Examining Common Factors Influencing Suburban Transit Ridership: A Case Study of Metro Stations outside Center-Shanghai

Chenwei Jin
RAC STUDIO
cwjin0207@163.com

Abstract

Research have suggested that TOD (transit-oriented development) in sub-centers with transit stations is an approach to increase foot traffic and the overall vitality of the area. However, current development of sub-centers mostly follows the established TOD model in CAZs (central activity zone), and most TOD studies do not deliberately distinguish or emphasize between CAZs and sub-centers. Therefore, this study aims to examine the influence of factors in the built-up area of Shanghai's urban sub-center on TOD-ness around the station. The four main factors are accessibility to private cars, land use, inner connectivity, and functionality. A regression model was introduced to validate the factors that significantly influence transit ridership. Results show that most of factors are relative to metro usage, with a lower correlation index, indicating the extent of the impact they bring varies remarkably with the urban context. while some factors have conversed influence as in urban centers.

Keywords: Shanghai, Metro ridership, TOD, Suburban, Transit station.

1. Introduction

As China's urban environment undergoes dramatic changes, transitioning from traditional mixed-use, high-density, pedestrian-oriented urban forms to the establishment of new towns and isolated, automobile-oriented superblocks on the city periphery (Cervero, 2013), the urban travel patterns in Shanghai have been significantly affected. The increasing preference for private car travel has led to city expansion along highways, resulting in longer average travel times and decreased accessibility to destinations (Niedzielski & Boschmann, 2014; Gakenheimer, 1999; Rode et al., 2017). In urban centers, limited available land, combined with the integration of railways and commercial, business, and office land use, promotes efficient and sustainable development. However, in areas outside the core, the construction of numerous toll-free urban expressways fragments the land, reducing development density. Despite this, due to Shanghai's explosive population growth, license plate restrictions, and driving costs, the metro remains a primary mode of transportation. Railway transit, being more environmentally friendly and efficient, coupled with Transit-Oriented Development (TOD), can encourage a shift from car use to railway, public transport, and walking by developing areas around stations, thus mitigating urban sprawl (Cervero & Sullivan, 2011).

The widespread adoption of TOD principles has spurred numerous studies focusing on metro stations and the impact of their surrounding built environments on ridership. However, most of these studies do not differentiate between metro stations in the city center and those outside it, and TOD development in sub-centers often follows the established models of city centers. For each transit station, the influence of local community factors on ridership can vary significantly with different city contexts. Therefore, it is crucial to independently study the factors influencing ridership at metro stations outside the city center. In other words, what are the common factors that influence suburban metro ridership? Are the criteria derived from TOD research and practices in city centers still applicable to suburban transit stations and their development? The results will not only provide policymakers, urban planners, metro operators, and developers with strategies to enhance vitality outside the city center but also offer development methods and valuable experiences within the Asian urban context for cities with lower urban levels or towards healthier development in megacities.

Accordingly, the aim of this study is to determine significant built-environment-related factors affecting metro usage outside central Shanghai. This study will first introduce principles for the selection of criteria and the study area, and further present the corresponding data collection and processing methods. Statistical methods will be employed to analyze the data

URI: https://hdl.handle.net/10125/109481
978-0-9981331-8-8

correlation, followed by a stepwise regression analysis using negative binomial regression models. Finally, we will discuss the results and summarize the conclusions and limitations of this study.

2. Literature review

Previous research has established comprehensive frameworks comprising metrics and variables that significantly impact transit ridership and transit-oriented development (TOD) characteristics. As noted by Singh et al. (2017), higher transit ridership is a strong indicator of a higher level of TOD. This study aims to identify the common attributes that influence suburban metro ridership. Consequently, our literature review focuses on several key aspects: the definition of pedestrian catchment areas, transit station and built environment factors, methodological approaches, and the specific emphasis of TOD research on suburban areas.

Most studies define the catchment area of a station based on the mean walking distance of pedestrians (Jun et al., 2015). Typically, a 250-800 meter buffer zone is used, corresponding to an 8-10 minute walk, which is considered a tolerable walking time (Jun et al., 2015; Zhao et al., 2013). According to the Shanghai Metro Communication Platform (2015), a 600-meter area can cover approximately 70% of the population in the service area.

The attributes of transit stations are generally described in terms of accessibility, which is often quantified by intersection density (Jun et al., 2015), road length (Zhao et al., 2013), distance to city centers (Zhao & Deng, 2013), and ease of bus transfer (Khosravi & Abrishami, 2024). Additionally, built environment attributes include land use (Choi et al., 2012), the number of parking lots (Kuby et al., 2004), walkability (Singh et al., 2017), and points of interest (POIs) (An & Tong, 2019).

In terms of analytical methodology, the majority of previous studies have utilized ordinary least squares (OLS) regression analysis (Kuby et al., 2004; Cervero & Murakami, 2008; Zhao et al., 2013). Other studies have employed Poisson models (Choi et al., 2012), negative binomial models (Khosravi & Abrishami, 2024), gradient boosting regression trees (Li et al., 2024), stepwise regression models, and multiple logit regression (Jun et al., 2015). Notably, studies that do not differentiate between stations inside and outside city centers typically report multiple regression models with R-square values ranging from 0.60 to 0.70. Canonical correlation analysis and principal component analysis are also widely used (Li et al., 2016; An & Tong, 2019).

Regarding TOD and transit ridership studies in suburban areas, Nigro et al. (2019) noted that very few studies explicitly focus on geographic areas characterized by medium or low population and activity densities. Nonetheless, Hrelja (2024) conducted a comparative study on TOD in medium-density contexts, arguing that each TOD project should be adapted to its specific context while maintaining the same planning processes as classic high-density urban TODs. The study also suggested that planning authorities need a long-term perspective, such as establishing welfare services within TODs. Cervero (2010) cautioned against over-parked suburban stations and high-standard parking conditions near suburban TODs, which may encourage private car ownership and usage, thus creating a 'vicious cycle' of supply and demand.

3. Methodology

3.1. Selection of criteria

Given that the factors influencing ridership vary according to the specific context of each station, this study aims to comprehensively identify the common factors affecting transit station usage outside the central urban area. To achieve this, it is necessary to consider both spatial and non-spatial aspects, as well as the development around the station and the transit infrastructure itself (Singh, 2015). Previous research has utilized traditional TOD evaluation dimensions, including Density, Diversity, and Design (Ewing et al., 2019; Woo, 2021). The Institute for Transportation & Development Policy (ITDP) also incorporates the dimensions of Shift, Densify, Mix, and Compact. Integrating these perspectives, we adopted a TOD evaluation framework developed from studies of five megacities in China, focusing on Node, Place, and Functionality (Su et al., 2021). Based on this framework, we selected four categories for our study: Land Use, Inner Connectivity, Functionality, and Private Car Accessibility.

3.2. Case study area

3.2.1. Unit of analysis. TOD focuses on the pedestrian user experience, aligned with Shanghai's current initiative to promote the 15-minute walkable neighborhood concept. We have chosen 600-meter radius buffers to approximate the 10-minute walking catchment area around metro stations. This approach accounts for potential detours caused by multiple exits at metro stations.

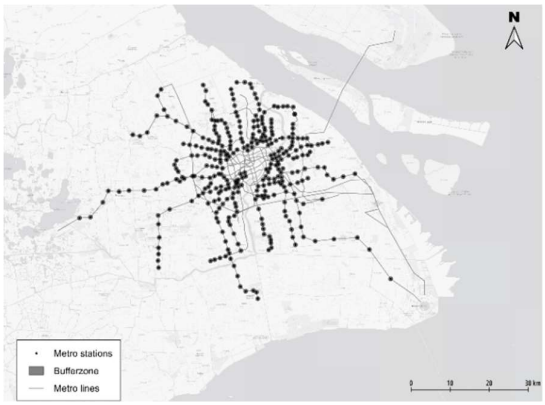

Figure 1. Study areas selected

3.2.2. Distinctions between city centers and suburban areas. Shanghai, as one of the largest and most dynamic cities in the world, exhibits a clear distinction between its city center and suburban areas. In terms of transportation modes, the city center relies more heavily on walking, cycling, and public transit due to its higher density and the frequent occurrence of road congestion. For instance, People's Square Station handles an average daily passenger flow of 120,000, making it the busiest station in the metro system. In contrast, suburban districts like Qingpu and Jiading rely significantly more on private vehicles. While these areas are connected to the city center via metro lines, public transportation in the suburbs is neither as frequent nor as convenient as in the city center. This reliance on cars is reflected in the higher rates of private vehicle ownership in these suburban areas.

Another notable characteristic of the city center is the concentrated and diverse land use, as it serves as the core of significant economic activities. For example, Huangpu District is home to a large number of adjacent residential areas, commercial districts, cultural centers, and government offices. In suburban areas, residential zones tend to dominate, while commercial activities are often confined to specific areas like shopping malls. From a traditional perspective on population density, Shanghai, as a polycentric city, has suburban sub-centers that sometimes exhibit population densities comparable to those of the city center, with similar metro ridership, as seen in districts like Minhang and Songjiang. However, these suburban areas generally lack the well-developed infrastructure found in the city center. As urban development in the city center approaches saturation, these suburban sub-centers present significant potential for future TODs, which underscores the importance of distinguishing between these areas.

3.2.3. Shanghai Central Activity Zone (CAZ). Although the distinctions between Shanghai's city center and suburban areas are significant across multiple dimensions, it becomes challenging to delineate a clear boundary in this polycentric city solely through data analysis. Thus, we turn our attention to the Shanghai government's urban planning white papers for further insight. *Shanghai Masterplan 2017-2035* introduced the concept of a Central Activity Zone (CAZ), which inherits the primary functions of the Central Business District (CBD) such as commerce and business, while also meeting the city's development needs by emphasizing functional diversity, adding vibrant elements, and enhancing the quality of the spatial environment. CAZ outlined in the planning documents aligns well with the traditional understanding of Shanghai's city center, encompassing core areas of Huangpu, Jing'an, and Xuhui districts, while also incorporating some forward-looking elements. Upon filtering the data, we found that the metro stations within the CAZ consistently rank in the top 20% of total metro ridership in Shanghai, further reinforcing the expectation of higher activity levels in the city center from a data perspective. Considering that the CAZ is a governmental guideline for Shanghai's future development, it will be more forward-looking and pragmatic to choose CAZ as the definition of city center. As a result, a total of 335 metro stations and their respective 600-meter radius catchment areas have been selected for the study.

4. Metrics Calculation

4.1. Land use

4.1.1. Land use percentage. Based on open-source land use classification data (Gong et al., 2020), land use is categorized into eleven subtypes including Residential, Office, Commercial Service, Industrial, Stations, Airport, Administrative, Educational, Medical, Sport and cultural, Park and greenspace.

4.1.2. Land use diversity. Renne (2009) indicated that land use diversity is a crucial factor in evaluating the degree of Transit-Oriented Development (TOD). Successful TODs often feature a balanced mix of various land use functions. To measure land use diversity, we employed the concept of entropy, which can be calculated using the following formula (Van Rij, Dekkers, & Koomen, 2008; Singh et al., 2017):

$$LU_d(i) = \frac{-\sum Q_{lu_i} ln\ (\ Q_{lu_i})}{ln(n)}$$

Where $Q_{lu_i} = \frac{S_{lu_i}}{S_i}, LU_d(i) =$ land use diversity in analysis area i, $lu_i =$ land use class (1,2, …, n) within analysis area i, Q_{lu_i} =Share of specific land use within the analysis area i, S_{lu_i} =Total area of the specific land use within the analysis area I, $S_i =$ Total area of analysis i.

4.2. Inner connectivity

4.2.1. Number of bus stops. We utilized open-source POI data to obtain the number of bus stops within a 600-meter catchment area, in order to calculate the density of bus services.

4.2.2. Shortest distance from bus stops. The distance to the nearest bus stop reflects the ease of metro-bus transfer at the station.

4.2.3. Road distance from city center. This metric measures the distance from each station to the city center along the road network, assessing the convenience of road transportation to the city center. The representative point for the city center is set at the intersection of the North-South Elevated Road and Yan'an Elevated Road.

4.2.4. Number of metro exits. The number of metro exits influences the size of the pedestrian catchment areas (Ped-Sheds) around a metro station, which reflects walkability. This, in turn, indicates the convenience for pedestrians in terms of walking to or from the station.

4.3. Functionality

4.3.1. Population. High population density increases the usage of transit services and promotes more walkable communities. Previous studies have also shown that higher population density correlates with a higher TOD score (Bach et al., 2006; Singh et al., 2017). This metric calculates the total population within a 600-meter catchment area.

4.3.2. Mean count inside populated raster. The metric calculates the average population count of residential cells within a 600-meter catchment area. Each cell corresponds to a 100m × 100m grid. Unlike the total population count, this metric specifically emphasizes the population density within residential areas (Chen, Xu, Ge, et al., 2024).

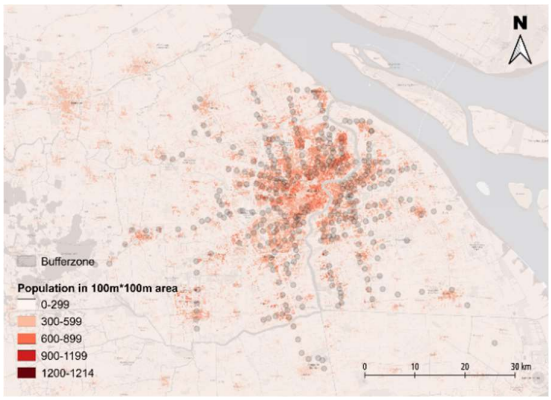

Figure 2. Raster population data with 100m*100m cells

4.3.3. Walkscores. Walkscore is a rating that assesses the pedestrian-friendliness of a street. In our study, we calculate the weighted average Walkscore for streets within the research area, considering the length of each street segment. This allows us to determine the average Walkscore at each specific location.

4.3.4. Number of reclassified POIs. This study employed POI data from Baidu Map. The data was then reclassified and filtered to 7 categories which basically cover the daily demand of passengers.

4.4. Private car accessibility

4.4.1. Road lengths. We leverage publicly available road data from Shanghai to reclassify roads into four distinct categories. We then calculate the length of each road type within the catchment area of each station to assess automobile accessibility. Previous research has also suggested a potential relationship between highway density and transit ridership (Zhao, Deng, Song, & Zhu, 2013).

Table 1. Reclassified categories of roads

Reclassified categories	Original categories
Center roads	Major Roads in Urban Centers
	Urban Center Roads
Primary roads	Arterial Roads
	Highway
	Major Arterial Roads in Suburban County Urban Centers
Secondary roads	Minor roads
Expressway	Toll-free expressway
	Expressway with tollbooths

4.4.2. Number of tollbooths. The number of tollbooths is indicative of the distance from the toll motorway and its density.

4.4.3. Distance from the nearest highway exit. Distance from the nearest highway exit reflects the extent to which the metro station is connected to Shanghai's expressway network, and thus, to some extent, the ease of access by private car.

4.4.4. Number of parking lots. Parking facilities within metro stations can provide convenience for commuters who switch from private cars to metro travel, thereby increasing metro ridership (Evans IV, Pratt, Stryker, & Kuzmyak, 2007).

4.5 Data processing

In addressing missing values, after manually inspecting the data on the map, we determined that all missing geographical information could be treated as zero. For outliers and missing values in metro ridership data, we excluded the affected station entirely. After performing the above calculations on the raw data using QGIS, we obtained the following data summary (as shown in the table). Given the need to compare data across different scales and units, we applied standardization before regression analysis to minimize the impact of absolute values on the significance of the regression model.

Table 2. Data Summary

Metrics	Max	Min	Mean	Std Dev	Unit
Healthcare	109.00	0.00	14.34	15.15	Count
Food	508.00	0.00	95.78	88.54	Count
Shopping	543.00	0.00	103.81	90.58	Count
EducationResearch	189.00	0.00	28.43	27.93	Count
FinanceInsurance	172.00	0.00	8.64	13.51	Count
Parking	241.00	7.00	54.64	36.59	Count
Toll	4.00	0.00	0.04	0.31	Count
LeisureRecreation	168.00	0.00	18.59	19.60	Count
CarService	130.00	0.00	13.07	13.93	Count
MetroExit	20.00	0.00	4.01	2.28	Count
OffRamp	10.00	0.00	0.48	1.26	Count
BusStop	17.00	1.00	5.85	2.90	Count
PopSum	37234.00	653.00	13193.87	8773.49	People
PopRasterMean	453.65	8.16	162.70	108.14	People
HighwayExitDist	6358.39	19.72	1300.25	937.82	M
BusStopDist	405.32	5.29	121.14	79.53	M
DistFromCenter	65099.78	3743.70	19345.63	10309.28	M
Walkscore	95.36	0.00	57.44	38.37	Percentage
EducationalPer	88.57	0.00	4.69	11.62	Percentage
ResidentialAreaPer	95.78	0.00	43.56	25.70	Percentage
ParkGreenspacePer	62.16	0.00	4.00	9.66	Percentage
AirportPer	98.66	0.00	1.27	8.16	Percentage
SportCulturalPer	70.61	0.00	4.41	10.79	Percentage
TransportationPer	35.34	0.00	0.47	3.16	Percentage
BusinessOfficePer	61.81	0.00	4.30	9.69	Percentage
AdministrativePer	67.19	0.00	3.06	8.63	Percentage
CommercialServicePer	45.99	0.00	2.49	6.24	Percentage
MedicalPer	52.00	0.00	1.77	5.98	Percentage
IndustrialPer	95.29	0.00	14.14	23.22	Percentage

Metrics	Max	Min	Mean	Std Dev	Unit
LandUseEntropy	0.99	0.10	0.73	0.17	M^2
CenterRoad_Length	7086.30	0.00	1784.81	1578.89	M
MainRoad_Length	3351.25	0.00	692.79	757.72	M
ExpressWay_Length	2189.39	0.00	265.41	466.50	M
SideRoad_Length	6385.74	0.00	1738.92	1254.29	M
TotalLength	8764.77	691.79	4481.94	1400.80	M
TotalRidership	75182.00	1765.00	22777.32	17488.71	People

5. Results and discussion

Based on a literature review, variables were categorized into four groups, and correlation analyses were conducted. Subsequently, to select a subset of variables with higher correlation, we employed negative binomial regression with stepwise regression.

5.1. Correlation analysis

While factor analysis or principal component analysis (PCA) were used in previous research (Cervero & Kockelman, 1997; Singh et al., 2017), the present study utilized a correlation analysis between the metrics and transit ridership. Based on the result of the normality test, we ran both Pearson and Spearman correlation analysis between the standardized metric, and Table 4 presents the result.

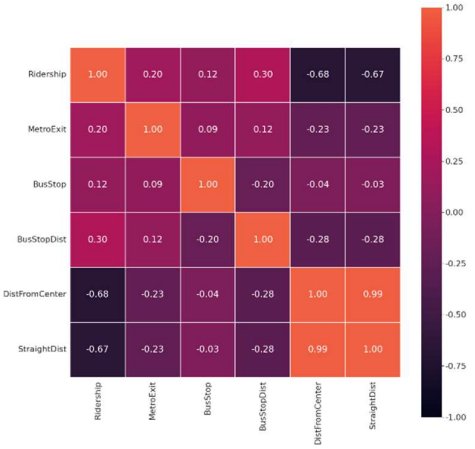

Figure 4. Pearson correlation heatmap of inner connectivity factors

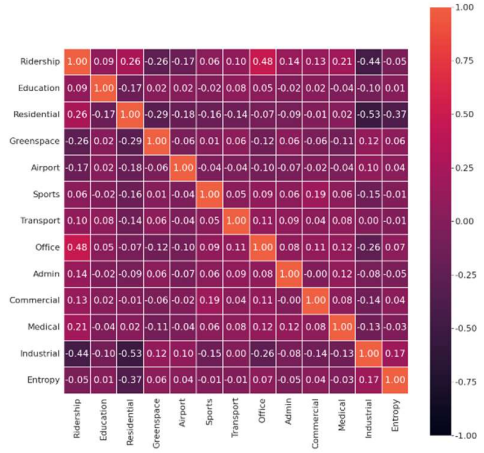

Figure 3. Spearman correlation heatmap of land use factors

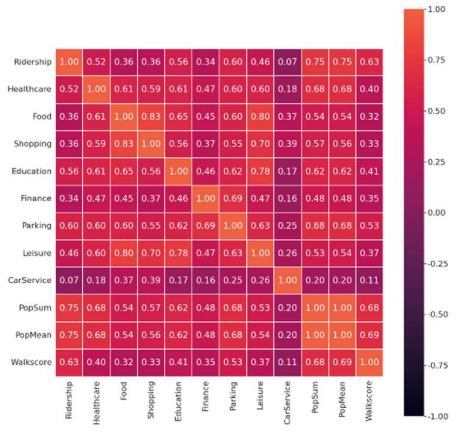

Figure 5. Pearson correlation heatmap of functionality factors

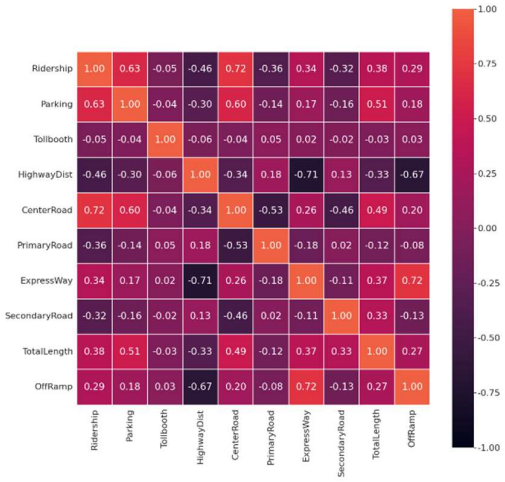

Figure 6. Spearman correlation heatmap of private car accessibility factors

Table 3. Correlated metrics from Pearson & Spearman correlation analysis(p<0.05)

	Metric	Correlation
Land use	IndustrialPer	-0.44
	Park&GreenspacePer	-0.26
	AirportPer	-0.17
	MedicalPer	0.21
	EducationalPer	0.09
	ResidentialAreaPer	0.26
	OfficePer	0.48
Inner connectivity	StraightDistFromCenter	-0.665
	RoadDistFromCenter	-0.68
	BusStop	0.12
	MetroExit	0.20
	ShortestBusStopDist	0.30
Functionality	Finance&Insurance	0.335
	Food	0.36
	Shopping	0.36
	Leisure&Recreation	0.46
	Healthcare	0.52
	Education&Research	0.56
	Walkscore	0.63
	Population	0.745
	PopulationRasterMean	0.745
Private car accessibility	HighwayExitDist	-0.46
	PrimaryRoadLength	-0.36
	SecondaryRoadLength	-0.32
	OffRamp	0.29
	ExpressWayLength	0.34
	TotalRoadLength	0.38
	Parking	0.63
	CenterRoad_Length	0.72

5.2. Negative binomial regression

Negative Binomial Regression is a widely used statistical model for count data, particularly suitable when the data exhibit overdispersion, meaning the variance is greater than the mean. Assuming dependent variable y is a count of the number of times an event occurs, following a negative binomial distribution, the model equation is as follows:

$$\ln \mu = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

where μ is the mean of y, the predictor variables $x_1, x_2, \ldots, x_p$ are given, and the population regression coefficients $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ are to be estimated.

In our study, we applied stepwise regression to select a subset of variables that most significantly affect the dependent variable, thereby minimizing the multicollinearity among independent variables. The coefficients in the negative binomial regression model, denoted as "coef", indicate the expected change in the natural logarithm of the dependent variable for a one-unit change in the predictor variable, while holding all other variables constant. In our research, these coefficients represent the relationship between changes in the predictor variables and the predicted number of metro station users. The results are presented in Table 4.

Table 4. Results from stepwise regression using Negative Binomial Regression

	coef	stderr	z	P>|z|
Intercept	2.900	0.020	144.805	0.000
RoadDistFromCenter	-0.270	0.038	-7.133	0.000
PopRasterMean	0.199	0.035	5.686	0.000
Walkscore	0.165	0.033	4.984	0.000
ResidentialPer	-0.034	0.025	-1.376	0.169
HighwayExitDist	-0.082	0.030	-2.781	0.005
Educatio	0.064	0.024	2.631	0.009
CenterRoad	0.073	0.029	2.501	0.012
Finance	-0.055	0.023	-2.373	0.018
OfficePer	0.047	0.021	2.245	0.025
EducationalPer	0.029	0.021	1.432	0.152
BusStop	-0.032	0.022	-1.496	0.135
OffRamp	-0.044	0.023	-1.938	0.053
Expressway	0.040	0.023	1.806	0.071

	coef	stderr	z	P>|z|
alpha	0.065	0.009	7.420	0.000

Regarding land use indicators, both educational and office land show a positive correlation with metro ridership. Conversely, residential land use exhibits a negative correlation with metro ridership.

For inner connectivity indicators, the distance to the city center has the most significant negative impact on ridership. Additionally, the number of bus stops shows a slight negative correlation.

Among functionality indicators, the average population per grid and Walkscore demonstrate the most significant positive impacts on ridership. The presence of educational research institutions also shows a slight positive correlation, while financial and insurance institutions exhibit a slight negative correlation with ridership.

In terms of private car accessibility indicators, the total length of toll-free expressways and central roads shows a slight positive impact on metro ridership. However, the distance to the nearest highway exit and the number of ramps exhibits a negative impact, indicating that proximity to highways might encourage private car usage over metro ridership.

5.3. Discussion

Correlation analysis and stepwise regression models have yielded several unexpected interesting results in our study since previous studies have not focused exclusively on suburban areas; instead, they have typically combined data from both city centers and suburbs in their analyses.

Firstly, for suburban stations, distance to the city center remains a significant factor influencing ridership. This may suggest that travel demand at suburban stations is more closely related to factors in the city center, for example, cultural centers or recreational services. Regarding residential functionality, high population density within residential areas and low proportion of residential land in the catchment area have a positive impact on transit ridership, reflecting the positive influence of high-density residential areas. This contrasts with the findings of An & Tong (2019) which implies that people would like to take various trips other than trips between home and work and that a significant number of trips occur outside home. It is possibly due to the difference in datasets because An & Tong (2019) utilized a mixture of city center data and suburban data. Also, population density shows significant impact in most previous studies, for example, according to Li et al. (2024), commuters are the majority of metro users which indicates that home-to-work trips are of significant amount especially in suburban areas, and this study provides another evidence supporting this commuter theory.

Secondly, contrary to the ideology of TOD development and previous research (Ewing & Handy, 2009; Singh et al., 2017) that promotes land use diversity for TODs, land use diversity did not show strong correlation with ridership. One of the reasons may be that the study area they choose is primarily inside city center where a diversified and compact land use can be easily adopted. Instead, specific types of land use and points of interest (POIs) showed correlation, such as commercial, office space, educational and research facilities, and healthcare services. Thus, it is important to increase the density of jobs (Cervero, 2006), office or commercial parcels, and healthcare services. This aligns with previous findings in city center that diverse travel needs are able to reduce imbalanced ridership between stations, improve overall system efficiency, and effectively increase ridership (Cervero & Murakami, 2008).

Thirdly, despite past research in Nanjing suggesting that bus connections can effectively increase metro station usage (Zhao et al., 2013), both Pearson correlation and regression model results show that outside the city center, high number of bus stops and short distance from metro stations weaken metro usage. Previous study and TOD index calculation (Khosravi & Abrishami, 2024) confirms that convenient transfers between the metro and buses effectively promote metro usage (Zhao et al., 2013). The empirical divergence may be because in suburban areas, with limited metro coverage and accessibility, buses serve a part of short-distance travel needs traditionally met by metro, while also diverting some citywide travel needs due to lack of express service and incomplete metro network which results in detours. In another words, bus travels faster in suburban area than in city center, therefore it has the capability to compete against metro, which is another proof that results concluded under city center context and suburban context can differ drastically.

Fourthly, in terms of the road system, the density of toll-free expressways and the distance to the nearest highway exit promote metro station usage. The result confirms previous research suggesting that expressway density may be related to transit ridership (Zhao et al., 2013) because it increases the accessibility of metro stations by driving cars. However, most of the previous studies focus on street network strongly associated with walking, for example, Cervero (2002) argues that convenient access to metro station and walkable street pattern increases ridership, which does not align with results under suburban context in this study. Conversely, the density of main and secondary roads on the ground is negatively correlated with ridership. In Shanghai's suburban area, the door-to-door travel time

by car is probably shorter that by metro, therefore the result possibly indicates that ground-level urban roads promote private car travel for short-distance, local-wide trips.

Fifthly, Walkscore significantly influences metro ridership, aligning with the original concept of TOD and almost every previous study, where a pedestrian-friendly neighborhood within a short walking distance of regional transportation services is envisioned (Calthorpe, 1993). The result suggests that in suburban areas, walkability around the station is as important as in city center.

Finally, the pseudo R-square of the final model is approximately 0.2, indicating that traditional TOD indicators included in this study may not effectively characterize ridership data for non-central city stations, or that influencing factors vary at each station depending on specific contexts. Therefore, it may be hard to find and determine a universal variable which strongly affects metro ridership in suburban areas.

In conclusion, this article, by exploring the unique scope of the study, advances the applicability of TOD theory in suburban and sub-center contexts. It offers a fresh perspective on TOD in suburban environments, addressing a gap in the existing TOD literature related to suburban areas.

6. Implications and limitations

The findings of this study, differing in some aspects from previous research, aim to explore the key factors influencing Transit-Oriented Development (TOD) in non-central areas. While the concept of TOD remains consistent, each station exhibits unique characteristics under their current contexts. Therefore, the significance of this study lies in providing policymakers and developers with a different perspective to understand TOD in suburban areas, since development patterns, even evaluation methods of TOD inside CAZ can not be copied and pasted to suburban scenarios.

For example, based on this research, transportation planning should pay more attention to the competition and cooperation among metro, buses, and private cars. Meanwhile, in land-use planning, it is crucial to balance commercial, office, residential, and healthcare land uses according to certain urban context, in order to optimize transit ridership and overall urban development. However, no matter what the urban context is, walkability remains crucial.

Although this study is specific to Shanghai, its findings can be applicable to other high-density cities with similar historical trends of railway development lagging behind urban expansion, leading to challenges in sustainable urban development.

Regarding data and methodology, the comprehensiveness of the data can be further improved. The ridership data used in this study represents only a single day (March 6, 2023, Monday) and includes only weekday data. Given the significant differences in ridership patterns between weekdays and weekends (An & Tong, 2019), future research should consider these variations to provide a more comprehensive understanding of metro station usage.

Moreover, for data with excessive zero values, such as in this study, a Zero-Inflated Negative Binomial regression model might be a more appropriate choice over the negative binomial regression model used here, which can handle over-dispersed count data but not zero inflation.

In conclusion, this study provides valuable insights into the dynamics of metro station catchment areas outside Shanghai's central business district. It offers significant implications for urban planning and development strategies, emphasizing the need for balanced land use and comprehensive transportation planning to promote sustainable urban development. Policymakers and urban planners can use these insights to optimize transit ridership and better implement the concept of TOD particularly in suburban area for a more vibrant and efficient city.

7. References

An, D., Tong, X., Liu, K., & Chan, E. H. (2019). Understanding the impact of built environment on metro ridership using open source in Shanghai. Cities, 93, 177-187.

Bach, B., van Hal, E., de Jong, M. I., & de Jong, T. M. (2006). Urban Design and traffic; a selection form Bach's toolbox. Stedenbouw en verkeer; een selectie uit de gereedschapskist van Bach. Crow.

Beijing City Lab, (2022), Data 44, Walk score of all main Chinese cities. http://www.beijingcitylab.com.

Cervero, R., & Sullivan, C. (2011). Green TODs: marrying transit-oriented development and green urbanism. International Journal of Sustainable Development & World Ecology, 18(3), 210–218.

Cervero, R. B. (2013). Linking urban transport and land use in developing countries. *Journal of Transport and Land Use*, 6(1), 7–24.

Cervero, R., Adkins, A., & Sullivan, C. (2010). Are suburban TODs over-parked?. Journal of Public Transportation, 13(2), 47-70.

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

Choi, J., Lee, Y. J., Kim, T., & Sohn, K. (2012). An analysis of Metro ridership at the station-to-station level in Seoul. Transportation, 39, 705-722.

Ewing, R., Tian, G., Park, K., Sabouri, S., Stinger, P., & Proffitt, D. (2019). Comparative case studies: trip and

parking generation at Orenco Station TOD, Portland region and Station Park TAD, Salt Lake City region. Cities, 87, 48-59.
Gakenheimer, R. (1999). Urban mobility in the developing world. Transportation Research Part A: Policy and Practice, 33(7-8), 671-689.
Gong, P., Chen, B., Li, X. et al. (2020). Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. 65(3), 182-187.
Hossein Khosravi, Motahareh Abrishami, Majid Ramezani Mehrian, Brent Chamberlain (2024). The positive impact of transit-oriented-development characteristics on Metro Station usage: A case study of Tehran's metro stations and TOD index calculation, Cities, 148, 104840.
Hrelja, R., & Rye, T. (2024). Delivering Transit Oriented Development (TOD) in low to medium density contexts. Actor relationships and market conditions in smaller Swedish cities. International Journal of Sustainable Transportation, 18(3), 236-249.
Jun, M. J., Choi, K., Jeong, J. E., Kwon, K. H., & Kim, H. J. (2015). Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. Journal of transport geography, 48, 30-40.
Kuby, M., Barranda, A., & Upchurch, C. (2004). Factors influencing light-rail station boardings in the United States. Transportation Research Part A: Policy and Practice, 38(3), 223-247.
Li, L., Zhong, L., Ran, B., & Du, B. (2024). Analysis of the relationship between metro ridership and built environment: A machine learning method considering combinational features. Tunnelling and Underground Space Technology, 144, 105564.
Li, X., Liu, Y., Gao, Z., & Liu, D. (2016). Linkage between passenger demand and surrounding land-use patterns at urban rail transit stations: A canonical correlation analysis method and case study in Chongqing. International Journal of Transportation Science and Technology, 5(1), 10-16.
Niedzielski, M. A., & Eric Boschmann, E. (2014). Travel Time and Distance as Relative Accessibility in the Journey to Work. Annals of the Association of American Geographers, 104(6), 1156–1182.
Nigro, A., Bertolini, L., & Moccia, F. D. (2019). Land use and public transport integration in small cities and towns: Assessment methodology and application. Journal of Transport Geography, 74, 110-124.
Rode, P. et al. (2017). Accessibility in Cities: Transport and Urban Form. In: Meyer, G., Shaheen, S. (eds) Disrupting Mobility. Lecture Notes in Mobility. Springer, Cham.
Shanghai Urban Planning and Land Resource Administration Bureau, 2018, Shanghai Masterplan 2017-2035,
Shanghai Metro Communication Platform. (2015).http://www.shmetro.com/qa_info_show.jsp.
Shiliang Su, Hui Zhang, Miao Wang, Min Weng, Mengjun Kang (2021). Transit-oriented development (TOD) typologies around metro station areas in urban China: A comparative analysis of five typical megacities for planning implications, Journal of Transport Geography, 90,102939.
Singh, Y. J. (2015). Measuring transit-oriented development (TOD) at regional and local scales: a planning support tool.
Woo, J. H. (2021). Classification of TOD typologies based on pedestrian behavior for sustainable and active urban growth in Seoul. Sustainability, 13(6), 3047.