan Mingtai, Ye Sihui Research on the dynamic transmission of monetary factors and macro influence of housing prices in China —— Analysis based on TVP-SV-VAR model [J / OL]. *Research on Financial Development*, 2021 (1): 29-37. DOI: 10.19647/j.cnki.37-1462/f.2021.01.004.
- [2] Zhang Long, Shen Yingqi China housing price stability measurement, influencing factor identification and monetary policy "cross-cycle" design [J]. *Financial Science*, 2023 (8): 1-16.
- [3] Li Jiao, the heterogeneous fluctuation of urban housing prices in China: driving factors and action path [J]. *Journal of Chongqing University of Technology (Social Sciences)*, 2023, 37 (7): 93-109.
- [4] Yu Xiaofen, Zhan Dongsheng, Analysis of the spatial pattern and influencing factors of housing price index in 70 large and medium-sized cities in China [J / OL]. *Journal of Central China Normal University (Humanities and Social Sciences Edition)*, 2022,61 (1): 40-51. DOI: 10.19992/j.cnki.1000-2456.2022.01.006.
- [5] Sun Biao, Yangshan, the characteristics and influencing factors of housing prices in big cities under the action of multi-centers —— Take Hefei city as an example [J]. *Resources and Environment of the Yangtze River Basin*, 2021, 30 (7): 1538-1546.
- [6] Xu Danmeng, Li Xin, Zhang Suwen, Research on the spatial differentiation pattern and its influencing factors in Shenyang [J / OL]. *Human Geography*, 2021, 36 (6): 125-134. DOI: 10.13959/j.issn.1003-2398.2021.06.014.
- [7] DUCA J V, MUELLBAUER J, MURPHY A. What drives house price cycles? International experience and policy issues [J]. *Journal of Economic Literature*, 2021, 59(3): 773-864.
- [8] Wu Chao, Liu Pengyu, Nie Ke The multi-scale spatial relationship between housing price and influencing factors in Nanjing [J]. *Modern Urban Research*, 2021 (4): 93-98.
- [9] Empirical evidence from 31 provinces and 70 large and medium-sized cities in China [J / OL]. *Social Science in Nanjing*, 2021 (6): 26-38. DOI: 10.15937/j.cnki.issn1001-8263.2021.06.004.
- [10] Tu Jin, Jiang Wanchen, Leng Zhengxing The differentiation study of the influencing factors of the urban housing prices in China —— Based on the analysis of the big data of the second-hand housing market in Chengdu city [J / OL]. *Price Theory and Practice*, 2021 (10): 75-78. DOI: 10.19851/j.cnki.cn11-1010/f.2021.10.356.
- [11] Qiang Huanhuan, Wang Hui, Lei Shuyu Analysis on the spatial differentiation characteristics and influencing factors of housing prices in big cities based on MGWR model —— Take the main urban area of Nanjing as an example [J]. *Modern Urban Research*, 2024 (4): 109-116.
- [12] Chen Xiaoliang, Chen Standard, Wang Zhaorui, etc. The influencing factors of housing price differentiation among cities [J]. *Economic Theory and Economic Management*, 2024, 44 (2): 49-64.
- [13] Chen Xiaoliang, Cheng Shuo, Chen Cao, and other studies on the influencing factors of housing prices in first-tier cities based on machine learning methods [J]. *Nankai Journal (Philosophy and Social Sciences edition)*, 2023 (6): 146-163.
- [14] MAJ, CHENG J C, JIANG F, et al. Analyzing driving factors of land values in urban scale based on big data and non-linear machine learning techniques [J]. *Land use policy*, 2020, 94: 104537.
- [15] HO W K O, TANG B S, WONG S W. Predicting property prices with machine learning algorithms [J/OL]. *Journal of Property Research*, 2021, 38(1): 48-70. DOI:10.1080/09599916.2020.1832558.
- [16] YANG L, CHAU K W, SZETO W Y, et al. Accessibility to transit, by transit, and property prices: Spatially varying relationships [J]. *Transportation Research Part D: Transport and Environment*, 2020, 85: 102387.
- [17] YANG L, CHUX, GOUZ, et al. Accessibility and proximity effects of bus rapid transit on housing prices: Heterogeneity across price quantiles and space [J]. *Journal of Transport Geography*, 2020, 88: 102850.
- [18] BISHOP K C, KUMINOFF N V, BANZHAF H S, et al. Best Practices for Using Hedonic Property Value Models to Measure Willingness to Pay for Environmental Quality [J/OL]. *Review of Environmental Economics and Policy*, 2020, 14(2): 260-281. DOI:10.1093/leep/reaa001.
- [19] GUIFFARD J B. Valuing the virtual: The impact of fiber to the home on property prices in France [J]. *Telecommunications Policy*, 2024, 48(4): 102732.
- [20] Fu Yangjun, Qu Weidong advance and retreat: Research on urban Residence Decision of Floating Population —— Analysis the effect of multiple factors such as urban housing price and public service level [J / OL]. *Price Theory and Practice*, 2023 (7): 131-135. DOI: 10.19851/j.cnki.cn11-1010/f.2023.07.187.
- [21] Wang Jinfeng, Xu Chengdong: Principle and Prospect [J]. *Journal of Geography*, 2017,72 (1): 116-134.