DOCTORAL DISSERTATION

Research on spatial distribution pattern and interaction mechanism within rail transit station realm

September 2023

WANG DI

2020DBB404

Graduate School of Environmental Engineering

Dewancker Bart Lab

The University of Kitakyushu, Japan

Acknowledgements

Time flies, years pass by, as I prepare to bid farewell, memories remain vivid in my mind. The excitement of our first encounter, the support and assistance during challenging times, the laughter and joy during festive moments, the bittersweet experiences of learning, the hard work and dedication during busy days, and the vibrant colors of leisurely moments—all of these are the most precious treasures of the past three years. I extend my heartfelt gratitude to the teachers, seniors, classmates, and loved ones who have encouraged, supported, and helped me along the way.

First and foremost, I would like to express my gratitude to my doctoral supervisor, Professor Bart Julien Dewancker, for his meticulous guidance and assistance in various aspects of my studies, research, and life over the past three years. Professor Bart's profound knowledge and rigorous scholarly attitude have deeply influenced me, enabling continuous progress in my learning, and encouraging innovation in my research. From topic selection, outline revisions, initial draft modifications, to the final completion, Professor Bart has provided patient and meticulous guidance, often taking time out of his busy schedule to discuss with me. This has been a crucial factor in the successful completion of this paper. I am thankful for Professor Bart's teachings and care.

Furthermore, I would like to take this opportunity to express my deep gratitude to the members of the dissertation supervisory committee, Professor Yuji Ryu, and Professor Gao Weijun, for their invaluable comments, suggestions, and guidance, which have greatly contributed to the improvement of this paper.

Then, I would like to thank all the seniors and classmates for their help and concern, including Dr. Zuan Liu, Dr. Simin Yang, Dr. Shuo Chen, Dr. Jinming Wang, Dr. Jialu Dai, and Dr. Siswanti Zuraida. Your support has made my time in Bart Lab a joyful and loving experience.

I would also like to express my gratitude to my girlfriend, Luo Mingzhen. Your companionship during these three years in Japan has enriched my research, studies, and life.

Lastly, I am grateful to my family. You have always been my greatest support and source of strength. I appreciate your silent dedication. Although we are separated by thousands of miles, this familial bond is the most precious baggage on my journey, and everything you have taught me will accompany me along the way.

I would like to once again thank all the people who have encouraged, cared for, supported, and helped me. I wish you all happiness and good health!

Nomenclature

1. Transit-Oriented Development (TOD)
2. Station-City Integration Development (SCID)
3. Rail Plus Property (R+P)
4. The Belt and Road Initiative (BRI)
5. Point of Interest (POI)
6. Geographic Information System (GIS)
7. Joint Development (JD)
8. Global Positioning System (GPS)
9. Origin-Destination (OD)
10. Ordinary Least Squares (OLS)
11. Spatial Error Model (SEM)
12. Spatial Lagged Model (SLM)
13. Geographically Weighted Regression (GWR)
14. Multiscale Geographically Weighted Regression (MGWR)
15. Metro Rail Transit (MRT)
16. Station Integrated Development (SID)
17. Statistical Package for the Social Sciences (SPSS)
18. Akaike Information Criterion corrected (AICc)
19. Variance Inflation Factor (VIF)
20. Monthly Users Active (MUA)
21. Application Programming Interface (API)
22. OpenStreetMap (OPM)
23. Land Use Mix (LUM)

Abstract

The rapid development of urban rail transit has greatly improved the accessibility of areas around stations, attracting urban population and public resources to gather in these areas, making the station realms hotspots for current urban development and construction. In recent years, with transit-oriented development (TOD) gaining popularity among urban planners and managers worldwide, particularly in developing countries facing issues such as rapid urbanization, traffic congestion, and environmental pollution, the station realm has become a key area for urban planning and construction. However, China's urban rail transit started late compared with developed countries, the construction of the station realm lags behind the overall urban planning for a long time, and there is a mismatch between the function supply and people's demand. These problems will eventually lead to a low space utilization rate and a waste of public resources in the station realm. Furthermore, relevant urban studies for station realms are still in infancy, especially due to the difficulty in obtaining data on population activity and functional facility, which has resulted in relatively few studies on their spatial distribution patterns as well as interaction mechanisms. Therefore, this research will be based on geographic information system (GIS) and combined with multi-source data to establish the research methodology system tailored for the spatial distribution pattern and interaction mechanism of station realms. By conducting an empirical study on 16 stations of Xi'an Metro Line 2, this research achieves the following main results:

Chapter 1 is the Introduction, which focuses on the research background, literature review, research questions and purposes, as well as the content and technical route of the research. Firstly, the problems existing in the development process of China's urban rail transit station realm are introduced. Secondly, the importance of studying the spatial distribution pattern and interaction mechanism of station realms is discussed. Then, through the literature review of the research trends, theories, and methods of station realms, it is pointed out that research on population activities and functional facilities within station realms is still relatively scarce. Meanwhile, the rise of big data technology has made it possible to study urban space at the micro level, while GIS-based spatial analysis methods and models still dominate urban space research. Finally, the research questions and purposes are presented, and the research content and framework are concluded.

In Chapter 2, Research on relevant theoretical foundations. First, the scope and limitations of this study are clarified by defining the concepts of rail transit, station realm, urban spatial pattern, and interactive relationship, as well as population activity and functional facility. Secondly, the research related to the station realm clarifies the walking scale is the core elements for defining the scope in this thesis, as well as the function-oriented classification system was used to determine the type of station realm. Finally, the study of fundamental theories such as central place theory and urban catalyst theory as well as the comparison of related cases provide the basis for the study of the spatial distribution characteristics and interaction mechanism of the station realm in this thesis.

In Chapter 3, The Construction of research methodology framework. The research methodological framework was constructed based on the research idea of "station realm scope - station realm type - data collection - spatial distribution pattern - interaction mechanism", and the survey methods, measurement methods and analysis methods used in this study were specified.

In Chapter 4, Spatial Distribution Pattern Analysis of Xi’an Metro Line 2 Station Realm. 16

stations of Xi'an Metro Line 2 are selected as the research objects. focuses on the scope and classification results of each study station realm and analyzes the distribution characteristics of population activities and functional facilities. In terms of spatial distribution patterns, population activity in each station realm is primarily characterized by commuting during working day, with the morning peak mainly concentrated between 08:00 and 10:00 and the evening peak mainly distributed between 17:00 and 19:00. On off day, there is no significant morning or evening peak phenomenon, but rather a continuous fluctuation between 09:00 and 18:00, with population activity during this time interval generally higher than on off day. Overall, population activity in different types of station realms is in a distribution statue of aggregation, and the degree of aggregation on working day is generally lower than on off day. In terms of the spatial distribution pattern of functional facilities, except for government agencies and public service facilities, other categories of facilities are in a distribution statue of aggregation in most of the research station realms. Among them, the aggregation degree of catering services, living services, and shopping facilities ranks among the top three. In addition, although there are significant differences in the spatial distribution patterns of population activities and functional facilities between the different station realms, the overall spatial location of the aggregation of population activities and functional facilities is basically the same, which indicates that there is a certain mutual influence relationship between them.

In Chapter 5, Research on the correlation of population activities and functional facilities within station realm. This study analyzed the correlation between population activities and functional facilities in different station realms using Spearman's correlation test. The results of the analysis showed that catering service facility, living service facility, shopping facility, and transportation service facility show a significant positive correlation with population activity in all the research station realm. While Government agencies and public service facility show no significant correlation with population activities in most of the station realms. In addition, based on the significance of the correlation coefficients, the categories of facilities that have no significant correlation with population activities in each research station realm were excluded, thus establishing a variables system for the next stage in the study of interaction mechanisms.

In Chapter 6, Research on Interaction Mechanism within Station Realm. First, the variables system of each research station realm was further optimized using multicollinearity and significance testing of OLS regression models. Second, the optimized variable system was regressed using the GWR model to analyze the degree of interaction between the population activities and different facilities in station realms. On this basis, the interaction sequence for each station realm was established, thus providing a theoretical basis for improving the allocation of public resources in this region. Specifically, the regression results indicate that transportation service facilities have a high degree of interaction with population activity in most of the study station realms. By increasing the number of transportation service facilities, the accessibility and convenience of transportation in the station realm can be effectively improved, thus attracting more population activity. In addition, the study also found that the scale of catering and shopping facilities in each station realm is higher than other facilities, but the degree of interaction with population activities tends to be lower. This suggests that over-sizing of catering and shopping facilities will be very limited in boosting population activity in the station realms and may even result in a waste of urban public resources. It is worth mentioning that medical and healthcare facilities have a high degree of interaction with population activities only in station realms with quality healthcare resources. Therefore, the equity of public healthcare resource allocation between station realms should be worthy of the attention of urban managers and researchers. Overall, the relevant analysis results for the interaction mechanism can help us understand the degree of demand for different categories of facilities by the population activities in the station realm, thus providing a reliable basis for improving the allocation of public resources in this region.

In Chapter 7, Conclusion and prospect. This section concludes all the key findings and provides prospects for future research.

The findings of this dissertation could be an important contribution for urban managers and planners to deeply understand the spatial distribution characteristics and laws of station realms at the micro level, especially the related results on the interaction mechanism between population activities and different categories of functional facilities, which can provide a basis for decision-making to improve the allocation of public resources in this region, so as to effectively avoid the waste of resources and the imbalanced distribution.

Keywords: rail transit station realm; population activity and functional facility; geographic information system (GIS); spatial distribution patterns; interaction mechanisms

Contents

Content of Figures

Content of Tables

Chapter 1. Introduction

Content

1.1 Research background

1.1.1 The rapid urban rail transit expansion leads to spatial development imbalance of station realm

By 2050 approximately two thirds of the world's population will be living in urban areas, bringing a distinct set of challenges and opportunities to the area of urban development. With its outstanding characteristics of high speed, high efficiency, and low energy consumption, urban rail transit plays an active role in promoting the modernization process of cities, improving the transportation environment, and guiding the sustainable development of urban space [1]. Most notably, it has guided and promoted the gathering of the urban population and functions around rail transit stations, thus forming a polycentric urban spatial structure [2]. Studies of rail transit systems in cities such as Hong Kong, Tokyo, and London have found that rail transit station-oriented development models such as TOD, station-city integration, and rail plus property integrate the location, scale, and function of the station, the surrounding urban layout, landscape design, and public service facilities, and other factors for integrated development and utilization, thereby improving the service efficiency of the functional space of the station area as a means to alleviate the growing activity needs of residents. Therefore, as the only channel linking rail transit with urban system, the station area has become the basic place of citizens' daily life and an important part of urban space. Especially in China, which is in the stage of high-speed urbanization, more and more cities have started to develop urban rail transit and invest a lot of urban resources in the development and construction of station realms, so as to improve regional economic benefits, optimize urban spatial structure and promote compact urban development.

However, compared with developed countries, China's urban rail transit started late, and the development concept is relatively backward (Figure 1- 1). In the process of line planning, design, as well as construction, there is a lack of attention to the comprehensive development of station surroundings and urban space, therefore the station is more like being rigidly "implanted" into the urban system, and it is difficult for rail transit system to guide urban development effectively and positively [3]. Thus, the construction of station realm has long lagged behind the urban master plan, thus leading to a mismatch between the growing demand for activities and the supply of urban functional facilities in the region [4,5,6]. These problems will eventually lead to a low space utilization rate and a waste of public resources in the station realm.

As a result, stations with excellent built environments can better play the node effect, thus attracting the most important commercial, cultural, medical, and other resources and a large number of jobs to gather around them, becoming the concentration of modern urban functions and employment. This aggravates the divergence of spatial distribution patterns among stations in the urban rail transit system and further leads to the imbalance of spatial development of stations. This phenomenon is particularly evident in northwest China, where the degree of development of various urban regions is relatively uneven. To solve this problem, domestic and foreign scholars have conducted extensive and in-depth studies on the relationship between urban rail transit and urban space. Most of the studies believe that the characteristics of residents' activities around the rail transit line correspond to the layout of urban functions, and the residents' activities in a specific space and time can produce some corresponding characteristics of urban functional use. Therefore, understanding the interaction between the distribution of population activities and urban functional facilities is the key to solving the imbalance of spatial development in the station realm.

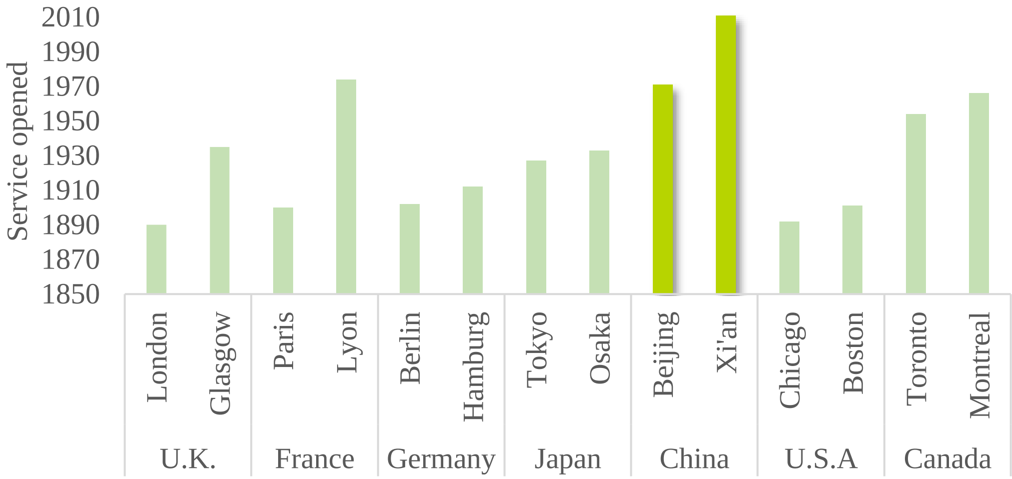


Figure 1- 1. Opening times of urban rail transit in different countries

1.1.2 Importance of research on spatial distribution patterns and interaction mechanisms in station realms

In the Global Urban Competitiveness Report published by UN-Habitat, it is stated that the main components of a city are population, activity, and space. The characteristics of urban space are mainly reflected in spatial aggregation, spatial connection, and spatial sharing [7]. Studies have shown that the impact of rail transit stations on the surrounding urban spatial pattern is mainly reflected in three aspects: spatial form, spatial distribution, and spatial density [8]. Therefore, since the morphology and density of station realm space have remained stable for some time after the station was built, considering that station realm space distribution is essentially an important stage in cognizing the interaction between population activities and urban functional facilities, it is particularly important to understand the spatial aggregation state and distribution characteristics of the station realm population activities and functional facilities through spatial distribution pattern analysis. At the same time, along with the continuous improvement and innovation of urban spatial analysis technology system, we need to carry out quantitative analysis of the spatial distribution pattern of station realm from theoretical development and practical application on a new operable platform, grasp the spatial and temporal changes in the distribution of population activities, determine the spatial correspondence between it and the layout of urban functions, provide quantitative basis for further revealing the interaction mechanism between population activities and functional facilities so as to improve the scientific and operable countermeasures for optimizing the spatial distribution pattern of station realm.

In-depth analysis shows that there is a close connection between population activities and urban functional facilities in the rail transit station realm, i.e. the completion of rail transit stations effectively improves the accessibility of the areas around the stations, thus attracting the gathering of population and functional facilities, along with the increase in population activities and the increasing demand for space for living, working and leisure, which in turn promotes the mixed and diversified development of urban functions in the station area[9,10,11]. However, in the current quantitative research on the relationship between population activities and functional facilities in station realms, there are problems such as complicated research model construction, high data demand, low precision, and long acquisition period. In particular, the activity behavior of human individuals themselves in time and space is very complex, which makes it difficult to study the spatial and temporal distribution characteristics of population activities at the microscopic scale. In recent years, with the development of information communication and positioning technology, urban planning analysis technology gradually moves from "small sample" to "large sample" research. Big data can provide more accurate spatiotemporal data of individual residents' behaviors from different perspectives, making it possible to study the distribution characteristics of population activities in the station area at the micro level. At the same time, the big data of Point of Interest (POI) of urban functional facilities based on electronic maps can accurately and objectively reflect the spatial layout of urban functional facilities, thus making it possible to conduct fine-grained quantitative research between population activities and functional facilities in the station realm. In addition, according to the characteristics of big data resources and geographic information analysis technology to form a complementary research approach, it can make the spatiotemporal characteristics of population activities more easily expressed in geographic space.

Therefore, based on geographic information system (GIS) combined with the use of big data analysis technology will become an effective way to identify the spatial distribution characteristics of population activities and functional facilities in the rail transit station realm, as well as to reveal the interaction mechanism between them.

1.1.3 The development status of urban rail transit in Xi'an

Xi'an is in the middle of the People's Republic of China and is one of the national central cities, with a population of 12,952,900[12]. Xi'an is not only the starting point of the Silk Economic Belt in China's "The Belt and Road Initiative" policy, but also an important city in the New Asia-Europe Continental Bridge and the Yellow River Basin, and the economic, scientific, and technological, educational, financial, cultural and commercial center of the Guanzhong City Group. In February 2018, China's National Development and Reform Commission released the Guanzhong Plain City Cluster Development Plan, which proposes to support Xi'an's development as an international metropolis with an international integrated transportation hub.

As of December 2022, there are 8 metro lines in operation in Xi'an, with a total operational mileage of 279 km; there are 176 stations, including 17 interchange stations (Figure 1- 2). In addition, Metro Line 1 Phase 3, Line 2 Phase 2, Line 8, Line 10, Line 15, and Line 16 are under construction. In June 2019, the National Development and Reform Commission's approval of the third phase of urban rail transit construction planning in Xi'an (2018-2024) stated that a rail network with 12 lines in operation and a total length of 423 km will be formed by 2024, while the proportion of rail transit to public transport trips will reach 60% [13]. Thus, the construction of urban rail transit in Xi'an has entered the stage of network development, which will greatly improve the travel conditions of residents, increase the accessibility of areas along the line, and strengthen the connection between the areas around each station point, thus having a significant impact on the development of urban space.

However, compared to international metropolises such as Shanghai, Hong Kong, and Tokyo where rail transit is well developed, there is still a big gap in urban rail transit in Xi'an. Not enough attention has been paid to the overall development of station areas. According to public data, the area covered by 500 meters of rail transit stations in Xi'an only accounts for 12.2% of the total area of urban construction land [14]. This indicates that the construction of rail transportation in Xi'an is lagging the current situation of urban development.

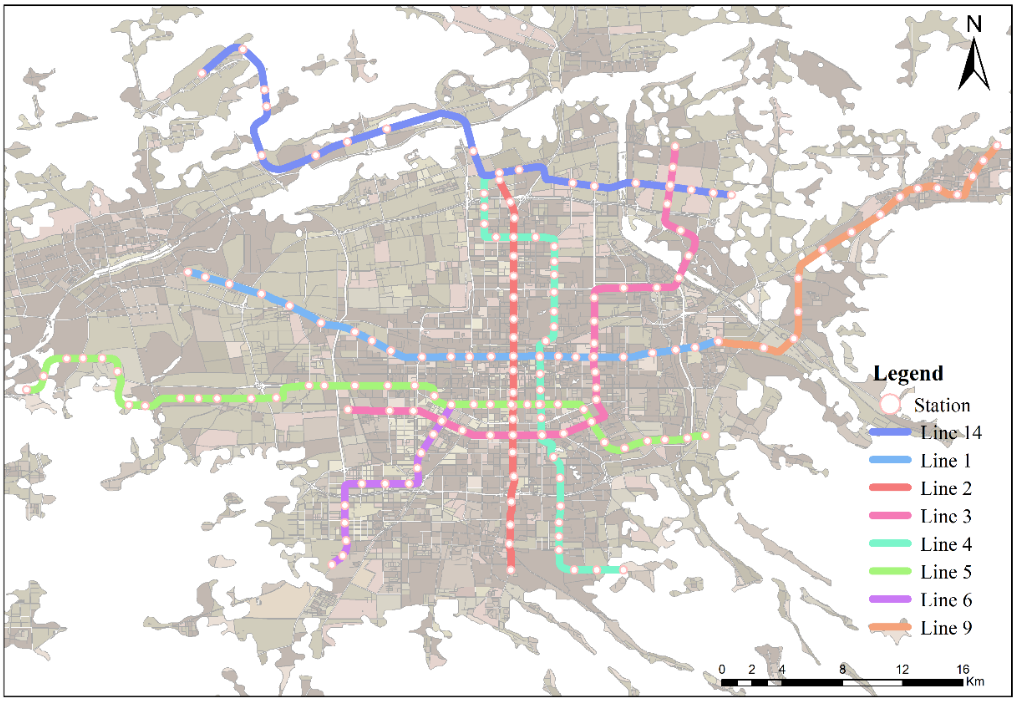


Figure 1- 2. Xi'an rail transit operation route map (As of November 1, 2022)

1.2 Literature review

The study of spatial distribution pattern and interaction mechanism of rail transit station realm belongs to cross-discipline, and its research focuses on several related fields such as urban geography and traffic geography, urban planning, and traffic planning. This paper introduces the research results of spatial distribution pattern and interaction mechanism of rail transit station realm at the present stage from three aspects: research progress, theoretical research, and methodological research.

1.2.1 Research trends in rail transit station realm

In this section, this study uses VOSviewer software [15] to visually analyze [16]the relevant literature in order to summarize the current research trends regarding rail transit stations. The bibliometric analysis followed a three-step approach: collecting literature data, analyzing literature data, and visualizing bibliometric associations. The literature data were analyzed mainly from the Web Of Science core collection database. The literature was searched by using the Web Of Science core collection database with the keywords "topic" = "rail transit station realm" or " Around rail transit station" was used as the keyword, and the period was selected from 2000 to 2022, and 2167 articles were retrieved. As can be seen from the Figure 1- 3, the number of articles issued has been maintaining a trend of year-on-year growth, especially in the latter year of 2017 ushered in a substantial increase, which indicates that the scale and quality of research on the rail transit station realm has continued to increase, and that the attention of research on the rail transit station realm has increased significantly since 2018.

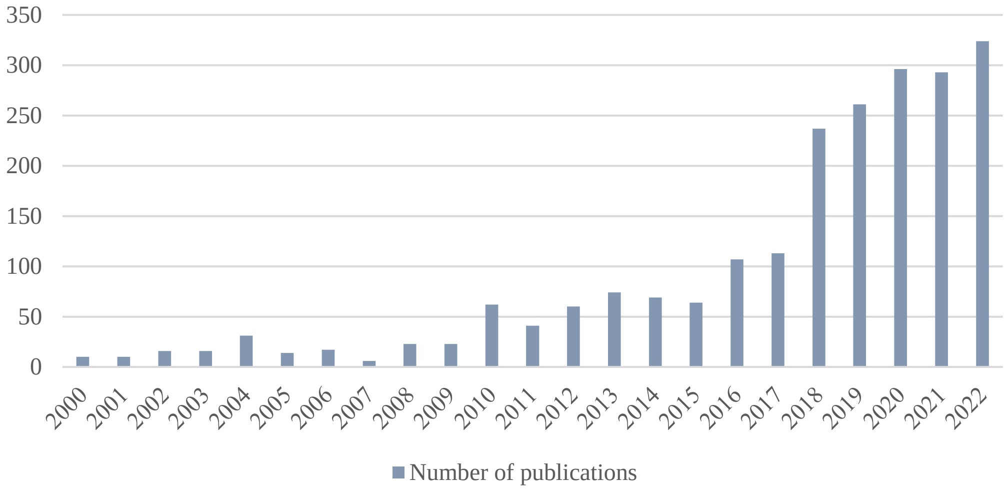


Figure 1- 3. Statistics on the number of articles published in the Web of science core database for the rail transit station realm

A total of 42,635 keywords (including author keywords and index keywords) used in the collected literature and 10 research studies were used as the minimum threshold for screening, i.e., a keyword must have been used in at least 15 research studies [17]. This minimum threshold yielded 415 eligible keywords for further cluster analysis. In the keyword clustering network, the nodes represent the keywords, while the size represents the frequency of the nodes in the literature and the color represents the different clusters, i.e., research results under the same cluster have thematic similarity. As shown in the Figure 1- 4, the keywords are distinguished by three different colors (red, green and blue), which indicate that they form three different clusters.

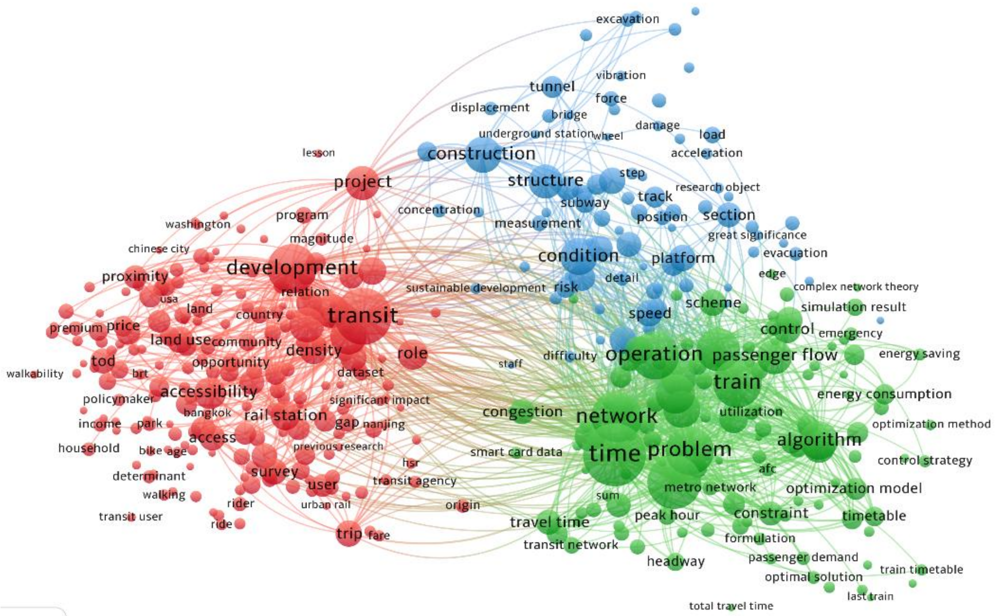


Figure 1- 4. Network visualization of keywords co-occurrence (weights=occurrences)

The specific results of the cluster analysis statistics are shown in Table 1- 1, according to the attributes of the keywords these three clusters can be interpreted as (a) red - station realm theory study; (b) green - station realm passenger flow study; (c) blue - -station realm planning and design studies. Among them, the red cluster includes keywords related to station realm theory, such as development concept, impact relationship, accessibility, etc. The green cluster focuses on the study of station realm passenger flow influencing factors, including time, passengers, line network, and operations. The blue cluster focuses more on the study of station realm planning and design, involving architecture, structures, indicators, platforms, etc.

Table 1- 1. Description of co-occurring keyword clusters

Cluster	Research theme	Total	Five most frequently used keywords	Occurrences	Average publication year	Average citations
Red	Theoretical	207	Transit	516	2017.4	10.91
			Development	392	2017.2	9.39
			Relationship	232	2018.3	11.90
			Transit station	217	2017.9	8.90
			Accessibility	183	2019.1	10.10
Green	Passenger flow	83	Time	606	2017.5	11.90
			Passenger	468	2017.7	11.05
			Network	415	2017.7	9.79
			Operation	386	2017.1	8.85
			Problem	384	2017.3	11.19
Blue	Planning and Design	125	Condition	239	2016.3	8.23
			Construction	231	2015.9	4.11
			Structure	205	2016.2	7.32
			Parameter	143	2018.3	8.32
			Platform	134	2016.7	10.01

The density analysis provides an understanding of the current research hotspots in the literature about the rail station realm. Each point on the plot of the density analysis is filled with a color based on the density of the elements surrounding that point; the greater the density secret, the closer it is to red; conversely, the lower the density, the closer it is to blue. The density size depends on the number of elements in the surrounding area and the importance of those elements. The density view can be used to quickly observe the important areas and the density of knowledge and research in each area. Figure 1- 5 shows the visualization results of the density analysis, and interaction, accessibility analysis, land use intensity, and passenger flow analysis are the research directions that have received the most attention in the field of rail transit station realm in recent years.

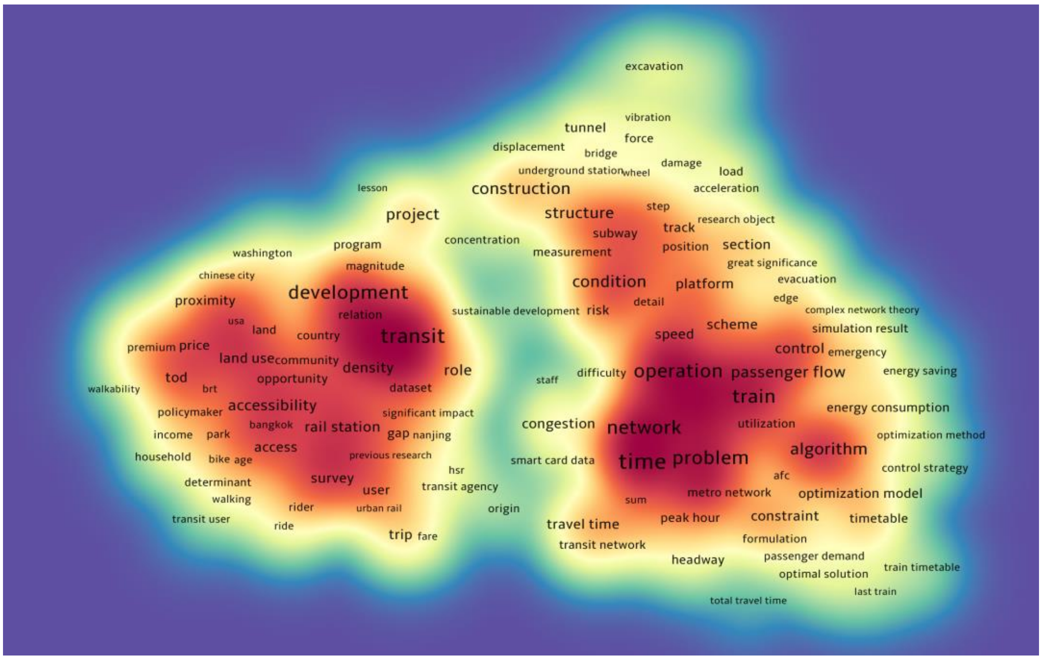


Figure 1- 5. Density analysis of keyword frequency

1.2.2 Review of theoretical research

(1) Theoretical study of spatial distribution pattern in station realm

The spatial distribution pattern is one of the significant characteristics of the rail transit station area, reflecting the distribution and organization of various spatial elements [18]. These spatial elements include population, buildings, functional facilities, land use, etc., and their spatial distribution pattern reflects geographical differences and regional characteristics [19]. Among them, C Mulders Kusumo (2005) investigated the retail layout and spatial configuration around railway stations in the cities of Delft and Leiden in the Netherlands [20]. According to Cai Guotian (2004), the construction and operation of rail transit not only affects the price of land along the route, but also affects the distribution of functional spaces such as retail and commercial [21]. Wang Xifu (2005) summarized and analyzed the land use types along Nanjing Metro Line 1, and concluded that rail transit has spatial and typological differentiation effects on land use of different nature [22]. Taking Shanghai's rail transit as an example, Wang Huimin (2006) studied and analyzed the functional spatial layout of the station areas along Shanghai's rail transit lines, pointing out that the residential functional space spreads outward in a bead chain pattern along the lines, and its influence on the residential space increases in scope as it gets farther from the city center [23]. Zemp S (2011) analyzed the spatial distribution characteristics of land use and transport services around railway stations in Switzerland to classify the 1,300 railway stations in Switzerland [24]. Liu Shiqi (2014) analyzed the land use structure, land utilization rate, information entropy, and equilibrium indexes within 2000m of each station and different ring zones in Beijing to reveal the spatial distribution characteristics of land use around different stations [25]. Wu Huaguo (2016) classified rail transit station areas into residential, commercial center, and transportation hub types according to their functional orientation, and explored the land use composition, development intensity, and spatial layout of station areas in a categorical manner [26]. Wu Yongxi (2017) categorized and identified the spatial pattern of productive service industries in Xi'an's rail transit station area by combining factors such as urban land development, infrastructure support, industrial environment and urban location around the rail transit stations [27]. Yuan Ming (2018) quantified the spatial distribution characteristics of the multi-level commercial space in the station area of rail transit stations in the core urban area of Shanghai [28]. Niu S (2019) analyzed the land use characteristics around rail transit TOD stations in Singapore and found that the city center hierarchy and the new town development concept led to differences in the land use characteristics of TOD stations [29]. Yu Yang (2022) analyzed the spatial distribution patterns of the clustering degree, spatial distribution, and business composition of station realm commercial spaces in three urban-level subway stations in Chengdu, including Chunxi Road, Tianfu Square, and Mule City, as examples [30].

(2) Theoretical study of the spatial interaction of the station realm

Most of the current studies focus on the interaction between land use intensity, land price, urban space, accessibility, and other factors and passenger flow in the station realm, while the interaction between population activities and urban functions is not sufficiently studied, especially under the time-varying characteristics of the population.

In terms of land prices, Cervero (1994) studied the positive impact of rail transit on office type real estate development in joint development (JD) and came to the basic conclusion that office rents in JD are directly proportional to the patronage of the rail system [31]. Gadziński J, Radzimski A (2016) A survey of residents and rent prices in the vicinity along the Rapid Tram in Poznań, Poland, and further analysis of the impact of the PST on travel behavior, housing choice and satisfaction, and apartment prices in the vicinity along the line, found that about 20% of respondents in close proximity to a PST station indicated that they would be willing to pay more for an apartment close to the station [32].

In terms of land use, Zhou, J. and Xu, J.G. (2002) studied the mutual influence relationship between residents' travel patterns and land use types in the station area of Shanghai Metro Line 3, pointing out that rail transit promotes the clustering of population around stations, which leads to the development of residential and commercial land, but shows rejection of industrial land, thus concluding that rail transit causes the divergence of land use properties [33].Pan H, Zhang M (2008) The impact of rail transit on land use in Shanghai reveals that high development intensity land use occurs in areas more accessible to residents near stations, and further elaborates on the impact on land prices around stations [34]. Calvo F (2013) illustrates the impact of the Madrid Metro on population settlement and land use around stations by analyzing the relationship between population, land use, and transportation systems in the vicinity of Madrid Metro stations in Spain [35]. Chan S (2013) studied the impact of intensity of different nature of sites on passenger flow and found that for every 10% increase in population density during morning peak hours, inbound passenger flow increased by 7.4%, and for every 10% increase in commercial and administrative office space, outbound passenger flow increased by 6%, and that increasing intensity, diversity, and level of connectivity to public transit are effective means of increasing passenger flow [36]. H Sung (2014) explored the impact of land use on ridership at rail stations in the Seoul Metropolitan Area through accessibility and found that land use intensity within the coverage area of 750 m radius of the station was positively associated with ridership, while land use diversity was not associated with ridership [37]. Wang Yajie (2018) studied the interaction between land use and passenger flow in the Beijing subway station realm and found that functional elements, locational elements, and socioeconomic elements all have an impact on the interaction between volume ratio and passenger flow [38].

In terms of accessibility, the Cervero study shows that residents around the subway are two to five times more likely to ride the subway than in other areas, suggesting that accessibility attracts more riders [39]. Guan Chenghe (2019) explored the spatial distribution of high-rise buildings in Shanghai and their relationship with public transportation development by examining the number of high-rise buildings within the 500-m buffer zone of metro stations and the time distance to these stations [40]. Cai Chaoyang (2019) studied the reciprocal feedbacks between accessibility and spatial distribution of land in the rail transit station area in Tianjin, where accessibility within the rail transit station area has a large impact on commercial and office space, with the most significant impact on the capacity of commercial space [41].

In addition, some studies also point out that intensity or mix elements do not have a significant impact on passenger flow. Wang (2012) conducted a regression analysis of passenger flow and land use elements at 30 metro stations in Shanghai and found that land use intensity and mix and accessibility did not have a significant effect on passenger flow, while bus service was more closely related to passenger flow [42]. By analyzing the effect of land use mix on subway station passenger flow, Junfang Li (2016) found that the land use mix of 109 stations in Tokyo, Japan, had a weak effect on passenger flow, while the ratio of number of jobs to residential population showed a significant positive correlation with outbound passenger flow [43]. Using Shanghai metro station passenger flow data, Jin Yu (2015) used a multinomial logit regression model to find that factors such as station area population, number of jobs, proportion of commercial land to commercial service facility land, number of bus stops around the station, and the number of years the station has been open significantly affect station passenger flow distribution patterns [44].

1.2.3 Review of methodological research

(1) Methodological study of spatial distribution pattern identification

Not only qualitative analysis but also quantitative analysis is needed in the study of spatial distribution pattern of rail transit station realm. Firstly, it is necessary to determine whether there is a pattern in the spatial distribution of elements, i.e., whether the spatial distribution is clustered, scattered, or random. Then it is also necessary to analyze the type of spatial aggregation (the aggregation of things of the same type, or the aggregation of things of different types), the location, size, and shape of the aggregation area, etc.; finally, it is also necessary to analyze the reasons for the aggregation. The literature shows that the widely used spatial pattern identification methods are mainly based on the spatial statistical tools provided by geographic information system software (ArcGIS), including Average Nearest Neighbor, multi-distance spatial clustering analysis (Ripley's K Function), kernel density analysis, global spatial autocorrelation (Moran's I, Getis-Ord General G), and local spatial autocorrelation (Anselin Local Moran's I, Getis-Ord Gi*), as shown in Table 1- 2.

In terms of spatial point pattern identification, Botte M and Olaru D (2012) used kernel density estimation to analyze household travel diaries collected within three unique TOD areas in Perth, Western Australia, revealing the spatial distribution of activities characterizing travel behavior within TOD areas [45]. Using the average nearest neighbor index, Bingqing Che (2020) analyzed the spatial point data and population data of public service facilities such as bus stops, banks, schools, and hospitals to reveal their spatial distribution characteristics [46]. Sui, Hongxin (2020) used multi-source data information data and the mean nearest neighbor of spatial statistical tools to measure the urban function facilities in the rail transit station realm, and the study showed that the urban functions in the rail transit station realm space showed a clustering distribution pattern [47]. Xing Zhaomin (2020) conducted a hotspot analysis of GPS location data within the metropolitan rail transit station areas using a kernel density analysis tool to reveal the spatiotemporal travel characteristics of residents [48]. Liu Wen (2022) used the Wujiaochang station area of Shanghai's rail transit as a research object to analyze the kernel density of POI data of land development in the past 10 years using ArcGIS platform, and further discussed the correlation between rail transit station area policies and land use [49].

In terms of the spatial distribution pattern of spatial autocorrelation, Mi Diao (2015) used spatial autocorrelation to analyze accessibility around the subway in the city of Boston and further assessed the impact of the property values of individual households in the subway system [50]. Laura Eboli (2018) rated the service satisfaction or importance of Milan rail stations and analyzed the spatial patterns of each attribute varying across regions and time using a spatial autocorrelation tool [51]. Xia Xuanxuan (2021) used global autocorrelation to analyze the spatial and temporal characteristics of population, economy, and spatial urbanization around the rail transit lines in Beijing [52]. Wijayanto Y (2022) using the Moran's Index putting on data PT KCI from 2014 through 2020, spatial data, and population data, the density level of Jabodetabek KRL stations and their routes as well as the spatial pattern of Jabodetabek KRL in Indonesia were investigated. [53].

Table 1- 2. Statistics on identification method of spatial distribution pattern of station realm

Technology Classification	Application Scenarios	Statistical indicators
Spatial point pattern identification	Determine the spatial distribution pattern of point things or phenomena.	Average Nearest Neighbor Ripley’s K
	Generate density distribution maps based on the statistical data of the elements to visualize their spatial distribution	Kernel Density Analysis
Spatial autocorrelation analysis	Whether a certain type of thing or phenomenon causes the appearance of similar or dissimilar things or phenomena in the vicinity, i.e., whether the space is autocorrelated.	Moran’s I Getis-Ord General G
	Find the spatial location of an unusual gathering of a certain type of thing or phenomenon in order to facilitate the analysis of the cause of the gathering.	Anselin Local Moran’s I Getis-Ord Gi^*

The above study found that spatial pattern analysis methods have been widely applied to rail transit station realms, and domestic and foreign scholars have used various spatial pattern analysis methods to identify the spatial distribution characteristics of elements such as residential activities, functional facilities, and land prices. Table 1- 2 summarizes the application scenarios and statistical indicators of various current spatial pattern analysis methods. The analysis methods and statistical indicators differ according to the research objects and data. Therefore, how to select the analysis methods and statistical indicators required for the study is the focus of this paper to establish the research framework of the spatial distribution pattern of station realms.

(2) Methodological study of data technology analysis

In terms of the choice of technical analysis methods, domestic and foreign scholars have mainly conducted correlation studies based on a large amount of travel data with land use and urban functional facilities data. The research idea is developed in three parts: "data collection and survey - correlation analysis of target factors - analysis of results and prediction or strategy". The technical analysis method of this study is based on travel data collection and survey, i.e., data on human mobility and precise geographical data on functional facilities. Traditional mobile data can no longer support the study of rail transit station realms at a more micro level. With the development of information and communication and location-based navigation technologies, analysis techniques are gradually moving from "small sample" to "large sample" studies. Big data provides important information for the study of residents' spatial activities from different perspectives, and it is possible to use cell phone data, public transportation data, social media data, etc. to obtain spatial and temporal data of residents' individual behaviors and use the data to study the distribution characteristics of population activities at the micro level. The literature review shows that the attention and the number of results on the relationship between the role of demographic activities and urban functions combined with big data analysis techniques have been on the rise in recent years.

Among them, Francesco Calabrese (2014) used mobile phone trajectory data to investigate the mobility patterns of mobile phone users within the Boston metropolitan area [54]. Mi Kyeong Kim (2017) used public transportation card data to identify the passenger flow distribution patterns and station influence zone characteristics of Seoul Metropolitan Area subway stations [55]. Xu Xinyue (2018) analyzed the route choice behavior of subway passengers based on the Automatic fare collection (AFC) data of Beijing Metro Line 1, Line 2, Line 4 and Line 6 [56]. Huang Jie (2018) calculated the travel time and OD matrix (Origin-Destination Matrix) of 4.31 million smart transit card data to study the temporal and spatial distribution characteristics of its passenger flow, taking the Beijing subway as an example [57]. Shen Lifan (2018) studied the relationship between reasonable walkable range-built environment and rail commuting at rail stations using subway swipe card data, taking 44 rail stations in Beijing as an example [58]. Ana Condeço-Melhorado (2020) used Twitter's geolocation data to analyze the activity patterns of Twitter users during the 2016 Rio Olympics [59]. Kovács, Z (2021) used Twitter's geolocation data to analyze user activity patterns to study the spatiotemporal characteristics of users' use of the city [60].

In addition, the use of electronic maps of points of interest (POI) to identify urban functional facilities has become one of the mainstream methods to study the functional layout characteristics of cities. Sun Shixi (2018) used an analysis of Baidu Map POI to characterize the land use composition around Beijing's rail transit stations [61].Shu Bo (2019) conducted a quantitative empirical study on the structural characteristics of urban functions along five subway lines in operation in Chengdu City by using the kernel density analysis and inductive comparative analysis method of ArcGIS software using Gode map POI data, and found that the distribution of urban functions around rail stations is not exactly in a circle pattern [62]. Pengjun Zhao (2021) used cell phone signaling data and POI data to determine the range of residents' living circles and the accessibility of service facilities, and analyzed the spatial matching relationship between them and their geographical differentiation characteristics, using Beijing as an example [63].

From the published literature, big data usually has the advantages of rich sources, large sample size, and real-time, which can objectively and truly reflect the characteristics of the object. In the study of the spatial distribution pattern and interaction of station realms, big data analysis methods can provide researchers with more detailed required data and can reflect the behavioral activity distribution of residents in the station area space and the preference degree of different functional facilities through the analysis of visualized data.

(3) Methodological study of spatial relationship model

At present, in order to clarify the relationship between the elements of the station realm researchers have used different spatial regression models to study, and mainly divided into two categories: global spatial regression models and local spatial regression models, as shown in Table 1- 3. The global spatial regression models mainly include Ordinary Least Squares (OLS), Spatial Error Model (SEM) and Spatial Lagged Model (SLM), among which OLS is the most commonly used regression method to assess the relationship between urban spatial elements [64,65]， and it is also the starting point of all spatial regression analysis [66]. OLS regression models have been widely used to study the interaction between rail stations and surrounding elements such as land price [67]、community structure [68]、infrastructure [69]. SEM is a spatial regression model based on a linear regression model that introduces a spatial error term to account for the effects of spatial dependence [70]. SLM introduces spatial lagged values to explain the variables [71]. Usually, SEM and SLM need to be used together to test the significance of both by building a spatial autocorrelation model, if the error value of the model is significant, choose SEM, if the Lagged value is significant, choose SLM, if neither is significant, then choose OLS.

In terms of local spatial regression models, they currently include geographically weighted regression (GWR), multiscale geographically weighted regression (MGWR), and geographically temporal weighted regression (GTWR). Among them, GWR is the basis of MGWR and GTWR, and is an important method for local analysis in modern spatial econometrics. GWR is a local regression method based on kernel density estimation and weighted least squares, focusing on explaining spatial non-stationarity and heterogeneity of explanatory variable parameters [72]. GWR is commonly used by most researchers because of its compatibility with discrete data, large data or real-time update data, and its ease of interpretation and operation [73]. For example, MF Dziauddin (2015) GWR estimated the impact of the Greater Kuala Lumpur LRT transit system on the value of residential properties along the route in Malaysia [74]. Shaoying Li (2020) used a GWR model to characterize the interaction between passenger flow and the built environment in the rail station realm by integrating multiple sources of spatial big data such as POI, social media, and building footprint data [75].

AS Fotheringham built the MGWR model by adding equations that can calculate the bandwidth at different scales to the GWR model [76]. This improves the performance and application of the GWR model, because classical GWR assumes that all modeled processes operate at the same spatial scale. Since MGWR incorporates multiple distances as an influencing factor, it is more suitable for studying geographical phenomena in the macro-spatial domain or the relationships between things that act on each other. GTWR is a GWR model based on temporal weights proposed by Huang B in 2010, which forms a temporal weight matrix by embedding temporal interval distances in the kernel function [77]， GTWR makes the research more reasonable and relevant, but it is not widely used due to its complex calculation process and the high accuracy of the data required, as well as the need for specific tools to assist the calculation.

Table 1- 3. Analysis of the advantages and limitations of the spatial relationship model

Model	Running Logic	Advantages	Limitations
OLS	Find the best functional match for the data by minimizing the squared sum of errors	Easy to operate. Highly interpretable Wide range of applications Low computational cost	Assumption Limits Sensitive to outliers Multiple covariance
SEM	First calculate the error values of the spatial units adjacent to the parameters, and then use generalized least squares to estimate the other parameters	Predictions can be made on the data Considers spatial correlation Model stability	Complex models Assume many constraints Cannot handle highly autocorrelated data
SLM	First calculate the lagged values of the spatial units adjacent to the parameters, and then use generalized least squares to estimate the other parameters	Consider spatial correlation Accurate parameter estimation Capable of estimating heterogeneity	High data requirements Complex models High computational costs
GWR	Using the different spatial locations of each element, calculate the bandwidth, then calculate the weight value of the spatial location of each element, and then use the regression equation to calculate	High accuracy of results Consider the spatial location properties Visualization of results	Not applicable to small data sets Cannot process multi-point data
MGWR	Improve GWR by calculating the bandwidth of each independent variable	Improved model accuracy Suitable for large geographic scales	Requires auxiliary tool High computational costs Limited interpretability
GTWR	Based on GWR, regression analysis was performed on the elements of each time period	The weight of time variation in geographic space is considered	Requires auxiliary tool High computational cost Limited interpretability

In recent years, quantitative statistical analysis and multiple regression models have dominated the analysis of the interrelationship between different elements in the study of rail transit station realm, and simple comparative analysis has been gradually withdrawn, which also reflects the advancement of the research related to the travel patterns of station realm residents and the characteristics of urban built environment from simple to complex and from local to comprehensive processes.

1.2.4 Summary

The study of the spatial distribution pattern and interactive interrelationship of rail transit station realms is multi-level, systematic and interdisciplinary in character.

From the theoretical research, scholars have conducted extensive and in-depth theoretical research on the spatial distribution pattern of rail transit station realm in terms of land use characteristics, industrial distribution characteristics and individual city function spatial distribution characteristics, but few studies have identified and integrated different categories of city function spatial distribution characteristics. At the same time, the current research on the interrelationship between the elements of station realms is mainly carried out in two directions: passenger flow and land use, and the research paradigm follows a combination of theory and empirical evidence, mainly using indicators such as land price, land use function and land use intensity to study the interaction with passenger flow. The study also found that the public space around the rail station is a special space dominated by the activity trajectory of residents, and people are the main actors of activities in the public space [78]. The characteristics of residents' activities along the rail line correspond to the layout of urban functions, and the residents' activities in a specific time and space can produce some corresponding urban functional use characteristics [79]. At present, most studies tend to focus on the temporal distribution characteristics of passenger flows, and do not explore the spatial and temporal distribution characteristics of population activities within the station realm in depth.

In terms of methodological research, among the spatial pattern identification methods, many scholars have used the spatial pattern analysis tools provided by ArcGIS to identify the spatial distribution characteristics of elements such as station realm land use, urban functional facilities, and passenger flow according to the type of research data. And the big data analysis technology based on multi-source heterogeneous data has been a research hotspot in recent years, forming a complementary research approach based on the characteristics of different data resources, and multi-source heterogeneous data makes the spatiotemporal behavioral characteristics of population activities easier to express in space. In the spatial regression models, quantitative statistical analysis and multiple regression models dominate, and simple comparative analysis has been withdrawn.

From the empirical cases, the spatial distribution patterns, and interactions of rail transit stations in different cities have different characteristics because the planning and construction of urban rail transit are influenced by multiple aspects such as geographical location, policy guidance and technical conditions. At present, most of the research results are mainly focused on rail transit in developed or large cities. Especially in China, the research mainly focuses on the rail transit systems of large cities in the southeast region, such as Beijing, Shanghai and Guangzhou. And there are few in-depth studies on the rail transit station realm of cities in the northwest region. Therefore, the research result of this paper, which takes Xi'an Metro Line 2 as the research object, not only fills the gap in the theoretical system of urban rail transit station realm research in China, but also has important significance for the planning and development of urban rail transit station realm in northwest China and along the " One Belt and One Road" in the future.

1.3 Research question and purpose

1.3.1 Research question

This study focuses on the following three questions：

(1) How to identify the distribution pattern of population activities and functional facilities within the station realm at the micro level?

(2) What is the degree of correlation between population activity and different categories of facilities within the station realm? What categories of facilities have a positive correlation with population activity?

(3) What are the interaction mechanisms between population activities and different categories of facilities within station realm? How to improve the allocation of public resources in the station realm through interaction mechanisms?

1.3.2 Research purpose

The research purpose of this paper are mainly the following three points:

(1) To analyze the spatial distribution status and locations of population activities and functional facilities within station realm. Based on geographic information systems (GIS), utilizing multi-source big data and multiple spatial analysis methods, the aim is to identify the spatial distribution characteristics of population activities and functional facilities within the station realm at the micro level, thus contributing to the enrichment of the theoretical system of spatial-related research in the station realm.

(2) To determine the degree of correlation between population activities and various facilities within station realm, as well as establish the variable system for spatial relationship model. By analyzing the degree of correlation between population activities and facilities such as catering, shopping, and transportation in the station realm, the aim is to establish a reliable system of variables for the studying the interaction mechanism between population activities and functional facilities.

(3) To explore the interactive relationship between population activities and various facilities within station realms, and to propose development suggestions for different facilities based on the degree of interaction. The degree of interaction between population activities and various facilities in different types of station realms was obtained through regression models, and a sequence of interactions was established, thus providing a basis for improving the allocation of public resources in station realm.

1.4 Research content and framework

This paper, supported by research background, literature review, and relevant theoretical studies, establishes a research framework for the spatial distribution patterns of urban rail transit station areas. It employs a complementary strategy of big data and traditional survey data, as well as a combination of qualitative and quantitative analysis techniques. Furthermore, taking Xi'an Metro Line 2 as the empirical object, it explores the spatial distribution patterns of station realms and summarizes the distribution models of these areas. It further reveals the interaction mechanisms between "population activities" and "functional facilities". The goal is to comprehensively understand the characteristics and patterns of station area spatial distribution from a micro perspective, provide decision-making basis for urban managers and designers to improve the allocation of public resources and the living environment in station areas, and achieve the sustainable development of urban space supported and guided by rail transit. The research content of thesis is listed below, with the research framework is shown in Figure 1- 6.

In Chapter 1, focuses on the research background, literature review, research questions and purposes, as well as the content and technical route of the research.

In Chapter 2, the interpretation of core concepts, clarify the core elements that define the scope of the station realm and determine the classification system that delineates the type of station realm. The theoretical studies related to the spatial distribution pattern and interaction, and a comparison of the characteristics of station realm development modes.

In Chapter 3, the research methodological framework is constructed. It mainly includes the method of defining the scope and type of the station realm, the method of collecting data on population activities and functional facilities, the method of analyzing the spatial distribution pattern of the station realm, and the regression model selected to study the interaction mechanism.

Chapter 4 is the first part of the empirical study, and 16 stations of Xi'an Metro Line 2 are selected as the research objects. Firstly, the scope and type of each research station realm are defined based on the actual survey data. Secondly, the spatial distribution characteristics of population activities and functional facilities in each station realm were analyzed using location-based information data. Finally, the location relationship between them was explored.

Chapter 5 is the second part of the empirical study, which analyzes the correlation between population activities and various categories of functional facilities in each research station realm and establishes a system of variables through the significance of the correlation coefficients.

Chapter 6 is the third part of the empirical study, which mainly utilizes the ordinary least squares (OLS) model and the geographically weighted regression (GWR) model to conduct regression analysis on the constructed variable system, so as to analyze the degree of interaction between the population activities and different functional facilities in the station realm, and to establish the interaction mechanism. Meanwhile, a detailed discussion around the main findings of the study is presented, and further development suggestions are made to improve the allocation of public resources in the station realm.

In Chapter 7, summarizes the main conclusions of this study and suggests research directions for future research.

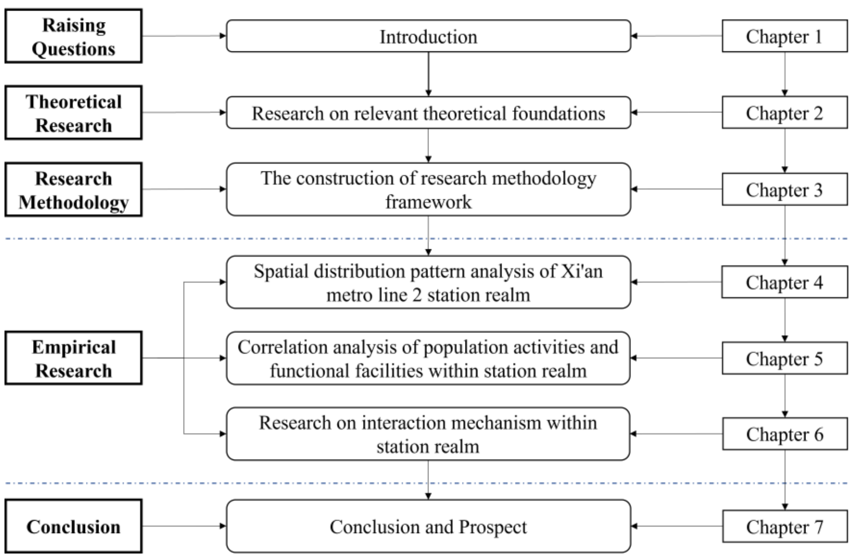


Figure 1- 6. Research framework

References

[1] The urban rail development handbook[M]. World Bank Publications, 2018.

[2] Dong ZF, Liu XM. Research progress on the impact of rail transportation on urban spatial structure in China[J]. Settlement, 2017 (4): 136-139.

[3] Gu Min. Current situation and outlook of urban rail transit development in China[J]. China Railway, 2011 (10): 53-56.

[4] Jin H, Jin F, Zhu H. Measuring spatial mismatch between public transit services and regular riders: a case study of Beijing[J]. ISPRS International Journal of Geo-Information, 2019, 8(4): 186.

[5] Sun Q, Hao H E, Sun L. Research on Supply and Demand Matching of Urban Public Transport Based on IC Card and GPS Data[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2020, 587(1): 012025.

[6] Wang B, Liu C, Zhang H. Where are equity and service effectiveness? A tale from public transport in Shanghai[J]. Journal of Transport Geography, 2022, 98: 103275.

[7] Marco Camia, Pengfei Ni. Global Urban Competitiveness Report (2019-2020). UN - Habitat,2020.

[8] Wang Di. Investigation on the Spatial Structure of Xi'an Metro Line 2 Station Area[D]. Master's thesis. Chang'an University, 2020.

[9] Wu Yongxi. Research on the spatial pattern of producer services in Xi'an rail transit stations [D]. Master's thesis. Xi'an University of Architecture and Technology, 2017.

[10] Bo Shu, Yang Chen, Jin Cui. Preliminary Study on the City's Functional Structure of Subway Station's Surrounding Area Under the TOD Model: Empirical Analysis Based on POI Data Along Chengdu Metro [J]. Huazhong Architecture, 2019, 37(5): 79-83.

[11] AlQuhtani S, Anjomani A. Do rail transit stations affect the population density changes around them? The case of Dallas-Fort Worth Metropolitan Area[J]. Sustainability, 2021, 13(6): 3355.

[12] Shaanxi Provincial Bureau of Statistics. 2021. Available online: http:// tjj.shaanxi.gov.cn/ tjsj/ndsj/tjgb/ qs_444/202105/t20210528_2177393.html (accessed on 1 June 2021).

[13] National Development and Reform Commission of China. "Guiding Opinions on the Priority Development of Public Transportation in Cities." State Development [2012] No. 64.

[14] Xi'an TOD site big data release. https://mp.weixin.qq.com/s/W3MtDyH9973kxfi88m4wIQ

(accessed on 20 September 2022).

[15] Centre for Science and Technology Studies. VOSviewer[J]. 2018.

[16] Van Eck N J, Waltman L, Dekker R, et al. A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS[J]. Journal of the American Society for Information Science and Technology, 2010, 61(12): 2405-2416.

[17] Wang P, Zhu F, Song H, et al. A bibliometric profile of current science between 1961 and 2015[J]. Current Science, 2017: 386-392.

[18] Bertolini L. Spatial development patterns and public transport: the application of an analytical model in the Netherlands[J]. Planning practice and research, 1999, 14(2): 199-210.

[19] Huang Yaping, Urban Planning. Urban Spatial Theory and Spatial Analysis [M]. Southeast University Press, 2002.

[20] Mulders-Kusumo C. Is a railway station a central urban place? Spatial configuration study of retail distribution pattern around railway stations[C]//5th International Space Syntax Symposium. 2005: 201-210.

[21] Cai Guotian. Study on the Spatial Impact of Rail Transit Construction on Retail Commercial Activities in Guangzhou [D]. Doctor's thesis. Guangzhou: South China Normal University, 2004.

[22] Wang XF, Xu JG, Li YF. A study of land use differentiation under the potential impact of urban rail transit construction in Nanjing[J]. Human Geography, 2005(3): 112-116.

[23] Wang Huimin, Zhang Zhenguo, Lin Tao. A study on the spatial distribution characteristics of Shanghai residences under the influence of rail transportation[J]. Modern Urban Research, 2006(7): 65-70.

[24] Zemp S, Stauffacher M, Lang D J, et al. Classifying railway stations for strategic transport and land use planning: Context matters! [J]. Journal of transport geography, 2011, 19(4): 670-679.

[25] Liu Shiqi,Guo Jing,Li Ruoxi,et al. Analysis of land use characteristics around typical stations of rail transit in Beijing[J]. Urban Development Research,2014,21(4):66-71.

[26] Wu Huaguo. Impact of urban rail transit stations on land use types around stations from the perspective of TOD[J]. Urban Geography. 2016.(4):92.

[27] Wu Yongxi. Research on the spatial pattern of producer services in Xi'an rail transit stations. Master's degree, Xi'an University of Architecture and Technology, Xi'an, 2017.

[28] Yuan M. Analysis of commercial space distribution characteristics and influencing factors in

the station area of rail transit stations in core urban areas of Shanghai[J]. Urban Rail Transit Research,2018,21(07):1-4+9.

[29] Niu S, Hu A, Shen Z, et al. Study on land use characteristics of rail transit TOD sites in new towns—taking Singapore as an example. Journal of Asian Architecture and Building Engineering, 2019, 18(1), 16-27. DOI: 10.1080/13467581.2019.1586712

[30] Yu Yang, Zhou R, Wu Bingxiao, et al. Evolution Mechanism and Optimization Path of Commercial Space in Metro Station Area under TOD Guidance:Three Cases in Chengdu [J]. Planner, 2022, 4: 107-114.

[31] R.Cervero. Rail Transit and Joint Development: Land Market Impacts in Washington, D.C. and Atlanta[J]. Journal of the American Planning Association, 1994, 60(1):83-94

[32] Gadziński J, Radzimski A. The first rapid tram line in Poland: How has it affected travel behaviours, housing choices and satisfaction, and apartment prices? Journal of Transport Geography, 2016, 54, 451-463. DOI: 10.1016/j.jtrangeo.2015.11.001

[33] Zhou J, Xu JG. Corridor effect of rail transit and urban land use analysis--Take Shanghai Rail Transit Mingzhu Line (Phase I) as an example[J]. Urban Rail Transit Research, 2002(1): 77-81.

[34] Pan H, Zhang M. Rail transit impacts on land use: Evidence from Shanghai, China. Transportation Research Record. 2008, 2048(1), 16-25. DOI: 10.3141/2048-03

[35] Calvo F, de Oña J, Arán F. Impact of the Madrid subway on population settlement and land use. Land use policy.,2013, 31, 627-639. DOI： 10.1016/j.landusepol.2012.09.008

[36] Chan S, Miranda-Moreno L. A station-level ridership model for the metro network in Montreal, Quebec[J]. Canadian Journal of Civil Engineering, 2013, 40(3): 254-262.

[37] Sung H, Choi K, Lee S, et al. Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership. Journal of Transport Geography, 2014, 36, 134-140. DOI：10.1016/j.jtrangeo.2014.03.013

[38] Wang Yajie. Research on the Interaction between Land Use of Station Areas and Passenger Flow of Subway in Beijing [D]. Doctor’s thesis. Tsinghua University, 2018.

[39] Cervero R. Transit-based housing in California: evidence on ridership impacts[J]. Transport Policy, 1994, 1(3): 174-183.

[40] Guan Chenghe. Spatial distribution of high-rise buildings and its relationship to public transit development in Shanghai[J]. Transport Policy, 2019, 81: 371-380.

[41] Chen Shaopei. Research on the Collaborative Development of Regional Accessibility and Spatial Distribution of Tianjin Rail Transit Station [D]. Doctor’s thesis. Tianjin University of Architecture and Technology, 2019.

[42] Zhao Zhaosheng, Wang Miao. Analysis of the evolutionary characteristics of Shanghai's rail transit network[J]. Urban Bridges and Flood Control, 2012 (9): 1-5.

[43] Li JF, Yao MF, Ji F, et al. Impact of land use mix on passenger flow of rail transit stations[J]. Journal of Tongji University: Natural Sciences Edition, 2016, 44(9): 1415-1423.

[44] Jin Yu. Research on the time-varying characteristics of passenger flow in urban rail transit stations and its influencing factors--Shanghai as an example[J]. Modern Urban Research, 2015 (6): 13-19.

[45] Botte M, Olaru D. Geo-spatial analysis of activity spaces in a TOD environment-Tracking impacts of rail transport policy using kernel density estimation[J]. Road & Transport Research: A Journal of Australian and New Zealand Research and Practice, 2012, 21(1): 64-81.

[46] CHE Bingqing, DU Yan, JIAN Xiaobin. Different Service Levels of Urban Bus Stops in Jiangsu Province [J]. Areal Research and Development,2020,39(05):53-57.

[47] Sui Hongxin, Wang Jianghao, Liu Yinxi, Deng Yu. Urban Functional Spatial Succession Trajectory and Determination of Renewal Mode in Rail Transit Station Area——Multiple Cases Comparisons and Sustainable Enlightenment [J]. Remote Sensing Technology and Application, 2020,35(02):302-314.

[48] Xing Zhaomin. Analysis on Spatiotemporal Characteristics of Combined rail Transit and DiDi Journey Based on Big Data Analyzing ——A Case Study of Chengdu City [D]. Master’s thesis. Shandong University of Technology, 2020. DOI: 10.27276/d.cnki.gsdgc.2020.000263.

[49] Liu W, Qiu Y, Lou H. The policy and practice of land development in rail transit station area based on ARCGIS take Shanghai Wujiaochang as an example[C]//2022 International Conference on Big Data, Information and Computer Network (BDICN). IEEE, 2022: 292-300.

[50] Diao M. Selectivity, spatial autocorrelation and the valuation of transit accessibility[J]. Urban studies, 2015, 52(1): 159-177.

[51] Eboli L, Forciniti C, Mazzulla G. Spatial variation of the perceived transit service quality at rail stations[J]. Transportation Research Part A: Policy and Practice, 2018, 114: 67-83.

[52] Xia X, Li H, Kuang X, et al. Spatial–Temporal Features of Coordination Relationship between Regional Urbanization and Rail Transit—A Case Study of Beijing[J]. International journal of

environmental research and public health, 2022, 19(1): 212.

[53] Wijayanto Y, Fauzi A, Rustiadi E. Spatial Patterns Analysis of Jabodetabek Electric Rail Transportation Using Spatial Autocorrelation Approach[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2022, 950(1): 012082.

[54] Calabrese F, Diao M, Di Lorenzo G, et al. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example[J]. Transportation research part C: emerging technologies, 2013, 26: 301-313.

[55] Kim M K, Kim S P, Heo J, et al. Ridership patterns at subway stations of Seoul capital area and characteristics of station influence area[J]. KSCE Journal of Civil Engineering, 2017, 21: 964-975.

[56] Xu X, Xie L, Li H, et al. Learning the route choice behavior of subway passengers from AFC data[J]. Expert systems with applications, 2018, 95: 324-332.

[57] Huang J, Wang J E, Jin H T, et al. 2018. Investigating spatiotemporal patterns of passenger flows in the Beijing metro system from smart card data[J]. Progress in Geography, 37(3): 397-406. DOI: 10.18306/dlkxjz.2018.03.010

[58] Shen Lifan, Wang Ye, Zhang Chun, et al. Study on the relationship between the built environment and rail commuting within reasonable walking accessibility of rail stations--Example of 44 rail stations in Beijing[J]. Journal of Geography,2018,73(12):2423-2439.

[59] Condeço-Melhorado A, Mohino I, Moya-Gómez B, et al. The Rio olympic games: A look into city dynamics through the lens of Twitter data[J]. Sustainability, 2020, 12(17): 7003.

[60] Kovács, Z.; Vida, G.; Elekes, Á.; Kovalcsik, T. Combining Social Media and Mobile Positioning Data in the Analysis of Tourist Flows: A Case Study from Szeged, Hungary. *Sustainability* **2021**,*13*, 2926. https://doi.org/10.3390/su13052926

[61] Sun Shixi. Research on the Dynamic Development of Beijing's New Town Occupation and Residence Space under the Growth of Rail Transit Network [D]. Master's thesis. Northern Polytechnic University, 2018.

[62] Shu Bo, Chen Yang, Cui Jin, Chen Lusong. Preliminary Study on the City's Functional Structure of Subway Station's Surrounding Area Under the TOD Model: Empirical Analysis Based on POI Data Along Chengdu Metro [J]. Huazhong Architecture,2019,37(05):79-83. DOI: 10.13942/j.cnki.hzjz.2019.05.020

[63] ZHAO Pengjun, LUO Jia, HU Haoyu. Spatial match between residents' daily life circle and public service facilities using big data analytics:A case of Beijing [J]. Progress in

Geography,2021,40(04):541-553.

[64] Dark S J. The biogeography of invasive alien plants in California: an application of GIS and spatial regression analysis[J]. Diversity and Distributions, 2004, 10(1): 1-9. DOI:10.1111/j.1472-4642.2004. 00054.x

[65] Mahara G, Wang C, Yang K, et al. The association between environmental factors and scarlet fever incidence in Beijing region: using GIS and spatial regression models. International journal of environmental research and public health, 2016, 13(11): 1083. DOI: 10.3390/ijerph13111083

[66] Hamilton L C. Regression with graphics: A second course in applied statistics[J]. 1992.

[67] Zhong H, Li W. Rail transit investment and property values: An old tale retold[J]. Transport Policy, 2016, 51: 33-48.

[68] Kahn M E. Gentrification trends in new transit-oriented communities: Evidence from 14 cities that expanded and built rail transit systems[J]. Real estate economics, 2007, 35(2): 155-182.

[69] Baum-Snow N, Kahn M E, Voith R. Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000 [with comment] [J]. Brookings-Wharton papers on urban affairs, 2005: 147-206.

[70] Baltagi B H, Song S H, Koh W. Testing panel data regression models with spatial error correlation[J]. Journal of econometrics, 2003, 117(1): 123-150.

[71] Lambert D M, Brown J P, Florax R J G M. A two-step estimator for a spatial lag model of counts: Theory, small sample performance and an application[J]. Regional science and urban economics, 2010, 40(4): 241-252.

[72] Fotheringham A S, Brunsdon C. Local forms of spatial analysis[J]. Geographical analysis, 1999, 31(4): 340-358.

[73] Wang Yuanfei, He Honglin. Spatial data analysis methods [M]. Science Press, 2007.

[74] Dziauddin M F, Powe N, Alvanides S. Estimating the effects of light rail transit (LRT) system on residential property values using geographically weighted regression (GWR)[J]. Applied Spatial Analysis and Policy, 2015, 8: 1-25.

[75] Li S, Lyu D, Huang G, et al. Spatially varying impacts of built environment factors on rail transit ridership at station level: A case study in Guangzhou, China[J]. Journal of Transport Geography, 2020, 82: 102631.

[76] Fotheringham A S, Yang W, Kang W. Multiscale geographically weighted regression (MGWR)[J]. Annals of the American Association of Geographers, 2017, 107(6): 1247-1265.

[77] Huang B, Wu B, Barry M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices[J]. International journal of geographical information science, 2010, 24(3): 383-401.

[78] Zhuo Xuan. Study on Vitality of Public Space in the Vicinity of Rail Transit Stations Based on Space Syntax——Take Xuzhou Rail Transit Line 1 as an Example. Master’s degree. China University of Mining and Technology, Jiangsu Province, 2018.

[79] C.Baker, Trackong Washington's metro [J]. American Demographics, 1983, (11): 30-35, 46.

Chapter 2. Research on relevant theoretical foundations

Content

2.1 Core concept identification

2.1.1 Urban rail transit

Urban rail transit refers to a rail transportation system constructed, operated, and serving the urban areas or surrounding regions, using railway technology, powered by electricity [1]. It facilitates urban residents' travel and follows specified routes, stations, operating hours, and fares. Compared to other modes of transportation, urban rail transit exhibits characteristics such as high capacity, high speed, high efficiency, low pollution, and low noise. It is an essential component of modern urban public transportation systems and a symbol of urban modernization and sustainable development.

In 1863, the world's recognized first underground railway, the "London Metropolitan Railway," was opened, marking the birth of urban rail transit [2]. In the early and mid-20th century, electrically powered trains rapidly replaced steam locomotives, propelling urban rail transit into a phase of sustained and rapid development. However, from the mid to late 20th century, due to the development of the automobile industry, wars, massive investments, and other factors, the development of rail transit stagnated. It wasn't until after World War II when the excessive growth of private automobiles, urban sprawl, and increasing traffic congestion became pressing issues that urban rail transit once again received attention, leading to the construction of rail transit systems in many cities.

From 1925 to the 1940s, some major cities in Europe and Japan had already achieved large-scale and networked operations of rail transit. In the 1990s, cities in Asia and South America also experienced rapid development in urban rail transit. Today, urban rail transit has evolved into a state of diverse coexistence with multiple types. Currently, China is also in a period of rapid development in urban rail transit construction.

There are various types of urban rail transit systems worldwide, and due to the different aspects of rail transit systems or lines in different regions, the classification criteria for urban rail transit also vary. This paper, based on the "Classification Criteria for Urban Public Transportation," combines the operational characteristics, technological features, and other factors of various rail transit modes. As shown in Table 2- 1, the urban rail transit was classified into seven categories: Metro Rail Transit (MRT), Light Rail Transit (LRT), Monorail Transit, Tramway, Maglev Train, Automatic Guideway Transit (AGT), and Suburban Rail Transit (SRT) [3].

As shown in Table 2- 2, among the 9,018.3 km of urban rail transit lines in operation throughout China in 2021, the three main types of rail transit are as follows: metro lines cover 7,305 km, suburban rail covers 577 km, and tram lines cover 543.6 km [4]. This indicates that subway systems hold a dominant position in the urban rail transit systems of major cities in China and are also the primary mode of operation in the current urban rail transit system in Xi'an city.

In conclusion, based on the current usage of urban rail transit in China and the consistency with the topic of discussion in this study, the term "urban rail transit" specifically refers to metro systems serving the urban areas. However, other types of rail transit will be mentioned and discussed in the case studies presented in this paper.

Table 2- 1. Classification and comparison of urban rail transit in China

Classification	MRT	LRT	Monorail Transit	Tramway	Maglev Train	AGT	SRT
Type	Underground, elevated and above ground	Fully enclosed overhead, above ground or underground	Closed or dedicated lane elevated	Closed surface lane	Enclosed or semi-enclosed elevated	Fully enclosed underground or elevated	Underground or elevated, fully enclosed
Construction cost	High	Medium	Low	Low	High	Medium	High
Capacity (Million passengers/h)	4.5～7	1～3	1～3	⩽1	1.5～3	⩽1	2.5～3.5
Operating Speed (km/h)	⩾35	25～35	30～35	23～25	60～85	⩾25	120～160
Formation (Vehicle)	4～10	2～6	4～6	1～2	4～10	2～6	⩽10

Table 2- 2. Operating mileage statistics of China urban rail transit in 2021 [4]

Operating Mileage (Km)	MRT	LRT	Monorail Transit	Tramway	Maglev Train	AGT	Suburban Rail Transit
China	7305	285.6	143.8	543.6	58.7	104.6	577
Xi'an	228.2	--	--	--	--	--	29.3

2.1.2 Rail transit station influence realm

With the gradual integration of rail transit lines and stations into cities, a series of new phenomena and rules have emerged, along with a range of urban spaces influenced or radiating from these stations. The areas surrounding these stations have become hotspots of urban development and focal points of public activities, playing a crucial role in supporting the functioning of urban spaces. Currently, there is no unified terminology or definition in existing international research. As shown in the Table 2- 3, scholars from various fields such as transportation, urban planning, and economic geography have proposed different definitions from different perspectives.

In summary, from the perspective of urban planning and transportation planning, this paper uses the term "rail transit station realm" to describe this specific urban space, referred to as "station realm." It is defined as "the urban space centered around rail transit stations, where human activities and urban elements are influenced by the proximity to these stations within a certain range." It represents a specific territorial concept within the city. The size and scope of the rail transit station's influence realm are of great significance for urban transportation planning and development. The detailed explanation of how to define the station realm will be discussed in Section 2.2.

Table 2- 3. Related expressions of the rail transit station realm

Perspective	Definition Title	Meaning
Transportation Perspective	Traffic attraction area (Liu Huimin,2003)	1.Walking accessibility of rail stations.
	Passenger flow radiation area (Guo Peng,2007)	
	Passenger connection area (Dai Jie,2009)	
	Coverage area (Wibowo,2010)	
	Catchment area (Guerra et al， 2012)	2.The passenger flow generation area created by the station.
	Pedestrian Catchment Area (Zhao,2013)	
	Last mile area (Nie Lei,2013)	
	Service area (El-Geneidy,2014)	
Urban Development Perspective	TOD area (Calthorpe,1993)	Urban space where population, functional layout, land use and other urban elements are affected by a certain range of the station.
	Transit village (Bernick et al,1997)	
	TOD rational area (Guo Tao,2008)	
	Metro station area (Huang Jun,2008)	
	Station area (Chorus et al,2011)	
	Influencing realm of rail station (Zhu Dongzhu,2015)	
Economic Perspective	TJD area (Cervero et al,2002)	The area with the greatest range of market returns or where the station has a greater impact on surrounding property values.
	Metro impact area (Liu Gui Wen,2007)	

2.1.3 Urban spatial distribution pattern and interactive relationship

DL. Foley (1964) argued that urban space should be understood from two aspects: form and process. Form refers to spatial distribution patterns, while process refers to the mechanisms of spatial functioning [5].

Based on Foley's conceptual framework, MM. Webber (1964) proposed three elements of urban space: material elements, which involve the positional relationships of physical elements; activity elements, which pertain to the spatial distribution of various human activities; and interaction elements, which encompass various flows in the city, such as human flow, goods flow, and traffic flow [6]. Therefore, the form of urban space refers to the spatial distribution patterns of material and activity elements, while the process refers to the interactions between these elements.

According to Huang Yaping et al. (2002), urban space serves as a carrier and container that supports and accommodates urban activities. Urban space encompasses the distribution and interactions of all urban elements (both material and non-material) within the urban region and constitutes a dynamic system or collection that evolves over time [7].

(1) Urban Spatial distribution pattern

The study of spatial distribution patterns plays a crucial role in urban planning and management as it provides scientific evidence and theoretical support. Through the analysis of spatial distribution patterns, a deep understanding of urban spatial structure, functional distribution, and development trends can be achieved. This knowledge can then be used to formulate targeted planning and management measures to promote sustainable urban development. Spatial distribution patterns are a manifestation of human social activities in a geographical environment and are an important research subject in urban planning and management. The development of this field can be summarized into three stages:

1. Regional differentiation stage (late 19th century to the 1940s): This stage focused on studying the influence of geographical environmental factors on spatial distribution patterns. Research primarily examined the impact of natural geographic factors on population, industry, agriculture, transportation, as well as the distribution patterns of urban and rural areas. Research methods mainly involved regional comparisons and empirical studies.

2. Social location stage (1950s to 1980s): During this stage, with urbanization and industrialization processes, attention shifted towards the influence of socio-economic factors on spatial distribution patterns. Key research topics included urban functional zoning, urban form, and urban spatial structure. Research methods mainly involved regional analysis, statistical analysis, and empirical studies.

3. Spatial structure stage (1990s to present): In this stage, with the acceleration of urbanization and globalization processes, researchers delved deeper into the study of urban spatial distribution patterns. Major research topics included intra-urban spatial structures, spatial connections between cities, and the evolution of urban spatial distribution patterns. Research methods primarily involved spatial analysis and geographic information system technologies.

In summary, the urban spatial distribution pattern refers to the regularity, concentration, decentralization, and variability of various phenomena, elements, or objects within the city over a certain period and scope. In this context, the key elements are population and urban facilities, and the spatial distribution pattern of these elements reflects variability and characterization.

(2) Urban interactive relationship

EW. Soja, an early scholar exploring the relationship between urban materiality and social space, introduced the concept of "The socio-spatial dialectic," which was an early achievement in directly studying the relationship between urban materiality and social space. Soja believed that people both create and are shaped by urban space, actively changing and adjusting the environment to meet their needs and reflect their values. Simultaneously, humans gradually adapt to the natural environment and the people around them. Soja (1980) explained this as an ongoing two-way process in urban development, where people, while creating and transforming urban space, are also influenced and controlled by the spaces they inhabit and work in [8].

C. Pickvance and others discussed the relationship between urban renewal and social space. C. Pickvance (1985) argued that the main change in urban renewal is material renewal rather than economic revitalization. Social conflicts centered around public resources serve as the primary mechanism for inner-city renewal. Various social groups with shared interests form territorial coalitions based on local areas (communities) to acquire more investment and achieve improvements in the local material environment [9].

In recent years, as urbanization has entered the stage of stock development, research on the relationship between urban material space and social space has often focused on more micro-level perspectives. The research has become more specific and detailed, primarily concentrating on studies of gentrification and the reconfiguration of urban communities and neighborhoods [10]. Additionally, studies have been conducted on the role of urban material spaces such as streets, squares, and shopping centers in promoting social integration [11,12] as well as the relationship between the development of gated communities and urban social development [13].

In summary, the urban spatial interaction of this topic refers to the two-way process between population activities and functional urban facilities. It involves assessing their interdependence through quantitative analysis and spatial modeling. The methodology aims to propose an optimization strategy for urban functional layout based on the allocation of public resources, so as to improve the physical environment and achieve the sustainable development of the station realm.

2.1.4 Population activity and functional facility

Wikipedia defines Human Population as the number of humans in a geographical area, and the subjects that study populations include sociology and geography. It is the fundamental category of population research. The United Nations Department of Economic and Social Affairs (2015) defines population statistics as the collection of statistical data on population events and the characteristics of individuals and related persons involved in those events. Population statistics are of immeasurable value for planning, monitoring, and evaluating various schemes, such as those related to social security, education, public housing, and more [14]. E. Diener et al. (1980) define activities as the sum of actions that are united by a common purpose and fulfill a certain social function. Activities consist of purpose, motivation, action, and commonality and have a complete structural system [15]. In specific terms, activities refer to actions undertaken by individuals with a common purpose, involving purposeful actions.

Urban functions refer to the uses and purposes of various urban spaces, including social, residential, commuting, administrative, and productive activities. Urban functions encompass facilities and services that enable economic operations [16]. They also include material engineering facilities that provide public services for social production and residents' daily lives. Urban functions are the general material conditions necessary for the survival and development of society. As urban functions become more specialized, functional facilities not only include public infrastructure such as power grids, communication (telephony, television, internet, postal services), water supply, gas supply, transportation (roads such as highways, railways, airports, ports), commonly referred to as basic infrastructure, but also include social undertakings such as education, technology, emergency services, healthcare, law, culture, sports, waste management, collectively known as "social infrastructure."

In summary, in this article, population activities refer to the number of purposeful actions carried out by individuals within a station realm at a specific moment or within a day. It is not merely a simple statistical figure but also possesses detailed geographic location attributes, allowing for a visual representation of the spatial distribution of population activities. Functional facilities mainly refer to the quantity of social infrastructure such as commerce, healthcare, dining, etc., within the station realm. Data with geographic location attributes can effectively reflect the agglomeration characteristics of urban functions in a specific area at a micro level.

2.2 Fundamental research related to the station realm

2.2.1 The scope of station realm

The delineation of the station realm boundary is a prerequisite for conducting any research on station area space. The reasonable definition of the station realm boundary is crucial to ensure the reliability of the analytical results of the entire study.

The walking scale is usually the core element to define the scope of rail transit station realm. At present, the method of defining the scope of the rail transit station realm is usually a circular area formed by taking the rail transit station as the center of the circle and the walking distance of 10 min to 15 min as the radius [17]. According to the general normal human walking speed of about 1.0 m/s ~ 1.2 m/s, of which 10 min is 600 m ~ 720 m, 15 min walking distance of 900 m ~ 1080 m. However, in practice, since the scale standard is not a fixed value, the station area range value tends to vary from city to city. For example, the walking distance of station realm in Canadian cities is mostly between 300m and 900 m, while it is 400 to 800 m in the U.S. [18,19]. Hyungun Sung delineated the Seoul rail transit station impact area with a reasonable walking range of 500 m [20]. In Shenzhen, China, some scholars proposed that the public habitual walking time is mostly concentrated in 6~10 minutes interval walking range is about 400~700m based on the survey [21]. In addition, the spatial extent is defined differently in other cities in China, but most of them are within this range (as shown in the Table 2- 4).

Table 2- 4. List of methods for defining scope of station realm

Author	Region	Method
Du Caijun, Jiang Yukun [22]	Beijing	Average walking time: 9.9 min Walking scale: 900 m
Huang Libin[23]	Shanghai	Average walking time: 10 min Walking scale: 600 m
Huang Weidong, Su Qianqian [24]	Hangzhou	Average walking time: 5 min ~ 10 min Walking scale: 400 m ~ 800 m
Wang Jinyuan, Zheng Xian, et al. [25]	Shenzhen	Average walking time: 6 min ~ 10 min Walking scale: 400 m ~ 700 m
Ning Zhang, Jie Dai, et al. [26]	Nanjing	Average walking time: 6 min ~ 10 min Walking scale: 904 m
Duan Degang, Zhang Fan [27]	Xi’an	Average walking time: 10 min Walking scale: 800 m
Guo Daqi, Gao Feng, et al. [28]	Shenyang	Average walking time: 10 min Walking scale: 800 m

However, some scholars believe that the rail transit station realm is not simply delineated by abstract geometric radius, it is also influenced by topographic space, road pattern, plot function and other factors [49].

In summary, in order to make the research results more accurate and practical, this study will investigate and study the pedestrian and environmental characteristics around each station of Xi'an Metro Line 2, calculate the pedestrian scale based on the actual findings on pedestrian walking time and walking speed, and make further optimization based on the characteristics of parcel boundaries and roads, finally determine the influence realm range of each station.

2.2.2 The types of station realm

Through in-depth literature analysis, it is found that there are not many studies specifically focused on the classification of rail transit station realms. However, research on the classification of rail transit stations can serve as a reference for selecting the division criteria in this study, and the classification methods used in these studies can also provide insights for this section of the research. The common classification criteria for rail stations are typically node-oriented classification, place-oriented classification, and function-oriented classification.

Node-oriented classification primarily focuses on the transportation functions of stations as the main criterion for division. As a high-capacity public transportation mode in urban areas, rail transit has the advantages of speed and convenience, making it an important component of urban connectivity and transportation hubs. Therefore, this type of classification emphasizes the role of rail stations in the urban rail transit system and their relationships with other public transportation modes.

Place-oriented classification emphasizes that various urban activities around rail transit stations are supported by land use. The nature of land use is the main factor determining the types of urban spatial functions within the station realm. Thus, this classification method uses urban land use functions as the criteria, including the location factors of rail transit stations in the city.

Function-oriented classification categorizes stations based on the land use around the stations and their roles in specific functional areas. The advantage of function-oriented classification is that it can rationalize land use and contribute positively to urban planning.

(1) Classification of rail transit stations under node-oriented approach

In the node-oriented classification, which focuses on transportation functions, rail transit stations are classified based on their location, architectural layout, structural type, and alignment of the rail lines [29]. Stations can be categorized according to their location as underground stations, surface stations, or elevated stations. Based on the architectural layout, stations can be classified as shallow-buried or deep-buried. The structural types can be rectangular box-shaped underground stations or elliptical tunnel stations. Depending on their operational nature, stations can be classified as intermediate stations, hub stations, transfer stations, terminal stations, vehicle depots, and parking facilities. Finally, stations can also be categorized based on the alignment of the rail lines as island platform stations, side platform stations, or a combination of both island and side platforms.

Guangzhou, Shanghai, Shenzhen, Beijing, and other major cities divide rail transit stations into 3-4 grades according to a series of indicators such as the number of rail transit lines connected to the station, the mode of interchange traffic related to the station, and the land attributes of the area where the station is located[30,31],The details are shown in the Table 2- 5.

Table 2- 5. Node-oriented station hierarchy in different cities

City	Grading indicators	Grading profile and criteria
Guangdong	Number of connected rail lines	Major Interchange Hub: Nodes connecting three urban-level or two regional-level rail lines. Interchange Station: Nodes connecting two urban rail lines. General Station: Other regular rail transit stations.
Shanghai	Interchange Modes of Transfer Land Use Attributes of Interchange Hub Areas	Passenger Hub Station: Rail transit hubs connecting with large-scale external transportation hubs. Bus Hub Station: Rail transit hubs located at major conventional bus hubs, line junctions, or CBD areas. Bus Interchange Station: Rail transit hubs connecting with regular bus hubs. General Interchange Station: Rail transit stations connecting with regular bus stops.
Shenzhen	Types of interchange modes Land Use Attributes of Interchange Hub Areas	Integrated Interchange Hub: Rail transit hubs located at the junctions of major conventional bus hubs, external transportation hubs, or city primary and secondary centers. Major Interchange Hub: Rail transit hubs located at the junctions of regular bus hubs or central areas of districts. General Interchange Hub: Rail transit hubs connecting with regular bus stops.
Beijing	Types of Interchange Modes Number of Interconnected Rail Transit Lines	Level 1 Hub: Rail transit hubs connecting with large-scale external transportation hubs. Level 2 Hub: Interchange hubs between rail transit lines and interchange hubs connecting rail transit with multiple regular bus lines. Level 3 Hub: Rail transit stations connecting with regular bus stops.

(2) Classification of rail transit stations under place-oriented approach

Under the place-oriented classification, rail transit stations are classified based on the surrounding area, aiming to clarify the purpose of land development and reduce the workload of land use categorization. However, due to the consideration of multiple indicators during the classification process, the resulting classification results may be more ambiguous.

Zheng Ming believes that there are two planning patterns for subway stations by combining land use: a. Strong Center, where the central unit undergoes high-intensity development, forming an intensive urban complex with the station; b. Weak Center, where stations are independently located such as standalone at-grade or elevated stations, or underground stations located at intersections [32]. Fu Bofeng et al., taking the land use within the influence range of suburban rail transit stations as the starting point, comprehensively consider the transportation function and place characteristics of the stations, and divide suburban rail transit stations into 7 major categories and 11 subcategories

[33]. At Tongji University, Hui Ying categorizes station areas into four types: Public Center Area, Transportation Hub Area, Mature Residential Area, and Urban Periphery Area based on the characteristics and functions of the station areas [34]. Zhang Yunan, starting from the micro-scale urban spatial form of the stations, proposes several forms of the core area between urban rail transit stations and surrounding areas, including station-centric urban blocks, node-centric stations, integrated complexes blending stations with urban spaces, nodes developing around the urban center, and various levels of transportation hubs inside and outside the city [35].

（3） Classification of rail transit stations under function-oriented approach

From the perspective of guiding land development in station areas, function-oriented classification is more suitable. According to urban functional positioning, rail transit stations are generally classified into three major types: central stations, residential stations, and hub stations [36]. As shown in the Table 2- 6, with the continuous improvement of technological conditions, researchers have refined and enriched the classification indicators, greatly enhancing the accuracy of the classification results for rail transit stations.

Table 2- 6. Different scholars have classified stations based on functional orientation

Author	Classification basis	Classification results
Huang Jun	Transportation Function: Traffic Type (A) + Function Type (B) Urban Function: Spatial Function (C) + Regional Level Entity (D)	Station Area Spatial Type: A1 + B1 + C1 + D1
Wang Xialu	Station Area Functional Layout Development Intensity	Public Center Type Station, Transportation Hub Type Station, Residential Community Type Station, Landscape Open Type Station
Duan Degang, Zhang Fan	Functionalities of Surrounding Areas Transportation Function Land Use Functions	Residential Type Station, Public Type Station, Commercial/Service Type Station, Transportation Type Station
Zhao Chang	Urban Function Transportation Function	Planned Type, Residential Type (Peripheral Residential Type, Core Residential Type), School District Type, Center Type (Peripheral Center Type, Core Center Type), Hub Type, Industrial Type
Ministry of Housing and Urban-Rural Development	Classification of Rail Transit Network Land Use Function Transportation Service Scope Service Level	Hub Station (Class A), Center Station (Class B), Cluster Station (Class C), Special Control Station (Class D), Terminal Station (Class E), General Station (Class F)
Liu Jiachen	Urban Function, Transportation Function of the Station, Population Density	Public Commercial Service Type, High Population Density General Type, Low Population Density General Type, Industrial Warehouse Type, Higher Education Type, Planned Development Type

In summary, different research purposes result in different station classification approaches. Based on the station classification methods, there is currently no comprehensive standard to determine the type of each station. Particularly, for the study of rail transit station realms as part of urban space, it is important to focus on describing the urban spatial characteristics within the station area. Therefore, based on the synthesis of indicators from the three classification systems: node-oriented, place-oriented, and function-oriented, this study combines the characteristics of rail transit stations themselves and the urban space. The classification of rail transit station realms in this study mainly includes two major indicators: transportation functional space and urban functional space. Transportation functional space primarily considers the transportation environment characteristics of the station itself and the station area, while urban functional space primarily considers the spatial form and density characteristics of the station realm. The specific indicator classification is shown in the Table 2- 7. Additionally, to classify the Xi'an Metro Line 2 station realm objectively and comprehensively, this study adopts cluster analysis to analyze and categorize the various indicators.

Table 2- 7. Classification index of rail station realm

First-level indicators	Second-level indicators	Third-level indicators
Transportation Functional Space	Station conditions	Number of Transit Lines
		Number of stations entrances and exits
	Station realm transportation environment	Number of bus stops
		Number of transportation service facilities
Urban Functional Space	Station realm spatial form	Boundary form
		Network form
		Building form
	Station realm spatial density	Population density
		Land use intensity

2.3 Theoretical research of urban spatial distribution

2.3.1 Geographic information system theory

Geographic Information System theory believes that the analysis of urban spatial distribution pattern can be done with the help of Geographic Information System (GIS) technology, through collecting, storing, processing and analyzing the spatial data of the city, so as to understand the urban spatial pattern and urban development trend more comprehensively and precisely.

The history of GIS theory can be roughly divided into the following stages：

(1) The pioneering period of spatial analysis (1960s to early 1970s): Roger Tomlinson, a geographer, introduced the concept of Canada Land Inventory, the world's first geographic information management system that enabled digital map production and management [37]。Subsequently, he first introduced the concept of Geographic Information System in his paper published in 1967 and described the basic principles and applications of GIS. [38]。

(2) The initial establishment of GIS theory (mid-1970s to early 1980s): In the mid-1970s, GIS theory gradually gained the attention of academia and industry, and theoretical research continued to deepen. basic concepts and methodological systems of GIS were initially established, and basic

technologies such as map projection, spatial data structure, data acquisition and data processing gradually matured. [39]。

(3) The period of widespread application of GIS (1980s to mid-1990s): The application of GIS technology gradually expanded to many fields, including land use, urban planning, resource management, environmental monitoring and so on. At the same time, GIS software tools gradually developed, commercial software such as ArcGIS, MapInfo, etc. came out one after another.

(4) The development period of geographic information science (mid-1990s to present): GIS theory gradually penetrated into geographic information science, GIS became an emerging science integrating data, methods, technologies and theories, and the integration and development of related technologies such as remote sensing, spatial statistics, multimedia technology, etc. made the scope of GIS applications expand continuously [40]. With the rise of the Internet and open-source software, OpenGIS has started to become a new research direction. It proposes a new GIS framework and model, aiming to promote the rapid development of GIS applications and development.

In conclusion, the theoretical research and practical development of GIS have been continuously deepened, from the initial mapping to geographic information science, from computer applications to the Internet era, the scope and field of application of GIS have been expanded, providing strong support for the sustainable development of urban space.

2.3.2 Spatial analysis technology

Spatial analysis can refer to the analysis of spatial data or the spatial analysis of data. The former focuses on the analysis of non-spatial characteristics of spatial objects and phenomena. Robert Haining's work specifically emphasizes the sampling design, statistical compilation, data simulation, and analysis of statistical properties of spatial statistical data. He believes that a central issue is how to use mathematical (statistical) models to describe and simulate spatial phenomena and processes, transforming geographical models into mathematical models to facilitate quantitative description and computational processing [41]。 Wang Jinfeng et al. then define it as a technique for analyzing spatial data based on the location and morphology of geographic objects with the aim of extracting and transmitting spatial information. Spatial analysis is the main feature of the Geographic Information System [42]. Zhang Kequan and Guo Renzhong, on the other hand, focus on statistical analysis methods, particularly multivariate statistical analysis methods, for handling spatial data. These analyses typically do not consider the spatial characteristics of the data as restrictive factors. The specific spatial locations described by spatial data do not play a constraining role in these analyses [43]。

From this perspective, the analysis of spatial data does not fundamentally differ from general data analysis. However, the interpretation of the results of spatial data analysis inevitably relies on geographic space. In most cases, the results of the analysis are described and represented in the form of maps. Therefore, although the spatial positions of the data sampling points are not explicitly considered during the analysis process, the analysis still describes spatial processes and reveals spatial patterns and mechanisms.

Spatial analysis of data involves studying spatial objects in terms of their spatial positions, relationships, and other aspects, with the aim of providing quantitative descriptions of spatial phenomena [44,45]. From the perspective of information extraction, spatial analysis of data is a process of description and explanation, involving the extraction of features and computation of parameters. The main methods used in spatial analysis include spatial statistics, graph theory, topology, and computational geometry, with the primary task being the description and analysis of spatial structures. The main contents of spatial analysis include spatial location, spatial distribution, spatial morphology, spatial distance and spatial relationship[46]， the relevant definitions and analysis methods are shown in Table 2- 8.

Table 2- 8. Content, definition, and analysis methods of spatial analysis

Main content	Definition	Analysis method
Spatial location	Transmission of positioning information of spatial objects with the assistance of spatial coordinate system	Spatial superposition analysis
Spatial distribution	Group positioning information of the same type of spatial objects, including distribution, trend, comparison, etc.	Mean Nearest Neighbor Analysis Multi-distance spatial clustering analysis Kernel density analysis
Spatial form	Geometric form of spatial objects	The one-dimensional measure of area and perimeter Fractional dimensionality
Spatial distance	Proximity of spatial objects to each other	Network analysis
Spatial relationship	Correlation, orientation, similarity, correlation, etc. of spatial objects	Spatial autocorrelation analysis

2.4 Theoretical research of urban spatial interaction relations

2.4.1 Central place theory

The Central Place Theory was initially proposed by German geographer Walter Christaller in 1933. In his work "Central Places in Southern Germany," Christaller proposed that a city is a collection of central areas, each consisting of a residential area, a service center, and a market region [47]. The size and function of a city depend on its central position. Hoffman's theory provided important theoretical support and guidance for urban geography and urban planning.

Building upon Christaller's work, American geographer Brian J. L. Berry further developed and expanded the Central Place Theory in the 1950s. He introduced three important concepts: central place hierarchy, central place function, and market areas. Applying Christaller's theory to the study of American cities, Berry found a close relationship between the spatial distribution and functional differentiation of cities and factors such as central position, market area size, and hierarchy.

Furthermore, in the 1960s and 1970s, the Central Place Theory underwent further development and revisions. American geographers William Garrison and Lawrence Curran introduced two new concepts: marginal analysis and expansion analysis [48,49]. They argued that the spatial distribution and development of cities depend on the influence range of central places, the size and demand of markets, and the level of transportation and communication technology [50].

With the globalization and development of information technology, the Central Place Theory gradually shifted towards a global perspective, emphasizing urban competition, cooperation, and the global status and power of cities. In the 1980s, British geographer Peter J. Taylor and others introduced the New Economic Geography [51]. This theory emphasized cities as centers of global economy and culture, studying the development and transformation of cities in the context of globalization, as well as the connections and interactions between cities and their surrounding regions and other cities. The theory has been widely applied and developed in the fields of urban geography and globalization studies. The core ideas of the New Economic Geography include:

1. Global Cities: The theory views cities as integral components of globalization, studying the status, influence, and functions of global cities, as well as the connections and interactions between them [52].

2. Network Space: The theory recognizes that the development of information and communication technologies has changed the spatial relationships between cities and regions, creating a global network space in which cities and regions play important roles [53].

3. Spatial Organization: The theory emphasizes the spatial organization and functional differentiation of cities, viewing cities as complex systems composed of multiple centers and sub-centers, each with its own specific function and influence range [54].

4. Geopolitics: The theory places urban development and transformation within the context of geopolitics, studying the relationships and mutual influences between cities and nations/regions, as well as the roles and functions of cities in geopolitics [55].

Additionally, research methods in the New Economic Geography include case studies, statistical analysis, GIS technology, network analysis, and others.

2.4.2 Urban catalyst theory

The concept of "catalyst" in urban context, borrowed from the chemical concept of a catalyst, refers to elements that can trigger a series of chain reactions and stimulate the development of other elements within a city, without undergoing any changes themselves throughout the process [56]. The concept of "urban catalyst" was formally introduced by American scholars Wayne Attoe and Donn Logan in 1989 in their book "American Urban Architecture: Catalysts in the Design of Cities." Attoe and Logan proposed that any element in a city that can inspire and catalyze urban development can be considered as a "catalyst." By introducing these elements as catalysts, they can mutually stimulate and coordinate various chain reactions, which can manifest in social, economic, and political aspects, as well as in the form and architecture of a specific urban space [57].

These catalyst elements that trigger a series of chain reactions in the city can be tangible elements, such as a building, a green space, a plaza, or a transportation mode, as well as intangible elements, such as architectural trends or development patterns. As long as they can facilitate the interaction of urban elements, catalyze urban development, and have an impact, they can be regarded as catalysts. The theory of "urban catalyst" differs from traditional urban design theories as it guides the development and construction of urban space in a bottom-up and gradual manner. By controlling the focal points of urban growth, it plays a dual role of stimulation and guidance [58]. Urban catalysts emerge during the process of urban development and, in turn, influence and contribute to the continuous improvement and development of urban spatial structure and functionality.

The catalytic process is shown in the Figure 2- 1. Actions (represented by shaded lines), and the regulating aspects of the process are represented by the dotted lines around the shaded lines. Whether the urban space is developed, restored, or renewed catalyzes other actions, which in turn drive other actions. Of course, each action is also limited in some way so that the reaction does not disrupt the overall space of the city.

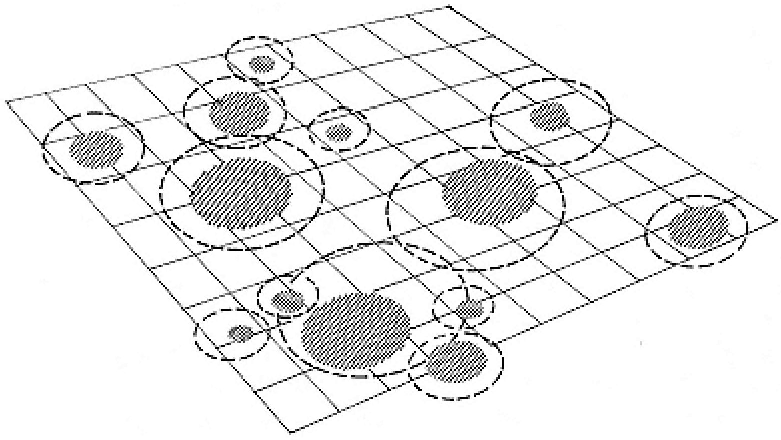


Figure 2- 1. Diagrammatic representation of the catalytic process [56]

Urban public transport stations are the places where the highest concentration of people naturally gathers in the city, inevitably leading to a significant amount of commercial activity and promoting the development of the surrounding areas. Urban rail transit, with its high capacity for passenger transportation, acts as a catalyst through its stations. The ability of rail transit stations to handle large volumes of passenger flows helps create an attractive urban space, providing opportunities for increased interaction and activities, and guiding the city towards a healthy and sustainable development.

The guiding role of urban rail transit in spatial structure has gradually become evident during the rapid development of rail transit. With rail transit stations as the anchor, the development and construction of commercial, office, residential, and public service areas in the surrounding regions have generated new business opportunities, driving a series of developments and updates in station realms and promoting their spatial development. Therefore, rail transit stations, as important catalysts in the construction of new urban areas or the revitalization of old ones, simultaneously promote the overall development of the city through their individual growth [59]. They provide a method for gradually achieving urban planning goals, fully harnessing their catalytic role in guiding the development of urban functions and urban structures and are of significant importance in guiding urban spatial development.

2.5 Theoretical research on development model of urban rail transit station

2.5.1 Transit-oriented development model

Transit-Oriented Development (TOD) is a public transportation-oriented development model pioneered in 1993 by Peter Calthorpe, an American architect and urban planner who is a representative of New Urbanism [60]. He defined TOD as a comprehensive community of commercial, cultural, educational, and residential developments within a walking radius of 400 to 800 meters (5 to 10 minutes), using a grid of roads, mixed use of functions, appropriate development density, and walkability within the residential area (Figure 2- 2). In urban planning, TOD is a type of urban development that maximizes the amount of residential, business and leisure space within walking distance of public transport [61]. It promotes a symbiotic relationship between dense, compact urban form and public transport use [62].

Although TOD theory was initially adopted primarily as a response to the unrestricted sprawl of individual cities in the United States after World War II, as a response to the traditional automobile use-driven development model [63]. However, as urbanization levels are increasing rapidly in countries around the world, more and more cities or regions are turning to TOD as a means of improving their urban fabric. TOD has become a comprehensive strategy for merging land use and transportation networks, enhancing compact urban forms, and promoting public transportation use by reducing the use of private vehicles.

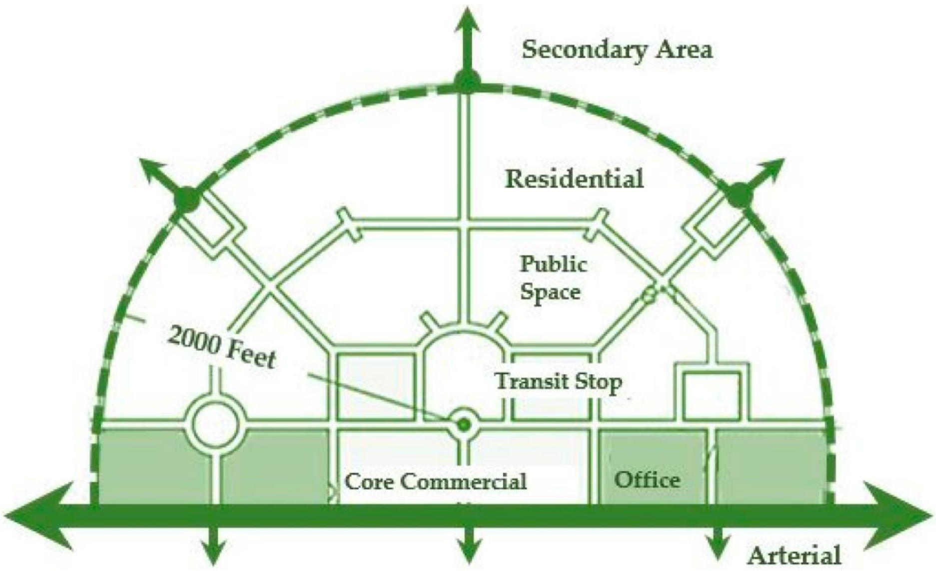


Figure 2- 2. Illustration of transit-oriented development (TOD) [64]

In 2005, the City of San Francisco implemented the Transbay Redevelopment Project with the aim of transforming the Transbay area into a mixed-use, transit-oriented, and sustainable development model. During this time, the old Transbay terminal and elevated ramp will be transformed into the Salesforce Transit Center, a modern, integrated transit hub.

The Salesforce Transit Center is located one block south-east of Market Street, a primary commercial and transportation artery in San Francisco (Figure 2- 3). The Salesforce Transit Center, with a total built-up area of approximately 14 square meters, serves as a hub for 11 regional bus systems and street-level municipal transit services (Figure 2- 4). With the completion of the underground tunnel, it will provide regional commuter rail and high-speed rail services for the residents of San Francisco, eventually extending its reach across the entire state of California.

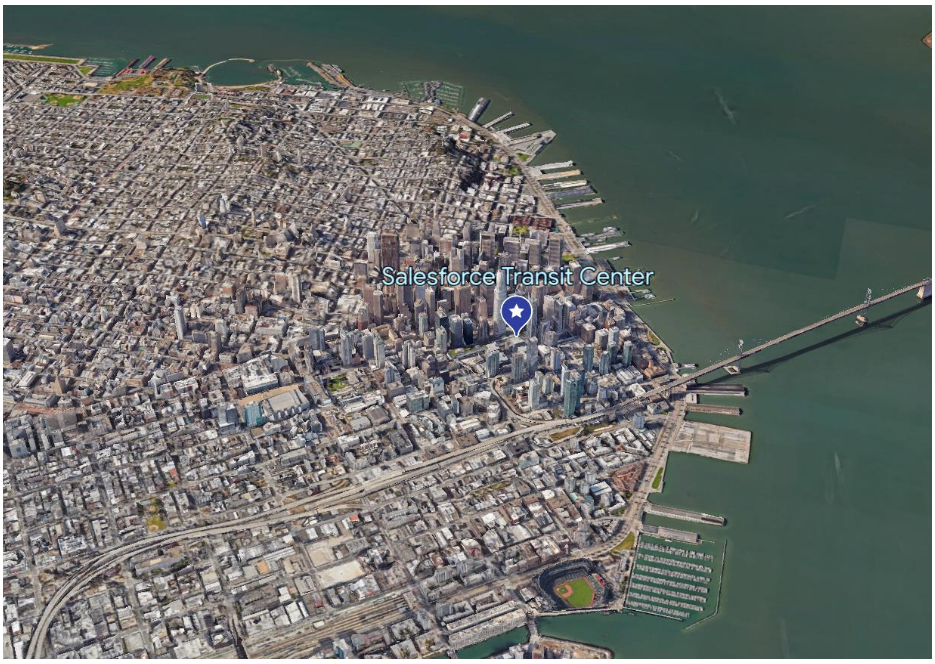


Figure 2- 3. Salesforce Transit Center location in San Francisco

Figure 2- 4. Salesforce Transit Center's main building form [65]

The Transit-Oriented Development (TOD) model led by the Salesforce Transit Center has had a significant impact on the economy and communities in the Transbay area. To date, the regional development has created approximately 650,000 square meters of commercial and office space, 4,200 housing units (including 35% permanently affordable housing), 1,000 hotel rooms, and a nearly 9,000 square meter retail district, forming a truly multifunctional urban center. Simultaneously, since 2017, high-tech research and development companies, including Google, Uber, Facebook, Dropbox, Airbnb, and Amazon, have progressively established their presence, making the Transbay area a high-tech hub of San Francisco.

Bounded by Mission, Howard, Beale and Second streets, the four-block long multi-modal Salesforce Transit Center’s transit operators provide daily transit services within and to San Francisco, the North and East Bays along with the Peninsula. As show in the Figure 2- 5, the Transit Center serves AC Transit, Greyhound, San Francisco Municipal Transportation Agency (Muni), Golden Gate Transit, Paratransit, and WestCat Lynx, with proximity to SamTrans, Amtrak, Muni Metro, and BART. Bus operations are found on the ground floor Bus Plaza (between Fremont and Beale streets) and the third floor Bus Deck. In the future, the Transit Center will serve Caltrain and California High Speed Rail when rail service is connected from the 4th and King Street station and ultimately, the State’s statewide rail service system.

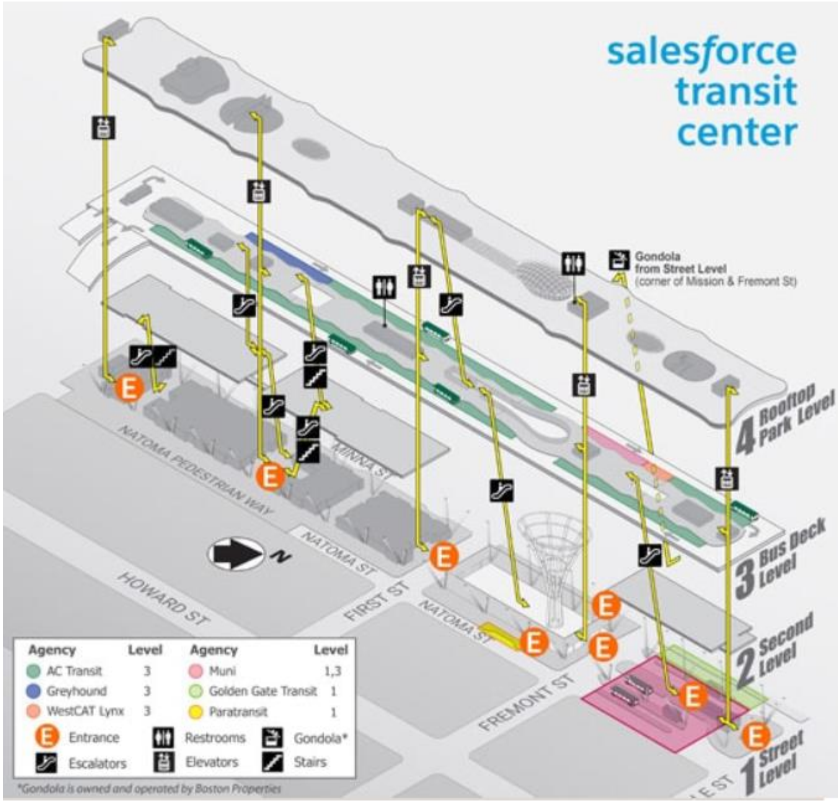


Figure 2- 5 Salesforce Transit Center's Stereoscopic Transfer System [66]

2.5.2 Station-city integrated development model

"Station-City Integration" is an urban planning concept that aims to integrate transportation hubs with urban buildings, enhancing the sustainability and efficiency of the city through spatial and functional integration [67]. Station-city integration development goes beyond the development of transportation hubs and their surrounding areas. It simultaneously integrates the "human factor" by combining multiple modes of transportation and engages in the overall planning and development of residential and commercial areas along the rail transit lines. As shown in the Figure 2- 6, the basic approach of development and construction can be summarized as the integrated development of office, residential, and transportation infrastructure, with the goal of creating a comprehensive urban development that blends the station with the surrounding city. Various techniques are employed, such as expanding the development area with bus transit systems, constructing interdisciplinary public spaces and pedestrian networks, and strategically placing facilities at terminal and intermediate stations to attract passenger flow. Station-city integration development not only allows operators and developers of rail transit to achieve economic benefits but also amplifies the agglomeration effect of hub stations, thereby enhancing the city's attractiveness.

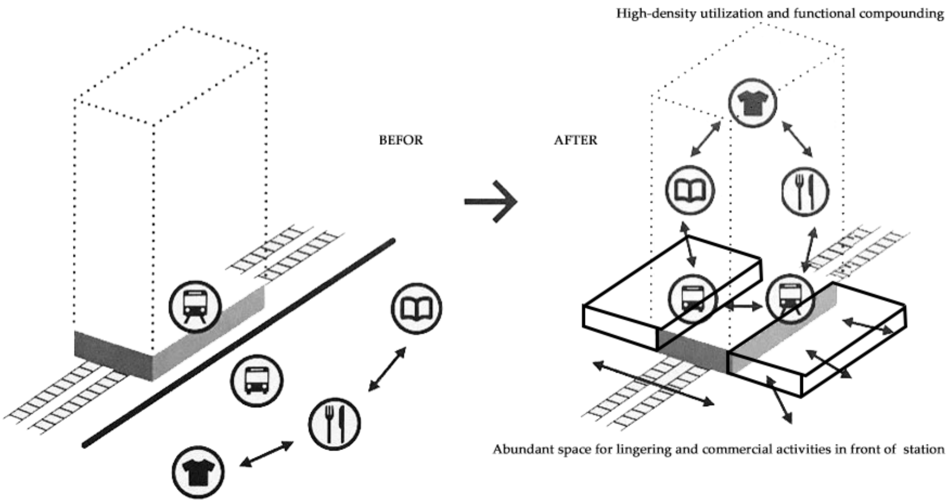


Figure 2- 6. Cluster Development Model Centered on Rail Transit Station

The development of station-city integration in Japan, guided by rail transit, has undergone multiple transformations and has been deeply implemented in various aspects of rail transit planning, design, construction, and operation. Japan is also one of the earliest, most widely applied, and most developed regions in the world when it comes to the practice of "station-city integration" development model. With the significant development of rail transit, the development model of rail transit stations in Japan has become increasingly refined and has evolved to the TOD 4.0 model, which emphasizes the integrated development from "station-city" to "station-city-people" [68]. Through the collaborative development of rail transit construction and urban development, compact cities with a walking distance of approximately 800 meters have been formed [69]. For example, Shibuya Station in Tokyo (Figure 2- 7), through a series of redevelopment projects, has reconfigured transportation infrastructure while integrating urban functional spaces through the concept of an "Urban Core." This integration has facilitated smooth connectivity in public transportation, linking high-rise complexes consisting of offices, music theaters, event halls, and commercial facilities to multiple railway lines, making Shibuya Station one of Tokyo's most representative commercial districts.

The three typical measures of station-city integration that showcase the three-dimensional utilization of space are overhead development, vertical transportation, and station complexes. Overhead development is a common practice that involves utilizing space in layers. Its purpose is not only to save land and space but also to bridge fragmented spaces and activate the city's "negative space." Station complexes refer to integrated complexes that combine multiple functions with station construction. In the underground section, resources such as subway stations, underground parking lots, and underground commercial facilities are shared, making full use of the underground space of the subway station. Underground commercial facilities are located on the basement level, while underground parking is situated on the sub-basement level. An underground pedestrian system is connected to public parking lots and underground commercial facilities. In the above-ground section, there are usually bus stops, commercial areas, offices, and other functions, creating a multifunctional space that reinforces the status of the transportation hub.

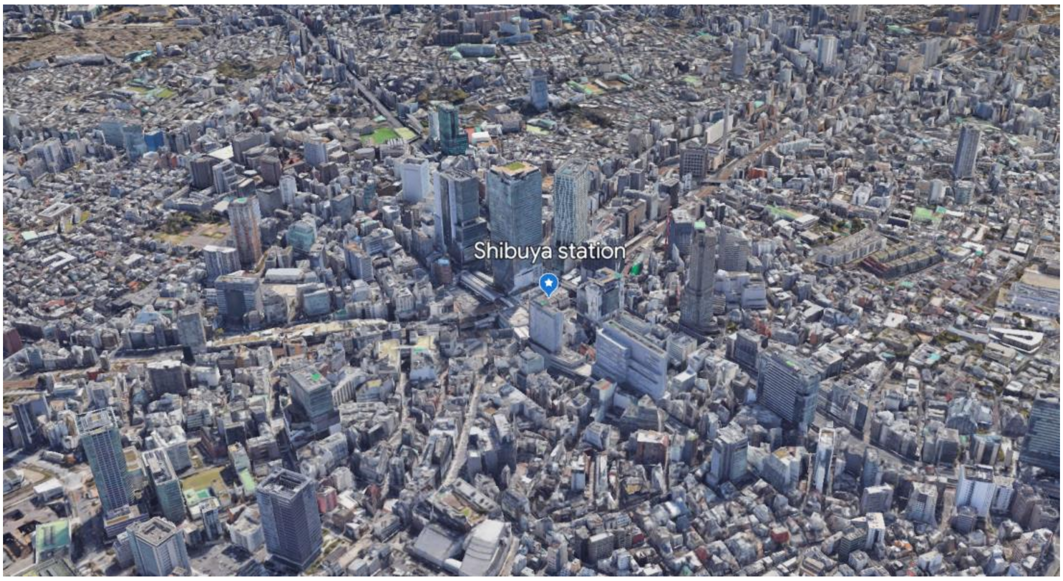


Figure 2- 7. Shibuya Station's Space Location in Tokyo

An example is shown in Figure 2- 8, during the integrated construction of Shibuya Station, an elevated road was constructed to reduce the division caused by the railway lines. Additionally, to manage complex pedestrian flows, the station area has planned multiple elevated and underground pedestrian walkways, including underground passages, surface-level walkways, and elevated walkways. This approach not only saves land resources and enhances spatial connectivity but also facilitates convenient travel while shaping a vibrant urban space. Moreover, within the Shibuya Station area, there are numerous corridors and walkways connecting various commercial spaces. Especially noteworthy are the walkways above the railway tracks, which link both sides of the railway. Passengers can easily traverse between different shopping centers, facilitating a mutually beneficial interaction between transportation and commercial activities.

Figure 2- 8. Station-city Integration at Tokyo Shibuya Station [70]

2.5.3 Rail plus property model

The "Rail plus Property" (R+P) Model is a development concept proposed by the MTR Corporation in Hong Kong, aiming to promote urban spatial integration and functional intensification centered around railway stations. Hong Kong has a total area of approximately 1,100 km², but due to its mountainous terrain, the urban land area is only around 200 km². This makes Hong Kong one of the most densely populated cities in the world, with a population density of 7,125 people/km² as of 2020 [71]. According to statistics from 2018, Hong Kong residents make approximately 13.216 million trips on public transportation daily, with railways accounting for 43.9% of the total trips. Furthermore, nearly 50% of the population resides within a 500-meter radius of a subway station [72]. Hong Kong effectively utilizes the R+P model to ensure the normal functioning of a densely populated city with limited land resources.

The R+P model is derived from Transit-Oriented Development (TOD), where residential developments are constructed above or near railway stations. This creates the "Rail plus Property" (R+P) model, as illustrated in Figure 2- 9. With intensive development and the influence of the terrain in established areas, Hong Kong's urban space has transitioned from a single-plane city to a three-dimensional city, responding to the demands of high-density development. This spatial characteristic synergizes with the development of the public transportation system.

Kowloon Station is one of the representatives of the "R+P" model development in the Hong Kong subway system. The subway tracks and the buildings above form an integrated block complex. The underground levels consist of the subway station and high-speed rail station, the ground level includes bus and taxi stands, and the second to third floors house commercial shops. Above that are enclosed high-rise residential buildings, providing a total of 6,390 housing units. The top-level platform of the podium serves as a central landscaped garden for the residential area and the only public open space. The entire Kowloon Station complex offers all the essential life services, from transportation to residential and office spaces, making it a self-sufficient and independent neighborhood.

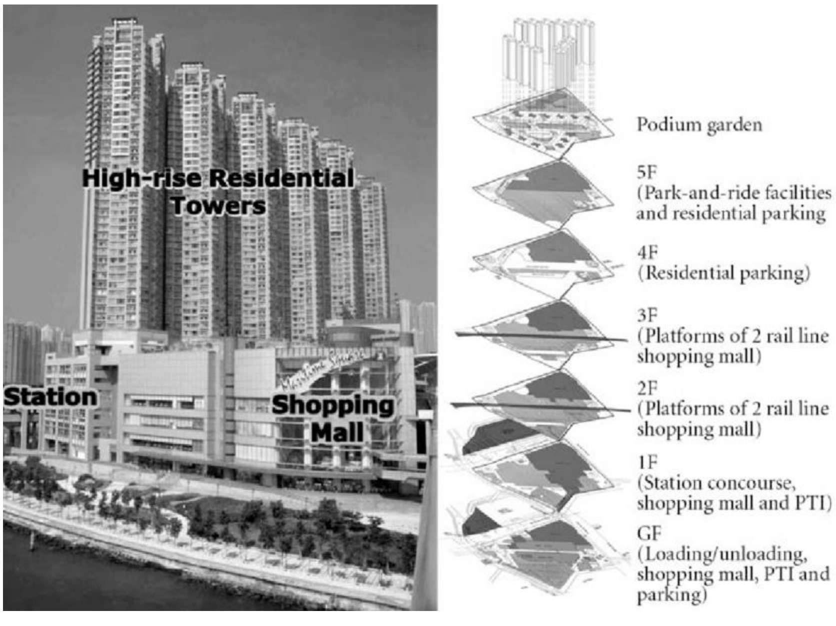


Figure 2- 9. Maritime Square Residential-retail Development atop Tsing Yi Station [73]

The "R+P" model tightly integrates Kowloon Station with the properties above (commercial and residential areas), creating a compact spatial form centered around the rail station that promotes urban spatial integration and functional intensification (Figure 2- 10). Additionally, the 18 to 35-

story residential towers form a prominent silhouette within the massive architectural ensemble.

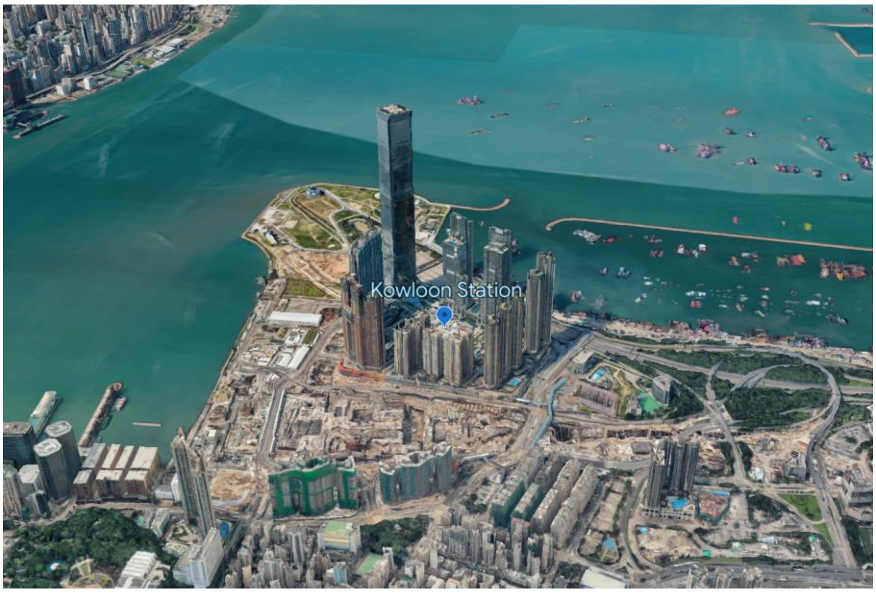


Figure 2- 10. Kowloon Station's Space Location

The Kowloon Station neighborhood covers an area of 135,000 square meters, with a total floor area of 1.7 million square meters and a high plot ratio of 1:12 [74]. The development of the Kowloon Station neighborhood in a high-density and compact mode with the "R+P" model effectively addresses the constraints of Hong Kong's high population density and scarce land resources. As shown in Figure 2- 11, the spatial layout within the Kowloon Station area has undergone a transformation from a single "flat" dimension to a three-dimensional configuration. The Kowloon Station neighborhood integrates community services, offices, hotels, and transportation facilities, among other public amenities, into a comprehensive public space. It resembles a self-contained mini city that is completely separated from the outside. The three-dimensional layout provides great convenience for the users of the properties above Kowloon Station.

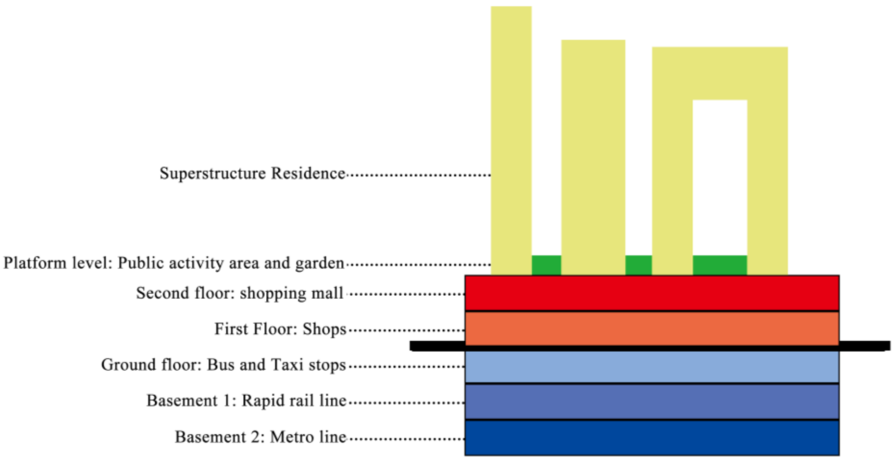


Figure 2- 11. Kowloon Station's Stereoscopic Transfer System

2.5.4 The station integrated development model

Station Integrated Development (SID) refers to the coordinated planning, development, and construction of both the railway station and its surrounding area, creating a highly integrated urban layout [75]. Many cities in France, such as Paris, Lyon, and Lille, have pioneered the organic integration of high-speed rail hubs with urban architectural complexes, forming transportation and building complexes within the city. Lyon Part-Dieu station is the largest integrated transportation hub in Lyon, serving national transportation and being the largest train station in France outside the Île-de-France region. With an area of 60 hectares (Figure 2- 12), Lyon Part-Dieu station area features a large-scale urban central business district (CBD) with 64,380 square meters of commercial space, 216 hotels, and 700 parking spaces. It gathers 40% of Lyon's office buildings and accommodates approximately 85,000 people passing through the station each day [76].

The Lyon Part-Dieu station area is an important urban gateway in Lyon, with several iconic buildings that significantly influence the city's architectural landscape. The Part Dieu Tower, completed in 1977 and standing at a height of 164 meters, and the Oxygen Tower, completed in 2010 and standing at a height of 166 meters, together form the tallest skyline in Lyon. According to statistics from the Lyon Part-Dieu station area official website (http://www.lyon-partdieu.com), the residential population in the area was approximately 20,000 people in 2016, with a population density of 17,600 people per square kilometer. The land use intensity in this area is much higher than in surrounding areas, making it one of the true centers of Lyon.

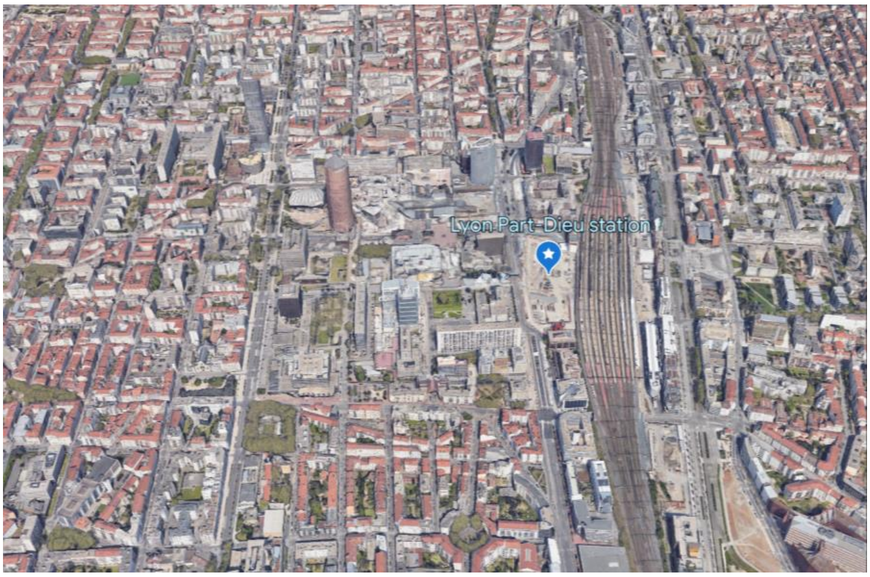


Figure 2- 12. Lyon Part-Dieu Station's Space Location

The station area encompasses various functions, including office spaces, residential areas, commercial establishments, and leisure facilities, creating an intensive and multi-functional spatial layout (Figure 2- 13). The current office buildings in the station area cover an approximate area of 1 million square meters, and there are 12 hotels of different levels and 8 hotel-style apartments, making it one of the most important business and trade negotiation areas in Lyon. Lyon Part-Dieu station area is also one of the largest shopping centers in Europe, with an annual visitor flow of 54.75 million people. Additionally, the area features traditional food markets, a concert hall, Lyon Central Library, sports arenas, and other public service facilities, providing comprehensive support for the needs of employees, residents, and tourists in the Lyon Part-Dieu station area.

Figure 2- 13. Functional Zoning of Lyon Part-Dieu Station Area

2.6 Chapter summary

This chapter is research on relevant theoretical foundations. First, the scope and limitations of this study are clarified by defining the concepts of rail transit, station realm, urban spatial pattern and interactive relationship, as well as population activity and functional facility. Secondly, the research related to the station realm clarifies the walking scale is the core elements for defining the scope in this thesis, as well as the f function-oriented classification system was used to determine the type of station realm. Finally, the study of fundamental theories such as central place theory and urban catalyst theory as well as the comparison of related cases provide the basis for the study of the spatial distribution characteristics and interaction mechanism of the station realm in this thesis.

References

[1] China Association of Metros. Classification of urban rail transit (T/CAMET 00001-2020) [S]. 2020.09.

[2] Awdry C. Encyclopaedia of British railway companies[M]. Stephens, 1990.

[3] Ministry of Construction of the People's Republic of China. Standard for Classification of China Urban Public Transportation (CJJ/T 114-2007) [S]. Beijing: China Architecture and Building Press, 2007.07.

[4] Feng Aijun. Statistics and Development Analysis of Urban Rail Transit in China in 2021[J]. Tunnel Construction (English and Chinese), 2022, 42(2): 336-341.

[5] Foley D L. An approach to metropolitan spatial structure[J]. Explorations into urban structure,

1964: 21-78.

[6] Webber M M. The urban place and the nonplace urban realm[M]//Explorations into urban structure. University of Pennsylvania Press, 2016: 79-153.

[7] Huang Yaping, Urban Planning. Urban Spatial Theory and Spatial Analysis [M]. Southeast University Press, 2002.

[8] Soja E W. The socio-spatial dialectic[J]. Annals of the Association of American geographers, 1980, 70(2): 207-225.

[9] Pickvance C. The rise and fall of urban movements and the role of comparative analysts[J]. Environment and Planning D: Society and Space, 1985, 3(1): 31-53.

[10] Walks R A, Maaranen R. Gentrification, social mix, and social polarization: Testing the linkages in large Canadian cities[J]. Urban Geography, 2008, 29(4): 293-326.

[11] Staeheli L A, Mitchell D. USA's destiny? Regulating space and creating community in American shopping malls[J]. Urban studies, 2006, 43(5-6): 977-992.

[12] Sauter D, Huettenmoser M. Liveable streets and social inclusion[J]. Urban Design International, 2008, 13(2): 67-79.

[13] Lemanski C. Spaces of exclusivity or connection? Linkages between a gated community and its poorer neighbour in a Cape Town master plan development[J]. Urban studies, 2006, 43(2): 397-420.

[14] Principles and Recommendations for a Vital Statistics System. New York: United Nations Department of Economic and Social Affairs.2015. https://unstats.un.org/unsd/demographic-social/Standards-and-Methods (Access on October 15, 2022).

[15] Diener E, Lusk R, DeFour D, et al. Deindividuation: Effects of group size, density, number of observers, and group member similarity on self-consciousness and disinhibited behavior[J]. Journal of Personality and Social Psychology, 1980, 39(3): 449.

[16] O'sullivan A, Sheffrin S M. Economics: Principles in action[J]. 2003.

[17] Calthorpe P. The next American metropolis: Ecology, community, and the American dream[M]. Princeton architectural press, 1993.

[18] Planning Commission TOD Committee of Fairfax County. Walking Distance Research[R]. Fairfax: Fairfax County Planning Commission, 2006.

[19] O’Sullivan S, Morrall J. Walking Distances to and from Light-rail Transit Stations[J].

Transportation Research Record, 1996(1): 19-26

[20] Hyungun Sung,Keechoo Choi,Sugie Lee,SangHyun Cheon. Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership[J]. Journal of Transport Geography,2014,36.

[21] Wang JY, Zheng X, Mo YK. Construction of density zoning and determination of volume ratio for rail transit TOD development--Shenzhen rail transit line 3 as an example[J]. Urban Planning, 2011(4): 30-35.

[22] Du Caijun,Jiang Yukun. Exploration on the law of connecting urban rail transit with other transportation modes[J]. Urban Rapid Transit, 2005(03):50-54.

[23] Huang, Libin. Evaluation of the impact of rail transit stations on regional development in large cities[D]. Master’s thesis. Tongji University, 2006.

[24] Huang Weidong,Su Xixi. Research on the construction mode of public transport community based on TOD theory--Hangzhou as an example[J]. Journal of Urban Planning,2010(S1):151-156.

[25] Wang JY, Zheng X, Mo YK. Construction of density zoning and determination of volume ratio for rail transit TOD development--Shenzhen rail transit line 3 as an example[J]. Urban Planning, 2011(4): 30-35.

[26] Zhang N, Dai J, Zhang XJ. Walking connection range of rail transit stations based on multinomial logit model[J], Urban Transportation Research,2012,15 (5):46-49.

[27] Duan Degang,Zhang Fan. Study on the classification of urban rail stations from the perspective of land use optimization--Xi'an Metro Line 2 as an example[J]. Urban Planning,2013,37(9):39-45.

[28] Guo Daqi, Gao Feng, Li Xiaoyu . Research on the land development mode around urban rail transit stations: an example of land development around Shenyang metro stations [C] Proceedings of China Urban Planning Annual Conference. Beijing: China Construction Industry Press, 2014.

[29] Ma Deqin, Lin Anlin. Underground Railways and Light Rail Transit [M]. Southwest Jiaotong University Press.2003:66-71.

[30] Yan Xiaopei, Zhou Suhong, Mao Jiangxing. Transportation system and land use in high-density development cities: A case study of Guangzhou [M]. Beijing: Science Press,2006.

[31] Xie Dis. Spatial integration around urban rail transit stations [D]. MS thesis. Nanjing University, 2011.

[32] Zheng M. Urban Development in the Age of Rail Transit [M]. China Railway Press, 2006.

[33] Fu Baofeng, Wu Jiao Rong, Chen Xiao Hong. Research on the classification method of suburban rail stations[J]. Journal of Railway, 2008, 30(6): 19-23.

[34] Huiying. Planning and Construction of Urban Rail Transit Station Areas[J]. Journal of Urban Planning, 2002(4).37-43.

[35] Zhang Yunan. A preliminary study on the development of urban rail transit and urban space integration[J]. Central China Architecture, 2011 (10): 89-91.

[36] Zheng Wenhan. Exploring the development intensity of different types of rail transit station areas[J]. Urban Development Research, 2008 (S1): 93-95.

[37] Coombs D B, Thie J. The Canadian land inventory system[J]. Planning the Uses and Management of Land, 1979, 21: 909-933.

[38] Tomlinson R F. An introduction to the geographic information system of the Canada Land Inventory[J]. Department of Forestry and Rural Development, Ottawa, Canada, 1967.

[39] Dangermond J. Trends in GIS and comments[J]. Computers, Environment and Urban Systems, 1988, 12(3): 137-159.

[40] Guan, Y. Research on the development of GIS technology and its application. Applied Mechanics and Materials, 2014,522: 1017-1020.

[41] Haining R. Spatial data analysis in the social and environmental sciences[M]. Cambridge University Press, 1993.

[42] Wang JF, et al. Spatial analysis [M]. Science Press, 2006.

[43] Zhang KQ, Guo RZ. Mathematical modeling of thematic mapping [M]. Beijing: Surveying and Mapping Press, 1991.

[44] Unwin D J, Doornkamp J C. Introductory spatial analysis[M]. London: Methuen, 1981.

[45] Bailey T C, Fotheringham S, Rogerson P. A review of statistical spatial analysis in geographical information systems[J]. Spatial analysis and GIS, 1994: 13-44.

[46] Zhong Yexi. Accessibility study of urban spatial pattern evolution [M]. Southeast University Press, 2012.

[47] Christaller W. Central places in southern Germany[M]. Prentice-Hall, 1966.

[48] Garrison, W. L. A theory of central places. Papers and Proceedings of the Regional Science Association, 1967,19: 45-68.

[49] Curran, L. M. The extension of central place theory to urban hierarchies. Geographical Analysis, 1973, 5: 226-237.

[50] Berry B J L. Cities as systems within systems of cities[J]. Papers in regional science, 1964, 13(1): 147-163.

[51] PL Knox, PL Knox, PJ Taylor, PL Knox. World cities in a world-system[M]. Cambridge University Press, 1995.

[52] Sassen S. The global city: Introducing a concept[J]. Brown J. World Aff., 2004, 11: 27.

[53] Taylor P, Derudder B. World city network: a global urban analysis[M]. Routledge, 2015.

[54] Derudder B, Taylor P, Ni P, et al. Pathways of change: Shifting connectivities in the world city network, 2000—08[J]. Urban studies, 2010, 47(9): 1861-1877.

[55] Scott, Allen J., ed. Global city-regions: trends, theory, policy[M]. OUP Oxford, 2001.

[56] Nilsen E T, Rundel P W, Sharifi M R. UC Press E-Books Collection, 1982-2004[J].

[57] Attoe W, Logan D. American urban architecture: catalysts in the design of cities [M]. California: University of California Press,1989.

[58] Bohannon C L. The urban catalyst concept[D]. Virginia Tech, 2004.

[59] Hong L, Jiang Y, Mo Z, et al. Research of the impacts of the rail transit on the development of the commercial space[C]//2015 12th International Conference on Service Systems and Service Management (ICSSSM). IEEE, 2015: 1-4.

[60] Calthorpe P. The Next American Metropolis: Ecology, Community, and the American Dream[M]. New York: Princeton Architecture Press, 1993.

[61] Cervero R. Transit-oriented development in the United States: Experiences, challenges, and prospects[J]. 2004.

[62] Encyclopedia of the City[M]. Taylor & Francis, 2005.

[63] Yu W-B, Wang Z, Meng H-N. The "unitary community" in China vs. the TOD community in the United States[J]. Urban Planning, 2007, 31(5): 57-61.

[64] Nafi S, Furlan R, Grosvald M, et al. Transit-oriented development in Doha: the case of the Al Sadd neighborhood and Hamad hospital metro station[J]. Designs, 2021, 5(4): 61.

[65] Salesforce Transit Center. https:// industrial. sherwin-williams.com/na/us/en/coil-extrusion/ media-center/ case-studies/ salesforce-transit-center.html (accessed on November 05, 2022).

[66] Salesforce Transit Center. https://salesforcetransitcenter.com (accessed on November 05, 2021).

[67] Maeda, Atsuo. “Aiming for the Creation of New Space: Development of Tokyo Station City”. JR East Technical Review, 10 (2007).

[68] Nikken Sekkei ISCD Study Team. (2014). Integrated station-city development: The next advances of TOD. China Building Industry Press.

[69] Yue Y, Chang J. Station-City Integration: Urban Space Ecological Transformation Research Based on Rail Transit[J]. Human-Centered Urban Planning and Design in China: Volume II: Urban Design and Mobility, 2021: 41-64.

[70] Creating a vibrant Shibuya full of charm and realizing a three-dimensional urban neighborhood. https://www.nikken.co.jp/cn/projects/mixed_use/shibuya_scramble_square_the_first_phases.html (access on October 22, 2022)

[71] The World Bank Population Density Database. https://data. worldbank.org.cn/indicator/EN. POP.DNST?locations=HK (Accessed on November 1, 2022).

[72] Yin Ziyuan.. From “Rail + Property” to Compact Rail City: Story behind Statistical Data of MTR [J]. The Architect, 2018, No.195(5): 55-60.

[73] Cervero R, Murakami J. Rail and property development in Hong Kong: Experiences and extensions[J]. Urban studies, 2009, 46(10): 2019-2043.

[74] Xue Quli, Zhai Hailin, Chen Beiying. The floating city island on the subway station: A case study of Kowloon station development in Hong Kong[J]. Journal of Architecture,2010(7):82-86.

[75] Zheng J. and Yan Kefei. Research on the integrated development mode of rail stations based on TOD concept - taking Changsha Binjiang New City as an example[J]. Proceedings of the Workshop on Transportation Planning for Large Cities in China

[76] Yu Xiaohua, Zheng Jian, Zhong Ping. A case study of Site Integrated Development "SID" - A visit to Lyon High Speed Rail Hub [J]. Traffic and Transportation, 2012,28(1): 40-42.

Chapter 3. The construction of research methodology framework

Content

Introduction

This chapter is the construction of the research methodological framework, focusing on five parts, including station realm scope, station realm type, source and access to positioning data, spatial distribution pattern, and interaction mechanism. First, based on the relevant literature study in Chapter 2.2, the method of determining the scope station realm scope and the classification method of station realm type is determined. Second, the reasons for the selection, sources and access methods of positioning data representing population activities and functional facilities are explained. Third, based on the attribute characteristics of the data, the spatial distribution pattern analysis methods suitable for population activities and functional facilities were selected respectively. Finally, the analysis methods for exploring the interaction mechanism are mainly based on correlation analysis and spatial regression modeling. The research methodological framework constructed in this chapter is to provide technical support for the next empirical research.

3.1 Defining method of station realm scope

3.1.1 Technology route

In this study, based on the methods used by many scholars in the past to determine the scope of the influence realm of rail transit stations, the relevant data to be investigated in this section of the study have been clarified through the elemental analysis in Chapter 2.2. Therefore, this survey uses questionnaire survey method, observation record method, and electronic map measurement method to investigate the data on pedestrian walking time, walking speed, and environmental characteristics around each station of Line 2. The specific technology route is shown in Figure 3- 1.

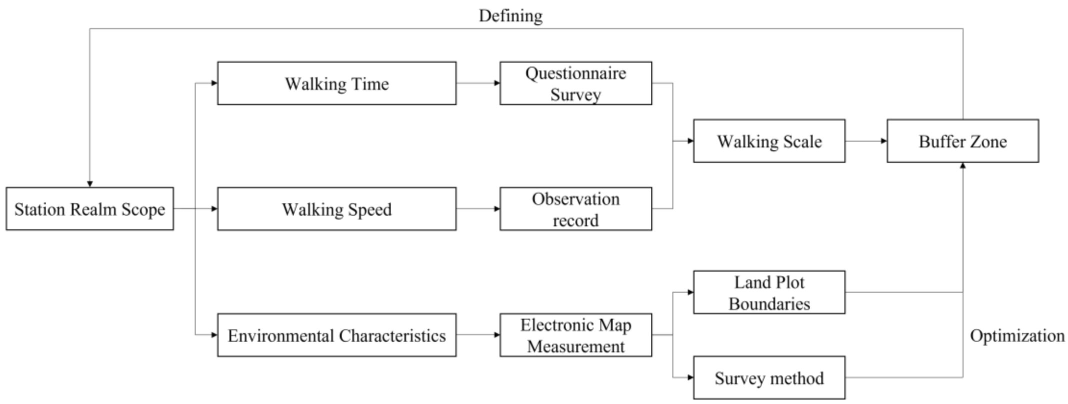


Figure 3- 1. Technology route of defining station realm scope

3.1.2 Survey method of data

The specific survey contents and methods are as follows:

(1) Survey walking time

The thesis will use a questionnaire to investigate the maximum acceptable walking time for pedestrians around the station. The surveyors were located near the entrances and exits of the subway stations and distributed questionnaires to the pedestrians, the style of which is shown in Appendix A. In this study, the survey data will be counted and analyzed using Excel, and the highest percentage of time will be selected as the maximum acceptable walking time for pedestrians, which will be set as T_p.

(2) Measuring walking speed

Set the observation distance to 30 meters, use the electronic stopwatch to observe the walking time of pedestrians around the station, and calculate the walking speed based on the observation distance and walking time, the specific equation is:

$$V_p = L/T_o \qquad 3\text{- }1$$

Where, V_p is the walking speed of a single sample, L is the observation distance, T_o is the walking time required for the pedestrian to pass the observation distance. Use Excel to count the sample walking speed and calculate its arithmetic mean, set as V_m.

(3) Access geographic data

In this study, land boundary lines and urban road networks are used to adjust the boundaries of the station realms, whose geographic data are mainly derived from OpenStreetMap (OSM) - an electronic map that can be edited online and exported by anyone [1]. As show in Figure 3- 2, the geographic information data of land boundary line and road network around the stations were exported in Shapefile format by accessing OSM, and then they were imported into ArcGIS to establish a geo-database and saved (Figure 3- 3).

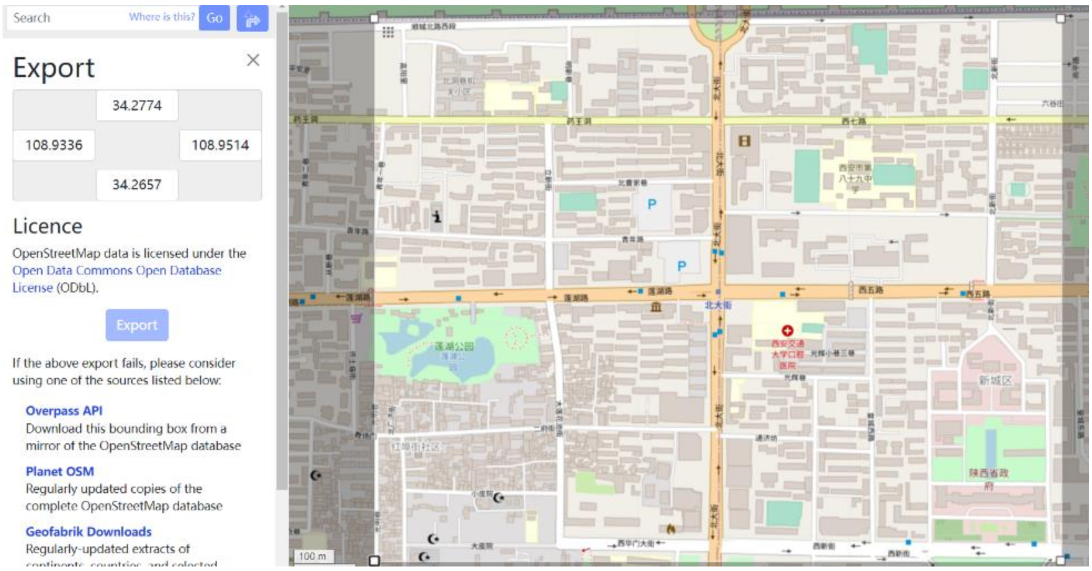


Figure 3- 2. Exporting geographic data from the OSM

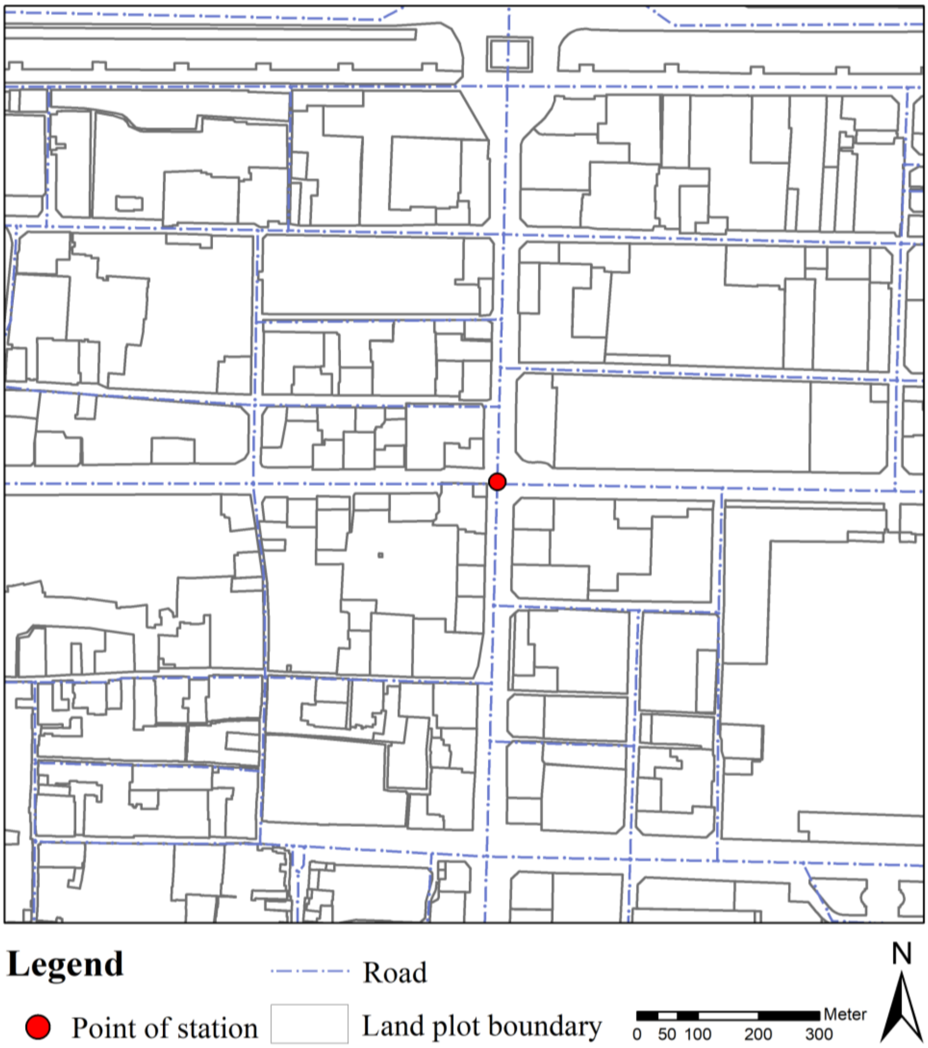


Figure 3- 3. Geo-database in ArcGIS

3.1.3 The defining method for station realm scope

First, from the survey data, the obtained pedestrian travel data are walking time and walking speed respectively, so the walking distance can indicate the walking scale, then the walking scale is calculated as follows:

$$L_p = T_p \times V_m \qquad 3\text{- }2$$

Where, L_p is the pedestrian's walking scale, T_p is the pedestrian's walking time, V_m is the pedestrian's walking speed.

Second, the calculated walking scale is used as the radius and the station is the center of the circle, and ArcGIS buffer analysis is used to generate a circular buffer for each station, as shown in Figure 3- 4. a.

Finally, the boundary of the buffer zone is optimized according to the characteristics of the land plot boundaries and roads within the buffer zone, so as to obtain the extent of the station influence realm of Xi'an Metro Line 2 as shown in Figure 3- 4. b.

Figure 3- 4. The scope optimization process of the station realm. (a) Before optimization, (b) After optimization

3.2 Classification method of station realm type

3.2.1 Technology route

The results of the literature study from Chapter 2.2.2 show that the classification system of rail transit stations selected for different research purposes is different, especially in the selection of indicators that mainly consider the passenger flow, the number of station entrances and exits, the number of bus lines around the station, and other transportation function factors. However, for the station realm which belongs to the urban space at the micro level, the urban spatial form and spatial density around the station are also important characteristic elements [2,3]. Therefore, based on considering the transportation function elements, this paper adds the indicators such as building form, road network form, population density and land use intensity into the classification system, then uses cluster analysis to classify the station realm, so as to make the classification results more detailed and objective. The specific technology route is shown in Figure 3- 5.

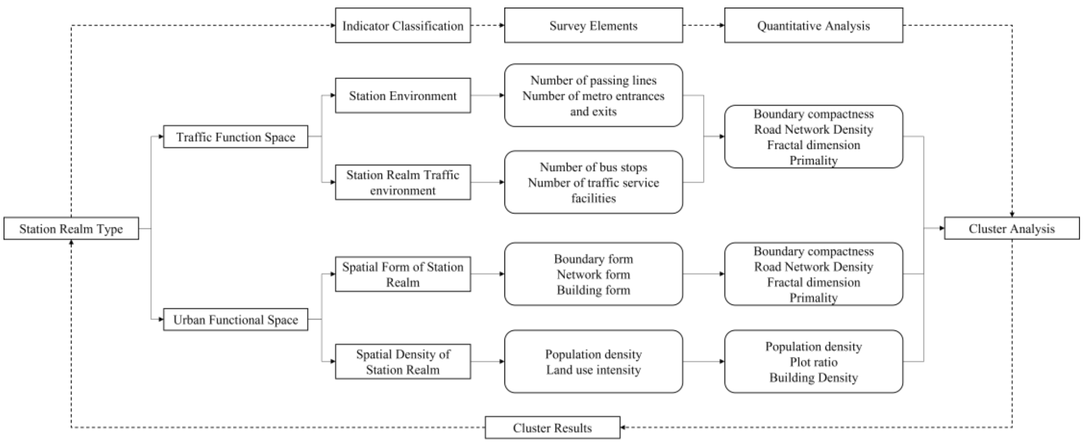


Figure 3- 5. Technology route of station realm type classification

3.2.2 Survey methods for classification indicators

According to the literature research on survey methods, the selection of survey methods should be easy to operate, widely applicable, and have practical value. Based on the research findings related to the classification of rail transit station realm in Section 2.2.2, this study has determined the various indicators shown in Table 3- 1 and their corresponding survey methods.

1. Electronic map measurement: This method involves using electronic maps to assist in measuring objects that cannot be surveyed on-site.

2. Photo sampling: The survey of building facades is recorded in the form of photographs, which are numbered and well-documented for ease of processing in later stages.

3. On-site recording: Through on-site visits, public spaces, bus stops, and other locations within the station realm are marked. Information such as the number of floors and land use types of buildings within the station realm is recorded. The on-site records are then compared with electronic maps for calibration.

4. Literature survey: Relevant reports, yearbooks, and other documents are searched to gather statistical data such as population figures.

5. GIS calculation: Basic data is processed using a ArcGIS platform to obtain secondary data, primarily used for calculating area, perimeter, and other measurements in bulk.

Table 3- 1. Station realm spatial classification indicators and survey methods

Classification Indicators	Secondary indicators	Survey Methods
Station conditions	Interchangeable rail lines	On-site recording
	Station entrances and exits	On-site recording
Station realm traffic environment	Number of bus stops	On-site recording
	Quantity of transportation facilities	On-site recording
Spatial form	Station realm size	GIS calculation
	Perimeter of the station realm outline	GIS calculation
	Length of road networks at various levels	Electronic map measurement
	Number of building floors	On-site recording
	Building facade outlines	Photographic sampling GIS calculation
	Building height	On-site recording
Spatial density	Resident population	Statistical yearbooks
	Land use types	On-site recording
	Land area	GIS calculation
	Building area	GIS calculation
	Building site area	GIS calculation

3.2.3 Quantification method of classification indicators

(1) Quantification method of traffic space

The data of station condition indicators, such as the number of entrances and exits at metro stations and the number of transit lines, can be obtained directly from the official metro website.

(2) Quantification method of spatial form

In terms of spatial form, the urban planar form typically refers to the geometric shape of urban land in the plane space. The compactness ratio of the outer contour shape of the city is considered a crucial concept that reflects the spatial form of the city [4,5,6]. It can serve as an indicator representing the compactness and fullness of the urban built-up area, used for comparing and analyzing urban forms. Generally, during the rapid expansion phase of the rail transit station realm, the compactness of its outer contour shape decreases. When the area transitions to an internal infill and redevelopment phase, its compactness tends to increase. There are various methods to calculate the compactness of urban forms, with the compactness formulas proposed by Cole (1964) and Batty (2001) being the most widely used. In this study, the circularity method based on perimeter is employed [7], Its measurement equation is:

$$C_I = 2\sqrt{\pi A}/P \quad 3\text{-}3$$

where, the C_I (Compaction Index) represents the compactness of the city within the station realm, where A is the station realm's area and P is the perimeter of the station realm boundary. The C_I value ranges between 0 and 1, with a value of 1 indicating a perfectly circular shape, representing the highest level of compactness. As C_I approaches 0, the shape of the area becomes more linear and less compact.

Urban transportation relies heavily on roads, which serve as the most essential infrastructure for the operation of urban transportation systems. Therefore, the urban road network forms the backbone of a city, and the structural form of roads determines the network morphology of the city. Road network density refers to the network morphology composed of roads of different functions, levels, and locations within a certain range, with a certain density and appropriate form [8]. The road network density indicator reflects both the quantity and level of road construction within a given area and to some extent reflects the rationality and balance of road network layout [9,10]. It is considered one of the ideal indicators for evaluating regional network morphology. Road network density can generally be calculated by dividing the length of the road network centerlines within the station realm by the station realm's size. This calculation includes main roads, secondary roads, and local roads within the urban road network but excludes roads within residential areas. The equation is as follows:

$$P = L/A \quad 3\text{-}4$$

Where P is the density of the road network, L is the total length of the road network and A is the area of the station realm. A higher value of P indicates a denser road network within the station realm.

The combination and distribution of buildings in urban space form the skyline, which is the overall image of the city characterized by its undulating, varied heights, and clear distinction between primary and secondary elements. A distinctive skyline with these characteristics contributes to the unique architectural form of the region [11,12]. Since the birth of fractal theory in 1977, it has been widely applied in the field of urban spatial morphology. The fractal dimension is the primary parameter used to describe a fractal [13]. It represents the size of the space occupied by the fractal set in Euclidean space and reflects the complexity of the set. The greater the fractal dimension of a fractal object, the more complex and intricate it is, while a smaller fractal dimension corresponds to a smoother object. Therefore, based on the comparative analysis of relevant cases in Chapter 2.3, a mature and developed station realm skyline should exhibit undulations, a well-balanced density, and a rich hierarchical structure. It should consist of a main group of buildings and several secondary building groups, with each group arranged in a varied and well-balanced manner. Such an undulating and richly layered skyline would obtain a relatively high fractal dimension value.

The box-counting dimension method is a widely used approach for calculating fractal dimensions and is among the most applied methods. Its calculation logic involves counting the number of boxes (grid cells) occupied by the fractal shape at different sizes. The method then calculates the ratio between the number of boxes and the corresponding scale of the boxes. Finally, a linear regression is performed to determine the slope, which is used to calculate the fractal dimension. The equation is as follows:

$$D_B F = \lim_{\delta \to 0} \frac{\log N_\delta(F)}{-\log \delta} \qquad 3\text{- }5$$

where F is the skyline contour of the study area, $D_B F$ is the fractal dimension of the skyline contour, δ is the side length of each box of the grid, and $N_\delta(F)$ is the number of boxes occupied by the skyline contour. Due to the huge statistics of box number, fractal calculation software is usually used to calculate the box dimension. In this paper, we calculate the fractal dimension of the skyline of the rail transit station realm the help of MATLAB mathematical software.

The prominence ratio reflects the degree to which the tallest building group in the city skyline stands out visually. It is an important indicator for assessing the recognizability and dominance of the core building group within a region [14]. In Kevin Lynch's book "The Image of the City," he proposed five elements that contribute to the recognizability of a city, with landmarks being one of them [11]. This means that a higher prominence ratio results in a stronger visual and landmark effect of the core high-rise building group, leading to better recognizability and dominance. The prominence ratio is calculated by dividing the difference between the height of the tallest building and the average height of all buildings by the height of the tallest building. The equation is as follows:

$$D = (H_{max} - H_p)/H_{max} \qquad 3\text{- }6$$

where, D is the prominence ratio of the building group within the station realm, H_{max} is the height of the tallest building, and H_p represents the average building height. The higher the ratio, the higher the prominence ratio, indicating a greater visual prominence of the tallest building group in relation to the average building height.

(3) Quantification method of spatial density

Population density refers to the average number of people living in a certain area of land during a specific period. It is an important indicator for assessing the population distribution in an area and is commonly used to measure the population distribution in a country, city, or region. Population density assumes that the population is evenly distributed within the specified geographic area.

Therefore, the smaller the scope of population density calculation, the more accurately it reflects the actual population distribution. The equation is as follows:

$$D_p = N_p / A \qquad 3\text{-}7$$

Where D_p is Population density within the station realm, N_p is Population count within the station realm, A is Area of the station realm (in square kilometers), A higher D_p value indicates a higher population density within the station realm, indicating a more densely populated area.

The different types of land use in cities carry out different activities, resulting in varying levels of traffic generation. In urban geography, land use mix can be quantified using principles from information theory, where the magnitude of entropy represents the degree of mixing. A higher entropy value indicates a more balanced distribution of land uses in an area, while a lower value suggests a more concentrated allocation of land uses and a lower degree of land use mix [15,16,17]. Referring to Knaap's calculation method for land use mix [18], the formula for calculating land use mix is as follows:

$$S = -\sum_{i=1}^{n} P_i \ln P_i \qquad 3\text{-}8$$

where S is entropy value representing the degree of land use mix, Pi is proportion of the area occupied by the i-th land use type related to travel, relative to the total land area of the area, n is number of different land use types related to travel within the sample area.

Land use intensity refers to the spatial distribution and intensity of various components that make up a city, such as population, employment, buildings, economic activities, and social structure, on a per unit area basis [19,20]. Plot ratio and building density are commonly used indicators to reflect land use intensity within a certain area [21]. Plot ratio represents the ratio of total built-up area to net land area within a region. The formula is as follows:

$$P_r = A_b / A_l \qquad 3\text{-}9$$

where, P_r is Plot ratio within the station realm, A_b is total built-up area of the individual plots within the area, A_l is total land area of the individual plots within the area.

Building density refers to the proportion of the total footprint area of buildings to the land area they occupy within a certain range (%). It reflects the open space ratio and the degree of building intensity within a specific land area.

$$D_b = A_m / A_l \qquad 3\text{-}10$$

where:

D_b - Building density within the station realm,

A_m - Total footprint area of the individual plots within the area,

A_l - Total land area of the individual plots within the area,

A higher D_b value indicates a higher proportion of building footprint area to land area within the station realm, resulting in a lower open space ratio.

The land use types in this study are based on the "China Urban Land Classification and Planning Construction Land Standards"[22]. The specific subcategories and categories of land use are shown in Table 3- 2.

Table 3- 2. Code for classification of urban land use and planning standards of development land

Code	Name	Secondary Code	Name
R	Residential land	R2	Second-class residential land
		R3	Third-class residential land
A	Administration and public services land	A1	administrative office land
		A2	cultural and recreational land
		A3	Education and research land
		A4	Sports land
		A5	Medical and health land
		A6	Social welfare land
		A7	Cultural relics and monuments land
		A9	Religious land
B	Commercial and business facilities land	B1	Commercial land
		B2	Business land
		B3	Recreation and sports land
		B4	Land for public utility business outlets
		B9	Other service facilities
M	Industrial and manufacturing land		
S	Street and transportation facilities land		
W	Logistics warehousing land		
U	Municipal utilities land		
H	Development land		
G	Green space and square land	G1	Public Green Space
		G2	Protective green space
		G3	Square land

3.2.4 The method of cluster analysis method

At present, there have been scholars using the method of cluster analysis, selecting different classification indexes to classify urban rail transit stations for research [23,24,25]. This provides a quantitative and more detailed method for station classification, which is more operable in practice.

In statistics, clustering analysis is often used to group similar individuals together into larger categories. During the analysis process, based on the similarity of several or multiple characteristics collected from individuals, clustering analysis automatically identifies and assigns individuals into different categories, forming clusters. Within the same cluster, individuals share high similarity in terms of their properties and features, while there are significant differences between individuals from different clusters.

For this clustering analysis, the K-means clustering algorithm [26] is used. The basic principle of the K-means algorithm is as follows: given a dataset and the desired number of clusters k (specified by the user), the algorithm repeatedly assigns data points to k cluster categories based on a chosen distance function. The general steps of the K-means algorithm are as follows:

1. Randomly select k objects as the initial cluster centers.

2. Calculate the distance between each object and the seed cluster centers and assign each object to the nearest cluster center.

3. After the assignment of individuals, calculate the mean of all individuals within each cluster as the new center for that cluster.

4. Recalculate based on the current cluster centers and repeat the process until all individuals are assigned to their respective clusters.

5. This process is repeated until no further assignment can be made.

Before clustering the data, it is necessary to calculate the similarity measure. In this study, the squared Euclidean distance is used as the similarity measure. The specific formula is as follows:

$$d_{ij} = \sum_{k=1}^{m} (x_{ik} - x_{jk})^2 \quad 3\text{- }11$$

where d_{ij} is the distance between the *i*-th sample and the *j*-th sample; x_{ik},x_{jk} are the *k*-th indicator in the *i*-th and *j*-th samples, respectively.

3.3 Methods of collecting data on population activity and functional facility

3.3.1 Technology route

In this study, data on population activity and facility amenities were collected within the boundaries of each station realm. The population activity data was collected using the WeChat mini applet called EasyGO, while the facility amenities data was collected using the API provided by AutoNavi Map. The collected data was stored in CSV format. To process the data, the point conversion tool in ArcGIS was utilized, and the appropriate coordinate system was set. Finally, a database was established for each station realm.

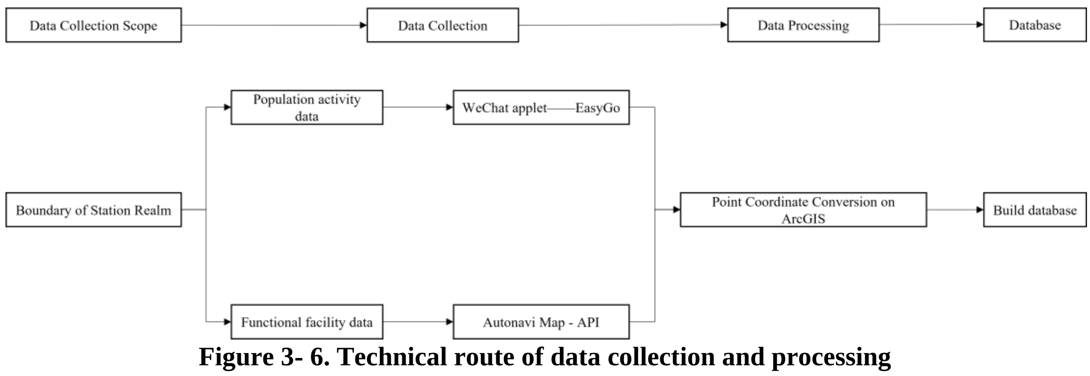


Figure 3- 6. Technical route of data collection and processing

3.3.2 Population activity data

Mobile phone signaling technology is usually used to extract OD data of residents' traffic travel [27], but the distribution density of communication base stations is small (the coverage radius of urban areas is generally 100~500m) [28]. Using the base station to send and receive data may cause spatial positioning errors, which in turn affects the accuracy of crowd identification. The AFC system data is only applicable to the statistics of the total passenger flow of the rail transit network, and the real-time passenger flow distribution cannot be obtained [29]. In addition, the use of social media platforms such as Facebook, Twitter, and Weibo to collect analytical data related to population activity trajectories is not sufficient to represent the activity characteristics of people of all ages since most of the contributors are young people [30]. The Wi-Fi signaling technology needs to plan and deploy many signals equipment to ensure full coverage of signals in the station realm region, in order to ensure the accuracy of data collection, but greatly increase the research cost [31].

In this study, we choose the mobile terminal program developed by WeChat, EasyGO applet. WeChat is an instant messaging software launched by China's Tencent on January 21, 2011, that supports mobile operating systems such as Android and IOS. EasyGO data has the following advantages:

(1) Strong real-time performance. The data of WeChat comes from the GPS positioning information of the mobile intelligent terminals of the crowd, it has the characteristics of dynamic update and real-time feed-back [32].

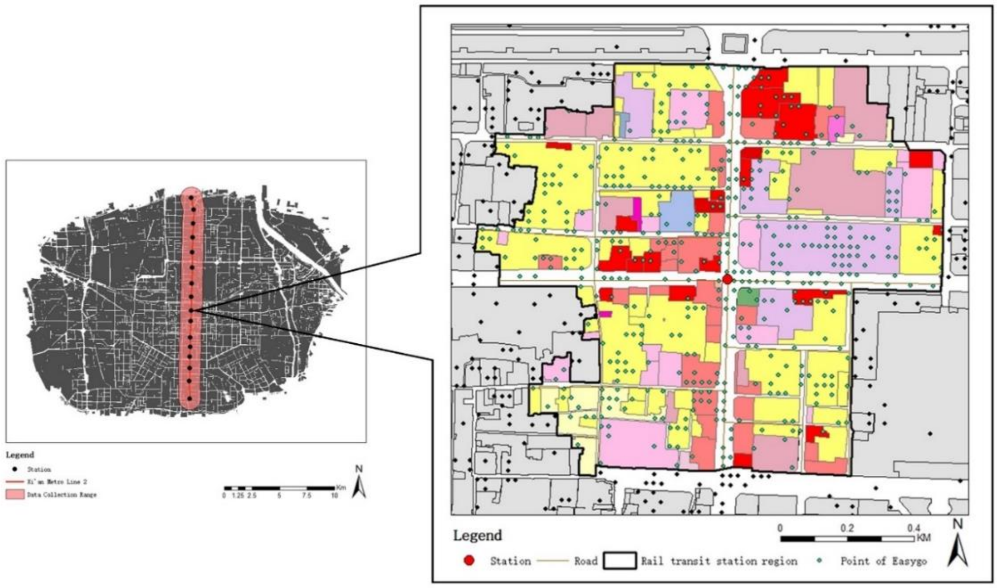
(2) High accuracy. The extraction range of the data of EasyGO is only 25m×25m grid, which is more suitable for research at the micro-scale [33].

(3) High coverage. The positioning data of EasyGO mainly comes from WeChat, which covers a wider population in China and has strong utilization value. As of June 30, 2021, the software's monthly users active (MUA) exceeded 1.2 billion, making it the most active social software in China [34]. In addition, WeChat accounts for about 69.2% of China's communication software [35], excluding children, elementary school students and the elderly who rarely use mobile phones, the data can fully cover various activities of the population of all ages. It has been shown that there is a high correlation coefficient of 0.9 between Tencent location data and the residential population [36].

(4) Easy availability. EasyGO directly provides an API interface to make data easier to obtain directly.

Therefore, by recording the real-time location information of WeChat users, it gives the real-time pedestrian location in the form of spatial point data, which has the characteristics of low acquisition cost, high spatial resolution, and real-time dynamic change. Therefore, the above points are the reasons why the data of EasyGO can reflect the spatial distribution characteristics of population activity in this research.

The EasyGO data for population activities are sourced from the Tencent Location Big Data Service window (https://heat.qq.com/index.php). As shown in the Figure 3- 7, the data is collected at the boundaries of each station realm. All the above data were converted from raw data to point data based on latitude and longitude information using ArcGIS platform and were summarized and integrated according to the names of the stations to prepare for the next stage.



(a) (b)

Figure 3- 7. Data collection area and partial data. (a) Data collection area of Metro Line 2; (b) Point data distribution of population activities.

3.3.3 Functional facilities data

(1) Data source

POI (Point of interest) data mainly includes spatial information and attribute information data of geographic entities closely related to people's lives. It can quickly and intuitively obtain the distribution of various functions in urban space, which is superior to traditional research data [37]. In addition, different from the previous studies on the relationship between rail stations and surrounding urban land through the nature of land use, POI data is more microscopic and has more advantages in data accuracy, which can more intuitively reflect the spatial agglomeration characteristics of urban functions [38]. In this study, POI reflects the spatial distribution of urban functions in the station realm and conducts correlation analysis with the spatial activity characteristics of pedestrians.

The data for POIs in this paper are sourced from the AutoNavi Map's Public Application Programming Interface (API) (https://lbs.amap.com).

(2) Data collection

The POI data were also collected within the boundaries of each research station realm, ten categories that are closely related to daily life were selected according to the primary classification codes based on the National Economic Industry Classification (GB/T 4754-2017) issued by the General Administration of Quality Supervision, Inspection and Quarantine of China [39]. The quantity statistics of POI data are shown in Appendix B.

All the above data were converted from raw data to point data based on latitude and longitude information using ArcGIS platform and were summarized and integrated according to the names of the stations to prepare for the next stage.

3.4 Analysis methods of spatial distribution pattern

3.4.1 Technology route

Based on the literature research on spatial distribution pattern recognition methods in section 1.2. and the detailed data description in section 3.3, In this study, the choice of analytical methods for identifying spatial patterns is based primarily on the attribute values assigned to the data. Since the population data for travel suitability consists of point data with attribute values, where each point represents a specific population count, spatial autocorrelation analysis is more suitable. On the other hand, the POI data from Amap (Gaode Maps) for functional facilities do not have attribute values, with each point representing a single facility, making spatial point pattern recognition analysis more appropriate. The specific technical approach is illustrated in the Figure 3- 8.

Regarding the spatial distribution pattern of population activity in station realms, the first step is to conduct global spatial autocorrelation analysis to assess the spatial distribution status of population activity. If there is evidence of clustering, the specific type of clustering can be identified. Subsequently, local spatial autocorrelation analysis is used to identify spatial distribution location for different types of population activity clusters.

For the spatial distribution pattern of functional facilities in station realms, the average nearest neighbor method is employed to assess the distribution status of facilities in the station realm. Kernel density analysis is then utilized to visualize the distribution location and provide an intuitive representation of the spatial distribution characteristic of facilities.

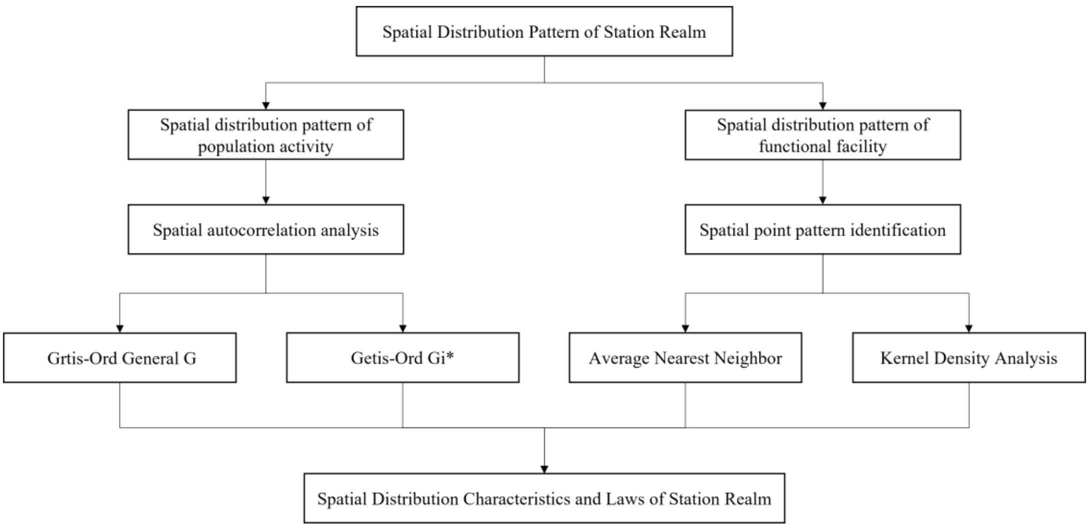


Figure 3- 8. Technology route of spatial distribution pattern

3.4.2 Spatial correlation analysis method of population activities

(1) Global spatial autocorrelation analysis method

Currently, global spatial autocorrelation statistics mainly include global Moran's I statistic and High/Low Clustering (Getis-Ord General G statistic). Compared to the global Moran's I statistic, high/low clustering not only identifies whether the data are spatially clustered or dispersed, but also determines the type of clustering or dispersion, e.g., whether it is a high-value clustering or a low-value clustering [40,41]. Therefore, this study will use the High/Low Clustering to identify distribution status of population activities in the station realm. The formula is as follow:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} x_i x_i}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_i}, \forall j \neq i \qquad 3\text{-}12$$

Where x_i and x_i are attribute values for features i and j, and $w_{i,j}$ is the spatial weight between feature i and j. n is the number of features in the dataset and $\forall j \neq i$ indicates that feature i and j cannot be the same feature.

High/Low Clustering tests the statistical significance of spatial autocorrelation using z_G- score [42]. As shown in Figure 3- 9, a positive z_G- score indicates high/high clustering, while a negative value indicates low/low clustering, meaning high population activity clustering with high population activity or low population activity clustering with low population activity. The z_G- score for the statistic is computed as:

$$z_G = \frac{G - E[G]}{\sqrt{V[G]}} \qquad 3\text{-}13$$

Where:

$$E[G] = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}}{n(n-1)}, \forall j \neq i \qquad 3\text{-}14$$

$$V[G] = E[G^2] - E[G]^2 \qquad 3\text{-}15$$

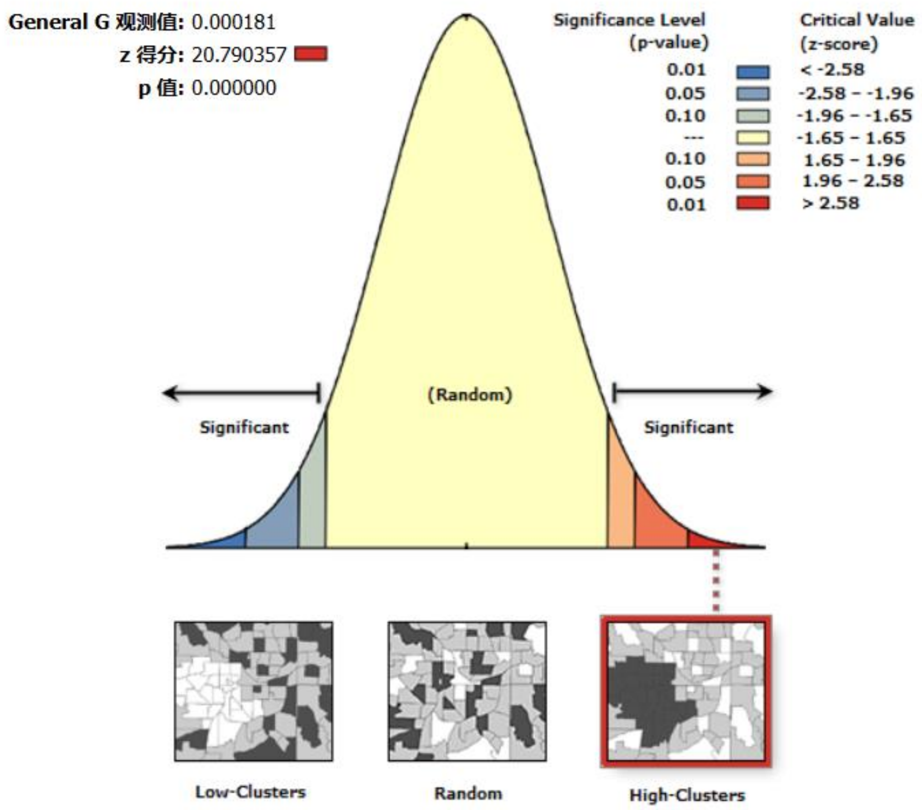


Figure 3- 9. Results of high/low clustering statistics (Grtis-Ord General G) analysis

(2) Local spatial autocorrelation analysis method

The local spatial autocorrelation analysis mainly consists of Local Moran's I [43] and Hot Spot Analysis (Getis-Ord Gi*) [44]. Each method has its own advantages: Hot Spot Analysis can accurately detect clustering areas, while Local Moran's I generally have larger identification bias for the clustering extent and can only roughly identify the center of clustering areas, with the identified range usually smaller than the actual range [45,46]. Under the same conditions, Hot Spot Analysis has a better ability to identify high-value clusters compared to Local Moran's I [47]. Therefore, this study primarily utilizes Hot Spot Analysis with the Getis-Ord Gi* index to identify the spatial distribution areas of various clusters in station realms' population activities.

The equation for calculating the Getis-Ord Gi* statistic is as follows:

$$G_i^* = \frac{\sum_{j=1}^{n} w_i x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{s\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \quad 3\text{- }16$$

Where x_j represents the attribute value of feature j, $w_{i,j}$ represents the spatial weight between feature i and feature j, and n is the total number of features.

And：

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n} \quad 3\text{- }17$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - (\bar{X})^2} \quad 3\text{- }18$$

The Gi* statistic returned for each feature in the dataset is a z_{Gi}- score. For statistically significant positive z_{Gi}- score, the larger the z-score is, the more intense the clustering of high values (hot spot). For statistically significant negative z_{Gi}- score, the smaller the z-score is, the more intense the clustering of low values (cold spot) [48].

3.4.3 Spatial point pattern identification method

Spatial point pattern identification techniques such as average nearest neighbor and kernel density analysis have been widely applied to study the spatial distribution characteristics of urban elements [49,50,51,52]. Since spatial point pattern recognition techniques can only analyze data without attribute values, and POI data is in line with this requirement, this study will use average nearest neighbor and kernel density analysis to study the spatial distribution status and characteristics of functional facilities within the station realm.

(1) Average nearest neighbor

The Average Nearest Neighbor refers to the average distance between points [53]. This analysis method compares the average distance between nearest neighboring point pairs with the average distance between nearest neighboring point pairs in a random distribution pattern to determine the spatial pattern [54]. Therefore, this study will use the average nearest neighbor to determine the spatial distribution status of functional facilities within the station realm.

The Average Nearest Neighbor ratio is given as:

$$ANN = \frac{\overline{D}_O}{\overline{D}_E} \quad 3\text{-}19$$

Where $\overline{D}_O$ is the observed mean distance between each feature and its nearest neighbor:

$$\overline{D}_O = \frac{\sum_{i=1}^{n} d_i}{n} \quad 3\text{-}20$$

and $\overline{D}_E$ is the expected mean distance for the features given in a random pattern:

$$\overline{D}_E = \frac{0.5}{\sqrt{n/A}} \quad 3\text{-}21$$

In the above equations, d_i equals the distance between feature i and its nearest neighboring feature, n corresponds to the total number of features, and A is the area of a minimum enclosing rectangle around all features, or it's a user-specified Area value.

The average nearest neighbor Z_A-score for the statistic is calculated as:

$$Z_A = \frac{\overline{D}_O - \overline{D}_E}{SE} \quad 3\text{-}22$$

where:

$$SE = \frac{0.26136}{\sqrt{n^2/A}} \quad 3\text{-}23$$

The positive or negative z_A-score in the analysis results can be used to determine whether the spatial distribution of functional POI data within the influence zone of the station is random, clustered, or dispersed. In this analysis, a negative z_A-score represents clustering, while a positive z_A-score represents dispersion.

(2) Kernel density analysis

The Kernel Density analysis in ArcGIS 10.7 platform can visualize the clustering of points based on the values and distribution of input point features, providing an indication of the overall distribution pattern in the area. The various types of functional POI data from Amap are all geolocated point data. Therefore, Kernel Density analysis can be used to investigate the spatial clustering of functional facilities within the station realm, including the location of the clusters.

The Kernel Density analysis is a method that utilizes the values and distribution of input point features to display the clustering pattern of points and reflect the overall distribution of an area. It generates a continuous raster grid (heatmap) that represents the density of the points [55]. The Kernel Density analysis has been widely applied in urban spatial research [56,57,58]. In this study, the Kernel Density analysis is employed to explore the spatial distribution of pedestrian activities and functional facilities within each station realm. The calculation formula is as follows:

$$Density = \frac{1}{(radius)^2}\sum_{i=1}^{n}\left[\frac{3}{\pi}\cdot pop_i\left(1-\left(\frac{dist_i}{radius}\right)^2\right)^2\right] \quad 3\text{- }24$$

For $dist_i < radius$

where: $i = 1, \ldots, n$ are the input points. Only points in the sum are included if they lie within the radius distance of the (x, y) position. $dist_i$ is the distance between point i and the (x, y) position.

3.5 Analysis methods of interaction mechanism

3.5.1 Technology route

As shown in the Figure 3- 10, this paper explores the interaction mechanism follows three steps:

Step 1, Spearman correlation analysis will be used to analyze the correlation between Population activity and functional facility and further establish the system of variables.

Step 2, Ordinary least squares (OLS) model will be used to test the variables system for multicollinearity and significance.

Step 3, The final optimized variable system will be used in the Geographically weighted regression (GWR) model to explore the interactive relationship between population activities and different facilities.

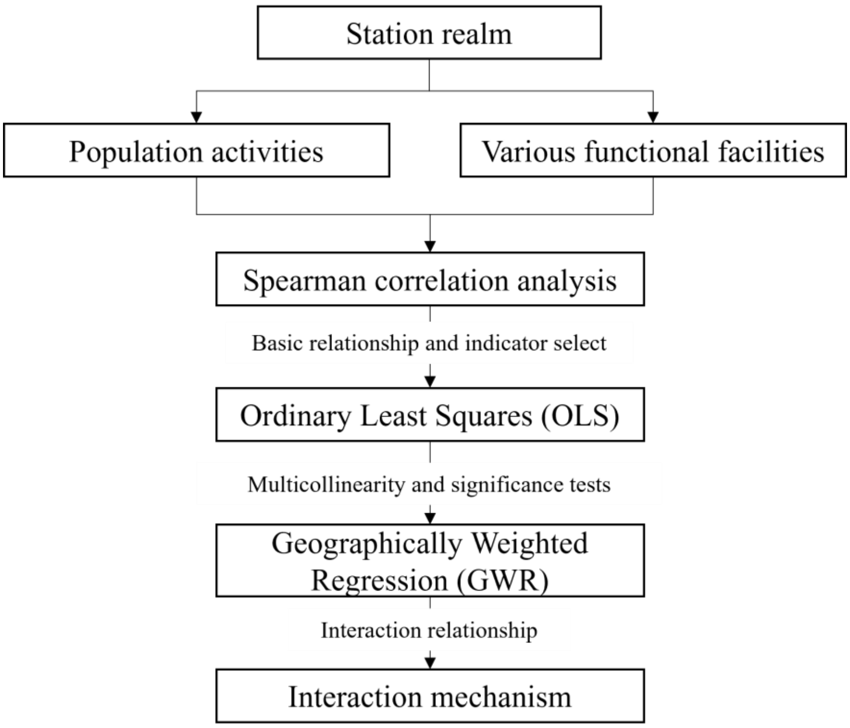


Figure 3- 10. Flowchart of spatial relationship model of the station realm

It is necessary to specify that before constructing a GWR model, it is common practice to first build an OLS model to identify variables that do not significantly explain population activity. The results of the OLS model can help identify and exclude non-significant variables, thereby improving the fit of the GWR model. This screening process aids in selecting the most explanatory combination of explanatory variables, leading to an improved fit of the GWR model.

3.5.2 Method of correlation analysis

By plotting scatter plots of the data, it is possible to observe a monotonic relationship between urban functional facilities and population activities. Since both sets of data, functional facilities, and population activities, do not pass the normality test, it is not appropriate to use Pearson correlation analysis [59,60]. However, as both variables are continuous and exhibit a monotonic relationship, the non-parametric test method of Spearman correlation analysis is used.

For a sample of size *n*, where the *n* original data points of *X* and *Y* are transformed into rank data *R(X)* and *R(Y)* respectively, the correlation coefficient r_s is calculated as follows:

$$r_s = \rho R(X), R(Y) = \frac{cov[R(X_I), R(Y_I)]}{\sigma R(X_I), R(Y_I)} \quad 3\text{- }25$$

where ρ is represents the coefficient, which is calculated using rank variables, $cov[R(X_I), R(Y_I)]$ denotes the covariance of the rank variables, $\sigma R(X_I), R(Y_I)$ refer to the standard deviations of the rank variables. A larger value of r_s indicates a stronger correlation between the variables.

3.5.3 Ordinary Least Squares (OLS) model

In this study, the role of the OLS model is mainly to diagnose multicollinearity in the system of variables. Multicollinearity is assessed by the variance inflation factor (VIF) value and if the VIF is less than 10, it indicates the non-existence of multicollinearity [61,62]. Secondly, based on the regression coefficients of the OLS model, variables with insignificant coefficients can also be further excluded.

OLS regression is a commonly used traditional (non-spatial) statistical method and serves as the starting point for spatial regression analysis. This model allows us to identify the factors that have a significant impact on the dependent variable from the set of independent variable indicators constructed for the dependent variable. When the sign associated with the coefficient is positive, the coefficient is positively related to the dependent variable, and conversely, it is negatively related. In addition, probabilities (p-values) can help us find which variables are statistically significant ($p < 0.05$) and thus deserve further analysis. The expression for OLS regression is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_n x_n + \varepsilon \quad 3\text{- }26$$

where y is the dependent variable, x_n is the explanatory variable, β_n is the regression coefficient, and ε is the random error. In this study, the number of active people in the study station realm is the dependent variable, and the number of POIs in each functional facility is the explanatory variables.

3.5.4 Geographically Weighted Regression (GWR) model

In 1996, Brunsdon et al. introduced the geographically weighted regression (GWR) technique [63], which provides an intuitive and practical approach for analyzing spatial heterogeneity and non-stationarity [64]. It has become an important method in local spatial statistical analysis. In 1970, Tobler proposed the first law of geography, stating that "Everything is related to everything else, but near things are more related than distant things [65]." This law highlights the crucial characteristic of spatial objects and their attribute relationships, with the degree of association decaying as spatial distance increases.

The GWR technique incorporates the first law of geography into local spatial statistical methods. It performs regression analysis separately for independently sampled analysis points, obtaining spatially varying regression coefficients that correspond to each specific spatial location. These estimated coefficients quantify the spatial heterogeneity characteristics by varying with different spatial positions. Relevant literature suggests that the GWR model has an advantage in its ability to address spatial heterogeneity (spatial heterogeneity refers to the inhomogeneity and complexity of the spatial distribution of urban elements) compared to the spatial lag model (SLM) and the spatial error model (SCM) [66]. Meanwhile, the GWR model has a good performance in highlighting the spatial nature of the data [67], especially in terms of compatibility with discrete data, positioning big data or real-time updated data [68]. Therefore, the regression results of the GWR model are used in this study as the main basis for revealing the interaction mechanism between population activities and functional facilities within the station realm.

The basic form of the GWR model can be expressed as follows:

$$y_i = \beta_o(u_i, v_i) + \sum_{\mathrm{k}=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \qquad (i = 1,2,3, \dots, n) \qquad 3\text{-}27$$

where y_i is the value of the dependent variable at location i, $x_{ik}(k = 1,2 \dots n)$ is coordinates of regression analysis point i, (u_i, v_i) is the intercept term that varies with the coordinates in $\beta_o(u_i, v_i)$ is the intercept term, $\beta_o(u_i, v_i)(k = 1,2 \dots n)$ is the regression coefficient.

3.6 Chapter summary

In this chapter, the research methodological framework is constructed around five main aspects: (1) Taking pedestrian scale as the core and using field investigations and mathematical statistics, the methods for defining the station realm scope are clarified. (2) Using transportation functional space and urban functions as primary indicators, four secondary indicators and fourteen investigation indicators are identified. The methods for data collection and analysis for each indicator are determined, and cluster analysis is conducted to classify the station realms into different types. (3) the sources and collection methods of population activity data and functional facility data based on location information are determined. (4) Based on the characteristics of the data, spatial autocorrelation analysis is determined as the method for identifying the distribution patterns of population activities in station realms. Spatial point pattern recognition techniques are used for recognizing the spatial distribution patterns of functional facilities in station realms. (5) The research route of "Spearman correlation analysis - OLS model - GWR model" is proposed to explore the interaction mechanism between population activities and functional facilities. In summary, the survey methods, measurement methods and analysis methods used in this study are specified around the construction idea of "station realm scope - station realm type - data collection - spatial distribution pattern - interaction mechanism".

Reference

[1] Frederick Ramm, Jochen Topf, Steve Chilton. OpenStreetMap: Using and Enhancing the Free Map of the World. UIT Cambridge. 2011.

[2] Zacharias J, Zhang T, Nakajima N. Tokyo Station City: The railway station as urban place[J]. Urban Design International, 2011, 16: 242-251.

[3] Zhuang X, Zhang L, Lu J. Past—Present—Future: Urban spatial succession and transition of rail transit station zones in Japan[J]. International Journal of Environmental Research and Public Health, 2022, 19(20): 13633.

[4] Lima Neto G B, Gerbal D, Marquez I. The specific entropy of elliptical galaxies: an explanation for profile-shape distance indicators? [J]. Monthly Notices of the Royal Astronomical Society, 1999, 309(2): 481-495.

[5] Camagni R, Gibelli M C, Rigamonti P. Urban mobility and urban form: the social and environmental costs of different patterns of urban expansion[J]. Ecological economics, 2002, 40(2): 199-216.

[6] Liu Jiyuan, Zhang Zengxiang, Zhuang Dafang, et al. Spatial and temporal characteristics of land use change in China in the 1990s and its causes[J]. Geography Research, 2003, 1.

[7] Li Xiaojiang, Jin Xiaochun, Shi Yanan, Zhai Jian et al. Spatial analysis and auxiliary decision-making methods in urban and rural planning preparation [M]. China Construction Industry Press, 2016.

[8] Song Chenghao. Research on the density of road network under standard road network conditions[D]. Master's thesis. Chang'an University, 2013.

[9] A. Friedman, V. Cammalleri. Reducing Energy, Resources and Construction Waste through Effective Residential Unit Design [J]. Building Research & Information.1994, 22(2): 103-108.

[10] Lyu X, Han Q, Vries B D. Procedural modeling of urban layout: population, land use, and road network [J]. Transportation Research Procedia, 2017, 25:3337-3346.

[11] Lynch K. Reconsidering the image of the city[M]. Springer US, 1984.

[12] TRIEB M. Urban design: theory and practice[M]. Stuttgart Press,1974.

[13] Wu Bing, Ge Zhaopan. Application of fractal theory in geographic information science research[J]. Geography and Land Studies, 2002, 18(3): 23-26.

[14] Yuan Hang. Study on the optimization strategy of skyline adjustment on both banks of Han

River in Hanzhong City[D]. Master's thesis. Xi'an University of Architecture and Technology,2017.

[15] Frank L D, Pivo G. Impacts of mixed use and density on the utilization of three modes of travel: single occupant vehicle, transit, and walking [J]. Washington d c: Transportation Research Record, 1995:13-42.

[16] Qian Linbo. Spatial distribution of urban land use mix and residents' travel[J]. Urban Research,2000(03):7-10.

[17] Zhang M, Zhao P. The impact of land-use mix on residents' travel energy consumption: New evidence from Beijing [J]. Transportation Research Part D Transport & Environment, 2017, 57:224-236.

[18] Knaap Gerrit-Jan, Song Yan, Nedovic-Budic Zorica. Measuring Patterns of Urban Development: New Intelligence for the Waron Sprawl [J]. Local Environment, 2007,12(3): 239-257.

[19] Zhou Suhong, Yang, Lijun. Urban traffic under the influence of urban development intensity[J]. Journal of Urban Planning,2005, (2):75-80.

[20] Peng, Lu, Huapu. Analysis of the impact of urban density on traffic demand based on spatial analysis[J]. Transportation System Engineering and Information,2007(04):90-95.

[21] Zheng Meng, Zhang Xiaodong. Determining the appropriate development intensity of land based on traffic carrying capacity: An example of detailed control planning in central Beijing[J]. Urban Transportation,2008,6(5):15-18

[22] Ministry of Housing and Urban-Rural Development of the People's Republic of China. Urban land classification and planning construction land standards [M]. Beijing: China Construction Industry Press, 2011:6-9.

[23] Xing Hekefan. The Rational land development of Strength surrounding the tracksites [D]. Master's Theses. Nanjing Forestry University, 2015.

[24] He Xin, Li Ke. Classification of urban rail transit stations based on cluster analysis method [J]. Information & Communications, 2015 (07): 36-37.

[25] Xia Xue, Gai Jingyuan. Classification of urban rail transit stations and points and analysis of passenger flow characteristics based on K-Means clustering algorithm [J]. Modern Urban Transit, 2021 (04): 112-118.

[26] MacQueen J. Classification and analysis of multivariate observations[C]//5th Berkeley Symp. Math. Statist. Probability. Los Angeles LA USA: University of California, 1967: 281-297.

[27] WHITE J, WELLS I. Extracting origin destination information from mobile phone data. Eleventh International Conference on Road Transport Information and Control, London, UK, 19-21 March 2002. IEE conference publication. 2002, 30-34. DOI: 10.1049/cp:20020200

[28] Zufen L, Lei Y, Yong G, et al. Extraction method of temporal and spatial characteristics of residents' trips based on cellular signaling data. Transport Research 2016, 2(1), 51-57.

[29] Xu Ruihua, Xu YS. A real-time prediction method for passenger flow distribution of urban rail lines. Journal of Tongji University: Natural Sciences Edition 2011, 39(6), 857-861.

[30] Zhang, X.; Sun, Y.; Chan, T.O.; Huang, Y.; Zheng, A.; Liu, Z. Exploring Impact of Surrounding Service Facilities on Urban Vibrancy Using Tencent Location-Aware Data: A Case of Guangzhou. Sustainability 2021, 13, 444.

[31] Li Luning, Liu Menghang, Li Qiang, et al. A new support of human activity research: Characteristics, research status, and application prospects of Wi-Fi data. Progress in Geography 2021, 40(11), 1970-1982.

[32] Duan Y M, Liu Y, Liu X H, et al. Measuring polycentric urban structure using Easygo big data: A case study of Chongqing metropolitan area. Progress in Geography 2019, 38(12), 1957-1967.

[33] Shen Lifan, Wang Ye, Zhang Chun, Jiang Dongrui, Li He. Study on the relationship between the built environment and rail commuting in reasonable walkable range of rail stations--Example of 44 rail stations in Beijing. Journal of Geography 2018,73(12), 2423-2439.

[34] Tencent's financial performance for the second quarter and mid-year of 2021. Available online: https:// static.www.tencent.com/ uploads/ (accessed on 25 September 2021).

[35] Analysis Report on the Development of China's Mobile Internet Industry in Q3 2020. Available online: http://www.100ec.cn/ detail--6575116.html (accessed on 25 September 2021).

[36] Liu, Z.; Qian, J.; Du, Y.; Wang, N.; Yi, J.; Sun, Y.; Ma, T.; Pei, T.; Zhou, C. Multi-level Spatial Distribution Estimation Model of the Inter-regional Migrant Population Using Multi-source Spatio-temporal Big Data: A Case Study of Migrants from Wuhan during the Spread of COVID-19. J. Geo-Inf. Sci. 2020, 22, 147–160.

[37] Long Y, Shen Z. Discovering functional zones using bus smart card data and points of interest in Beijing[M]//Geospatial analysis to support urban planning in Beijing. Springer, Cham, 2015: 193-217. DOI: 10.1007/978-3-319-19342-7_10

[38] Shu B, Yang C, Cui J, Chen L. Preliminary Study on the City's Functional Structure of Subway Station's Surrounding Area Under the TOD Model: Empirical Analysis Based on POI Data Along

Chengdu Metro [J]. Huazhong Architecture 2019, 5: 79-83.

[39] National Economic Industry Classification (GB/T 4754-2017) [S]. The General Administration of Quality Supervision, Inspection and Quarantine of China, 2017.

[40] Abdulhafedh A. A novel hybrid method for measuring the spatial autocorrelation of vehicular crashes: Combining Moran's Index and Getis-Ord G i* statistic[J]. Open Journal of Civil Engineering, 2017, 7(02): 208.

[41] Eboli L, Forciniti C, Mazzulla G. Spatial variation of the perceived transit service quality at rail stations[J]. Transportation Research Part A: Policy and Practice, 2018, 114: 67-83.

[42] Esri A P. How High/Low Clustering (Getis-Ord General G) works[J].

[43] Anselin L. Local indicators of spatial association—LISA[J]. Geographical analysis, 1995, 27(2): 93-115.

[44] Songchitruksa P, Zeng X. Getis–Ord spatial statistics to identify hot spots by using incident management data[J]. Transportation research record, 2010, 2165(1): 42-51.

[45] Zahran E, Shams S, Said S, et al. Validation of forest fire hotspot analysis in GIS using forest fire contributory factors[J]. Syst. Rev. Pharm, 2020, 11: 249-255.

[46] Liu W, Ma L, Smanov Z, et al. Clarifying Soil Texture and Salinity Using Local Spatial Statistics (Getis-Ord Gi* and Moran's I) in Kazakh–Uzbekistan Border Area, Central Asia[J]. Agronomy, 2022, 12(2): 332.

[47] Songchitruksa P, Zeng X. Getis–Ord spatial statistics to identify hot spots by using incident management data[J]. Transportation research record, 2010, 2165(1): 42-51.

[48] ESRI E. How hot spot analysis (Getis-Ord Gi*) works[J]. ArcGIS Pro Tool Reference, 2012.

[49] Adolphson M. Kernel densities and mixed functionality in a multicentered urban region[J]. Environment and Planning B: Planning and Design, 2010, 37(3): 550-566.

[50] Shafabakhsh G A, Famili A, Bahadori M S. GIS-based spatial analysis of urban traffic accidents: Case study in Mashhad, Iran[J]. Journal of traffic and transportation engineering (English edition), 2017, 4(3): 290-299.

[51] Melyantono S E, Susetya H, Widayani P, et al. The rabies distribution pattern on dogs using average nearest neighbor analysis approach in the Karangasem District, Bali, Indonesia, in 2019[J]. Veterinary world, 2021, 14(3): 614.

[52] Thompson A E, Walden J P, Chase A S Z, et al. Ancient Lowland Maya neighborhoods:

Average Nearest Neighbor analysis and kernel density models, environments, and urban scale[J]. PloS one, 2022, 17(11): e0275916.

[53] Ebdon, David. Statistics in Geography. Blackwell, 1985.

[54] Mitchell, Andy. The ESRI Guide to GIS Analysis, Volume 2. ESRI Press, 2005.

[55] Yu W, Ai T. Visualization and analysis of POI points in cyberspace supported by kernel density estimation method[J]. Journal of surveying and mapping, 2015, 44(01): 82-90.

[56] Leslie T F. Identification and differentiation of urban centers in Phoenix through a multi-criteria kernel-density approach[J]. International Regional Science Review, 2010, 33(2): 205-235.

[57] Kang Y, Cho N, Son S. Spatiotemporal characteristics of elderly population’s traffic accidents in Seoul using space-time cube and space-time kernel density estimation[J]. PLoS one, 2018, 13(5): e0196845.

[58] Iseki H, Eom H. Impacts of rail transit accessibility on firm spatial distribution: Case study in the metropolitan area of Washington, DC[J]. Transportation Research Record, 2019, 2673(11): 220-232. DOI: 10.1177/0361198119844464.

[59] Hauke J, Kossowski T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data[J]. Quaestiones geographical, 2011, 30(2): 87-93.

[60] Eboli L, Forciniti C, Mazzulla G. Spatial variation of the perceived transit service quality at rail stations[J]. Transportation Research Part A: Policy and Practice, 2018, 114: 67-83.

[61] Menard, S. Applied Logistic Regression Analysis, 2nd ed; Sage: Newbury Park, CA, USA, 2002.

[62] Bryant F B, Yarnold P R. Principal-components analysis and exploratory and confirmatory factor analysis[J]. 1995.

[63] Brunsdon C, Fotheringham A S, Charlton M E. Geographically weighted regression: a method for exploring spatial nonstationary [J]. Geographical analysis, 1996, 28(4): 281-298.

[64] Brunsdon C, Fotheringham S, Charlton M. Geographically weighted regression[J]. Journal of the Royal Statistical Society: Series D (The Statistician), 1998, 47(3): 431-443.

[65] Tobler W R. A Computer Movie Simulating Urban Growth in the Detroit Region[J]. Economic Geography,1970,46(2): 234-240.

[66] Gao C, Feng Y, Tong X, et al. Modeling urban growth using spatially heterogeneous cellular automata models: Comparison of spatial lag, spatial error and GWR[J]. Computers, Environment

and Urban Systems, 2020, 81: 101459.

[67] Urban, R. C., & Nakada, L. Y. K. (2021). GIS-based spatial modelling of COVID-19 death incidence in São Paulo, Brazil. Environment and Urbanization, 33 (1), 229–238.

[68] Wang Yuanfei, He Honglin. Spatial data analysis methods [M]. Science Press, 2007.

Chapter 4. Spatial distribution pattern analysis of Xi'an metro line 2 station realm

Content

Introduction

This chapter presents the empirical study based on the theoretical framework established in Chapter 3. Taking the 16 stations of Metro Line 2 as an example, the ranges of each station realm are defined, and the types of station realms are determined. The identification of station realm spatial distribution patterns mainly focuses on two dimensions: "population activities" and "functional facilities". Through data collection and processing, a large-scale investigation data system supported by population activities and functional facilities is established, which includes data from the EasyGO and functional facilities data from Amap. Spatial autocorrelation analysis and spatial point pattern recognition techniques are used to identify the spatial distribution patterns of population activities and functional facilities in each station realm of Metro Line 2, to understand the spatiotemporal distribution characteristics of population activities and the distribution characteristics of functional facilities.

4.1 Study area and object

4.1.1 Overview of the study area

The Xi'an Metro Line 2 was opened for operation on September 16, 2011. It spans a total length of 26.7 kilometers and currently has 21 stations (Figure 4- 1). Metro Line 2 is the first urban rail transit line opened in Xi'an and northwest China [1].

Figure 4- 1. Xi'an Metro Line 2 and the location of each station

As shown in Figure 4- 2, Line 2 runs north to south, starting from Xi'an North Station in Weiyang District in the north, passing through Lianhu District, Xincheng District, Beilin District, Yanta District, and ending at Weigu South Station in Chang'an District in the south. It is also the core route that overlaps with the city's central axis and north-south urban development axis, linking the administrative center, transportation hubs, university gathering areas, scientific research areas, commercial centers, and several tourist attractions in Xi'an in a tight chain, thus strongly supporting Xi'an's urban spatial development strategy [2].

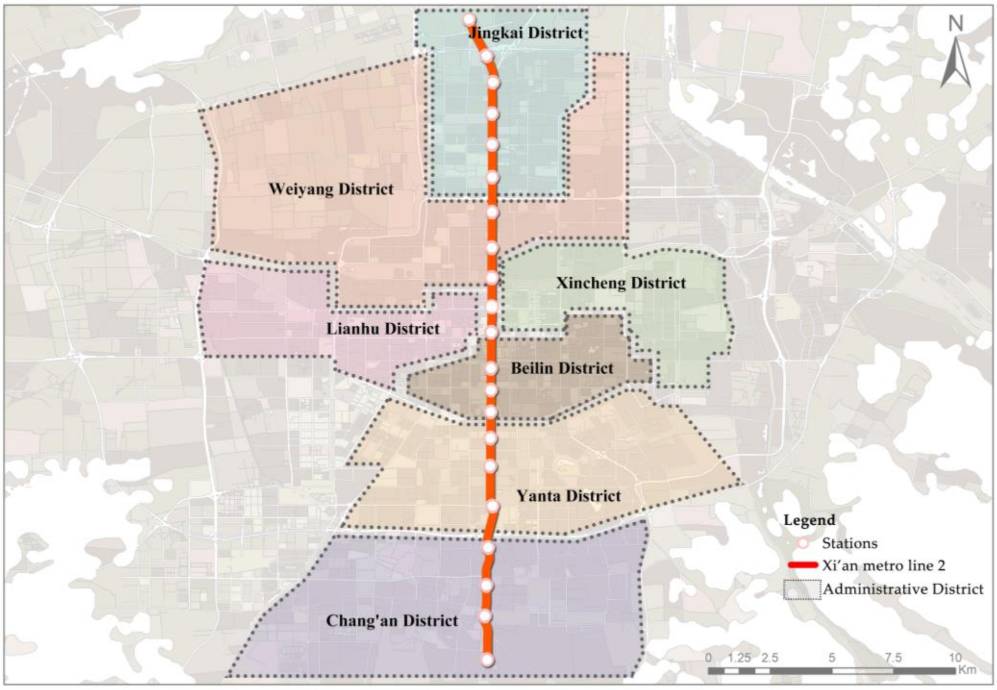


Figure 4- 2. Distribution of Metro Line 2 stations in Xi'an city administrative districts

The highest single-day passenger flow can reach 1,257,600 passengers [3]. Therefore, Metro Line 2 is the main support for the north-south spatial expansion of the city and the main north-south passenger flow corridor in the central city of Xi'an. Usually, residents in the rail transit station realm tend to prefer rail transit travel. The incoming and outgoing passenger flow of a station often reflects the spatial aggregation of the population in its vicinity, i.e., the station with higher incoming and outgoing passenger flow has a higher density of human flow in its station realm and the more frequent population activities in space [4,5]. As shown in the figure, the incoming and outgoing passenger flows for one natural week (November 5-November 12, 2022) at six stations of Metro Line 2 are counted, and the results show that there is a significant variability in the incoming and outgoing passenger flows at each station of Metro Line 2, and this phenomenon is reflected at different times at the same station.

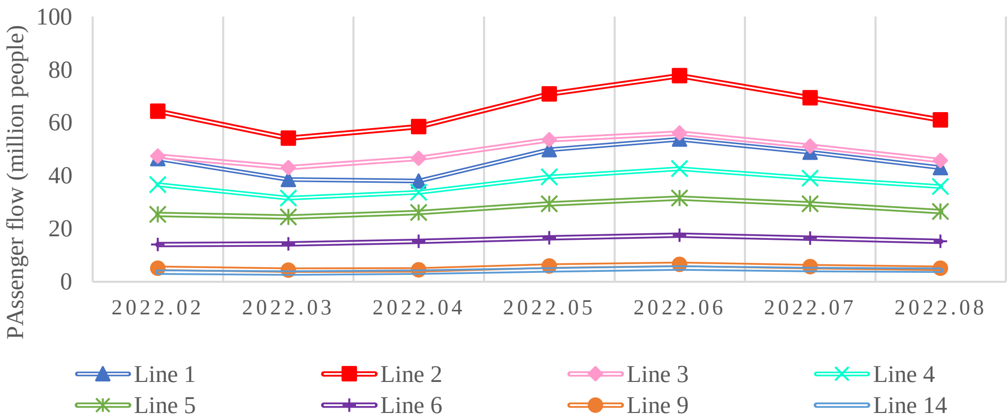


Figure 4- 3. Monthly average daily passenger flow of Xi'an Metro in 2022 [6]

To sum up, Metro Line 2 is a very important research case, which has a profound impact on the urban spatial pattern along the line and around the stations, and the characteristics of the spatial distribution of population activities and urban functions in the station realm and the inner interaction mechanism between them revealed by the in-depth study have strong implications for the future planning and development of the rail transit station realm in Xi'an and the whole Northwest China.

4.1.2 Study object selection

The selection criteria for the study objects in this study consider the following three conditions:

(1) Urban functions of the stations:

Comparing the results of related studies on rail transit station realms in major cities such as Beijing, Shanghai, and Shenzhen, it is evident that the selected objects should have universality and generalizability. If a station serves multiple urban functions, its spatial characteristics in the station realm become more specific and complex. As a result, the research findings may lack universality and generalizability.

(2) Urban location of the stations:

Based on the relevant theoretical foundations mentioned earlier, the rail transit station realms studied in this paper belong to the inner space of the city. Stations located in the outskirts of the city are often situated in the urban-rural transition zones, where the level of urbanization is generally lower, and there may be inadequate provision of facilities and sparse population. Therefore, the selection of research objects must consider the urban location of the stations.

(3) Development status of the surrounding areas:

The research objects should be chosen in areas surrounding stations where the built environment is mature. Considering data availability, if the surrounding area is under construction or in the initial stages of development, it not only increases the difficulty of obtaining data but also reduces the accuracy of the obtained data. This, in turn, affects the accuracy of the research in identifying the spatial distribution characteristics and patterns in the station realms.

4.1.3 Selected station

Through the preliminary investigation, the basic information statistics of each station of Xi'an Metro Line 2 are shown in Table 4- 1. Five stations, Xi'an Bei (Bei Ke) Station, San Yao Station, Feng Xi Yuan Station, Hang Tian Cheng Station, and Wei Qu Nan Station, will be excluded for the following reasons:

(1) Xi’an Bei Station

The Xi’an Bei Station serves as a large comprehensive transportation hub, integrating railway passenger transport, long-distance and medium-distance bus services, urban rail transit, buses, taxis, and private vehicles [7]. Due to its complex functionality in connecting Xi'an with external transportation, it is excluded as a research object.

(2) San Yao Station, Feng Xi Yuan Station, Hang Tian Cheng Station, and Wei Qu Nan Station

These stations are located on the outskirts of the city, beyond the third ring road (Figure 4- 4). Additionally, the surrounding areas of these stations are still under development, which poses challenges in obtaining survey data. Therefore, these stations are also excluded from the research.

Table 4- 1. Basic information about Xi'an Metro Line 2

Abbreviations	Name	Interchange lines	Administrative District	Urban Traffic Location
BKZ	BEI KE ZHAN	Metro line 4 Metro line 14		Outside the Third Ring Road
BY	BEI YUAN	——		
YDGY	YUN DONG GONG YUAN	——		
XXZX	XING ZHENG ZHONG XIN	Metro line 4	Weiyang District	Third Ring Road
FCWL	FENG CHENG WU LU	——		
STSG	SHI TU SHU GUAN	Metro line 8 (In progress)		
DMGX	DA MING GONG XI	——		
LSY	LONG SHOU YUAN	——	Lianhu District	Second Ring Road
AYM	AN YUAN MEN	——		
BDJ	BEI DA JIE	Metro line 1	Lianhu District Xincheng District	
ZL	ZHONG LOU	Metro line 6 (Extension Line)		First ring road
YNM	YONG NING MENG	——		
NSM	NAN SHAO MEN	Metro line 5	Beilin District	Second Ring Road
TYC	TI YU CHANG	——		
XZ	XIAO ZHAI	Metro line 3		
WYJ	WEI YI JIE	——		Third Ring Road
HZZX	HUI ZHAN ZHONG XIN	Metro line 8 (In progress)	Yanta District	
SY	SAN YUAN	——		
FXY	FENG XI YUAN	——		
HTC	HANG TIAN CHENG	Metro line 15 (In progress)	Changan District	Outside the Third Ring Road
WQN	WEI QU NAN	——		

Therefore, considering the above factors, the study focuses on the 16 stations within the inner area bounded by the third ring road, which is also the main urban area of Xi’an. The 16 stations include Bei Yuan station to Hui Zhan Zhong Xin station.

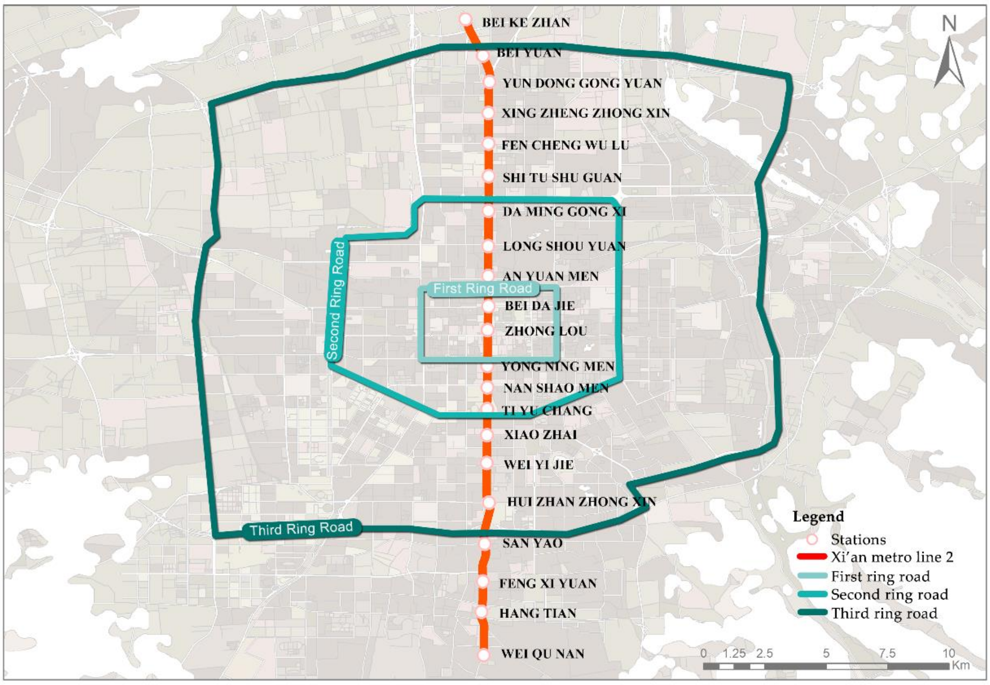


Figure 4- 4. The location of stations on Xi'an Metro Line 2

4.2 Defining scope of station realm

4.2.1 Survey of pedestrian walking scale

The pedestrian scale of pedestrians around study stations was measured according to the method proposed in Section 3.1.

(1) Firstly, questionnaires were distributed to pedestrians around each station of Xi'an Metro Line 2 (50 copies for each station, 800 copies in total, 774 valid questionnaires, 96.8% effective rate) to investigate the acceptable maximum walking time of pedestrians. As shown in the Figure 4- 5, the highest percentage of walking time was 10 min (36.82%), followed by 8 min (26.49%) and 15 min (16.54%). Therefore, based on the relevant literature research and the actual demand of this paper, 10 min was selected as the walking time around the 16 station of Xi’an Metro Line 2.

Figure 4- 5. Pedestrian walking time

(2) Secondly, the observation distance was set to 30 meters, and 30 pedestrians were randomly selected around each station (480 in total at 16 stations), and the time of pedestrians passing the observation distance was measured by a stopwatch timer. The walking speed was calculated for all measurements according to Equation 3- 1, and the results are shown in Figure 4- 6. The arithmetic mean of the walking speed of 480 pedestrians was calculated to be 1.18m/s, which was set as the walking speed of pedestrians around the 16 stations of Xi'an Metro Line 2.

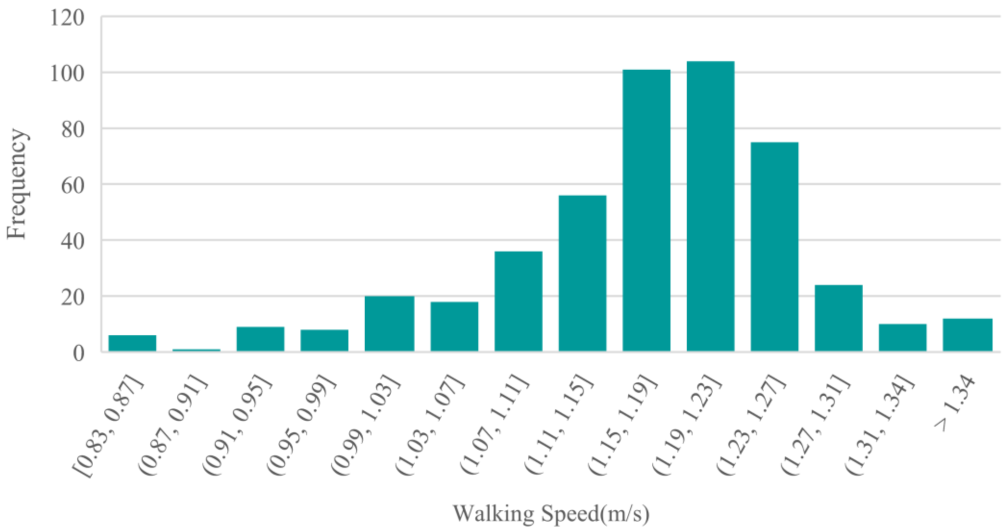


Figure 4- 6. Pedestrian walking speed

(3) According to Equation 3- 2, the walking scale of pedestrians around the station of Metro Line 2 is calculated to be about 708m. Then, the walking scale is used as a radius to create circular shape for each research station on ArcGIS, as shown in Figure 4- 7, for example.

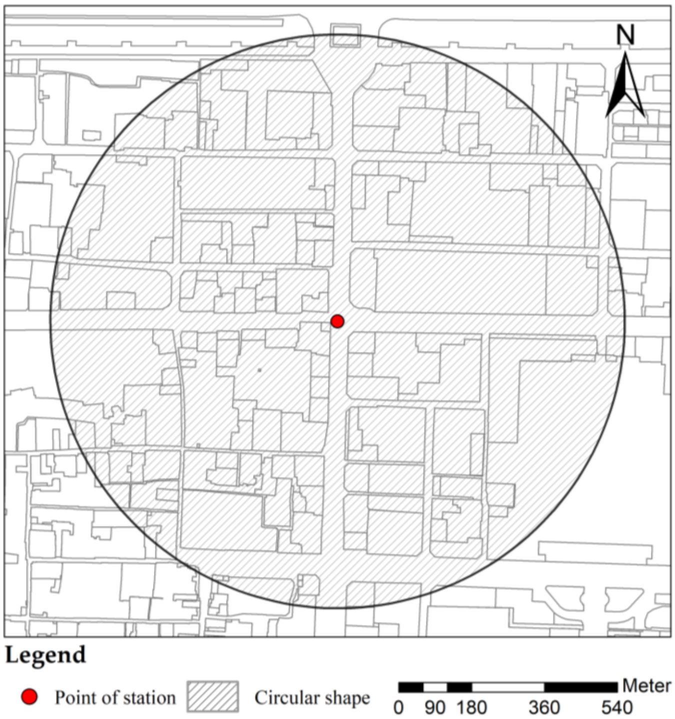


Figure 4- 7. Create circular shape for study station on ArcGIS

4.2.2 Definition result

Based on Section 3.1, this study will also use the urban road network and land boundary line to adjust the circular shape of each study station to make the boundaries of the station realm more realistic. The data of urban road network and land boundary line were obtained from OSM (https://www.openstreetmap.org, access time: August 10, 2022).

Finally, the shape of the adjusted boundary for each study station is shown in Figure 4- 8, thus completing the definition of the station realm in this study.

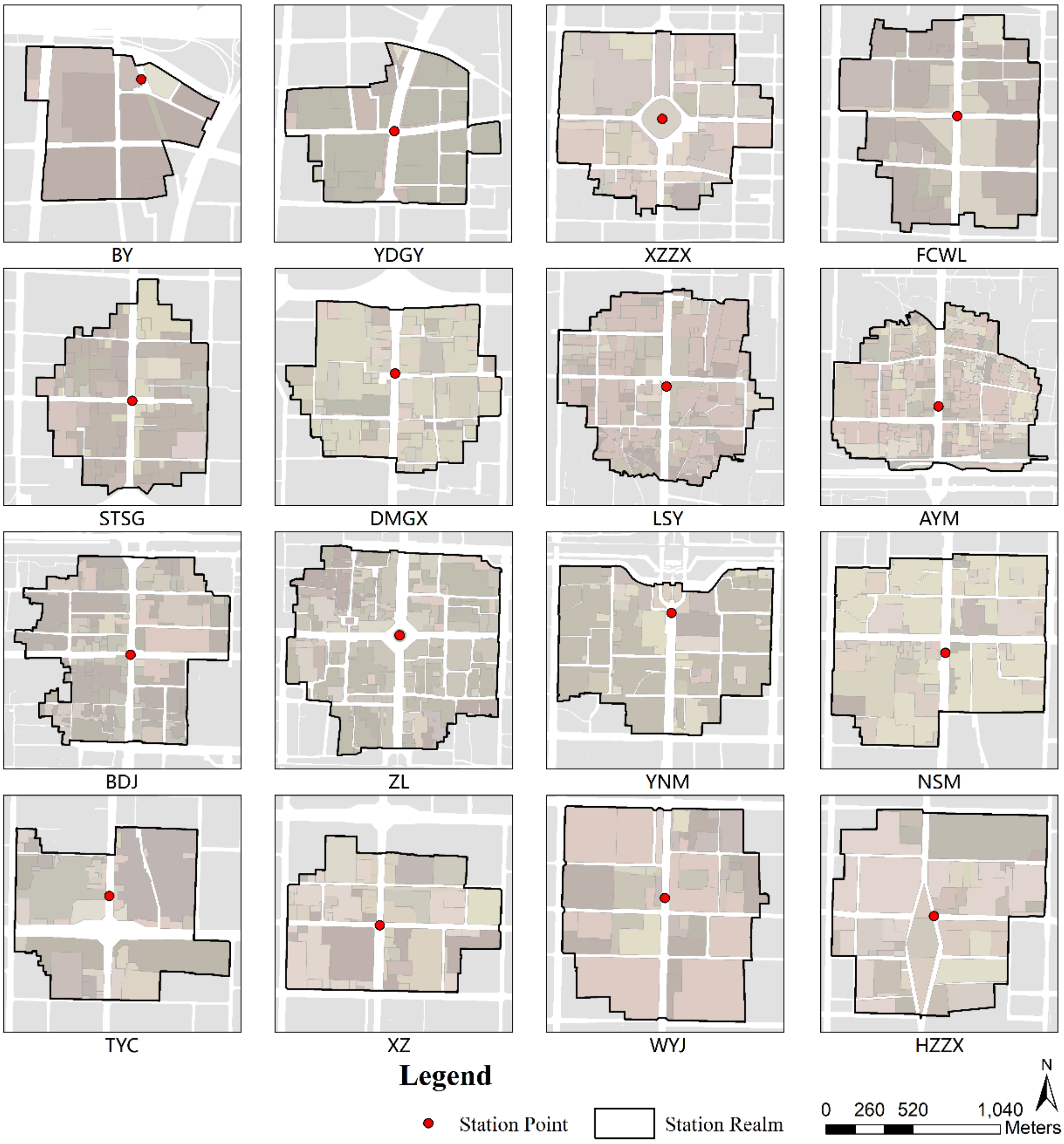


Figure 4- 8. Scope of the 16 research station realms

4.3 Classification of station types

4.3.1 Analysis of station realm traffic function environment

Through the field survey of each index of traffic function environment, the data statistics of transit lines, the number of metros entrances and exits, the number of bus stops, and the number of traffic facilities within each station realm of Line 2 were conducted, at the same time, a database was created for each station realm on ArcGIS. The visualization results are shown in the Figure 4- 9.

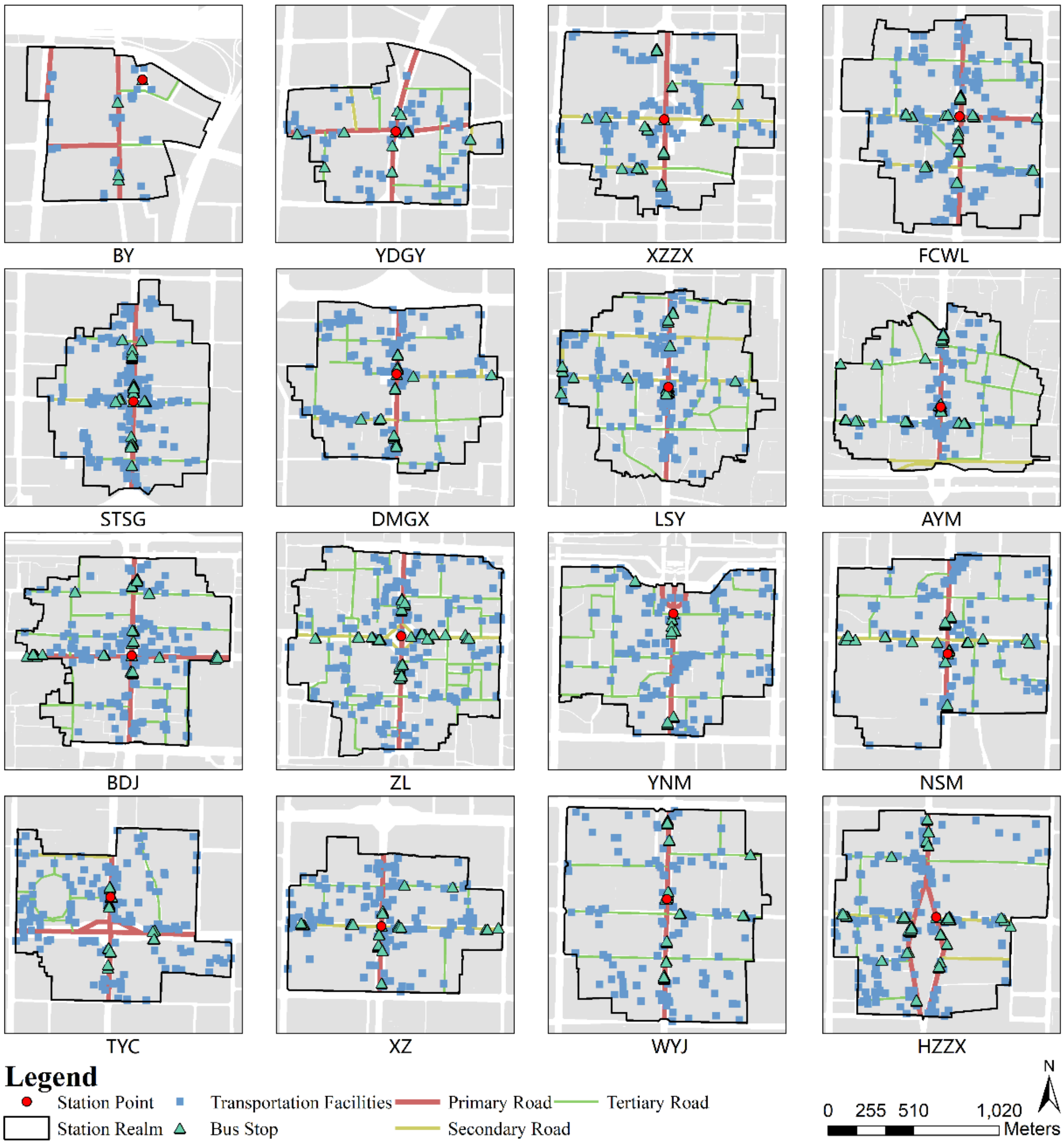


Figure 4- 9. Visualization of the traffic function environment within each station realm

4.3.2 Analysis of spatial form within station realm

(1) Boundary form characteristics of station realm

Based on the previous definition of the influence range of key research stations on Xi'an Metro Line 2, the area and perimeter data of each key research station realm were statistically analyzed. Table 4- 2 presents the results. The BDJ station realm has a compactness value of 0.62, indicating a high degree of irregularity in the spatial form of the station realm boundary. Currently, the increase in the urban built-up area within the station realm is mainly through external expansion.

The WYJ station realm has the highest compactness value of 0.82, indicating the least irregularity in the spatial form of the station realm boundary. The increase in the urban built-up area within this station realm is primarily through infill development between the edges of the built-up area. By considering the location characteristics of each station realm, it can be observed that the compactness of the station realm boundaries along Xi'an Metro Line 2 generally increases from the central city area to the inner ring road and then to the outer ring road. This indicates a transition from internal expansion to internal infill and redevelopment stages for the station realms from the central city area to the inner ring road and then to the outer ring road.

Table 4- 2. Boundary compactness statistics for 16 station realms

Code	Area of station realm (m^2)	Perimeter of station realm (m)	Compactness ratio
BY	753687.48	4041.60	0.76
YDGY	1262594.20	5639.68	0.71
XXZX	1979003.61	6750.02	0.74
FCWL	1711534.97	6484.81	0.72
STSG	1481082.21	6087.63	0.71
DMGX	1325517.71	5384.35	0.76
LSY	1542765.02	6370.60	0.69
AYM	1370970.00	6177.55	0.67
BDJ	1236478.83	6378.27	0.62
ZL	1432799.43	5641.73	0.75
YNM	1112872.54	5100.71	0.73
NSM	1355405.55	5081.50	0.81
TYC	1460523.75	6131.07	0.70
XZ	1394105.80	5437.03	0.77
WYJ	1656312.09	5586.42	0.82
HZZX	1952449.43	6194.52	0.80

(2) Road network form characteristics

Through survey methods such as field survey and photo observation, this study collected data for each grade of road network in 16 station realms, mainly counted the length of each grade, and visualized them using ArcGIS10.7 platform. As shown in Figure 4- 10, the network forms of each station realm can be summarized into four basic structural forms: conventional square network, end-of-range network, strip network, and free network.

The road network density of each key study station realm is shown in Table 4- 3, where the road network density of the XZZX station realm is 3.34, DMGX station realm is 5.47, LSY station realm is 7.93, BDJ station realm is 8.16, TYC station realm is 4.70, and WYJ station realm is 4.49.It can be seen that the overall road network density in the station realm of Metro Line 2 shows the characteristic of gradually decreasing from the high city center to the outside.

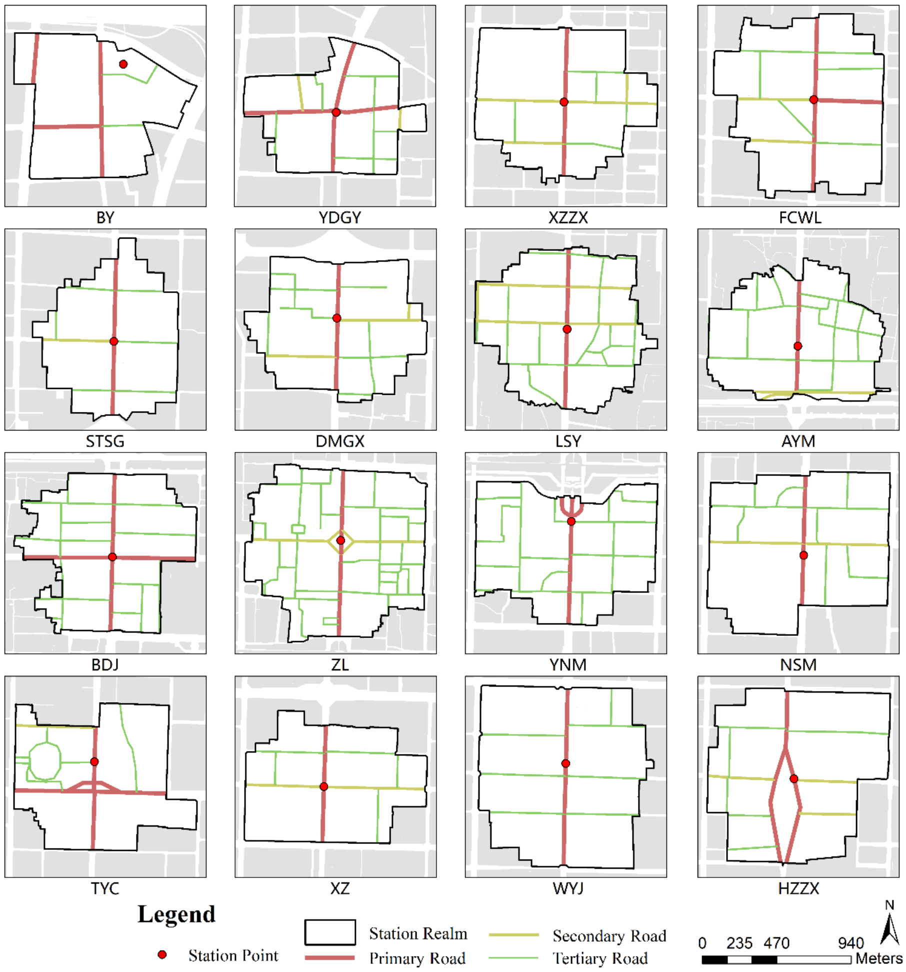


Figure 4- 10. Grid pattern in each station realm

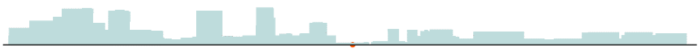
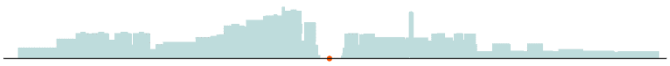
Table 4- 3. Statistics of road network density in each station realm

Code	Road network length (m)	Road network density
BY	2322.82	3.08
YDGY	6117.99	4.85
XXZX	6044.35	3.05
FCWL	6349.70	3.71
STSG	4762.16	3.22
DMGX	6264.27	4.73
LSY	10388.90	6.73
AYM	9321.87	6.80
BDJ	8997.23	7.28
ZL	12227.23	8.53
YNM	7695.06	6.91
NSM	5819.61	4.29
TYC	6493.74	4.45
XZ	4830.04	3.46
WYJ	5469.94	3.30
HZZX	7226.68	4.15

(3) Architectural form characteristics

In this paper, the current information of the building facade area (north-south direction), number of floors and height of the buildings in each research station realm was collected through field survey and photo recording, and the 3D analysis technology of ArcGIS10.7 platform was used to visualize the building facade profile of the station realm (Appendix C). The fractal dimension was calculated by MATLAB software, and the specific results are shown in Table 4- 4. The overall building height of each key research station realm basically shows the characteristics of decreasing contours from high to low centered on the station.

Table 4- 4. Statistics of fractal dimension of building contours in each station realm

Code	Skyline	Fractal dimension
BY		1.7835
YDGY		1.6893
XXZX		1.6139
FCWL		1.6896
STSG		1.7247
DMGX		1.7216
LSY		1.7185
AYM		1.7163
BDJ		1.6424
ZL		1.6738
YNM		1.7008
NSM		1.6543

TYC		1.6446
XZ		1.6505
WYJ		1.7290
HZZX		1.7338

The results of the dominance degree calculation for the architectural form within each station realm are shown in Table 4- 5. Among them, the Sports Stadium station realm has the highest dominance degree (0.87), indicating that the core high-rise buildings within this station realm are relatively prominent in terms of their visual impact and landmark significance. They exhibit good recognizability and commanding presence. On the other hand, the ZL station realm has the lowest dominance degree (0.71) due to height restrictions imposed by regulations on historical block architecture (not exceeding 36 meters in height). This limitation results in a moderate level of recognizability and commanding presence for the high-rise buildings within this station realm.

Table 4- 5. Building height and building primacy in each station realm

Code	Highest building floor	Average building floor	Average building height (m)	Building primacy
BY	28	5.75	17.24	0.79
YDGY	35	6.75	20.26	0.81
XXZX	31	4.74	14.21	0.85
FCWL	31	6.94	20.81	0.78
STSG	31	7.34	22.02	0.76
DMGX	33	5.88	17.63	0.82
LSY	33	5.85	17.54	0.82
AYM	32	4.95	14.86	0.85
BDJ	20	4.96	14.88	0.75
ZL	16	4.57	13.70	0.71
YNM	32	6.35	19.05	0.80
NSM	32	5.97	17.90	0.81
TYC	51	6.86	20.58	0.87
XZ	37	5.20	15.59	0.86
WYJ	30	5.64	16.92	0.81
HZZX	32	5.83	17.48	0.82

4.3.3 Analysis of spatial density within station realm

This study conducted data statistics on the current land use composition within each station realm of Line 2 in Xi'an City, based on the CAD maps of current land use in Xi'an City and the shapefile data of current buildings obtained from OpenStreetMap. A database was established in the ArcGIS 10.7 platform to facilitate the calculation of population density, land mixture degree, and land use intensity for each station realm.

(1) Population density

In this study, the EasyGO applet was used to calculate the total daily population activity within each station realm, which was then used to calculate population density. The statistical results are shown in Figure 4- 11. Among them, ZL station realm had the highest population density, followed by XZ station realm, while AYM station realm had the lowest population density.

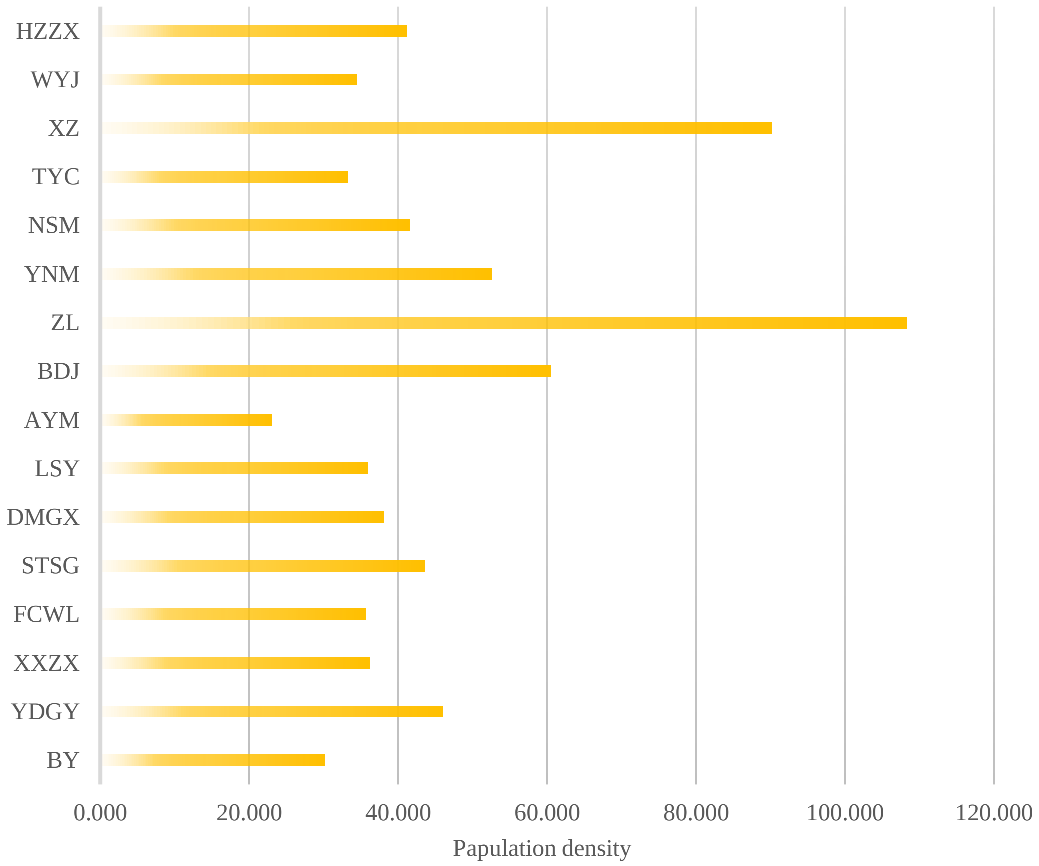


Figure 4- 11. Population density statistics within each station realm

(2) Land use mix

Based on Figure 4- 12, this study conducted a statistical analysis of the current land use composition within each station realm of Line 2. The land use max was calculated using Formula 3-8, and the specific results are shown in Figure 4- 13. Among them, XZZX station realm had the highest land use max, while YDGY station realm had the lowest land use max.

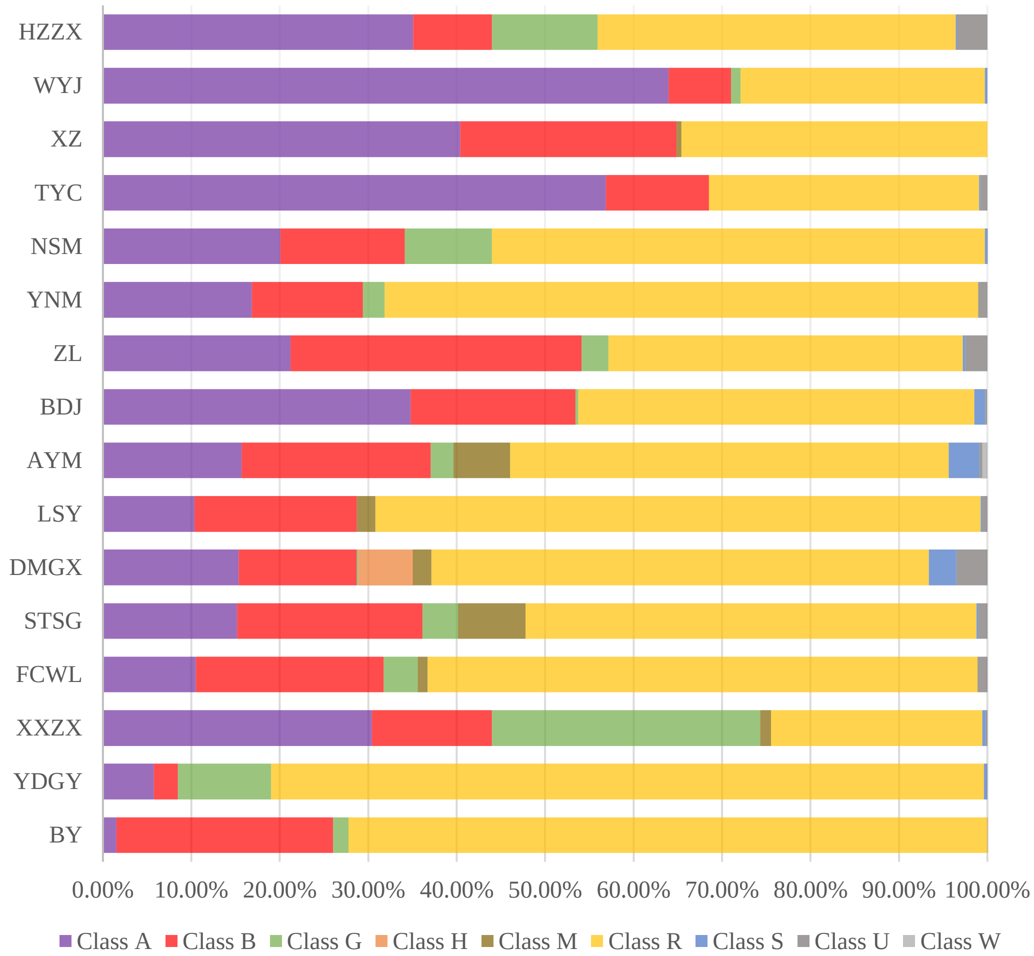


Figure 4- 12. Statistics on the percentage of land use types within each station realm

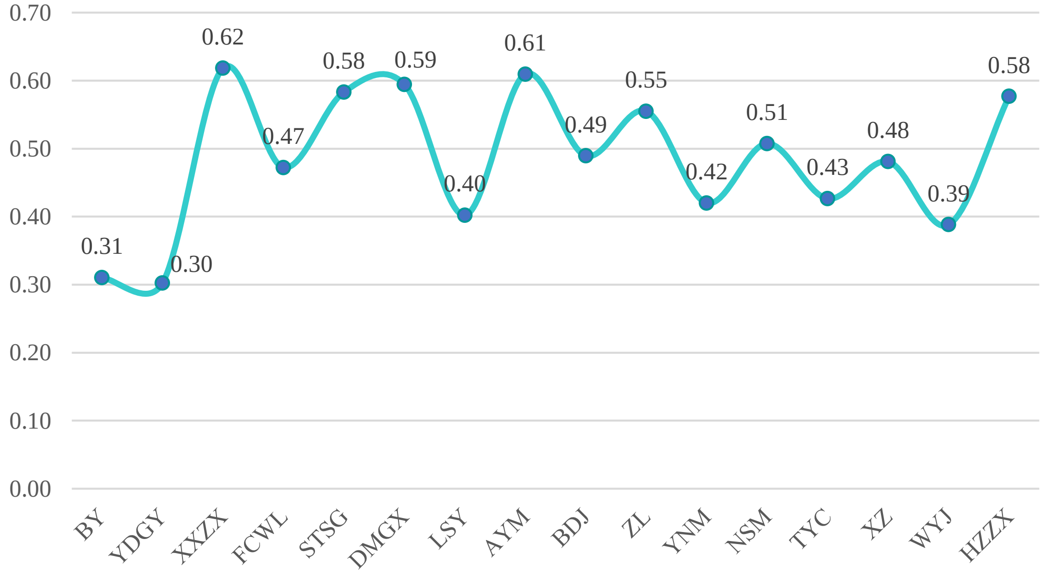


Figure 4- 13. Statistics of land use mix within each station realm

(3) Building density

In this study, the building density within each station realm was calculated using the raster calculation tools in ArcGIS 10.7. The results are shown in Figure 4- 14, where darker colors indicate higher building density in the corresponding land parcels.

Figure 4- 14. Building density statistics within each station realm

From Figure 4- 15, it can be observed that the average building density is highest in the ZL station realm, with a value of 133.69%. This indicates that the land vacancy rate is low, and the level of building intensity is very high in this station realm. On the other hand, the BY station realm has the lowest average building density at 15.28%, suggesting a relatively higher land vacancy rate and lower building intensity in this area.

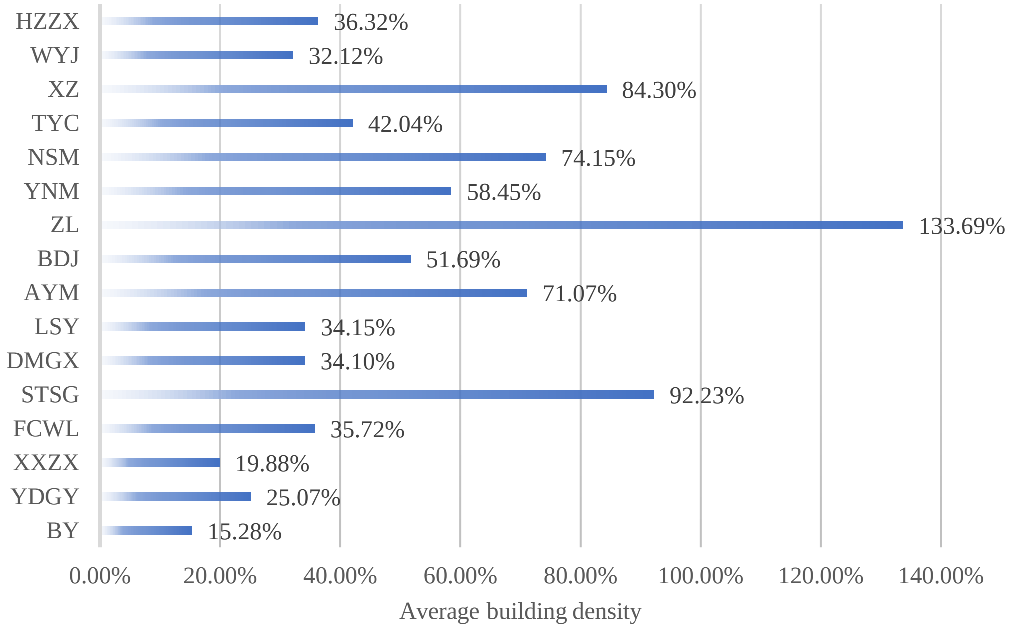


Figure 4- 15. The average building density of each station realm

(4) Plot ratio

This study calculated the plot ratio (PR) for each station realm of Line 2 using ArcGIS 10.7. Figure 4- 16 illustrates the results, where darker shades of red indicate higher PR values. According to Figure 4- 17, which presents the average PR for each station realm, the XZ station realm has the highest average PR of 6.37, followed by the ZL station realm with an average PR of 5.05. The BY station realm has the lowest average PR.

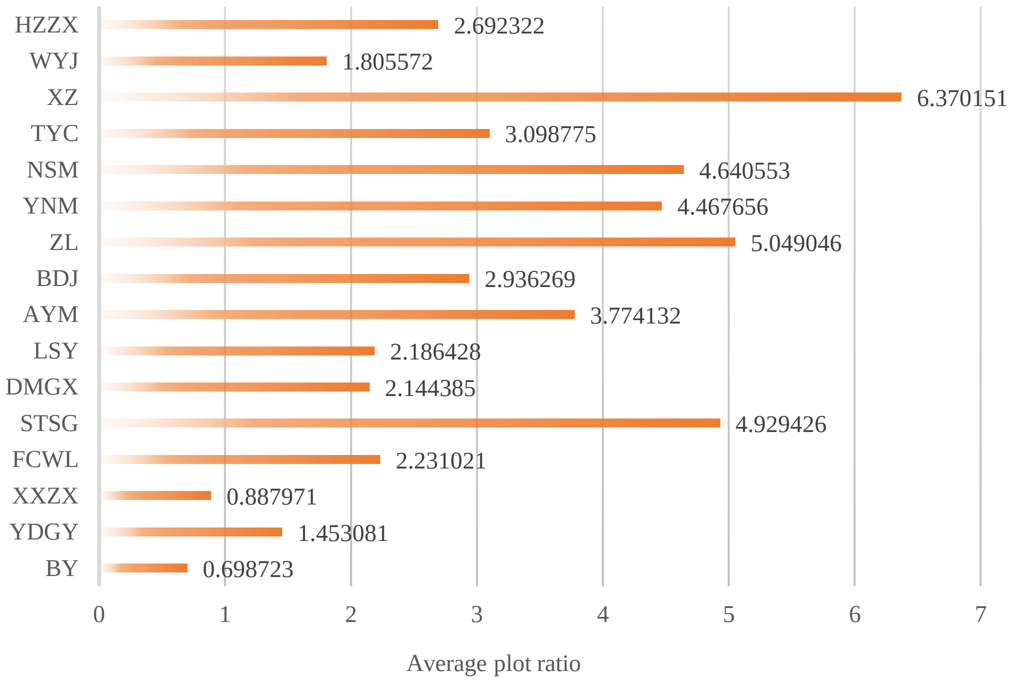


Figure 4- 16. The average plot ratio of each station realm

Figure 4- 17. Statistics of plot ratio within each station realm

4.3.4 Cluster analysis results

In order to facilitate the subsequent analysis and research, the 16 station realms are represented by codes M1 to M16 in this paper, M1 for BY station, and M16 for HZZX station. The codes of each indicator are represented by C1 to C12, as shown in the Table 4- 6.

Table 4- 6. Code for each level indicators

Secondary Indicators	Tertiary Indicators	Code
Station environment	Passing lines	C01
	Number of bus stop	C02
Station traffic environment	Number of transportation facilities	C03
	Metro entrances and exits	C04
Spatial form of station realm	Compactness ratio	C05
	Road network density	C06
	Fractal dimension	C07
	Building primacy	C08
Spatial density of station realm	Population density	C09
	land use mix	C10
	Average building density	C11
	Average plot ratio	C12

The individual indicators are summarized by station code, as shown in the Table 4- 7.

Table 4- 7. Statistics of indicators for each station realm

Code	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12
M01	1	3	37	4	0.76	3.08	1.78	0.79	30.17	0.31	0.15	0.70
M02	1	15	90	4	0.71	4.85	1.69	0.81	45.91	0.30	0.25	1.45
M03	2	63	129	8	0.74	3.05	1.61	0.85	36.12	0.62	0.20	0.89
M04	1	68	181	4	0.72	3.71	1.69	0.78	35.63	0.47	0.36	2.23
M05	1	61	193	4	0.71	3.22	1.72	0.76	43.60	0.58	0.92	4.93
M06	1	44	138	4	0.76	4.73	1.72	0.82	38.12	0.59	0.34	2.14
M07	1	53	192	4	0.69	6.73	1.72	0.82	35.93	0.40	0.34	2.19
M08	1	70	106	4	0.67	6.8	1.72	0.85	23.07	0.61	0.71	3.77
M09	2	67	158	4	0.62	7.28	1.64	0.75	60.42	0.49	0.52	2.94
M10	2	63	196	11	0.75	8.53	1.67	0.71	108.32	0.55	1.34	5.05
M11	1	35	152	4	0.73	6.91	1.70	0.8	52.50	0.42	0.58	4.47
M12	2	52	123	6	0.81	4.29	1.65	0.81	41.61	0.51	0.74	4.64
M13	1	46	169	4	0.7	4.45	1.64	0.87	33.21	0.43	0.42	3.10
M14	2	52	148	6	0.77	3.46	1.65	0.86	90.19	0.48	0.84	6.37
M15	1	58	114	4	0.82	3.30	1.73	0.81	34.41	0.39	0.32	1.81
M16	1	101	174	4	0.8	4.15	1.73	0.82	41.16	0.58	0.36	2.69

(1) Data standardization

In the process of cluster analysis, variables of different magnitudes must be standardized for subsequent cluster analysis because the variables have different magnitudes. The basic idea of data standardization is to calculate the obtained data according to the mathematical concept and reduce the values of the original variables to a relatively small interval to facilitate categorization and comparison. In this paper, we use the standardized fraction normalization method with the following equation:

$$x_{ij} = \frac{x_{ij} - \mu_{ij}}{\sigma_{ij}} \qquad 4 - 1$$

where μ_j is the mean of the j-th column of sample data, σ_j is the standard deviation of the j-th column of sample data.

where μ_j is calculated as follows:

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \qquad 4 - 2$$

where σ_j is calculated as follows:

$$\sigma_j = \left[\frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \mu_j\right)^2\right]^{\frac{1}{2}} \qquad 4 - 3$$

The standardization results are shown in the Table 4- 8.

Table 4- 8. Statistics of each indicator after standardized treatment

Code	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12
M01	-0.65	-2.21	-2.47	-0.47	0.47	-1.04	1.99	-0.40	-0.74	-1.70	-1.18	-1.46
M02	-0.65	-1.69	-1.24	-0.47	-0.47	-0.03	-0.03	0.07	-0.04	-1.79	-0.87	-1.00
M03	1.44	0.43	-0.34	1.54	0.09	-1.05	-1.82	1.02	-0.48	1.34	-1.03	-1.34
M04	-0.65	0.65	0.86	-0.47	-0.28	-0.68	-0.03	-0.63	-0.50	-0.13	-0.52	-0.52
M05	-0.65	0.34	1.14	-0.47	-0.47	-0.96	0.64	-1.11	-0.15	0.95	1.25	1.13
M06	-0.65	-0.41	-0.13	-0.47	0.47	-0.10	0.64	0.31	-0.39	1.05	-0.58	-0.58
M07	-0.65	-0.01	1.12	-0.47	-0.85	1.03	0.64	0.31	-0.49	-0.81	-0.58	-0.55
M08	-0.65	0.74	-0.87	-0.47	-1.23	1.07	0.64	1.02	-1.06	1.24	0.59	0.42
M09	1.44	0.61	0.33	-0.47	-2.17	1.35	-1.15	-1.34	0.60	0.07	-0.01	-0.09
M10	1.44	0.43	1.21	3.06	0.28	2.05	-0.48	-2.29	2.74	0.66	2.58	1.20
M11	-0.65	-0.80	0.19	-0.47	-0.09	1.14	0.20	-0.16	0.25	-0.62	0.18	0.84
M12	1.44	-0.05	-0.48	0.54	1.42	-0.35	-0.92	0.07	-0.24	0.26	0.68	0.95
M13	-0.65	-0.32	0.58	-0.47	-0.66	-0.26	-1.15	1.49	-0.61	-0.52	-0.33	0.01
M14	1.44	-0.05	0.10	0.54	0.66	-0.82	-0.92	1.25	1.93	-0.03	1.00	2.00
M15	-0.65	0.21	-0.69	-0.47	1.61	-0.91	0.87	0.07	-0.56	-0.91	-0.65	-0.78
M16	-0.65	2.11	0.70	-0.47	1.23	-0.43	0.87	0.31	-0.26	0.95	-0.52	-0.24

(2) Clustering factor extraction

Before clustering, a homogeneity analysis was conducted on the standardized variable factors, using Principal Component Analysis (PCA) to extract common factor variables from the 12 variables, thereby improving the accuracy of the observed data.

The KMO test and Bartlett's test were performed, as shown in Table 4- 9, using 12 site factor variables and 16 site variables for factor analysis. Based on the test results, it can be observed that the KMO test value is 0.641, indicating good intercorrelation among the variables. The p-value is 0.00, suggesting that the data structure is reasonable and the initial factor variables are suitable for factor analysis.

Table 4- 9. KMO and Bartlett's test

KMO test value		0.641
	Chi-Square	114.665
KMO and Bartlett's test	*df*	66
	P value	0.000

Principal Component Analysis (PCA) was used to extract common factors, and the extraction results are shown in Table 4- 10. The results of factor extraction indicate that all factor extraction values are above 0.5, indicating that they effectively capture the information from the initial variables.

Table 4- 10. Communality statistics

Code	Initial value	Extracted value
C01	1.000	0.886
C02	1.000	0.841
C03	1.000	0.715
C04	1.000	0.768
C05	1.000	0.898
C06	1.000	0.747
C07	1.000	0.901
C08	1.000	0.608
C09	1.000	0.860
C10	1.000	0.753
C11	1.000	0.877
C12	1.000	0.622

Based on the initial variables of the 12 factors, four variables were extracted with eigenvalues greater than 1. The specific factor analysis variances are shown in Table 4- 11.

Table 4- 11. Total variance explained

Factor	Eigen value (Unrotated)			Eigen value (Rotated)		
	Total	% of Variance	Cumulative % of Variance	Total	% of Variance	Cumulative % of Variance
1	4.656	38.802	38.802	2.984	24.870	24.870
2	1.789	14.908	53.711	2.530	21.083	45.953
3	1.723	14.358	68.068	2.403	20.025	65.977
4	1.309	10.908	78.976	1.560	12.999	78.976
5	0.883	7.355	86.331			
6	0.584	4.866	91.198			
7	0.424	3.535	94.733			
8	0.282	2.352	97.085			
9	0.168	1.400	98.486			
10	0.109	0.909	99.395			
11	0.063	0.527	99.922			
12	0.009	0.078	100.000			

Extraction method: Principal component analysis

In summary, factor analysis extracted a total of 4 common factors with eigenvalues greater than 1 from the 12 variables. Based on the cumulative results, these four common factors can explain 78.976% of the information in the initial sample, indicating that the factor extraction falls within a reasonable range. The extracted common factors can effectively explain the initial variables.

To better interpret the extracted public factors, the factor loading matrix is rotated using orthogonal transformation, which can simplify the structure of the rotated factor loading matrix [8,9]. Table 4- 12 shows the rotated factor loading matrix and Table 4- 13 shows the component score coefficient matrix. Therefore, for the common factor F1, the loading values of road network density (C06), population density (C09), building density (C11), and plot ratio (C12) are relatively close and range from 0.533 to 0.839, indicating a positive correlation. However, the loading value of building prominence (C08) is negatively correlated (-0.739). Therefore, the factor F1 represents station realms with an average distribution of building heights, high road network density, high floor area ratio, and high population and building densities. F1 can be defined as the station realm scale factor.

For the common factor F2, the loading values of transit lines (C01), subway entrances/exits (C04), and population density (C09) are relatively close and positively correlated. The loading value of fractal dimension (C07) is negatively correlated. Therefore, F2 represents station realms with a high number of transit lines and station entrances/exits, leading to high population density. F2 can be defined as the station realm pedestrian flow factor.

Table 4- 12. Factor loading (Rotated)

Code	Factor loading			
	F1	F2	F3	F4
C01		0.900		
C02			0.915	
C03			0.717	
C04		0.697		
C05				0.944
C06	0.533			-0.675
C07		-0.886		
C08	-0.739			
C09	0.780	0.502		
C10			0.832	
C11	0.829			
C12	0.637			

For the common factor F3, the loading values of bus station quantity (C02), transportation facility quantity (C03), and land use mix (C10) are relatively high and close. This indicates a well-connected transfer environment and high land use mix in the station realms. F3 can be defined as the station realm transfer factor.

For the common factor F4, the loading value of boundary compactness (C05) is significantly positive, while the road network density (C06) is negatively correlated. This suggests that the station realms are in a state of internal functional renewal or external expansion. Therefore, F4 can be defined as the station realm development factor.

Table 4- 13. Component Score Coefficient Matrix

Code	Component			
	F1	F2	F3	F4
M01	-0.13	-0.88	-2.45	0.83
M02	-0.34	0.00	-1.88	-0.56
M03	-1.77	2.12	0.39	0.35
M04	-0.33	-0.60	0.52	-0.25
M05	0.82	-1.08	1.09	0.20
M06	-0.39	-0.57	0.25	0.37
M07	-0.15	-0.77	0.03	-1.21
M08	-0.42	-0.59	0.86	-0.83
M09	0.10	0.82	0.00	-2.27
M10	2.93	1.04	0.11	-0.09
M11	0.69	-0.53	-0.44	-0.54
M12	0.12	1.05	0.01	1.17
M13	-1.07	0.35	0.15	-0.68
M14	0.49	1.44	0.10	1.12
M15	-0.24	-0.78	-0.37	1.35
M16	-0.32	-1.01	1.63	1.04

（3）Cluster analysis results

Using the extracted 4 common factors as clustering variables, the data was input into the SPSS software for cluster analysis. To achieve a more reasonable classification of rail transit station realms, clustering analyses were performed with different values of K, set as 3, 4, 5, and 6. Considering the output results and the existing situation, the clustering result with K=5 is determined as the final station realm classification in this study. The clustering result is presented in Table 4- 14.

Table 4- 14 Cluster analysis results of Metro Line 2 station realm

Code	Cluster	Distance
M01	1	0.878
M02	1	0.878
M03	3	1.608
M04	4	0.888
M05	5	1.123
M06	5	0.713
M07	4	0.621
M08	4	0.81
M09	4	1.702
M10	2	0
M11	4	1.209
M12	3	0.776
M13	4	1.078
M14	3	0.914
M15	5	1.212
M16	5	1.075

4.3.5 Classification result of station realm types

Based on the above analysis, the station realms in this study are divided into 5 categories. The specific classification results are shown in Table 4- 15.

Category 1 includes BeiYuan station (M01) and YunDongGongYuan station (M02). The classification is mainly based on the station realm development factor. These station realms are characterized by their distant geographic location from the city center and the absence of significant central clusters. The population density, land-use mix, and land use intensity are not high, indicating that they are in the stage of development. Therefore, they are classified as under development station realms.

Category 2 includes ZhongLou station (M10). The classification is primarily based on the station realm scale factor, followed by the station realm pedestrian flow factor and transfer factor. These station realms have a high number of transit lines and subway entrances, extensive bus connections, complete transportation facilities, and are located in the city center. They have high population density, land-use mix, and land use intensity. Therefore, they are classified as city center station realms.

Category 3 includes XingZhengZhongXin station (M03), NanShaoMen station (M12), and XiaoZhai station (M14). The classification is primarily based on the station realm pedestrian flow factor, followed by the station realm development factor and transfer factor. These station realms have a high number of transit lines and subway entrances, high population density, a large number of bus stops and transportation services, and a high level of development and land-use mix. Therefore, they are classified as high-density hub station realms.

Category 4 includes FengChengWuLu station (M04), LongShouYuan station (M07), AnYuanMen station (M08), BeiDaJie station (M09), YongNingMen station (M11), and TiYuChang station (M13). The classification is primarily based on the station realm transfer factor. These station realms have a large number of bus connections, a significant number of transportation services, and a high land-use mix. Therefore, they are classified as transfer station realms.

Category 5 includes ShiTuShuGuan station (M05), DaMingGongXi station (M06), WeiYiJie station (M15), and HuiZhanZhongXin station (M16). The classification is primarily based on the station realm transfer factor and development factor. These station realms have a significant number of bus and transportation services and a higher number of residential areas, resulting in lower road network density and relatively moderate development. Therefore, they are classified as general residential station realms.

Table 4- 15. Xi'an Metro Line 2 station realm classification results

Station realm Type	Number	Station realm
Under development type	2	BY, YDGY
Urban center type	1	ZL
High population density Hub type	3	XXZX, NSM, XZ
Interchange type	6	FCWL, LSY, AYM, BDJ, YNM, TYC
General residential type	4	STSG, DMGX, WYJ, HZZX

4.4 Data collection

The data for population activities and POIs in this paper are sourced from the Tencent Location Big Data Service window (https://heat.qq.com/index.php) and the AutoNavi Map's Public Application Programming Interface (API) (https://lbs.amap.com). These applications can collect all the data in the delimited area and export them to csv files. The details are shown in Table 4- 16.

Table 4- 16. The specific metadata of the obtained original dataset

Data attributes	EasyGO	AutoNavi map
Location	Longitude	Longitude
	Latitude	Latitude
Time information	Date	--
	Time period	--
Ancillary information	Number of Activity people	Place name
Data categorization		Place address
Data ID	Date	City functional nature

4.4.1 Population activity data

In this study, the Easygo data were collected from November 17th and November 20th, 2021 (17th is working day and 20th is off day) using the boundary of each research station realm as the data collection scope, and the collection time was from 7:00 a.m. to 22:00 p.m. (according to the Xi'an metro operation time), and were collected including raw data in CSV format such as activity date, activity number, longitude, and dimension of pedestrians. A total of 12,283 pieces of population activity data were collected for this study, and the specific statistics for each research station realm are shown in the Figure 4- 18.

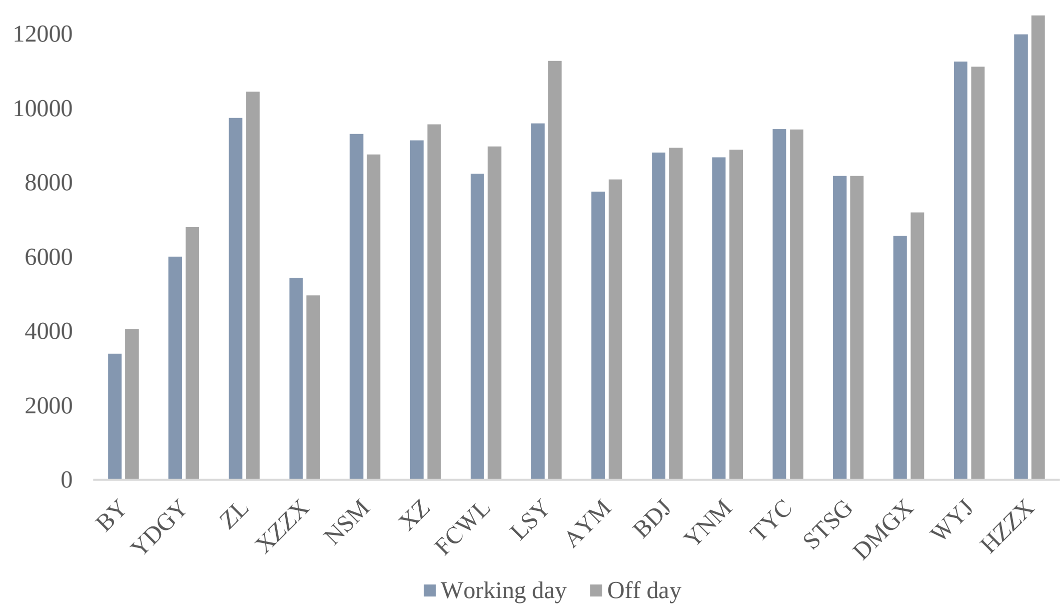


Figure 4- 18. Statistics on the number of population activities in each research station realm

4.4.2 Functional facility data

The POI data were also collected within the boundaries of each research station realm, and the collection time was December 30, 2021, with a total of 37,012 pieces. The quantity statistics of POI data are shown in Figure 4- 19.

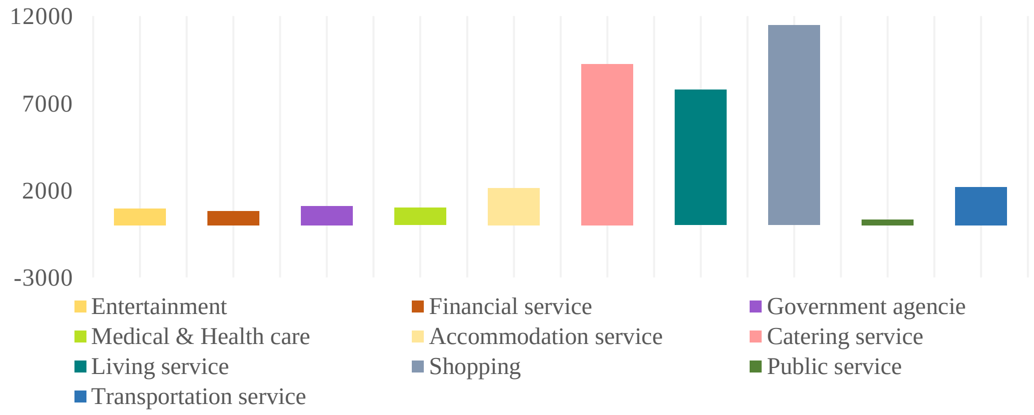


Figure 4- 19. Statistics on the number of various facilities in each research station realm

4.5 Analysis of spatial distribution pattern within station realm

4.5.1 Overall station realm spatial distribution pattern of Xi'an Metro Line 2

This study first identified the overall spatial distribution pattern of the Xi'an Metro Line 2 from a global perspective to understand the basic distribution characteristics of population activities and functional facilities within the station realms. As shown in the Figure 4- 20, ArcGIS was used to create a 3D model of the urban space within the Xi'an Metro Line 2 stations. All spatial distribution pattern recognition and subsequent interactive mechanism research in this study were based on this spatial scope.

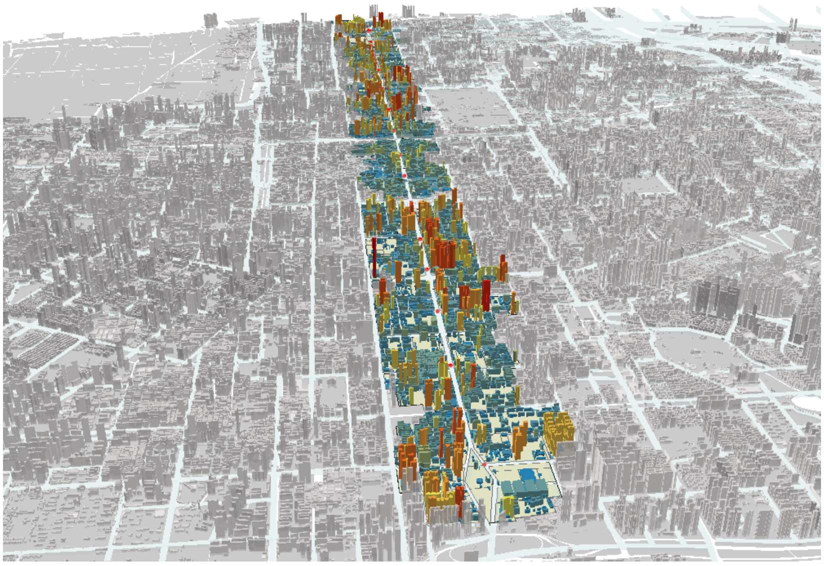


Figure 4- 20. 3D visualization of Xi'an Metro Line 2 station realm

(1) The spatial distribution pattern of population activities within Metro Line 2 station realm

The temporal distribution of population activities on working day and off day is shown in the Figure 4- 21. On working day, population activities on Metro Line 2 are mainly characterized by commuting, resulting in distinct peak hours. The morning peak is concentrated between 08:00 and 10:00, while the evening peak is mainly distributed between 17:00 and 19:00. On off day, there is no obvious morning peak in the temporal distribution of population activities within the Metro Line 2 stations. Instead, there is a continuous fluctuation between 09:00 and 18:00. Additionally, the population activity level on off day within this time range is higher than on working day, indicating that people prefer to engage in activities between 09:00 and 18:00 on off day.

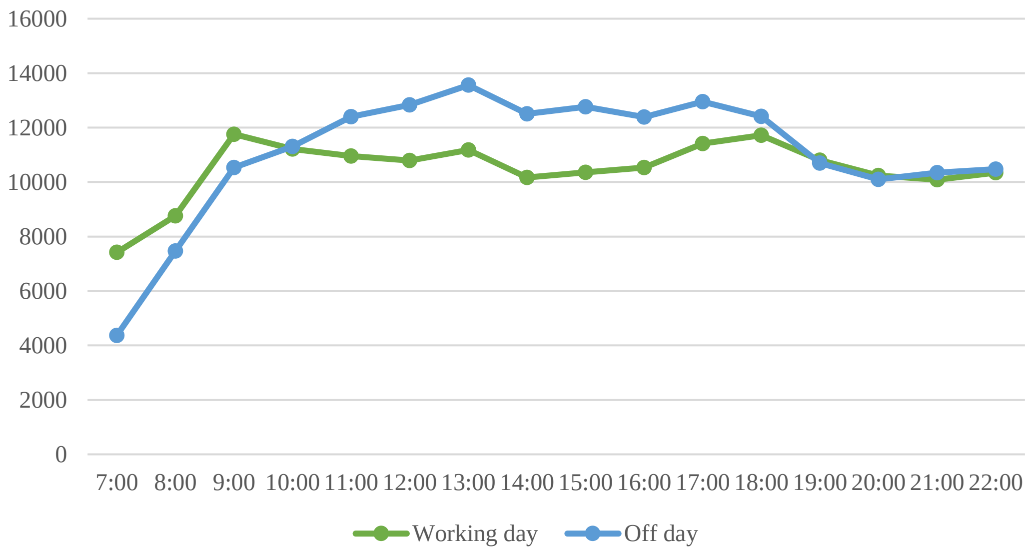


Figure 4- 21. Statistics on single-day population activities within Metro Line 2 station realm

According to the global spatial autocorrelation analysis results (Table 4- 17), the Z_I-Score of Moran's I indicates that the population activities within the station realms of Metro Line 2 are in a clustered spatial distribution state on both working day and off day. The Z_I-Score on working day is slightly higher than on off day, indicating that the degree of clustering of population activities within the station realms is relatively higher on working day than on off day. In addition, the Z_I-Score of the high-low clustering analysis for both working day and off day is positive, indicating that the spatial clustering pattern of population activities within the station realms is characterized by high/high clustering, meaning that areas with high population activity levels tend to cluster together.

Table 4- 17. Results of global spatial autocorrelation analysis in Metro Line 2 station realm

Date	Z_I-Score	Z_G-Score
Working day	136.749	134.949
Off day	135.722	134.179

The Figure 4- 22 represents the visualization of the local spatial autocorrelation analysis (hotspot analysis) results. It shows that there is significant spatial heterogeneity in the distribution of population activities within the Metro Line 2 station realm, particularly in the XXZX station realm, NSM station realm, and HZZX station realm, where both high/high clustering and low/low clustering phenomena are observed.

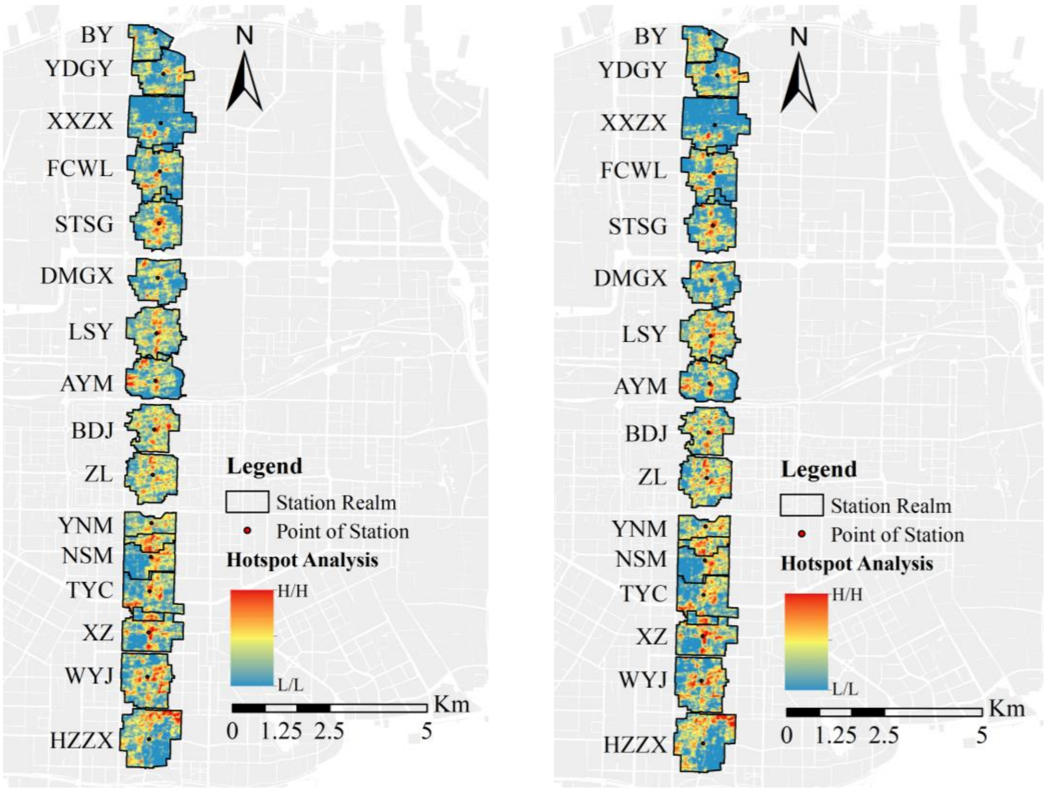


Figure 4- 22. Hotspot analysis of single-day population activity within station realm of Xi'an Metro Line 2. (a) Working day. (b) Off day

The specific statistical results are presented in the Table 4- 18, which indicates that population activities exhibit a higher degree of clustering on working day but with a lower proportion of high/high clustering compared to off day. This suggests that on working day, the spatial distribution of population activities within the stations is more concentrated, while on off day, the spatial distribution of population activities is more widespread.

Table 4- 18. Hotspot analysis statistics of population activities in Metro Line 2 station realm

Date	Percentage of high/high concentration	Max. GiZ-Score	Min. GiZ-Score
Working day	39.50%	10.581	-2.070
Off day	42.91%	10.360	-2.098

(2) Spatial distribution pattern of functional facilities in Line 2 station realms

The results of the average nearest neighbor analysis for each type of facility are illustrated in Figure 4- 23. It can be observed that all types of functional facilities exhibit an aggregated distribution pattern within the station realms of Metro Line 2. Shopping facilities have the highest level of clustering, followed by dining services, while public service facilities show the lowest degree of clustering.

Using the kernel density analysis tool, we conducted further research on the overall spatial distribution of functional facilities in the Metro Line 2 station realms, as well as the spatial distribution of different types of facilities. As shown in the Figure 4- 24, the functional facilities in the station realms of Metro Line 2 exhibit an overall spatial distribution pattern resembling a "central line with six hearts." The facilities are primarily clustered between the FCWL station realm, STSG station realm, LSY station realm, ZL station realm, YNM station realm, and NSM station realm, with XZ station realm encompassing these six realms.

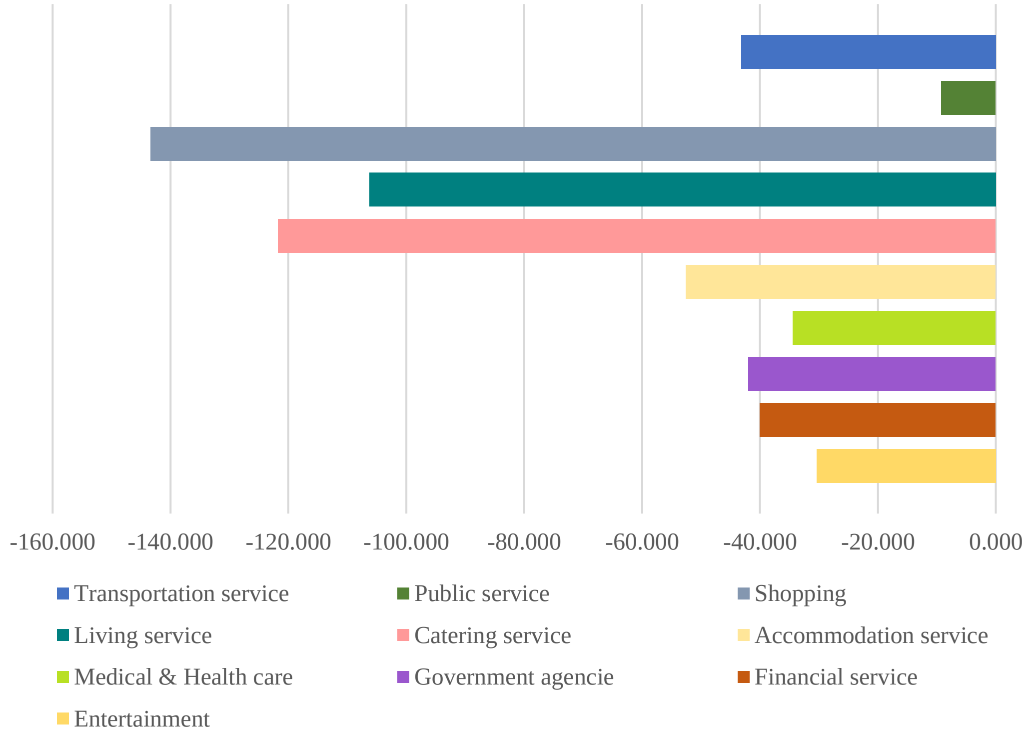
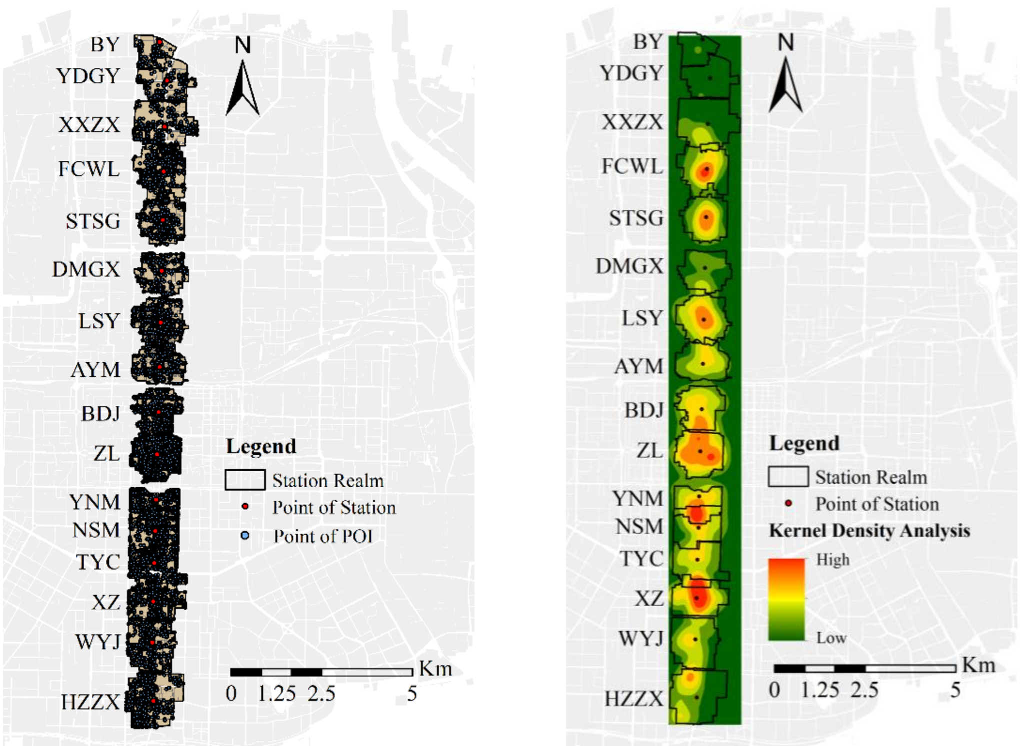


Figure 4- 23. Average Nearest Neighbor Analysis of POI within Metro Line 2 station realm



(a) POI point distribution (b) Kernel density analysis

Figure 4- 24. Spatial distribution of POIs within Xi'an Metro Line 2 station realm

The kernel density analysis of different types of functional facilities, as shown in Figure 4- 25, reveals significant variations in the distribution patterns across different station realms. Among them, the ZL station realm serves as the central area of Metro Line 2. Except for medical and health facilities, other types of functional facilities exhibit a high level of clustering within this realm. Particularly, entertainment facilities, accommodation services, and public service facilities are predominantly concentrated within the ZL station realm. Additionally, compared to other types of facilities, transportation service facilities show a relatively balanced distribution across the entire Metro Line 2 station realms.

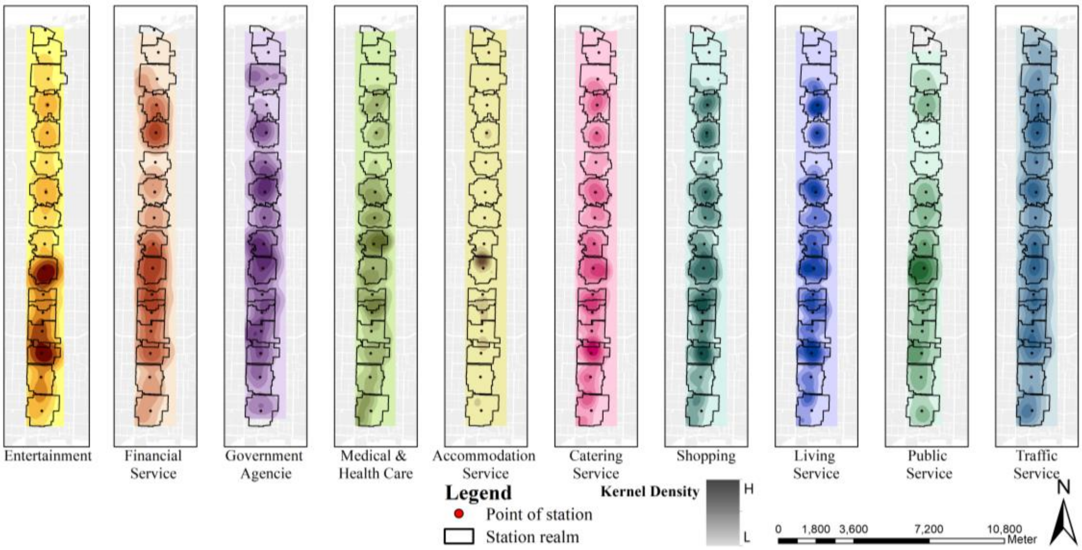


Figure 4- 25. Kernel density analysis of various functional facilities

4.5.2 Analysis of spatial distribution pattern within Under development type station realm

(1) Spatial distribution pattern of population activity

In terms of time distribution, the statistical results of population activity in BY station realm and YDGY station realm on working day and off day are shown in the Figure 4- 26 and Figure 4- 27. On working day, population activity in BY station realm and YDGY station realm is mainly characterized by commuting, with noticeable peaks during morning and evening rush hours. The morning peak is concentrated between 08:00 and 10:00, while the evening peak is mainly distributed between 17:00 and 19:00. On off day, there is no significant morning peak in population activity in BY station realm. The number of population activities continues to rise between 07:00 and 09:00 and starts to decline after 09:00. In YDGY station realm, population activity shows a continuous fluctuating pattern between 11:00 and 17:00. Overall, the population activity in BY station realm and YDGY station realm during the time period of 09:00 to 18:00 on off day is higher than on working day.

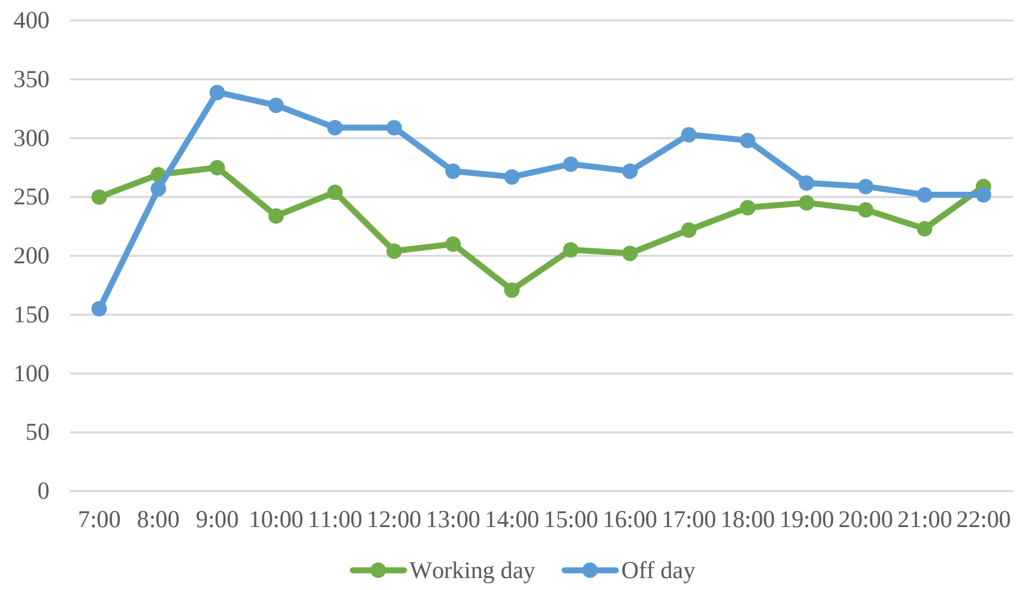


Figure 4- 26. Statistics on number of single-day population activities within YDGY station realm

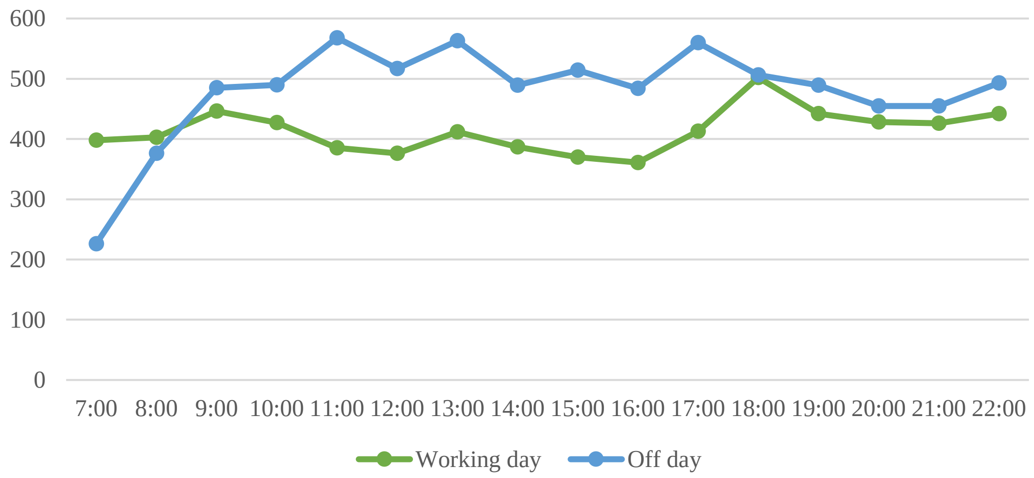


Figure 4- 27. Statistics on number of single-day population activities within BY station realm

In terms of spatial distribution, based on the results of the High/Low clustering analysis (Table 4-19), both BY station realm and YDGY station realm show positive Z_G-Score for population activity on working day and off day. This indicates that the spatial clustering type of population activity in these two station realms is high/high clustering, meaning a high number of population activities are concentrated in areas with high population activities.

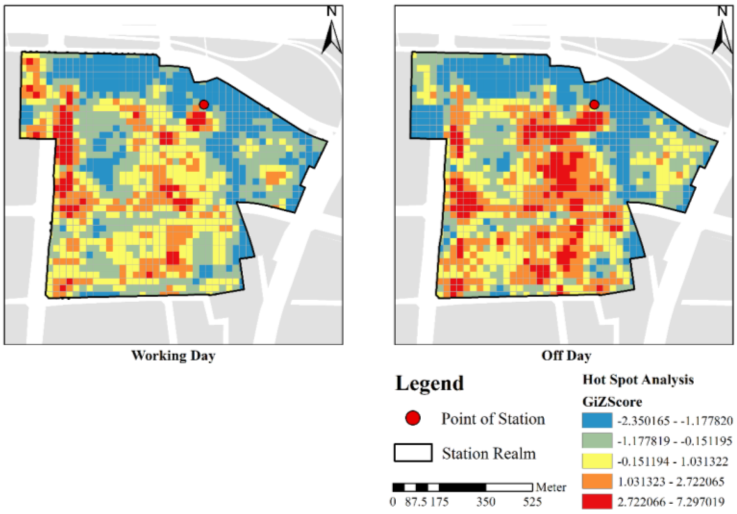
Table 4- 19. High/low cluster analysis statistics of single-day population activities within under development type station realm

Code	Number of people		z_G- score	
	Working day	Off day	Working day	Off day
BY	3397	4055	20.04	20.79
YDGY	6007	6798	27.07	27.48

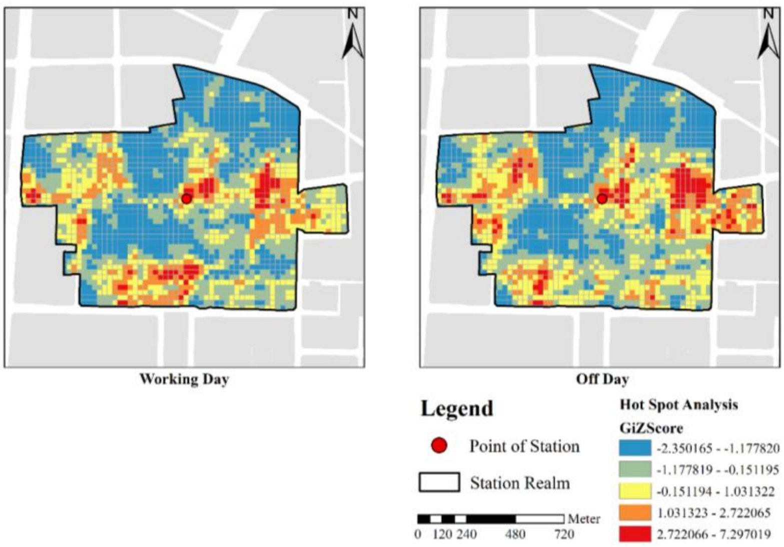
The Figure 4- 28 presents the visualization of the hotspot analysis results. It can be observed that population activity in the BY station realm is mainly concentrated in the southwest area of the station, while in the YDGY station realm, population activity is primarily concentrated in the east. Furthermore, considering the statistical results from Table 4- 20, the proportion of high/high clustering in both the BY station realm and YDGY station realm is higher on off day compared to working day. This indicates that on off day, the range and degree of population activity are larger than on working day.

Table 4- 20. Statistics of the percentage of high/high z_G- score in station realm

Code	Percentage of H/H z_G- score	
	Working day	Off day
BY	43.13%	46.72%
YDGY	42.20%	43.46%



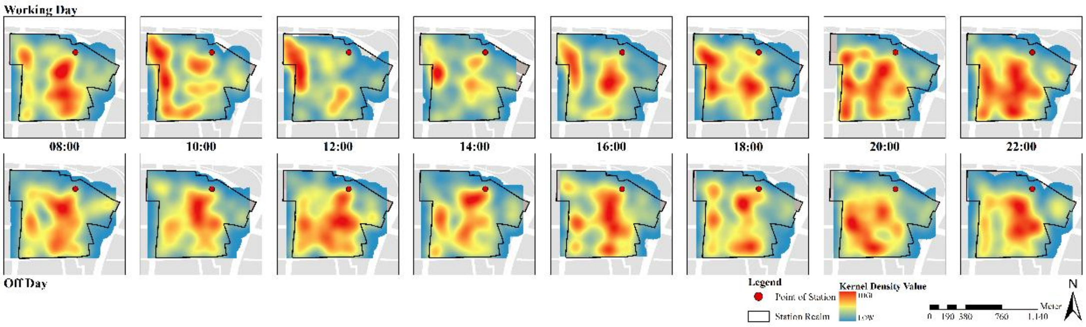
(a) BY station realm



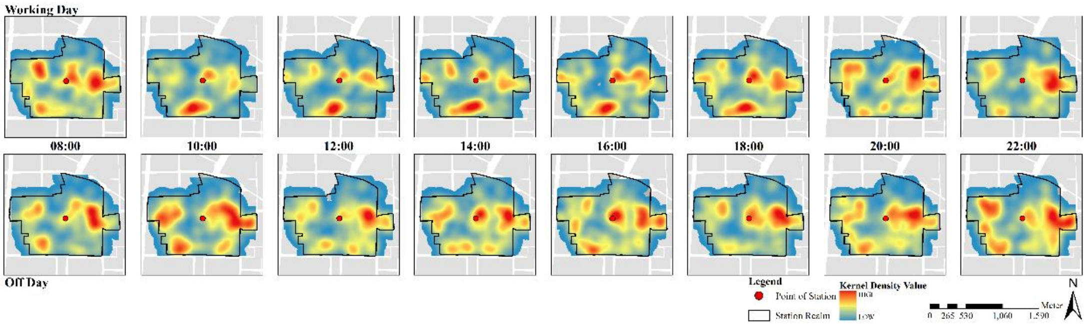
(b) YDGY station realm

Figure 4- 28. Hot spot analysis of single-day population activity within under development type station realm

In addition, this study also conducted hotspot analysis on population activity at different time periods on working day and off day in the BY station realm and YDGY station realm. As shown in the Figure 4- 29, the distribution of population activity within the station realms exhibits different characteristics as it varies with time.



(a) BY station realm



(b) YDGY station realm

Figure 4- 29. Temporal and spatial distribution of population activities within under development type station realm

(2) Spatial distribution pattern of functional facility

The analysis results for functional facility within the BY station realm and YDGY station realm are shown in the Table 4- 21. In the BY station realm, only accommodation facilities, catering services, living services, shopping facilities, and transportation services exhibit clustered distribution patterns, with dining services showing the highest level of clustering. In the YDGY station realm, accommodation facilities, catering services, living services, shopping facilities, and transportation services also exhibit clustered distribution patterns, with shopping services showing the highest level of clustering.

Table 4- 21. Average Nearest Neighbor Analysis of POI within under development type station realm

POI types	BY station realm		YDGY station realm	
	Number	Average Nearest Neighbor (Z-score)	Number	Average Nearest Neighbor (Z-score)
Entertainment	5	3.62	10	2.14
Financial service	3	14.14	6	4.27
Government agency	2	125.69	3	30.55
Medical & Health care	6	1.69	6	0.11
Accommodation service	22	-2.66	19	-0.13
Catering service	48	-6.63	37	-5.96

Living service	49	-2.62	75	-7.58
Shopping	42	-5.32	83	-10.25
Public service	0	0.00	3	7.77
Transportation service	37	-5.77	90	-5.13

Figure 4- 30 illustrates the visual representation of the kernel density analysis results for the distribution of functional facilities within BY station realm and YDGY station realm. It can be observed that in BY station realm, functional facilities are primarily clustered in the southern part of the station, exhibiting a high-density distribution. On the other hand, in YDGY station realm, functional facilities are mainly concentrated in the eastern and southern parts of the station, also demonstrating a high-density distribution. This indicates significant regional variations in the spatial distribution of functional facilities within these two station realms.

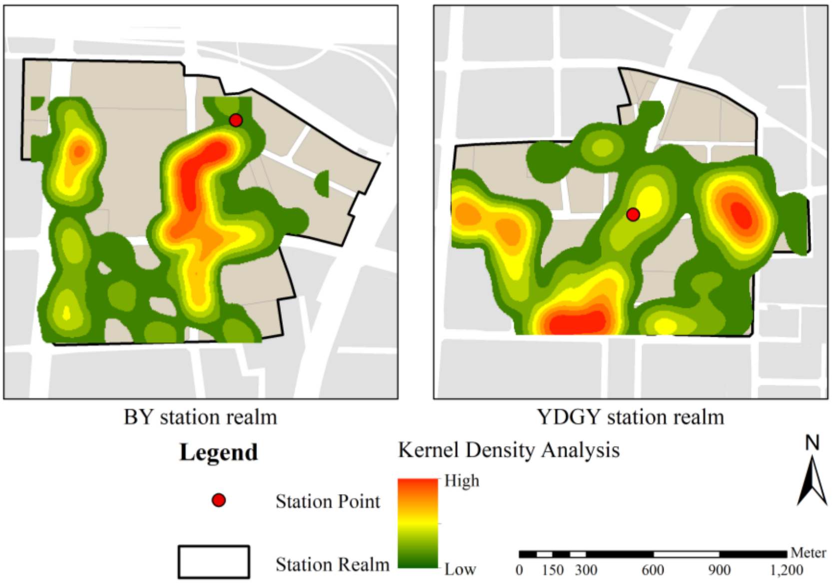


Figure 4- 30. Kernel density analysis of functional facilities within station realm

4.5.3 Analysis of spatial distribution pattern within Urban center type station realm

(1) Spatial distribution pattern of population activity

The statistical results of population activities in ZL station realm on both working day and off day are presented in Figure 4- 31. On working day, population activities in ZL station realm exhibit a commuting pattern, with clear peak hours. The morning peak hour is concentrated between 08:00 and 10:00, while the evening peak hour is mainly distributed from 16:00 to 18:00, with higher population activity during the evening peak hour compared to the morning peak hour. On off day, there is a distinct unimodal distribution of population activities in ZL station realm. Throughout the day, population activities steadily increase between 07:00 and 13:00 and start to decline after 13:00. Overall, ZL station realm experiences higher population activity levels on off day between 10:00 and 19:00 compared to working day.

Regarding the spatial distribution, based on the results of the High/Low clustering analysis (Table 4- 22), both working day and off day show positive Z_G-Score for population activities in ZL station realm, indicating a high/high clustering pattern. This implies a concentration of population activities with high population density and clustering in the spatial distribution within ZL station realm.

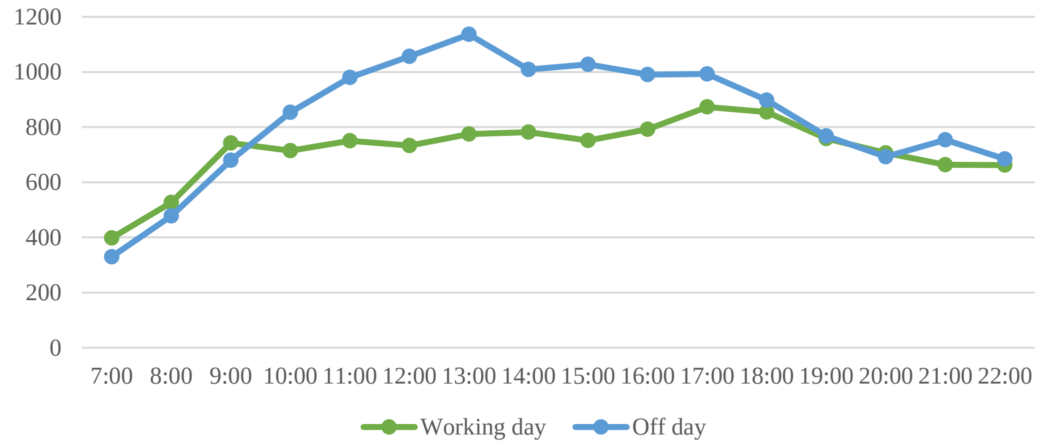


Figure 4- 31. Statistics of single-day population activities within ZL station realm

Table 4- 22. High/low cluster analysis results of single-day population activities within ZL station realm

Analysis method	Indicators	Date	
		Working day	Off day
Numerical statistics	Number of people	9739	10447
High/Low Clustering Analysis	Z_G-Score	24.57	27.15

Figure 4- 32 shows the visualization of hotspot analysis results, it can be observed that the areas of population activity aggregation within ZL station realm are mainly concentrated within a radius of 500 meters around the station. Additionally, considering the statistical results from the table, the proportion of high/high clustering in ZL station realm is 44.11% on working day and 46.35% on off day. This indicates that on off day, the range and clustering intensity of population activities are greater compared to working day.

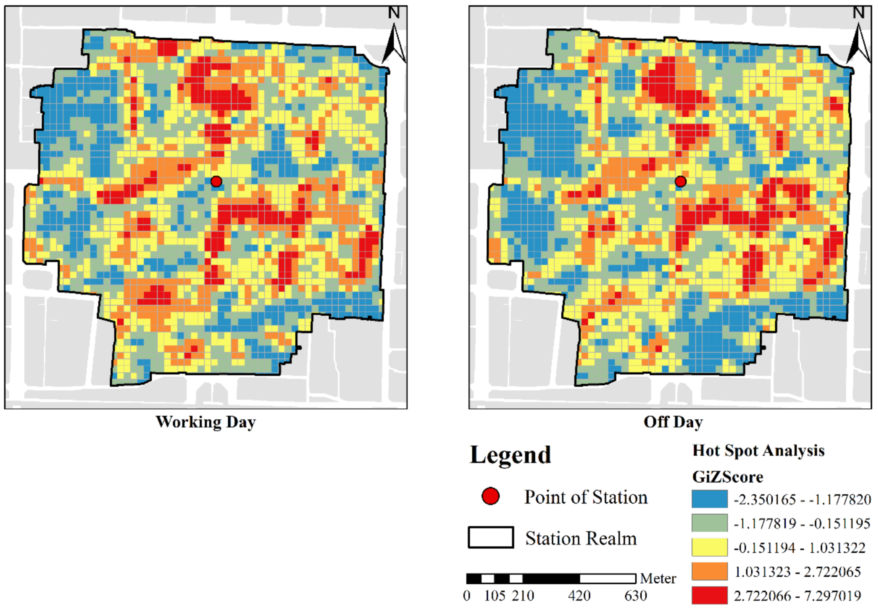


Figure 4- 32. Hot spot analysis of single-day population activity within ZL station realm

In addition, this study also conducted hotspot analysis on population activity at different time periods on working day and off day in the ZL station realm. As shown in the Figure 4- 33, the distribution of population activity within the station realms exhibits different characteristics as it varies with time.

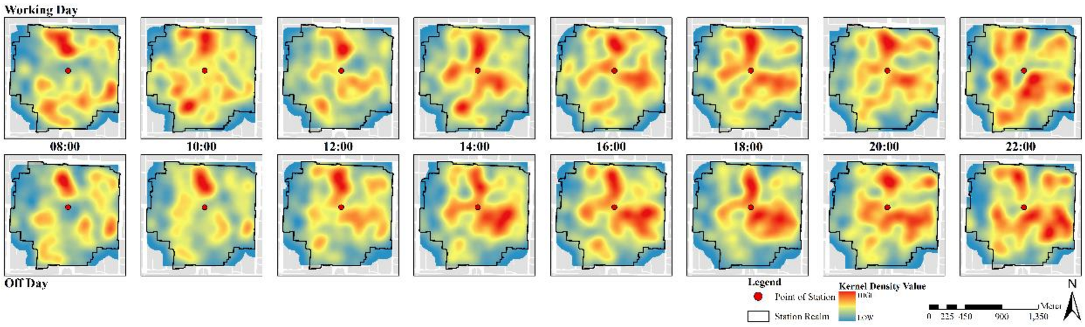


Figure 4- 33. Temporal and spatial distribution of population activities within ZL station realm

The average nearest neighbor analysis results for various types of facilities within ZL station realm are presented in the Table 4- 23, indicating that all types of facilities exhibit an aggregated distribution pattern. Among them, shopping services show the highest level of aggregation, followed by catering services and accommodation services.

Table 4- 23. Average Nearest Neighbor Analysis of POIs within ZL station realm

POI types	Number of POIs	Average Nearest Neighbor (Z score)
Entertainment	174	-12.30
Financial service	118	-12.80
Government agency	140	-12.22
Medical & Health care	105	-9.25
Accommodation service	583	-22.47
Catering service	1293	-36.09
Living service	1049	-29.79
Shopping	1634	-47.35
Public service	82	-4.85
Transportation service	196	-10.72
Total	5374	-19.78

In terms of spatial distribution characteristics, the visualization of the kernel density analysis results for facility distribution within the station realm shows that facilities mainly aggregate in the western and northern areas of ZL station realm (Figure 4- 34). It can be determined that the aggregated areas of facilities are also distributed within a radius of 500 meters from the station.

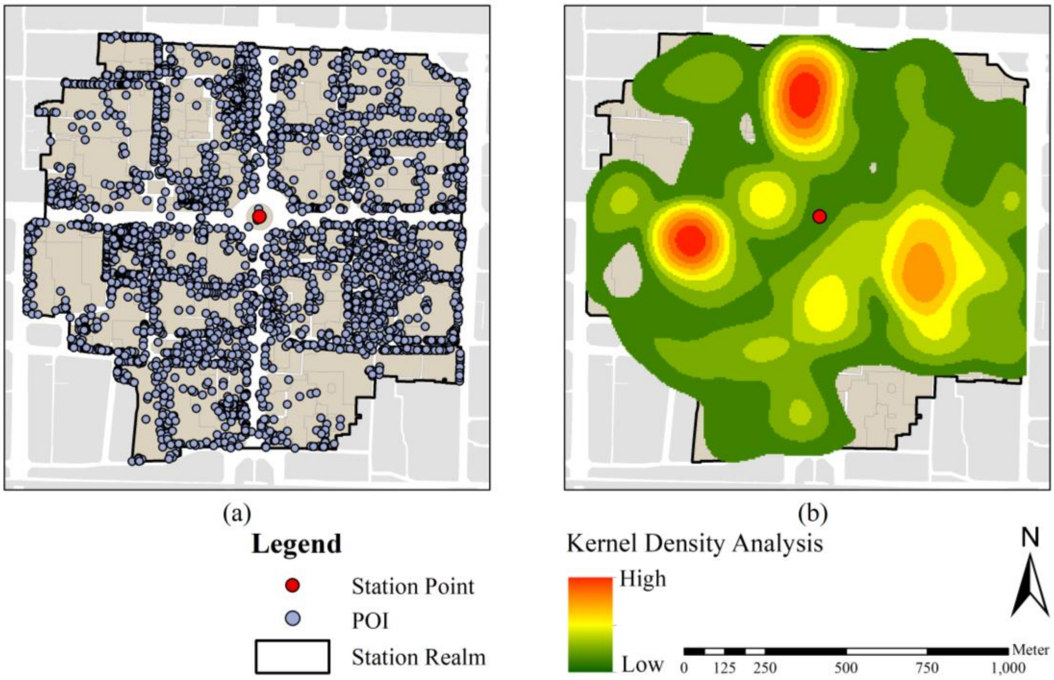


Figure 4- 34. Kernel density analysis of functional facilities within ZL station realm

4.5.4 Analysis of spatial distribution pattern within High population density hub type station realm

High population density hub type station realm including XZZX station realm, NSM station realm and XZ station realm.

(1) Spatial distribution pattern of population activities within station realms

Regarding the temporal distribution, the statistical results of population activities within the station realms during working day are shown in the Figure 4- 35. XZZX station realm exhibits peak hours during the morning rush from 08:00 to 10:00 and the afternoon rush from 15:00 to 17:00. NSM station realm experiences peak hours from 09:00 to 11:00 and 17:00 to 19:00. XZ station realm shows three peak periods, namely 08:00 to 10:00, 12:00 to 14:00, and 16:00 to 18:00.

On off day, the temporal distribution of population activities within the station realms is illustrated in the Figure 4- 36. XZZX station realm displays a clear morning peak between 08:00 and 10:00. The population activities in NSM station realm gradually increase from 07:00 to 12:00 and then fluctuate. In XZ station realm, the population activities steadily rise from 07:00 to 15:00 and then show a declining trend after 15:00.

Overall, there are significant differences in the number of population activities between working day and off day in XZZX station realm, NSM station realm, and XZ station realm. Particularly, XZ station realm experiences a substantial increase in population activities on off day.

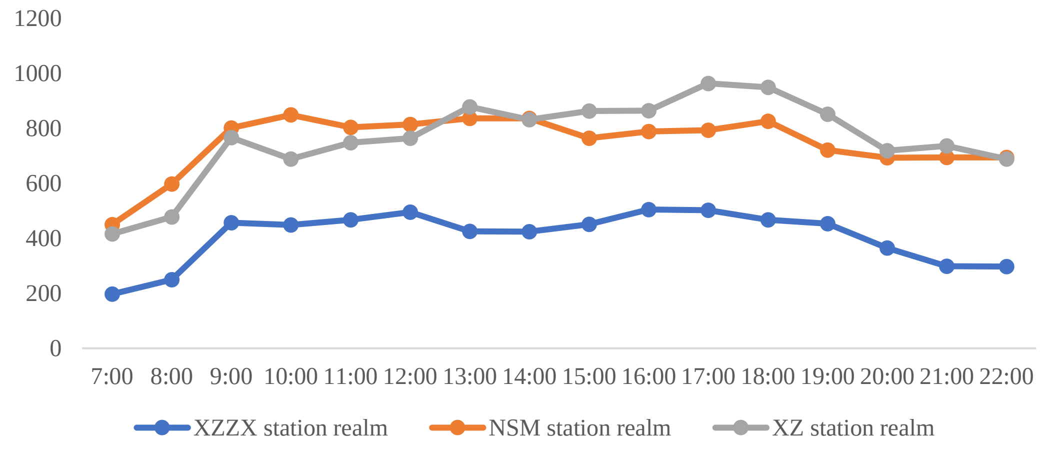


Figure 4- 35. Statistics of working day population activities within High population density hub type station realm

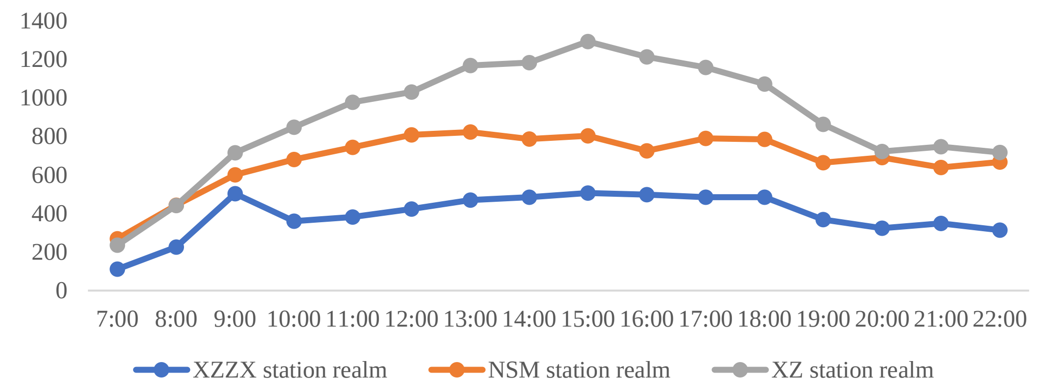


Figure 4- 36. Statistics of off day population activities within High population density hub type station realm

In terms of spatial distribution, based on the results of the High/Low clustering analysis (Table 4-24), the z_G- score for population activities in XZZX station realm, NSM station realm, and XZ station realm are positive for both working day and off day. This indicates that the spatial clustering type of population activities in these three station realms is high/high clustering, meaning there is a high quantity of population activities and a high degree of spatial clustering of population activities.

Table 4- 24. High/low cluster analysis statistics of single-day population activities within high population density-hub type station realm

Code	Number of people		z_G- score	
	Working day	Off day	Working day	Off day
XZZX	5435	4965	41.73	40.29
NSM	9307	8757	36.63	36.17
XZ	9133	9566	36.12	37.68

The Figure 4- 37 displays the visualization of hotspot analysis results. It can be observed that in XZZX station realm, population activities mainly cluster in the southwest area of the station realm. In NSM station realm, population activities are primarily concentrated in the northern and western parts of the station. In XZ station realm, population activities are mainly concentrated around the station and in the northern part of the station realm.

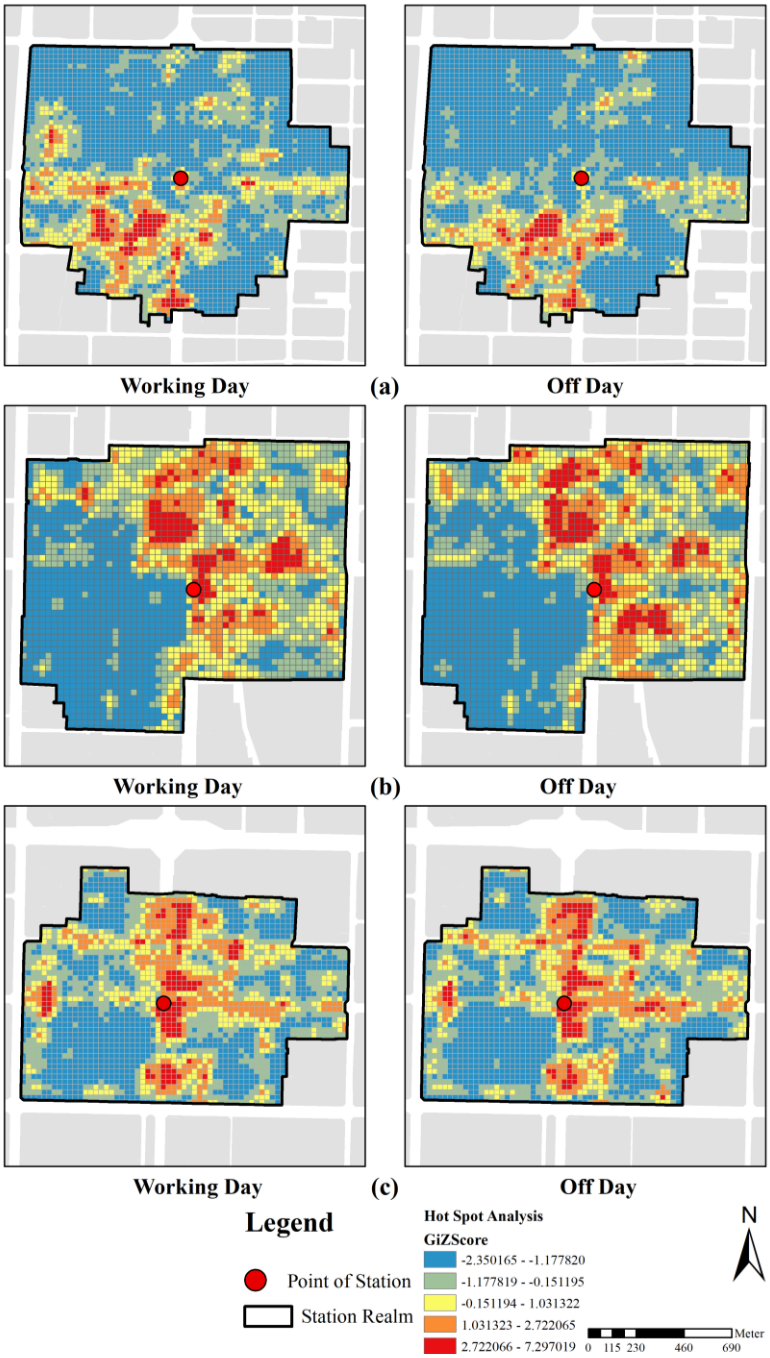


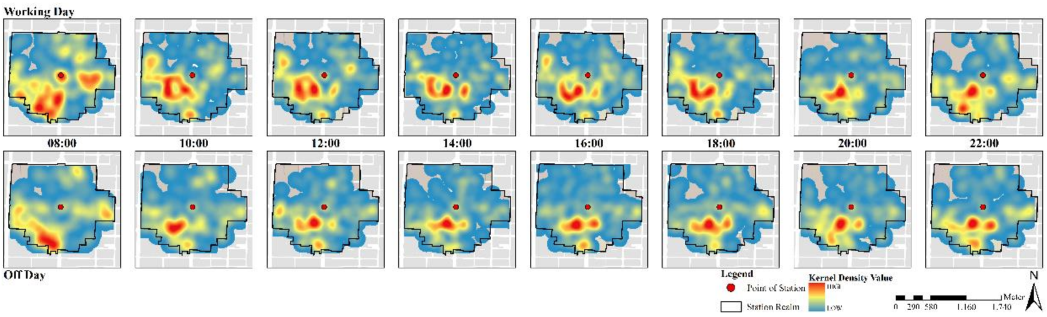
Figure 4- 37. Hot spot analysis of population activity within high population density-hub type station realm

Furthermore, considering the statistical results (Table 4- 25), the proportion of high/high z_G-score for both working day and off day in XZZX station realm and XZ station realm is nearly equal. However, in NSM station realm, the proportion of high/high clustering of population activities on off day is slightly higher than on working day.

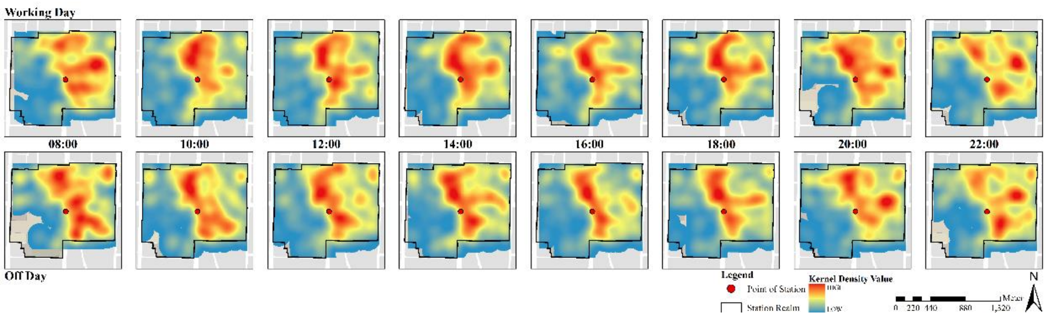
Table 4- 25. Statistics of the percentage of high/high z_G- score in high population density-hub type station realm

Code	Percentage of H/H z_G- score	
	Working day	Off day
XZZX	33.65%	33.46%
NSM	42.08%	43.10%
XZ	39.33%	39.12%

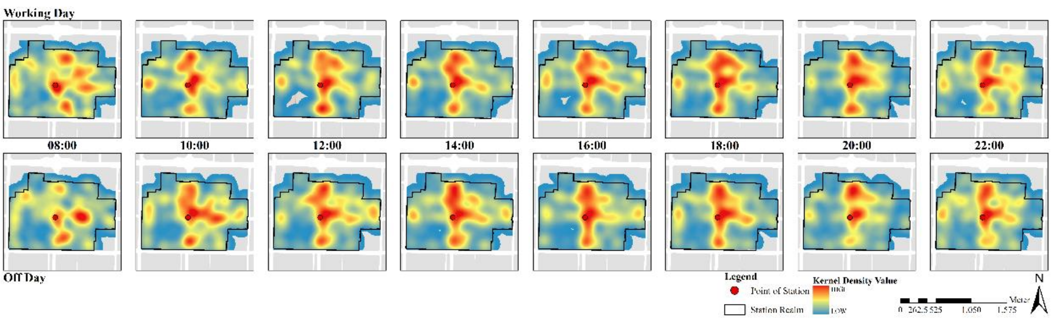
In addition, this study also conducted hotspot analysis on the spatial distribution characteristics of population activities in XZZX station realm, NSM station realm, and XZ station realm during working days and off day at different time intervals, as shown in Figure 4- 38. The distribution pattern of population activities in the station realms exhibits varying characteristics as the time changes.



(a) XZZX station realm



(b) NSM station realm



(c) XZ station realm

Figure 4- 38. Temporal and spatial distribution of population activities within high population density-hub type station realm

(2) Spatial distribution pattern of functional facility within station realm

The average nearest neighbor analysis results for various functional facilities within XZZX station realm, NSM station realm, and XZ station realm are shown in Table 4- 26. Among them, all types of facilities in XZ station realm exhibit a clustered distribution pattern, while public service facilities in XZZX station realm and NSM station realm show a dispersed distribution pattern. Additionally, shopping facilities have the highest level of clustering in these three station realms.

Table 4- 26. Average Nearest Neighbor Analysis of POI within high population density-hub type station realm

POI types	XZZX		NSM		XZ	
	Number	Z score	Number	Z score	Number	Z score
Entertainment	31	-1.23	110	-6.80	165	-13.11
Financial service	34	-3.43	61	-8.89	78	-11.56
Government agency	70	-9.23	95	-10.21	69	-8.33
Medical & Health care	26	-5.18	43	-4.17	65	-7.20
Accommodation service	40	-4.27	171	-14.05	202	-15.36
Catering service	302	-20.45	730	-31.78	1299	-43.73
Living service	163	-12.60	530	-26.41	960	-39.80
Shopping	389	-27.03	788	-34.84	1549	-52.59
Public service	14	2.74	22	0.69	39	-4.03
Transportation service	129	-11.12	169	-8.95	149	-7.02
Total	1198	-9.18	3643	-16.678	4575	-20.273

Based on the spatial distribution characteristics, as shown in Figure 4- 39, the areas of functional facility clustering within XZZX station realm, NSM station realm, and XZ station realm are generally consistent with the areas of population activity clustering. Functional facilities in XZZX station realm mainly cluster in the southwestern part of the station realm, while in NSM station realm, they are primarily distributed in the northern part of the station. In XZ station realm, functional facilities cluster around the station and in the northern area.

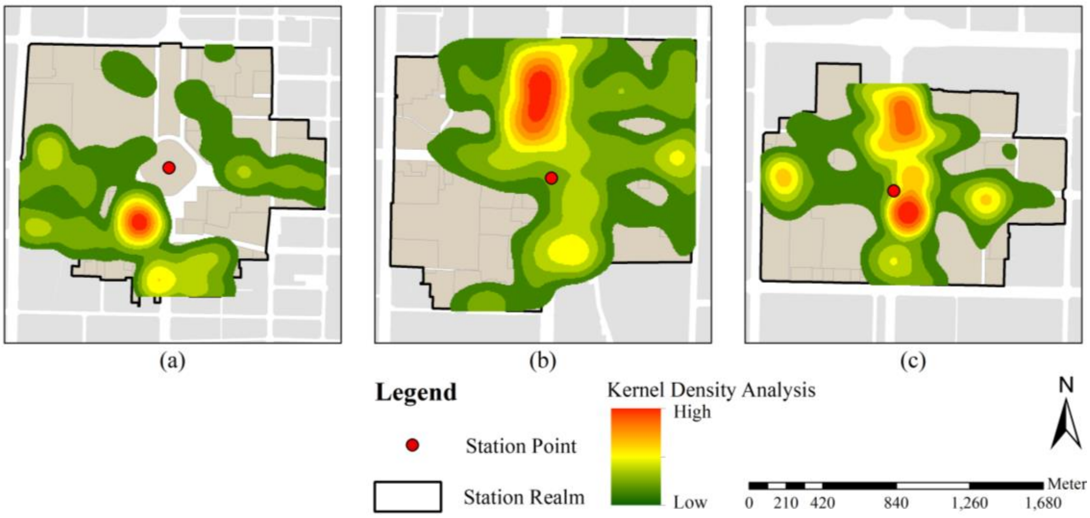


Figure 4- 39. Kernel density analysis of functional facilities within station realm. (a) XZZX station realm. (b) NSM station realm. (c) XZ station realm

4.5.5 Analysis of spatial distribution pattern within Interchange type station realm

(1) Spatial distribution pattern of population activities within station realms

In terms of time distribution, the statistics of population activity during working day in the station realms are shown in the Figure 4- 40. Generally, the morning peak of population activity occurs between 08:00 and 10:00 across all six station realms. However, the evening peak varies among them. For instance, in LSY station realm, the evening peak of population activity appears between 20:00 and 21:00, while in BDJ station realm, it occurs between 16:00 and 18:00. The other station realms generally experience their evening peaks between 17:00 and 19:00. It is worth noting that both FCWL station realm and BDJ station realm exhibit a population activity peak between 12:00 and 14:00.

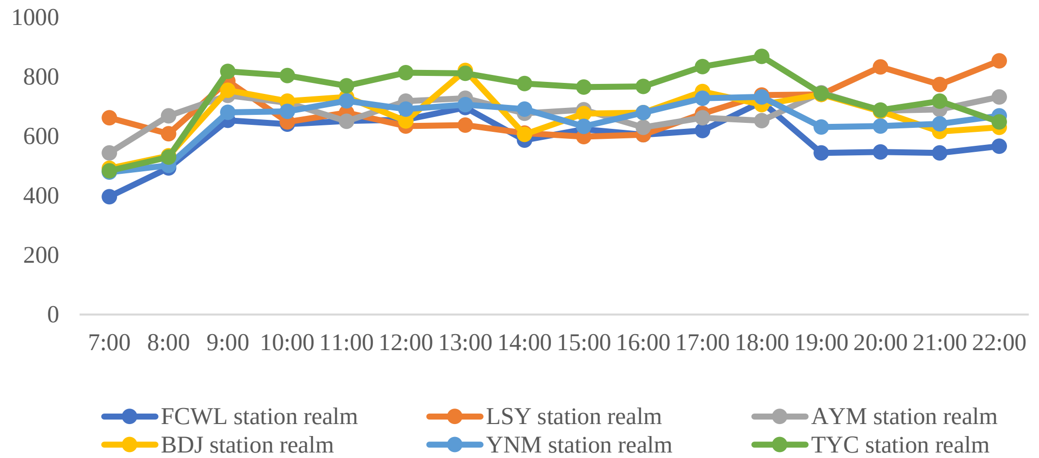


Figure 4- 40. Statistics on number of working day population activities within interchange type station realm

On off day, as shown in the Figure 4- 41, the first peak of population activity in all six station realms appears between 12:00 and 14:00, while the second peak mainly occurs between 17:00 and 19:00. In LSY station realm, the second peak occurs between 16:00 and 18:00. Overall, FCWL station realm, LSY station realm, AYM station realm, BDJ station realm, and YNM station realm have higher population activity levels on off day compared to working day.

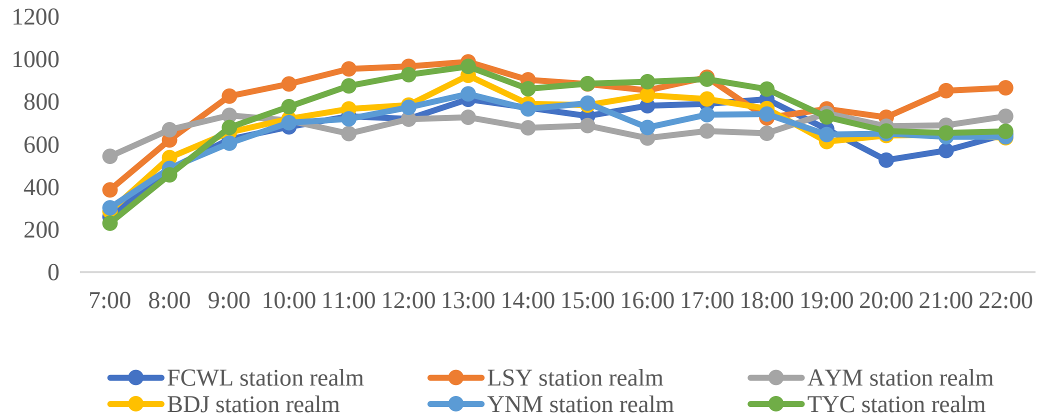


Figure 4- 41. Statistics on number of off day population activities within interchange type station realm

On spatial distribution, according to the results of the High/Low clustering analysis (Table 4- 27), the z_G- score for population activity in the six interchange-type station realms are positive for both working day and off day. This indicates that the spatial clustering type of population activity in these station realms is high/high clustering, meaning there is a high concentration of population activity and a high level of clustering of population activity.

Table 4- 27. High/low cluster analysis statistics of population activities within interchange type station realm

Code	Number of people		z_G- score	
	Working day	Off day	Working day	Off day
FCWL	8240	8974	36.32	36.58
LSY	9594	11278	27.24	26.23
AYM	7760	8088	33.74	32.38
BDJ	8808	8934	25.47	25.13
YNM	8679	8885	26.85	24.54
TYC	9434	9429	33.87	34.74

In terms of spatial distribution, as shown in Figure 4- 42 and Figure 4- 43, the population activity in XZZX station realm mainly clusters in the southwestern part of the station realm. NSM station realm exhibits a concentration of population activity in the northern and western parts of the station, while XZ station realm shows a significant concentration of population activity in the vicinity of the station and the northern part of the station realm.

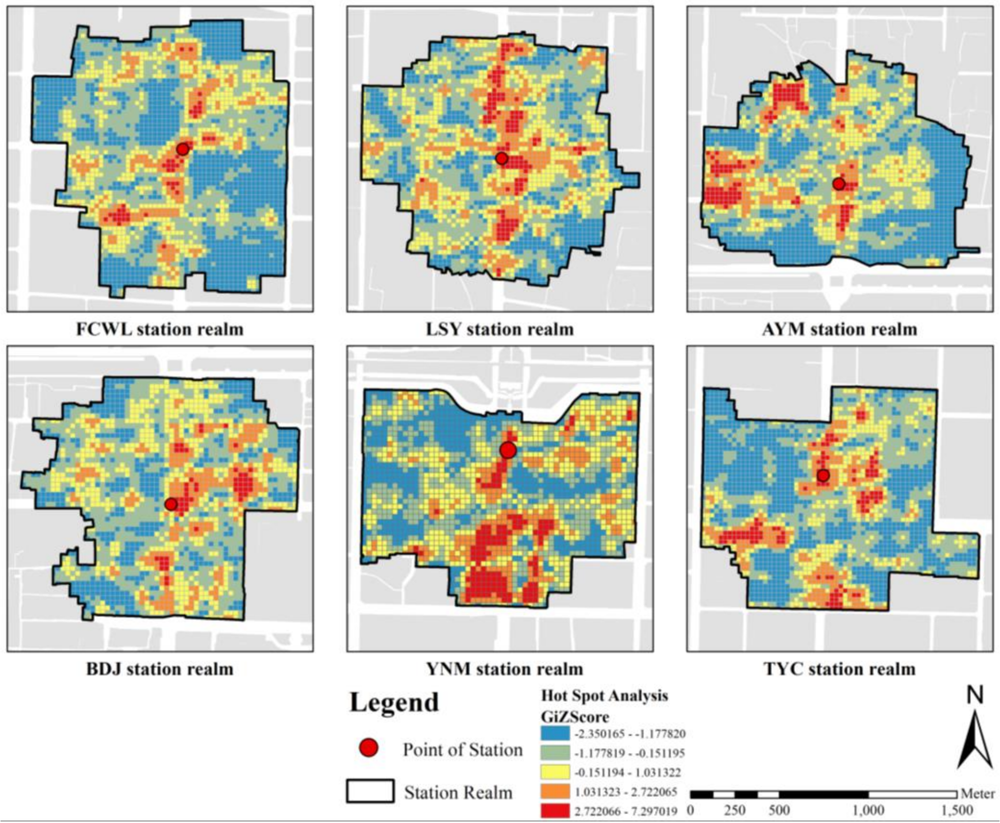


Figure 4- 42. Hot spot analysis of population activity within interchange type station realm on working day

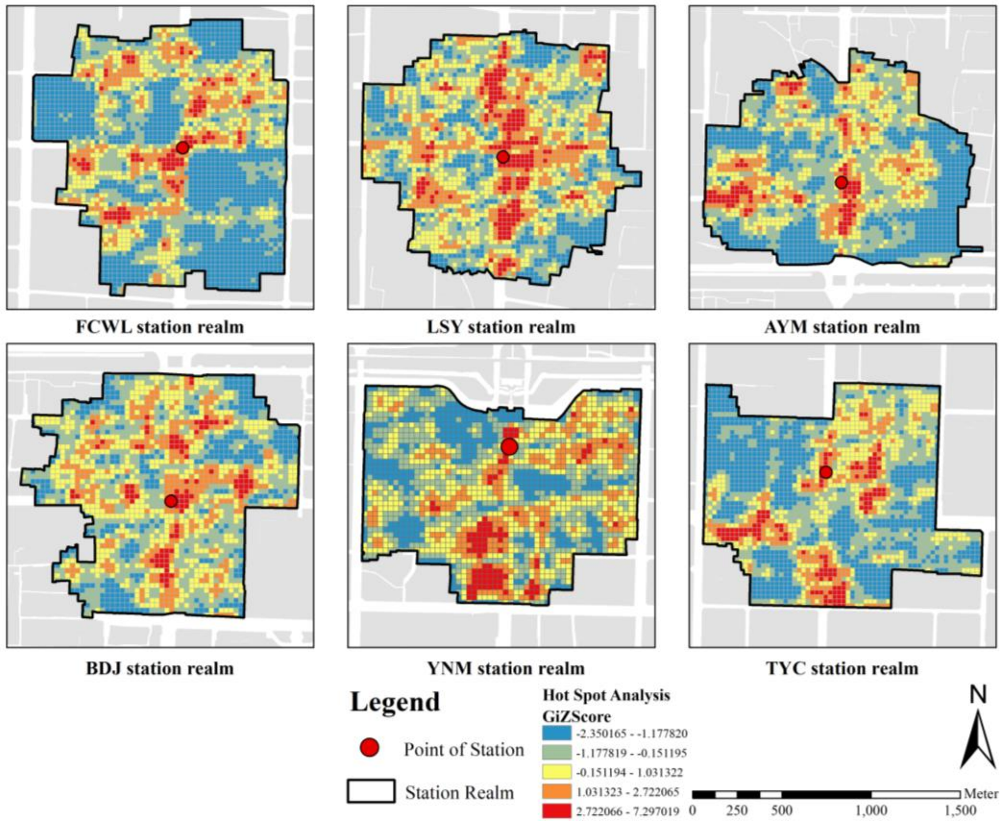


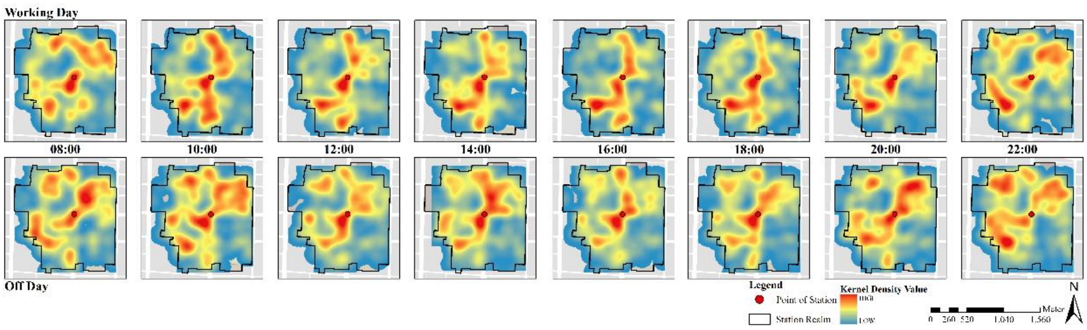
Figure 4- 43. Hot spot analysis of population activity within interchange type station realm on off day

Analyzing the proportion of high/high clustering of population activity in each station realm (Table 4- 28), it can be observed that on off day, all six station realms have a higher proportion of high/high clustering based on the z_G- score compared to working day. LSY station realm, BDJ station realm, and YNM station realm show a significant increase in the proportion of z_G- score for high/high clustering on off day, indicating a larger coverage of population activity clustering during off day in these three station realms compared to working day.

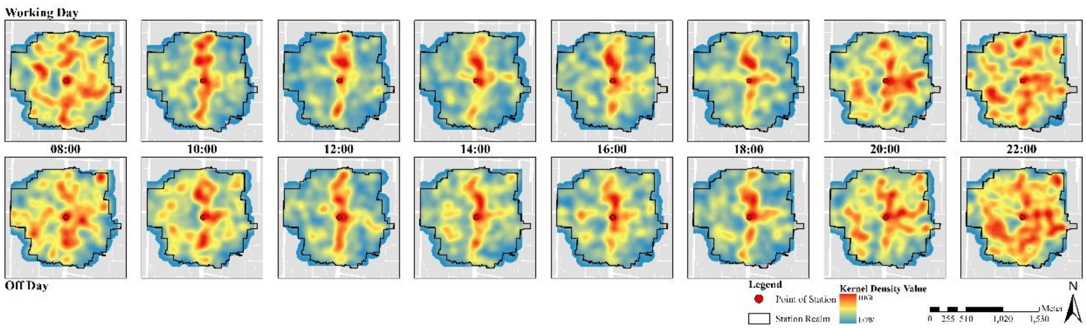
Table 4- 28. Statistics of the percentage of high/high z_G- score in interchange type station realm

Code	Percentage of H/H z_G- score	
	Working day	Off day
FCWL	39.26%	40.65%
LSY	40.59%	44.05%
AYM	40.50%	41.11%
BDJ	42.06%	45.35%
YNM	39.84%	45.59%
TYC	38.39%	39.36%

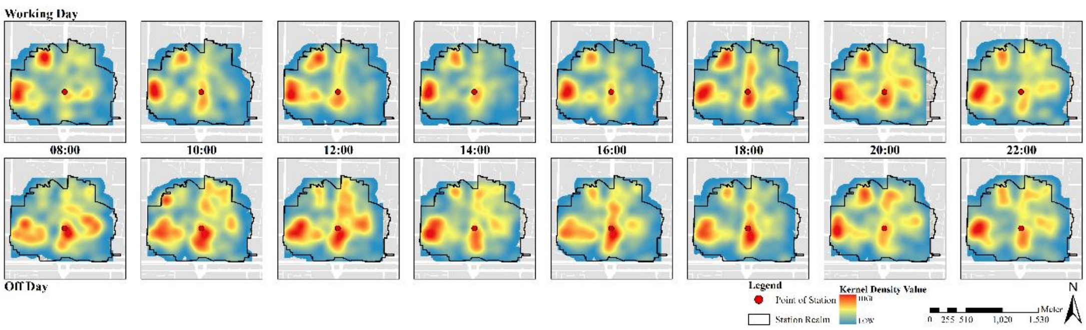
In addition, this study also conducted hotspot analysis on the spatial distribution of population activities in FCWL station realm, LSY station realm, AYM station realm, BDJ station realm, YNM station realm, and TYC station realm during working day and off day at different time intervals, as shown in Figure 4- 44. The distribution patterns and locations of population activities within each station realm exhibited varying characteristics with the changing time.



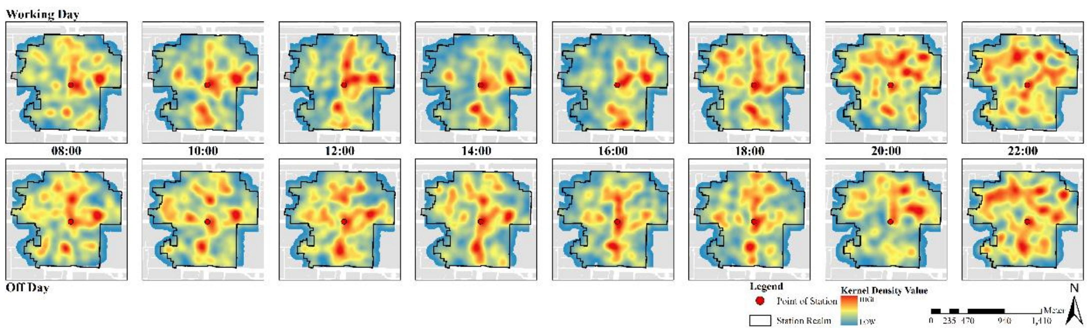
(a) FCWL station realm.



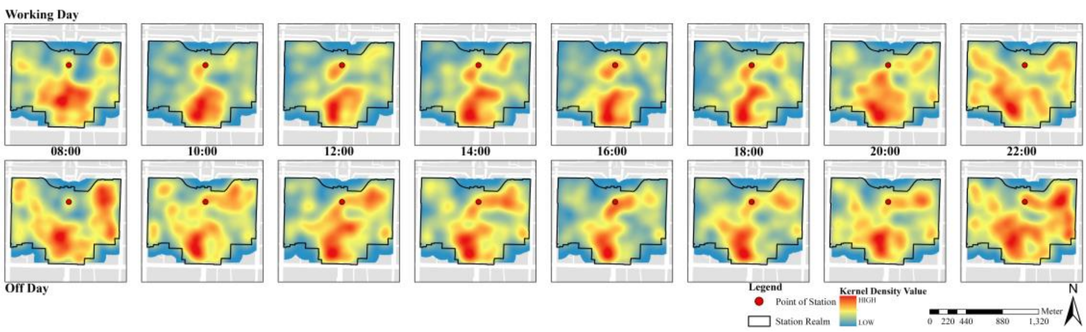
(b) LSY station realm



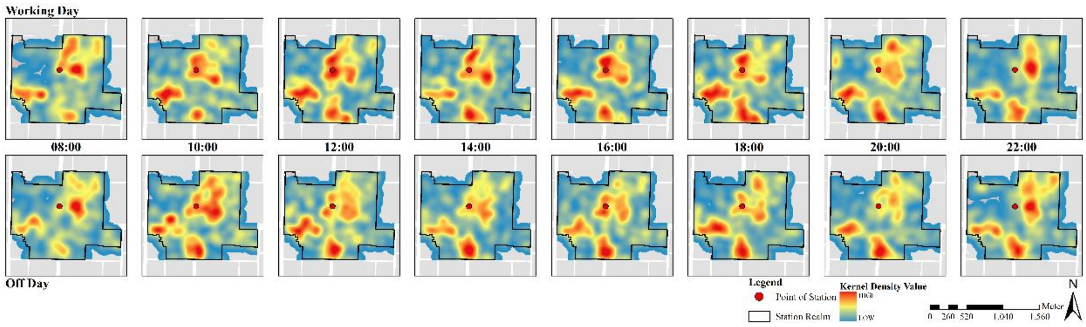
(c) AYM station realm



(d) BDJ station realm.



(e) YNM station realm



(f) TYC station realm

Figure 4- 44. Temporal and spatial distribution of population activities within interchange type station realm

(2) Spatial distribution pattern of functional facilities in the station realm

The average nearest neighbor analysis results for various functional facilities in FCWL station realm, LSY station realm, AYM station realm, BDJ station realm, YNM station realm, and TYC station realm are shown in Table 4- 29. It can be observed that functional facilities in FCWL station realm exhibit a clustered spatial distribution, while the public service facilities in the other station realms show a dispersed distribution. Additionally, YNM station realm has the highest level of clustering for dining service facilities, while shopping facilities exhibit the highest level of clustering in the other station realms.

Regarding spatial distribution characteristics, Figure 4- 45 visualizes the kernel density analysis results for functional facilities in interchange-type station realms. Functional facilities in FCWL station realm and LSY station realm exhibit a clustered distribution around the stations. In AYM station realm, functional facilities are mainly concentrated in the northern part of the station realm, while in BDJ station realm, YNM station realm, and TYC station realm, functional facilities are primarily concentrated in the southern part of the station realm.

Table 4- 29. Average Nearest Neighbor Analysis of POIs within interchange type station realm

POI types	FCWL	LSY	AYM	BDJ	YNM	TYC
Entertainment	-6.50	-3.67	-2.91	-0.03	-5.78	-6.80
Financial service	-10.65	-7.03	-7.04	-10.88	-14.87	-8.89
Government agency	-3.41	-13.77	-5.41	-14.21	-9.28	-10.21
Medical & Health care	-5.90	-6.82	-5.38	-14.59	-8.79	-4.17
Accommodation service	-5.32	-7.22	-7.73	-8.97	-14.41	-14.05
Catering service	-29.50	-30.84	-19.95	-30.63	-35.81	-31.78
Living service	-24.48	-26.78	-24.29	-24.63	-34.79	-26.41
Shopping	-43.17	-40.20	-26.04	-41.55	-34.45	-34.84
Public service	-0.73	2.10	0.96	0.94	0.69	0.69
Transportation service	-8.37	-12.43	-9.98	-7.95	-9.29	-8.95
Mean	-13.80	-14.67	-10.78	-15.25	-16.68	-14.54

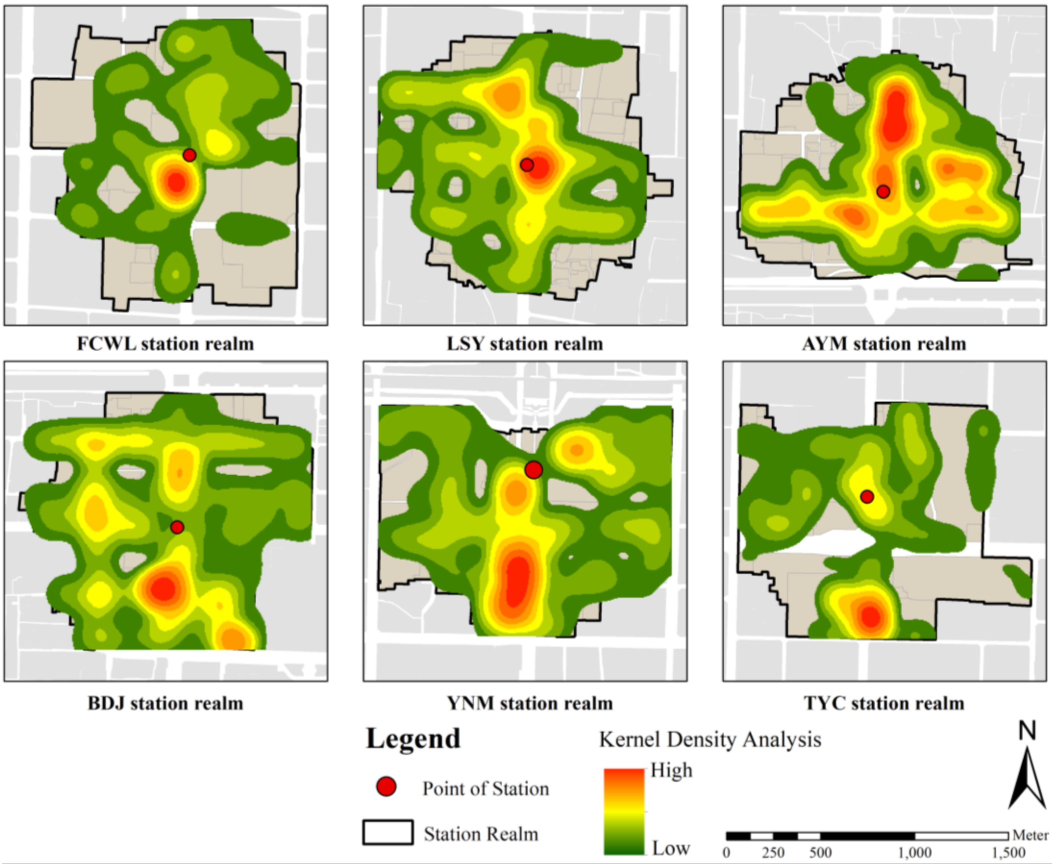


Figure 4- 45. Kernel density analysis of off day population activity within interchange type station realm

4.5.6 Analysis of spatial distribution pattern within General residential type station realm

General residential-type station realms generally include STSG station realm, DMGX station realm, WYJ station realm, and HZZX station realm.

(1) Spatial distribution pattern of population activities within station realms

In terms of time distribution, the population activities in these station realms during working day are shown in the Figure 4- 46. STSG station realm, WYJ station realm, and HZZX station realm exhibit clear commuting patterns, with morning peaks occurring between 08:00 and 10:00, and evening peaks between 17:00 and 19:00. DMGX station realm shows a more stable time distribution, with only an evening peak observed between 16:00 and 18:00.

On off day, as shown in the Figure 4- 47, STSG station realm experiences a second peak between 17:00 and 19:00. DMGX station realm has a peak in population activities only between 10:00 and 13:00. WYJ station realm shows a peak in population activities between 10:00 and 14:00. HZZX station realm experiences a peak in population activities between 16:00 and 18:00.

Overall, during the time period from 11:00 to 18:00, the population activities in STSG station realm, DMGX station realm, WYJ station realm, and HZZX station realm are higher on off day compared to working day.

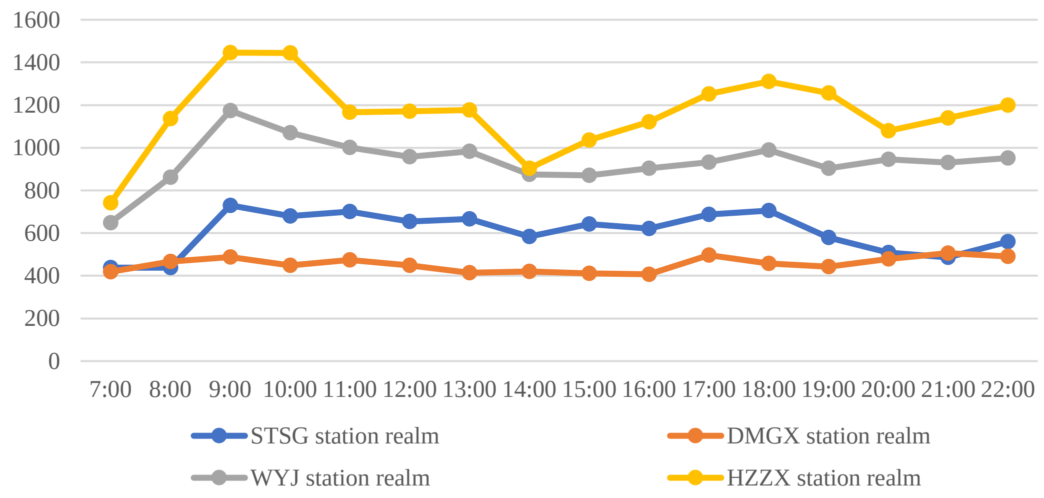


Figure 4- 46. Statistics on number of population activities within general residential type station realm on working day

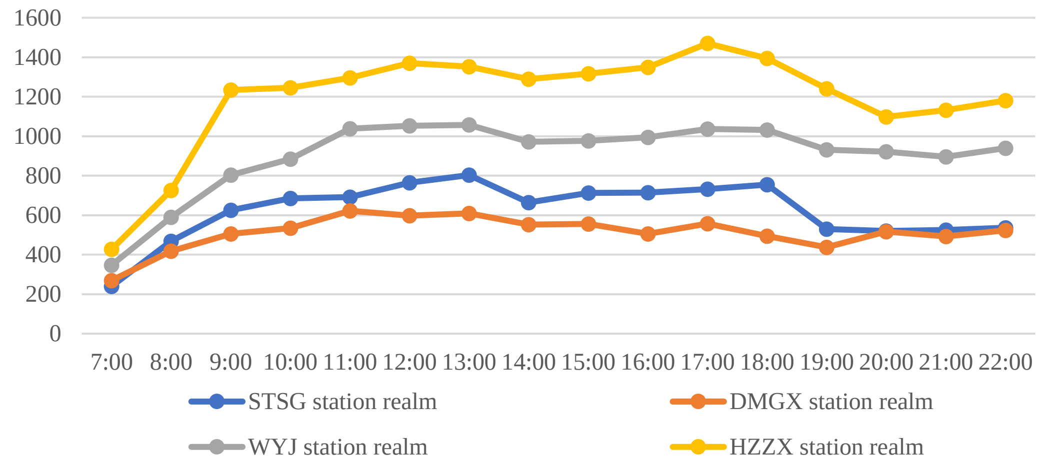


Figure 4- 47. Statistics on number of population activities within general residential type station realm on off day

According to the results of the High/Low Clustering Analysis (Table 4- 30), the z_G- score for population activities in the four general residential type station realms are positive for both working day and off day. This indicates that the spatial clustering type of population activities in these four station realms is high/high clustering, meaning a high number of population activities and their clustering.

Table 4- 30. High/low cluster analysis statistics of population activities within general residential type station realm

Code	Number of people		z_G- score	
	Working day	Off day	Working day	Off day
STSG	8176	8181	36.30	34.78
DMGX	6571	7195	31.62	30.50
WYJ	11257	11119	35.80	35.21
HZZX	11989	12498	43.16	42.69

Figure 4- 48 and Figure 4- 49 displays the visualization of the hotspot analysis results. It shows that population activities in the STSG station realm exhibit a clustered distribution around the station realm. In contrast, the distribution of population activities in the DMGX station realm and WYJ station realm is relatively dispersed in the clustered areas. However, in the HZZX station realm, there is a noticeable variation in the distribution of population activities. The population activities are mainly concentrated in the northeast part of the station realm, while the clustering level of population activities in the southeast part of the station realm is significantly lower.

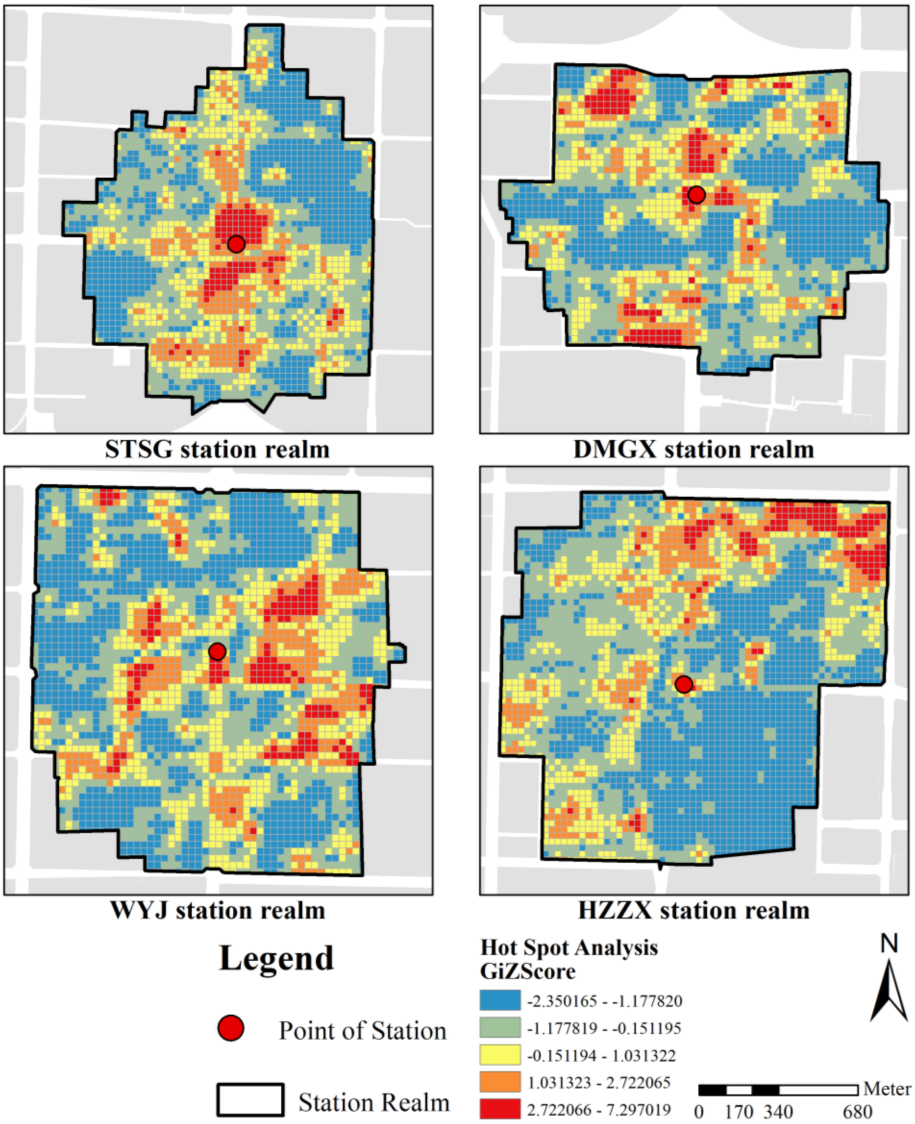


Figure 4- 48. Hot spot analysis of population activity within general residential type station realm on working day

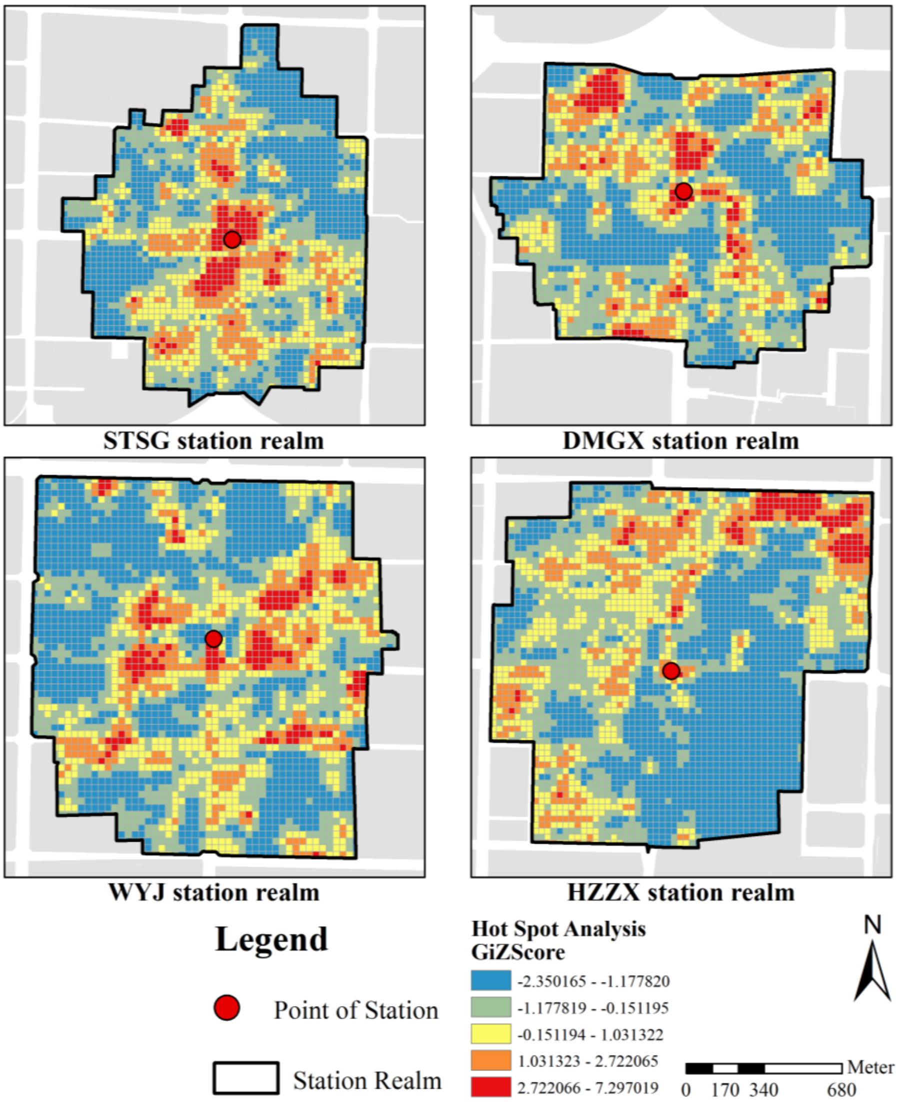


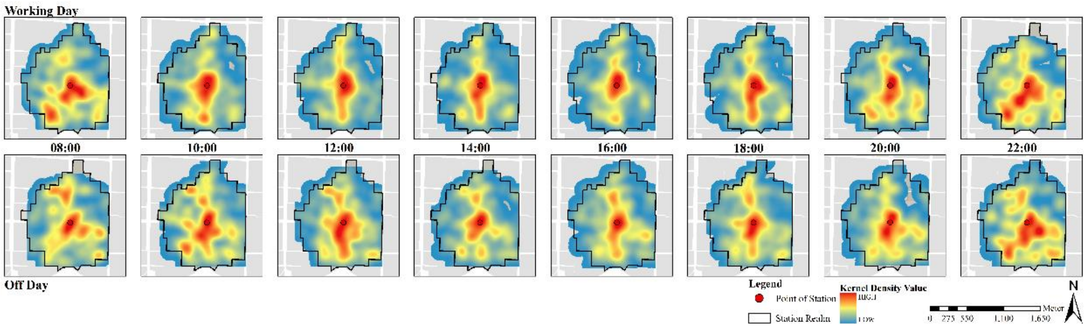
Figure 4- 49. Hot spot analysis of population activity within general residential type station realm on off day

According to the statistical results of the hotspot analysis (Table 4- 31), the proportion of high/high clustering z_G- score for population activities in the STSG station realm, WYJ station realm, and HZZX station realm is higher on off day than on working day. However, in the DMGX station realm, the proportion of z_G- score for population activities is lower on off day compared to working day. This indicates that the area covered by population activity clusters in the DMGX station realm decreases on off day.

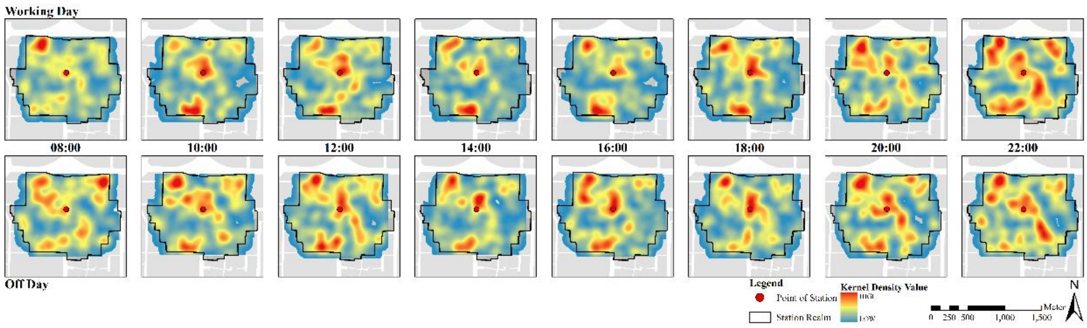
Table 4- 31. Statistics of the percentage of high/high z_G- score in general residential type station realm

Code	Percentage of H/H z_G- score	
	Working day	Off day
STSG	39.13%	41.02%
DMGX	41.93%	41.62%
WYJ	39.80%	40.89%
HZZX	39.49%	41.71%

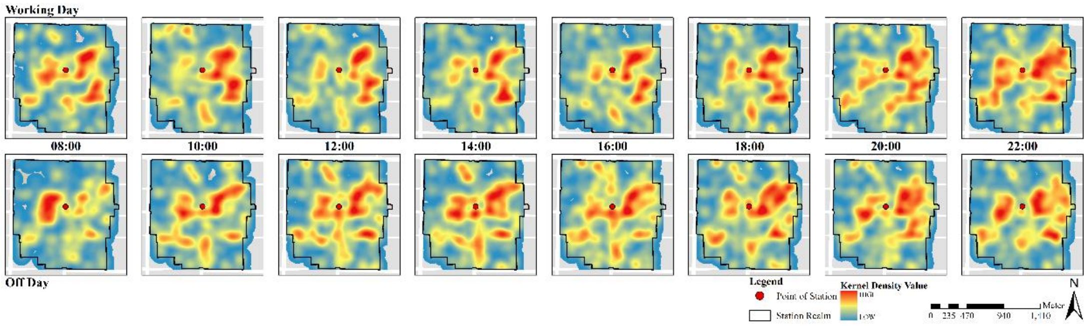
In addition, this study conducted hotspot analysis on the spatial distribution of population activities in the STSG station realm, DMGX station realm, WYJ station realm, and HZZX station realm during working day and off day, as shown in the Figure 4- 50. The distribution patterns and locations of population activities within each station realm exhibit varying characteristics over time.



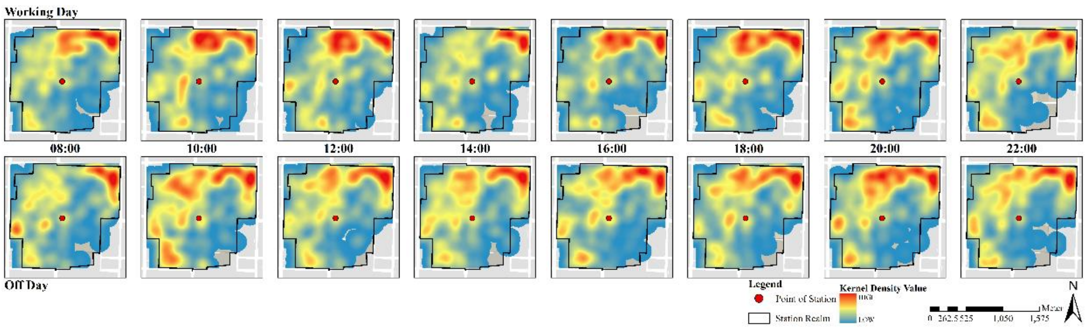
(a) STSG station realm



(b) DMGX station realm



(c) WYJ station realm



(d) HZZX station realm

Figure 4- 50. Temporal and spatial distribution of population activities within general residential type station realm

(2) Spatial distribution pattern of functional facility within station realm

As shown in Table 4- 32, the public service facilities in the STSG station realm, DMGX station realm, WYJ station realm, and HZZX station realm exhibit a dispersed distribution pattern. Furthermore, the catering service facilities show the highest level of agglomeration in the DMGX station realm, WYJ station realm, and HZZX station realm, while the STSG station realm demonstrates the highest level of agglomeration for shopping facilities in terms of spatial distribution.

Table 4- 32. Average Nearest Neighbor Analysis of POIs within general residential type station realm

POI types	STSG	DMGX	WYJ	HZZX
Entertainment	-4.28	-3.49	-5.06	-4.58
Financial service	-13.26	-1.51	-7.40	-5.01
Government agency	-5.55	-10.71	-10.57	-7.65
Medical & Health care	-5.40	0.21	-3.93	-8.34
Accommodation service	-12.00	-1.56	-8.71	-11.26
Catering service	-25.00	-19.98	-27.98	-29.23
Living service	-30.32	-12.51	-18.68	-24.81
Shopping	-33.46	-17.83	-25.97	-28.72
Public service	6.22	138.43	0.07	1.06
Transportation service	-11.78	-9.08	-5.00	-9.21
Mean	-4.28	-3.49	-5.06	-4.58

In terms of spatial distribution characteristics, as shown in the Figure 4- 51, the functional facilities in the STSG station realm exhibit a spatial distribution pattern that clusters around the station. In the DMGX station realm, the functional facilities primarily concentrate in the southern part of the station realm. In the WYJ station realm, the functional facilities mainly cluster in the western and southern parts of the station realm. As for the HZZX station realm, the functional facilities are primarily distributed in the western side of the station realm, with the agglomeration area located in the northwestern part of the station realm.

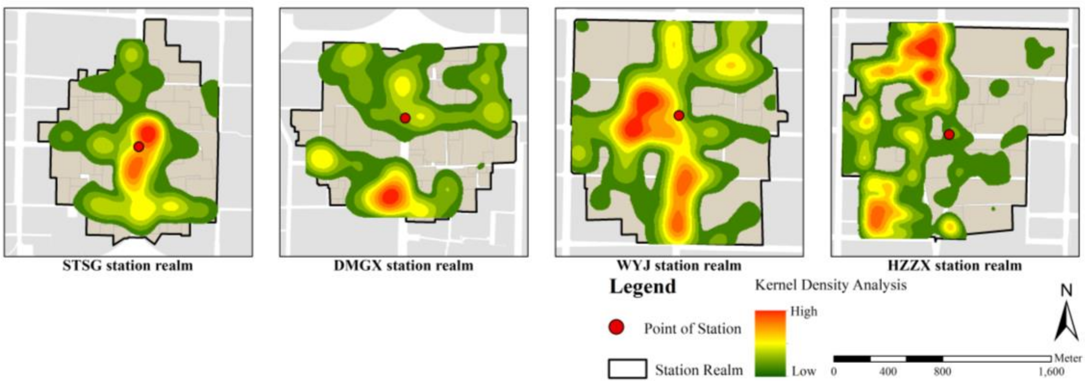


Figure 4- 51. Kernel density analysis of off day population activity within general residential type station realm

4.6 Discussion

4.6.1 The overall spatial distribution pattern of the station realm of Xi'an Metro Line 2

The time distribution of population activity in the stations of Line 2 of the Xi'an subway, as shown in Figure 4- 52 and Figure 4- 53, reveals distinct commuting patterns on working day. The morning peak occurs between 08:00 and 10:00, while the evening peak is concentrated between 17:00 and 19:00. On off day, there is no clear morning peak in population activity within Line 2 station realms. Instead, activity fluctuates continuously between 09:00 and 18:00, with higher population activity levels observed on off day. This indicates a preference for activities during the 09:00-18:00 time period on off day. Additionally, the overall population activity on Line 2 of the Xi'an subway exhibits an increasing trend from north to south throughout the day.

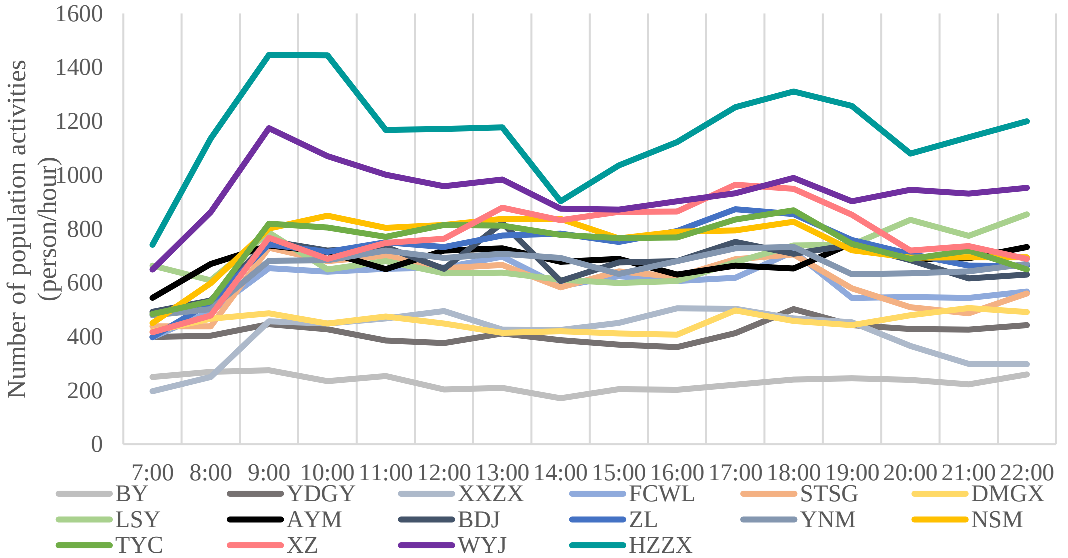


Figure 4- 52. Temporal distribution of working day population activities in each station realm

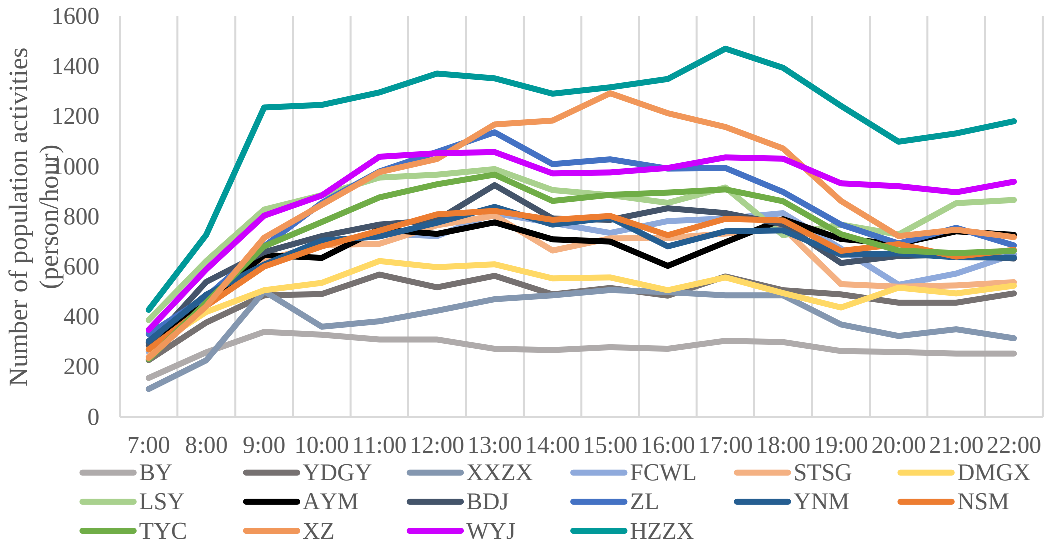


Figure 4- 53. Temporal distribution of off day population activities in each station realm

Based on the global spatial autocorrelation analysis results, the population activity in the Line 2 station realms is characterized by high/high clustering, indicating a high level of population activity and spatial clustering of population activity on both working day and off day. The visualization of local spatial autocorrelation analysis (hotspot analysis) results shows significant spatial variations in population activity distribution within the Line 2 station realms, particularly in the XXZX, NSM, and HZZX station realms, where high/high clustering and low/low clustering phenomena are evident. Furthermore, the statistical results indicate that the degree of population activity clustering and the proportion of high/high clustering are lower on working days compared to off day. This suggests that on working days, population activity is more concentrated spatially, while on off day, population activity is more widely dispersed across the station realms.

In terms of the spatial distribution pattern of functional facilities, as shown in Table 4- 33, the top three facilities in terms of quantity are shopping, catering service, and living service. Overall, various types of functional facilities in Line 2 subway station realms exhibit a clustered distribution pattern, with shopping facilities showing the highest level of clustering, followed by catering service and living service.

Table 4- 33. Statistics on the number and concentration of top three functional facilities

POI types	Number	Average Nearest Neighbor (*Z*-score)
Catering service	9258	-121.786
Living service	7772	-106.314
Shopping	11475	-143.426

Based on the overall kernel density analysis results of functional facilities in the station realms, there is a spatial distribution pattern of "one line and six cores" within Line 2 subway station realms. Specifically, functional facilities are primarily clustered between the FCWL, STSG, LSY, ZL, YNM, and NSM station realms, with XZ station realm playing a role as well. The kernel density analysis results for different categories of functional facilities indicate that the ZL station realm is the central area of Line 2 subway, where, except for Medical & Health care facilities, other functional facilities exhibit a high level of clustering. Entertainment facilities, accommodation services, and public services show the highest level of clustering within the ZL station realm. Entertainment facilities only exhibit a clustered distribution in the ZL and XZ station realms. Additionally, compared to other types of facilities, transportation services are relatively evenly distributed throughout the Line 2 station realms.

4.6.2 Spatial distribution pattern of each type of station realm

In terms of the time distribution of population activities, the population activities in each station realm generally exhibit the commuting characteristics of peak hours during working day, and the population activity levels during 10:00-18:00 on off day are higher than on working day. In terms of spatial distribution, the population activities in various station realms show a clustered distribution pattern during both working day and off day, with a high/high clustering type, indicating a high level of population activity and clustering.

Regarding the spatial distribution of functional facilities, although there are significant differences in the spatial distribution characteristics of different types of facilities in different station realms, there are still some commonalities. Except for the ZL station realm, public service facilities are dispersed in the other 15 station realms, while dining services, daily necessities services, and shopping facilities are among the top three in terms of their clustering levels in each station realm.

In terms of the distribution of population activities, as shown in the Figure 4- 54, there are significant differences in the clustering positions between station realms. The spatial distribution patterns of most station realms do not exhibit the geographical characteristics proposed by many studies, which describe a concentric distribution pattern around the stations [10,11]. This indicates that there is some deviation between theoretical expectations and the actual development of station realms. However, this study also found that in the Xi'an Metro Line 2, the population activities, and functional facilities in eight station realms are concentrated within a radius of 500 meters from the stations. These station realms include the Urban center type ZL station realm, the High population density hub type XZ station realm, the Interchange type FCWL station realm, LSY station realm, AYM station realm, BDJ station realm, and the General residential type STSG station realm, WYJ station realm. This to some extent reflects the agglomeration effect of stations on the surrounding environment, which is consistent with some related research conclusions [12,13]. Additionally, through comparison, it is observed that there is a certain spatial correspondence between the clustering positions of population activities and functional facilities in each station realm. However, further analysis is needed to understand the spatial relationship and interaction between the two.

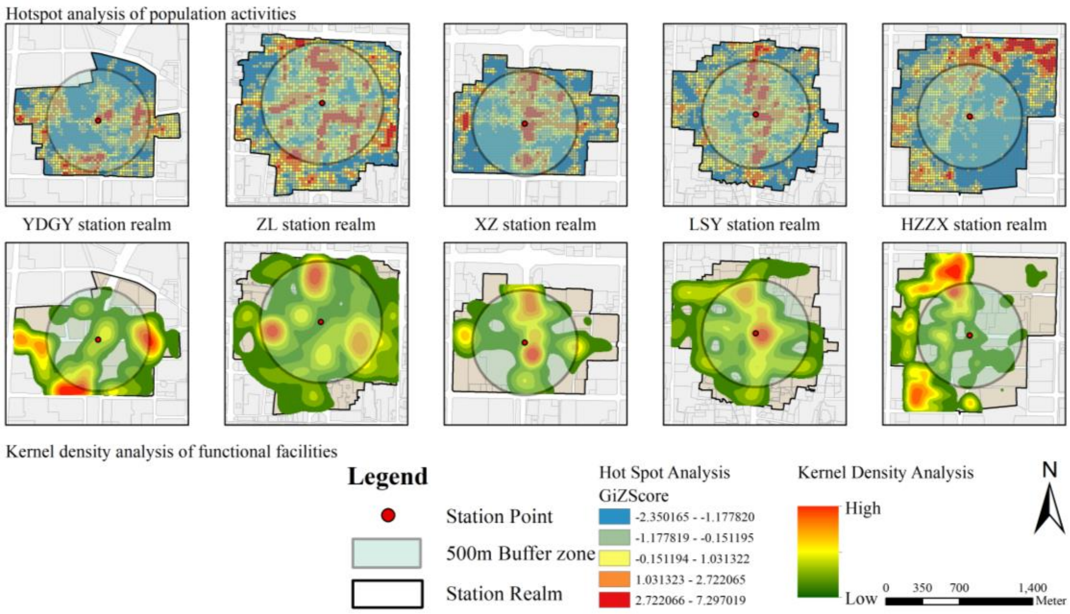


Figure 4- 54. Spatial correspondence between population activities and functional facilities within station realm of Xi’an Metro Line 2

4.6.3 Spatial Correspondence of Station Realm

Based on the analysis of the spatial distribution patterns of station realms in Xi'an Metro Line 2, as shown in Figure 4- 55, the distribution areas of population activities and the clustering areas of functional facilities in the station realms are generally consistent. This indicates a clear spatial correspondence between high population activity clusters and areas with a high concentration of functional facilities. In other words, areas with a high concentration of population activities tend to coincide with places where functional facilities are highly clustered.

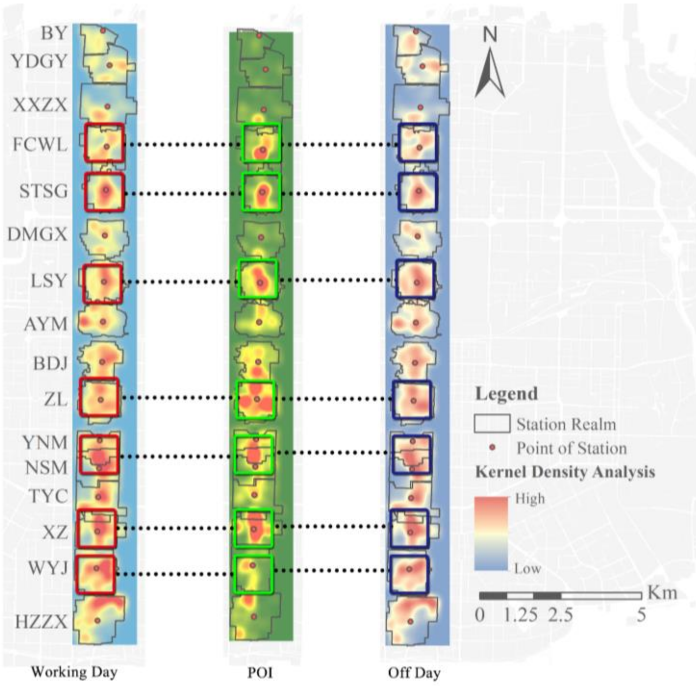


Figure 4- 55. Spatial location relationship between population activities and functional facilities in the station realm of metro line 2

From a local perspective, as shown in Figure 4- 56, the distribution areas of population activity clusters and the clustering areas of functional facilities also correspond to each other in the spatial context of each type of station realm. However, there are cases where the distribution areas of population clusters do not match the areas where functional facilities are clustered. For example, in the HZZX station realm, population activities are clustered in the northeast, but the results of kernel density analysis show that the clustering area of functional facilities is in the northwest of the station realm. This suggests the presence of a certain category of functional facility in the HZZX station realm that strongly attracts population activities or has a high degree of interaction with them. In the subsequent research, spatial relationship models will be used to further analyze and evaluate this phenomenon.

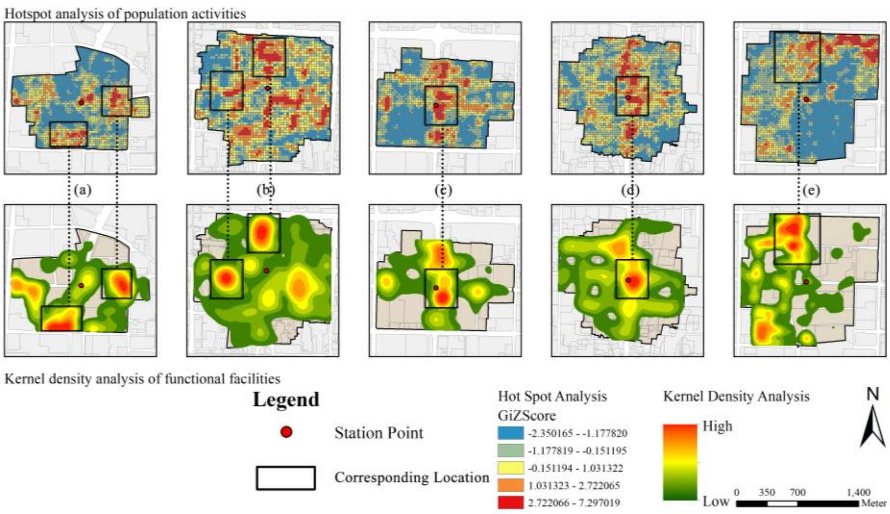


Figure 4- 56. Spatial location relationship between population activities and functional facilities in each type of station realm

4.7 Chapter summary

In this chapter, the spatial distribution patterns of the overall station realm of Xi'an Metro Line 2 and its 16 station realms were studied based on field investigations and big data analysis techniques. Firstly, the scope and types of station realms were defined through field investigations and statistical analysis. Secondly, data collection and integration analysis were conducted from two dimensions: "spatiotemporal behavior of population activities" and "functional facilities." This allowed for the exploration of the spatiotemporal distribution characteristics of population activities and the distribution characteristics of functional facilities in the overall station realm of Xi'an Metro Line 2 and different station realms. The study investigated the support and guidance patterns of population activities on functional facilities, and preliminary correspondences were identified between the spatiotemporal distribution of population activities and the distribution characteristics of various types of functional facilities within different station realms. These findings provide data support for the subsequent analysis of the relationship between population activities and functional facilities.

Reference

[1] Xi’an metro. Available online: https://en.wikipedia.org/wiki/Xi%27an_Metro (accessed on 5 October 2022).

[2] Duan D, Zhang F. Study on classification of urban rail transit stations from the perspective of land use optimization: A case study on Xi’an subway line 2. City Planning Review 2013, 37(9), 39-45.

[3] Official account of Xi'an Evening News. https: // bai jia hao. baidu.com/ s?id= 1632431736440353328 & wfr = spider & for= pc (accessed on 5 October 2022).

[4] Li Bo. Analysis to land use optimization of the rail transit station region based on two-way passenger flow balance guidance in Xi’an city. Master’s thesis. Chang’an University. 2020.

[5] Li Weijia. Study on Land Use Optimization of Xi’an Rail Transit Stations under the Influence of Time-varying Passenger Flow. Master’s thesis. Chang’an University. 2020.

[6] Xi'an Metro Official Weibo Account. https://weibo.com/xianditie (accessed on 14 January 2023).

[7] Xi'anbei Railway Station. https:// zh.wikipedia.org/ wiki/ Xi'anbei Railway Station (accessed on 20 September 2022).

[8] Xin Tingye, Zhang Haifeng, Bai Chenglong, Liu Mengxuan. Analysis on Social and Economic Correlative Factors of Land Use Change in Xining City [J]. Natural Resource Economics of China, 2017, 30(03): 48-51.

[9] Xu Wei, Zhang Changjiang, Ma Genghua, Li Rui, Deng Pingxin. Urban Rail Transit Site

Classification Based on K-means Clustering [J]. Journal of Guizhou University (Natural Sciences), 2018, 35 (06): 106-111. DOI: 10.15958/j.cnki.gdxbzrb.2018.06.18.

[10] Huang H. The land-use impacts of urban rail transit systems[J]. Journal of Planning Literature, 1996, 11(1): 17-30.

[11] Shi Y, Hu X Y, Yang J Y. Research of station zone spatial circle and internal mechanism in TOD pattern: Taking Nanjing Metro Line 2 Jiqingmen Station as example[C]//Applied Mechanics and Materials. Trans Tech Publications Ltd, 2013, 409: 861-866.

[12] Higgins C D, Ferguson M R, Kanaroglou P S. Light rail and land use change: Rail transit's role in reshaping and revitalizing cities[J]. Journal of Public Transportation, 2014, 17(2): 93-112.

[13] Bao X. Urban rail transit present situation and future development trends in China: Overall analysis based on national policies and strategic plans in 2016–2020[J]. Urban Rail Transit, 2018, 4(1): 1-12.

Chapter 5. Correlation analysis of population activity and functional facility within station realm

Content

Introduction

In this chapter, the study will analyze the correlation between population activities and functional facilities in the station realm on both working day and off day. The aim is to explore the degree and variations of the correlation between population activities and various types of functional facilities. Specifically, functional facilities that exhibit a significant positive correlation will be selected as explanatory variables, in preparation for the establishment of the spatial relationship model in the subsequent steps. The research framework for this chapter is illustrated in Figure 5- 1.

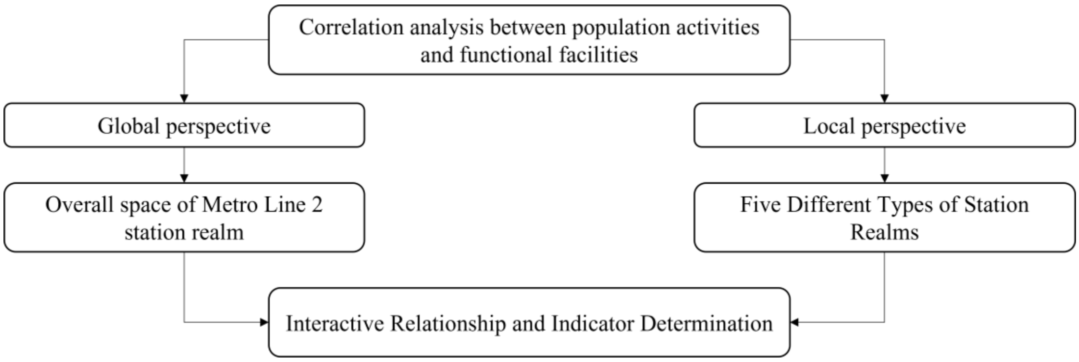


Figure 5- 1. Research framework

5.1 Methodology

Although the analysis of the spatial distribution patterns of population activities and functional facilities in the station realms of Xi'an Metro Line 2 in the previous chapter suggests a certain spatial correspondence, the specific quantitative relationship needs to be determined through correlation analysis. In this chapter, the Spearman correlation analysis method, using the Statistical Package for the Social Sciences (SPSS) [1], will be employed to calculate the correlation matrix and determine the degree of correlation between population activities and various types of functional facilities within different station realms.

5.1.1 Selection of experimental subjects

The selection criteria for the experimental objects should meet the following two conditions:

1. Representative and Universal Selection:

The selection of experiment objects should primarily consider station realms with representative spatial patterns. Otherwise, selecting objects with highly similar features would not provide intuitive characteristics and significance for comparison. Additionally, the selected objects must cover all types of station realms along the Xi'an Metro Line 2 to ensure more convincing analysis results.

2. Different Urban Segments:

The selected objects should be located in different urban segments, taking into account various sections along Metro Line 2 in the city. It is not advisable to choose consecutive or concentrated station realms. This approach aims to sample different urban segments and achieve more distinct results in correlation analysis.

In summary, following the criteria, this study selects the following station realms as the experimental objects: YunDongGongYuan Station, ShiTuShuGuan station, Longshouyuan Station, BeiDaJie station, ZhongLou Station, NanShaoMen station, Xiaozhai Station, and WeiYiJie Station. The specific information of each experimental subjects is shown in Table 5- 1.

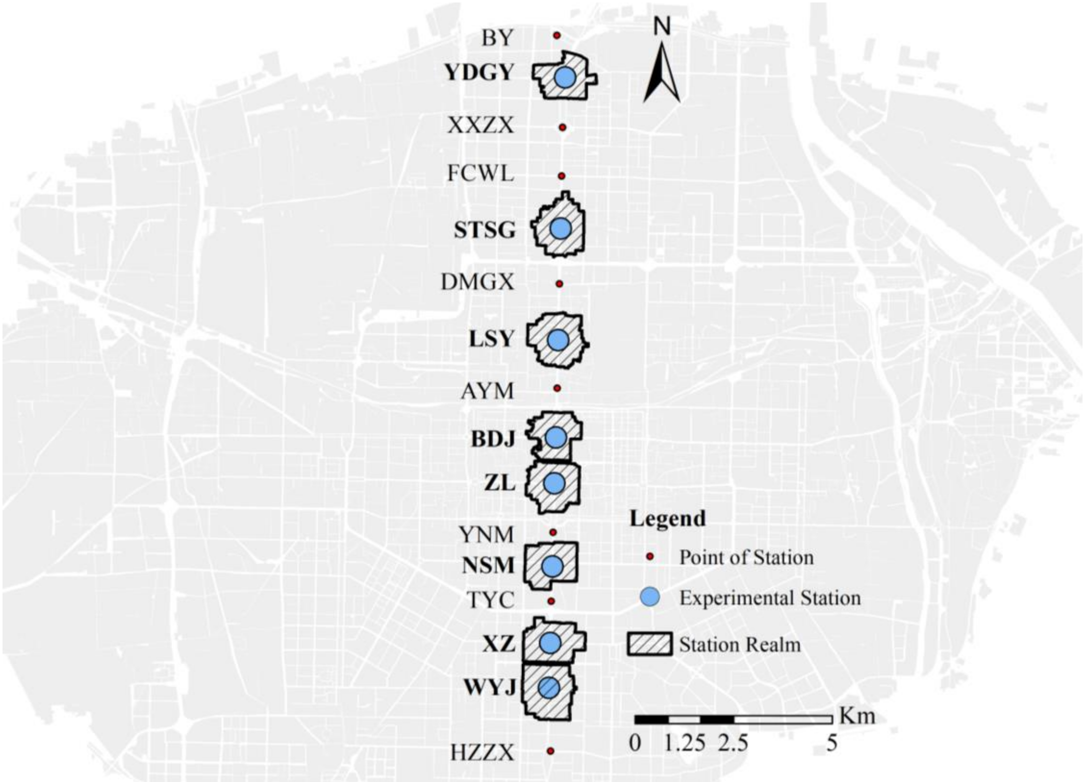


Figure 5- 2. The location of experimental station realms

Table 5- 1. Experimental subjects of Xi'an Metro Line 2

Code	Station name	Station realm type
M02	YunDongGongYuan (YDGY)	Under development type
M10	ZhongLou (ZL)	Urban center type
M12	NanShaoMen (NSM)	High population density Hub type
M14	XiaoZhai (XZ)	
M07	LongShouYuan (LSY)	Interchange type
M09	BeiDaJie (BDJ)	
M05	ShiTuShuGuan (STSG)	General residential type
M15	WeiYiJie (WYJ)	

5.1.2 Variable selection

Before conducting the correlation analysis between population activities and functional facilities in this study, it is necessary to provide a conceptual definition and establish the value range for the results of the Spearman correlation analysis model. The correlation coefficient, denoted as r_s, represents the correlation between variables in the research sample and indicates the magnitude of the correlation ($-1 \leqslant r_s \leqslant 1$) [2]. A value closer to 1 indicates a stronger positive correlation between the variables, while a value closer to -1 indicates a stronger negative correlation. Based on the reference to relevant research literature [3,4], this study adopts a significance level of $p < 0.05$ as the criterion for determining significance. The p-value is a test statistic used to determine the significance of the correlation between variables. A p-value < 0.05 or $0.05 < p < 0.1$ (a relatively lenient significance level) indicates a significant relationship between the two variables [5]. In addition, for the convenience of data statistics and analysis, this study designates the quantity of population activity on working day and off day as WD and OD, respectively. Various types of functional facilities are denoted as F01 to F10. The specific details as show in Table 5- 2 and Table 5- 3.

Table 5- 2. The code for quantity of population activity on working day and off day

Date	Code
Working day	WD
Off day	OD

Table 5- 3. The code for POIs

POI types	Code
Entertainment	F01
Financial Service	F02
Government Agency	F03
Medical & Health Care	F04
Accommodation Service	F05
Catering Service	F06
Living Service	F07
Shopping	F08
Public Service	F09
Transportation Service	F10

5.1.3 Processing of data

The specific data processing procedures for population activity and functional facility data in this study are as follows:

(1) Constructing database

Using the ArcGIS platform, a spatial dataset is created for each experimental station realm, and EasyGo point data representing population activity and POI point data representing functional facilities are imported separately. The total population activity on working days and off day within each experimental station realm is calculated, along with the quantity of various types of functional facilities. The data is then stored in the data set in Shp format.

(2) Build fishing network

Using the boundary of each experimental station realm, a rectangular grid with a cell size of 25 m * 25 m (corresponding to the sampling scale of EasyGo point data [6]) is created using the fishnet tool.

(3) Establishing spatial connections

The spatial join tool in ArcGIS is utilized to match the attribute values of connected features with target features based on their relative spatial positions. In this study, the previously calculated population activity data and functional facility data are connected to the established grid based on the station realm type. This ensures that each ID field in the grid contains attribute values for population activity and the quantity of various types of functional facilities.

(4) Exporting data

The spatially connected grid data is exported as a dBASE format table file.

(5) Conducting correlation analysis

The exported table files for each research station realm are imported into the SPSS software for Spearman correlation analysis between the quantity of population activity and functional facilities within each station realm.

5.2 Correlation analysis of population activities and functional Facilities within Metro Line 2 station realm

5.2.1 Results of correlation analysis between population activities and functional facilities

Based on the correlation matrix analysis in Table 5- 4, it can be observed that there is a positive correlation between the population activity in station realms of the Xi'an Metro Line 2 and the functionality facilities on both working days and off days, with a *P*-value < 0.05.

This indicates a significant positive relationship between population activity and functionality facilities. Moreover, the positive correlation coefficient on off days is slightly higher than that on working days.

Table 5- 4. Statistics of correlation coefficients between population activities and functional facilities in Metro Line 2 station realm

		WD	OD
POI	Correlation coefficient	0.243**	0.246**
	P-value	0.000*	0.000*

Notes:
Asterisks next to numbers indicate statistically significant (p < 0.05).
** Represents significant correlation.
* Represents value less than 10-6.

5.2.2 Results of correlation analysis between population activities and various functional facilities

According to the results of correlation analysis in Table 2-1, it can be observed that there is a positive correlation between population activity in station realms of the Xi'an Metro Line 2 and various types of functionality facilities on both working days and off days, with a *P*-value < 0.05. This indicates a significant positive relationship between population activity and functionality facilities.

Specifically, on both working days and off days, the highest correlation with population activity is observed for catering service facilities, followed by living service facilities, while government agencies exhibit the lowest positive correlation with population activity.

When comparing across working days and off days, it is found that the positive correlation coefficients between population activity and catering service facilities, living service facilities, and shopping facilities are higher on off days compared to working days. However, there is a slight decrease in the positive correlation coefficients for financial service facilities, government agencies, and transportation service facilities on off days. The remaining four categories of service facilities show similar correlation coefficients on both working days and off days.

Table 5- 5. Statistics of correlation coefficients between population activities and various functional facilities in Metro Line 2 station realm

Code	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
F01	0.080**	0.000	0.080**	0.000
F02	0.074**	0.000	0.067**	0.000
F03	0.033**	0.000	0.017**	0.001
F04	0.085**	0.000	0.085**	0.000
F05	0.109**	0.000	0.110**	0.000
F06	0.178**	0.000	0.188**	0.000
F07	0.170**	0.000	0.177**	0.000
F08	0.167**	0.000	0.173**	0.000
F09	0.035**	0.000	0.034**	0.000
F10	0.100**	0.000	0.095**	0.000

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

5.3 Correlation analysis of population activities and functional facilities within different types of station realm

5.3.1 Results of correlation analysis between population activities and functional facilities in different types of station realm

Based on the correlation analysis results shown in Table 5- 6, it can be observed that there is a significant positive correlation (P-value < 0.05) between population activity and functional facilities in all the tested station realms during both working days and off days.

In terms of station realm types, the XZ station realm, characterized by a high population density and hub, exhibits the highest positive correlation coefficient between population activity and functionality facilities. On the other hand, the YDGY station realm, which is still under development, shows the lowest correlation coefficient.

Furthermore, except for the YDGY station realm and STSG station realm, the correlation coefficients between population activity and functional facilities in other station realms increase on off days compared to working days.

Table 5- 6. Statistics of correlation coefficients between population activities and various functional facilities in different types of station realm

Code	WD		OD	
	Correlation coefficient	*P*-value	Correlation coefficient	*P*-value
YDGY	0.156**	0.000*	0.143**	0.000*
ZL	0.220**	0.000*	0.242**	0.000*
NSM	0.284**	0.000*	0.292**	0.000*
XZ	0.339**	0.000*	0.343**	0.000*
LSY	0.226**	0.000*	0.235**	0.000*
BDJ	0.195**	0.000*	0.209**	0.000*
STSG	0.325**	0.000*	0.291**	0.000*
WYJ	0.181**	0.000*	0.203**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

5.3.2 Results of correlation analysis between population activities and various functional facilities in different types of station realm

(1) YDGY station realm

From the correlation matrix analysis in Table 5- 7, it can be observed that the YDGY station realm, there is a significant positive correlation between population activity and four types of functional facilities: catering services, living services, shopping services, and transportation services. Among these, transportation services exhibit the highest correlation coefficient with population activity. The correlation coefficients between living services and population activity, as well as shopping services and population activity, increase on off days compared to working days.

Table 5- 7. Statistics of correlation coefficients between population activities and various functional facilities in YDGY station realm

Code	WD		OD	
	Correlation coefficient	*P*-value	Correlation coefficient	*P*-value
F01	0.007	0.730	-0.010	0.628
F02	0.018	0.398	0.035	0.102
F03	0.016	0.446	0.014	0.516
F04	0.032	0.135	0.016	0.460
F05	0.033	0.127	0.017	0.425
F06	0.054**	0.012	0.054**	0.013
F07	0.086**	0.000*	0.091**	0.000*
F08	0.051**	0.000*	0.058**	0.000*
F09	0.013	0.536	0.016	0.471
F10	0.107**	0.000*	0.092**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.

* Represents value less than 10-6.

(2) ZL station realm

According to Table 5- 8, in the ZL station realm, there is a negative correlation between government agencies and population activity. There is no correlation between public service facilities and population activity, while financial service facilities show a significant positive correlation with population activity only on working days. The remaining six types of functional facilities exhibit a significant positive correlation with population activity, with catering services showing the highest correlation coefficient. Additionally, accommodation services, catering services, living services, and shopping services all demonstrate an increase in positive correlation coefficients with population activity on off days.

Table 5- 8. Statistics of correlation coefficients between population activities and various functional facilities in ZL station realm

Code	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
F01	0.069**	0.001	0.053**	0.010
F02	0.044**	0.032	0.026	0.195
F03	-0.013	0.527	-0.032	0.117
F04	0.057**	0.005	0.047**	0.020
F05	0.122**	0.000*	0.140**	0.000*
F06	0.177**	0.000*	0.187**	0.000*
F07	0.141**	0.000*	0.174**	0.000*
F08	0.131**	0.000*	0.170**	0.000*
F09	0.021	0.308	0.026	0.209
F10	0.080**	0.000*	0.065**	0.001

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

(3) NSM station realm

As shown in Table 5- 9, within the NSM station realm, apart from government agencies and public service facilities, the other eight types of functional facilities exhibit a significant positive correlation with population activity. Among them, catering services have the highest correlation coefficient, while medical and health care facilities have the lowest correlation coefficient. Additionally, accommodation services, catering services, living services, and shopping services all show an increase in positive correlation coefficients with population activity on off days.

Table 5- 9. Statistics of correlation coefficients between population activities and various functional facilities in NSM station realm

Code	WD		OD	
	Correlation coefficient	*P*-value	Correlation coefficient	*P*-value
F01	0.171**	0.000*	0.140**	0.000*
F02	0.066**	0.002	0.052**	0.000*
F03	0.020	0.347	0.028	0.191
F04	0.064**	0.002*	0.036	0.089

F05	0.121**	0.000*	0.132**	0.000*
F06	0.151**	0.000*	0.178**	0.000*
F07	0.114**	0.000*	0.133**	0.000*
F08	0.168**	0.000*	0.169**	0.000*
F09	0.015	0.473	0.019	0.359
F10	0.138**	0.000*	0.116**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

(4) XZ station realm

As shown in Table 5- 10, within the XZ station realm, government agencies do not exhibit any correlation with population activity. However, the other nine types of functional facilities show a significant positive correlation with population activity. Among them, catering services have the highest correlation coefficient with population activity both on off days and working days., followed by shopping services and living services. Public service facilities have the lowest correlation coefficient with population activity.

Table 5- 10. Statistics of correlation coefficients between population activities and various functional facilities in XZ station realm

Code	WD		OD	
	Correlation coefficient	*P*-value	Correlation coefficient	*P*-value
F01	0.178**	0.000*	0.174**	0.000*
F02	0.097**	0.000*	0.103**	0.000*
F03	0.059	0.134	0.023	0.275
F04	0.082**	0.000*	0.090**	0.000*
F05	0.179**	0.000*	0.188**	0.000*
F06	0.299**	0.000*	0.295**	0.000*
F07	0.229**	0.000*	0.242**	0.000*
F08	0.261**	0.000*	0.262**	0.000*
F09	0.044**	0.032	0.060**	0.014
F10	0.126**	0.000*	0.143**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

(5) LSY station realm

As shown in Table 5- 11, within the LSY station realm, there is no correlation between medical and health care facilities as well as public service facilities and population activity. This could be due to a small sample size or the absence of a true correlation between the two variables. Further research is needed to study the relationship trend between them before making definitive conclusions. On off days, government agencies exhibit a significant negative correlation with population activity. The remaining seven types of functional facilities show a significant positive correlation with population activity, with catering services having the highest correlation coefficient.

Table 5- 11. Statistics of correlation coefficients between population activities and various functional facilities in LSY station realm

Code	WD		OD	
	Correlation coefficient	*P*-value	Correlation coefficient	*P*-value
F01	0.105**	0.000*	0.095**	0.000*
F02	0.067**	0.001	0.070**	0.000*
F03	0.005	0.794	-0.046**	0.020
F04	0.037	0.062	0.038	0.054
F05	0.095**	0.000*	0.107**	0.000*
F06	0.184**	0.000*	0.190**	0.000*
F07	0.153**	0.000*	0.157**	0.000*
F08	0.160**	0.000*	0.183**	0.000*
F09	0.028	0.159	0.019	0.320
F10	0.098**	0.000*	0.094**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

(6) BDJ station realm

As shown in Table 5- 12, within the BDJ station realm, there is no correlation between entertainment facilities, financial service facilities, government agencies, and public service facilities with population activity. The remaining six types of functional facilities exhibit a significant positive correlation with population activity. On working days, medical and health care facilities have the highest correlation coefficient with population activity, followed by shopping service facilities, while living service facilities have the lowest correlation. On off days, shopping service facilities have the highest correlation coefficient with population activity, and accommodation service facilities have the lowest positive correlation.

Table 5- 12. Statistics of correlation coefficients between population activities and various functional facilities in BDJ station realm

Code	WD		OD	
	Correlation coefficient	*P*-value	Correlation coefficient	*P*-value
F01	0.023	0.293	0.025	0.257
F02	0.042	0.053	0.031	0.146
F03	0.042	0.050	0.022	0.301
F04	0.131**	0.000*	0.125**	0.000*
F05	0.096**	0.000*	0.096**	0.000*
F06	0.122**	0.000*	0.141**	0.000*
F07	0.083**	0.000*	0.115**	0.000*
F08	0.129**	0.000*	0.155**	0.000*
F09	0.007	0.752	0.004	0.842
F10	0.115**	0.000*	0.106**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.

* Represents value less than 10-6.

(7) STSG station realm

As shown in Table 5- 13, within the STSG station realm, except for medical and health care facilities and public service facilities, the other eight types of functional facilities exhibit a significant positive correlation with population activity on both working days and off days. Among them, living service facilities have the highest positive correlation coefficient with population activity, followed by dining service facilities, while government agencies have the lowest correlation.

Table 5- 13. Statistics of correlation coefficients between population activities and various functional facilities in STSG station realm

Code	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
F01	0.130**	0.000*	0.122**	0.000*
F02	0.138**	0.000*	0.119**	0.000*
F03	0.089**	0.000*	0.066**	0.001
F04	0.051	0.210	0.018	0.374
F05	0.165**	0.000*	0.151**	0.000*
F06	0.227**	0.000*	0.215**	0.000*
F07	0.249**	0.000*	0.235**	0.000*
F08	0.222**	0.000*	0.206**	0.000*
F09	0.032	0.116	0.079	0.320
F10	0.178**	0.000*	0.156**	0.000*

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

(8) WYJ station realm

As shown in Table 5- 14, within the WYJ station realm, government agencies do not exhibit any correlation with population activity, while entertainment facilities only show a significant positive correlation on off days. The other eight types of functional facilities have a significant positive correlation with population activity on both working days and off days, with shopping service facilities having the highest correlation coefficient with population activity.

Table 5- 14. Statistics of correlation coefficients between population activities and various functional facilities in WYJ station realm

Code	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
F01	0.027	0.164	0.040	0.136
F02	0.063**	0.001	0.045**	0.009
F03	0.032	0.088	0.002	0.899
F04	0.071**	0.000*	0.073**	0.000*
F05	0.111**	0.000*	0.127**	0.000*
F06	0.119**	0.000*	0.138**	0.000*

F07	0.112**	0.000*	0.116**	0.000*
F08	0.144**	0.000*	0.166**	0.000*
F09	0.044**	0.021	0.058**	0.012
F10	0.043**	0.024	0.045**	0.019

Notes:
Asterisks next to numbers indicate statistically significant ($p < 0.05$).
** Represents significant correlation.
* Represents value less than 10-6.

5.4 Discussion

5.4.1 The variability of correlation at different aspects

In terms of time, the positive correlation between population activity and functional facilities on off days is generally higher than on working days in most station realms. This is primarily due to the typical activity patterns of urban populations. On off days, people usually have more time for shopping, entertainment, and other activities. However, it is worth noting that in the YDGY station realm, which is in the under-development type, and the STSG station realm, which is of a general residential type, the positive correlation between population activity and functional facilities is higher on working day than on off days. This could be attributed to the inadequate configuration of functional facilities in these two station realms, which fail to meet the increased activity demands of people on off days, thereby forcing them to seek additional service facilities in other areas. Additionally, it is also possible that some residents in these two station realms prefer to stay at home and refrain from going out on off days.

As shown in Table 5- 15, there are significant variations in the correlation coefficients between different station realms. However, both on working day and off days, the XZ station realm exhibits the highest correlation between population activity and functional facilities, while the YDGY station realm has the lowest correlation. It is worth noting that although the ZL station realm, which is of the urban center type, has the highest number of functional facilities, its correlation coefficient with population activity ranks fourth. Additionally, for example, both the ZL and XZ station realms have lower population activity counts than the WYJ station realm, but their correlation coefficients are higher than that of the WYJ station realm. Therefore, while summarizing the overall patterns, it is important to conduct specific analyses of individual phenomena that emerge.

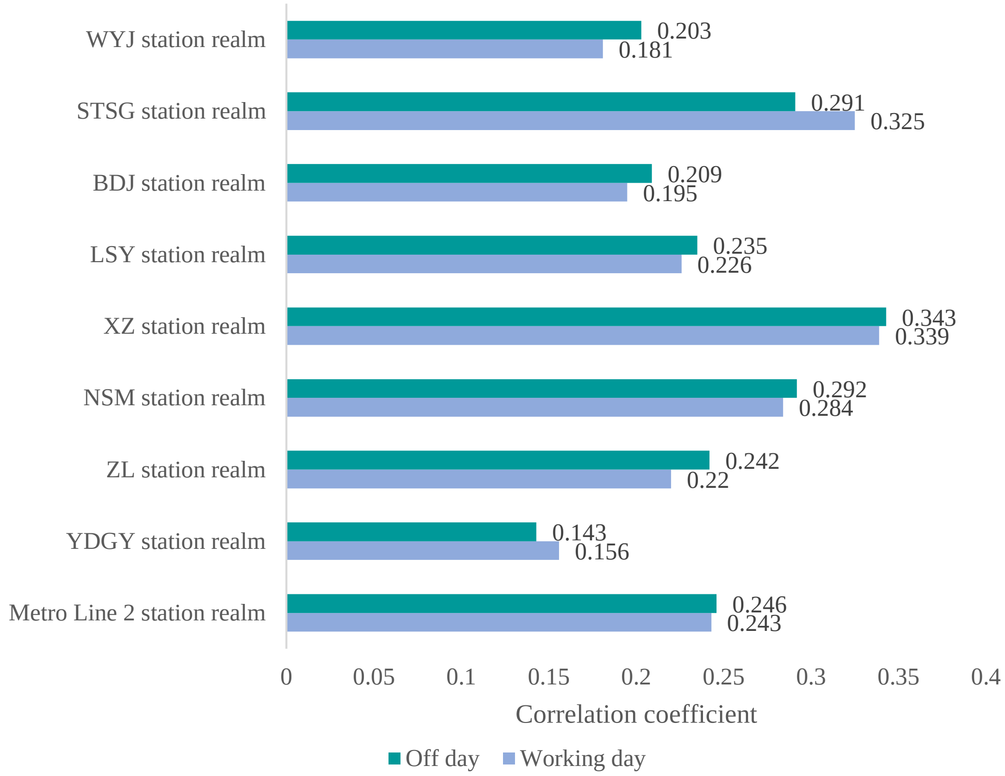


Figure 5- 3. Positive correlation intensity between population activity and functional facilities in station realm during working day and off day

Table 5- 15. Positive correlation intensity and quantity statistics of population activities and functional facilities in station realm

Name of station realm	Working day		Off day		Number of functional facilities
	Population numbers	Correlation coefficient	Population numbers	Correlation coefficient	
Metro Line 2	167748	0.243	177125	0.246	37012
YDGY	6007	0.156	6798	0.143	332
ZL	9739	0.143	10447	0.242	5374
NSM	9307	0.284	9133	0.292	3125
XZ	9133	0.339	9566	0.343	4575
LSY	9594	0.226	11278	0.235	3358
BDJ	10798	0.195	11209	0.209	3163
STSG	8176	0.325	8181	0.291	2443
WYJ	14998	0.181	14468	0.203	1846

From the perspective of functional facility categories, the correlation coefficients between population activity and three types of facilities—namely, catering services, living services, and shopping—are consistently ranked among the top three in each station realm of Xi'an Metro Line 2. This indicates that catering services, living services, and shopping are important factors influencing the quantity and spatial distribution of population activity in the station realms. Additionally, public service facilities exhibit significant positive correlation only in the XZ station realm, while there is no correlation in other station realms. Government agencies show positive correlation with population activity only in the STSG station realm.

5.4.2 Determining quantitative indicators for the spatial relationship models

(1) Under development type

Based on the results of the correlation analysis, as shown in Table 5- 16, there is a significant positive correlation between population activity and the following four categories of functional facilities: catering service, living service, shopping, and transportation service, both on working days and off days. Therefore, this study will consider these four categories of functional facilities as the primary indicators for the spatial relationship model of the YDGY station realm.

Table 5- 16. Quantitative indicators of spatial relationships of YDGY station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Catering	0.054**	0.012	0.054**	0.013
Living	0.086**	0.000*	0.091**	0.000*
Shopping	0.051**	0.000*	0.058**	0.000*
Transportation	0.107**	0.000*	0.092**	0.000*

(2) Urban center type

Based on the results of the correlation analysis, there is no correlation between population activity and government agencies as well as public service facilities on both working days and off days. Therefore, these two categories are excluded. This study will consider the remaining eight categories of functional facilities as the primary indicators for evaluating the spatial relationship model of the ZL station realm, as shown in Table 5- 17.

Table 5- 17. Quantitative indicators of spatial relationships of ZL station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Entertainment	0.069**	0.001	0.053**	0.010
Financial	0.044**	0.032	—	——
Medical & Health care	0.057**	0.005	0.047**	0.020
Accommodation	0.122**	0.000*	0.140**	0.000*
Catering	0.177**	0.000*	0.187**	0.000*
Living	0.141**	0.000*	0.174**	0.000*
Shopping	0.131**	0.000*	0.170**	0.000*
Transportation	0.080**	0.000*	0.065**	0.001

(3) High population density with hub type

Based on the results of the correlation analysis for the NSM station realm, there is no correlation between government agencies and public service facilities with population activity. However, medical & health care facilities show a significant positive correlation with population activity only on working days. Therefore, as shown in Table 5- 18, this study will exclude government agencies and public service facilities, while considering the remaining eight categories of service facilities as the primary indicators for evaluating the spatial relationship model of the NSM station realm on working days. For the model indicators on off days, the evaluation will include seven categories of service facilities, excluding government agencies, medical and health care facilities, and public service facilities.

Table 5- 18. Quantitative indicators of spatial relationships of NSM station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Entertainment	0.171**	0.000*	0.140**	0.000*
Financial	0.066**	0.002	0.052**	0.000*
Medical & Health care	0.064**	0.002*	0.036	0.089
Accommodation	0.121**	0.000*	0.132**	0.000*
Catering	0.151**	0.000*	0.178**	0.000*
Living	0.114**	0.000*	0.133**	0.000*
Shopping	0.168**	0.000*	0.169**	0.000*
Transportation	0.138**	0.000*	0.116**	0.000*

Based on the correlation analysis results for the XZ station realm, except for government agencies, population activity shows a significant positive correlation with the other nine categories of functional facilities. Therefore, as shown in Table 5- 19, this study will consider these nine categories of functional facilities as the primary indicators for evaluating the spatial relationship model of the XZ station realm.

Table 5- 19. Quantitative indicators of spatial relationships of XZ station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Entertainment	0.178**	0.000*	0.174**	0.000*
Financial	0.097**	0.000*	0.103**	0.000*
Medical & Health care	0.082**	0.000*	0.090**	0.000*
Accommodation	0.179**	0.000*	0.188**	0.000*
Catering	0.299**	0.000*	0.295**	0.000*
Living	0.229**	0.000*	0.242**	0.000*
Shopping	0.261**	0.000*	0.262**	0.000*
Public service	0.044**	0.032	0.060**	0.014
Transportation	0.126**	0.000*	0.143**	0.000*

(4) Interchange type

Based on the correlation analysis results for the LSY station realm, it is evident that there is a significant positive correlation between population activity and seven categories of functional facilities, except for government agencies, Medical & Health care, and public services. Therefore, as shown in Table 5- 20, these seven categories of functional facilities will be considered as the primary indicators for evaluating the spatial relationship model of the LSY station realm.

Table 5- 20. Quantitative indicators of spatial relationships of LSY station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Entertainment	0.105**	0.000*	0.095**	0.000*

Financial	0.067**	0.001	0.070**	0.000*
Accommodation	0.095**	0.000*	0.107**	0.000*
Catering	0.184**	0.000*	0.190**	0.000*
Living	0.153**	0.000*	0.157**	0.000*
Shopping	0.160**	0.000*	0.183**	0.000*
Transportation	0.098**	0.000*	0.094**	0.000*

Based on the correlation analysis results for the BDJ station realm, it is evident that there is a significant positive correlation between population activity and six categories of functional facilities, including medical and health care, accommodation services, catering services, living services, shopping, and transportation services, on both working days and off days. Therefore, as shown in Table 5- 21, these six categories of functional facilities will be considered as the primary indicators for evaluating the spatial relationship model of the BDJ station realm.

Table 5- 21. Quantitative indicators of spatial relationships of BDJ station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Medical & Health care	0.131**	0.000*	0.125**	0.000*
Accommodation	0.096**	0.000*	0.096**	0.000*
Catering	0.122**	0.000*	0.141**	0.000*
Living	0.083**	0.000*	0.115**	0.000*
Shopping	0.129**	0.000*	0.155**	0.000*
Transportation	0.115**	0.000*	0.106**	0.000*

(5) General residential type

Based on the correlation analysis results for the STSG station realm, it is observed that there is a significant positive correlation between population activity and eight categories of service facilities, excluding medical and health care facilities and public service facilities. Therefore, as shown in Table 5- 22, these eight categories of functional facilities will be considered as the primary indicators for evaluating the spatial relationship model of the STSG station realm.

Table 5- 22. Quantitative indicators of spatial relationships of STSG station realm

Functional facilities	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Entertainment	0.130**	0.000*	0.122**	0.000*
Financial	0.138**	0.000*	0.119**	0.000*
Government agency	0.089**	0.000*	0.066**	0.001
Accommodation	0.165**	0.000*	0.151**	0.000*
Catering	0.227**	0.000*	0.215**	0.000*
Living	0.249**	0.000*	0.235**	0.000*
Shopping	0.222**	0.000*	0.206**	0.000*
Transportation	0.178**	0.000*	0.156**	0.000*

Similarly, within the WYJ station realm, there is a significant positive correlation between population activity and eight categories of functional facilities, excluding entertainment facilities and government agencies. Thus, as presented in Table 5- 23, these eight categories of functional facilities will be regarded as the primary indicators for assessing the spatial relationship model of the WYJ station realm.

Table 5- 23. Quantitative indicators of spatial relationships of HZZX station realm

Code	WD		OD	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Entertainment	0.027	0.164	0.040**	0.036
Financial	0.063**	0.001	0.045**	0.009
Medical & Health care	0.071**	0.000*	0.073**	0.000*
Accommodation	0.111**	0.000*	0.127**	0.000*
Catering	0.119**	0.000*	0.138**	0.000*
Living	0.112**	0.000*	0.116**	0.000*
Shopping	0.144**	0.000*	0.166**	0.000*
Public service	0.044**	0.021	0.058**	0.012
Transportation	0.043**	0.024	0.045**	0.019

5.5 Chapter summary

In this chapter, the study employed Spearman correlation analysis to examine the relationships between overall population activity and various types of functional facilities in the station realms of Xi'an Metro Line 2. It also investigated the correlation levels between population activity and different types of functional facilities in each station realm. Based on the criterion of significant positive correlation, interaction indicators were selected for different types of station realms. The "Under development" type YDGY station realm includes four types of functional facilities: catering service, living Service, shopping, and transportation service. The "Urban center" type ZL station realm includes seven types of functional facilities: entertainment, medical & health care, accommodation service, catering service, living service, shopping, and transportation service. The "High population density with hub" type XZ station realm includes nine types of functional facilities, excluding government agencies. The "Interchange" type LSY station realm includes seven types of functional facilities, excluding government agencies, Medical & Health Care, and public service facilities. The "General residential" type HZZX station realm includes eight types of functional facilities, excluding government agencies and public service facilities. These research findings provide a quantitative indicator system for the construction of the subsequent interactive mechanism framework and spatial relationship model.

References

[1] Lee H S, Lim J H. Statistical package for the social sciences[J]. JypHyunJae Publication, 2013.

[2] Schober P, Boer C, Schwarte L A. Correlation coefficients: appropriate use and interpretation[J]. Anesthesia & analgesia, 2018, 126(5): 1763-1768.

[3] Choi S C. Tests of equality of dependent correlation coefficients[J]. Biometrika, 1977, 64(3):

645-647.

[4] Cowles, M., & Davis, C. On the Origins of the .05 Level of Statistical Significance. American Psychologist, 1982, 37(5), 553–558.

[5] Xiao C, Ye J, Esteves R M, et al. Using Spearman's correlation coefficients for exploratory data analysis on big dataset[J]. Concurrency and Computation: Practice and Experience, 2016, 28(14): 3866-3878.

[6] Zong L, He S, Lian J, et al. Detailed mapping of urban land use based on multi-source data: a case study of lanzhou[J]. Remote Sensing, 2020, 12(12): 1987.

Chapter 6. Research on interaction mechanism within station realm

Content

Introduction

This chapter is based on the results of the previous chapter's correlation analysis. It selects the service facilities that have a significant positive correlation with population activity as explanatory variables and the number of population activity on working days and off days as the dependent variable. An OLS (Ordinary Least Squares) global regression model is established to analyze the basic interaction between population activity and functional facilities in different types of station realms. Furthermore, for the explanatory variables with good significance, a GWR (Geographically Weighted Regression) local regression model is constructed to further explore the key variables that dominate the interaction relationship.

The chapter also investigates the differences in the interaction between population activity and functional facilities across different station realms from three perspectives: time, type, and scale. Leveraging the theory of supply and demand relationships, it summarizes the interaction mechanisms between population activity and various types of functional facilities, and provides recommendations for the future development of rail transit station realms. The research framework of this chapter is illustrated in Figure 6- 1.

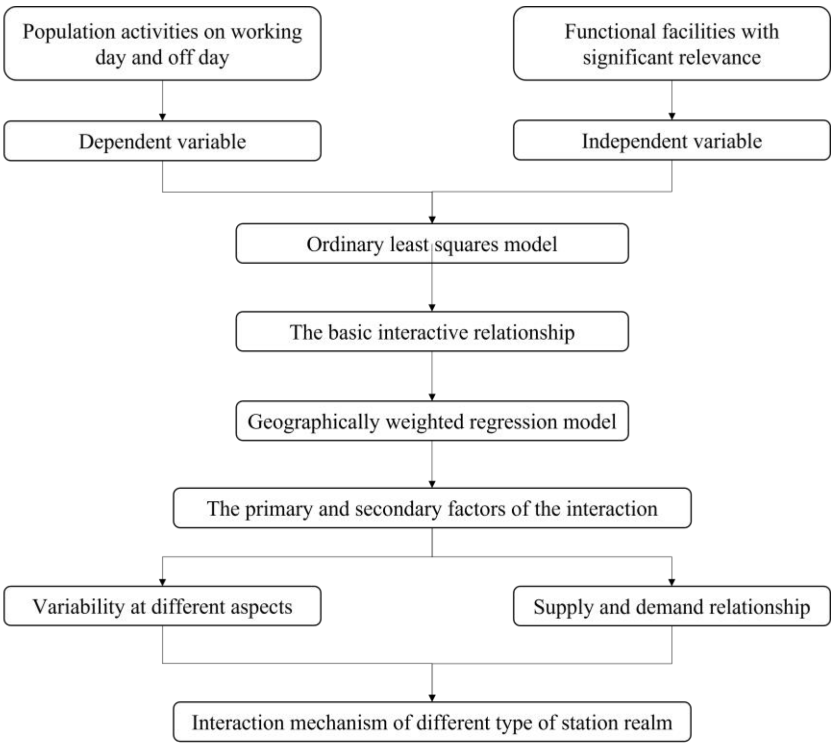


Figure 6- 1. Research framework

6.1 Methodology

6.1.1 Parameterization of the models

In this study, the dependent variable in both the OLS and GWR regression models is the total number of population activities within each experimental station realm on either working days or off days. The specific variable codes for the dependent variable are presented in Table 6- 1.

Table 6- 1. The dependent variable code of each experimental station realm

Station realm	Code: Number of working day population activities	Code: Number of off day population activities
YNGY	Y_WD	Y_OD
ZL	Z_WD	Z_OD
NSM	N_WD	N_OD
XZ	X_WD	X_OD
LSY	L_WD	L_OD
BDJ	B_WD	B_OD
STSG	S_WD	S_OD
WYJ	W_WD	W_OD

The explanatory variables in the OLS and GWR regression models are the quantities of various functional facilities within each station realm. The selection of explanatory variables in this study is mainly based on the results of the correlation analysis in sections 5.3 and 5.4, focusing on the functional facility categories that have a significant positive correlation with population activities within each experimental station realm. The specific variable codes for the explanatory variables in each station realm are presented in Table 6- 2.

Table 6- 2. Independent variable code of each experimental station realm

Name of station realm	POI types	Code
YDGY station realm	Catering service POI	Y_01
	Living service POI	Y_02
	Shopping service POI	Y_03
	Transportation service POI	Y_04
ZL station realm	Entertainment POI	Z_01
	Financial service POI	Z_02
	Medical & Health care POI	Z_03
	Accommodation service POI	Z_04
	Catering service POI	Z_05
	Living service POI	Z_06
	Shopping service POI	Z_07
	Transportation service POI	Z_08
XZ station realm	Entertainment POI	X_01
	Financial service POI	X_02
	Medical & Health care POI	X_03
	Accommodation service POI	X_04
	Catering service POI	X_05
	Living service POI	X_06
	Shopping service POI	X_07
	Public service POI	X_08
	Transportation service POI	X_09
LSY station realm	Entertainment POI	L_01
	Financial service POI	L_02

	Accommodation service POI	L_03
	Catering service POI	L_04
	Living service POI	L_05
	Shopping service POI	L_06
	Transportation service POI	L_07
WYJ station realm	Financial service POI	W_01
	Medical & Health care POI	W_02
	Accommodation service POI	W_03
	Catering service POI	W_04
	Living service POI	W_05
	Shopping service POI	W_06
	Public service POI	W_07
	Transportation service POI	W_08

6.1.2 OLS modeling

The specific operational procedure for conducting regression analysis of population activities and functional facilities using the OLS tool in the ArcGIS platform in this study is outlined as follows:

1. Input

In order to more accurately assess the dynamic influence relationship between functional facilities and population activities. Based on the boundaries of each study station realm, a grid with an image element scale of 25*25 m was created on ArcGIS platform as the spatial statistical unit for the regression analysis in this paper (which is consistent with the collection scale of our population activity data).

The dependent variable (population activity) and the explanatory variable (functional facility) were connected to this grid for OLS regression analysis.

2. Checking the performance of the regression model

Model performance is evaluated using the multiple R-squared and adjusted R-squared values. These metrics measure the goodness-of-fit of the model and range from 0.0 to 1.0. The adjusted R-squared value, which takes into account the complexity of the model (number of variables), is usually slightly smaller than the multiple R-squared value.

The AICc (Akaike Information Criterion, AIC corrected) is used to assess and compare the performance of regression models. Models with lower AICc values better fit the observed data. If the difference in AICc values between two models is greater than 3, the model with the lower AICc value is considered superior.

The significance of the model is evaluated using the joint F-statistic and joint chi-square statistic. The joint F-statistic is reliable only when the Koenker (BP) statistic is not statistically significant. If the Koenker (BP) statistic is significant, the significance of the entire model should be determined based on the joint chi-square statistic. The null hypothesis for both tests is that the explanatory variables have no effect in the model. For a confidence level of 95%, p-values less than 0.05 indicate statistical significance of the model.

3. Interpretation of OLS results:

The variance inflation factor (VIF) is used to measure the redundancy among explanatory variables. Explanatory variables associated with VIF values greater than 10 are typically removed from the regression model one by one.

Coefficients are used to interpret the relationships between each explanatory variable and the dependent variable. The sign of the coefficient indicates the direction of the relationship. A negative sign indicates a negative relationship between the explanatory variable and the dependent variable. The coefficient reflects the expected change in the dependent variable for a one-unit change in the associated explanatory variable, with all other variables held constant.

The T-test is used to assess the statistical significance of individual explanatory variables. When the probability or robust probability (p-value) is small, the likelihood that the coefficient is zero is also small.

4. Assessment of stationarity

The Koenker (BP) statistic, which is a standardized Breusch-Pagan statistic, is used to test whether the explanatory variables in the model have consistent relationships with the dependent variable in both geographic space and data space. If the model is consistent in geographic space, the spatial process represented by the explanatory variables will also exhibit consistency across different locations in the study area. If the model is consistent in data space, the relationship between the predicted values and each explanatory variable will not vary with changes in the explanatory variable values (indicating the absence of heteroscedasticity in the model). For a confidence level of 95%, p-values less than 0.05 indicate statistically significant heteroscedasticity or non-stationarity in the model.

The above is the detailed operational procedure for conducting regression analysis of population activities and functional facilities using the OLS tool in the ArcGIS platform, as described in this study.

6.1.3 GWR modeling

The basic construction process of GWR modeling using the GWR tool in the ArcGIS platform for regression analysis of population activities and functional facilities in each study station realm is as follows:

1. Input:

In the input feature parameters, include the field for the dependent variable and the fields representing the explanatory variables. Identify features in the data that contain missing values in either the dependent or explanatory variables and use the fill missing values tool to complete the dataset.

2. Model type:

The GWR tool provides three types of regression models: continuous, binary, and count. In statistical literature, these regression types are referred to as Gaussian, logistic, and Poisson, respectively. Based on the measurement and range of values of the dependent variable in this study, the Gaussian model type is selected.

3. Bandwidth selection:

The bandwidth, also known as the neighborhood, is the distance range or number of neighboring features used for each local regression equation. It is a crucial parameter to control the smoothness of the model. The shape and extent of the analyzed neighborhood depend on the input parameters of neighborhood type and neighborhood selection method.

The neighborhood type parameter can be based on either the number of neighboring features or the distance range. When using the number of neighboring features, the neighborhood size is specified as a function of the specified number of neighboring features, resulting in smaller neighborhoods in dense feature locations and larger neighborhoods in sparse feature locations. When using the distance range, the neighborhood size for each feature in the study area remains constant, resulting in more features included in the neighborhood in dense feature locations and fewer features in sparse feature locations.

The neighborhood selection method parameter specifies how the neighborhood size is determined (actual distance used or number of neighboring features). When selecting the neighborhood through the golden search or manual interval option, the basis is always to minimize the value of the Akaike Information Criterion (AICc).

In this study, the golden search method is chosen to determine the optimal value of the distance range or number of neighboring features parameter. The golden search first identifies the maximum and minimum distances and iteratively tests the AICc at different distances between them. Once the number of features in the dataset exceeds 1000, the maximum distance is the distance at which each feature has at most 1000 neighboring features. The minimum distance is the distance at which each feature has at least 20 neighboring features. If the number of features is less than 1000, the maximum distance is the distance at which each feature has n/2 neighboring features (half of the features as neighboring features), and the minimum distance is the distance at which each feature has at least 5% of n neighboring features (5% of the features in the dataset as neighboring features). The golden search determines the distance or number of neighboring features with the lowest AICc as the neighborhood size.

4. Setting the local weight scheme:

The strength of GWR is that it can apply geographic weights to the features used in each local regression equation. Features that are farther away from the regression point have smaller weights, resulting in less influence on the regression results for the target feature, while features that are closer to the regression point have larger weights. A kernel is used to determine the weights, which is a distance decay function that determines the rate at which the weights decrease with increasing distance. The GWR tool provides two kernel options in the local weight scheme parameter: Gaussian and bisquers.

Due to the Gaussian weight scheme ensuring that each regression feature has multiple neighboring features, it increases the chances of variations in those neighboring feature values. This helps avoid the issue of spatial autocorrelation in geographically weighted regression. Therefore, the Gaussian weight scheme is selected to assign weights to each regression feature. The Gaussian weight scheme assigns a weight value of 1 to the regression feature (feature i), and as the distance between surrounding features (feature j) and the regression feature increases, the weight becomes smooth and gradually decreases. The Gaussian weight scheme never reaches zero, but for features far from the regression feature, the weight may be very small and have almost no impact on the regression.

6. Output:

In addition to the regression residuals, the output table includes parameter fields such as observed and predicted values of the dependent variable, condition number (COND), local R2, explanatory variable coefficients, and standard errors. It also reports parameters such as intercept, standard error of the intercept, coefficients, standard errors for each explanatory variable, predictions, and residuals.

Regression coefficients: They represent the strength and type of relationship between the explanatory variables and the dependent variable. If the regression coefficient is positive, it indicates a positive relationship between the explanatory variable and the dependent variable; otherwise, a negative relationship exists. If the relationship is strong, the regression coefficient will be relatively large, while it will be close to 0 for weak relationships.

R^2 can be used to assess the goodness of fit. Its value varies between 0.0 and 1.0, with higher values indicating a better fit. This value can be interpreted as the proportion of the dependent variable variance covered by the regression model.

$AdjR^2$ is the adjusted R-squared value, which compensates for the effect of the number of variables in the model, so the adjusted R2 value is typically smaller than the R2 value.

AICc can be used to assess model performance and compare regression models. Considering model complexity, models with lower AICc values will provide a better fit to the observed data.

$Sigma^2$ is the least-squares estimate of the residual variance (square of the standard deviation). Smaller values of this statistic are preferable. This value represents the normalized sum of squared residuals (sum of squared residuals divided by the effective degrees of freedom of the residuals). $Sigma^2$ is used in AICc calculation.

6.2 Regression results and analysis of the OLS model

For OLS regression results it is first necessary to calculate the variance inflation factor (VIF) to avoid multicollinearity among the explanatory variables of the model and to exclude variables with VIF values greater than 10 [1]. Secondly, the R-squared of the regression results also needs to be tested, and if the value is greater than 0.6, it indicates a good fit for the model. For the P-value of regression results, a regression coefficient is considered significant when $P < 0.05$ [2]. This means that statistically, we can reject the null hypothesis and conclude that there is a genuine association between the regression coefficient and the dependent variable, rather than it being due to random error.

6.2.1 OLS regression results of each station realm

(1) YDGY station realm

The OLS regression results for the YDGY station realm are shown in Table 6- 3, and the VIF values for all explanatory variables are below 10. This indicates that there is no significant multicollinearity among the variables. Additionally, the adjusted R-squared for the model established on working days is 0.6810, and for off days, it is 0.7140. These values suggest that the model fits the data well, indicating that the selected combination of explanatory variables has a good explanatory power for the dependent variable.

Table 6- 3. The explanatory variables of the VIF of YDGY station realm

Code	VIF
Y_01	1.049174
Y_02	1.118599
Y_03	1.150115
Y_04	1.003770

Table 6- 4 presents the OLS regression results for the model. Both the coefficients of living services and transportation services show significance for both working days and off days. Therefore, the interaction model between population activities and functional facilities in the YDGY station realm on working days and off days can be expressed as follows:

$$Y_W = 2.9280 + 1.6989 \times \mathrm{Y02} + 1.5124 \times \mathrm{Y04} \qquad 6\text{- }1$$

$$Y_O = 3.3849 + 1.6905 \times \mathrm{Y02} + 1.9742 \times \mathrm{Y04} \qquad 6\text{- }2$$

Where Y_W is the total population activity in the YDGY station realm on working days, Y_O is the total population activity on working days, Y02 represents the number of living services, and Y04 represents the number of transportation services.

Table 6- 4. The OLS regression results for the YDGY station realm

Independent variable	Dependent variable			
	Y_WD		Y_OD	
	Coefficient	*P-value*	Coefficient	*P-value*
Intercept	2.9280	0.000000**	3.3849	0.000000**
Y_01	1.1381	0.147526	2.4442	0.144393
Y_02	1.6986	0.000137**	1.6905	0.001208**
Y_03	0.3134	0.406285	0.1042	0.814612
Y_04	1.5124	0.000132**	1.9742	0.000025**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(2) ZL station realm

The VIF results for the OLS regression of the ZL station realm, as shown in Table 6- 5, it shows that all explanatory variables have VIF values less than 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model established on working days is 0.63409, and for off days is 0.71392, suggesting a good fit of the model.

Table 6- 5. The explanatory variables of the VIF of ZL station realm

Code	VIF
Z_01	1.002448
Z_02	1.038727
Z_03	1.033428
Z_04	1.079613
Z_05	1.255517
Z_06	1.264698
Z_07	1.211425
Z_08	1.010687

The OLS regression results for the ZL station realm are presented in Table 6- 6. The regression coefficients for accommodation facilities, catering services, shopping facilities, and transportation services are all statistically significant on both working days and off days. However, the regression coefficient for medical and health care services is significant only on working days and not on off days. The regression coefficient for living services is significant on off days but not on working days. Based on these results, the interaction equation between population activity and functional facilities in the ZL station realm on working days and off days can be expressed as follows:

$$Z_W = 4.1941 + 1.2653 \times Z03 + 0.4481 \times Z04 + 0.2279 \times Z05 + 0.1686 \times Z07 + 1.2017 \times Z08 \quad \text{6- 3}$$

$$Z_O = 4.7826 + 0.6542 \times Z04 + 0.3699 \times Z05 + 0.331 \times Z06 + 0.262 \times Z07 + 0.8524 \times Z08 \quad \text{6- 4}$$

Where: Z_W is the total population activity in the ZL station realm on working days, Z_O is the total population activity on off days, Z02 represents the quantity of medical and health care facilities, Z03 represents the quantity of accommodation facilities, Z04 represents the quantity of catering services, Z05 represents the quantity of living services, Z06 represents the quantity of shopping facilities, and Z07 represents the quantity of transportation services in the ZL station realm.

Table 6- 6. The OLS regression results for the ZL station realm

Independent variable	Dependent variable			
	Z_WD		Z_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	4.1941	0.000000**	4.7826	0.000000**
Z_01	0.1732	0.592388	-0.1145	0.777830
Z_02	0.0975	0.537931	—	—
Z_03	1.2653	0.001554**	0.2311	0.644199
Z_04	0.4481	0.000000**	0.6542	0.000000**
Z_05	0.2279	0.003122**	0.3699	0.000142**
Z_06	0.1404	0.187876	0.3306	0.013404**
Z_07	0.1686	0.000000**	0.2620	0.000000**
Z_08	1.2017	0.000220**	0.8524	0.035699**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(3) NSM station realm

As shown in Table 6- 7, all the explanatory variables have VIF values below 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model on working days is 0.6141, and on off days it is 0.6399, indicating that the selected combination of explanatory variables explains the dependent variable well.

Table 6- 7. The explanatory variables of the VIF of NSM station realm

Code	VIF
N_01	1.143303
N_02	1.049968
N_03	1.036397
N_04	1.826333
N_05	2.343824
N_06	2.433788
N_07	1.432870
N_08	1.034472

The OLS regression results for the NSM station realm are shown in Table 6- 8. Among the variables, only the regression coefficients of entertainment facilities, shopping services facilities, and transportation services facilities are statistically significant, both on working days and off days. Additionally, the catering services facility has a significant positive regression coefficient only on off days. Therefore, the interactive relationship can be represented by the following equation:

$$N_W = 4.590375 + 3.988998 \times N01 + 0.646578 \times N07 + 3.493961 \times N08 \qquad 6\text{-}5$$

$$N_O = 4.204225 + 2.486632 \times N01 + 0.516433 \times N05 + 0.646578 \times N07 + 3.493961 \times N08 \qquad 6\text{-}6$$

Where: N_W is the total population activity in the NSM station realm on working days, N_O is the total population activity on off days, N01 represents the quantity of entertainment facilities, N05 is the quantity of catering facilities, N07 represents the quantity of shopping facilities, N08 represents the quantity of transportation services in the NSM station realm.

Table 6- 8. The OLS regression results for NSM station realm

Independent variable	Dependent variable			
	N_WD		N_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	4.590375	0.000000**	4.204225	0.000000**
N_01	3.988998	0.000000**	2.486632	0.000367**
N_02	0.589181	0.374523	0.254917	0.679775
N_03	0.651641	0.223042	—	—
N_04	0.524970	0.167754	0.436278	0.218102
N_05	0.312630	0.072137	0.516433	0.001450**
N_06	-0.202454	0.130922	-0.142650	0.251487
N_07	0.646578	0.000007**	0.541029	0.000044**
N_08	3.493961	0.000000**	2.70064	0.000007**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(4) XZ station realm

The VIF results of the OLS regression for the XZ station realm are shown in Table 6- 9. All the explanatory variables have VIF values below 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model on working days is 0.7871, and on off days it is 0.7630, indicating a good fit of the model and suggesting that the selected combination of explanatory variables explains the dependent variable well.

Table 6- 9. The explanatory variables of the VIF of XZ station realm

Code	VIF
X_01	1.192279
X_02	1.075916
X_03	1.145246
X_04	1.024140
X_05	2.043808
X_06	2.049021
X_07	1.459824
X_08	1.008882
X_09	1.045759

The OLS regression results for XZ station realm are shown in Table 6- 10, Among the variables, accommodation services, dining services, shopping services, and transportation services exhibit positive and significant regression coefficients on both working days and off days. The regression coefficient for entertainment facilities is only positively significant on working days. Additionally, living services facilities show a significant negative regression coefficient on off days. Therefore, the interactive relationship between population activity and functional facilities in the XZ station realm on working days and off day can be represented by the following equation:

$$X_W = 4.408006 + 0.816378 \times X01 + 0.350034 \times X04 + 0.774405 \times X05 + 0.253428 \times X07 + 20447416 \times X09 \quad 6\text{- }7$$

$$X_O = 5.067337 + 0.503266 \times X04 + 0.980967 \times X05 - 0.245065 \times X06 + 0.390971 \times X07 + 3.130025 \times X09 \quad 6\text{- }8$$

Where X_W is the total population activity in the XZ station realm on working days, X_O is the total population activity on off day. X01 is the quantity of entertainment facilities, X04 is the quantity of accommodation services, X05 is the quantity of catering services, X06 is the quantity of living services facilities, X07 is the quantity of shopping services, X09 is the quantity of transportation services facilities in XZ station realm.

Table 6- 10. The OLS regression results for the XZ station realm

Independent variable	Dependent variable			
	X_WD		X_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	4.408006	0.000000**	5.067337	0.000000**
X_01	0.816378	0.030991**	0.865511	0.081154
X_02	0.515782	0.469110	0.980156	0.294197
X_03	0.214575	0.810740	1.031404	0.294197

X_04	0.350034	0.018093**	0.503266	0.009568**
X_05	0.774405	0.000000**	0.980967	0.000000**
X_06	-0.135322	0.090119	-0.245065	0.019279**
X_07	0.253428	0.000050**	0.390971	0.000002**
X_08	0.431237	0.585278	0.834247	0.420826
X_09	2.447416	0.000050**	3.130026	0.000075**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(5) LSY station realm

The VIF results of the OLS regression for the LSY station realm are shown in Table 6- 11. All the explanatory variables have VIF values below 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model on working days is 0.73682, and on off days it is 0.69318, indicating a good fit of the model and suggesting that the selected combination of explanatory variables explains the dependent variable well.

Table 6- 11. The explanatory variables of the VIF of LSY station realm

Code	VIF
L_01	1.136245
L_02	1.020754
L_03	1.186348
L_04	1.411381
L_05	1.352967
L_06	1.339626
L_07	1.013178

Table 6- 12 displays the OLS regression results for the LSY station realm. The regression coefficients for financial services, accommodation services, catering services, shopping services, and transportation services are all positively significant on both working days and off days. Therefore, the interactive relationship between population activity and functional facilities in the LSY station realm on working days and off days can be represented by the following equation:

$$L_W = 4.317336 + 3.006991 \times L02 + 1.041471 \times L03 + 0.715467 \times L04 + 0.557436 \times L06 + 1.786846 \times L07 \quad (6\text{- }9)$$

$$L_O = 3.585013 + 4.316638 \times L02 + 4.316638 \times L03 + 1.081538 \times L04 + 0.401561 \times L06 + 1.691191 \times L07 \quad (6\text{- }10)$$

Where L_W is the total population activity in LSY station realm on working day, L_O is the total population activity on off day. L02 is the quantity of financial services facilities, L03 is the quantity of accommodation services, L04 is the quantity of catering services, L06 is the quantity of living services facilities, L07 is the quantity of transportation services facilities.

Table 6- 12. The OLS regression results for the LSY station realm

Independent variable	Dependent variable			
	L_WD		L_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	4.3173	0.000000**	3.5850	0.000000**
L_01	0.5508	0.443173	0.2227	0.765160

L_02	3.0070	0.000071**	4.3166	0.000000**
L_03	1.0415	0.010216**	1.0816	0.010185**
L_04	0.7155	0.000000**	0.5272	0.000098**
L_05	0.2196	0.137467	0.3155	0.248769
L_06	0.5574	0.000000**	0.4016	0.000003**
L_07	1.7869	0.000001**	1.6912	0.000008**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(6) BDJ station realm

The VIF results of the OLS regression for the BDJ station realm are shown in Table 6- 13. All the explanatory variables have VIF values below 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model on working days is 0.6308, and on off days it is 0.6147, indicating a good fit of the model and suggesting that the selected combination of explanatory variables explains the dependent variable well.

Table 6- 13. The explanatory variables of the VIF of BDJ station realm

Code	VIF
B_01	1.044839
B_02	1.068353
B_03	1.175305
B_04	1.286978
B_05	1.272553
B_06	1.031302

Table 6- 14 presents the OLS regression results for the BDJ station realm. On working days, the regression coefficients for medical and health care facilities, catering services, and transportation services are positively significant. On off days, the regression coefficients for medical and health care facilities, accommodation services, catering services, and transportation services are positively significant. Therefore, the interactive relationship between population activity and functional facilities in the BDJ station realm on working days and off days can be represented by the following equation:

$$B_W = 4.5070 + 2.1487 \times B01 + +0.5842 \times B03 + 1.3123 \times B06 \qquad 6\text{- }11$$

$$B_O = 4.6608 + 1.7867 \times B01 + 1.5448 \times B02 + 0.5251 \times B03 + 1.1401 \times B06 \qquad 6\text{- }12$$

Where B_W is the total population activity in BDJ station realm on working day, B_O is the total population activity on off day. B01 is the quantity of medical and health care facilities, B02 is the quantity of accommodation services, B03 is the quantity of catering services, B06 is the quantity of transportation services facilities in the BDJ station realm.

Table 6- 14. The OLS regression results for the BDJ station realm

Independent variable	Dependent variable			
	B_WD		B_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	4.5070	0.000000**	4.6608	0.000000**
B_01	2.1487	0.000000**	1.7867	0.001321**
B_02	0.4717	0.245601	1.5448	0.009016**

B_03	0.5842	0.000068**	0.5251	0.013064**
B_04	-0.2833	0.160914	-0.2549	0.385822
B_05	0.1471	0.097086	0.1980	0.124654
B_06	1.3123	0.016257**	1.1401	0.011044**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(7) STSG station realm

The VIF results of the OLS regression for the STSG station realm are shown in Table 6- 15. All the explanatory variables have VIF values below 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model on working days is 0.6599, and on off days it is 0.7234, indicating a good fit of the model and suggesting that the selected combination of explanatory variables explains the dependent variable well.

Table 6- 15. The explanatory variables of the VIF of STSG station realm

Code	VIF
S_01	1.441120
S_02	1.256115
S_03	1.021219
S_04	1.303485
S_05	1.425181
S_06	1.805255
S_07	1.250721
S_08	1.030349

Table 6- 16 presents the OLS regression results for the STSG station realm. On working days, the regression coefficients for catering services, shopping services, and transportation services are positively significant. On off days, the regression coefficients for catering services, living services, shopping services, and transportation services are positively significant. Therefore, the interactive relationship between population activity and functional facilities in the STSG station realm on working days and off days can be represented by the following equation:

$$S_W = 3.5439 + 1.7827 \times S01 + 1.1405 \times S03 + 0.4585 \times S05 + 0.1688 \times S07 + 2.2465 \times S08 \quad \text{6- 13}$$

$$S_O = 3.3091 + 0.6723 \times S04 + 0.5302 \times S05 + 0.3313 \times S06 + 0.2347 \times S07 + 2.5440 \times S08 \quad \text{6- 14}$$

Where S_W is the total population activity in STSG station realm on working day, S_O is the total population activity on off day. S05 is the quantity of catering services facilities, S06 is the quantity of living services, S07 is the quantity of shopping services, S08 is the quantity of transportation services facilities in the STSG station realm.

Table 6- 16. The OLS regression results for the STSG station realm

Independent variable	Dependent variable			
	S_WD		S_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	3.5439	0.000000**	3.3091	0.000000**
S_01	1.7827	0.011961**	0.6270	0.372681

S_02	0.6731	0.129526	1.6595	0.141627
S_03	1.1405	0.005601**	0.5086	0.165275
S_04	0.4605	0.224122	0.6723	0.009271**
S_05	0.4585	0.013963**	0.5302	0.000280**
S_06	0.1264	0.318368	0.3313	0.000815**
S_07	0.1688	0.005144**	0.2347	0.000001**
S_08	2.2465	0.000000**	2.5440	0.000000**

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

(8) WYJ station realm

The VIF results of the OLS regression for the WYJ station realm are shown in Table 6- 17. All the explanatory variables have VIF values below 10, indicating the absence of multicollinearity among the variables. Additionally, the adjusted R^2 for the model on working days is 0.6009, and on off days it is 0.6842, indicating a good fit of the model and suggesting that the selected combination of explanatory variables explains the dependent variable well.

Table 6- 17. The explanatory variables of the VIF of WYJ station realm

Code	VIF
W_01	1.038049
W_02	1.014184
W_03	1.032932
W_04	1.075704
W_05	1.371271
W_06	1.167010
W_07	1.379308
W_08	1.047104
W_09	1.021893

Table 6- 18 presents the OLS regression results. On working days, the regression coefficients for medical & health care services, catering services, and living services are positively significant. On off days, the regression coefficients for catering services and living services are positively significant. Therefore, the interactive relationship between population activity and functional facilities in the STSG station realm can be represented by the following equation:

$$W_W = 5.095837 + 2.351781 \times W02 + 0.5536 \times W04 + 0.791246 \times W05 \quad 6\text{- }15$$

$$W_O = 4.816953 + 0.959781 \times W04 + 0.742302 \times W05 \quad 6\text{- }16$$

Where W_W is the total population activity in WYJ station realm on working day, W_O is the total population activity on off day. W02 is the quantity of medical & health care facilities, S04 is the quantity of catering services, S05 is the quantity of living services in the WYJ station realm.

Table 6- 18. The OLS regression results for the WYJ station realm

Independent variable	Dependent variable			
	W_WD		W_OD	
	Coefficient	P-value	Coefficient	P-value
Intercept	5.095837	0.000000**	4.816953	0.000000**
W_01	2.351781	0.041279**	0.633504	0.559356

W_02	1.226950	0.335287	1.435845	0.231285
W_03	0.643173	0.320848	1.113414	0.068134
W_04	0.553600	0.022801**	0.959781	0.000034**
W_05	0.791245	0.022499**	0.742302	0.023038**
W_06	0.175639	0.500361	0.160098	0.514273
W_07	-0.822482	0.661984	0.228337	0.897448
W_08	0.665353	0.407944	0.829422	0.273444

** indicates that the coefficients are significantly correlated at the 5% statistical level ($p < 0.05$)

6.2.2 Significant variables in the regression results of the OLS model

(1) Under development type station realm

Based on the regression results using the Ordinary Least Squares (OLS) model, this study selected the service facilities with statistically significant regression coefficients in the YDGY station realm as variables to establish the Geographically Weighted Regression (GWR) model. The specific selection results are presented in Table 6- 19.

Table 6- 19. Significant variables in the regression results of the OLS model for the YDGY station realm

Station realm	Types of facility	WD	OD
YDGY	Living service	Y_02	Y_02
	Transportation service	Y_04	Y_04

(2) Urban center type station realm

Based on the regression results using the OLS model, this study selected the service facilities with statistically significant regression coefficients in the ZL station realm as variables to establish the GWR model. Specifically, for working days, the selected variables included medical and health care facilities, accommodation services, catering services, shopping services, and transportation services. For off days, the selected variables included accommodation services, catering services, living services, shopping services, and transportation services. The specific selection results are presented in Table 6- 20.

Table 6- 20. Significant variables in the regression results of the OLS model for the ZL station realm

Station realm	Types of facility	WD	OD
ZL	Medical & Health care	Z_03	——
	Accommodation service	Z_04	Z_04
	Catering service	Z_05	Z_05
	Living service	——	Z_06
	Shopping service	Z_07	Z_07
	Transportation service	Z_08	Z_08

(3) High population density Hub type station realm

Based on the regression results using the OLS model, this study selected the service facilities with statistically significant regression coefficients in the high population density hub station realms as variables to establish the GWR model. Specifically, in the NSM station realm, the selected variables included entertainment facilities, catering services, shopping services, and transportation services. In the XZ station realm, the selected variables included entertainment facilities, accommodation services, catering services, shopping services, and transportation services. The specific selection results are presented in Table 6- 21.

Table 6- 21. Significant variables in the regression results of the OLS model for the NSM and XZ station realm

Station realm	Types of facility	WD	OD
NSM	Entertainment	N_01	N_01
	Catering service	——	N_05
	Shopping service	N_07	N_07
	Transportation service	N_08	N_08
XZ	Entertainment	X_01	——
	Accommodation service	X_04	X_04
	Catering service	X_05	X_05
	Shopping service	X_07	X_07
	Transportation service	X_09	X_09

(4) Interchange type station realm

Based on the regression results using the OLS model, this study selected the service facilities with statistically significant regression coefficients in the Interchange type station realms as variables to establish the GWR model. Specifically, in the LSY station realm, the selected variables included financial services, accommodation services, catering services, shopping services, and transportation services. In the BDJ station realm, the selected variables included medical and health care services, accommodation services, catering services, and transportation services. The specific selection results are presented in Table 6- 22.

Table 6- 22. Significant variables in the regression results of the OLS model for the LSY and BDJ station realm

Station realm	Types of facility	WD	OD
LSY	Financial service	L_02	L_02
	Accommodation service	L_03	L_03
	Catering service	L_04	L_04
	Shopping service	L_06	L_06
	Transportation service	L_07	L_07
BDJ	Medical & Health care	B_01	B_01
	Accommodation service	——	B_02
	Catering service	B_03	B_03
	Transportation service	B_06	B_06

(5) General residential type station realm

Based on the regression results using the Ordinary Least Squares (OLS) model, this study selected the service facilities with statistically significant regression coefficients in the General residential type of station realms as variables to establish the Geographically Weighted Regression (GWR) model. Specifically, in the STSG station realm, the selected variables included entertainment, government agency, accommodation services, catering services, living service, shopping services, and transportation services. In the WYJ station realm, the selected variables included financial services, catering services, and living services. The specific selection results are presented in Table 6- 23.

Table 6- 23. Significant variables in the regression results of the OLS model for the STSG and WYJ station realm

Station realm	Types of facility	WD	OD
STSG	Entertainment	S_01	——
	Government agency	S_03	——
	Accommodation service	——	S_04
	Catering service	S_05	S_05
	Living service	——	S_06
	Shopping service	S_07	S_07
	Transportation service	S_08	S_08
WYJ	Financial service	W_01	——
	Catering service	W_04	W_04
	Living service	W_05	W_05

6.3 Regression results and analysis of the GWR model

6.3.1 Under development type station realm

The adjusted R^2 for the GWR (Geographically Weighted Regression) model established for the YDGY station realm on working days is 0.7084, and on off days it is 0.7406. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 24 provides the descriptive statistics of the GWR model regression results for the YDGY station realm. The regression coefficients for the living service facilities and transportation service facilities show both positive and negative values, indicating the presence of spatial heterogeneity in the interaction between these two types of functional facilities and population activity. By calculating the average values of the regression coefficients for each variable, we can determine the relative degree of interaction between functional facilities and population activity. It is found that, both on working days and off days, the degree of interaction is highest for transportation service facilities, followed by living service facilities. In terms of time, the level of interaction for both living service facilities and transportation service facilities decrease on off days compared to working days. Therefore, in the YDGY station realm, transportation service facilities emerge as the key variable driving the interaction between functional facilities and population activity.

On the spatial distribution pattern of the regression coefficients, the Figure 6- 1 illustrates this analysis, where darker colors represent higher positive regression coefficients between service facilities and population activity. Within the YDGY station realm, the regression coefficients for living service facilities exhibit a distribution pattern indicating lower levels of interaction with population activity as the distance from the station decreases, both on working days and off days. On the other hand, the regression coefficients for transportation service facilities display a decreasing pattern from the west to the east, suggesting that the western region of the YDGY station realm exhibits a higher level of interaction between population activity and transportation service facilities.

Table 6- 24. Descriptive statistics of GWR regression results for YDGY station realm

Date	Coefficient	Dependent variable	
		Y_02	Y_04
Y_WD	Mean	1.1457	1.4321
	ABS (mean)	1.472897	1.60303
	Max	8.5462	6.1396
	Min	-2.4056	-1.9627
Y_OD	Mean	0.9750	1.2252
	ABS (mean)	1.2598	1.5575
	Max	6.9807	6.7863
	Min	-2.3369	-2.9808

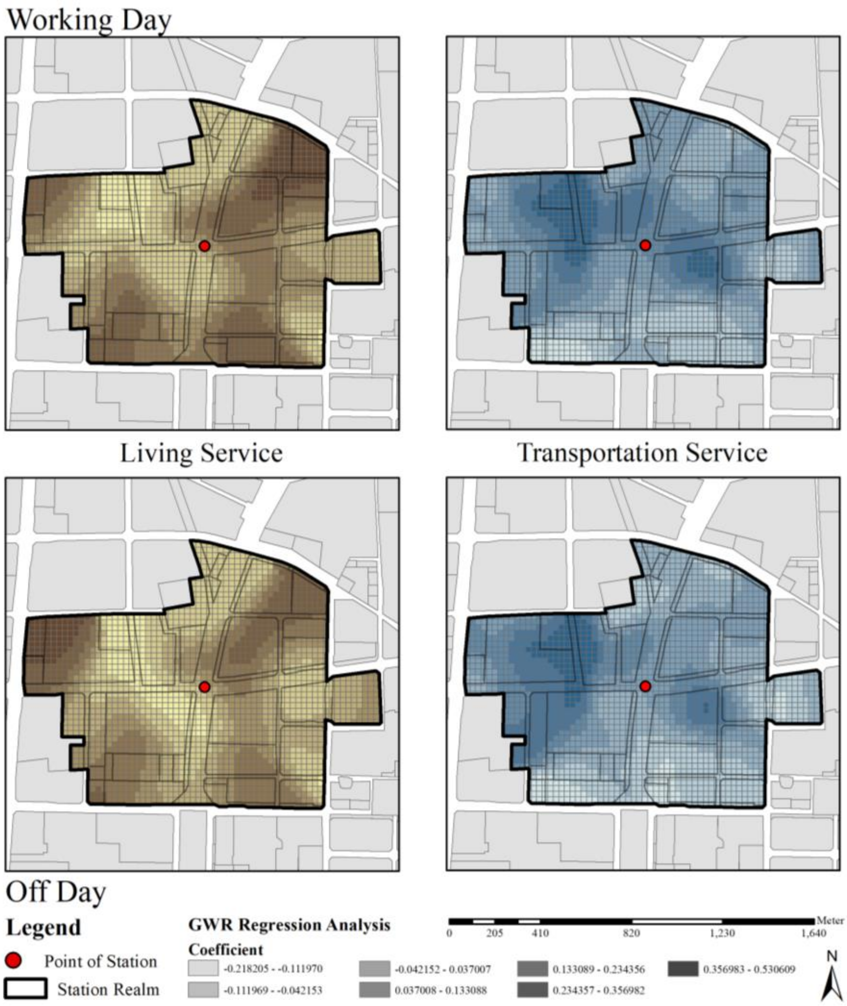


Figure 6- 2. Visualization of GWR regression results for YDGY station realm

6.3.2 Urban center type station realm

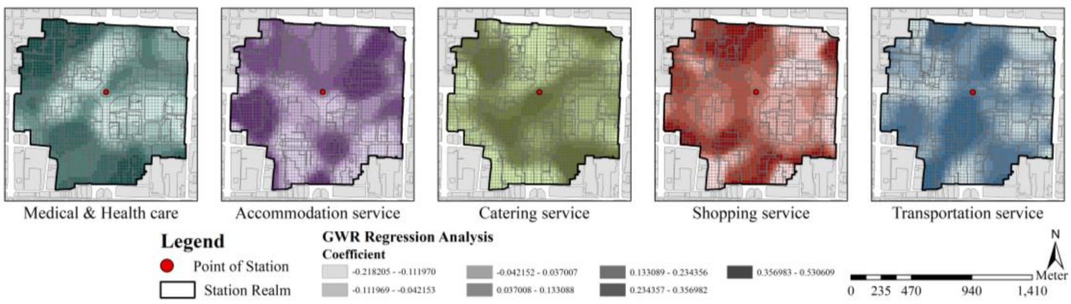
The adjusted R^2 for the GWR (Geographically Weighted Regression) model established for the ZL station realm on working days is 0.7325, and on off days it is 0.8307. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 25 provides the descriptive statistics of the GWR model regression results for the ZL station realm. It can be observed that the regression coefficients for the different types of service facilities exhibit both positive and negative signs, indicating the presence of spatial heterogeneity in the interaction between these six types of facilities and population activity within the station realm. Based on the average values of the regression coefficients, on working days, transportation service facilities show the highest level of interaction with population activity, followed by medical and health care facilities, while shopping service facilities exhibit the lowest level of interaction. On off days, the interaction level with transportation service facilities remains the highest, followed by accommodation service facilities, with shopping service facilities showing the lowest level of interaction. This suggests that transportation service facilities play a key role in driving the interaction with population activity within the ZL station realm. Furthermore, in terms of temporal variations, compared to working days, the interaction level between population activity and accommodation, catering, living, and shopping service facilities all show an increase, while the interaction level with transportation service facilities decreases.

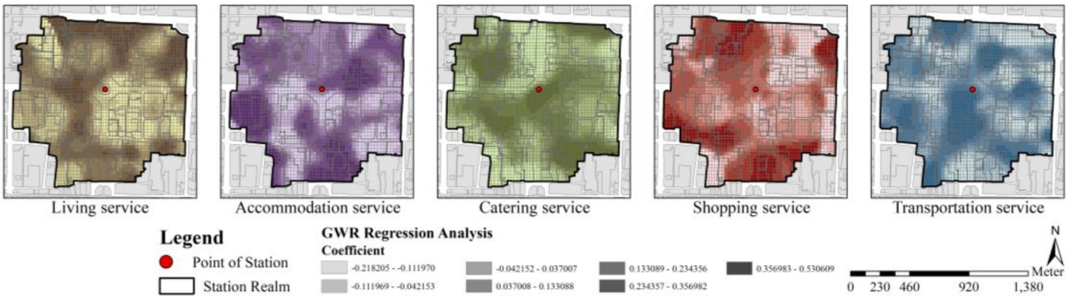
Table 6- 25. Descriptive statistics of GWR regression results for ZL station realm

Date		Dependent variable					
	Coefficient	Z_03	Z_04	Z_05	Z_06	Z_07	Z_08
Z_WD	Mean	2.2619	0.7721	1.5008	—	0.6255	4.6165
	Max	34.9524	14.9270	10.0985	—	5.9858	24.1159
	Min	-21.494	-14.3004	-6.1440	—	-9.3664	-11.1857
Z_OD	Mean	—	1.7874	1.6872	1.3995	1.0195	3.4260
	Max	—	18.0608	13.2798	24.2217	13.1310	27.3705
	Min	—	-12.2548	-9.0972	-10.0806	-11.5889	-13.8928

In terms of the spatial distribution patterns of the regression coefficients, as shown in Figure 6-3, there is evident spatial heterogeneity in the interaction levels between these six types of service facilities and population activity within the ZL station realm. Particularly, medical and health care facilities exhibit a spatial pattern where the interaction level with population activity decreases as the distance to the station decreases.



(a) Working day



(b) Off day

Figure 6- 3. Visualization of GWR regression results for ZL station realm

6.3.3 High population density Hub type station realm

(1) NSM station realm

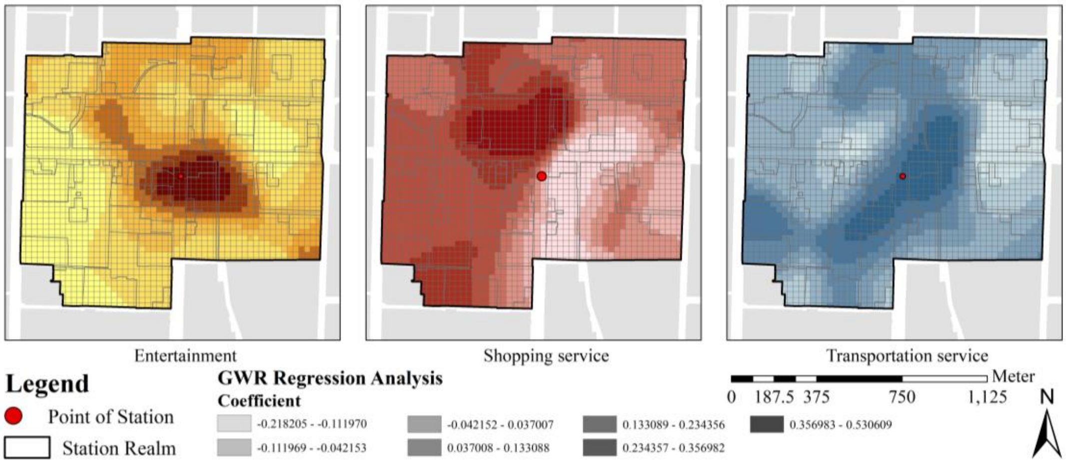
The adjusted R^2 for the GWR (Geographically Weighted Regression) model established for the NSM station realm on working days is 0.8199, and on off days it is 0.7964. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 26 presents the descriptive statistics of the GWR model regression results for the NSM station realm. It can be observed that the regression coefficients of various service facilities have both positive and negative signs, indicating the presence of spatial heterogeneity in the interaction between these four types of facilities and population activities within the station realm. In terms of the average values of the regression coefficients, on both working days and off days, entertainment facilities exhibit the highest positive interaction with population activities, followed by transportation services. On the other hand, shopping facilities generally exhibit a negative interaction with population activities. This suggests that entertainment facilities play a key role in the interaction between population activities and the NSM station realm. Furthermore, in terms of temporal variations, the degree of interaction for entertainment facilities and transportation services all decreased on the off days.

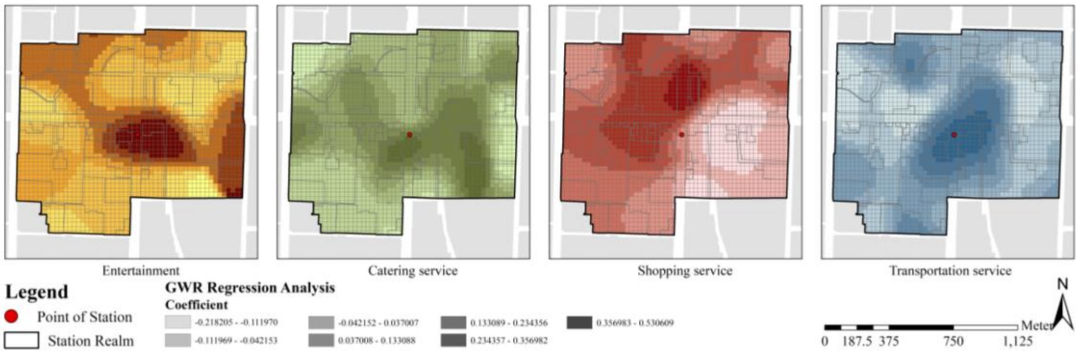
Table 6- 26. Descriptive statistics of GWR regression results for NSM station realm

Date		Dependent variable			
	Coefficient	N_01	N_05	N_07	N_08
N_WD	Mean	3.9363	——	-1.2283	1.9835
	Max	25.4423	——	8.8806	9.4760
	Min	-1.7846	——	-7.9873	-4.3218
N_OD	Mean	3.0701	0.5085	-0.8359	1.2234
	Max	17.4684	2.6082	9.1606	9.8493
	Min	-13.2922	-1.2989	-7.2485	-2.6223

In terms of the spatial distribution patterns of the regression coefficients, as shown in Figure 6- 4, the interaction levels of these four types of service facilities exhibit significant spatial heterogeneity within the NSM station realm. Specifically, entertainment facilities and transportation services facilities show a distribution pattern where the closer they are to the station, the higher interaction level with population activities.



(a)



(b)

Figure 6- 4. Visualization of GWR regression results for NSM station realm. (a) Working day. (b) Off day

(2) XZ station realm

The adjusted R^2 for the GWR (Geographically Weighted Regression) model established for the XZ station realm on working days is 0.8206, and on off days it is 0.8341. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 27 provides the descriptive statistics of the GWR model regression results for the XZ station realm. It can be observed that the regression coefficients of the different service facilities have both positive and negative signs, indicating the presence of spatial heterogeneity in the interaction between these facilities and population activities within the station realm. By examining the average values of the regression coefficients, on working days, accommodation services have the highest positive interaction level with population activities, followed by transportation services. On the other hand, the interaction levels of shopping services and catering services with population activities are relatively lower. On off days, transportation services have the highest positive interaction level with population activities, followed by accommodation services. This suggests that in the XZ station realm, accommodation services play a key role in the interaction with population activities on working days, while transportation services are crucial for the interaction on off days. Furthermore, in terms of temporal variations, the interaction levels between accommodation

services, shopping services, and transportation services with population activities increase on off days. Entertainment facilities only exhibit a positive interaction with population activities on off days, while the impact of catering services remains relatively stable.

Table 6- 27. Descriptive statistics of GWR regression results for XZ station realm

Date		Dependent variable				
	Coefficient	X_01	X_04	X_05	X_07	X_09
X_WD	Mean	1.1428	1.8048	0.5564	0.4043	1.7727
	Max	17.0751	40.8892	7.7080	4.6148	11.5254
	Min	-27.2881	-3.4334	-1.3878	-0.3628	-3.9562
X_OD	Mean	——	2.3137	0.5052	0.6243	2.4404
	Max	——	40.2463	5.0880	4.4112	15.6221
	Min	——	-6.8033	-1.4666	-0.8131	-3.7951

Regarding the spatial distribution pattern of regression coefficients, as shown in Figure 6- 5, there is evident spatial heterogeneity in the degree of interaction between these five types of service facilities and population activities within the XZ station realm. Particularly, entertainment facilities exhibit a distribution pattern where the closer they are to the station, the lower their interaction level with population activities. This suggests that areas closer to the station have relatively less interaction between entertainment facilities and population activities. Additionally, the interaction level of shopping facilities shows a decreasing distribution pattern from the south to the north within the station realm. This means that the southern region has a higher interaction level between shopping facilities and population activities compared to the northern region. This spatial distribution pattern may be influenced by factors such as population distribution, road networks, and the layout of commercial development within the XZ station realm.

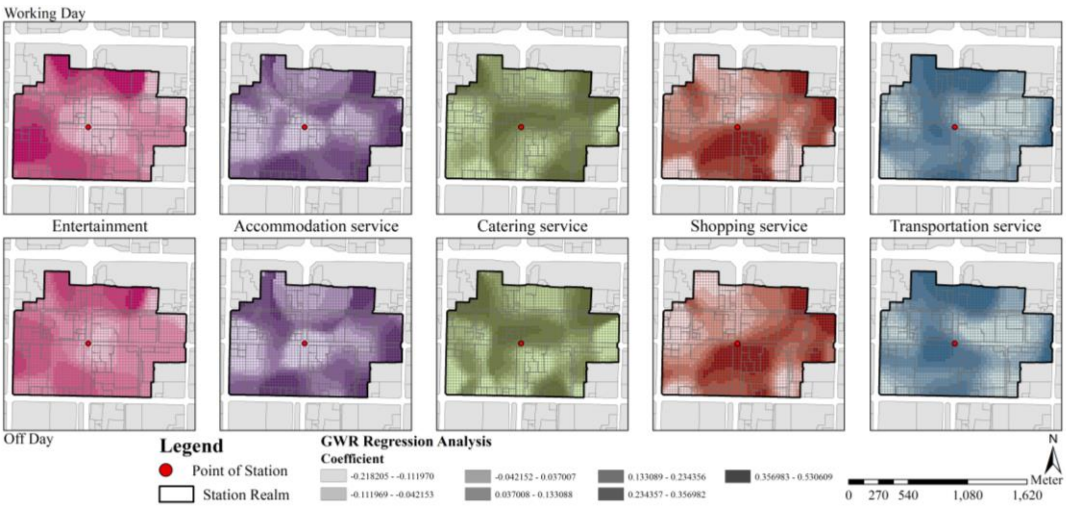


Figure 6- 5. Visualization of GWR regression results for XZ station realm

6.3.4 Interchange type station realm

(1) LSY station realm

The adjusted R^2 for the GWR (Geographically Weighted Regression) model established for the LSY station realm on working days is 0.7485, and on off days it is 0.7795. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 28 displays the descriptive statistics of the GWR model regression results for the LSY station realm. It can be observed that the regression coefficients of various service facilities have both positive and negative signs, indicating spatial heterogeneity in the interaction between these five types of service facilities and population activities within the station realm. Looking at the average values of the regression coefficients, financial services exhibit the highest positive interaction level with population activities on both working days and off days, followed by transportation services. On the other hand, the interaction levels of catering services and shopping services with population activities are relatively lower.

This implies that financial services play a dominant role in the interaction between functional facilities and population activities within the LSY station realm. Furthermore, in terms of temporal variation, the interaction levels of shopping services and transportation services with population activities increase on off days, while financial services and accommodation services show a slight decrease. The interaction level of catering services remains relatively stable with minor changes.

Table 6- 28. Descriptive statistics of GWR regression results for LSY station realm

Date		Dependent variable				
	Coefficient	L_02	L_03	L_04	L_06	L_07
L_WD	Mean	2.5414	1.2116	0.5069	0.4201	1.4944
	Max	9.9361	4.6161	1.2803	1.7116	5.0358
	Min	-2.3499	-1.5475	-0.0370	-1.0844	-0.9012
L_OD	Mean	2.4922	1.1010	0.5957	0.6415	1.6120
	Max	9.4623	5.0498	1.5794	1.6082	5.3814
	Min	-2.6325	-2.3505	-0.4742	0.0605	-1.3402

Regarding the spatial distribution pattern of the regression coefficients, as shown in Figure 6- 6, there is evident spatial heterogeneity in the interaction levels between these five types of service facilities and population activities within the LSY station realm.

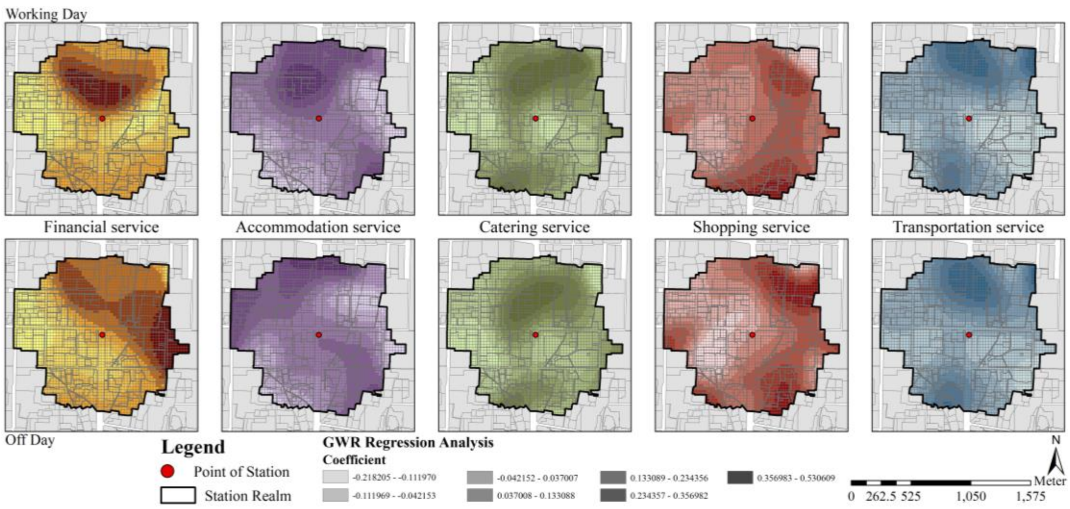


Figure 6- 6. Visualization of GWR regression results for LSY station realm

(2) BDJ station realm

The adjusted R^2 for the GWR (Geographically Weighted Regression) model established for the BDJ station realm on working days is 0.7232, and on off days it is 0.6633. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 29 presents the descriptive statistics of the GWR model regression results for the BDJ station realm. It can be observed that the regression coefficients of various service facilities have both positive and negative signs, indicating the presence of spatial heterogeneity in the interaction between these four types of service facilities and population activities within the station realm.

From the average values of the regression coefficients, on working days, the interaction between transportation services and population activities is the highest, followed by medical and health care facilities, while the interaction between catering services and population activities is the lowest. On off days, the interaction between medical and health facilities and population activities is the highest, followed by transportation services, while the interaction between catering services and population activities is the lowest.

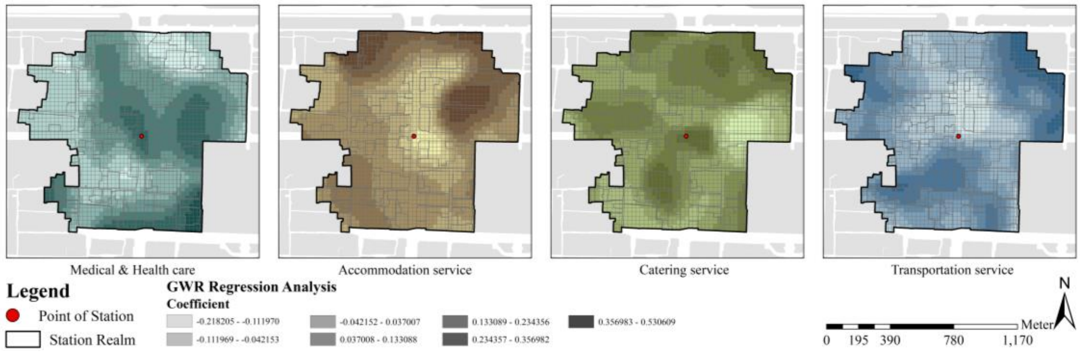
This suggests that transportation services are the key variable dominating the interaction between functional facilities and population activities on working days in the BDJ station realm, while medical and health care facilities play a crucial role in the interaction on off days. Furthermore, in terms of temporal variations, apart from transportation services, the interaction levels between medical and health care facilities, accommodation services, and catering services with population activities all increase on off days.

Table 6- 29. Descriptive statistics of GWR regression results for BDJ station realm

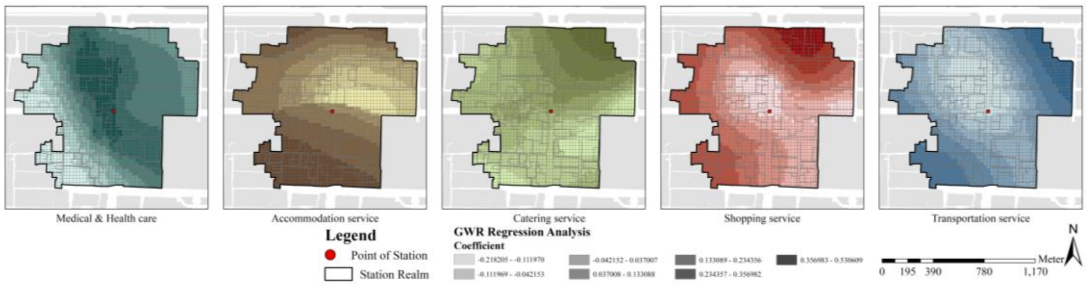
Date	Coefficient	Dependent variable			
		B_01	B_02	B_03	B_06
B_WD	Mean	1.0022	——	0.5480	1.2472
	Max	10.9586	——	2.7195	6.6855
	Min	-3.5232	——	-4.6841	-2.7819
B_OD	Mean	1.6169	0.8330	0.6174	1.0045
	Max	2.7574	2.9115	1.5582	3.4696
	Min	0.1444	-2.6352	-0.1456	-0.5390

Regarding the spatial distribution pattern of the regression coefficients, as shown in Figure 6- 7, there is evident spatial heterogeneity in the interaction levels between these five types of service facilities and population activities within the BDJ station realm.

In addition, there are significant differences in the spatial distribution patterns of the regression coefficients for various service facilities between working days and off days in the BDJ station realm. This indicates that there is greater population mobility in the BDJ station realm, leading to distinct variations in the interaction between population activities and functional facilities at different times. The differences may be attributed to the varying activity demands and behavioral patterns of people on working days and off days, resulting in different levels of interaction between different types of service facilities and population activities. These variations may be influenced by factors such as population mobility, consumer habits, and social culture on working days and off days. This finding emphasizes the importance of considering the characteristics of population activities at different time periods in the planning and allocation of functional facilities in station realms to accommodate population mobility and changing demands.



(a)



(b)

Figure 6- 7. Visualization of GWR regression results for BDJ station realm. (a) Working day. (b) Off day

6.3.5 General residential type station realm

(1) STSG station realm

The adjusted R^2 for the GWR model established for the STSG station realm on working days is 0.6852, and on off days it is 0.7904. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

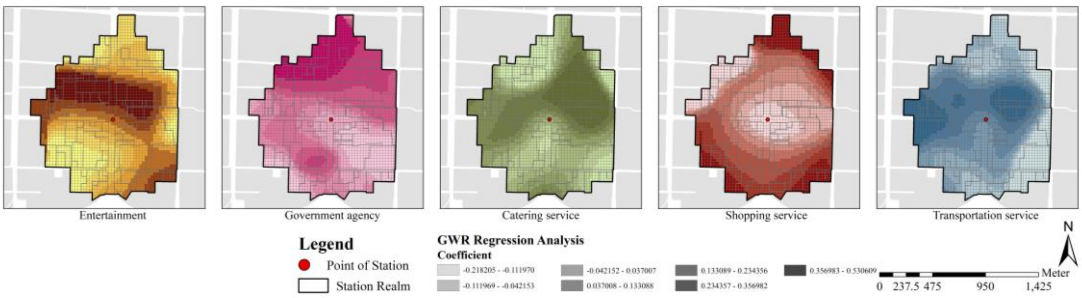
Table 6- 30 presents the descriptive statistics of the GWR model regression results for the STSG station realm. It can be observed that the regression coefficients for various service facilities have both positive and negative signs, indicating the existence of spatial heterogeneity in the interaction between these seven types of service facilities and population activities within the station realm. Based on the average values of the regression coefficients, on working days, entertainment facilities exhibit the highest positive interaction with population activities, followed by transportation services, while catering and shopping facilities show relatively lower levels of interaction. On off days, accommodation services demonstrate the highest positive interaction with population activities, followed by transportation services, while catering facilities exhibit the lowest level of interaction. These findings suggest that entertainment facilities play a crucial role in the interaction between functional facilities and population activities on working days in the STSG station realm, whereas accommodation services are key variables in this interaction on off days. Furthermore, in terms of temporal variation, the interaction levels between population activities and accommodation

services, living services, and shopping services all increase on off days.

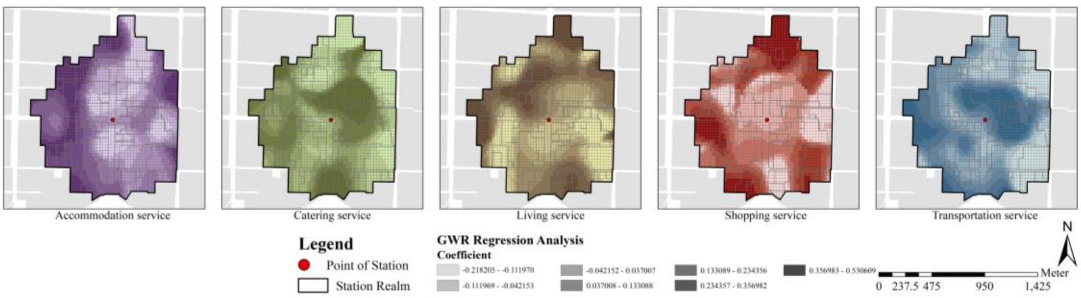
Table 6- 30. Descriptive statistics of GWR regression results for STSG station realm

Date		Dependent variable						
	Coefficient	S_01	S_03	S_04	S_05	S_06	S_07	S_08
S_WD	Mean	2.6136	0.9897	——	0.4276	——	0.4490	1.5636
	Max	8.6523	13.5439	——	1.7617	——	3.1715	5.5497
	Min	-1.4672	-3.6593	——	-0.9706	——	-0.6659	-4.2359
S_OD	Mean	——	——	1.3746	-0.1987	0.7288	0.8312	1.1176
	Max	——	——	11.4023	3.6774	18.4312	17.4094	5.5886
	Min	——	——	-9.0509	-47.7611	-1.6956	-1.2238	-7.7288

Regarding the spatial distribution pattern of the regression coefficients, as shown in Figure 6-8, there is evident spatial heterogeneity in the interaction levels between these five types of service facilities and population activities within the STSG station realm. Particularly, shopping facilities exhibit a distribution pattern where the closer they are to the station, the lower their interaction level with population activities. This suggests that areas closer to the station have relatively less interaction between shopping facilities and population activities.



(a)



(b)

Figure 6- 8. Visualization of GWR regression results for STSG station realm. (a) Working day. (b) Off day

(8) WYJ station realm

The adjusted R^2 for the GWR model established for the WYJ station realm on working days is 0.7477, and on off days it is 0.7230. These values indicate a good fit of the model, suggesting that the selected combination of explanatory variables has a strong explanatory power for the dependent variable.

Table 6- 31 presents the descriptive statistics of the GWR model regression results for the WYJ station realm. The regression coefficients for various types of service facilities exhibit both positive and negative signs, indicating the presence of spatial heterogeneity in the interaction between these three types of service facilities and population activities within the station realm. Based on the average values of the regression coefficients, it is evident that living services demonstrate the highest positive interaction with population activities on both working days and off days, followed by catering services, while financial services exhibit the lowest level of interaction. This suggests that transportation services play a dominant role in the interaction between functional facilities and population activities in the WYJ station realm. Additionally, in terms of temporal variation, the correlation between catering services and population activities increases on off days, while living services show minimal changes.

Table 6- 31. Descriptive statistics of GWR regression results for WYJ station realm

Date		Dependent variable		
	Coefficient	W_01	W_04	W_05
W_WD	Mean	-0.4334	0.7389	1.2046
	Max	10.2878	12.2594	11.3001
	Min	-8.1581	-2.7757	-2.1018
W_OD	Mean	——	1.0688	1.21305
	Max	——	10.0578	13.5295
	Min	——	-1.8956	-2.1889

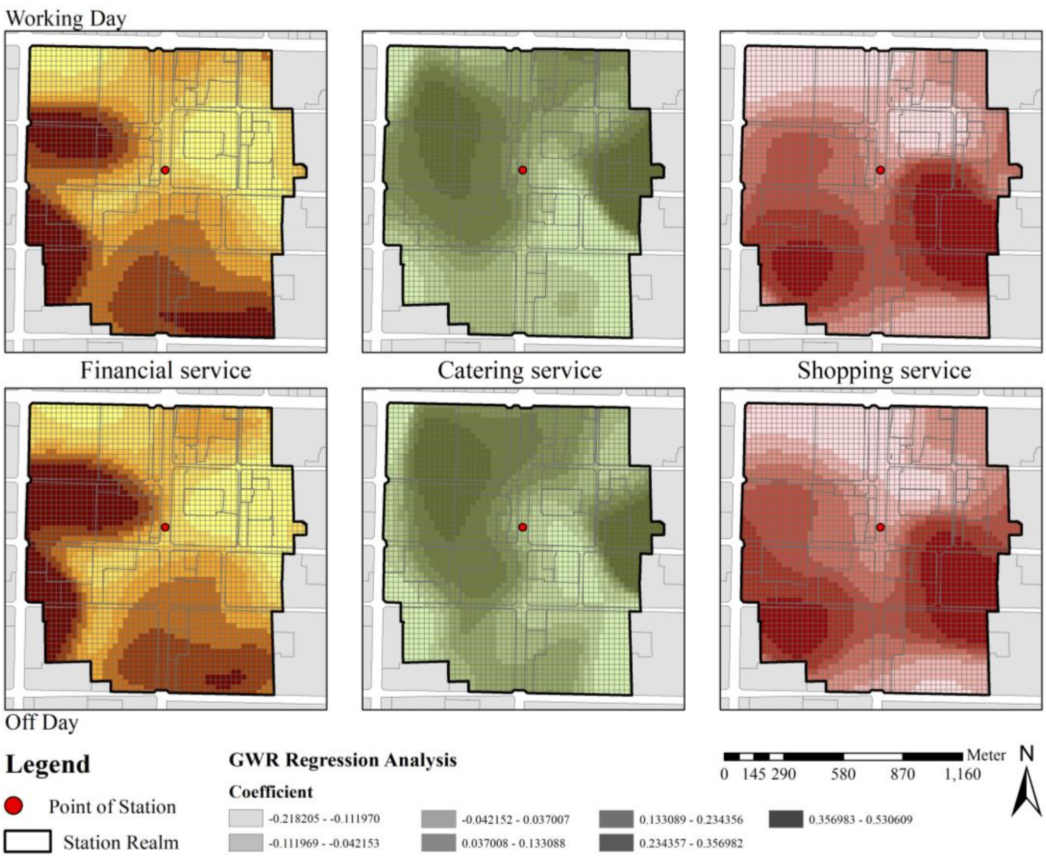


Figure 6- 9. Visualization of GWR regression results for WYJ station realm

Regarding the spatial distribution pattern of the regression coefficients, as shown in Figure 6-9, there is evident spatial heterogeneity in the interaction levels between these three types of service facilities and population activities within the WYJ station realm. These three service facilities in the WYJ station realm do not exhibit a distribution pattern where their correlation with population activities increases as they are closer to the station. Instead, these facilities are dispersed in a clustered shape within the WYJ station realm.

6.4 Discussion

6.4.1 Model performance comparisons

Although the OLS and GWR models provide similar regression results in terms of the overall characteristics of the interaction relationships, it is necessary to compare their model performance to quantify the degree of interaction between population activity and functional facilities in each station realm. As shown in Table 6- 32, comparing the Adjusted R^2 and AIcc of the models reveals that the GWR models have higher Adjusted R^2 values and lower AIcc values compared to the OLS model. This indicates that the GWR models have better fit and performance than the OLS model, demonstrating their superior explanatory power in describing the spatial distribution characteristics of micro-level station realm interaction relationships. The GWR models highlight the spatial aspects of the data, which is consistent with the findings of many scholars in current research [3,4,5].

However, the GWR model cannot be used alone due to its inability to diagnose multicollinearity among explanatory variables. The fundamental reason for this is that the GWR model does not address the issue of multicollinearity, whereas the global OLS model does. Additionally, any explanatory variables that are not significant in the OLS regression should be removed before using the GWR model to avoid introducing errors in the regression results. Therefore, in this study, the OLS regression results serve as a reference, while the GWR regression results serve as the standard for evaluating the degree of interaction between population activity and functional facilities in each station realms.

Table 6- 32. Performance Comparison of OLS Model and GWR Model

Station realm	OLS model				GWR model			
	WD		OD		WD		OD	
	Adjusted R^2	AIcc	Adjusted R^2	AIcc	Adjusted R^2	AIcc	Adjusted R^2	AIcc
YGGY	0.6810	1252	0.7140	1954	0.7084	785	0.7406	703
ZL	0.6341	1501	0.7139	1610	0.7325	627	0.8307	650
NSM	0.6141	1308	0.8199	1141	0.8199	979	0.7964	782
XZ	0.7871	1657	0.7630	1783	0.8206	969	0.8341	722
LSY	0.7378	1575	0.6932	1677	0.7485	659	0.7795	634
BDJ	0.6308	1355	0.6147	1657	0.7232	698	0.6633	842
STSG	0.6599	1514	0.7234	1396	0.6852	613	0.7904	439
WYJ	0.6009	1962	0.6842	1632	0.7417	967	0.7230	941

6.4.2 Establishment and application of interaction sequences

Overall, based on the mean regression coefficients, we obtained a sequence of the degree of interaction between population activities and various service facilities in Table 6-6, and the interaction mechanism of each station realm was finally established. Therefore, the optimization and allocation of public resources according to the sequence of interaction relations can effectively fulfill people's demands and thus achieve the matching of supply and demand in station realm.

Table 6- 33. The interaction sequence of each experimental station realm

Station realm	Working day	Off day
YDGY	F10 > F07	F10 > F07
ZL	F10 > F04 > F06 > F05 > F08	F10 > F05 > F06 > F07 > F08
NSM	F01 > F10 > F08	F01 > F10 > F06 > F08
XZ	F05 > F10 > F01 > F06 > F08	F10 > F05 > F08 > F06
LSY	F02 > F10 > F05 > F06 > F08	F02 > F10 > F05 > F08 > F06
BDJ	F10 > F04 > F06	F04 > F10 > F05 > F06
STSG	F01 > F10 > F03 > F08 > F06	F05 > F10 > F08 > F07 > F06
WYJ	F07 > F06 > F02	F07 > F06

Based on the study of the interaction relationships discussed earlier, it is evident that there exist complex mechanisms between population activity and functional facilities in station realms, involving various aspects such as time, type, scale, and space. To better understand the interaction mechanisms between population activity and functional facilities in station realms, this study adopts the concept of "supply and demand" from microeconomics to explore the "supply-demand relationship" between them. By analyzing the operational principles, a better grasp of the spatial functioning of station realms can be achieved, providing theoretical support for spatial development recommendations for different types of station realms.

In the context of rail transit station realms, the demand refers to the distribution of population activity in a certain period on different types of service facilities within the station realm. On the supply side, it refers to the configuration and distribution of various types of functional facilities. The supply-demand relationship, therefore, refers to the degree of interaction between these two aspects. Thus, the supply-demand relationship between population activity and functional facilities in station realms can be further explained as follows: different types of service facilities that positively interact with population activity within different station realms will generate different levels of population activity occurrence and attraction. Conversely, the various demands generated by population activities for different travel purposes can further adjust the configuration of functional facilities. Additionally, the research findings of this study also indicate that the supply-demand relationship between population activity and functional facilities tends to exhibit dynamic characteristics, which is an important component of the station realm system. The supply-demand relationship is depicted in Figure 6- 10.

In summary, addressing the supply of various functional facilities in station realms is not simply a matter of adding more facilities based on the quantity of population activity. Instead, it should be demand-driven, considering the diverse activity needs generated by the population at different times. This approach allows for the planning of the appropriate quantity and distribution of functional facilities, leading to more rational allocation of public resources and improved utilization efficiency of station realm spaces.

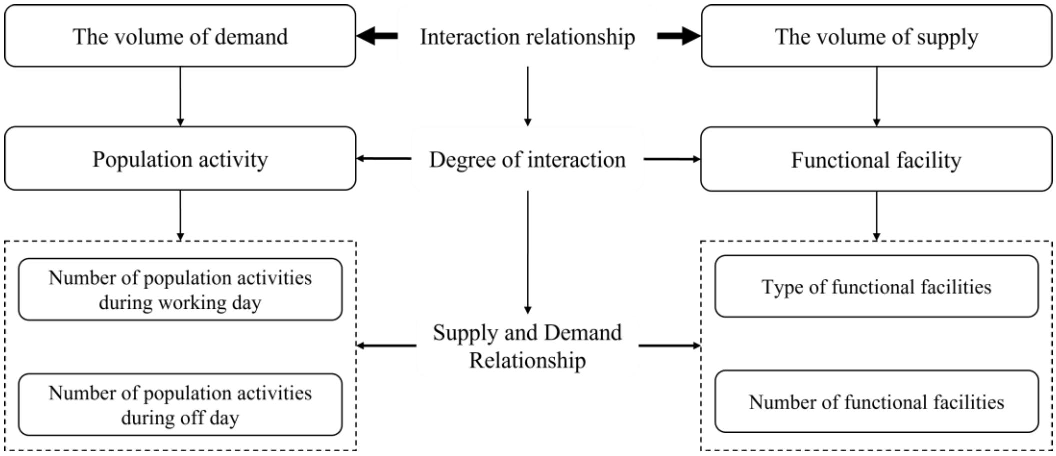


Figure 6- 10. Supply and demand relationship between population activities and functional facilities in station realm

6.4.3 The variability of interaction mechanisms at time dimension

The regression results from different types of station realms indicate that the degree of positive interaction between population activities and various service facilities varies between working days and off days. For example, in the STSG station realm on working days, there is no interaction between population activities and accommodation services or living services. However, on off days, there is a positive interaction between population activities and these two types of facilities, especially with a high degree of interaction with accommodation services.

In the YDGY station realm, the interaction between transportation services and living services with population activities is higher on working days than on off days. Based on empirical investigations, developing station realms often lack sufficient functional facilities to meet residents' various daily needs. Therefore, on off days, residents intentionally travel to city centers or well-established areas to engage in various activities, leading to increased usage of transportation services and higher levels of interaction between transportation services and population activities.

In the urban center type of ZL station realm, the interaction between accommodation services and shopping services with population activities increases on off days, while the interaction between transportation services and population activities decreases. Based on empirical investigations, popular tourist attractions such as the Bell Tower, Drum Tower, Muslim Quarter, and Beilin Museum exist in the ZL station realm. Therefore, on off days, the increase in the number of tourists promotes a higher degree of interaction between accommodation services, shopping services, and population activities. Since tourists typically explore and visit these attractions on foot, the interaction between transportation services and population activities decreases.

In the high population density hub type of station realms, specifically in the XZ station realm, the interaction between entertainment, accommodation services, shopping services, and transportation services with population activities increases on off days. This suggests an increase in the number of population activities utilizing these four types of facilities on off days. Through empirical investigations and interpretation of urban planning, the Xiazhai station realm has become the central commercial area in the southern part of Xi'an. Moreover, it is home to numerous tourist attractions, universities, and large commercial complexes. Therefore, on off days, the Xiazhai station realm becomes the busiest commercial area in the southern part of Xi'an.

Additionally, the study found that the key variables influencing the positive interaction in the station realms also vary between working days and off days. As shown in Table 6- 34, in the ZL station realm, transportation services are the key variables on working days, while accommodation services become the key variables on off days. In the XZ station realm, accommodation services are the key variables on working days, while transportation services become the key variables on off days. In the BDJ station realm, the key variables shift from transportation services on working days to medical and health care services on off days. This observation indicates that the degree of interaction between population activities and various functional facilities in different types of station realms varies over time. Moreover, even within the same type, the key variables may change. This finding emphasizes the importance of considering the characteristics of population activities during different time periods in planning and allocating functional facilities in station realms to adapt to the changing population mobility and demands.

Table 6- 34. Key variables of interaction relationships within different types of station realms

Type of station realm	Case	Working day	Off day
Under development type	YDGY	Transportation	Transportation
Urban center type	ZL	Transportation	Accommodation
High population density Hub type	NSM	Entertainment	Entertainment
	XZ	Accommodation	Transportation
Interchange type	LSY	Financial	Transportation
	BDJ	Transportation	Medical & Health care
General residential type	STSG	Entertainment	Accommodation
	WYJ	Living	Living

Therefore, the research results contribute to urban managers' understanding of the most urgent service facility needs in people's daily activities within different types of station realms from a dynamic perspective. This understanding enables the formulation of more targeted planning measures and more effective improvements in the allocation of urban public resources. Specifically, driven by the key variables, efforts can be made to develop and attract populations by focusing on the development of these key variables. As the number of population activities increases, the demand for urban functional facilities becomes more diverse, thereby driving the development of other service facilities until a functionally mixed-use station realm is formed.

6.4.4 The variability of interaction mechanisms at category dimension

Based on the regression results of the GWR model, the study summarized the frequency of occurrence of service facility types that have a positive interaction with population activities within different types of station realms, as shown in the Table 6- 35. Among the five categories and eight station realms, transportation services have the highest frequency of positive interaction with population activities (14 occurrences), followed by catering services (12 occurrences) and shopping services (11 occurrences), while government agencies have the lowest frequency (1 occurrence). This indicates that within station realms, transportation services, catering services, and shopping services are the most frequently used urban functional facilities in people's daily lives, which is consistent with the findings of previous related research [6,7].

Table 6- 35. Service facilities with positive interaction with population activities in different types of station realms

Type of station realm	Case	Positive interactive functional facilities: Working day	Positive interactive functional facilities: Off day
Under development type	YDGY	Living service Transportation service	Living service Transportation service
Urban center type	ZL	Medical & Health care Accommodation service Catering service Shopping service Transportation service	Accommodation service Catering service Shopping service Living service Transportation service
High population density Hub type	NSM	Entertainment Shopping service Transportation service	Entertainment Catering service Shopping service Transportation service
	XZ	Entertainment Accommodation service Catering service Shopping service Transportation service	Accommodation service Catering service Shopping service Transportation service
Interchange type	LSY	Financial service Accommodation service Catering service Shopping service Transportation service	Accommodation service Catering service Shopping service Transportation service
	BDJ	Medical & Health care Accommodation service Catering service Transportation service	Medical & Health care Accommodation service Catering service Shopping service Transportation service
General residential type	STSG	Financial service Government Agency Catering service Shopping service Transportation service	Accommodation service Living service Shopping service Transportation service
	WYJ	Catering service Living service	Catering service Living service

Specifically, transportation services have become the key variables influencing the interaction between population activities in the YDGY station realm, ZL station realm, XZ station realm, LSY station realm, and BDJ station realm. Through empirical investigations, it was found that the number of transportation services is generally higher in these five station realms. For example, BDJ station is a transfer station between Metro Line 2 and Line 1, ZL station is a transfer station between Metro Line 2 and Line 6, and XZ station is a transfer station between Metro Line 2 and Line 3. Therefore, in people's daily lives, the metro has become an indispensable means of transportation, and the convenient transfer between stations and other public transportation can effectively improve the intra-station realm accessibility, attracting more population to engage in activities in these areas.

In terms of the spatial pattern of regression coefficients, as shown in Figure 6- 11, only transportation facilities exhibit a distribution pattern where the closer they are to the station, the higher their interaction with population activities, such as ZL station realm, NSM station realm, and STSG station realm. On the other hand, the regression coefficients of other service facility types show significant spatial heterogeneity within the station realm. This finding further emphasizes the importance of transportation services in the planning and development of urban rail transit station realms, particularly in terms of convenient connections between stations and various transportation modes, as well as internal accessibility within the station realm.

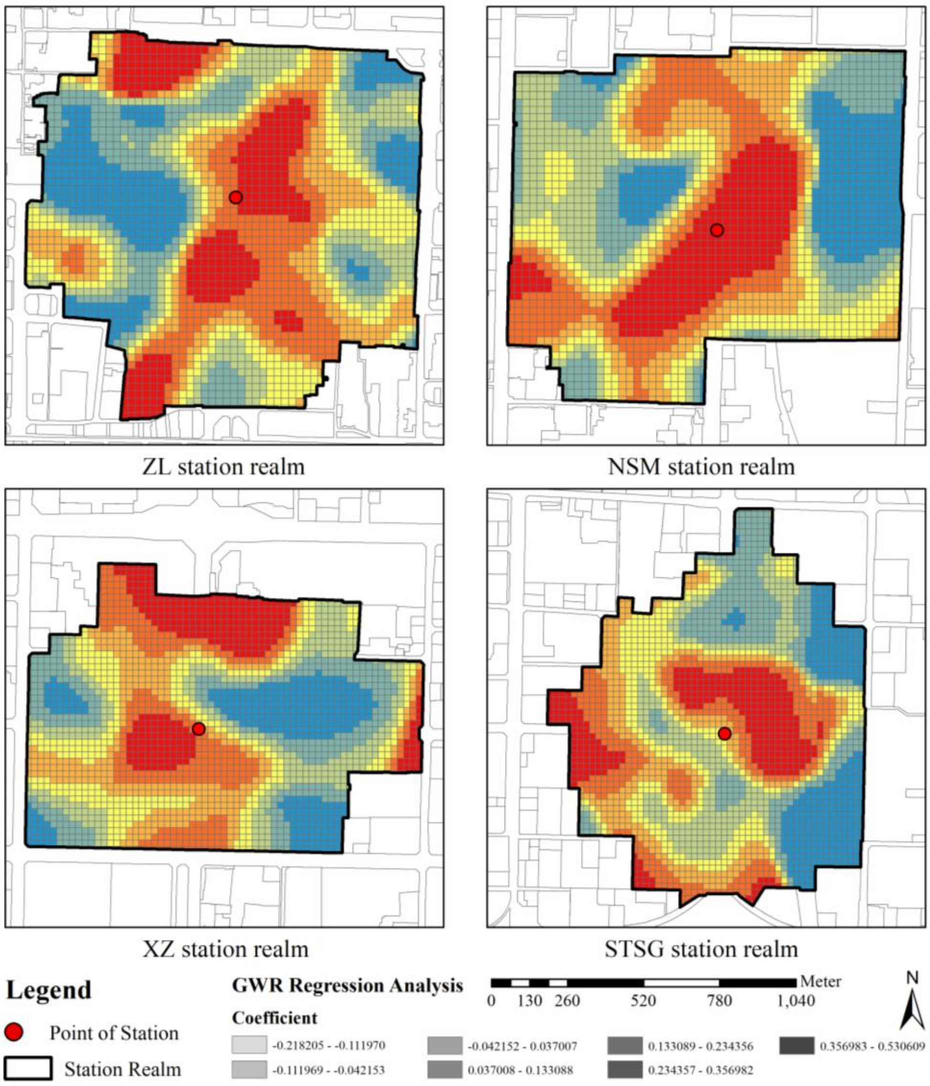


Figure 6- 11. Distribution patterns of regression coefficients of transportation service facilities in different types of station realms

Therefore, in the current context of urban traffic congestion and limited land resources, improving transportation services can effectively strengthen the connections between different types of service facilities within the station realm. This, in turn, can meet the diverse activity needs of people and enhance the utilization of urban functions. By prioritizing the enhancement of transportation services, urban planners and policymakers can create a more integrated and efficient station realm, ensuring seamless mobility and promoting sustainable urban development.

It is worth noting that medical and health care facilities show a high level of positive interaction with population activities only in the BDJ station realm. Through field investigations, it was found that the BDJ station realm is home to several large public hospitals, as well as numerous medical support facilities and pharmacies. In the current environment, there is an increasing demand for quality healthcare environments and medical services, which leads people to gather in station realms with relatively good medical resources. The rapid transportation and high capacity of the urban rail transit system further exacerbate this phenomenon, resulting in increased burden on hospitals and reduced medical experience for patients. Balancing the allocation of public healthcare resources in station realms and addressing the social equity issues arising from this situation should be of concern to urban managers and researchers.

6.4.5 The variability of interaction mechanisms at scale dimension

In previous studies, most scholars have taken a macro perspective, assuming a positive correlation between the quantity of functional facilities and population activities [8,9]. That is, the higher the quantity of a particular type of facility, the greater its interaction with population activities, and consequently, the higher the number of people using that facility. However, this study conducted a comparative analysis of the quantity of functional facilities and their regression coefficients within each station realm. The findings revealed that a higher quantity of facilities within a station realm does not necessarily indicate a higher level of interaction with population activities.

For example, as shown in Table 6- 36, although there is a positive interaction between catering services as well as shopping facilities with population activities in most station realms, and their quantity is usually higher than other service facilities, the degree of interaction they generate is relatively low. Therefore, for urban managers, solely focusing on the development of traditional catering businesses and shopping services is no longer sufficient to effectively enhance the population vitality and regional economy of station realms, particularly considering the rapid growth of the catering delivery and online shopping industries.

Furthermore, the regression coefficients of the same type of service facility may vary between working days and off days. This indicates that at a more micro level of the station realm, the interaction between population activities and functional facilities is often dynamic. Moreover, in practical planning and development, simply pursuing an increase in the quantity of functional facilities is not an effective way to promote population activities. On the contrary, it may lead to significant waste of public resources and reduced spatial efficiency within the station realm. This is particularly evident in centrally located station realms within land-constrained areas and densely populated hub station realms.

Therefore, when formulating planning and development measures for station realms, it is important to consider the interactive mechanism between population activities and functional facilities within each type of station realm. This approach ensures the coordinated development of station realm planning and construction while meeting the diverse activity needs of people.

Table 6- 36. Comparison of the number of service facilities and their degree of interaction in different types of station realms

Station realm	Types of facility	Number of facilities	Regression coefficient (GWR)	
			Working day	Off day
ZL	Medical & Health care	105	0.7395	-0.0648
	Accommodation service	583	1.0393	1.4718
	Catering service	1293	0.4505	0.4929
	Shopping service	1634	0.3030	0.5526
	Transportation service	196	1.3868	0.8882
XZ	Entertainment	165	1.1428	1.2706
	Accommodation service	202	1.8048	2.3137
	Catering service	1299	0.5564	0.5052
	Shopping service	1549	0.4043	0.6243
	Transportation service	149	1.7727	2.4404
BDJ	Medical & Health care	211	0.9821	1.5783
	Accommodation service	153	0.7486	0.7245
	Catering service	781	0.5146	0.4959
	Shopping service	1080	——	0.2409
	Transportation service	158	1.2636	0.9067
WYJ	Financial service	36	-0.4334	-0.3145
	Catering service	570	0.7389	1.0688
	Living service	332	1.2046	1.21305

6.5 Chapter summary

This chapter provides a comprehensive analysis of the vertical and horizontal interaction principles and magnitudes between population activity and key elements of functional facilities in different types of station realms along the Xi'an Metro Line 2. Based on the findings of the correlation analysis in Chapter 5, an OLS model was established to comprehensively study the interaction relationship between population activity and functional facilities in the station realms. Significant explanatory variables were identified and selected from each station realm. Subsequently, a GWR model was constructed to further analyze the level of interaction between population activity and various types of functional facilities within different station realms. The GWR model helped identify the key variables that dominate the interaction relationships. Lastly, the chapter explored the differences in interaction relationships among various station realms at different levels. Based on the findings, recommendations were made for the spatial development direction of future station realms.

By conducting these analyses, this chapter provides insights into the interaction mechanisms between population activity and functional facilities in different types of station realms along Line 2, allowing for a deeper understanding of the dynamics and spatial patterns of these interactions. The research outcomes also offer valuable suggestions for the future spatial development of station realms.

References

[1] James G, Witten D, Hastie T, et al. An introduction to statistical learning[M]. New York: springer, 2013.

[2] Pohlman J T, Leitner D W. A comparison of ordinary least squares and logistic regression[J]. 2003.

[3] Sassi M. OLS and GWR approaches to agricultural convergence in the EU-15[J]. International Advances in Economic Research, 2010, 16: 96-108.

[4] Ma X, Zhang J, Ding C, et al. A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership[J]. Computers, Environment and Urban Systems, 2018, 70: 113-124.

[5] R. C., & Nakada, L. Y. K. GIS-based spatial modelling of COVID-19 death incidence in São Paulo, Brazil. Environment and Urbanization, 2021,33(1), 229–238.

[6] Zhang, X.; Sun, Y.; Chan, T.O.; Huang, Y.; Zheng, A.; Liu, Z. Exploring Impact of Surrounding Service Facilities on Urban Vibrancy Using Tencent Location-Aware Data: A Case of Guangzhou. Sustainability 2021, 13, 444. DOI:10.3390/su13020444.

[7] Wang X, Zhang Y, Yu D, et al. Investigating the spatiotemporal pattern of urban vibrancy and its determinants: Spatial big data analyses in Beijing, China. Land Use Policy, 2022, 119,106162. DOI: 10.1016/j.landusepol.2022.106162

[8] Whittemore A H, BenDor T K. Talking about density: An empirical investigation of framing[J]. Land use policy, 2018, 72: 181-191.

[9] Zhang X, Sun Y, Chan T O, et al. Exploring impact of surrounding service facilities on urban vibrancy using Tencent location-aware data: A case of Guangzhou[J]. Sustainability, 2021, 13(2): 444.

Chapter 7. Conclusion and prospect

Content

7.1 Conclusion

7.1.1 Spatial distribution pattern of population activities and functional facilities in the station realm

(1) Spatiotemporal distribution pattern of station realm population activities

In terms of temporal distribution characteristics, population activities in the station realms are primarily characterized by commuting patterns on working days. As a result, distinct morning and evening peaks are observed. During the morning peak, population activities are concentrated between 08:00 and 10:00, while the evening peak mainly occurs between 17:00 and 19:00. On off days, there is no clear morning and evening peak phenomenon in the temporal distribution of population activities. Instead, activities fluctuate continuously between 09:00 and 18:00. Furthermore, the population activity levels within the 09:00-18:00 timeframe on off days are generally higher than those on working days. This indicates that people prefer to engage in activities during this time period on off days. In Xi'an, despite the increasing proportion of public transportation usage, with the subway accounting for over 50% in 2021, urban traffic congestion remains a serious issue. Therefore, understanding the variations in active periods of population activities can help city managers accurately and effectively optimize the operation time and structure of public transportation, thereby alleviating urban traffic congestion and related issues.

Regarding spatial distribution, the global spatial autocorrelation analysis results indicate that population activities in different types of station realms exhibit an overall clustering pattern, characterized as high/high clustering. However, local spatial autocorrelation analysis (hotspot analysis) reveals significant variations in the distribution of population activity clusters among different station realms. Most station realms exhibit dispersed areas of population activity clusters and do not show a concentric distribution pattern around the stations, indicating a certain deviation between theoretical expectations and actual development of station realms. However, the study also identifies eight station realms where population activities exhibit a higher degree of clustering within a radius of 500 meters from the stations. These station realms include ZL, XZ, FCWL, LSY, AYM, BDJ, STSG, and WYJ. This to some extent reflects the agglomeration effect of stations on population activities. Additionally, influenced by the variations in population activity levels between working days and off days, population activities tend to have a higher degree of clustering on off days compared to working days.

(2) Spatial distribution pattern of station realm functional facilities

The results of average nearest neighbor analysis indicate significant variations in the clustering levels of different types of functional facilities across station realms. However, within each station realm, the quantities and clustering levels of catering services, living services, and shopping facilities generally rank among the top three. In addition, public service facilities only exhibit clustering distribution in the ZL station realm. This is mainly because the quantities of public service facilities such as newsstands, public toilets, and rest areas are generally low in other station realms, and under normal circumstances, these types of public facilities do not have a strong attraction to pedestrians.

The kernel density analysis results show that functional facilities in the studied station realms generally exhibit clustered distributions. However, only six station realms (BY, XZ, FCWL, LSY, STSG, WYJ) demonstrate the characteristic of higher facility density closer to the stations. In fact, compared to developed countries, China's rail transit development started relatively late, and stations are often constructed in already mature built-up areas [1]. This has limited the spatial development of station realms due to factors such as public policies, transportation supply, station location, and land use structure [2,3]. Therefore, in most station realms, the desired spatial distribution patterns advocated by TOD and station-city integration, characterized by station-centered, high-density, and mixed-use development, have not been fully realized.

(3) Spatial correspondence between station realm population activities and functional facilities

The analysis of station realm spatial distribution patterns reveals a certain spatial correspondence between population activities and functional facilities within station realms, which is consistent with most previous research findings [4,5,6]. It suggests that areas with a high concentration of functional facilities tend to attract a higher population density. However, a notable difference is observed in several station realms along Metro Line 2, where there is a mismatch between areas of population aggregation and areas of single-service facility concentration. For example, in the HZZX station realm, there is a population aggregation in the northeast area, while the kernel density analysis shows a concentration of functional facilities in the northwest area of the station realm. In the ZL station realm, the aggregation areas of population activities and functional facilities demonstrate good spatial correspondence in the western and northern parts of the realm, but no such correspondence exists in the southeastern part of the realm.

(4) Data considerations

In contrast to previous studies that have utilized big data sources such as mobile signaling [7,8], social media [9,10], and Wi-Fi signaling technologies [11] to explore population distribution characteristics, this study employs more accurate and timely Yi Chu Xing positioning data, which provides researchers with a reliable data source for identifying dynamic distribution patterns of population activities from a micro-perspective. Additionally, existing literature suggests that POI (Point of Interest) data can effectively reflect the functional layout of urban areas [12]. Therefore, the use of EasyGo real-time positioning data and Amap POI data in this study helps researchers obtain more compelling analytical results.

7.1.2 Correlation between population activities and functional facilities in station realm

The results of the Spearman correlation analysis indicate that there are variations in the correlation between population activities and various service facilities within different types of station realms. For instance, in the YDGY station realm (under development), population activities are significantly and positively correlated with four types of service facilities: catering services, living services, shopping services, and transportation services. In contrast, in the XZ station realm (high population density hub), population activities exhibit significant positive correlations with all nine types of service facilities except for government agencies. Specifically, catering services, living services, shopping services, and transportation services show significant positive correlations with population activities in all station realms. On the other hand, public service facilities demonstrate significant positive correlations with population activities only in the XZ station realm. Government agencies show positive correlations with population activities only on working days in the STSG station realm. Furthermore, the variations in the degree of correlation also manifest in temporal changes. For example, in the NSM station realm, medical and wellness facilities are significantly and positively correlated with population activities on working days, but not on off days. In the ZL station realm, entertainment facilities in the WYJ station realm show significant positive correlations with population activities only on off days. These findings suggest that the relationship between population activities and service facilities within station realms is dynamic in nature.

7.1.3 Interaction mechanism of population activities and functional facilities in station realm

(1) The GWR model provides a better explanation for the interactive relationship between population activities and functional facilities in station realms, emphasizing the spatial aspect of the data.

In terms of the overall characteristics of the interaction, comparing the Adjusted R^2 and AIcc of the models, it can be concluded that the GWR models constructed for each station realm have better goodness of fit and performance than the OLS model. Therefore, the GWR model in this study demonstrates better explanatory power and is more suitable for describing the spatial distribution characteristics of station realm interactions at the micro-level, highlighting the spatial aspect of the data. This finding is consistent with current research results in spatial relationships [13,14,15]. However, this does not mean that the GWR model can be used alone. The fundamental reason is that it cannot diagnose multicollinearity issues among the explanatory variables. Therefore, the global OLS model is used to identify insignificant explanatory variables, which should be removed from the GWR model to avoid errors in the regression results.

(2) The interactive relationship between population activities and functional facilities in station realms exhibits dynamic characteristics.

Firstly, the regression results indicate that the degree of interaction between population activities and various types of functional facilities varies over time, and even the key variables driving the interaction may change within the same type of station realm. For example, in the urban center type of ZL station realm, the interaction between accommodation facilities and population activities increases on off days, while the interaction between transportation facilities and population activities decreases. In the XZ station realm, accommodation facilities are the key variables on working days, while transportation facilities play a dominant role on off days. The BDJ station realm transitions from transportation facilities as the dominant factor on working days to Medical & Health care facilities on off days.

Secondly, in terms of station realm types, station realms located on the urban periphery tend to have fewer types of functional facilities with a positive interactive relationship. For example, in the YDGY and WYJ station realms, there are only two types of functional facilities with a positive interactive relationship with population activities. In high population density-hub station realms, entertainment, shopping, and transportation facilities all have a positive interactive relationship with population activities. In interchange type of station realms, accommodation, dining, and transportation facilities all have a positive interactive relationship with population activities. In general, residential station realms, catering, and living service facilities both have a positive interactive relationship with population activities.

In summary, the spatial relationship model constructed in this study effectively explains the dynamic characteristics of the interactive relationship between population activities and various functional facilities in different types of station realms. The research findings contribute to the understanding of the dynamic demands generated by people's daily activities in different types of station realms, aiding urban managers in more effectively improving the allocation of urban public resources from a dynamic perspective.

(3) The transportation service facilities have become a key factor influencing population activities in station realms.

The regression results indicate that among the various study station realms, transportation services exhibit the highest frequency of positive interactive relationships with population activities (14 occurrences), making them the key variables dominating the interaction between functional facilities and population activities in multiple station realms (with the highest degree of positive interaction). Furthermore, the spatial distribution pattern of regression coefficients reveals that only transportation services exhibit a distribution characteristic where proximity to the station corresponds to a higher level of interaction with population activities. Thus, transportation services have become a crucial factor influencing population activities in station realms.

In contemporary China, traffic congestion and limited land resources are widespread challenges in most cities. Concurrently, the development of station realms often lags the overall urban construction process. Therefore, enhancing the scale of transportation services, improving the connectivity between subway stations and other modes of transportation, and enhancing the internal accessibility of station realms are effective approaches to elevate the level of transportation services in these realms. By doing so, the connections between different service facilities can be strengthened, attracting a greater volume of population activities and ultimately enhancing the utilization of urban functionalities.

(4) The regression coefficients of catering services and shopping facilities are inversely proportional to their scale.

The research results demonstrate that, within each station realm, both catering services and shopping facilities have larger scales compared to other service facilities. However, their positive interaction with population activities is generally low. This indicates that catering services and shopping facilities have limited impact on promoting population activities. Furthermore, with the rapid rise of the food delivery and online shopping industries, traditional catering services and physical retail stores will be further influenced. Therefore, in station realm planning and development, simply pursuing an accumulation of functional facilities is not an effective approach to enhance population activities. On the contrary, it may result in significant waste of public resources and reduce the efficiency of station realm utilization. For urban managers, it is essential to fully understand the interactive relationship between population activities and functional facilities in different types of station realms and grasp the most urgent needs of residents. This understanding will enable the formulation of spatial development strategies that are more aligned with reality and achieve coordinated development of station realm planning and construction.

(5) The rational allocation and equity of public healthcare resources in station realms need to be emphasized.

The research findings indicate that medical and healthcare facilities only exhibit a higher level of positive interaction with population activities in station realms where healthcare resources are superior. For instance, in the BDJ station realm, medical and healthcare facilities have a significant positive interaction with population activities, and during off days, they have become a key variable influencing population activities. Through on-site investigations, it has been observed that there are multiple large-scale or comprehensive public hospitals, as well as numerous medical supporting facilities and pharmacies in the BDJ station realm. In the current context of environmental changes, there is a significant increase in people's demand for high-quality medical environments and healthcare standards, which prompts them to seek medical treatment in station realms with relatively good healthcare resources. The characteristics of rapid metro transportation and high passenger volume will further exacerbate this phenomenon, leading to increased burden on healthcare institutions in the area and decreased healthcare experience for patients. Therefore, it is crucial for urban managers and researchers to address how to balance the allocation of public healthcare resources in each station realms and the associated issues of social equity.

7.2 Research innovation points

This study treats the station realm of rail transit as a complete and independent urban spatial unit, breaking the limitations of traditional urban spatial boundaries. It adopts a more micro perspective to understand the spatial distribution patterns and temporal dynamics of population activities and urban functions, thus expanding the research scope of urban spatial structure. The study combines urban planning, transportation, and geography, and establishes a systematic research framework focusing on station realm scope, station realm types, and spatial distribution patterns. It achieves quantitative analysis of the spatial distribution of population activities and urban functional layout within station realms, enriching the theoretical and methodological system of station realm spatial research in rail transit.

In terms of data, unlike big data such as mobile phone signaling, social media, and WIFI signaling technology, this study uses the positioning data of the EasyGO with higher accuracy and a larger user group [16,17], which provides a reliable source of data for the researchers to analyze the characteristics of the dynamic distribution of the population activities in the station realm at the micro level. Meanwhile, existing literature shows that POI data based on electronic map location information has important application value for understanding urban phenomena [18], estimating public resource allocation [19,20], and identifying functional areas of the city [21], especially to help better understand the complex interconnections between people and places [22]. Therefore, this study uses EasyGO data and AotoNavi Map's POI data to identify and analyze the spatial distribution characteristics of population activities and functional facilities, which can more accurately reveal the interaction mechanism between population activities and functional facilities in the station realm, which could provide a reliable basis for improving the allocation of public resources and realizing the matching of supply and demand in this region.

The study establishes a spatial relationship model for the interaction between population activities and functional facilities within station realms. Through correlation analysis, it constructs OLS global regression models for variables with significant tests to understand the basic interaction relationships between population activities and functional facilities in different types of station realms. Then, by establishing GWR local regression models, it further identifies the key variables that dominate the interaction between population activities and various types of service facilities. The GWR model effectively explains the mechanisms of population activities and functional facilities and exhibits good generality and performance. It reveals the general patterns of the interaction between population activities and functional facilities at different periods, types, and scales, providing technical support for future research on optimizing station realm spaces.

7.3 Prospect

This research has made some important progress in the study of urban spatial distribution patterns within rail transit station realms, especially in revealing the intrinsic interaction mechanism between population activities and functional facilities. However, due to the differences in the economic, cultural, and natural environments of different cities, the spatial characteristics presented in the station areas may vary. The conclusions of this study will also have some limitations in the process of practical promotion and application. Therefore, future research can also be explored in depth in the following three directions:

(1) Future research can consider a comparative analysis of the spatial distribution characteristics of rail transit station realms in different cities, which could further improve the generalizability and application value of the research results.

(2) Future research can consider conducting individual analysis of the interaction between a particular category of the facility with population activity, which could help support more specific and targeted data for improving the allocation of public resources in the station realm.

(3) Future research could consider expanding the form of positioning data, for example, by adding information such as age, income, and occupation to population activity data, thus further improving the rationality of the spatial relationship model.

Reference

[1] Cervero R. Linking urban transport and land use in developing countries. Journal of transport and land use, 2013, 6(1): 7-24.

[2] Ye Z, Chen Y, Zhang L. The analysis of space use around Shanghai metro stations using dynamic data from mobile applications. Transportation research proedria, 2017, 25: 3147-3160. DOI: 10.1016/j.trpro.2017.05.353

[3] Liang Y, Song W, Dong X. Evaluating the space use of large railway hub station areas in Beijing toward integrated station-city development. Land, 2021, 10(11): 1267. DOI: 10.3390/land10111267

[4] Zhong Weijing, Wang De, Xie Dongcan, Yan Longxu. Dynamic characteristics of Shanghai's population distribution using cell phone signaling data[J]. Geographic Research, 2017, 36(5): 972-

984

[5] Liu T, Zhou W, Cao Yingui. Study on the distribution of urban functional areas and population activities in Shenyang[J]. Journal of Geoinformation Science, 2018, 20(7): 988-995.

[6] Shi, Y.; Yang, J.; Shen, P. Revealing the Correlation between Population Density and the Spatial Distribution of Urban Public Service Facilities with Mobile Phone Data. ISPRS Int. J. Geo-Inf. 2020, 9, 38.

[7] Huang Q, Yang Y, Xu Y, et al. Citywide road-network traffic monitoring using large-scale mobile signaling data. Neurocomputing, 2021, 444: 136-146.

[8] Zhao Pengjun, Luo Jia, Hu Haoyu. Spatial match between residents' daily life circle and public service facilities using big data analytics: A case of Beijing. Progress in Geography, 2021, 40(4): 541-553.

[9] Condeço-Melhorado, A.; Mohino, I.; Moya-Gómez, B.; García-Palomares, J.C. The Rio Olympic Games: A Look into City Dynamics through the Lens of Twitter Data. Sustainability 2020, 12, 7003. https://doi.org/10.3390/su12177003

[10] Zhu, W.; Ma, D.; Zhao, Z.; Guo, R. Investigating the Complexity of Spatial Interactions between Different Administrative Units in China Using Flickr Data. Sustainability 2020,12, 9778. https://doi.org/10.3390/su12229778

[11] Zola E, Barcelo-Arroyo F. A comparative analysis of the user behavior in academic WiFi networks[C]//Proceedings of the 6th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks. 2011: 59-66.

[12] Li J, Li J, Yuan Y, et al. Spatiotemporal distribution characteristics and mechanism analysis of urban population density: a case of Xi'an, Shaanxi, China. Cities, 2019, 86, 62-70. DOI：10.1016/j.cities.2018.12.008.

[13] Sassi M. OLS and GWR approaches to agricultural convergence in the EU-15[J]. International Advances in Economic Research, 2010, 16: 96-108.

[14] Ma X, Zhang J, Ding C, et al. A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership[J]. Computers, Environment and Urban Systems, 2018, 70: 113-124.

[15] R. C., & Nakada, L. Y. K. GIS-based spatial modelling of COVID-19 death incidence in São Paulo, Brazil. Environment and Urbanization, 2021,33(1), 229–238.

[16] Zong, Leli, et al. "Detailed mapping of urban land use based on multi-source data: A case study

of Lanzhou." Remote Sensing, 2020, 12(12): 1987.

[17] Wu, Chao, et al. "Using street view images to examine the association between human perceptions of locale and urban vitality in Shenzhen, China." Sustainable Cities and Society, 2023, 88: 104291.

[18] Jiang, Shan, et al. "Mining point-of-interest data from social networks for urban land use classification and disaggregation." Computers, Environment and Urban Systems, 2015, 53: 36-46.

[19] Zhang, Xiuyuan, Shihong Du, and Zhijia Zheng. "Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with VHR images and POI data." ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161: 1-12.

[20] Wu, Xueling, Ruiqi Mao, and Xiaojia Guo. "Equilibrium of tiered healthcare resources during the COVID-19 pandemic in China: a case study of Taiyuan, Shanxi Province." International Journal of Environmental Research and Public Health, 2022, 19(12): 7035.

[21] Gao S, Janowicz K, Couclelis H. Extracting urban functional regions from points of interest and human activities on location-based social networks[J]. Transactions in GIS, 2017, 21(3): 446-467.

[22] Psyllidis A, Gao S, Hu Y, et al. Points of Interest (POI): a commentary on the state of the art, challenges, and prospects for the future[J]. Computational Urban Science, 2022, 2(1): 20.

Appendix

北九州市立大学

THE UNIVERSITY OF KITAKYUSHU

Graduate School of Environmental Engineering

Date: / /

Questionnaire

My name is Di Wang, a doctoral student from The University of Kitakyushu. Currently, I am conducting a survey on the walking time of pedestrians. I would be very grateful if you could accept this survey and answer the following questionnaire. The data from this survey will be used only for my doctoral dissertation.

A. Do you regularly use urban rail transportation (metro/subway, light rail, tram, etc.) in your daily travel activities?

- ☐ Every day
- ☐ 5 - 6 times a week
- ☐ 3 - 4 times a week
- ☐ 1 - 2 times a week
- ☐ Never

B. What mode of transportation do you usually use to arrive at the rail transit station?

- ☐ Walking
- ☐ Bicycle
- ☐ Bus
- ☐ Other, for example: _______________

C. What is your maximum acceptable walking time for arriving at the rail transit station?

- ☐ 5 minutes
- ☐ 8 minutes
- ☐ 10 minutes
- ☐ 15 minutes
- ☐ More than 15 minutes

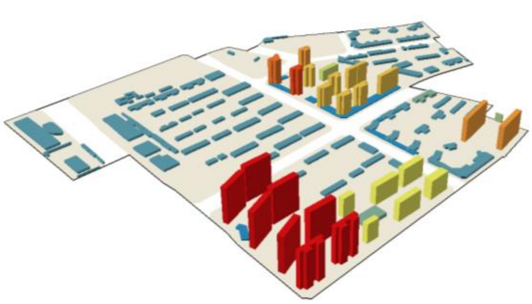
Thank you for your cooperation.

Appendix A. Questionnaire on maximum acceptable walking time for pedestrians

Appendix B. Definition and classification of POI data

Serial number	Primary classification	Secondary classification	Tertiary classification
1	Entertainment	Culture	Museums, science and technology museums, libraries, etc.
		Leisure and Recreation	Parks, playgrounds, KTV, etc.
		Sports and Fitness	Swimming pools, badminton courts, gymnasiums, etc.
2	Financial Service	Bank	Bank management, banks, credit unions, etc.
		Insurance	Insurance management, insurance, etc.
		Securities	Securities management, investment, guarantee, etc.
3	Government Agency	Government	Provincial, city (local), district governments, etc.
		Public, Prosecution, Law	Public security, public prosecutor's office, court
		Taxation	State tax, local tax
		Social groups	Public groups at all levels
4	Medical & Health Care	Medical institutions	Hospitals, medical points, clinics
		Epidemic Prevention and Control	Epidemic prevention, medical examination
		Health Care Rehabilitation	Health care, rehabilitation
		Pharmacy	Pharmacies, medical equipment stores
5	Accommodation Service	Hotel	Express hotels, hotels, hostels, guest houses
6	Catering Service	Chinese restaurants	Local flavor restaurants, hot pot restaurants, etc.
		Foreign restaurants	Western restaurants, foreign cuisine restaurants
		Fast food restaurants	KFC, McDonalds, Pizza Hut, etc.
		Teahouse	Teahouse
		Beverage stores	Coffee shops, cold drink stores, dessert stores, etc.
7	Living Service	Water, electricity, gas services	Various types of life payment business hall, electricity maintenance, gas maintenance, etc.
		Living Services	Housekeeping, dry cleaners, wedding services, etc.

		Ticket office	Train ticket office, bus ticket office, etc.
		Store	Supermarkets, shopping malls
			Electrical appliances, digital products stores
8	Shopping		Tobacco, wine, tea, souvenir, and clothing stores
		Flowers, pets	Flower or pet stores
		Furniture, building materials	Furniture stores, building materials markets
		Wholesale market	All kinds of wholesale markets
		Public Transportation Stops	Bus stops, cab stands, public bicycle storage points, etc.
9	Transportation Service	Transportation parking facilities	Various types of parking lots
		Car Service	Gas stations, car rental, traffic vehicle management, etc.
10	Public Service	Public facilities	Newsstands, telephone booths, public restrooms, etc.



BY

YDGY

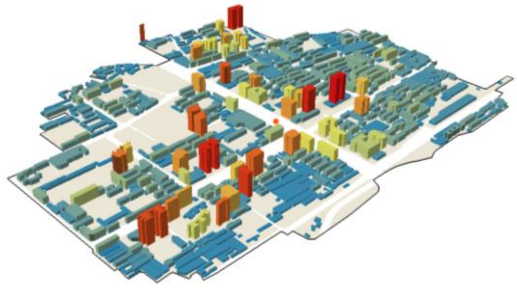
XZZX

FCWL

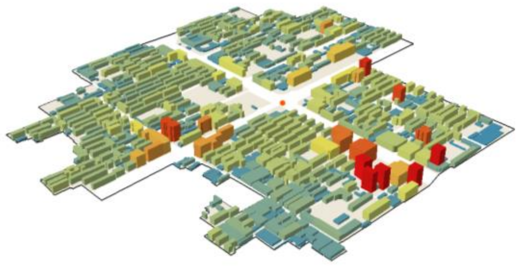
STSG

DMGX

LSY



AYM



BDJ

ZL

YNM

NSM

TYC

XZ

WYJ

HZZX

Appendix C. Building height in each station realm