

Research on the Urban Development Resilience and Sustainability Based on Data Analysis and Model Construction

Junpeng Yuan^{*, #}, Hao Yin[#], Jiongyi He[#], Lin Zhang[#], Yantong Li[#]

School of Mathematics and Information Sciences, Guangzhou University, Guangzhou 510006, China

*Corresponding author: Junpeng Yuan (Email: 13257516286@163.com)

[#]These authors contributed equally.

Abstract: This paper constructs a mathematical model based on the analysis of Point of Interest (POI) information and real estate sales data, aiming to provide scientific guidance for the managers and decision-makers of City 1 and City 2, and assist them in achieving high-quality sustainable development under multiple challenges. Regarding the problems of housing price prediction and housing stock estimation, in the context of a downward trend in housing prices, five indicators such as the greening rate, the ratio of building area to parking space, and parking space management fees are constructed. The results show that there are no significant differences in the number of households, the number of parking spaces, and underground parking fees between the two cities.

Keywords: Analytic Hierarchy Process (AHP); Sustainable Development; Resilience of Urban Development.

1. Introduction

According to statistics, 2022 marked the first time China experienced negative population growth, indicating a long-term downward trend in population size[1]. This demographic shift not only poses challenges to public resource allocation and social security systems but also has profound implications for the high-quality, sustainable development of many cities, particularly second- and third-tier cities[2]. Through detailed analysis and research of the current urban situation, we aim to provide a scientific basis for urban managers and policymakers to support cities in achieving high-quality, sustainable development amidst multiple challenges[3].

Wang and Meng demonstrates the efficacy of ML-supported flood risk assessment, combining flood susceptibility modeling with building function vulnerability analysis using Random Forest and CatBoost algorithms[4]. Complementary work by Zhang et al.[5] in Cities highlights the theoretical synergy between resilience and sustainability, advocating for integrated assessment models that balance short-term adaptive capacity with long-term ecological integrity. Li and colleagues[6] utilize geospatial simulation to optimize green infrastructure allocation in Shanghai, demonstrating a 22% reduction in urban heat island effects through ML-driven landscape connectivity analysis. Scholars propose participatory GIS platforms that integrate citizen

science data, as piloted in Nairobi's slum upgrading projects[7].

This paper first examines the maximum price declines in both cities in recent years, as housing prices are closely tied to community fundamentals. A scoring system is established based on six indicators: green coverage rate, building area-to-parking ratio, number of parking spaces, property management fees, surface parking fees, and underground parking fees. To quantitatively analyze urban service levels, this paper constructs a service level scoring model covering three domains: housing market and economy, environment and spatial utilization, and infrastructure. This paper integrates the Analytic Hierarchy Process (AHP) and generative artificial intelligence models to determine indicator weights. The indicator scores are combined with their corresponding weights to calculate domain scores, ultimately yielding a comprehensive city score to identify commonalities, characteristics, and the respective strengths and weaknesses of the cities.

2. House price prediction model

2.1. House price and score analysis

At the same time, based on the results obtained, this paper plots a scatter plot of house prices and total scores, it is shown in Figure 1 and 2.



Figure 1. Scattered relationship between house prices and scores in the two cities

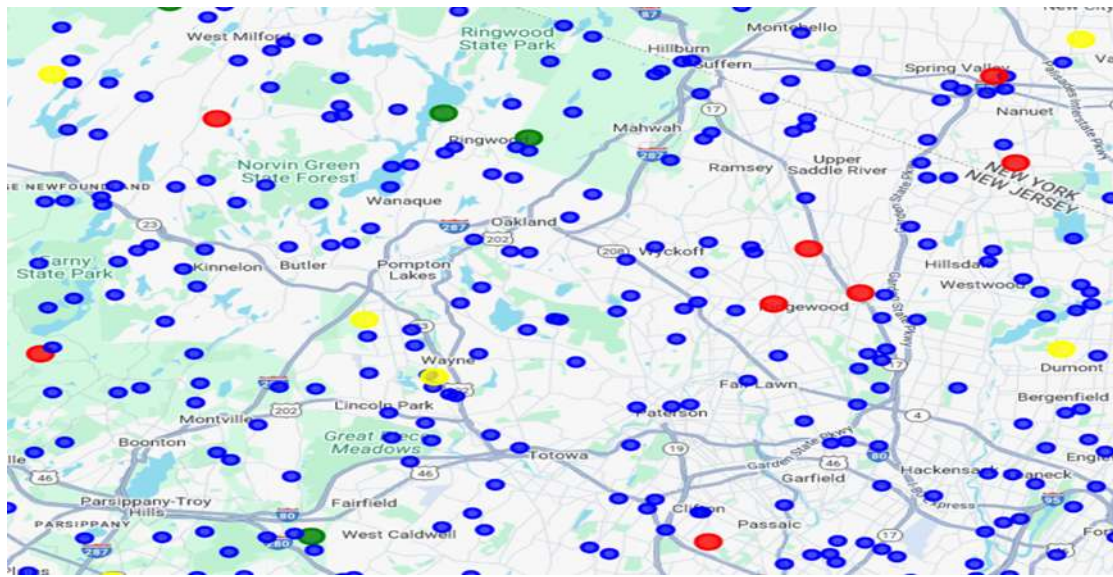


Figure 2. House Prices and Scores for City 1 (from left to right)

The blue color in the chart represents the low price (low score) area, the yellow color represents the middle price (middle score) area, and the red color represents the high price (high score) area. From the figure, it can be observed that the housing prices and scoring features of City 1 and City 2 exhibit a pattern of central concentration with sparse distribution at both ends, with relatively few high-priced (high-scoring) housing points. This indicates that most properties' scores and prices are concentrated in the middle range, with limited high-end properties, which aligns with reality and validates the rationality of the indicator framework constructed in this paper. Meanwhile, the figure shows outliers with extremely high scores, possibly due to the influence of data imputation for missing values, though this is not a common phenomenon. Although theoretically, scores should be positively correlated with housing prices, the actual data does not fully follow this trend, suggesting that other factors, such as urban economic conditions and population mobility, may also have a significant impact on housing prices.

Further comparative analysis reveals that in the left panel of Figure 2, a large number of low-cost properties (blue dots) are concentrated in the central area, with a wide and dense distribution. This indicates that City 1 has a significant number of low-cost properties, likely located in areas with lower development levels or slower development. In contrast,

a smaller number of mid-priced properties (yellow dots) are concentrated in the more central areas of the map, which may represent relatively newer development project.

Meanwhile, the chart shows that yellow dots are most densely concentrated in the central area, indicating that scores in this region are generally at a medium level, likely corresponding to older urban areas or city centers with more mature infrastructure. Red high-score data points gradually increase on the periphery of the center, suggesting the emergence of new development projects or high-end communities. Thus, the interpretation of the score distribution closely aligns with the housing price distribution map.

Next, a local analysis of the housing price distribution map was conducted, as shown in Figure 3. Overall, blue (low-price) points dominate most areas, while yellow (mid-price) and red (high-price) points are relatively scarce and concentrated in the center and certain advantaged areas. This aligns with the general characteristics of urban housing price distribution. The pattern of high prices in the central area and low prices in peripheral areas is reasonable, consistent with the price gradient differences between urban and suburban areas. The occasional presence of high-price points in low-price concentrated areas also reflects the special case of some high-end development projects in otherwise ordinary regions.

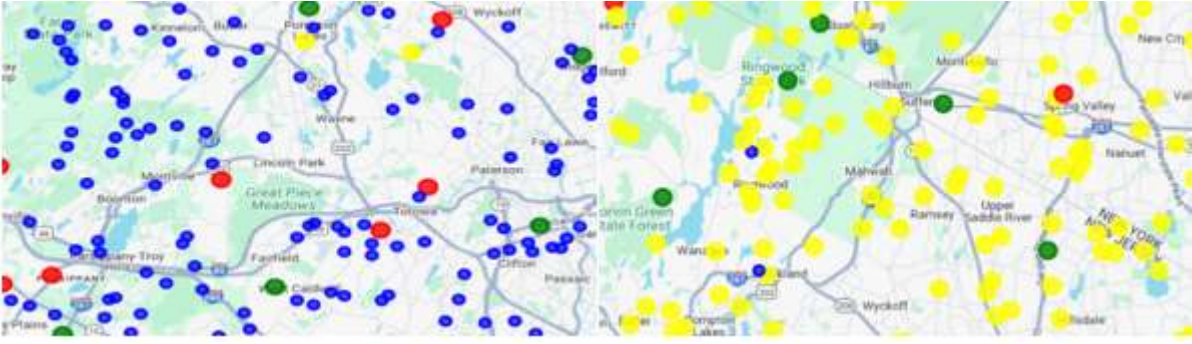


Figure 3. Local Analysis of City 1



Figure 4. Housing Price and Score Distribution in City 2 (Left to Right)

In the analysis of Figure 4, blue (low-price) dots are mainly in the periphery of the map. This aligns with the administrative planning of City 2 as these areas are often at the urban fringe or less - developed. Red (high - price) dots are around the yellow area and the sparse upper - right part of the right map, corresponding to key urban locations. Also, red (high - score) dots in the center of the left map further confirm our speculation that these areas may be high - end residential or high - quality living areas.

2.2. Steps for Calculating Housing Stock

In urban planning and real estate policy formulation, accurately estimating housing stock is crucial. This paper takes City 1 and City 2 as examples and proposes a housing stock estimation model based on population data to simplify the estimation process. The steps are as follows.

1. Calculation of the Number of Households.

$$H_{city} = \frac{P_{city}}{3} \quad (1)$$

where H_{city} represents the average number of households in the city, and P_{city} represents the total population of the city.

2. Calculation of the Number of Residences

$$N_{city} = H_{city} \times R_{city} \quad (2)$$

where N_{city} represents the housing stock of the city, and R_{city} represents the average housing ratio per household in the city.

3. Calculation of the Number of Residences

$$N_{city1} = \frac{P_{city1} \times 1.2}{3} \quad (3)$$

$$N_{city2} = \frac{P_{city2} \times 1.1}{3} \quad (4)$$

4. The complete model formula.

$$N_{city1} = \frac{910.19}{3} \times 1.2 \approx 364.08 \quad (5)$$

$$N_{city2} = \frac{360.41}{3} \times 1.1 \approx 132.15 \quad (6)$$

Based on the data, the average housing ratio per household in City 1 and City 2 respectively, with permanent populations of 9,101,900 and 3,604,100, and fertility rates of 0.88 and

1.19, respectively. Therefore, the average number of people per household in both cities is 3. Substituting the relevant data for the two cities into the calculation, the results are 364.08 and 132.15.

2.3. Construction of Scoring Indicators

In City 1 and City 2, this paper adopts three service level domains—housing market and economy, environment and spatial utilization, and infrastructure—and constructs the following indicators.

1. Housing Market and Economy: Price (USD), Property Management Fee (USD/square meter/month)

2. Environment and Spatial Utilization: Greenery Rate, Usable Floor Area Ratio.

3. Infrastructure: Parking Spaces

To determine the scores for each indicator of a city, the original data value $P_{i,j}$ for the indicator is used as the basis. Through the scoring function $S_{i,j}$ and the quantile $Q_{i,j}$, the indicator score is determined based on the quantile interval it falls into. The specific rules are as follows.

Low housing prices typically indicate higher housing affordability, which is friendly to most residents; whereas high housing prices impose significant economic pressure on residents, negatively impacting their quality of life[8].

Low property management fees can to some extent alleviate residents' economic burden, while high property management fees have the opposite effect.

A high greening rate helps improve residents' quality of life, reduce pollution, and enhance the ecological environment, while a low greening rate has limited impact on improving environmental quality[9].

A high floor area ratio (FAR) in a community indicates high land use intensity, typically associated with city centers or bustling areas, while a low FAR usually reflects low land use efficiency, often found in suburban or less developed areas[10].

Sufficient parking spaces can enhance residents' travel

convenience and quality of life, while insufficient parking spaces have the opposite effect.

For a field d containing m_d indicators, the total score $T_{i,d}$ of a city in field d can be obtained by summing and weighting the indicators.

$$T_{i,d} = \sum_{j=1}^{m_d} w_j S_{i,j} \quad (7)$$

To enhance the representativeness and rationality of the results, a method combining the Analytic Hierarchy Process (AHP) with the AIGC model is used to calculate the weights.

$$w_{i,j} = 0.7 * E_{i,j} + 0.3 * N_{i,j} \quad (8)$$

Where $E_{i,j}$ is the weight obtained through the Analytic Hierarchy Process, and $N_{i,j}$ is the weight obtained through

the AIGC model.

3. Solution of the Service Level Scoring Models

By substituting relevant data, the average scores of City 1 and City 2 in the fields of "Housing and Market Economy," "Environment and Space Utilization," and "Infrastructure," as well as the average scores for each indicator and the comprehensive scores of the urban communities, were calculated. To intuitively analyze the commonalities, characteristics, and strengths and weaknesses of City 1 and City 2, relevant charts were drawn in Figure 5 and 6.



Figure 5. Performance of City 1 and City 2 Across Indicators

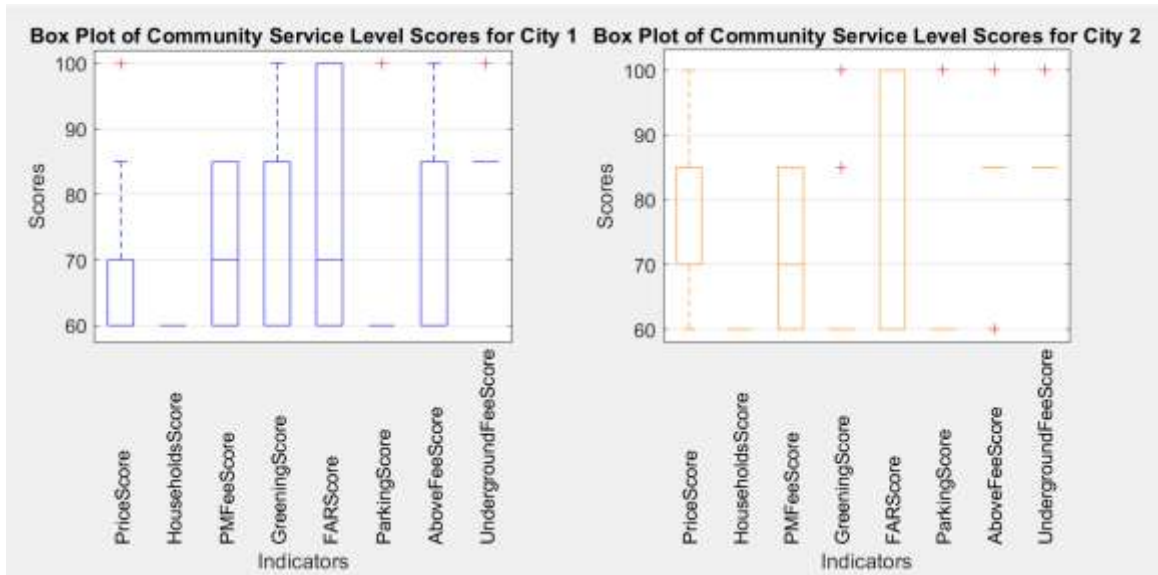


Figure 6. Comprehensive Scores of the Three

From the scoring data, we can obtain the following results regarding individual indicators, major fields, and total scores

In terms of household numbers, parking spaces, and underground parking fees, the score ranges of the two cities remain stable, indicating no significant differences in residential population, parking facility quality, or parking facility costs. Regarding property management fees, the two cities have similar scores on this indicator, suggesting that the property management service levels in both cities are comparable, with similar service content and quality.

Compared to City 2, City 1 has a wider score range for green coverage rate and higher green coverage scores, indicating greater variation in green coverage among its units, but some communities excel in greening levels, suggesting they have an advantage in enhancing residential comfort. In terms of housing prices, City 2 scores significantly higher than City 1, indicating that City 2's housing prices are generally lower than City 1's, demonstrating an advantage in reducing residents' economic pressure.

City 1 performs better in "environment and space

utilization," particularly in green coverage rate, reflecting higher standards and quality of living environment. However, City 1's higher housing prices suggest potential challenges in its real estate market. Additionally, City 1's scores are more concentrated, with most communities performing relatively stably but lacking high-scoring communities, indicating room for improvement in certain areas. City 2 demonstrates strong competitiveness in economic development and the real estate market, with lower housing prices that help alleviate residents' economic burdens. However, its environmental scores are slightly lower than City 1's, with further room for improvement in greening.

Both cities have their strengths and areas for improvement. Urban planners can develop more targeted optimization measures based on these analyses to further enhance the quality of life and overall competitiveness of the cities.

4. Conclusions

This paper develops housing price forecasting and housing stock models. By analyzing relevant data, specific forecasting results are obtained and provides a quantitative analysis of City 1 and City 2 from four perspectives: the housing market, the economy, the environment, spatial utilization, and infrastructure. The results show that the two cities are the same in terms of the number of households, parking spaces, underground parking fees, property management fees, and floor area ratio, and they offer similar living conditions and service levels. City 1 performs well in environmental aspects but has high housing prices and fewer high-scoring communities. City 2 is competitive in the economy and real estate market but needs improvement in greening.

References

- [1] Chen J, Ma H, Yang S, et al. Assessment of Urban Resilience and Detection of Impact Factors Based on Spatial Autocorrelation Analysis and Geo Detector Model: A Case of Hunan Province[J]. ISPRS International Journal of Geo-Information, 2023, 12(10):29.
- [2] Boglárka Eisinger Balassa, Réka Koteczki, Bence Lukács, et al. Comparative Analysis of the Use of Wine Pouch Based on Sustainability and Consumer Resilience Aspects[J]. strengthening resilience by sustainable economy and business - towards the SDGs, 2023.
- [3] Singh J, Hamid A A, J. Garza-Reyes. Supply chain resilience strategies and their impact on sustainability: an investigation from the automobile sector[J]. Supply Chain Management: An International Journal, 2023.
- [4] Omar León, Puente J, Fernandez I, et al. Challenges in Transitioning from Industry 4.0 to Industry 5.0: A Multicriteria Analysis[J]. Journal of Industrial Integration and Management, 2025, 10(01):109-151.
- [5] Ji M, Guo F, Zhou C, et al. Research on risk assessment and damage prediction model of typhoon disaster in Shandong Province based on TDRI[J]. 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT), 2023:934-940.
- [6] Masuda H, Nishimura T. Research on improving the resilience and mental activation of the elderly through the formation of narratives that embody resilience[J]. Abstracts of Annual Conference of Japan Society for Management Information, 2023:239-242.
- [7] Mirzayi M H. Developing Agility, Resilience, and Circular Economy Decision-Making Model Based on Data Envelopment Analysis for Evaluating Medical Equipment Suppliers[J]. Process Integration and Optimization for Sustainability, 2025, 9(2):681-693.
- [8] Akhavan M, Alivirdi M, Jamalpour A, et al. Impact of Industry 5.0 on the Construction Industry (Construction 5.0): Systematic Literature Review and Bibliometric Analysis[J]. 2025.
- [9] Jiang W, Wang K L, Miao Z, et al. Can telecommunications infrastructure enhance urban resilience? Empirical evidence from a differences-in-differences approach in China[J]. Environment, Development & Sustainability, 2025, 27(1).
- [10] Kaiser Z R M A, Deb A. Sustainable smart city and Sustainable Development Goals (SDGs): A review[J]. Regional Sustainability, 2025, 6(1).