Research Paper

Identification and cause analysis of food deserts in the old town

A case study of the old town of Nanjing

Li Chenhui, Southeast University, China

Abstract

Amidst the backdrop of an aging population, accessing healthy food has become a critical issue for maintaining food security. Areas where residents face difficulties in obtaining healthy food are termed "food deserts." For old urban areas, studying and identifying food deserts issues are vital for promoting health equity, planning the future development of old cities, and stimulating market vitality. With the support of big data, research on food deserts can now be analyzed at the street level. This paper takes the old city of Nanjing as an example to identify and analyze the formation of food deserts.Utilizing population data, food supply point data, this study first conducts a descriptive statistical analysis of the proximity and density of food supply points from communities to food supply points in the old urban area. It then employs the Two-Step Floating Catchment Area (2SFCA) method, considering the impact of age on food acquisition, to determine food deserts from both the supply and demand perspectives. Subsequently, the study uses the Mann-Whitney U test to explore the differences and connections between the proximity and density of food supply and built environmental characteristics such as the number of buildings in the block, the degree of land use mix, and public service facilities. Finally, Geographically Weighted Regression (GWR) is applied to investigate the relationship between the formation of food deserts and built environmental variables such as the number of buildings in the block, the degree of land use mix, and public service facilities. All data used in this study are publicly available online.The study reveals that there is a negative correlation between the number of buildings in the block and food deserts phenomena in most areas of the old city, while the degree of land use mix and the density of public service facilities are mostly positively correlated with food deserts. These findings provide valuable insights for urban planning and policy-making aimed at enhancing food accessibility and promoting health equity in old urban areas.

Keywords

Food deserts, old city,built environment

1. Introduction

In the mid-1990s, academia defined "food deserts" as geographical areas where a lack of healthy food retail outlets results in residents struggling to access healthy and affordable food within a certain distance (Cummins,2002).

According to the World Health Organization (WHO), a healthy diet can help prevent all forms of malnutrition, as well as non-communicable diseases (NCDs), including diabetes, heart disease, stroke and cancer. An unhealthy diet is one of the major risks to global health. In the context of China's aging population, access to healthy food has become an important issue to maintain food safety. For the old city, the research and identification of food deserts is of great significance for promoting health and equity, planning the future development of the old city, and stimulating the vitality of the old city market.

Properly distributed food supply points help reduce food waste and carbon emissions during transportation, supporting sustainable urban development. Urban planning should consider the reasonable distribution of food supply points, residential areas and work areas to reduce the impact on the environment. At present, with the support of big data, the study of food deserts have reached the street scale for analysis and research.

The current research on food deserts are mainly concentrated in the field of quantitative analysis. Traditional analysis methods mainly analyze the accessibility of food supply points, including buffer analysis (Kristian et al,2008; Clarke,2002)、kernel density analysis (Moore et al,2008) and other methods for density and proximity analysis. Some scholars also use network analysis (He et al,2019) to divide the arcs of street networks into homogeneous linear units for analysis, or use geographic grid methods to assign equal weights to grid population density. Thus, it is associated with the distribution of food deserts(Moore et al,2008). In addition, some scholars used road data to build a network analysis model and studied the accessibility of demand areas to food retail points through network analysis.

The current research on food deserts are mainly concentrated in the field of quantitative analysis. Traditional analysis methods mainly analyze the accessibility of food supply points, including buffer analysis, kernel density analysis and other methods for density and proximity analysis. Some scholars also use network analysis to divide the arcs of street networks into homogeneous linear units for analysis, or use geographic grid methods to assign equal weights to grid population density. Thus, it is associated with the distribution of food deserts. In addition, some scholars used road data to build a network analysis model and studied the accessibility of demand areas to food retail points through network analysis.

This study mainly studies the following issues. First, how to identify areas with weak food supply capacity (food deserts) in the old urban areas with high population density. In addition, with the support of big data, how to better consider the needs of the elderly population. Whether the relationship between food supply capacity and the number of block buildings, the mixing degree of block land use and the density of public service facilities is close at the block scale, and on this basis, how the built environment factors at the block scale affect the food supply capacity.

This study has theoretical and practical significance. First, an interdisciplinary approach is used, combining statistical and geographic data with urban planning. At the same time, the needs of vulnerable groups are taken into account, which in turn contributes to the construction of more equitable and harmonious cities. The distribution of food supply points is related to the health and quality of life of urban residents, and the rationality of the distribution of food supply points is related to the health and quality of life of urban residents, and is an important part of urban sustainable development. In addition, the rational distribution of food supply points is an important part of urban renewal and community development, and research can promote the sustainable development of old urban areas and the enhancement of community cohesion.

2. Research scope and data source

2.1. Research object and scope

The research object of this paper is the main block of the old city of Nanjing. According to the relevant documents on the protection of historical and cultural cities, the old city is within the three administrative regions of Xuanwu, Gulou and Qinhuai, and the core area of the ancient city basically covers this scope. Therefore, the main block of these three administrative regions is selected for research when the data is selected. Nanjing is one of the four ancient capitals of China, with a history of more than two thousand years. Located in the southwest of Jiangsu Province and the lower reaches of the Yangtze River, Nanjing is rich in historical, cultural and natural landscapes. There are four distinct seasons, abundant rain, belonging to the north subtropical humid climate. In 2023, the population of the elderly aged 60 and above in Gulou District is 307,100, and the population of the elderly aged 60 and above in Qinhuai District is 266,700, which is one of the regions with the largest elderly population in Nanjing. The aging of population in Gulou District, Qinhuai District and Xuanwu District is obvious.

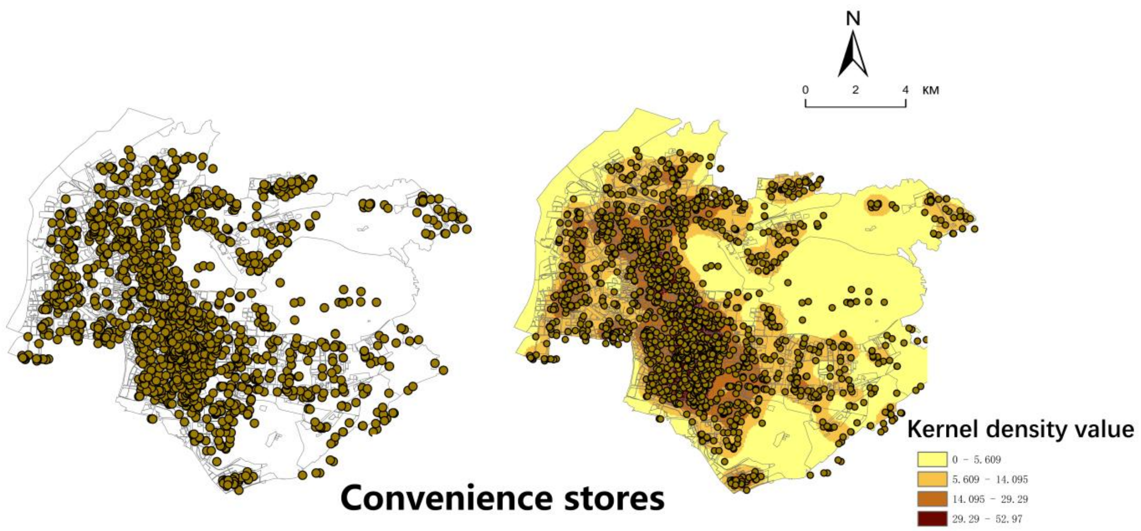
2.2. Data Sources

The population density data comes from data shared by Professor Chen Yuehong's team on the figshare platform(Chen et al,2024).A 100-m gridded population dataset of China’s seventh census in 2020 was generated by stacking ensemble learning and geospatial big data.

The population age data used in this study came from the website "worldpop", and the original data was raster data with a resolution of 100m. The site contains both 'Constrained' and 'Unconstrained' data, covering the years 2000 to 2020. This study extracted the number of people over 70 years old in each neighborhood within the 2020 study area of the 'Unconstrained' type.

The block data comes from the public data set "41 Urban blocks in 63 Chinese cities" released by Beijing City Lab, which extracts the largest construction land plots in the city as the range.Eleven attributes are available in the data, including ID of the block (PARCEL_ID), name of cities (CITY_NAME), number of buildings (BLD_Count), average floor number (BLD_FLOORN), average floor area ratio (BLD_FAR), average building density (BDL_bldDen)and so on (Long et al, 2019).The block data comes from the public data set "41 Urban blocks in 63 Chinese cities" released by Beijing City Lab, which extracts the largest construction land plots in the city as the range.

POI food supply point data from AmAP API2022 edition. The data content includes latitude and longitude, the name and type of the point. The food supply points in this study are defined as supermarkets, which are divided into department stores, supermarkets, convenience stores and shopping centers according to types, with a total of 2,567 stores. The food supply points and distribution involved in this study are shown in the figure below. In addition, POI data of interest points of public service facilities are also obtained. This study selects educational facilities, medical and health institutions, cultural and entertainment institutions, transportation stations, sports venues, social welfare and security facilities as public service points, so as to calculate the number variable of public service facilities in blocks.



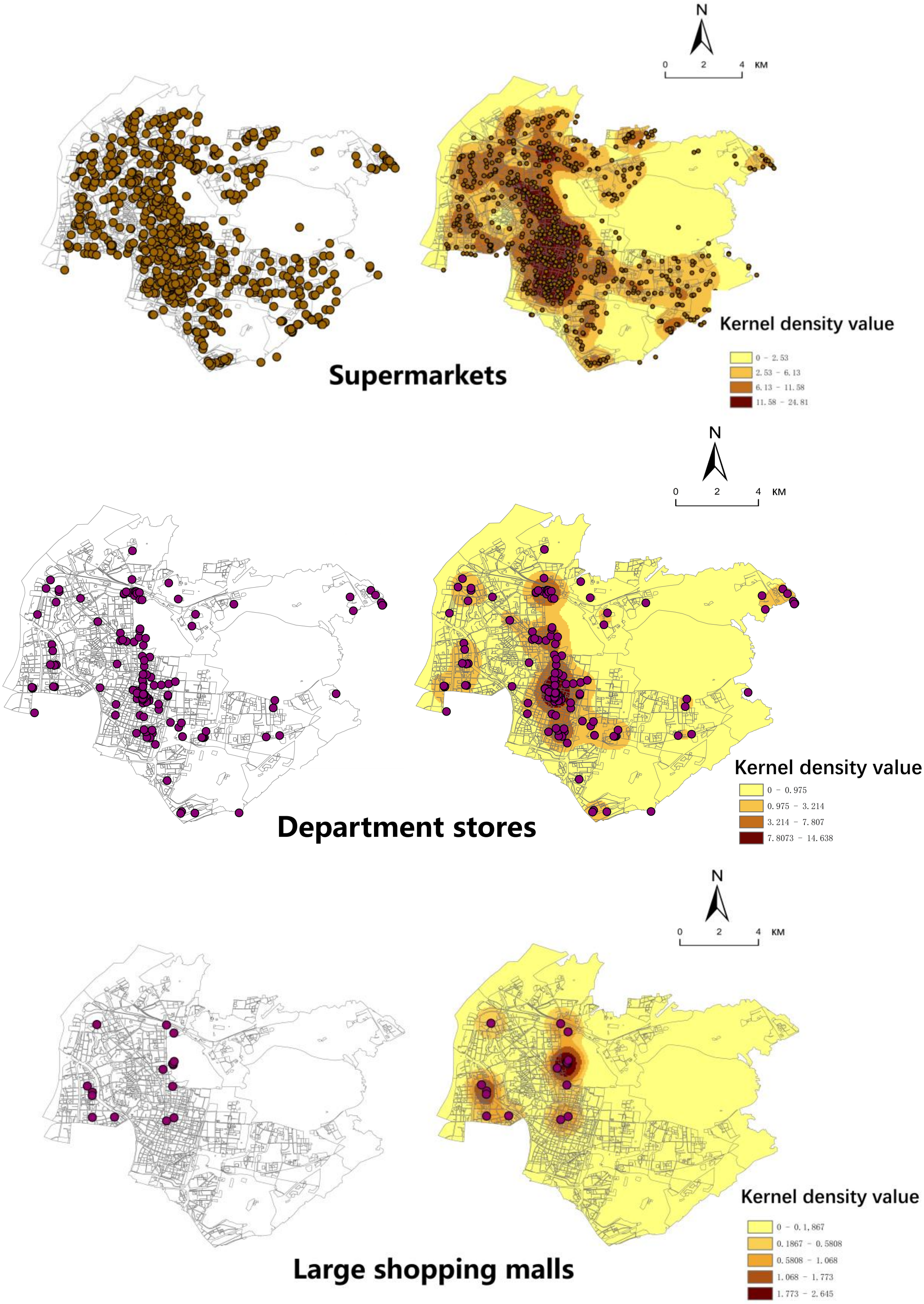


Figure 1. Distribution and kernel density estimation of four types of food supply points. Source: Author's own work.

Table 1. Statistical quantity, classification and proportion of four types of food supply points. Source: Author's own work.

Type	Number	Grade	Proportion
Large shopping malls	16	4	0.01
Department stores	134	3	0.05
supermarket	766	2	0.30
Convenience store	1650	1	0.64

3. Methodology

3.1. Food deserts identification based on two-step floating catchment area method (2SFCA)

$$A_i^F = \sum_{j \in \{d_{ij} \le d_0\}} R_j = \sum_{j \in \{d_{ij} \le d_0\}} \left[\frac{S_j}{\sum_{k \in \{d_{kj} \le d_0\}} D_k} \right]$$

The 2SFCA method was first proposed by Radke et al. (2000), further improved by Luo et al. (2003) and named as the Two-Step Floating Catchment Area (2SFCA) me*thod*. The method is to search the supply point and the demand point respectively in two steps. The first step is to focus on the food supply point, search the demand point within the threshold range, and then calculate the supply and demand ratio. The second step is to search all supply points within the threshold range, and then calculate the sum of the demand ratio of all supply points found.

In the formula, i represents the block and forms the central demand point k of the block; j represents the food supply point, which has been divided into four levels according to the food supply capacity; AiF represents the food supply capacity of block i calculated according to the Two-Step Floating Catchment Area (2SFCA) method, that is, the sum of the supply and demand ratio of food supply points in the corresponding search range of block i; dij is the radius of distance between block i and food supply point j; Rj is the ratio of the food supply capacity level corresponding to food supply point j to the population served within the search radius d0; Sj represents the supply scale of supply point j, which has been expressed as their respective supply grade according to their respective supply capacity; Dk represents the scale of demand at point k, which in this study refers to the number of people covered.

3.2. Mann-Whitney U test

The Mann-Whitney U test, also known as the Wilcoxon rank sum test, is a non-parametric statistical method used to compare the distribution of two independent samples. It is especially suitable for cases where the data does not meet the normal distribution or there are outliers, and is a non-parametric alternative to the T-test. This study uses this method to measure whether the number of buildings in the block, the degree of functional mixing of land use, and the number of public service facilities will affect the proximity and density of food supply points in the block.

3.3. Geographically weighted regression analysis

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i$$

Geographically Weighted Regression (GWR) is a local linear regression model that allows the regression coefficient to vary spatially, so as to capture the non-stationarity of spatial data. This model, proposed by Professor A. Stewart Fotheringham in 1996, is mainly used to deal with the problem of heterogeneity in spatial data, that is, the relationship between spatial variables may change with the change of spatial location. In the formula, yi is the dependent variable value at the block position (ui,vi), which in this study refers to the food supply capacity; xikis the KTH independent variable value at the location (ui,vi). This study selected three variables: number of buildings, mixed degree of land use functions, and number of public service facilities. βk(ui,vi) is the KTH regression coefficient at the block location (ui,vi), which is a function of geographical location; ϵi is the error term.

4.Results and conclusions

4.1. Results of food deserts identification

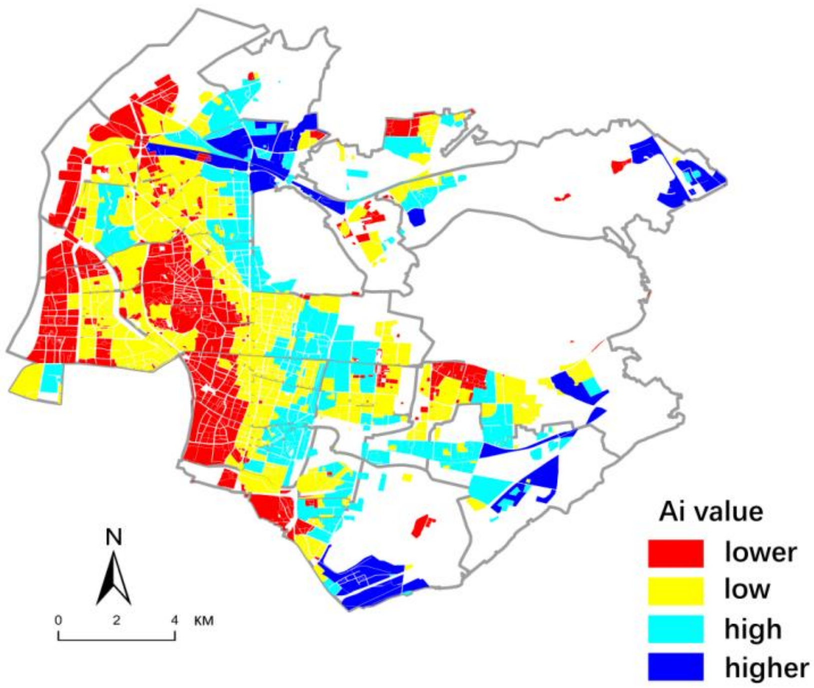


Figure 2. Map of district food supply capacity distribution. Source: Author's own work.

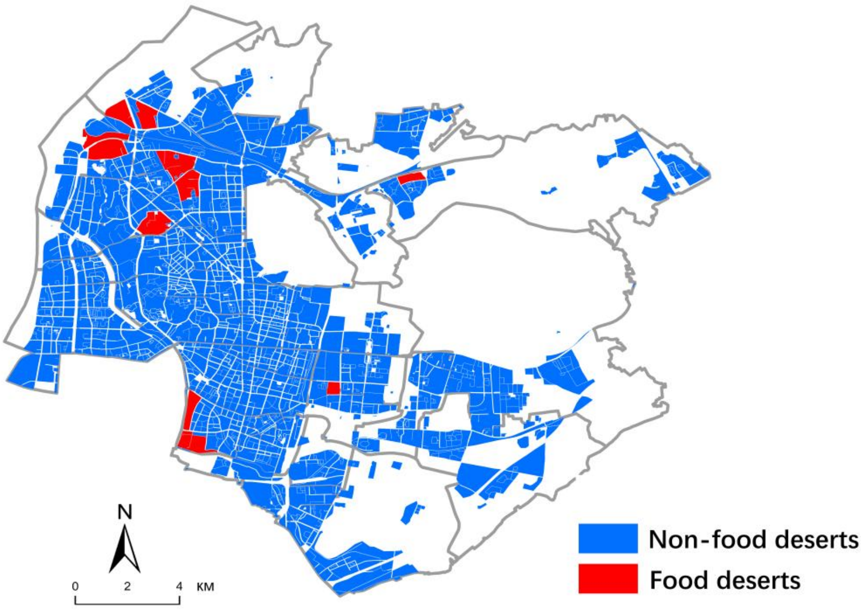


Figure 3. Results of food deserts analysis. Source: Author's own work.

As shown in Figure 1, it can be seen that the distribution and quantity of the four food supply points are different. The higher the supply capacity level, the smaller the quantity. Convenience stores and supermarkets are generally concentrated in the most densely populated areas of Qinhuai District and Gulou District, and large shopping malls with the highest food supply capacity can only be found in some streets of Gulou District. Next, department stores are the most dispersed.

The Two-Step Floating Catchment Area (2SFCA) method was used to obtain the distribution diagram of food supply capacity (Ai value), and the Ai value was divided into four hierarchical segments according to the natural breakpoint method of Arcgis: 0.28, 0.51, 0.98 and 1.91. Among them, below 0.28 is very low supply capacity, below 0.51 is low supply capacity, between 0.51 and 0.98 belongs to high food supply capacity, between 0.98 and 1.91 belongs to high food supply capacity. The areas with high food supply capacity are mainly distributed in Xiaoshi street, east of Xuanwuhu Street, south of Xiaoling Wei Street and south of Honghua Town.

The distribution of food supply capacity in blocks with more than 300 elderly people under the age of 70 was superimposed to screen out the distribution map of food deserts. It can be seen that Baotaqiao Street, Xiaguan Street, Jianning Road Street and Shuangtang Street in the southwest of Nanjing Old City have a small number of blocks belonging to the food deserts. Other blue areas are non-food deserts areas.

4.2. The result of cause analysis

The proximity of food supply points in a block is defined as: the distance between the center point of the block and the nearest food supply point; The density of food supply points in a block is defined as the number of supply points within a unit buffer area. The original hypothesis is that there is no significant difference in food supply density and proximity of blocks with different building numbers, different functional mixing degree and different public service density.

According to Mann-whitney Test, the food supply density and proximity of blocks with different building numbers, different functional mixing degree and different public service density are significantly different, which rejects the original hypothesis. Both sets of data do not satisfy the hypothesis of normal distribution. Firstly, descriptive statistical analysis is made on the data. In the following table, the quartile, average value and other values of food supply proximity and density corresponding to the three variable groups are counted respectively, and the difference in density and proximity between different groups can be clearly seen.

Table 2. Descriptive statistics of Mann-whitney Test. Source : Author's own work.

	Proximity			Density		
	Not stratified	Low number of buildings	High number of buildings	Not stratified	Low number of buildings	High number of buildings
25%	0	0	0	33	36	35
50%	29	67	50.8	56	77	58
Mean	79	90.6	67.3	69	82.3	103
75%	109	145.4	232.2	99	119	97
std.dev	126	8.5	47.8	44	15.2	35.9
		Low mixing degree of function	High mixing degree of function		Low mixing degree of function	High mixing degree of function
25%		0	0		31	43
50%		41	0		52	68
Mean		88.6	49		66	78
75%		125	60		93	110
std.dev		135.6	89		44.3	43.3
		Low number of public service facilities	High number of public service facilities		Low number of public service facilities	High number of public service facilities
25%		0	99		28	99
50%		54.6	126		43.5	126
Mean		100	21		48	121
75%		150.6	146		64	146
std.dev		141	39		27	36

The test results show (FIG. 4) that all the P values of the three groups of variables after the test of "high and low number of buildings", "high and low mixing degree of land use functions", and "high and low number of public service facilities" are less than 0.05 significance level, which means that statistically it can be proved that the above three groups of variables have significant differences in food supply density and proximity of blocks. In other words, the food supply density and proximity of blocks with different number of buildings, different functional mixing degree and different public service density are significantly different.

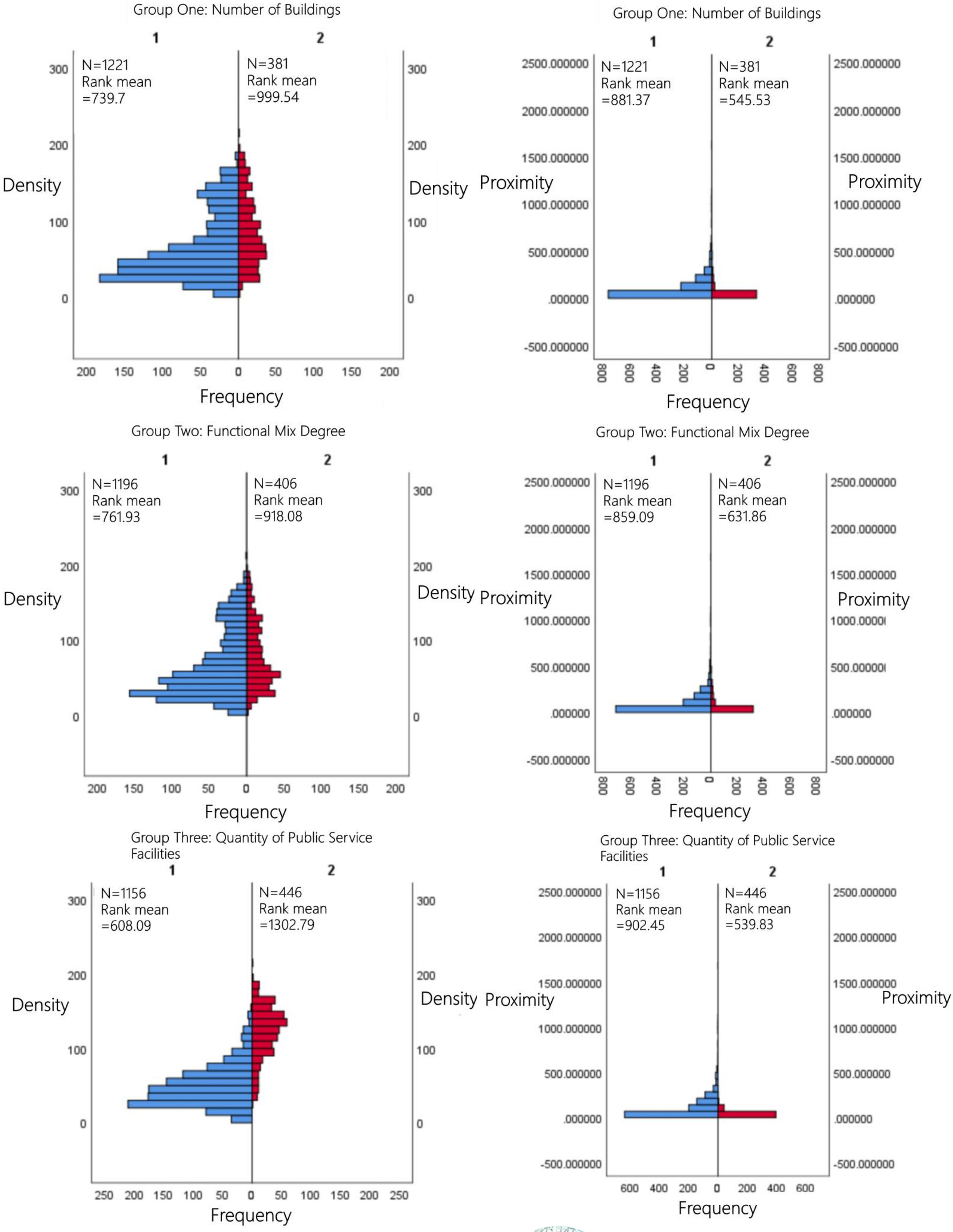

Group One: Number of Buildings
1
2
N=1221
Rank mean
=739.7
N=381
Rank mean
=999.54
Density
Density
Frequency
Group One: Number of Buildings
1
2
N=1221
Rank mean
=881.37
N=381
Rank mean
=545.53
Proximity
Proximity
Frequency
Group Two: Functional Mix Degree
1
2
N=1196
Rank mean
=761.93
N=406
Rank mean
=918.08
Density
Density
Frequency
Group Two: Functional Mix Degree
1
2
N=1196
Rank mean
=859.09
N=406
Rank mean
=631.86
Proximity
Proximity
Frequency
Group Three: Quantity of Public Service Facilities
1
2
N=1156
Rank mean
=608.09
N=446
Rank mean
=1302.79
Density
Density
Frequency
Group Three: Quantity of Public Service Facilities
1
2
N=1156
Rank mean
=902.45
N=446
Rank mean
=539.83
Proximity
Proximity
Frequency

Figure 4. Three groups of inspection results of "high and low number of buildings", "high and low mixing degree of land use functions" and "high and low number of public service facilities". Source: author.

After the multicollinearity (VIF) test, the multicollinearity of the variable is excluded. A geographically weighted regression model (GWR) was used to further verify the possibility of significant differences caused by the above three variables.

First, as shown in Figure 6 (a), the relationship between the functional mixing degree coefficient of blocks and the food supply capacity shows obvious spatial heterogeneity, and the correlation is mainly positive, that is, the higher the functional mixing degree of blocks, the stronger the food supply capacity. There is no significant difference in the spatial variation of absolute value of regression coefficient. The negative regression coefficient is mainly distributed in the southern part of the old city. It is speculated that the reason for this distribution relationship is that the functional mixing in the southern part of the city brings space congestion and land shortage, thus weakening the food supply capacity. And the positive is mainly distributed in the old city of Gulou district, especially the central gate street, Hunan Road street.

As shown in Figure 6 (b), the number of buildings and the food supply capacity are mainly negatively correlated, that is, the more buildings there are, the weaker the food supply capacity of the block. Speculation may be associated with the negative effects of more densely populated neighborhoods with a high number of buildings. The positive regression coefficient is mainly distributed in Xiaolingwei Street in the eastern part of the old city, which belongs to the science and innovation sector and has a moderate population distribution. The more buildings, the more food supply points, the stronger the supply capacity.

As shown in Figure 6 (c), the density of public service facilities and food supply capacity are mainly positively correlated, that is, the higher the density of public service facilities, the higher the food supply capacity. However, the negative regression coefficient is mainly distributed in the southeast of Xuanwu District, which may be due to the mismatch between public service facilities and residents' actual needs. Public service facilities may focus more on providing entertainment or leisure services, while the facilities related to food supply, such as supermarkets and wet markets, are insufficient. As a result, the more public service facilities, the weaker the food supply capacity.

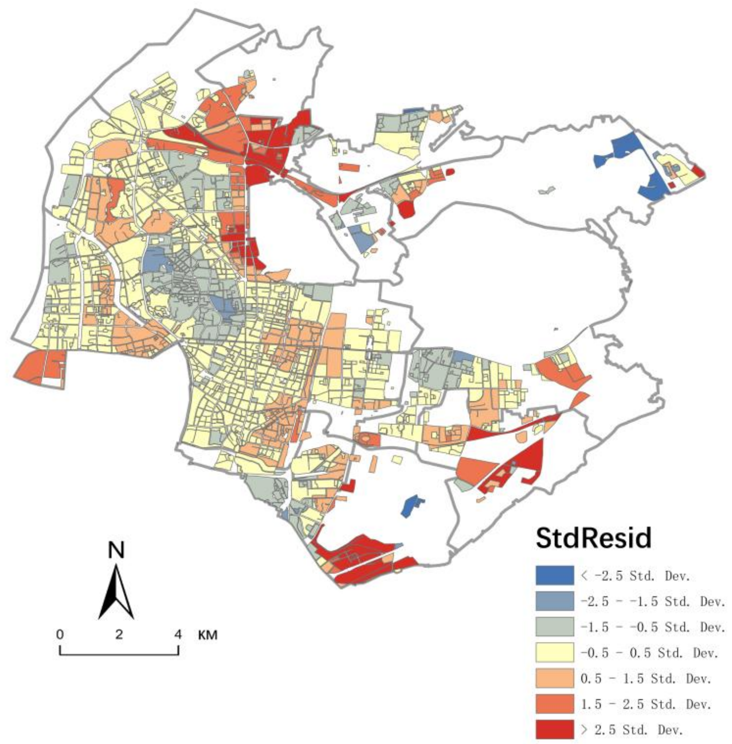
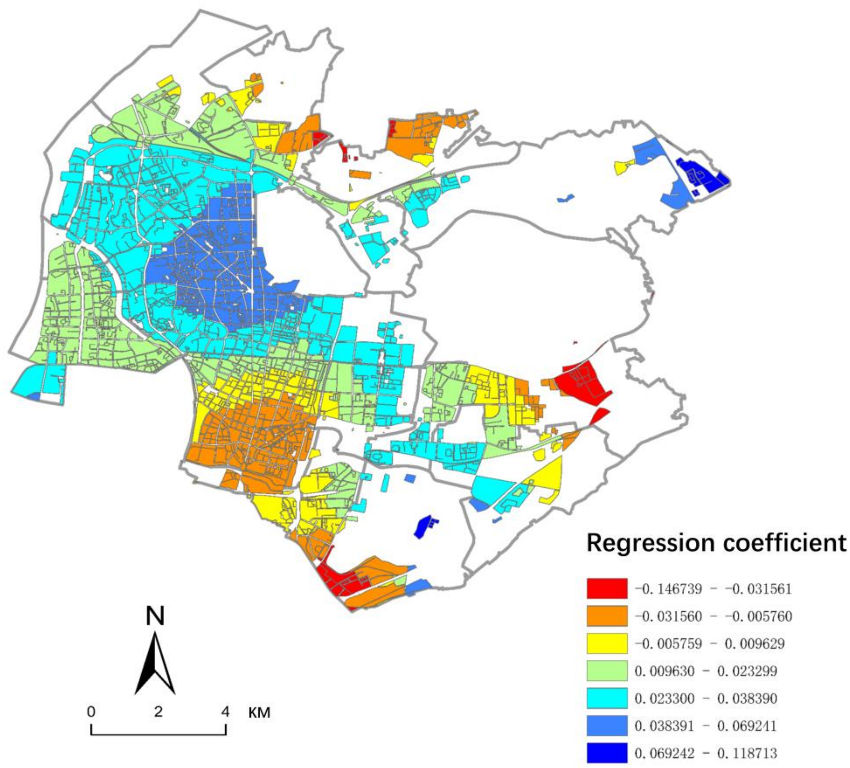
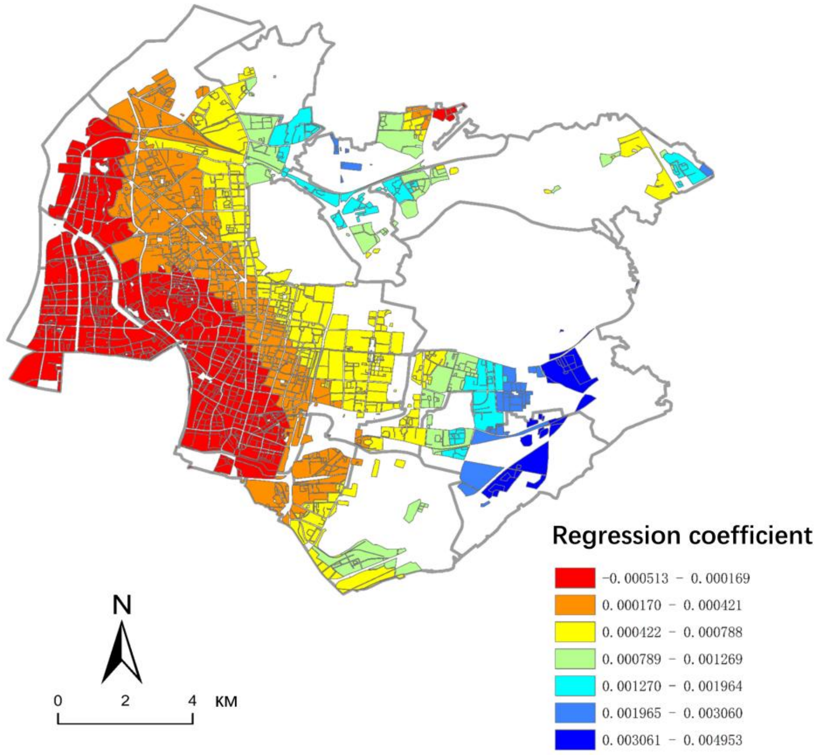


Figure 5. Global standard error. Source: Author's own work.





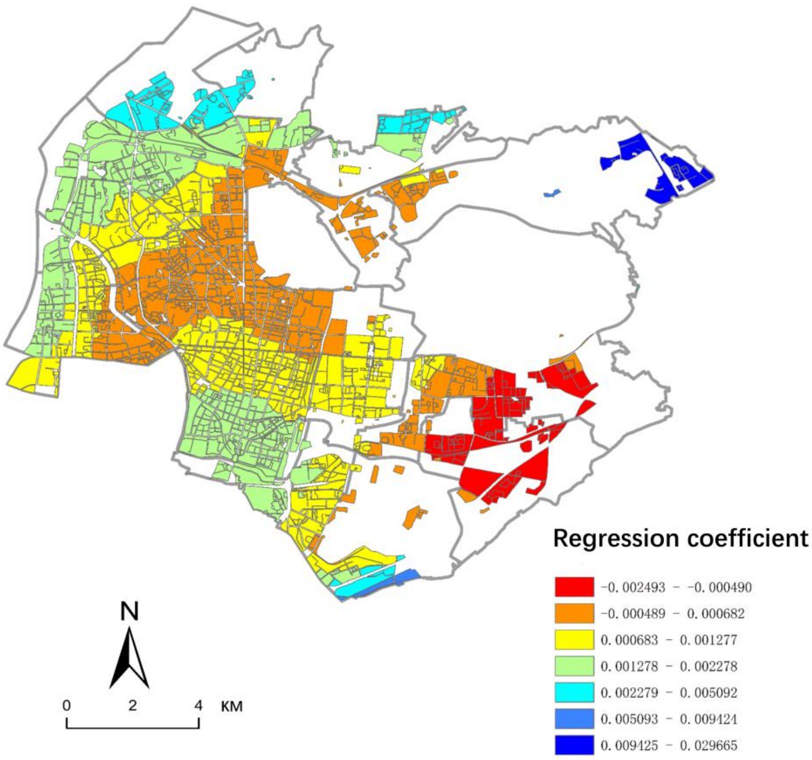


Figure 6. Mixed-degree coefficient of block function (a), regression coefficient of building number (b), regression coefficient of public service facility density (c). Source :Author's own work.

5. References

Chen, Y., Xu, C., Ge, Y., et al. (2024) 'A 100-m gridded population dataset of China's seventh census using ensemble learning and geospatial big data', Earth System Science Data Discussions, 2024, pp. 1–19.

Clarke, G. (2002) 'Deriving Indicators of Access to Food Retail Provision in British Cities: Studies of Cardiff, Leeds and Bradford', Urban Studies, 39(11), pp. 2041–2060.

Cummins, S. and Macintyre, S. (2002) '"Food deserts" – evidence and assumption in health policy making', BMJ, 325(7361), pp. 436–438.

He, Y.K., Wang, L. and Ai, T.H. (2019) 'A Study on Food Accessibility Supported by Street Network Linear Partition Model', Geography and Geo-Informatics Science, 35(1), pp. 50–57+1. [in Chinese]

Kristian, L. and Jason, G. (2008) 'Mapping the evolution of "food deserts" in a Canadian city: Supermarket accessibility in London, Ontario, 1961–2005', International Journal of Health Geographics, 7(1), p. 16.

Long, Y., Li, P. and Hou, J.X. (2019) 'Three-Dimensional Urban Form at the Block Level in Major Chinese Cities', Shanghai Urban Planning Review, (3), pp. 10–15. [in Chinese]

Moore, L.V., Diez Roux, A.V. and Brines, S. (2008) 'Comparing Perception-Based and Geographic Information System (GIS)-based characterizations of the local food environment', Journal of Urban Health, 85(2), pp. 206–216. doi:10.1007/s11524-008-9259-x.

Moore, L.V., Roux, A.V.D., et al. (2008) 'Associations of the local food environment with diet quality - a comparison of assessments based on surveys and geographic information systems: the multi-ethnic study of atherosclerosis', American Journal of Epidemiology, 167(8), pp. 917–924.

Smoyer-Tomic, E.K., Spence, C.J. and Amrhein, C. (2006) 'Food Deserts in the Prairies? Supermarket Accessibility and Neighborhood Need in Edmonton, Canada', The Professional Geographer, 58(3), pp. 307–326.