ScotiencePG
Science Publishing Group

How Do Land Use Types and Transport Conditions Affect Population Aggregation Around Suburban Metro Stations: A Case Study in Shanghai

Zhang Yunzheng

College of Architecture and Urban Planning, Tongji University, Shanghai, China

Email address:

1932113@tongji.edu.cn

To cite this article:

Abstract: With the acceleration of urban sprawl, many cities have extended their metro lines into peripheral areas. It is hoped that the expansion of transport network can promote the construction of less developed regions and help relieve overpopulation in city centers. When measuring the spatial quality of station catchment areas, land use and transport conditions are two main dimensions that should be paid attention to. But how they affect population distribution hasn't been carefully studied yet. Up to now, there have been a substantial amount of studies relating to the comparatively well-developed transit stations in city centers whereas suburban stations still require further exploration. To fill this gap, our study aims to learn about current land use types, transport conditions and population aggregation around 97 suburban metro stations in Shanghai based on available information and data from Baidu Map and employs a stepwise regression approach to examine how the environmental variables affect population aggregation. The study shows that six land use variables and three transport-related variables are significantly associated with population aggregation on weekday while four land use variables and three transport-related variables show their significance on weekend. By finding out the coefficients of different land use and transport variables, this study hopes to guide government and urban planners in future planning in order to attract more people to live and work around suburban metro stations.

Keywords: Suburban Metro Stations, Land Use Types, Transport Conditions, Population Aggregation

1. Introduction

Urban sprawl is an inevitable tendency around the world, and this phenomenon is easily noticeable especially in big cities. Population in city centers is growing rapidly, which brings great pressure to urban space and facilities. Due to the lack of developable land, insufficient supply of amenities and overcrowdedness of population in the urban cores, big cities are seeking for development opportunities in the urban periphery. Under this situation, decision makers try to attract more people to suburban areas in order to relieve the burden of city centers. To promote the development of remote areas, old metro lines are extended and new suburban metro lines are constructed. Stations along these metro lines are considered as catalytic points. Housing, offices, shopping malls and other urban functions have been gradually introduced into the metro station areas. In China, these stations are mainly placed in less

developed area, which is believed to have high development potential [1]. But up to now, most of the suburban metro stations are newly constructed and their surrounding areas have not been developed thoroughly.

Will suburban transit stations effectively shape the distribution of population immediately after they are constructed? What factors, and to what extent, influence the aggregation of people? These questions are still waiting for us to explore. Existing studies have generally revealed that the construction of transport systems do not always bring prosperity to the surrounding areas [2]. So we hypothesize that there must be some latent factors relating to this phenomenon.

In the study of station catchment areas, land use and transport conditions have been investigated by a large number of scholars. It is believed that spatial quality can be improved by controlling these two aspects. Dating back to the 1990s, Bertolini proposed the "node-place" model, in which "node" refers to transport and

"place" refers to land use [3]. Later, a growing number of scholars began to adopt this model in their studies to measure transport and land values of transit stations [4-6]. However, previous studies mainly focused on transit stations in city centers probably because of the already well-developed environment and easily available data. Compared to them, suburban metro stations haven't been explored in detail and should receive more attention since they will probably turn into central stations as urban expansion reaches the further stages. Unlike those stations in the city centers that are surrounded by high and dense buildings, suburban stations are always adjacent to low-density blocks. Therefore, we decide not to consider the "density" index of land use and only focus on land use types. Besides, transport condition is another possible aspect affecting people's aggregation around the stations, which should also be considered.

In this study, suburban metro stations in Shanghai were taken as examples. We collected land use types, transport conditions and population aggregation data from Baidu Map to learn about station catchment areas and used a stepwise regression model to find out the main explanatory factors as well as their impacts on population aggregation on working days and non-working days. The findings will provide guidance for authorities, urban planners and designers in making master plans and bolstering suburban metro development.

The remainder of the paper consists of the following sections. Section 2 is literature review, focusing on land use types, transport and population density of suburban metro station areas. Section 3 introduces the study area and how the data was collected and calculated. Section 4 presents the results of stepwise regression tests and explains the relationship between independent and dependent variables. The last section provides conclusions and policy implications.

2. Literature Review

2.1. Land Use Types of Suburban Transit Station Areas

Land use type is not a new concept in the study of transit station catchment areas and there have been a substantial amount of research relating to it. Scholars often use one or a few summative indicators to measure land use. For example, under Cervero's 3D principle, "diversity" is used to illustrate how many types of functions are included within a certain range of space [7]. For many studies, areas of different land use types are calculated respectively and they are merged together by a diversity formula to give an overall description of land use mix [8, 9]. Land use mix can be defined by different indexes, including entropy index, dissimilarity index, mix type index and area index. These indexes have different calculation methods and they can all represent how diversified the urban functions are in station catchment areas [10, 11]. Some scholars create sub-indexes under the "diversity" index. A study categorized land use types into residential, commercial, institutional and other functional areas and measured each type separately. But this categorization is not detailed enough and it does not reflect the influence of all the functions [12].

Among relevant studies, it can be witnessed that most of them put an eye on transit stations in the central part of the city, while a comparatively small proportion turn their attention to suburban stations. Some studies focused on one specific land use type of station catchment areas. Since the initial intention of developing suburban areas was to provide enough housing for residents, residential land and housing prices have been studied by a large proportion of scholars [13-15]. Another significant type, commercial land, also attracts some attention [16]. Besides, some scholars focus on offices and institutions. It is demonstrated that attractiveness of these functions is closely connected to accessibility of the station catchment areas. As for the transition of land use types, it is observed that in big cities like Shanghai and Chennai, residential growth is mostly visible at the periphery [17, 18]. As suburban transit station areas further develops, commercial and institutional uses begin to grow [19].

Some scholars tend to learn about land uses surrounding suburban transit stations in detail instead of focusing on merely one or two functions. For example, a study categorized and examined land uses around twenty stations located in suburban Washington, D. C. and Oakland through aerial photographs and field investigations, showing that urban functions such as residential, commercial and institutional land had been introduced into suburban station areas and people move from city centers to the periphery for new life and work [20]. Some scholars classified land uses into residential, institutional, light commercial and light industrial land [21] while in a study of Shanghai metro, stations were categorized into residential oriented stations, integrated functional oriented stations, employment oriented stations and employment-residence oriented stations based on land use types. These study didn't concentrate on stations located in the outlying areas [22].

It is pointed out in a paper that cities in developing countries have greater level of land use diversity, which requires further exploration [23]. At present, problems exist in the planning and design process of suburban transit stations in China. Compared with governments in other developing countries, Chinese government is the chief decision maker in the construction, development and implementation process of detailed land use planning [18, 24]. But the fact is that most suburban station areas haven't been planned well. Housing development usually accompanies the extension of subway lines simultaneously, but jobs and businesses do not. Some people are not willing to move out from city centers because they hope to have better access to living facilities. And it is also revealed that suburban transit stations won't have high level of development without an appropriate master plan [25]. Therefore, it's essential to allocate different types of land use scientifically.

2.2. Transport Conditions of Suburban Transit Station Areas

When evaluating the spatial quality of transit station areas,

transport characteristics are also taken into consideration. There have been groups of variables reflecting transport conditions. Some of them are about the attributes of individual stations. For example, Zhao et al. introduced a categorical dummy variable to indicate whether a station was elevated or underground. He and other scholars also considered whether the station was a transfer, terminal or typical station and demonstrated that the station type will significantly affect ridership [26, 27]. It was also suggested that the development of station catchment areas was related to how long the stations had been built. And the surrounding areas of metro stations would enter into a more advanced development stage as time went on [28]. Another criterion, number of entrances and exits, was also a possible determinant factor influencing station catchment areas and was included in a measurement framework [29].

Some factors can reflect the locations of stations. Distance to city center has been taken into account most frequently in previous studies and it turned out that this factor had a considerably negative influence on station ridership [30, 31]. Besides, average distances to adjacent stations, nearest commercial centers, schools or other facilities were also regarded as possible influential factors of station catchment areas [29].

In addition, some scholars also measured intermodal connection. When people get off the metro, they usually choose to take a bus if their destinations are not too close to the metro stations. It was shown that number of bus stops and bus lines had significant impacts on station ridership [32, 33]. In a study of TOD in Shanghai, average distance to bus stops within the station catchment area was calculated in the evaluation process [29]. And Zhao et al. also hypothesized that bicycle parking and riding spaces would influence ridership around metro stations [34].

2.3. Population around Suburban Transit Stations

In cities where there is comparatively more convenient suburban transport network, changes in population distribution are recorded. Extended transit lines attract urban population to the outlying areas, which helps to deal with the mismatch between population density and insufficient infrastructure in city centers. Such phenomenon exists in both the developing and the developed world. In Madrid, Barcelona, Seoul and some American metropolises, population in the vicinity of newly constructed suburban transit stations is growing distinctly while that of old transit station areas in central city experiences an evident decrease [35-38]. This can be concluded as "suburbanization of population" or "population decentralization".

Despite the fact that some suburban metro stations contribute to the prosperity of outer areas, other unsuccessful stations fail in fostering development and population growth. There are usually undeveloped lands around suburban stations, which provides no function and have no appealing power to people in main urban areas. Some station areas are covered with a great number of residential buildings, but jobs, businesses and other amenities are not introduced into these

areas sufficiently [39]. These areas turn into the so-called "ghost town" with high residential vacancy due to the overdevelopment of real estates and lag of facility construction, meaning that people are not willing to move there and the burden of urban overpopulation can't be relieved [40].

Population distribution related to transport development have aroused scholars' interest. Population density is usually considered as a component of the spatial measurement framework. For example, it was used to represent space use intensity and formed an important part of the TOD index of Shanghai metro stations [41]. Besides, the number of residents within a certain region was a part of the "node" index in the evaluation of subway station areas in Tehran [4]. Population can also be used as an explanatory variable correlated with some dependents. In a Korean study, it was included as an explanatory factor which generated positive influence on station ridership [32].

However, few studies take population as a dependent variable and try to find out its relationship with land use types and transport conditions in suburban areas. Therefore, studying this relationship can help urban planners better design the surrounding space of suburban transit stations and relieve the population stress of central city.

3. Methodology

3.1. Study Area

In the past decades, Shanghai, one of the largest cities in China, has experienced rapid urban expansion. Urban transport network played an important role in Shanghai's urban growth. In the early 1990s, Shanghai's first metro line was built. Later, to promote the development of suburban areas, new metro stations have been built in the outer areas. Shanghai's Outer Ring Road is commonly regarded as the boundary of urban core. This study focuses on metro stations located outside the Outer Ring Road, which are defined as suburban metro stations.

Up to August 1, 2020, there were totally 17 metro lines and 111 suburban metro stations in Shanghai. Since Hongqiao Airport Terminal 1, Hongqiao Airport Terminal 2, Hongqiao Railway Station, Songjiang South Railway Station, Pudong International Airport are used as intercity transfer nodes and their surrounding areas are mainly covered by transportation facilities, we excluded these five stations from our sample group. Besides, Xiaotang, Fengpu Avenue, Huanchengdong Rd., Wangyuan Rd., Jinhai Lake, Zhujiajiao, Oriental Land, Shuyuan, Lingang Avenue were also deleted due to lack of population aggregation data. Our final sample group was comprised of 97 suburban metro stations in total.

A station catchment area is generally defined as a circular area radiating from the station within walking distance. Its size varies in extant studies. Some scholars employed a comparatively small circle with a radius of 400m [42], while some applied a longer distance of 500 m or 600m [43, 44]. Plus, some researchers adopted the 10 minutes' walking circle,

of which the radius was approximately 800 m [45, 46]. Considering that metro stations on the outskirts are sparse and they may have longer walking distances, we decided to use the 800 m radius in our study.
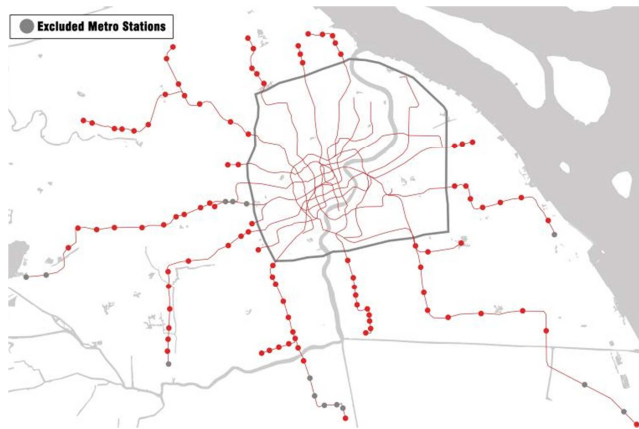

Figure 1. Suburban metro stations in Shanghai.

3.2. Data

3.2.1. Land Use Type Data

Baidu Map is a widely-used online map application in China. It can not only help people find their way, but also provide valuable geographic information by API technology and other built-in functions. Some extant studies used POI data to represent land use types in urban cores and tried to make categorizations accordingly. But in suburban areas, the combination of land uses is totally different, which is not reasonable for us to follow the previous categorizations. And this data source can't present any information of the percentage of agricultural land, vacant land or graveyard. Instead, Baidu Satellite Map, an attached function of Baidu Map, can show the distribution of different land use types. Considering the special land uses in suburban areas, in this study, they were categorized into 15 groups, including residential land, offices, commercial land, urban complexes (usually consisting of residential buildings, offices and shopping malls), industrial land, sports facilities, cultural&educational land, medical land, welfare, green space, logistics, transport facilities, agricultural land, graveyard and vacant land. We identified each land use type around suburban metro stations with the help of satellite maps and manually calculated how much percentage a certain land use type took

up within the 800 m circle.

Figure 4 shows the distribution of different types of land around 97 suburban metro stations in Shanghai. From the figure, we can clearly see that most station catchment areas are taken up by comparatively large proportion of residential land, which is well in line with the findings of previous studies. Offices, commercial land and urban complexes are also easily observed. But the percentages of them in each station are always small.
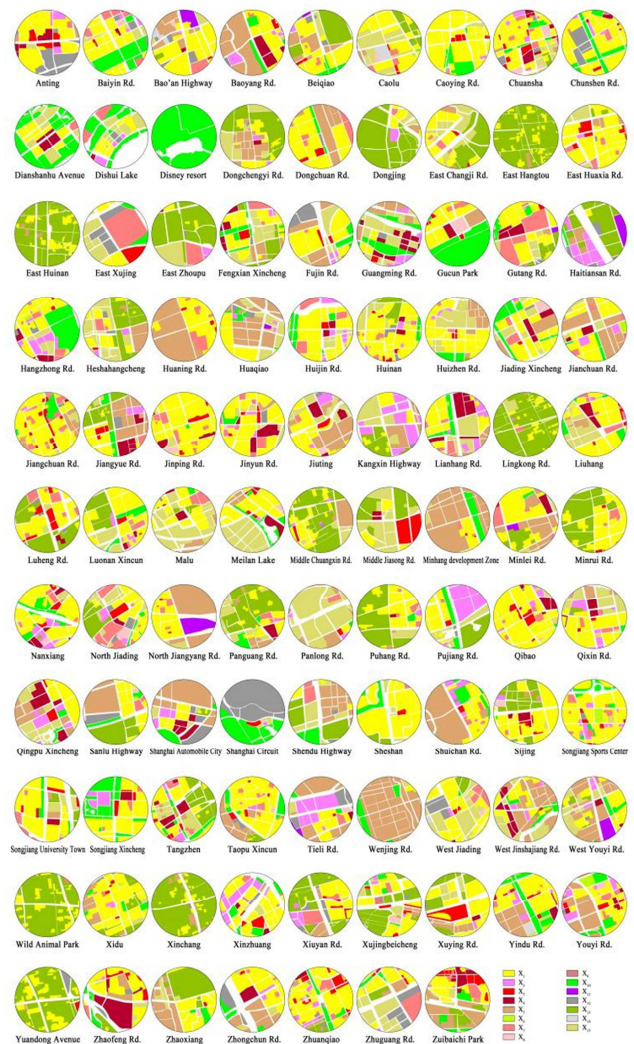

Figure 2. Land use types of suburban metro station areas in Shanghai.

Table 1. Descriptive result of land use types.

Variables	Min	Max	Mean	Std. dev	Observations
X_1 Residential	0	0.692	0.270	0.164	95
X_2 Offices	0	0.238	0.027	0.045	70
X_3 Commercial	0	0.220	0.018	0.024	75
X_4 Urban complex	0	0.220	0.025	0.037	57
X_5 Industrial	0	0.718	0.108	0.146	75
X_6 Sports	0	0.036	0.001	0.005	17
X_7 Cultural& educational	0	0.315	0.034	0.047	78
X_8 Medical	0	0.052	0.004	0.009	25
X_9 Welfare	0	0.019	0.000	0.002	2
X_9 Green land	0	0.809	0.061	0.109	73
X_{11} Logistics	0	0.121	0.004	0.016	8

Variables	Min	Max	Mean	Std. dev	Observations
X_{12} Transport facilities	0	0.566	0.023	0.065	60
X_{13} Agricultural	0	0.799	0.135	0.212	53
X_{14} Graveyard	0	0.049	0.001	0.006	4
X_{15} Vacant land	0	0.629	0.083	0.100	79

Table 1 shows the overall distribution conditions of different land use types. The 15 functions exist in some station catchment areas but are absent in others. Residential land, offices, commercial land, industrial land, green land and vacant land are observed in most of the suburban station areas. While welfare, logistics and graveyard are only located near a few stations. Nearly all the surveyed stations are surrounded by residential land, with only 2 exceptions. And the average percentage of residential land is the highest among all the types. Industrial and agricultural lands also cover a comparatively large part around suburban metro stations, but the presence of agricultural land varies a lot in different station catchment areas, and the area covered by industrial land also has significant differences.

Sports, medical land, welfare, logistics and graveyard only take up small parts, and of these functions welfare only exists around 2 stations. The maximum and standard deviation values of logistics are the largest compared with other 4 types, which means that in some areas its percentage is extremely small while in others the proportion is relatively large. The maximum values of industrial land and green land are both high. The mean value and standard deviation value of green land are lower than that of industrial land, suggesting that green lands are distributed more evenly near suburban metro stations.

Offices, commercial land, urban complexes, cultural&educational land and transport facilities take up a small part on average, but they are located around a majority of suburban metro stations. Among them, we can learn that commercial land has similar coverage in different station areas

from its smallest standard deviation value. Green land and vacant land have larger average coverage. But they don't show identical patterns in suburban station areas.

Overall, land use types are different in the investigated suburban metro station areas, which will probably have an influence on population aggregation.

3.2.2. Transport Data

Transport indicators were selected considering the availability of data. For each station, we chose 10 candidate explanatory variables relating to transport. Some refer to attributes of the stations, including station type (whether the station is elevated, underground or on the ground), whether it is a transfer station, whether it is a terminal station, when it was built, and the number of entrances and exits. We also collected some data reflecting the stations' locations. Some scholars have demonstrated that Shanghai presents a polycentric structure [47]. The general population distribution in the central area of Shanghai is dense and a drastic decline can be witnessed outside the Outer Ring Road, So we decided to take distance to the Outer Ring Road as an indicator instead of distance to city center. Besides, we also paid attention to average distance to adjacent metro stations. As for intermodal connection, passengers may prefer to transfer to a bus in suburban areas. So we chose 3 indicators about buses, including the number of bus stops, the number of bus lines and average distance to bus stops within station catchment area. All of the data mentioned above was collected through Baidu Map and its open platform.

Table 2. Descriptive result of transport conditions.

Variables		Min	Max	Mean	Std. dev
Y_1 Station type	elevated	0	1	0.588	0.495
	ground	0	1	0.031	0.174
	underground	0	1	0.381	0.488
Y_2 Transfer		0	1	0.041	0.200
Y_3 Terminal		0	1	0.175	0.382
Y_4 Years		2	24	9.113	4.780
Y_5 Entrances and exits		1	9	3.010	1.503
Y_6 Distance to the Outer Ring Road (km)		0.778	38.563	9.629	7.276
Y_7 Average distance to adjacent stations (km)		0.914	7.900	2.396	1.236
Y_8 Bus stops		1	30	11.701	6.450
Y_9 Bus lines		2	47	13.103	8.568
Y_{10} Average distance to bus stops (km)		0.124	0.624	0.453	0.091

Table 2 summarizes the descriptive statistics of suburban metro stations' transport conditions. One notable thing is that 1 categorical variable and 2 dummy variables are included in transport conditions. Station type is a categorical variable with 3 categories, so we should transform it into 3 dummy variables for subsequent regression analysis. For example, if a station has elevated platform, then the value of "elevated" is 1, or the value is 0. We can see that more than half of the the stations in suburban areas have elevated platforms while the second

largest proportion have underground platforms. Only about 3% of the stations have their platforms on the ground. Transfer and terminal are 2 dummy variables reflecting a station's function. Most of the stations are not transfer nodes, which indicates that suburban metro lines don't have too many intersections. And terminal stations take up about 17.5%. Among the 97 stations, Shendu Highway and Xinzhuang are both transfer and terminal stations.

3.2.3. Population Data

In previous studies, scholars studied population distribution based on traditional census data, which is an official data source [36]. But as electronic devices are becoming more and more widely-used and people's locations can be easily recorded, smart card data, mobile-based data and Baidu Heatmap data have been adopted in some research [48-50]. Census data is always at the district, city or provincial level. The range is too large if we just want to concentrate on station catchment area. Smart card data can show information of the number of people entering or exiting a metro station at different times of a day. But it is only at the station level and can't reflect how many people stay in the catchment area. Mobile-based data can show the precise locations of people, but this data source is limited to only a few institutions and it is difficult to get for the public. In this paper, population aggregation data was collected from Baidu Heatmap. Baidu Heatmap is an input function of Baidu Map. It is easily acquired and shows how people aggregate within a certain region. One thing that should be noticed is that Heatmap can not show the accurate number of population density. But it can present the approximate trend of population distribution and indicate the relative level by using color blocks. For example, in Figure 2, red represents the highest population aggregation level and blue represents the lowest. There are totally 7 colors, so we can use number 1 to 7 to reflect population aggregation level.
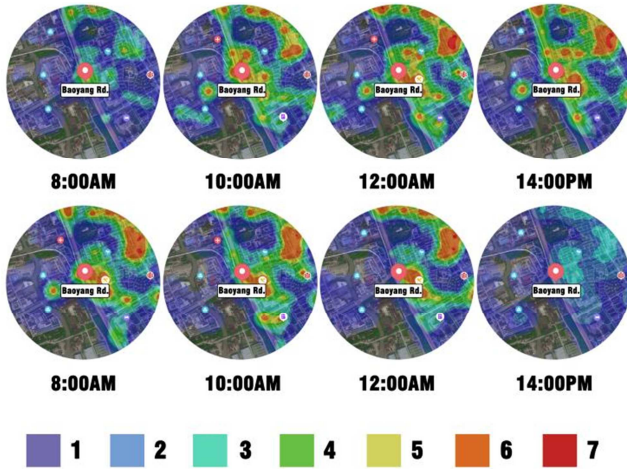

Figure 3. Heatmap values of population aggregation.

Since Heatmap data varies between working days and non-working days, we should consider population aggregation on weekday and weekend respectively. In the following study, We used Baidu Heatmap data on September 7, 2020 to represent population aggregation on weekday. And we calculated the average Heatmap data on September 12, 2020 and September 13, 2020 to represent population aggregation on weekend. Data was collected from 8:00 am to 22:00 pm each day, with 2-hour intervals. Population aggregation value at a certain time was determined by the following equation:

$$y=\sum nS_n, \ (n=1, \ 2, \ 3, \ 4, \ 5, \ 6, \ 7) \tag{1}$$

where y refers to population aggregation value at a certain time, n refers to population aggregation level of a certain color zone, S_n refers to the percentage a certain color zone takes up within the 800 m circle.

With the results calculated above, we can draw a line graph showing how population aggregation varied from 8:00 am to 22:00 pm. In the graph, fluctuations can be easily seen. And then we calculated the area enclosed by the population aggregation curve and X axis to show population distribution within a day. For example, in Figure 3, the area of blue color block is the representation of population aggregation of Huaning Rd. Metro Station on weekend.
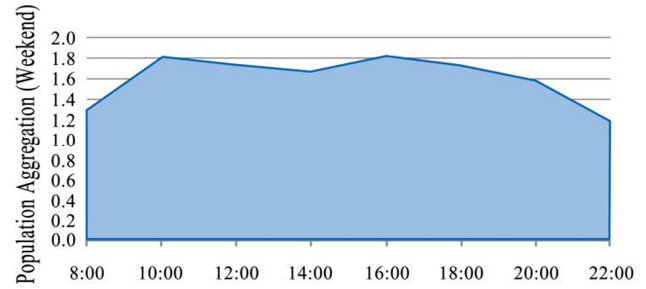

Figure 4. Population aggregation of Huaning Rd. on weekend.

Using the calculation method in the above section, we can find out the population aggregation level and draw a scatter plot accordingly. The scatter plot clearly indicates how many people gather around suburban metro stations on weekday and weekend. It is shown that most suburban metro station catchment areas can attract more people on weekend than on weekday. Population aggregation level for most of the station areas doesn't go beyond 40. But the values of 2 stations, Qibao and Xinzhuang, nearly reach 50, surpassing that of other stations. Some stations stay comparatively far away from the diagonal line, indicating that population distribution around them differs greatly between working day and non-working day. As for those close to the diagonal line, population aggregation doesn't show significant variance on different days.
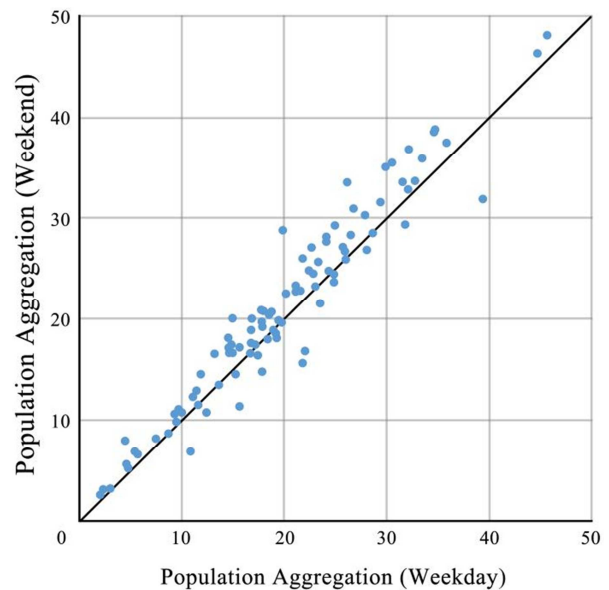

Figure 5. Population aggregation on weekday and weekend.

Figure 6 displays the geographic distribution of population on a map. The size of the circle represents how many people gather around the station within a day. The larger the circle is, the more people there are. It is distinct that people prefer to go to the west than to the east part of Shanghai. Most suburban metro stations situated close to the Outer Ring Road are more densely-populated comparatively. But some stations, though far away from the central area, also have the same trend. We can hypothesize that some other factors, in addition to distance, result in this phenomenon. Metro stations in the east, especially those located along Metro Line 16, cannot attract a great number of people. And we also use different colors to represent population variation between working day and non-working day. If more people stay in the catchment areas on weekday, then the color is red, or the color is green. From the map, we can directly see that Chuansha in Line 9 has the greenest color, indicating that there are far more people on weekend than on weekday. 4 stations near the Outer Ring Road, Tieli Rd. in Line 3, West Jinshajiang Rd. in Line 13, East Xujing in Line 2 and Kangxin Highway in Line 11, have red color and witness considerably higher level of population aggregation on weekday than on weekend, which is inconsistent with the majority of suburban metro stations.
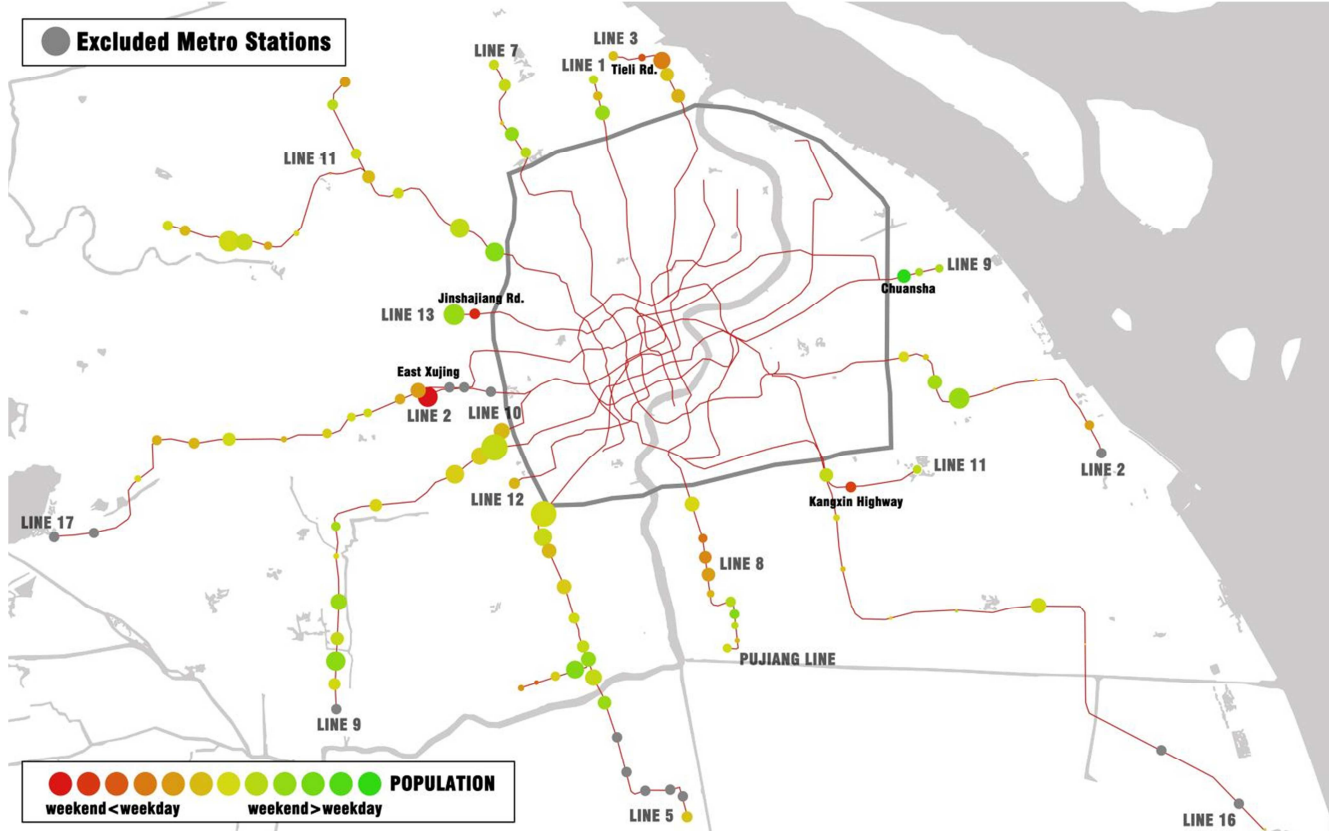

Figure 6. Map of population aggregation.

After the data collection and calculation process, a stepwise regression approach would be adopted to find out the relationship of population aggregation with land use types and transport conditions. It has been proven that stepwise regression can sort out the statistically significant independent variables and exclude those insignificant ones in the final model by screening all the explanatory variables, so it is applicable in our study [32].

4. Results and Discussion

We import the data of land use types, transport conditions and population aggregation into IBM SPSS Statistics 25.0. Since population aggregation shows different patterns within a week, we should carry out two stepwise regression analyses. Tables 3 and 4 show the regression results on weekday and weekend respectively.

Table 3. Stepwise regression result on weekday.

	Variables	Coef.	Std. err	t	Sig.
Land use types	X_1 Residential	13.242	3.588	3.690	0.000
	X_2 Offices	34.996	10.821	3.234	0.002
	X_3 Commercial	61.109	20.474	2.985	0.004
	X_4 Urban complex	56.878	12.918	4.403	0.000
	X_7 Cultural& educational	23.650	10.789	2.192	0.031
	X_{13} Agricultural	-10.896	2.571	-4.238	0.000

	Variables	Coef.	Std. err	t	Sig.
	Y_6 Distance to the Outer Ring Road (km)	-0.333	0.065	-5.091	0.000
Transport conditions	Y_8 Bus stops	0.204	0.107	2.374	0.020
	Y_9 Bus lines	0.206	0.080	2.583	0.011
(Constant)		12.221	1.509	8.097	0.000

Note: all variables are significant at the 0.05 level, and the overall accuracy of the regression is 77.6%.

***Table 4.** Stepwise regression result on weekend.*

	Variables	Coef.	Std. err	t	Sig.
	X_1 Residential	20.305	3.634	5.588	0.000
	X_3 Commercial	53.829	21.297	2.528	0.013
Land use types	X_4 Urban complex	58.034	13.495	4.300	0.000
	X_{13} Agricultural	-10.140	2.607	-3.890	0.000
	Y_6 Distance to the Outer Ring Road (km)	-0.378	0.066	-5.682	0.000
Transport conditions	Y_8 Bus stops	0.262	0.110	2.374	0.020
	Y_9 Bus lines	0.206	0.080	2.583	0.011
(Constant)		12.221	1.509	8.097	0.000

Note: all variables are significant at the 0.05 level, and the overall accuracy of the regression is 77.4%.

Stepwise regression can help remove variables that show little or no significance. The R square values of the regression models are 77.6% and 77.4% respectively. On weekday, 9 variables of land use types, including industrial land, sports, medical land, welfare, green land, logistics, transport facilities, graveyard and vacant land, and 7 variables of transport conditions, including station type, "transfer or not", "terminal or not", years of operation, number of entrances and exits, average distance to adjacent stations and average distance to bus stops didn't show much significance, so they were dropped out from the final model. From the stepwise regression results, we can clearly see that residential land, offices, commercial land, urban complex, cultural&educational land, agricultural land, distance to the Outer Ring Road, bus stops and bus lines are the main influential factors of population aggregation on working day. However, their coefficients differ a lot.

For land use variables, residential land, offices, commercial land, urban complex, cultural&educational land are positively correlated with population aggregation whereas agricultural land has negative impact. Among these independent variables, the coefficient of commercial land is the highest. This finding is well in line with previous studies, showing that shopping centers or other commercial venues contribute a lot to metro station ridership or population aggregation [32, 34]. The influence of urban complex is similar to that of commercial land, this may result from the fact that an urban complex also consists of a large amount of stores and convenient living facilities. Offices are also positively correlated with population because people need to go to work on weekdays. Residential land, which has been proved insignificant by some scholars [34], is the least important explanatory factor of all the positively related variables. This reveals that blindly increasing the construction of housing is not an effective way to attract people. Besides, cultural&educational land also has some positive effect, which is different from a previous study in Seoul [30].

However, offices and cultural&educational land no longer exist in the stepwise regression result on weekend. Perhaps they are excluded for the reason that employers and students don't need to go to work or go to school. On weekend, the correlation coefficient of residential land is higher than that on weekday, which can be assumed that people tend to stay close to where they live on weekends. Also, urban complex contributes slightly more to population distribution on weekend, even transcending the impact of commercial facilities. The coefficient of commercial land shows a decrease, indicating that fewer people choose to go shopping on weekend. This may be inconsistent with our common sense. Perhaps people in Shanghai prefer to stay at home after working for a week. They don't want to and don't need to commute to a commercial center because they can buy everything easily through online shopping websites. Besides, the influence of agricultural land remains stable on weekday and weekend.

As for transport conditions, only 3 variables, including distance to the Outer Ring Road, number of bus stops and bus lines, remain in the final regression results. Distance to the Outer Ring Road is negatively correlated with population aggregation and the effect is distinct. But this is a fixed variable that can't be controlled. Similar with other research findings, bus stops and bus lines are proved to be important factors that can help increase the number of people [51]. This suggests that an improvement of intermodal connection will boost the development of suburban metro station catchment areas and increase population gathering. Some transport-related characteristics, which were found to be key explanatory variables by other researchers, are not significant in explaining population aggregation around Shanghai's suburban metro stations. In some extant studies, whether a metro station is elevated, on the ground or underground influences ridership at the station level. But this study doesn't show its significance. Besides, in a study of Hong Kong's transit ridership, "years of operation" and "transfer or not" were proved to be significant variables, while in this study their impacts are not remarkable [51]. It was also revealed that in Seoul, number of entrances was a positive variable and distance to the closest station was a negative one [52]. However, they are dropped out in our final results.

5. Conclusion

This study attempts to find out land use types and transport conditions of suburban metro station catchment areas in Shanghai and sheds light on how they affect population aggregation. A stepwise regression model was conducted to examine the relationship based on available data.

The findings show that suburban metro stations are surrounded mainly by residential, industrial and agricultural land, but the percentage of each land use type varies among all the stations. Population aggregation on working days is mainly determined by 6 land use types, including residential land, offices, commercial land, urban complex, cultural&educational land and agricultural land. But offices and cultural&educational land don't show significant impacts on weekends. As for transport conditions, distance to the Outer Ring Road, the numbers of bus stops and bus lines are the main influential factors.

Some policy implications can be derived from our findings. For example, population distribution can be changed by controlling different land uses and improving transport infrastructure in station catchment areas. To attract more people, urban planners can rationally increase the ratio of residential land, offices, commercial land, urban complex and cultural&educational land and reduce agricultural land moderately if possible. It should be noticed that the impact of residential land is not so significant compared to other positively correlated land use types. So the traditional mass construction of housing doesn't effectively redistribute urban population. To solve the problem of crowdedness in city centers, shopping malls, cultural centers and other facilities should be built around suburban metro stations. Although distance to the Outer Ring Road negatively influences population, it is not a controllable factor. But improving intermodal transport, such as introducing more bus stops and bus lines, is probably a reasonable way.

There are some limitations in this study. During the data collection process, we adopted some manual methods, which might cause some inaccuracy. But minor imprecision is acceptable and it won't lead to serious error in the final result. Besides, this study was carried out based on a set of selected variables. It didn't take other indices, such as density, walkability, cyclability and accessibility into consideration. Since urban space is a complex system comprising of a great variety of elements, in the future, more independent variables can be included in the regression process to improve the goodness-of-fit and make the models more convincing.

References

[1] J. Yang, J. Chen, X. Le, and Q. Zhang, "Density-oriented versus development-oriented transit investment: Decoding metro station location selection in Shenzhen," Transp. Policy, 2016, vol. 51, pp. 93–102.

[2] K. Wang and M. Woo, "The relationship between transit rich neighborhoods and transit ridership: Evidence from the decentralization of poverty," Appl. Geogr., 2017, vol. 86, pp. 183–196.

[3] L. Bertolini, "Spatial development patterns and public transport: the application of an analytical model in the Netherlands," Plan. Pract. Res., 1999, vol. 14, pp. 199-210.

[4] S. Monajem and F. E. Nosratian, "The evaluation of the spatial integration of station areas via the node place model; an application to subway station areas in Tehran," Transp. Res. D. Transp. Environ., 2015, vol. 40, pp. 14-27.

[5] D. S. Vale, C. M. Viana, and M. Pereira, "The extended node-place model at the local scale: Evaluating the integration of land use and transport for Lisbon's subway network," J. Transp. Geogr., 2018, vol. 69, pp. 282-293.

[6] D. S. Vale, "Transit-oriented development, integration of land use and transport, and pedestrian accessibility: Combining node-place model with pedestrian shed ratio to evaluate and classify station areas in Lisbon," J. Transp. Geogr., 2015, vol. 45, pp. 70-80.

[7] R. Cervero and K. Kockelman, "Travel demand and the 3Ds: density, diversity, and design," Transp. Res. D. Transp. Environ., 1997, vol. 2, 199-219.

[8] D. E. Reusser, P. Loukopoulos, M. Stauffacher, and R. W. Scholz, "Classifying railway stations for sustainable transitions - balancing node and place functions," J. Transp. Geogr., 2008, vol. 16, pp. 191-202.

[9] G. Lyu, L. Bertolini, and K. Pfeffer, "Developing a TOD typology for Beijing metro station areas," J. Transp. Geogr., 2016, vol. 55, pp. 40-50.

[10] R. Bordoloi, A. Mote, P. P. Sarkar and M. Chunchu, "Quantification of land use diversity in the context of mixed land use," Procedia Soc. Behav. Sci., 2013, vol. 104, pp. 563–572.

[11] B. Qin, W. Zhu, J. Wang, and Y. Peng, "Understanding the relationship between neighborhood green space and mental wellbeing: A case study of Beijing, China," Cities, 2021, vol. 109, pp. 103039.

[12] K. T. K. Nyunt and N. Wongchavalidkul, "Evaluation of relationships between ridership demand and transit-oriented development (TOD) indicators focused on land use density, diversity, and accessibility: A case study of existing metro stations in Bangkok," Urban Rail Transit, 2020, vol. 6, pp. 56–70.

[13] H. Dong and M. Chunchu, "If You Build Rail Transit in Suburbs, Will Development Come?" J. Am. Plann. Assoc., 2016, vol. 82, pp. 316–326.

[14] R. Cervero and M. Duncan, "Benefits of Proximity to Rail on Housing Markets: Experiences in Santa Clara County," J. Public Trans., 2002, vol. 5, pp. 1–18.

[15] M. Boarnet and R. Crane, "L. A. story: A reality check for transit-based housing," J. Am. Plann. Assoc., 1997, vol. 63, pp. 189–204.

[16] L. Zhang and Y. Zhuang, "Analysis of multiple network accessibilities and commercial space use in metro station areas: An empirical case study of Shanghai, China," Int. J. High-Rise Build., 2019, vol. 8, pp. 49–56.

[17] L. Mejia-Dorantes, A. Paez, and J. M. Vassallo, "Transportation infrastructure impacts on firm location: the effect of a new metro line in the suburbs of Madrid," J. Transp. Geogr., 2012, vol. 22, pp. 236–250.

[18] Q. Shen, "Urban transportation in Shanghai, China: problems and planning implications," Int. J. Urban Reg. Res., 1997, vol. 21, pp. 589–606.

[19] S. P. Sekar and D. Gangopadhyay, "Impact of rail transit on land use and development: Case study of suburban rail in Chennai," J. Urban Plan. Dev., 2017, vol. 143, 04016038.

[20] H. Moon, "Land use around suburban transit stations," Transportation, 1990, vol. 17, pp. 67–88.

[21] P. Shen, L. Ouyang, C. Wang, Y. Shi, and Y. Su, "Cluster and characteristic analysis of Shanghai metro stations based on metro card and land-use data," Geo Spat. Inf. Sci., 2020, vol. 23, pp. 352–361.

[22] S. Pulugurtha and M. Agurla, "Assessment of models to estimate bus-stop level transit ridership using spatial modeling methods," J. Public Trans., 2015, vol. 15, pp. 33–52.

[23] R. Cervero, "Linking urban transport and land use in developing countries," J. Transp. Land Use, 2013, vol. 6, pp. 7–24.

[24] X. Wang, D. Tong, J. Gao, and Y. Chen, "The reshaping of land development density through rail transit: The stories of central areas vs. suburbs in Shenzhen, China," Cities, 2019, vol. 89, pp. 35–45.

[25] R. Cervero, J. Day, "Suburbanization and transit-oriented development in China," Transp. Policy, 2008, vol. 15, pp. 315–323.

[26] J. Zhao, W. Deng, "Relationship of walk access distance to rapid rail transit stations with personal characteristics and station context," J. Urban Plan. Dev., 2013, vol. 139, pp. 311–321.

[27] O. D. Cardozo, J. C. García-Palomares, and J. Gutiérrez, "Application of geographically weighted regression to the direct forecasting of transit ridership at station-level," Appl. Geogr., 2012, vol. 34, pp. 548–558.

[28] R. Tan, Q. He, K. Zhou, and P. Xie, "The effect of new metro stations on local land use and housing prices: The case of Wuhan, China," J. Transp. Geogr., 2019, vol. 79, 102488.

[29] Z. Li, Z. Han, J. Xin, X. Luo, S. Su, and M. Weng, "Transit oriented development among metro station areas in Shanghai, China: Variations, typology, optimization and implications for land use planning," Land use policy, 2019, vol. 82, pp. 269–282.

[30] K. Sohn and H. Shim, "Factors generating boardings at metro stations in the Seoul metropolitan area," Cities, 2010, vol. 27, pp. 358–368.

[31] S. Wang, L. Sun, J. Rong, S. Hao, and W. Luo, "Transit trip distribution model considering land use differences between catchment areas," J. Adv. Transp., 2016, vol 50, pp. 1820–1830.

[32] M. J. Jun, K. Choi, J. E. Jeong, K. H. Kwon, and H. J. Kim, "Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul," J. Transp. Geogr., 2015, vol. 48, pp. 30–40.

[33] M. Kuby, A. Barranda, and C. Upchurch, "Factors influencing light-rail station boardings in the United States," Transp. Res. Part A Policy Pract., 2004, vol. 38, pp. 223–247.

[34] J. Zhao, W. Deng, Y. Song, and Y. Zhu, "What influences metro station ridership in China? Insights from Nanjing," Cities, 2013, vol. 35, pp. 114–124.

[35] F. Calvo, J. Oña, and F. Arán, "Impact of the Madrid subway on population settlement and land use," Land use policy, 2013, vol. 31, pp. 627–639.

[36] J. Jin, D. Kim, "Expansion of the subway network and spatial distribution of population and employment in the Seoul metropolitan area," Urban Stud., 2018, vol. 55, pp. 2499–2521.

[37] N. Baum-Snow, "Did highways cause suburbanization? Q. J. Econ," 2007, vol. 122, pp. 775–805.

[38] M. À. Garcia-López, "Urban spatial structure, suburbanization and transportation in Barcelona," J. Urban Econ., 2012, vol. 72, pp. 176–190.

[39] S. Williams, W. Xu, S. B. Tan, M. J. Foster, and C. Chen, "Ghost cities of China: Identifying urban vacancy through social media data," Cities, 2019, vol. 94, pp. 275–285.

[40] H. J. Teng, C. O. Chang, C. M. Yu, and M. C. Chen, "Rent seekers vs. producers in cities: Contagious housing bubbles force housing price diffusion on urban overdevelopment," January 5, 2013. Available online: https://ssrn.com/abstract=2196798 or http://dx.doi.org/10.2139/ssrn.2196798

[41] Z. Ye, Y. Chen, and L. Zhang, "The Analysis of Space Use around Shanghai Metro Stations Using Dynamic Data from Mobile Applications," Transp. Res. Procedia, 2017, vol. 25, pp. 3147–3160.

[42] P. N. Chen and K. Karimi, "The impact of a new public transport system on neighbourhoods surrounding stations: The cases of Bermondsey and West Ham with the Jubilee Line extension in London," In 24th ISUF International Conference. Book of Papers, Valencia, Spain, September 27-29, 2017, Editorial Universitat Politècnica de València: Valencia, Spain, 2018, pp. 1449–1459.

[43] X. Li, Y. Liu, Z. Gao, and D. Liu, "Linkage between passenger demand and surrounding land-use patterns at urban rail transit stations: A canonical correlation analysis method and case study in Chongqing," Int. J. Transp. Sci. Technol., 2016, vol. 5, pp. 10–16.

[44] D. An, X. Tong, K. Liu, E. H. W. Chan, "Understanding the impact of built environment on metro ridership using open source in Shanghai," Cities, 2019, vol. 93, pp. 177–187.

[45] R. Huang, A. Grigolon, M. Madureira, and M. Brussel, "Measuring transit-oriented development (TOD) network complementarity based on TOD node typology," J. Transp. Land Use, 2018, vol. 11, pp. 305–324.

[46] Y. J. Singha, A. Lukmana, J. Flackea, M. Zuidgeestb, M. van Maarseveen, "Measuring TOD around transit nodes - Towards TOD policy," Transp. Policy, 2017, vol. 56, pp. 96–111.

[47] X. Liu, X. Yan, W. Wang, H. Titheridge, R. Wang, and Y. Liu, "Characterizing the polycentric spatial structure of Beijing Metropolitan Region using carpooling big data," Cities 2021, vol. 109, 103040.

[48] M. P. Pelletier, M. Trépanier, and C. Morency, "Smart card data use in public transit: A literature review," Transp. Res. Part C Emerg. Technol., 2011, vol. 19, pp. 557–568.

[49] C. Kang, Y. Liu, X. Ma, and L. Wu, "Towards Estimating Urban Population Distributions from Mobile Call Data," J. Urban Technol., 2012, vol. 11, pp. 3–21.

[50] W. Zhang, Y. Zhao, X. Cao, D. Lu, and Y. Chai, "Nonlinear effect of accessibility on car ownership in Beijing: Pedestrian-scale neighborhood planning," Transp. Res. D. Transp. Environ., 2020, vol. 86, 102445.

[51] B. P. Loo, C. Chen, and E. T. Chan, "Rail-based transit-oriented development: lessons from New York City and Hong Kong," Landsc. Urban Plan., 2010, vol. 97, pp. 202–212.

[52] H. Sung, K. Choi, S. Lee, and S. H. Cheon, "Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership," J. Transp. Geogr., 2014, vol. 36, pp. 134–140.