Revealing the determinants of the intermodal transfer ratio between metro and bus systems considering spatial variations

Citation for the original published paper (version of record):
Wu, P., Xu, L., Zhong, L. et al (2022). Revealing the determinants of the intermodal transfer ratio between metro and bus systems
considering spatial variations. Journal of Transport Geography, 104.
http://dx.doi.org/10.1016/j.jtrangeo.2022.103415

Journal of Transport Geography 104 (2022) 103415

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Revealing the determinants of the intermodal transfer ratio between metro and bus systems considering spatial variations

Pan Wu [a], Lunhui Xu [a], Lingshu Zhong [b,*], Kun Gao [b], Xiaobo Qu [c], Mingyang Pei [a,*]

[a] Department of Civil and Transportation Engineering, South China University of Technology, Guangzhou 510641, China
[b] Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg SE-412 96, Sweden
[c] School of Vehicle and Mobility, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Keywords:
Public transit
Transfer ratio
Spatial variations
Geographically weighted regression
Smartcard data
Multiple weather effects

ABSTRACT

Buses and metros are two main public transit modes, and these modes are crucial components of sustainable transportation systems. Promoting reciprocal integration between bus and metro systems requires a deep understanding of the effects of multiple factors on transfers among integrated public transportation transfer modes, i.e., metro-to-bus and bus-to-metro. This study aims to reveal the determinants of the transfer ratio between bus and metro systems and quantify the associated impacts. The transfer ratio between buses and metros is identified based on large-scale transaction data from automated fare collection systems. Meanwhile, various influencing factors, including weather, socioeconomic, the intensity of business activities, and built environment factors, are obtained from multivariate sources. A multivariate regression model is used to investigate the associations between the transfer ratio and multiple factors. The results show that the transfer ratio of the two modes significantly increases under high temperature, strong wind, rainfall, and low visibility. The morning peak hours attract a transfer ratio of up to 57.95%, and the average hourly transfer volume is 0.94 to 1.38 times higher at this time than in other periods. The intensity of business activities has the most significant impact on the transfer ratio, which is approximately 1.5 to 15 times that of the other independent variables. Moreover, an adaptative geographically weighted regression is utilized to investigate the spatial divergences of the influences of critical factors on the transfer ratio. The results indicate that the impact of a factor presents spatial heterogeneity and even shows opposite effects (in terms of positive and negative) on the transfer ratio in different urban contexts. For example, among the related socioeconomic variables, the impact of the housing price on the downtown transfer ratio is larger than that in the suburbs. Crowd density positively influences the transfer ratio at most stations in the northern region, whereas it shows negative results in the southern region. These findings provide valuable insights for public transportation management and promote the effective integration of bus and metro systems to provide enhanced transfer services.

1. Introduction

Developing high-quality public transportation systems is inevitable in metropolises with high population density (Gao et al., 2020). Transfers between different transportation modes have also become an integral part of the public transportation system (Seaborn et al., 2009). Herein, transfers refer to those between different public transportation modes. The definition of the two transfer modes of metro-to-bus and bus-to-metro is available in reference (Seaborn et al., 2009). The topological structure of the transfer system connects a metro network with a feeder bus network. The metro network is generally more efficient and has a greater capacity (Wang et al., 2018; Yang et al., 2015). The feeder bus network is flexible, featuring multiple lines and directions. In the system, many travelers cannot directly reach their destinations. It is often necessary to transfer one or more times, especially for those who travel a medium-long distance (Huang et al., 2019). Unpleasant transfers are potentially important factors that result in the negative experience of using public transit since they increase the travel time and reduce travel efficiency (Espino and Román, 2020; Schakenbos et al., 2016). Consequently, it is crucial to identify which factors impact transfer travel and understand the relationship between these factors and the transfer ratio. This is significant to provide passengers with a

* Corresponding authors.
E-mail addresses: 201810101746@mail.scut.edu.cn (P. Wu), lhxu@scut.edu.cn (L. Xu), lingshu@chalmers.se (L. Zhong), gkun@chalmers.se (K. Gao), peim@chalmers.se (M. Pei).

https://doi.org/10.1016/j.jtrangeo.2022.103415
Received 7 September 2021; Received in revised form 10 June 2022; Accepted 5 August 2022
Available online 18 August 2022

better transfer experience and subsequently increase the attractiveness of public transportation.

The transfer ratio is affected by multiple factors (such as weather) in real-time. However, most studies focus on the transfer penalty during the morning peak hours (Espino and Román, 2020; Navarrete and de Ortúzar, 2013). Few have qualitatively and quantitatively reported the determinants of the transfer ratio (Allard and Moura, 2018; Yang et al., 2015). Furthermore, none of these studies addresses weather factors, and they do not concentrate on the spatial variations in the transfer ratio. As a result, very little is known about how the impacts of various factors, such as the weather, on the transfer ratio vary in different urban contexts. The same factor probably has different effects on the transfer ratio in different regions. For example, the impact of rain on the transfer ratio downtown is weaker than that in the suburbs because there are more rain shelters in the city center than in remote areas. In addition, the transfer ratio and distribution of related factors in urban areas vary from station to station. It is necessary to consider spatial variations when exploring the transfer ratio explicitly. In the available literature, geographically weighted regression (GWR) analysis explicitly considers local effects, thus offering a reasonable explanation of spatial phenomena (Warf, 2014) and a promising method to reveal the underlying factors of citywide transfer.

Therefore, standing in the wake of existing research, this study aims to fill the above gaps by investigating the impacts of various factors on the transfer ratio between bus and metro systems based on multivariate data sources. First, the transfer ratio and transfer time are identified based on large-scale transaction data from automated fare collection systems in Shenzhen, China. Meanwhile, various influencing factors, including weather factors (e.g., rain, wind, visibility, and temperature), socioeconomic characteristics, the intensity of business activities, and built environment factors, are obtained from multivariate sources. Second, the influences of different socioeconomic characteristics, weather conditions, the intensity of business activities related factors, built environmental attributes, and transfer-related factors on the transfer ratio are explored with a multiple linear regression, which best utilizes multivariate data from multiple sources. This paper aims to identify which factors significantly impact the transfer ratio and determine their underlying influence. However, the multiple linear regression method can only determine which factors have a more significant impact on the transfer ratio and cannot analyze and account for the spatial heterogeneity of the same factor on the transfer ratio. Third, an adaptative GWR is further performed to reveal the divergences in the effects of socioeconomic factors, the intensity of business activities related factors, built environment factors, and transfer-related factors on the transfer ratio and provides more accurate modeling concerning the transfer ratio in various contexts. These findings can be applied to expand short-term public transportation scheduling and for future transfer station planning.

This paper is organized as follows. Section 2 gives a literature review of relevant studies. Section 3 mainly describes the study area and related data. Section 4 presents our models and elaborates on our analytical approaches. Section 5 discusses the analysis results and summarizes the significant findings of the research. Section 6 concludes the study and suggests future work.

2. Literature review

Transfers are considered a crucial part of public transportation that can concentrate ridership, serve passengers traveling medium-long distances and increase public transportation accessibility. However, transfers are inconvenient due to the extra travel time (Schakenbos et al., 2016). Accordingly, improving the performance of the transfer system and promoting the transfer experience is key to increasing the attractiveness of public transportation. It is essential to determine the influencing factors of transfer and their weights for further specific transfer system improvements.

Existing studies have widely investigated transfer penalties and captured the effects of various factors on transfers (Garcia-Martinez et al., 2018; Yang et al., 2015). Most studies focus on a subjective analysis of transfer based on limited sample data obtained by manual surveys, and they lack a systematic large-scale quantitative analysis. Furthermore, few studies have attempted to untangle the relationships between the transfer ratio and the potential determining factors to provide a comprehensive understanding for decision-makers to promote the sustainable development of urban public transportation. This paper extracts large-scale transfer data from smartcard transaction data to fill the research gap from the perspective of multisource data. It collects potential influencing factors from multiple platforms to explore the impact of various factors on the transfer ratio and specifically considers the impact of weather on transfer.

However, the impact of weather on transfer is rarely examined (Cascajo et al., 2019). In the current literature, the effect of weather on transfer is mentioned only by Iseki and Taylor (Iseki and Taylor, 2009). They conducted studies from a qualitative perspective, identified weather as one factor that impacts the transfer penalty, and pointed out that many passengers make transfers only in bad weather due to the extra waiting time. There is still no consensus on the impact of weather (such as rainfall, winds, and temperature) on transfer. In contrast, the impact of weather factors on travel demand for the public transportation system has been thoroughly studied. The impact of weather on the usage of public transportation is complex. Weather conditions have been reported to significantly affect ridership and travel demand (Arana et al., 2014; Singhal et al., 2014). For example, strong winds, rainfall, and temperature have a prominent effect on public transportation (Arana et al., 2014; Li et al., 2015; Miao et al., 2019; Singhal et al., 2014; Wei et al., 2019; Yang et al., 2021). Rainfall and low temperature are negatively correlated with ridership on weekends and leisure trips (Arana et al., 2014), which provides adequate references and inspires this article to explore the influence of weather factors on the transfer ratio.

In addition, public transportation systems are also affected by many other factors (Chakour and Eluru, 2016; Cools et al., 2010; Gao et al., 2021a, 2021b; Taylor et al., 2009). These pivotal factors are summarized in Table 1. Most studies have focused on the transfer time, weather, and built environment variables on public transit ridership (Garcia-Martinez et al., 2018), while few have concentrated on the effect of these factors on transfer. Furthermore, the estimation coefficients of various factors probably vary across space. For example, the impact of transfer time on the transfer ratio in the downtown area is lower than that in the suburbs. Because the well-developed public transportation network and dense stations are in the city center, the transfer passengers are not sensitive to variations in transfer time. Moreover, existing studies do not involve spatial analysis of these influencing factors associated with the transfer ratio. The related literature typically uses GWR models to analyze the spatial impact of various factors on ridership (Ma et al., 2018; Tu et al., 2018). The results of relevant studies are summarized in Table 2. They mainly include the three aspects of the influence of factors on transfer, the impacts of weather on ridership, and the influence of various factors on public ridership based on GWR analysis.

Table 2 shows that existing studies have mainly investigated the influence of average weather conditions on daily ridership (Guo et al., 2007). Some have explored the relationship between real-time weather variables and ridership (Singhal et al., 2014), accounting for the variation in travel demand caused by hourly weather fluctuations. Moreover, weather impacts on ridership vary depending on public transportation modes, regions, weather conditions, and dates (weekends or workdays). These studies provide critical references for analyzing the impact of weather on the transfer ratio in this study.

Furthermore, as shown in Table 2, transfer-related studies are usually based on survey data to characterize transfer behavior and conduct a subjective evaluation (Espino and Román, 2020; Schakenbos et al., 2016). They discuss the impact of the transfer time and station facilities,

Table 1
Summary of the factors considered in existing studies.

Independent variables	(Choi et al., 2012)	(Cardozo et al., 2012)	(Zhao et al., 2013)	(Singhal et al., 2014)	(Arana et al., 2014)	(Schakenbos et al., 2016)	(Zhou et al., 2017)	(Chen et al., 2019)	(Wu and Liao, 2020)	(Li et al., 2021)	(Gao et al., 2021a, 2021b)	This paper
Transfer time	–	–	–	–	–	●	–	–	–	–	–	●
Metro stations	–	–	–	–	–	●	–	–	–	–	●	–
Access/egress time/costs	–	–	–	–	–	●	–	–	–	–	–	–
Inbound/outbound ridership/points of interest (POIs)	–	–	–	–	–	–	–	–	–	●	–	–
Population	●	●	●	–	–	–	–	–	–	●	●	●
Employment	●	●	●	–	–	–	–	–	–	●	●	–
Morning peak	●	–	–	–	–	–	●	–	–	–	–	●
Evening peak	●	–	–	–	–	–	●	–	–	–	–	●
Non-workdays	●	–	–	●	●	–	●	–	●	–	–	●
House rent	–	–	–	–	–	–	–	–	–	–	–	●
Housing price	–	–	–	–	–	–	–	–	–	●	–	●
Income	–	–	–	–	–	–	–	–	●	–	–	–
Temperature	–	–	–	–	●	–	●	–	●	–	–	●
Wind speed	–	–	–	●	●	–	●	–	●	–	–	●
Relative humidity	–	–	–	–	●	–		–	●	–	–	–
Visibility/fog	–	–	–	●	–	–	–	–	–	–	–	●
Rainfall	–	–	–	●	●	–	●	–	●	–	–	●
Snow	–	–	–	●	–	–	–	–	–	–	–	–
Parking lots	–	–	–	–	–	–	–	●	–	●	●	–
Business buildings	●	–	–	–	–	–	–	●	–	–	–	–
Restaurant	–	–	●	–	–	–	–	●	–	–	–	●
Entertainment	–	–	●	–	–	–	–	–	–	–	●	–
Takeaway/average meal price of restaurants	–	–	–	–	–	–	–	–	–	–	–	●
Gross domestic product (GDP)	–	–	–	–	–	–	–	–	–	–	–	●
Schools	–	–	●	–	–	–	–	●	–	–	●	●
Bus stops	–	–	–	–	–	–	–	●	●	●	●	–
Feeder bus routes	–	–	●	–	–	–	–	●	–	–	–	●
Average feeder distance	–	–	–	–	–	–	–	–	–	–	–	●
Distance from the central business district (CBD)	–	–	●	–	–	–	–	●	–	●	–	●
Road density/land mixture	–	–	–	–	–	–	–	–	–	●	●	–
Fall/winter/spring	–	–	–	●	–	–	–	–	–	–	–	–

Note: "●" indicates that the parameter was used in the study, and "-" indicates that the parameter was not used in the study.

such as information availability and elevator availability, and evaluate the penalized factors (Navarrete and de Ortúzar, 2013; Schakenbos et al., 2016). However, the sample size obtained through surveys is relatively small. We can achieve only the average influence of these factors on transfers rather than the real-time impacts by mining survey data. Unlike these studies, this paper uses a high-quality multivariate dataset, including the real-time transfer ratio and transfer time extracted from smart data recorded by the automatic fare collection system. More importantly, the existing studies have discussed the spatial impact of the factors on transit ridership by using the GWR(Ma et al., 2018). However, they do not analyze the spatial impact of various factors on transfer ratio in different spaces. To fill this gap, this paper thoroughly analyzes the divergences of the influences of individual factors on the transfer in spatial regions using the GWR. In particular, the spatial difference in the impact of the same factor on the transfer ratio is revealed.

Accordingly, this study contributes to the literature in the following aspects.

- This study analyzes the factors influencing the transfer ratio using a massive smart card (SC) dataset and global positioning system (GPS) coordinate data at a fine-grained temporal scale. A comprehensive understanding of different factors that affect the transfer ratio at the station level is obtained.
- We explore the impact of real-time weather on transfer, which is highly important for understanding the relationship between weather and the transfer ratio. We identify the incentive/disincentive factors for transfer passengers, which can help decision-makers reduce the adverse effects of factors for the subsequent planning and construction of new metro stations.
- An adaptative GWR is utilized to investigate the spatial divergences of the influences of key factors on the transfer ratio between bus and metro systems. This paper analyzes the difference in the spatial influence of the same factors on the transfer ratio, which can help better understand the spatial heterogeneity of the determinants of the transfer ratio.

3. Study area and data

This study utilizes multisource datasets, including smartcard data and meteorological data, socioeconomic data, the intensity of business activities related data, and built environment data in Shenzhen. All data were collected in 2017 so that the multisource data could be integrated. This section provides an overview of the geographical background of the study area, its public transportation systems, and the sources of the datasets.

Table 2
Summary of previous studies on the relationship between various factors and transit travel.

Author &year	Subject of study	Data collection	Remarkable findings (important variables)	Precision	Period	Models
(1) Analysis of the influencing factors of public transportation transfer behavior						
(Espino and Román, 2020)	Transfers for bus users	SP, data collected from bus trips, $N = 2416$ observations	The transfer waiting time is critical.	–	–	Mixed logit models, latent class models
(Allard and Moura, 2018)	Transportation transfer quality effect the mode choice	SP survey, $N = 9976$	The study presented a framework for determining how variables influence perceived transportation, determining the number of travelers.	hourly	2014	Mixed logit models
(Garcia-Martinez et al., 2018)	Transfer penalties in multimodal public transportation	RP and SP survey, $N = 295$	Longer trips may be preferred over faster alternatives with transfers.	–	The morning peak period (7:00–10:00) for five days	Multinomial logit model
(Schakenbos et al., 2016)	Transfer disutility between bus/tram/metro and trains	SP, total $N = 1054$	The transfer time has a significant influence on transfer disutility.	–	–	Mixed logit models
(Navarrete and de Ortúzar, 2013)	Transfers of the metro and bus systems	A stated choice survey and data from Transantiago	Walking time and the transfer wait time are the most penalized types of time.	20 min	Peak hours, the months of July and August 2009	Mixed logit models
(2) Impacts of weather on transit ridership						
(Wu and Liao, 2020)	Weather and travel mode impact ridership	Questionnaire survey, metro ridership, weather data	Leisure travel is more affected by extreme weather. The weather has a significant influence on weekend travel.	daily	Jan. 1st, 2014, to Jun. 30th, 2018	Logit model
(Zhou et al., 2017)	Weather impacts on public transportation ridership	smartcard data, weather data (30 days)	Weather affects public transportation more than other factors.	hourly	The entire month of September 2014	Multivariate modeling approach
(Arana et al., 2014)	Weather impacts on transit ridership	Ridership data from a computer-aided dispatch system, weather data	Wind and rain lead to fewer trips, an increase in temperature causes an increase in the number of trips, weather indicators have negative effects on ridership.	daily	All weekends in 2010 and 2011, Oct. 1st, 2011, to Sept. 30th, 2012	Multilinear regression models
(3) Impacts of various variables on public ridership based on GWR analysis						
(Li et al., 2021)	Transfer distance (egress and access)	The Mobike trip dataset and metro smartcard dataset	The GWR model outperforms the ordinary least squares (OLS). The transfer distances are correlated with ridership, population density and distance from the CBD.	daily	(6:00–23:00) on 23 weekdays in August 2016	OLS, GWR
(Ma et al., 2018)	Transit ridership	Smartcard data, POI information	There are time-dependent effects of the built environment on ridership, significantly better goodness of fit was observed for geographically and temporally weighted regression (GTWR).	hourly	One month	OLS, GWR, GTWR
(Yu and Peng, 2019)	Ridesourcing demand	Ridesourcing trip data from transportation network	Geographically weighted Poisson regressions (GWPRs) improve the modeling fit. Built environment factors significantly impact ridesourcing demand.	daily	2016–2017	GWPR, spatial error (SER)
(Chen et al., 2019)	Metro ridership	Smartcard data, online POI data	A GWR model using the Minkowski distance (MD-GWR) achieves better goodness of fit. POI data have a better statistical performance.	hourly	From Sept. 4th, 2017 to Sept. 17th, 2017	OLS, GWR model using the Euclidean distance (ED-GWR), MD-GWR
(Tu et al., 2018)	Daily ridership of buses, the metro system, and taxis	Vehicle global positioning system (GPS) trajectories, smartcard data, spatial data	Employment, mixed land use, and road density significantly affect the ridership of each mode.	daily	From Sept. 24th to Sept. 30th, 2014	OLS, GWR
(Zhao et al., 2013)	Metro ridership	Annual average weekday station ridership, 2010 census data provided	Population, business/office floor area, the CBD, and the number of education buildings, entertainment venues, and shopping centers are related to ridership.	daily	Annual average weekday ridership at stations	OLS

Note: SP indicates a stated preference experiment, and RP indicates a revealed preference experiment.

3.1. Study area

Shenzhen is an appropriate study area because of the prevalent public transportation, the relatively high proportion of smartcard trips, and the dense public transportation network. Shenzhen includes ten administrative districts in 2017: Guangming District (GMD), Longhua District (LHAD), Baoan District (BAD), Nanshan District (NSD), Futian District (FTD), Luohu District (LHD), Yantian District (YTD), Longgang District (LGD), Pingshan District (PSD) and Dapeng District (DPD). According to the Shenzhen Statistical Yearbook 2018, at the end of 2017, the permanent population was 12.52 million, and the population density was 6234 persons per sq. km. The public transportation system consisted of nine metro lines with 166 stations and 992 bus routes with 17,430 buses. The total number of bus trips was 1654.25 million, dropping 11.44% compared with the previous year. In the past four years, the bus passenger volume has decreased continuously. In contrast, the total number of metro trips was 1655.45 million, 27.62% higher than the previous year. Metro ridership increased every year since some travelers may have moved from bus travel to metro travel. More importantly, most metro passengers are concentrated in some stations of metro lines 1 to 5. Consequently, this study explores these five metro lines (i.e., metro lines 1, 2, 3, 4, and 5), 52 metro stations, their surrounding bus stops, and their bus routes to avoid biases. The study area, metro lines, and metro stations are shown in Fig. 1.

October 2017 is chosen as the study period in this paper for two main reasons. First, October contains multiple dates, including 16 workdays, eight holidays (including National Day and Mid-Autumn Festival), and six ordinary weekends. The period is from 7:00 to 22:59, consistent with the operation time of all metro stations. Second, the weather in October was hot and changed significantly. The average temperature was 25.8 °C, with a temperature ranging from 17.6 °C to 34.2 °C. The total cumulative rainfall was 102.2 mm. Hence, the choice of this month makes it easier to capture the relationship between weather and transfer.

3.2. Data sources and data processing

This section describes the raw data sources, data processing, and statistical characteristics of the variables. The hierarchy of the variables and data sources are shown in Fig. 2. The six categories of independent variables consist of weather, socioeconomic, the intensity of business activities, built environment, transfer-related, and date variables. The spatial configuration and characteristics of the distribution of partially important variables regarding the weather, socioeconomic, intensity of business activities, and built environmental data are shown in Fig. 3. The weather variables change in real-time with temporal characteristics. The other variables vary from station to station, showing significant spatial variation in the study area. Therefore, it is imperative to discuss the spatial characteristics of these variables.

The datasets from multiple sources include one month of meteorological data from all weather stations in Shenzhen, socioeconomic data from the Shenzhen Statistical Yearbook, and one month of public transportation smartcard transaction data covering all cardholders. Additional data include the location coordinates of all metro stations and bus stops, built environment data, the intensity of business activities related data, and vector maps of the districts in Shenzhen. All data were collected from 7:00 to 22:59 Oct. 1st to Oct. 30th, 2017. Since some metro stations and bus routes operated later than 6:00 and closed earlier than 23:59, to obtain stable and reliable results, we selected the available data from 7:00 to 22:59.

The definitions of the variables are described in Table 3. The attributes considered in our analysis include dependent variables (transfer ratio) and independent variables. The independent variables mainly include weather, socioeconomic, intensity of business activities, built environment, transfer-related, and date variables. The detailed description of the independent variables and the dependent variables is as follows.

3.2.1. Independent variables

3.2.1.1. Weather variables. The weather data contained 480 records from the national meteorological administration of China, and raw data were collected at hourly intervals. The selected dataset includes temperature, wind, visibility, and rainfall. The four weather variables were extracted from the meteorological observation values recorded by the National Meteorological Administration of China in October 2017. The temperature in this paper is the highest temperature value per hour. The wind is defined as the average value observed during a given hour. Similarly, visibility is the minimum value in an hour. The raw rainfall is the cumulative hourly precipitation, and the rainfall is represented by a dummy variable coded 0 or 1 in this paper. The distribution of the four weather variables over one month is shown in Fig. A.1 of Appendix A.

3.2.1.2. Socioeconomic variables. The socioeconomic variables include house rent, housing price, and Geographical GDP. The Geographical GDP denotes the geographical gross domestic product (GDP) near each metro station, is calculated by the weighted average GDP in the adjacent administrative region and the distance from the metro station to the adjacent administrative region. The GDP of the administrative region came from the Shenzhen Statistical Yearbook 2018. The house rent and housing price were taken from the rental and housing prices published on major rental and housing sales websites (Lianjia.com, Anjuke) in October 2017.

3.2.1.3. The intensity of business activities related variables. Multiple factors in this study characterize the intensity of business activities near the metro station. The number of restaurants, the average meal price of restaurants, and the monthly sales of takeaways near each metro station were obtained from the Meituan app, Eleme app, and Public Comments app. The crowd density of metro stations and distance from the central business district (CBD) came from the Baidu Map official website.

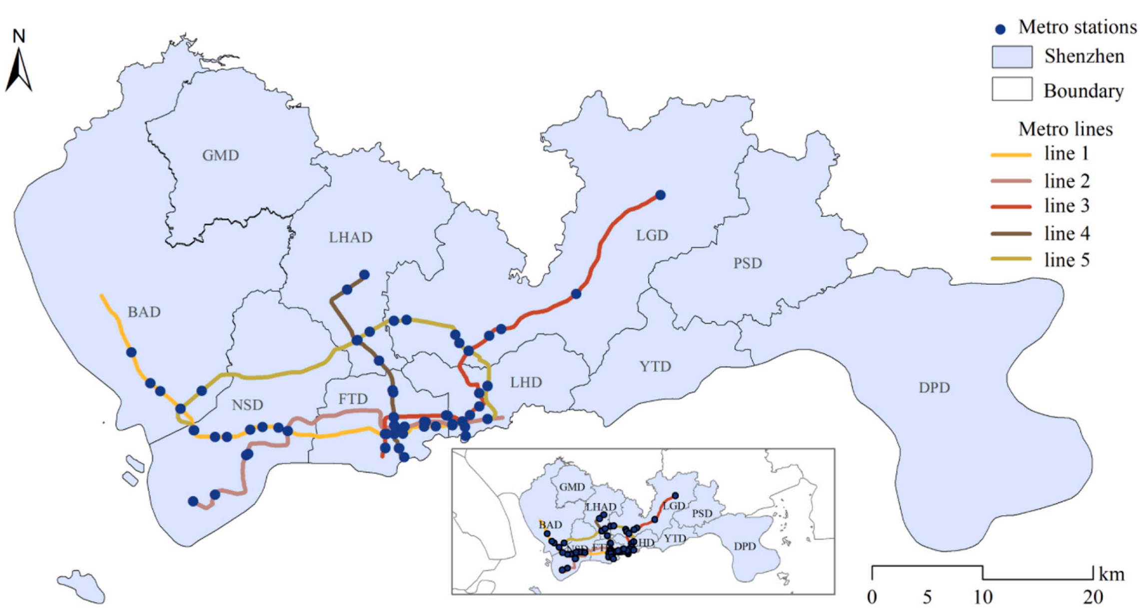

Fig. 1. Geographical location, metro stations, and metro lines in Shenzhen, China.

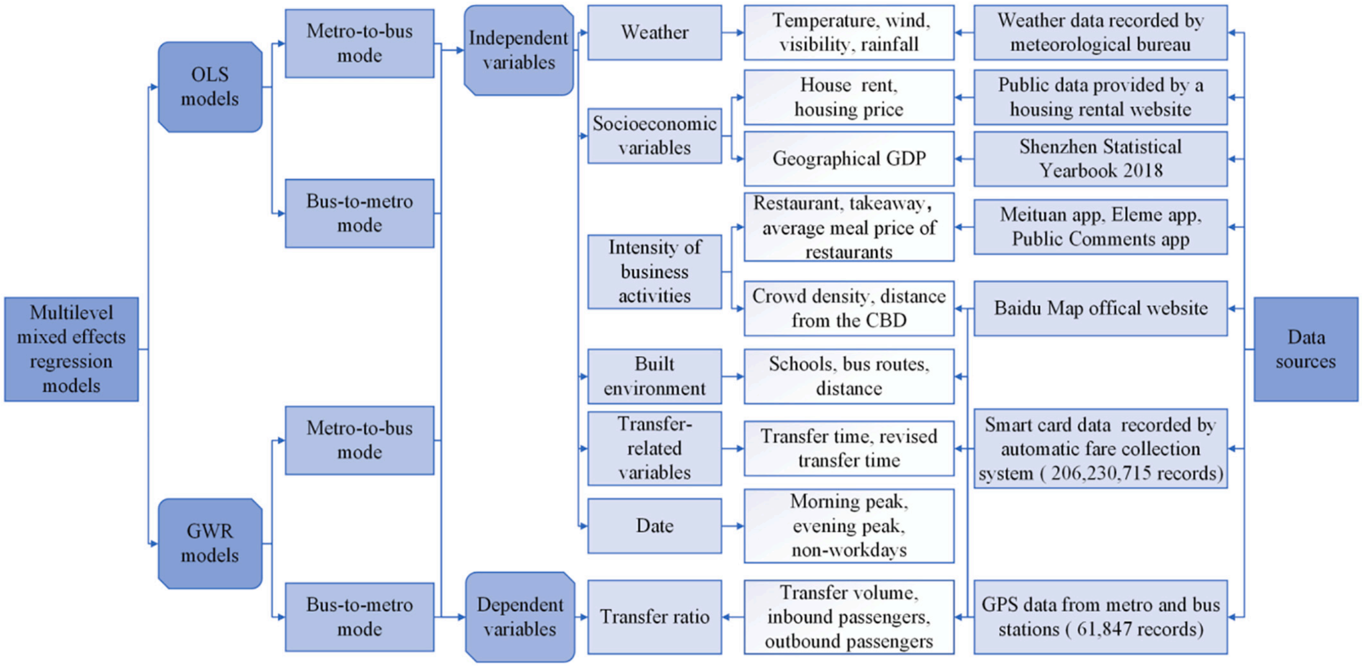

Fig. 2. The hierarchy of the variables in the models and data sources.

3.2.1.4. Built environment variables. The built environment factors consist of multiple variables. The number of schools was collected from the Baidu Map official website. The number of feeder bus routes and feeder distances between bus stops and metro stations was calculated by smartcard data and GPS data from metro stations and bus stops. The calculation method was based on references (Huang et al., 2019). The selected smartcard data from the bus and metro systems recorded passenger trips, including 2,026,230,715 swipe records. Every record contained details about the identification card number, card type, metro station, bus line number, and transaction timestamp. The GPS data from the metro and bus stops contained 61,847 records.

As shown in Table 3, bus routes 500 m denote the number of bus routes within 500 m of a metro station. Bus routes 500–1000 m represent the number of bus routes between 500 m and 1000 m near the metro station. Distance 500 m represents the average distance between bus stops and a metro station within 500 m of a metro station. Distance 500–1000 m denotes the average distance between bus stops and a metro station within 500 to 1000 m of a metro station. Among the above four variables, bus routes 500 m and distance 500 m are aggregated within 500 m from a metro station, while bus routes 500–1000 m and distance 500–1000 m are aggregated within 500 to 1000 m from a metro station. Because existing studies show that the service radius of metro stations and bus stops is 500 to 1000 m (Chakour and Eluru, 2016; Wang et al., 2018), some researchers have proposed the use of 1000 m as a limit for walking distance (Munizaga et al., 2014; Munizaga and Palma, 2012). Moreover, some related literature also considers the number of bus routes and the distance of bus stops from metro stations as independent variables to explore their effects on public transit (Hadas and Ranjitkar, 2012). Furthermore, the variables are categorized and analyzed according to the buffer distance (Chakour and Eluru, 2016). Therefore, we chose the number of bus routes and distance in the service range of 500 and 1000 m near the metro station as independent variables, which can indicate the number of bus routes connected near the metro and the average distance between the metro station and bus stops without utilizing quartiles, standard deviations, etc. In addition, existing studies have shown that there are also linear effects between these variables and metro ridership (Chen et al., 2019), and spatial models have also been used to explore the spatial effects of these variables on metro ridership. The main purpose of this paper is to reveal the determinants of the intermodal transfer ratio between metro and bus systems and to explore the local spatial effects of these variables on the transfer ratio by using a GWR. The GWR model can reveal the global linear influence and local spatial effects of the independent variables on the transfer ratio, although there are also nonlinear effects between built environment variables (bus routes and distance) and a metro station. Given the limitation of the length of the manuscript, we did not explore in this paper the nonlinear effect of these independent variables on the transfer ratio. However, in future research work, we will further explore the nonlinear effects between these variables and a metro station to further understand the logical relationship behind this.

3.2.1.5. Transfer-related variables. The transfer-related variables mainly include the transfer time (metro-to-bus mode) and revised transfer time (bus-to-metro mode). We mined passenger trip data and transfer data from smartcard and GPS data to understand passengers' travel behavior characteristics. The transfer process between the metro system and the bus system is shown in Fig. B.1 of Appendix B. The transfer identification method is based on references (Huang et al., 2019; Zhao et al., 2019, 2017). This paper has made some improvements on this basis and adopted a dynamic identification method considering spatiotemporal information. First, we calculate the transfer time of all transfer passengers for the metro-to-bus mode, which is the time difference between the time exiting the metro and the subsequent time boarding a bus. Second, the transfer time is calculated at an hourly scale with an upper bound of 40 min. Third, similarly, we also calculate the revised transfer time of all transfer passengers for the bus-to-metro mode, which is the time difference between the time when passengers board a bus and then subsequently enter the metro. Then, the transfer time is recorded at an hourly scale with an upper bound of 50 min. Finally, the 95th percentile of the filtered transfer time (revised transfer time) is regarded as the metro-to-bus (bus-to-metro) transfer time threshold.

3.2.1.6. Date variables. Fig. C.1 to Fig. C.3 in Appendix C. show significant differences in public transportation travel on different dates. It is meaningful to discuss the influence of different dates on the transfer ratio. As a result, based on the distribution characteristics of the transfer ratio, the dates are divided into four categories, namely, off-peak

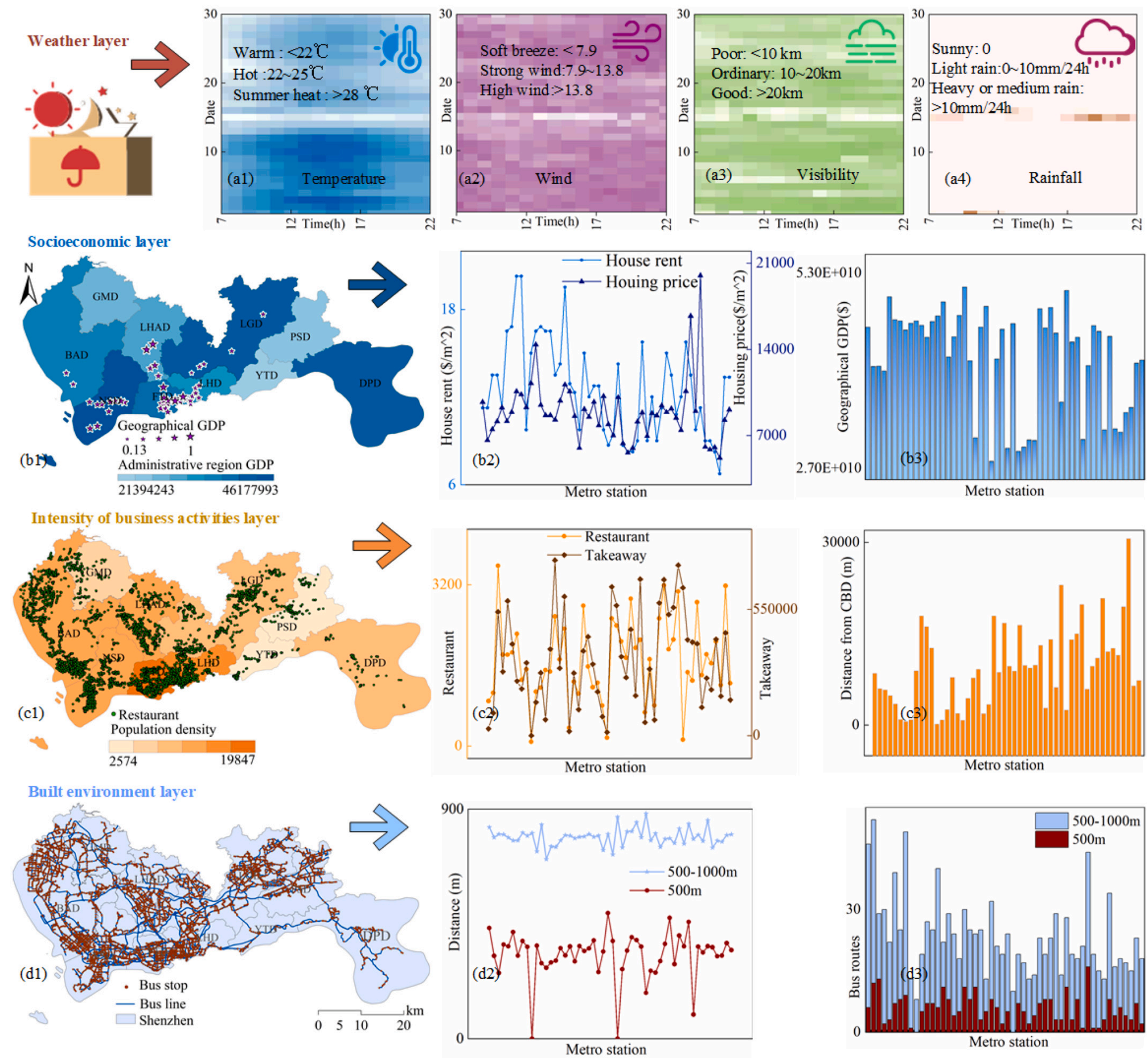

Fig. 3. The spatial configuration of the weather variables, socioeconomic variables, intensity of business activities related variables, and built environment variables in Shenzhen.

(baseline), morning peak, evening peak, and non-workdays, represented by a dummy variable taking the value of {0,1}.

3.2.2. *Dependent variables*

The hourly transfer ratio (dependent variables) equals the ratio of transfer volume to the number of inbound or outbound passengers. The number of hourly inbound and outbound passengers is derived from smartcard transaction data at metro stations. Furthermore, the transfer volume is filtered based on the transfer time (revised transfer time) threshold. There are two main reasons for choosing the transfer ratio as the dependent variable. First, the transfer ratio is more mobile and expandable than the transfer volume and can be used in studies related to transfer in all cities. Second, the transfer ratio can effectively reduce errors in the results caused by the transfer volume that is too large or too small at some metro stations. Therefore, the hourly transfer ratio of each metro station is chosen as the dependent variable in this paper. Fig. 4 below shows the characteristics of the spatial distribution of the dependent variable.

Accordingly, the definitions and descriptive statistics of the variables are described in Table 3. Most of the continuous variables are easy to understand and self-explanatory. However, the statistical characteristics of various variables are quite different. Most of the independent variables are continuous and have different dimensions. To eliminate the influence of different dimensions on the results, all continuous variables are standardized before performing the model analysis.

4. Methodology

A multivariate regression (MLR) is an effective method to estimate unknown factors, and the most widely used MLR method is ordinary least squares (OLS) (Tu et al., 2018). We use MLR to determine which factors significantly affect the transfer ratio and to identify the potential

Table 3
Definitions and descriptive statistics of variables.

Variables	Definitions	Unit	Mean	Sd.
Dependent variables				
Transfer ratio	The transfer ratio of the metro-to-bus mode is equal to the transfer volume divided by the number of outbound passengers.		0.094	0.09
	The transfer ratio of the bus-to-metro mode is equal to the transfer volume divided by the number of inbound passengers.		0.096	0.09
Independent variables				
Weather variables				
Temperature	Highest temperature per hour.	°C	27.01	3.45
Wind	Average wind speed per hour.	m/s	2.42	1.16
Visibility	Minimum visibility per hour.	m	32.02	10.89
Rainfall (dummy variable)	Rainfall exceeds 0.		0.073	0.26
Socioeconomic variables				
House rent	Average house rent near metro stations.	$\$/m^2$	12.57	3.26
Housing price	Average housing price near metro stations.	$\$/m^2$	8932.66	2585.91
Geographical GDP	Geographically weighted GDP near metro stations.	$	40,840,868,942	6,708,550,059
The intensity of business activities related variables				
Restaurant	The number of restaurants within 1000 m of a metro station.		1705.26	788.44
Average meal price of restaurants	Average meal price of restaurants within 1000 m of a metro station.	$/person	11.86	3.58
Takeaway	Average monthly sales of takeaway food within 1000 m of a metro station.		331,106	203,747
Crowd density	The hourly density of pedestrian flow near a metro station.		5.62	1.19
Distance from the CBD	Distance of a metro station from the CBD.	m	9600.29	6604.44
Built environment variables				
Schools	The total number of schools within 1000 m of a metro station.		8	3.36
Bus routes 500 m	The number of bus routes within 500 m of a metro station.		5.67	3.48
Bus routes 500–1000 m	The number of bus routes is between 500 m and 1000 m near the metro station.		17.93	8
Distance 500 m	Average distance between bus stops and metro stations within 500 m.	m	328.20	91.65
Distance 500–1000 m	Average distance between bus stops and metro stations within 500 m–1000 m.	m	791.85	34.43
Transfer-related variables				
Transfer time	Hourly difference time threshold from exiting the metro to boarding a bus.	min	29.47	6.71
Revised transfer time	Hourly difference time threshold from boarding a bus to boarding the metro.	min	40.42	8.18
Date variables (dummy variable)				
Morning peak	7–9 a.m. on weekdays in October 2017.		0.096	0.29
Evening peak	5–8 p.m. on weekdays in October 2017.		0.14	0.34
Non-workdays	Including ordinary weekends and national holidays, 1–8, 14, 15, 21, 22, 28, and 29 in October 2017.		0.46	0.50

Note:1. Sd. = standard deviation. 2. The rainfall and date variables are dummy variables, and the others are continuous variables. 3. For transfers from buses to the metro, the bus smartcard data contain only the bus boarding time and bus routes and lack the bus alighting time and stop information. The transfer time is revised for the bus-to-metro mode in this study, and defined as the time difference between boarding a bus and boarding the metro, including the previous trip's in-vehicle travel time.

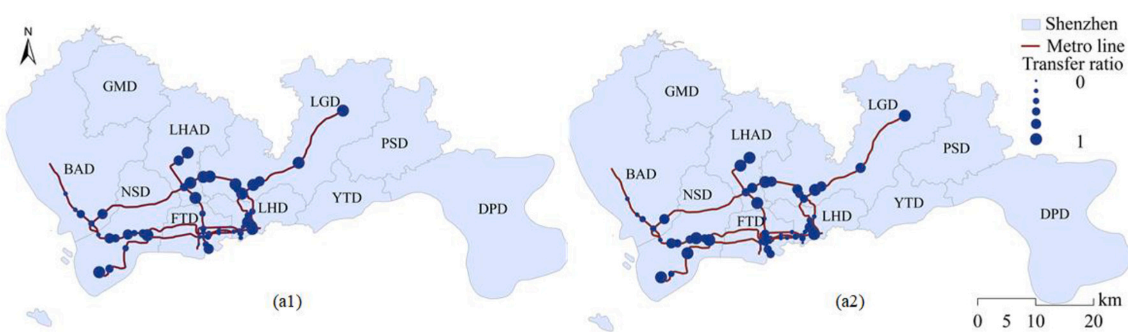

Fig. 4. Spatial distribution of the transfer ratio: (a1) the metro-to-bus mode and (a2) the bus-to-metro mode.

effects of various factors (including weather elements, socioeconomic variables, the intensity of business activities related variables, built environment variables, transfer-related variables, and date factors). However, MLR cannot analyze the spatial heterogeneity of various factors associated with the transfer ratio. GWR analysis explicitly considers local effects in the available literature, explaining spatial phenomena (Warf, 2014). Therefore, to further explore the disparity impact of significant variables on the transfer ratio in different regions, this paper uses GWR to further model the transfer ratio and various factors (including weather elements, socioeconomic variables, the intensity of business activities related variables, built environment variables, and transfer-related variables).

4.1. Global models

Theoretically, there is probably some correlation between variables. Before calculating the MLR models, possible collinearity between independent variables is diagnosed. The variance inflation factors (VIFs) are calculated for all models. A VIF <10 is usually considered acceptable, indicating almost no multicollinearity in the independent variables (Yan et al., 2019). Multilinear regression models based on OLS estimation are used to examine the influence of various factors on the transfer ratio. The formula is shown in Eq. (1):

$$\mathbf{y}_i = \beta_0 + \sum_{j=1}^{p} \beta_j \boldsymbol{\vartheta}_{ij} + \boldsymbol{\varepsilon}_i \tag{1}$$

where $\mathbf{y}_i$ denotes the dependent variable vector, representing the transfer ratio at metro station $i \in \{1, 2, \ldots n\}$. $\boldsymbol{\vartheta}_{ij}$ is the vector of independent variable j at metro station i, β_j denotes the estimated coefficient $j \in \{1, 2, \ldots p\}$, and ε_i is the residual term. Two MLR models are established to model the effects of different variables on the transfer ratio. As reproducible research, this part of the work is mainly implemented with the RStudio tool (Zheng, 2021).

4.2. Spatial autocorrelation test

Before using the spatial regression model, the spatial autocorrelation of the independent variables should be tested. Moran's I index is widely used to assess the spatial autocorrelation of independent variables (Chen et al., 2019).

Given the geographic locations of the metro stations (indexed by i or k), the Moran's I index of all independent variables $\boldsymbol{\vartheta}$ can be calculated to determine the global spatial autocorrelation. Moran's I index is calculated in Eq. (2):

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}} \bullet \frac{\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}(\vartheta_i - \overline{\vartheta})(\vartheta_k - \overline{\vartheta})}{\sum_{i=1}^{n}(\vartheta_i - \overline{\vartheta})^2} \tag{2}$$

where n is the number of metro stations, $\overline{\vartheta}$ is the mean value of the independent variable ϑ, and a_{ik} is the spatial weight between metro station i and metro station $k \in \{1, 2, \ldots n\}$. Moran's I index ranges from -1 to 1. If $I > 0$, then the independent variable has a positive spatial autocorrelation, if $I < 0$, then the independent variable has a negative spatial autocorrelation, and if $I = 0$, then the independent variable is spatially random.

4.3. Geographically weighted regression

Brunsdon et al. (1999) first proposed the GWR model to explore the spatial nonstationarity of spatial data. GWR is an extended form of linear regression used to model the spatially varying relationships of variables. A GWR model considers the spatial effects of independent variables and can adequately explain the variation in the variables across space. In this paper, to further explore the influence variability of the factor on the transfer ratio in different spaces, GWR is used to model the spatial relationship between the transfer ratio and various factors. The formula is as follows in Eq. (3):

$$\mathbf{y}_i(u_i, v_i) = \beta_0(u_i, v_i) + \sum_{j=1}^{p} \beta_j(u_i, v_i) \boldsymbol{\vartheta}_{ij}(u_i, v_i) + \boldsymbol{\varepsilon}_i(u_i, v_i) \tag{3}$$

where $\mathbf{y}_i(u_i, v_i)$ is the dependent variable vector, representing the transfer ratio at location (u_i, v_i). $\boldsymbol{\vartheta}_{ij}(u_i, v_i)$ is an independent variable vector at location (u_i, v_i), $\beta_j(u_i, v_i)(j = 0, 1, \ldots, p)$ is the estimated coefficient at location (u_i, v_i), and $\varepsilon_i(u_i, v_i)$ is the residual term at location (u_i, v_i).

According to the first law of geography presented by Tobler (Taylor and Mahmassani, 1997), the closest metro stations have a more significant correlation with each other. For a given geographic location (u_i, v_i), locally weighted least squares can be used to estimate $\beta_j(u_i, v_i)$ in Eq. (4):

$$min \sum_{i=1}^{n} \left[y_i(u_i, v_i) - \sum_{j=1}^{p} \beta_j(u_i, v_i) \vartheta_{ij} \right]^2 w_i(u_i, v_i) \tag{4}$$

where $w_i(u_i, v_i)_{i=1}^{n}$ is a spatial weight at location (u_i, v_i). Let $\boldsymbol{\beta}\ (u_i, v_i) = (\beta_0(u_i, v_i), \beta_1(u_i, v_i), \ldots, \beta_p(u_i, v_i))^T$. The local least squares estimate of $\boldsymbol{\beta}\ (u_i, v_i)$ at (u_i, v_i) is calculated based on Eq. (5):

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = \left(\boldsymbol{\vartheta}^T \mathbf{W}(u_i, v_i) \boldsymbol{\vartheta}\right)^{-1} \boldsymbol{\vartheta}^T \mathbf{W}(u_i, v_i) \mathbf{Y} \tag{5}$$

$$\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_p), \mathbf{X}_j = (x_{1j}, x_{2j}, \ldots, x_{nj})^T$$

$$\mathbf{Y} = (y_1, y_2, \ldots, y_n)^T$$

$$\mathbf{W}(u_i, v_i) = Diag(w_1(u_i, v_i), w_2(u_i, v_i), \ldots, w_n(u_i, v_i))$$

Based on reference (Li et al., 2021; Tu et al., 2018), the Gaussian kernel function is used to estimate the spatial effects of the variables in Eq. (6):

$$w_{ik} = \begin{cases} exp\left[-\left(\frac{d_{ik}}{r} \right)^2 \right], & d_{ik} < \mathrm{D} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where d_{ik} denotes the distance between metro stations i and k, and r is the bandwidth. The Gaussian kernel function is the most frequently used since it best fits the irregular spatial distribution of observations. A large bandwidth may underestimate the spatial effects, while a small bandwidth may result in overfitting. Accordingly, an adaptive bandwidth is used to mitigate the problem posed by the spatial variance. The optimal bandwidth is determined by finding the corresponding value that results in the minimum corrected Akaike information criterion (AICc), which can help avoid overfitting and obtain high-quality models.

5. Results and discussion

This section presents and discusses the impacts of various factors on the system-wide and station-level transfer ratio between different travel modes.

5.1. Results of the global models

Global models (MLR) were used to confirm which factors affected the transfer ratio for two separate transfer modes (metro-to-bus and bus-to-metro). In theory, different variables may be relevant. For example, a high wind speed might be associated with heavy rainfall (Zhou et al., 2017). Therefore, before developing the multivariate analysis models, the possible collinearity between the independent variables was examined. The examination results are shown in Table 4. The level of the VIFs was calculated to be no >4.02, which suggests that no strong multicollinearity existed among the variables.

To discriminate which factors significantly influenced the transfer ratio of the metro-to-bus mode and the bus-to-metro mode, this paper utilized MLR to explore the relationship between the transfer ratio and various factors. The results of the two modes perform well in the goodness of fit. Most independent variables are significant at the 0.01 confidence level. The independent variables can explain 40.37% and 34.03% of the variance in the transfer ratio. Therefore, the MLR model

Table 4
The results of MLR models.

Variables	The metro-to-bus mode				The bus-to-metro mode			
	Coefficient	t statistics	*p*-value	VIF	Coefficient	t statistics	p-value	VIF
Intercept	0.42 ***	48.14	<0.0001		0.25 ***	26.25	<0.0001	
Weather variables								
Temperature	0.05 ***	11.84	<0.0001	1.72	0.04 ***	8.85	<0.0001	1.69
Wind	0.03 ***	5.01	<0.0001	1.39	0.03 ***	5.24	<0.0001	1.39
Visibility	−0.02 ***	−4.12	<0.0001	1.70	−0.03 ***	−6.65	<0.0001	1.69
Rainfall	0.02 ***	4.24	<0.0001	1.46	0.01 **	3.13	0.002	1.46
Socioeconomic variables								
House rent	0.03 ***	4.48	<0.0001	3.18	0.02 **	2.62	0.009	3.15
Housing price	−0.18 ***	−30.21	<0.0001	1.79	−0.15 ***	−22.45	<0.0001	1.82
Geographical GDP	0.05 ***	12.04	<0.0001	2.14	0.04 ***	9.87	<0.0001	2.15
The intensity of business activities related variables								
Restaurant	0.02 ***	4.39	<0.0001	2.53	0.003	0.51	0.613	2.60
Average meal price of restaurants	−0.30 ***	−46.58	<0.0001	4.02	−0.31 ***	−43.54	<0.0001	3.91
Takeaway	−0.09 ***	−19.79	<0.0001	2.13	−0.07 ***	−13.36	<0.0001	2.16
Crowd density	0.07 ***	14.84	<0.0001	1.22	0.13 ***	26.52	<0.0001	1.25
Distance from the CBD	0.16 ***	27.83	<0.0001	2.39	0.15 ***	23.79	<0.0001	2.37
Built environment variables								
Schools	−0.02 **	−2.78	0.005	1.83	−0.01 *	−2.30	0.022	1.81
Bus routes 500 m	−0.03 ***	−7.80	<0.0001	1.38	-0.09 ***	−19.49	<0.0001	1.36
Bus routes 500–1000 m	0.12 ***	24.47	<0.0001	2.42	0.13 ***	23.14	<0.0001	2.38
Distance 500 m	−0.09 ***	−17.14	<0.0001	1.34	−0.15 ***	−26.87	<0.0001	1.41
Distance 500–1000 m	−0.11 ***	−24.08	<0.0001	1.28	−0.08 ***	−15.75	<0.0001	1.31
Transfer-related variables								
Transfer time	−0.20 ***	−40.90	<0.0001	1.21	/	/	/	/
Revised transfer time	/	/	/	/	0.03 ***	5.33	<0.0001	1.28
Date variables								
Morning peak	0.02 ***	5.32	<0.0001	1.42	0.05 ***	14.74	<0.0001	1.45
Evening peak	−0.01**	−3.08	0.002	1.32	- 0.004	−1.41	0.159	1.30
Non-workdays	−0.01***	−4.05	<0.0001	1.39	0.004 *	2.06	0.039	1.40
Diagnostic statistics								
Observations	23,817				22,787			
Multiple R-squared	0.4042				0.3409			
Adjusted R-squared	0.4037				0.3403			
Residual sum of squares	359.33				396.7			
F-statistic	768.7				560.7			
p-value	<0.0001				<0.0001			
AIC	−32,250.76				−27,591.91			
AICc	−32,250.71				−27,591.86			

Note: *, **, and *** denote significance at confidence levels of 95%, 99%, and 99.9%, respectively. '/' indicates that the variables are not considered in models. The rainfall and date variables in the above table are dummy variables, and the other variables are continuous. The continuous variables were normalized before the regression analysis.

provides a basic understanding of the variation in the transfer ratio. A detailed analysis is conducted below.

Weather factors are found to influence the transfer ratio significantly. Temperature, wind, and rainfall positively impact the transfer ratio, whereas visibility shows the opposite effect. The temperature has the largest impact on the transfer ratio among the weather variables and is approximately 1.66 to 4 times that of the other weather factors. Moreover, travelers are more willing to transfer under high temperatures, rainy conditions, low visibility, or strong wind weather since it is inconvenient and unsafe to ride a bicycle or electric bicycle. Thus, many people will choose feeder buses connected to the metro.

Socioeconomic factors have significant impacts on the transfer ratio. Housing price negatively influences the transfer ratio, whereas geographical GDP and house rent show the opposite effect. Among the related socioeconomic factors, housing price has the most significant impact on the transfer ratio, which is approximately 3.6 to 7.5 times that of the other socioeconomic factors, because housing price reflects the economic level of a region. If the housing price near metro stations is low, the metro station is usually far from the city center, or the transportation network is not convenient. Many residents need to travel medium-long distances. Consequently, the metro station has a higher transfer ratio.

The intensity of business activities related variables significantly influences the transfer ratio. The average meal price of restaurants and takeaway negatively impact the transfer ratio, whereas the crowd density and distance from the CBD show positive effects. Among the intensity of business activities related variables, the average meal price of restaurants has the largest impact on the transfer ratio, which is approximately 1.875 to 15 times that of the other intensity of business activities related factors. Moreover, among all the independent variables, the average meal price of restaurants has the most significant impact on the transfer ratio, approximately 1.5 to 31 times that of the other independent variables. Usually, a metro station that meets one of the following conditions will attract many transfer passengers and have a large transfer ratio. The metro station is far away from the CBD. Alternatively, the average meal price of restaurants near the metro station is low or the monthly sales of takeaways near the metro station are low. The outcome shows that the intensity of business activities has a nonnegligible impact on the transfer ratio, which should be considered in the subsequent planning of new metro stations.

The built environment variables are found to impact the transfer ratio significantly. Bus routes 500–1000 m positively impact the transfer ratio, whereas the other variables show the opposite effect. It is observed that the feeder distance and feeder bus routes within 500 to 1000 m of a metro station have a large impact on the transfer ratio, which is approximately 1.67 to 6 times that of the other built environment variables. More importantly, when the feeder distance between bus stops and the metro station is closer or when there are more feeder buses within 500 to 1000 m of the metro station, the transfer ratio is greater at the metro station.

The transfer-related variables significantly influence the transfer ratio. The transfer time negatively impacts the transfer ratio, whereas the revised transfer time shows the opposite effect. Moreover, the transfer time has a larger impact on the transfer ratio and is approximately 6.66 times the revised transfer time. For the metro-to-bus mode, shortening the transfer time at metro stations could effectively increase the transfer ridership. For the bus-to-metro mode, properly extending the revised transfer time will increase the number of transfer passengers. Because the revised transfer time includes the in-vehicle travel time of the bus, the transfer behavior occurs in the second half of the revised transfer time.

The date-related factors are found to influence the transfer ratio significantly. For the metro-to-bus mode, the morning peak hours positively affect the transfer ratio, whereas the evening peak and non-workdays show the opposite impact. For the bus-to-metro mode, the morning peak and non-workdays have a positive effect on the transfer ratio. Furthermore, the morning peak has an impact on the transfer ratio, which is approximately 2 to 12.5 times that of the other two date factors, because commuters are most concentrated during the morning peak hours. In contrast, during evenings and non-workdays, travelers have temporal flexibility. Consequently, the transfer ratio at metro stations is larger during morning peak hours and relatively smaller during other hours.

5.2. Results of the GWR models

Although we have determined which factors significantly impact the transfer ratio through the MLR model, MLR does not consider the different spatial effects of the factors associated with the transfer ratio. Therefore, GWR is further used to explore the different effects on the transfer ratio in space. Before the GWR analysis, we have examined the spatial autocorrelation correlation of the independent variables using the global Moran's I test. Since the weather and date variables did not show variability in the spatial distribution in the same city, the Moran's I indices of the weather and date variables are equal to 0, which is not listed in Table D1. The Moran's I test results of the other independent variables are shown in Table D1 of Appendix D. The Moran's I indices are not equal to 0. The p-values are <0.01. The above results indicate that the selected independent variables have significant spatial autocorrelation. Thus, it is essential to analyze the effects of these independent variables on the transfer ratio in spaces.

This paper used GWR to model the spatially varying relationships between the independent variables and the transfer ratio, as GWR analysis has performed very well in other related studies. Table D2 reports the estimated coefficients and their descriptive statistics for the two GWR models in Appendix D.

The indicators in the GWR models are compared with those in the MLR models in Table 5. The adjusted R^2 in the GWR models is 1.1 to 1.41 times higher than that in the MLR model. The AICc values in the GWR models are 1.0 times lower than those in the MLR models, which indicates that the GWR models outperform the traditional MLR models in the study, which is consistent with previous studies (Cardozo et al., 2012). The GWR results reveal that the actual effects of the factors associated with the transfer ratio are susceptible to the underlying spatial background. Details are discussed in the following section.

Table 5
Performance comparison between the MLR and GWR models.

Indicator	The metro-to-bus mode		The bus-to-metro mode	
	MLR	GWR	MLR	GWR
AIC	−32,250.76	−65,236.94	−27,591.91	−57,340.02
AICc	−32,250.71	−64,976.14	−27,591.86	−57,073.61
R^2	0.4042	0.8521	0.3409	0.8230
AdjustedR^2	0.4037	0.8502	0.3403	0.8205
Residual sum of squares	359.33	89.17	396.70	106.53

5.3. Discussion on spatially varying effects

In the GWR models, the estimated coefficients of the independent variables vary among metro stations. The studied metro stations are colored in Fig. 5 to Fig. 8 based on their estimated coefficients to better understand the spatially varying effects of the independent variables. These variables have significant effects on the transfer ratio. The variances in the influences of some important independent variables and the potential underlying reasons are discussed below.

Fig. 5 shows the spatially varying effects of the socioeconomic variables on the transfer ratio. Fig. 5(a1) and Fig. 5(a2) show the spatial variation in house rent on the transfer ratio. House rent is significantly associated with the transfer ratio in the two transfer modes. For the metro-to-bus mode, the effect of house rent on the transfer ratio is large in the LGD and BAD, with absolute coefficients ranging from 0.60 to 2.34. For the bus-to-metro mode, the negative effect of house rent on the transfer ratio is large in the LGD, with estimated coefficients ranging from −2.49 to −0.67. The outcome suggests that most metro stations with high house rent attract many transfer passengers in the central and southern areas, whereas some metro stations in the northern area show the opposite results. Because for most metro stations with high house rent, the public transportation network is usually well developed, which can gather many passengers and has high transfer ratio.

Fig. 5(b1) and b2) show the influence of the spatially varying housing price on the transfer ratio. The housing price is negatively correlated with the transfer ratio at most metro stations, while geographical GDP and house rent show the opposite effects. This may be ascribed to the fact that the housing price usually reflects the economic level of a region and the income of residents. Areas with higher housing prices are usually more economically developed and have a dense metro network with high accessibility. Most people can arrive at their destinations through the metro without transfer trips. Therefore, the transfer ratio in the area is smaller. Moreover, among the socioeconomic variables, the impact of housing price on the transfer ratio is the largest, with its maximum coefficient being approximately 3 to 12 times that of the other socioeconomic variables. The metro stations with large estimated coefficients of housing price that range from 5.11 to 14.29 are clustered in the city center. The results reveal that the transfer ratio in downtown areas is more susceptible to housing price fluctuations than the transfer ratio in suburban areas. The difference may be ascribed to the areas in the suburbs under a development stage with less dense commercial services. The public transportation network is less dense with low accessibility. Residents in the suburbs rely mainly on private cars to commute to various places and reduce the likeliness of using public transportation for long trips. Thus, the effect of housing price on the transfer ratio is greater in the city center than in the suburbs.

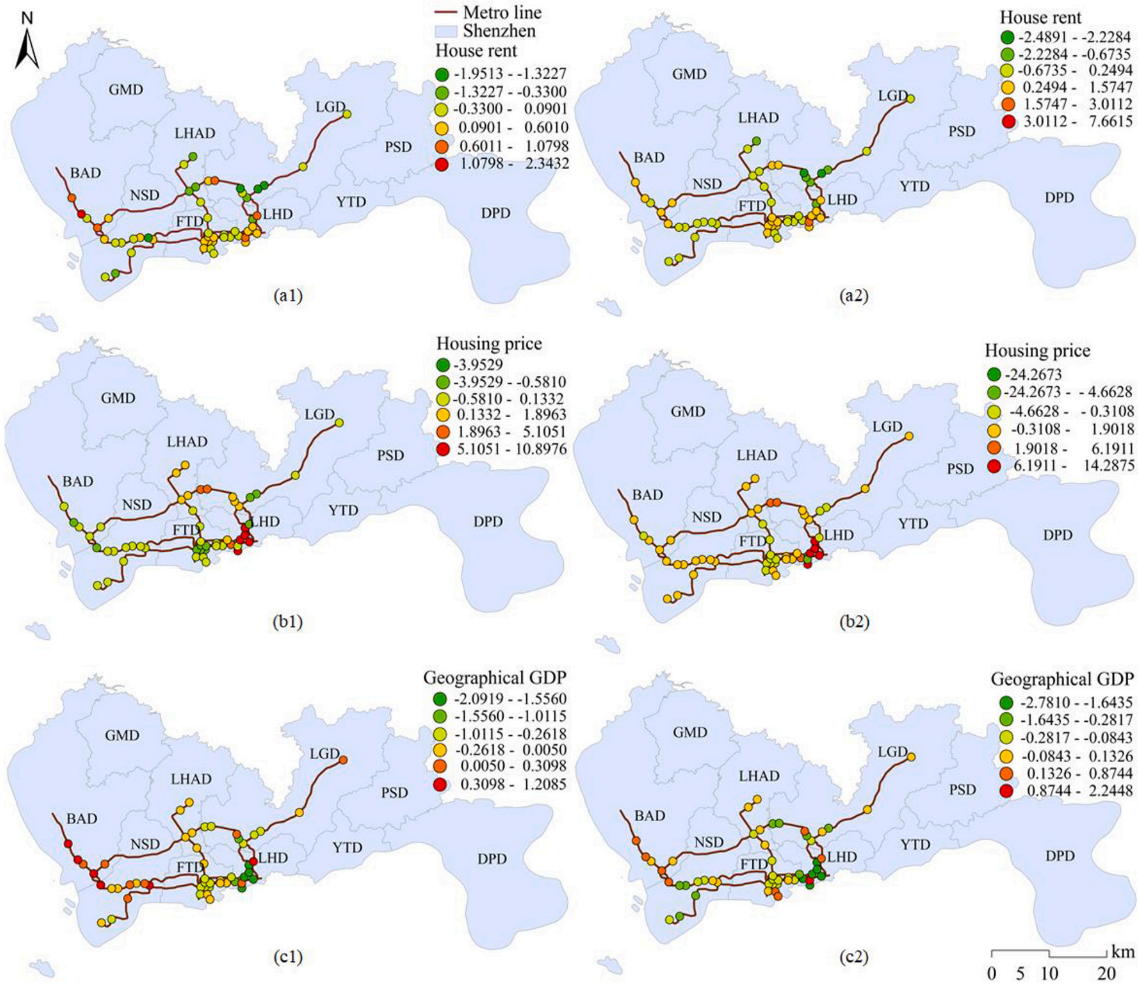

Fig. 5. Spatial distribution of the estimated coefficients of socioeconomic variables. (a1), (b1) and (c1) indicate the metro-to-bus mode; (a2), (b2) and (c2) indicate the bus-to-metro mode.

Fig. 5(c1) and (c2) show the spatially varying effect of geographical GDP on the transfer ratio. Geographical GDP is positively correlated with the transfer ratio at most metro stations because metro stations with a high GDP can attract more passengers far away from the stations, which ultimately increases the transfer ratio. For the two modes, the positive effect of geographical GDP is more significant in the BAD, with estimated coefficients ranging from 0.30 to 2.24. Moreover, the negative effect of geographical GDP is more financially beneficial in the LHD than in the other districts, with estimated coefficients ranging from −2.78 to −1.01. The results reveal that increasing the economic level in remote areas can increase the transfer ratio. Although the metro stations in the city center have different results, when the housing price near stations is higher, the transfer ratio is lower. A possible explanation is that downtown areas with high housing prices usually have convenient transportation and prosperous businesses, and people have less travel and transfer using public transportation.

Fig. 6 shows the spatially varying effects of the intensity of business activities related variables on the transfer ratio. Among them, restaurant, crowd density, and distance from the CBD are positively correlated with the transfer ratio at most metro stations. In contrast, the average meal price of restaurants and takeaways show the opposite effects. Moreover, the impact of the average meal price of restaurants on the transfer ratio is the largest in all independent variables.

Fig. 6(a1) and (a2) show the spatially varying effects of restaurants on the transfer ratio. Restaurants and the transfer ratio are positively correlated for most stations, with the estimated coefficients ranging from 0.05 to 1.57. They are negatively correlated at a few stations, with estimated coefficients ranging from −1.45 to −0.28. The restaurant has a more significant impact on the transfer ratio in the western and central regions, with the absolute values of the estimated coefficients ranging from 0.98 to 1.45. The results show that an increase in the number of restaurants can increase the transfer ratio for most stations, which may be because an increase in the number of restaurants attracts more population flow near the metro stations. A high flow of customer traffic is paramount for restaurants. Most restaurants will be concentrated in areas with high traffic flow. Usually, such areas have a well-developed transportation network with convenient transfer facilities, which attracts more transfer passengers and thus increases the transfer ratio.

Fig. 6(b1) and (b2) show the spatially varying effects of the average meal price of restaurants on the transfer ratio. The estimated coefficients have similar spatial distributions for the two modes. For most stations, the average meal price of restaurants and the transfer ratio are negatively correlated, with estimated coefficients ranging from −5.16 to −0.19. They are positively correlated at some metro stations in NSD and LGD, with estimated coefficients ranging from 0.14 to 3.27. The average meal price of restaurants has a more significant impact on the transfer ratio in the western and central regions, with the absolute values of the estimated coefficients ranging from 0.98 to 5.16. These results show that most stations with low average meal prices in restaurants can attract more transfer passengers because restaurants with lower prices usually attract more diners. The area where these restaurants are located can usually gather a large number of people with low-medium purchasing

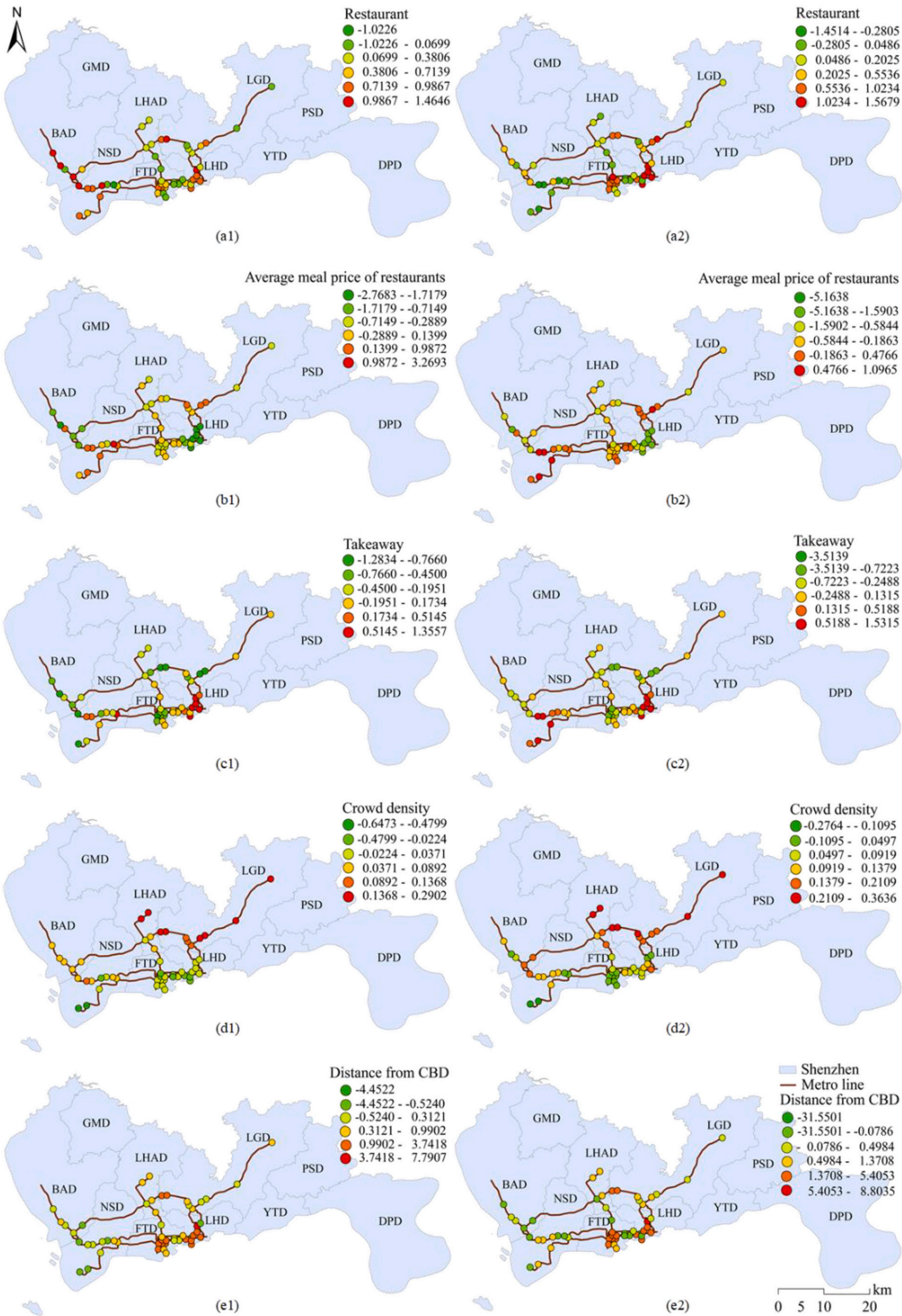

Fig. 6. Spatial distribution of the estimated coefficients of the intensity of business activities related variables. (a1), (b1), (c1), (d1) and (e1) indicate the metro-to-bus mode; (a2), (b2), (c2), (d2) and (e2) indicate the bus-to-metro mode.

power. Moreover, these people usually rely on public transportation for their daily travel. Thus, the metro station in the area has a larger transfer ratio. These may be the potential explanations for the larger effects of the average meal price of restaurants on the transfer ratio at these stations.

Fig. 6(c1) and (c2) show the spatially varying effects of takeaway on the transfer ratio. The estimated coefficients of takeaway have similar spatial distributions for the two modes. For most stations, takeaway and transfer ratio are negatively correlated, with estimated coefficients ranging from −3.51 to −0.20. In contrast, they are positively correlated at a few stations in the NSD and LHD, with estimated coefficients ranging from 0.13 to 1.53. The takeaway has a large impact on the transfer ratio in the western and central regions, with the absolute values of the estimated coefficients ranging from 0.52 to 3.51. The outcome shows that most stations with high monthly takeaway sales have a small transfer ratio since passengers have less travel demand. This may be ascribed to the fact that the areas with high monthly takeaway sales may be office areas, residential areas, or schools, where passengers have less travel demand and a smaller range of activities. The metro stations in areas with high takeaway sales attract relatively small passenger flows that use public transportation, which results in a smaller transfer ratio.

Fig. 6(d1) and (d2) show the spatially varying effects of the crowd density on the transfer ratio. The estimated coefficients have similar spatial distributions in the two modes. For most stations, crowd density positively impacts the transfer ratio, with estimated coefficients ranging from 0.04 to 0.36. However, it negatively impacts the transfer ratio at a few stations in the NSD and FTD, with estimated coefficients ranging from −0.65 to −0.02. Crowd density significantly impacts the transfer ratio in the northern and eastern regions, with the absolute values of the

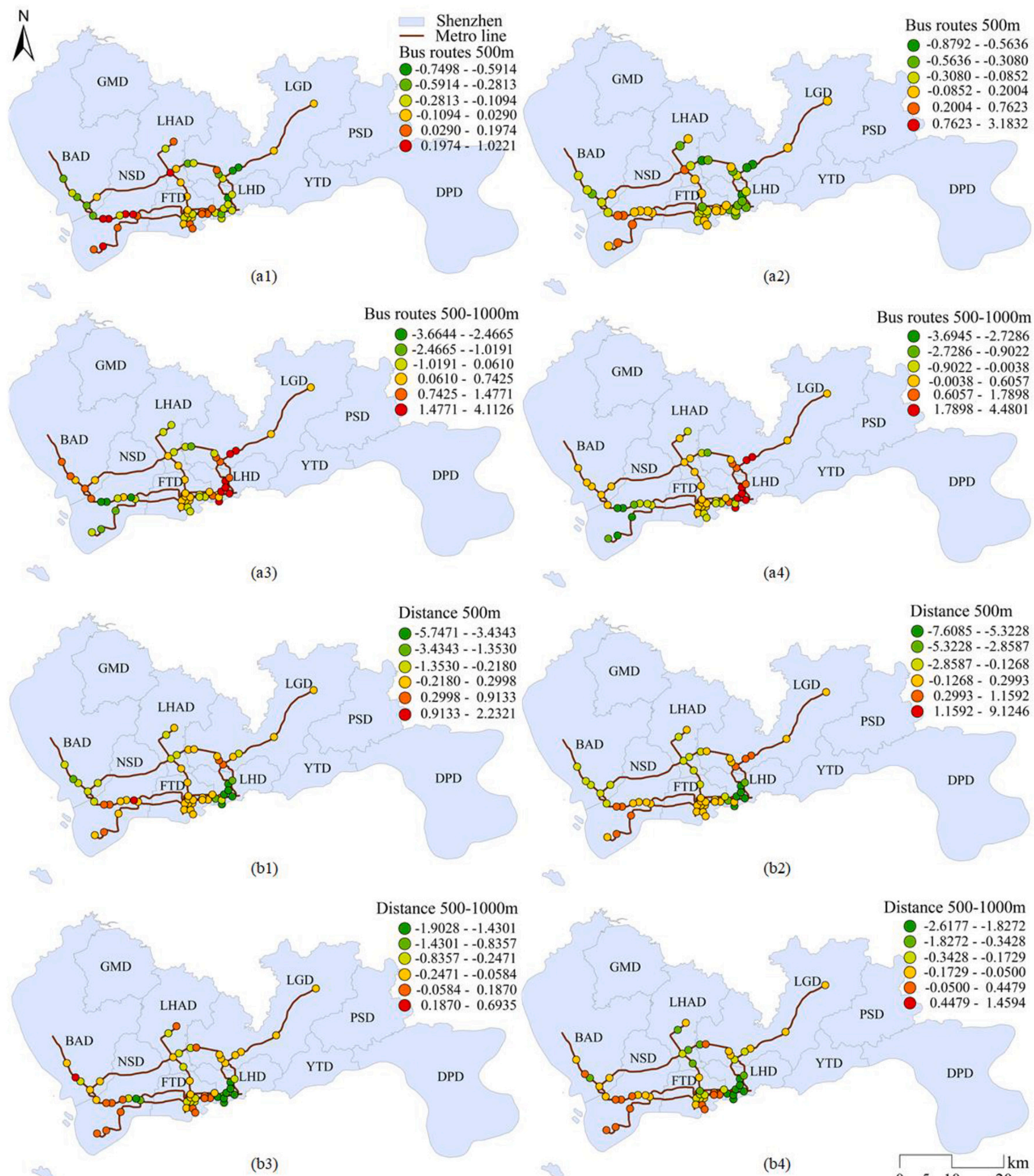

Fig. 7. Spatial distribution of estimated coefficients of built environment factors. (a1), (a3), (b1) and (b3) indicate the metro-to-bus mode; (a2), (a4), (b2) and (b4) indicate the bus-to-metro mode.

estimated coefficients ranging from 0.10 to 0.64. The results show that the transfer ratio is large for some stations with higher crowd density in the LHAD and LGD, and the opposite results are found at some stations in the NSD and FTD. The difference may be ascribed to the differential distribution of economic development levels. In the LHAD and LGD, the level of economic development has not been as good as that in the NSD and FTD in previous years. In the LHAD and LGD, the greater population density is usually concentrated in the less economically developed areas, where the transportation network is not sufficiently developed, a large number of passengers need to transfer to reach their destinations, and resulting in increased the transfer ratio. In contrast, the NSD and FTD have developed economic levels and high accessibility to the metro network. This is also true in areas with higher crowd density, where people can usually reach their destinations directly through the metro with fewer transfer trips, and thus there is a smaller transfer ratio.

Fig. 6(e1) and (e2) show the spatially varying effects of distance from the CBD on the transfer ratio. For most stations, the distance from the CBD positively impacts the transfer ratio, with estimated coefficients ranging from 0.31 to 8.80. However, it negatively impacts the transfer ratio at a few stations in the NSD and BAD, with estimated coefficients ranging from −31.55 to −0.08. Moreover, the estimated coefficients have similar spatial distributions in the two modes. The results reveal that transfer passengers living far from the CBD are more sensitive to fluctuations in the transfer ratio. For most metro stations, when the distance from the CBD is farther away, the transfer ratio is greater. This is because areas close to the CBD are usually economically developed, with a large population flow and high accessibility to the metro network. Therefore, people can usually take the metro to their destination, which results in less transfer passenger flow and a small transfer ratio. In contrast, areas farther away from the CBD usually have a lower economic level, and people usually cannot reach their destinations directly through the metro and need one or two transfer trips to reach their destinations. Therefore, there are more transfer passengers and a large transfer ratio.

Fig. 7 shows the spatial distribution of the coefficients of the built environment variables on the transfer ratio. Based on the variable coefficients, we find similar spatial distributions in the two modes.

Fig. 7(a1) to (a4) show the spatially varying effects of bus routes on the transfer ratio. Fig. 7(a1) and (a2) show the influence of bus routes 500 m on the transfer ratio. The figure indicates that bus routes 500 m are negatively associated with the transfer ratio at most metro stations, with the estimated coefficients ranging from −0.88 to −0.28. Because the metro station with many feeder bus routes within 500 m are usually located in commercial centers or public transportation hubs, where there is a well-developed transportation network, passengers can often arrive at their destinations directly through the metro or bus without transfer trips. Therefore, a metro station with more bus routes within 500 m has a smaller transfer ratio. Moreover, the impact of bus routes within 500 m on the transfer ratio is smaller for the metro-to-bus mode than for the bus-to-metro mode. This is attributed to the different ride order between the two transfer modes. The transfer passengers of the bus-to-metro mode are mainly from bus ridership; therefore, when there are more bus routes and more bus ridership, this results in a greater transfer ratio of the bus-to-metro mode. Because the transfer passengers of the metro-to-bus mode are mainly from metro ridership, the transfer ratio is less affected by fluctuations in bus routes 500 m. Additionally, bus routes 500 m are positively associated with the transfer ratio at a few stations in the NSD and FTD, with estimated coefficients ranging from 0.03 to 3.18. Compared to Fig. 7(a1) and (a2), (a3) and (a4) show the opposite results. The impact of bus routes 500–1000 m on the transfer ratio is overall much higher than that of bus routes 500 m, with the maximum coefficient of bus routes 500–1000 m being approximately 1.41 to 4.03 times that of bus routes 500 m. In addition, bus routes 500–1000 m are positively associated with the transfer ratio at most metro stations in the NSD, FTD, and LHAD, with estimated coefficients ranging from 0.06 to 4.48. Because transfer passengers within 500–1000 m from the metro station have a stronger transfer demand than those within 500 m from the metro station, these passengers have also a longer transfer distance to the metro stations. Moreover, the public transportation network is less dense, which is not convenient for people to travel by using the metro or bus. Passengers often need to transfer to reach their destinations. Therefore, an increase in the number of bus routes within 500–1000 m can increase the attractiveness of public transportation and thus increase the transfer ratio.

Fig. 7(b1) to (b4) show the spatially varying effects of feeder distance. The spatial distribution of the coefficients is very similar in the four figures. The distance is negatively associated with the transfer ratio at most stations, with estimated coefficients ranging from −7.61 to −0.05. This may be ascribed to the longer transfer distance, whereby the passengers' willingness to transfer is lower, thus the smaller the transfer ratio is. Furthermore, the impact of distance 500 m on the transfer ratio is much higher than that of distance 500–1000 m, with the maximum coefficient of distance 500 m being two times higher than that of distance 500–1000 m. The results show that when the feeder distance that connects metro stations and bus stops is longer, the transfer ratio at most stations is lower. Because most passengers transfer within 500 m from the metro station, especially in downtown and commercial areas, there are few transfer trips beyond 500 m from the metro station. Therefore, a metro station with a short transfer distance has a large transfer ratio.

Fig. 8(a1) and (a2) show the influence of the spatial variation in the transfer time and revised transfer time on the transfer ratio. For the metro-to-bus mode, the effect of the transfer time on the transfer ratio is negative for most metro stations in remote areas, with estimated coefficients ranging from −0.31 to −0.01. The positive effect of the transfer time on the transfer ratio occurs at a few stations in the center city, with estimated coefficients ranging from 0.01 to 0.04. Because the well-developed transportation networks are located in the city center with a high economic level and large passenger flow. A metro station with a longer transfer time has a greater transfer ratio. Meanwhile, most metro stations in remote areas have long transfer times, transportation networks that are not developed, longer bus headway, and fewer passengers who choose to transfer. Many passengers choose to travel by electric car or bicycle instead of the bus; thus, there is a lower transfer ratio. For the bus-to-metro mode, the positive coefficients of the revised transfer time are mainly distributed in LGD, FTD, and BAD and range from 0.04 to 0.29. The negative coefficients are mainly distributed in the LHAD, LHD, and NSD, ranging from −0.18 to −0.003. The outcome shows that the transfer time and the revised transfer time positively affect the transfer ratio at most stations in the central and western regions. This is because, unlike the transfer time of the metro-to-bus mode, the transfer time of the bus-to-metro mode includes the passenger's ride time on the first bus and the walking time from the bus stop to the metro station. Therefore, a metro station with a longer transfer time has a greater transfer ratio, which indicates that a metro station with a longer transfer time threshold in the downtown area can gather more transfer passengers.

Accordingly, the GWR models further reveal that the spatial pattern of the effects of given factors on the transfer ratio differs from a global view, which results in a complex spatial variation in the transfer ratio. The results indicate the spatial heterogeneity of the influence of the associated factors on the transfer ratio in the two modes. Furthermore, these factors significantly impact the transfer ratio because of the spatial heterogeneity of Shenzhen, China's spatial and socioeconomic structure, cooperation, and multimodal public transportation competition. In particular, the built environment variables and the intensity of business activities significantly affect the transfer ratio in the two modes. More importantly, this study provides a comprehensive understanding of the spatial variation in the transfer ratio that can help allocate available transportation resources and thus encourage a massive number of travelers to use the public transportation system. In addition, incorporating significant independent variables with the characteristics of the urban area can produce more robust quantifications of the transfer ratio

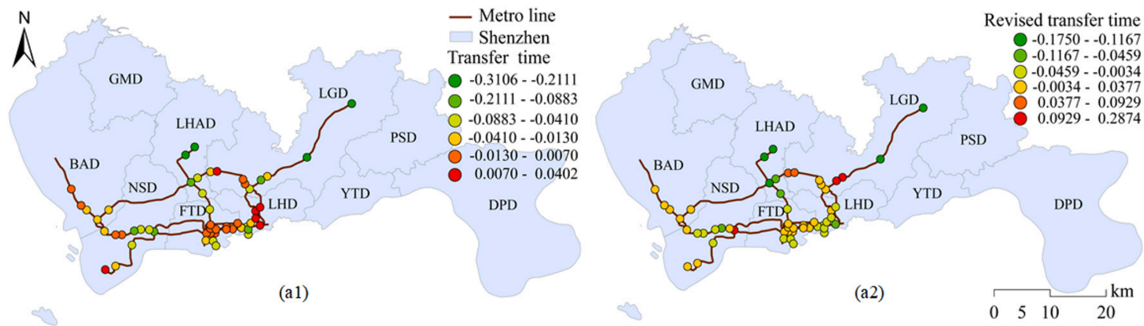

Fig. 8. Spatial distribution of the estimated coefficients of transfer-related variables: (a1) the metro-to-bus mode and (a2) the bus-to-metro mode.

to thus help regulate the supply of buses and metro stations.

6. Conclusion and implications

Understanding the transfer ratio of different travel modes is beneficial for pursuing a low-carbon comprehensive urban transportation system. However, limited efforts have been made to investigate the spatial variations in the transfer ratio between metro and bus systems and identify their influencing factors. This study endeavors to fill the gaps in relevant research by deciphering the spatial heterogeneity of the transfer ratio and their relationships with the various variables based on empirical analysis in Shenzhen, China. First, the transfer ratio and transfer-related factors are identified based on large-scale transaction data from automated fare collection systems. Meanwhile, various influencing factors, including weather, socioeconomic, the intensity of business activities, and built environment factors, are obtained from multivariate sources. Second, the influences of socioeconomic characteristics, weather conditions, the intensity of business activities, built environment attributes, date factors, and transfer-related factors on the transfer ratio are explored through MLR model. Third, an adaptative GWR is further performed to reveal the divergences in the effects of these factors on the transfer ratio and provide more accurate modeling concerning the transfer ratio in various contexts. The main contributions and findings can be summarized as follows:

- We extract the hourly transfer ratio distributions of different transfer modes from large-scale smartcard data and GPS data. We find notable spatial heterogeneity in the transfer ratio in different contexts. The transfer ratio is discussed from statistical and spatial perspectives.
- By leveraging multisource data, we obtain various weather data, transfer-related factors, socioeconomic factors, the intensity of business activities, and built environment factors in different urban areas and metro stations. MLR is utilized to model the associations of these factors with the transfer ratio for a global view. The results indicate that various factors indeed have significant effects on the transfer ratio. The intensity of business activities has the largest impact on the transfer ratio. The crowd density and distance from the CBD are positively related to the transfer ratio, which indicates that metro stations with higher degrees of these factors have a large transfer ratio. In contrast, the average meal price of restaurants and takeaway are found to be negatively associated with the transfer ratio. Socioeconomic factors including geographical GDP and house rent show a positive effect on the transfer ratio, whereas housing price negatively influences the transfer ratio, which suggests that metro stations with higher geographical GDP, house rent, and a lower housing price have a large transfer ratio. Built environment variables including bus routes within 500 m, distance 500 m, and distance 500–1000 m negatively impact the transfer ratio, whereas bus routes within 500–1000 m show the opposite effect. This indicates that metro stations with higher degrees of bus routes within 500 m, distance 500 m, and distance 500–1000 m have a smaller transfer ratio. Weather factors including temperature, wind, and rainfall positively impact the transfer ratio, whereas visibility shows the opposite effect, which implies that metro stations with a higher temperature, wind speed, and rainfall have a greater transfer ratio. Importantly, we find that the effects of some factors (e.g., crowd density) on the transfer ratio vary with different transfer modes.
- Last, GWR is employed to further investigate the spatial variations in the influences of various factors on the transfer ratio. The results demonstrate that notable variations exist in the influences of the built environment factors, transfer-related factors, intensity of business activities, and socioeconomic factors, whereas some factors such as the housing price, house rent, and geographical GDP present opposite effects on the transfer ratio in different spatial contexts. The results imply the necessity of considering the divergent effect of these factors in different geographical contexts. Comprehensive discussions are conducted to provide empirical explanations concerning the varying effects of transfer-related, socioeconomic, the intensity of business activities, and built environment factors on the transfer ratio. The GWR offers more accurate modeling of the transfer ratio and various associated variables from a local spatial perspective.

The above findings provide practical implications for the improvement, management, and planning of metro stations and buses. The spatial heterogeneity in the transfer ratio suggests that it is implausible to apply a spatially homogeneous transfer ratio to predict the intensity of spatial interactions between the metro and buses. In forecasting the spatial transfer ratio, it is necessary to develop customized models based on the specific characteristics of different urban spaces to obtain accurate results. Moreover, the results in MLR and GWR decipher the variations in the transfer ratio from a global perspective and a local geographical perspective. The obtained quantitative results such as the effects of different built environment factors, socioeconomic factors, and intensity of business activities could provide scientific support and referential values for the prediction of the transfer ratio, especially for areas where the metro station does not yet exist, whereby it is in its early stage or needs improvement. For instance, the MLR and GWR can be utilized to predict the transfer ratio of different areas based on their known built environment, socioeconomic factors, the intensity of business activities, and other statistics. This would serve accurate transfer ratio forecasting and thus supply allocations of buses and metro, which would be useful for achieving supply-demand balances of public transportation in different urban areas. These are also great practical considerations for bus companies to allocate and schedule buses to fulfill the transfer demand and thus realize a high coupling performance between buses and the metro system. This is particularly crucial for bus companies that have limited buses and need to improve utilization rates of a bus to reduce budgets. Moreover, insights into the effects of the intensity of business activities, socioeconomic factors, and built environment factors, such as the transfer distance and bus routes, offer empirical and quantitative instruments for evaluating the accessibility of transfer trips in different urban contexts. These explanations produce implications and directions for the planning of bus routes and bus stops that are aimed to enhance transfer ratio and promote establishment of an environmentally friendly and sustainable public transportation system.

However, there are still some limitations that need to be further explored. First, we only select Shenzhen as the study area. Although this study represents China, these variables may have different effects on the transfer ratio in other cities of the world. Subsequent similar studies will be conducted in other cities. Second, the sample size is limited to one month, which fails to consider seasonality for longitudinal analyses. Therefore, in future research, we will obtain a dataset with a longer available data period to further determine the relationship between the transfer ratio and seasonal weather conditions. Third, we discuss only the linear relationship between the independent variables and the transfer ratio, which may have a nonlinear relationship. Moreover, the interaction between various factors is not considered. Therefore, follow-up studies are suggested to explore the linear and nonlinear relationship between the independent variable and the transfer ratio, investigate the interaction between different factors simultaneously, and provide a deeper understanding of the impact of various factors on the transfer ratio. Fourth, the individual characteristics of transfer passengers, such as gender, age, income, etc., are not considered in the study. In a follow-up study, we will further investigate the effects of these factors on the transfer ratio. Last, as weather is timely, and hence, a spatiotemporal model will be used in a future study to explore the effects of weather factors on the transfer ratio.

Author statement

The authors confirm the contribution to the paper as follows: study conception and design: Lunhui Xu and Pan Wu; data collection: Pan Wu and Mingyang Pei; analysis and interpretation of results: Pan Wu, Mingyang Pei, Kun Gao; draft manuscript preparation: Lunhui Xu, Pan Wu, Mingyang Pei, Kun Gao, and Xiaobo Qu. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Competing Interest

The authors state that there are no competing interests associated with the publication of this study.

Acknowledgement

This study was funded by National Natural Science Foundation of China grant 11702099.

Appendix A

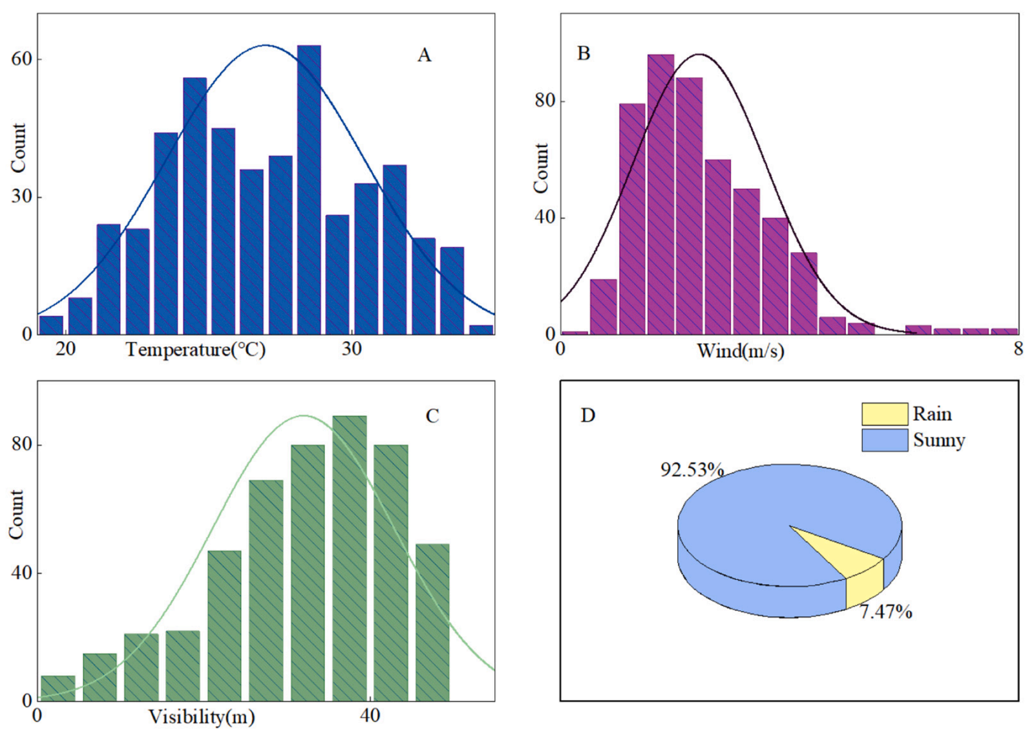

Fig. A.1. Distribution of the weather variables: (A) temperature, (B) wind, (C) visibility, and (D) rainfall (dummy variable).

Appendix B

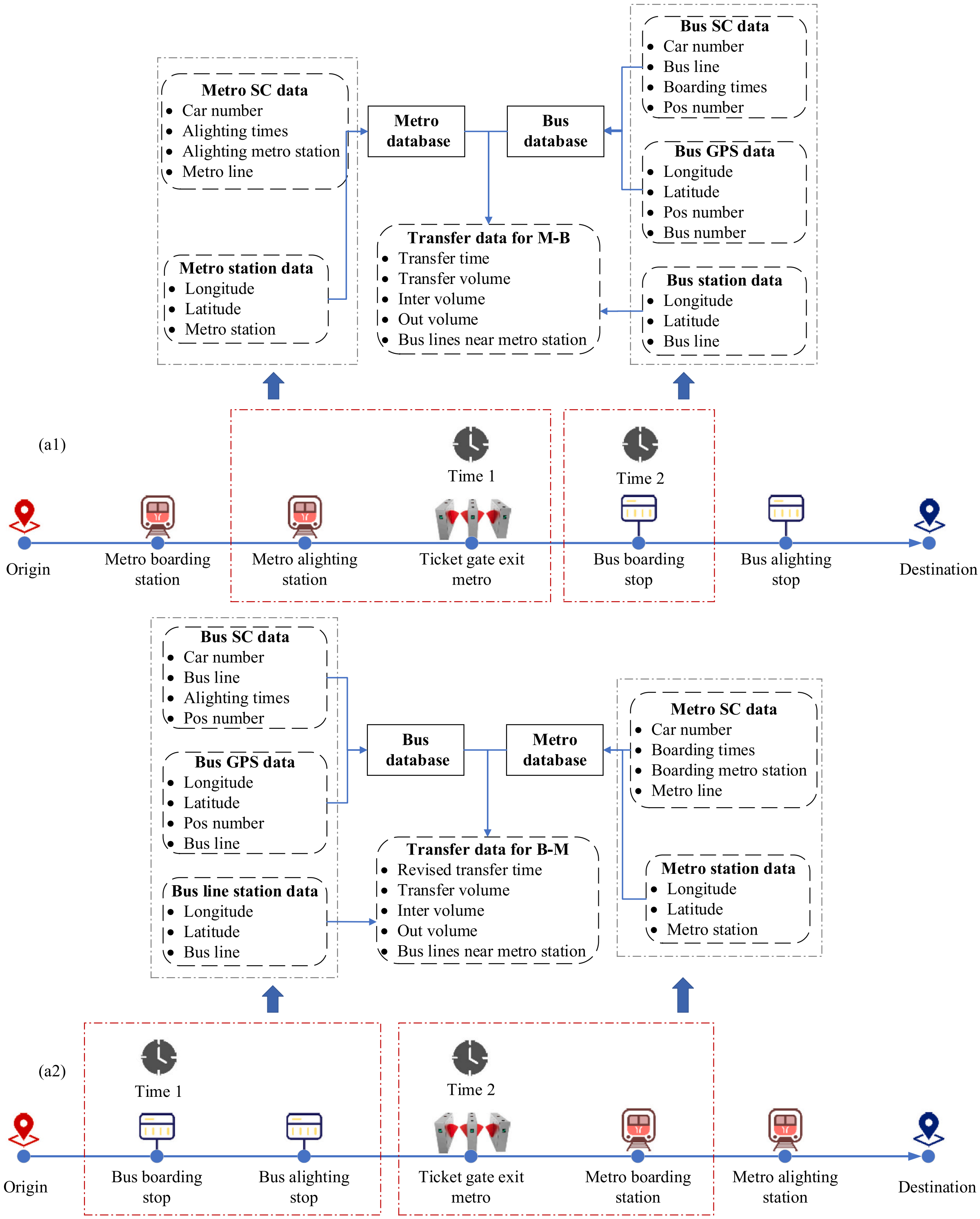

Fig. B.1. Schematic diagram of the transfer process between the metro and the bus systems: (a1) the metro-to-bus mode and (a2) the bus-to-metro mode.

Appendix C

See Fig. C.1, Fig. C.2, and Fig. C.3.

(1) Distribution of travel variables in the metro station

Taking Wuhe station (a typical metro station in Shenzhen) as an example, we analyze the real-time characteristics of the distribution of the travel variables on different dates.

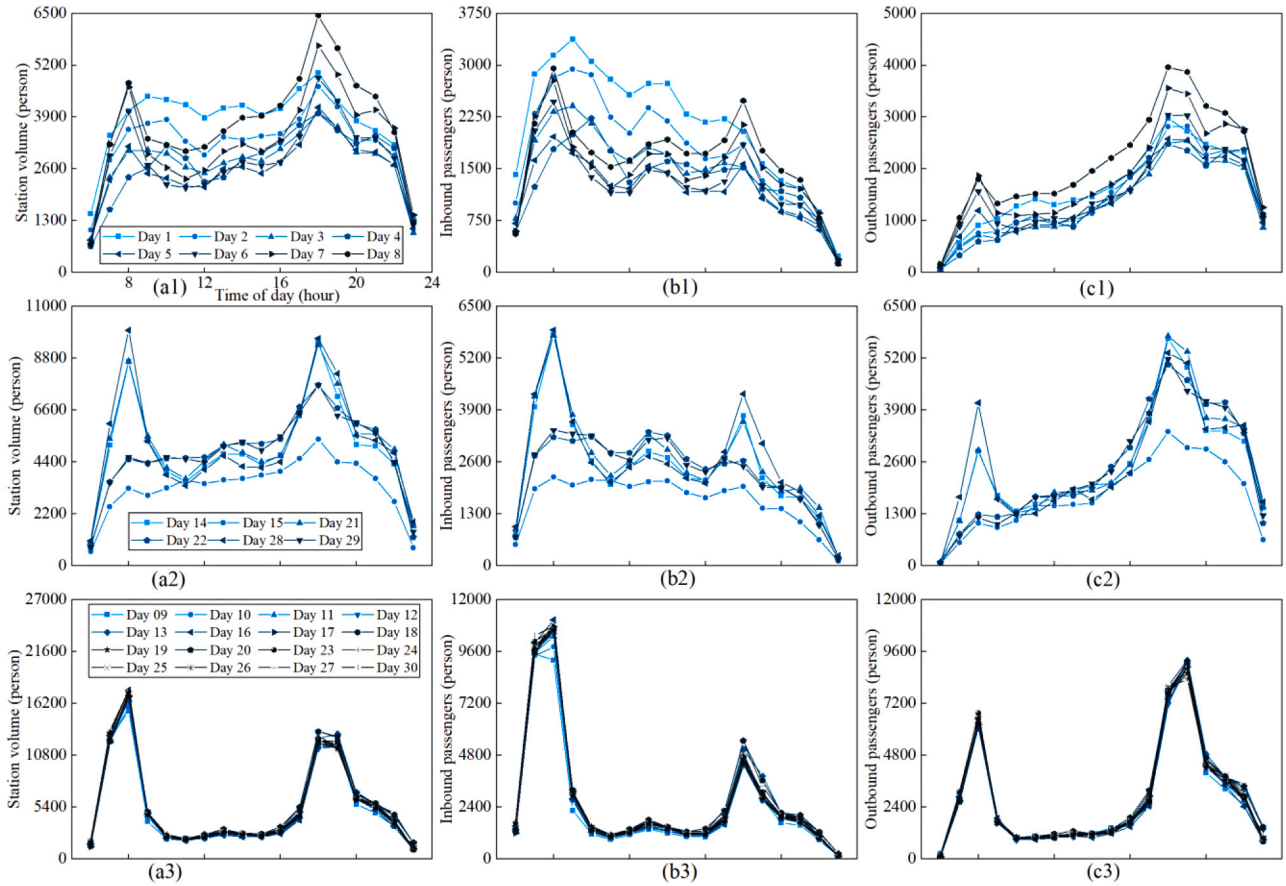

Fig. C.1. Distribution of metro travel data during holidays, weekends, and weekdays at Wuhe station: (a1), (a2), and (a3) the station volume; (b1), (b2), and (b3) the number of inbound passengers; and (c1), (c2), and (c3) the number of outbound passengers.

Fig. C.1 shows that the distribution of the station volume, the number of inbound, and the number of outbound passengers at Wuhe station differ enormously on weekdays, weekends, and holidays. Additionally, the characteristics of the daily distribution of the three variables on weekends and holidays are quite different. The distribution of the three variables on weekdays shows obvious peaks and flats, with the peak hours including the morning peak and evening peak.

(2) Distribution of transfer variables in the metro station

Taking Wuhe station as an example, we analyze the characteristics of the real-time distribution of the transfer indicators on different dates. Fig. C.2 shows that the transfer time, transfer volume, and transfer ratio at metro stations are distributed differently on weekends, holidays, and workdays for the metro-to-bus mode. Additionally, the distribution of the transfer time, transfer volume, and transfer ratio on workdays shows an obvious morning peak and evening peak. The characteristics of the distribution on weekdays are the same and consistent.

Fig. C.3 shows that the revised transfer time, transfer volume, and transfer ratio at the metro stations are distributed differently on weekends, holidays, and workdays for the bus-to-metro mode. When analyzing the revised transfer time, transfer volume, and transfer ratio, it is found that the distribution of workdays shows an obvious morning peak and evening peak. Moreover, the characteristics of the distribution of workdays are the same with good consistency. Therefore, the date variables are divided into the four categories of off-peak workdays, morning peak, evening peak, and non-workdays in this study.

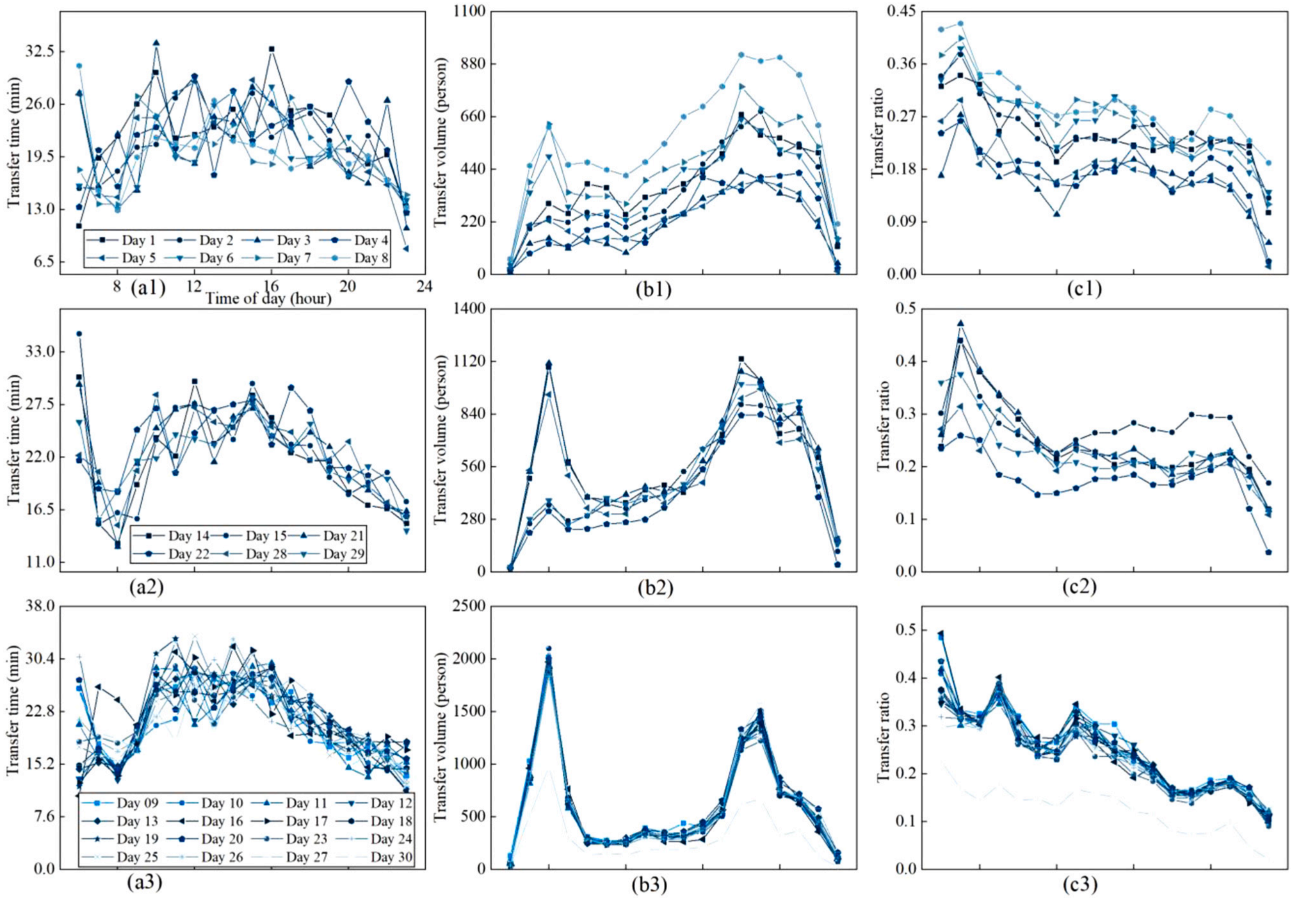

Fig. C.2. Distribution of transfer data on holidays, weekends, and weekdays for the metro-to-bus mode at Wuhe station: (a1), (a2), and (a3) the transfer time; (b1), (b2), and (b3) the transfer volume; and (c1), (c2), and (c3) the transfer ratio.

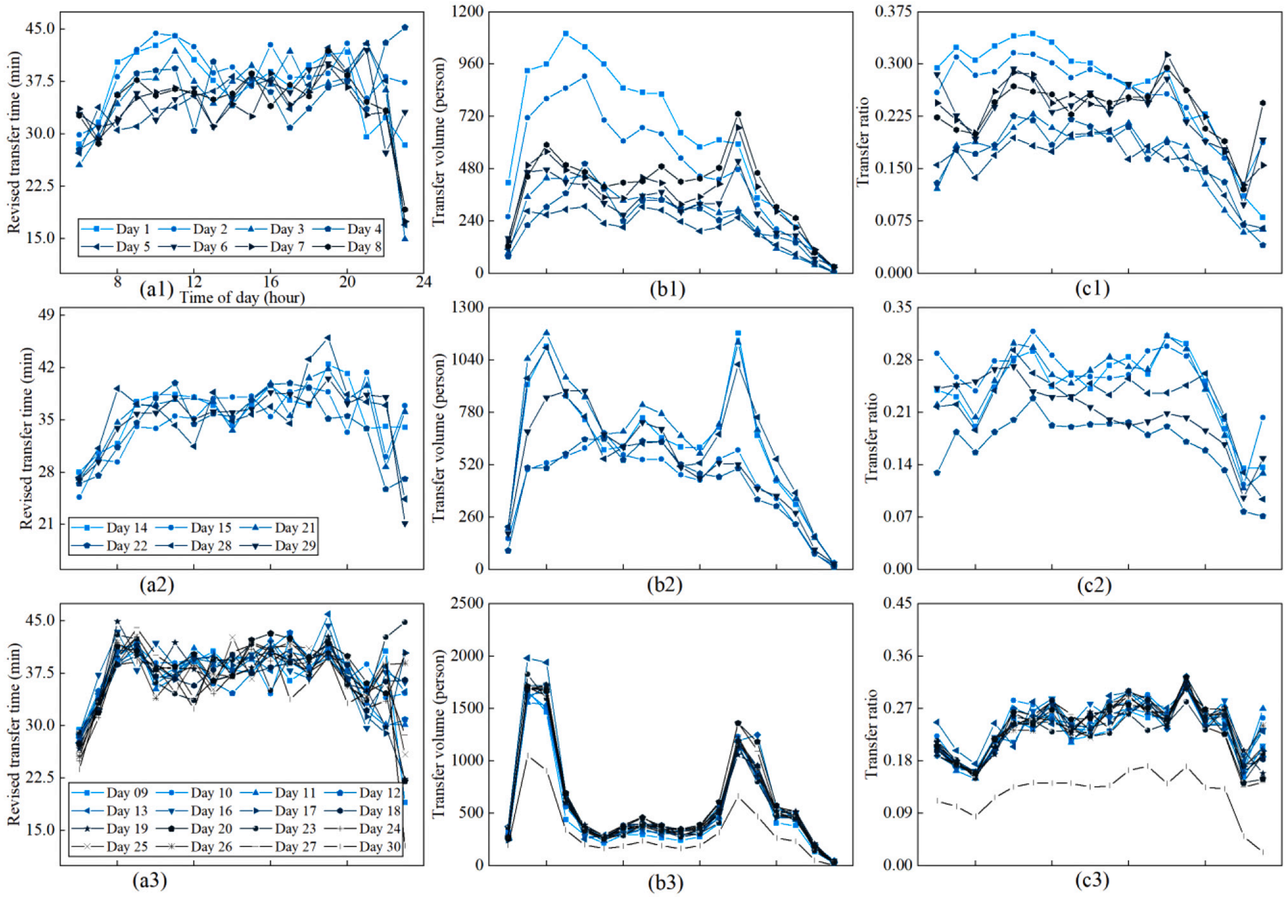

Fig. C.3. Distribution of transfer variables on holidays, weekends, and workdays for the bus-to-metro mode at Wuhe station: (a1), (a2), and (a3) the revised transfer time; (b1), (b2), and (b3) the transfer volume; and (c1), (c2), and (c3) the transfer ratio.

Appendix D

Table D1
Results of Moran's I test.

Variables	The bus-to-metro mode				The metro-to-bus mode			
	Moran's I index	Expected index	Sd.	p-value	Moran's I index	Expected index	Sd.	p-value
House rent	0.664	0	0.0002	0.001	0.666	0	0.0002	0.001
Housing price	0.612	0	0.0002	0.001	0.613	0	0.0002	0.001
Geographical GDP	0.55	0	0.0002	0.001	0.548	0	0.0002	0.001
Restaurant	0.233	0	0.0002	0.001	0.221	0	0.0002	0.001
Average meal price of restaurants	0.566	0	0.0002	0.001	0.573	0	0.0002	0.001
Takeaway	0.231	0	0.0002	0.001	0.229	0	0.0002	0.001
Crowd density	0.077	0	0.0002	0.001	0.072	0	0.0002	0.001
Distance from the CBD	0.852	0	0.0002	0.001	0.847	0	0.0002	0.001
Schools	0.321	0	0.0002	0.001	0.326	0	0.0002	0.001
Bus route 500 m	0.278	0	0.0002	0.001	0.258	0	0.0002	0.001
Bus route 500–1000 m	0.213	0	0.0002	0.001	0.182	0	0.0002	0.001
Distance 500 m	0.062	0	0.0002	0.001	0.08	0	0.0002	0.001
Distance 500–1000 m	0.1645	0	0.0002	0.001	0.169	0	0.0002	0.001
Transfer time	0.088	0	0.0002	0.001	/	/	/	/
Revised transfer time	/	/	/	/	0.143	0	0.0002	0.001

Table D2
Estimation of the GWR models.

Variables	The metro-to-bus mode					The bus-to-metro mode				
	Min.	25%	Median	75%	Max.	Min.	25%	Median	75%	Max.
(Intercept)	-1.31	0.04	0.41	0.57	2.50	-1.45	−1.01	0.10	0.6	4.03

(*continued on next page*)

Table D2 (*continued*)

Variables	The metro-to-bus mode					The bus-to-metro mode				
	Min.	25%	Median	75%	Max.	Min.	25%	Median	75%	Max.
Temperature	−0.02	−0.01	0.02	0.04	0.11	−0.04	0.01	0.04	0.07	0.14
Wind	−0.04	−0.02	−0.01	0.03	0.08	−0.03	−0.001	0.02	0.07	0.15
Visibility	−0.08	−0.03	−0.003	0.01	0.02	−0.12	−0.04	−0.02	−0.01	−0.001
Rainfall	−0.01	0.01	0.014	0.02	0.07	−0.003	0.004	0.01	0.02	0.04
House rent	−1.95	−0.19	0.02	0.41	2.34	−2.49	−0.12	0.06	0.47	7.66
Housing price	−3.95	−0.40	−0.08	0.62	10.9	−24.27	−0.54	0.08	0.55	14.29
Geographical GDP	−2.09	-0.38	−0.12	0.06	1.21	−2.78	−0.31	−0.11	0.04	2.24
Restaurant	−1.02	0.06	0.52	0.82	1.46	−1.45	0.02	0.15	0.87	1.57
Average meal price of restaurants	−2.77	−0.71	−0.35	0.09	3.27	−5.16	−0.73	−0.40	0.11	1.10
Takeaway	−1.28	−0.63	−0.21	−0.01	1.36	−3.51	−0.46	−0.10	0.43	1.53
Crowd density	−0.65	−0.01	0.02	0.09	0.29	−0.28	0.05	0.10	0.17	0.36
Distance from the CBD	−4.45	−0.03	0.48	2.23	7.79	−31.55	0.03	0.70	2.87	8.80
Schools	−2.32	−0.45	0.01	0.21	0.86	−3.11	−0.56	0.06	0.20	3.60
Bus routes 500 m	−0.75	−0.21	−0.05	0.11	1.02	−0.88	−0.33	−0.13	0.04	3.18
Bus routes 500–1000 m	−3.66	−0.12	0.23	1.10	4.11	−3.69	−0.13	0.23	0.60	4.48
Distance 500 m	−5.75	−0.39	0.06	0.18	2.23	−7.61	−0.37	0.08	0.21	9.12
Distance 500–1000 m	−1.90	−0.38	−0.20	−0.04	0.70	−2.62	−0.36	−0.14	0.01	1.46
Transfer time	−0.31	−0.06	−0.02	0.0002	0.04	/	/	/	/	/
Revised transfer time	/	/	/	/	/	−0.17	−0.03	0.001	0.02	0.29
Morning peak	−0.05	−0.02	0.01	0.05	0.15	−0.05	0.02	0.03	0.07	0.23
Evening peak	−0.11	−0.03	−0.001	0.02	0.18	−0.06	−0.03	−0.01	0.004	0.10
Non-workdays	−0.04	−0.01	−0.004	0.002	0.05	−0.03	−0.01	0.002	0.01	0.05

References

Allard, R.F., Moura, F., 2018. Effect of transport transfer quality on intercity passenger mode choice. Transp. Res. Part A Policy Pract. 109, 89–107. https://doi.org/10.1016/j.tra.2018.01.018.

Arana, P., Cabezudo, S., Peñalba, M., 2014. Influence of weather conditions on transit ridership: a statistical study using data from smartcards. Transp. Res. Part A Policy Pract. 59, 1–12. https://doi.org/10.1016/j.tra.2013.10.019.

Brunsdon, C., Fotheringham, A.S., Charlton, M., 1999. Some notes on parametric significance tests for Gwr. J. Reg. Sci. 39, 497–524.

Cardozo, O.D., García-Palomares, J.C., Gutiérrez, J., 2012. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. Appl. Geogr. 34, 548–558. https://doi.org/10.1016/j.apgeog.2012.01.005.

Cascajo, R., Lopez, E., Herrero, F., Monzon, A., 2019. User perception of transfers in multimodal urban trips: a qualitative study. Int. J. Sustain. Transp. 13, 393–406. https://doi.org/10.1080/15568318.2018.1476632.

Chakour, V., Eluru, N., 2016. Examining the influence of stop level infrastructure and built environment on bus ridership in Montreal. J. Transp. Geogr. 51, 205–217. https://doi.org/10.1016/j.jtrangeo.2016.01.007.

Chen, E., Ye, Z., Wang, C., Zhang, W., 2019. Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data. Cities 95, 102359. https://doi.org/10.1016/j.cities.2019.05.028.

Choi, J., Lee, Y.J., Kim, T., Sohn, K., 2012. An analysis of metro ridership at the station-to-station level in Seoul. Transp. (Amst). 39, 705–722. https://doi.org/10.1007/s11116-011-9368-3.

Cools, M., Moons, E., Creemers, L., Wets, G., 2010. Changes in travel behavior in response to weather conditions: do type of weather and trip purpose matter? Transp. Res. Rec. 22–28 https://doi.org/10.3141/2157-03.

Espino, R., Román, C., 2020. Valuation of transfer for bus users: the case of gran Canaria. Transp. Res. Part A Policy Pract. 137, 131–144. https://doi.org/10.1016/j.tra.2020.05.003.

Gao, K., Yang, Y., Sun, L., Qu, X., 2020. Revealing psychological inertia in mode shift behavior and its quantitative influences on commuting trips. Transp. Res. Part F Traffic Psychol. Behav. 71, 272–287. https://doi.org/10.1016/j.trf.2020.04.006.

Gao, K., Yang, Y., Li, A., Li, J., Yu, B., 2021a. Quantifying economic benefits from free-floating bike-sharing systems: a trip-level inference approach and city-scale analysis. Transp. Res. Part A Policy Pract. 144, 89–103. https://doi.org/10.1016/j.tra.2020.12.009.

Gao, K., Yang, Y., Li, A., Qu, X., 2021b. Spatial heterogeneity in distance decay of using bike sharing: an empirical large-scale analysis in Shanghai. Transp. Res. Part D Transp. Environ. 94, 102814 https://doi.org/10.1016/j.trd.2021.102814.

Garcia-Martinez, A., Cascajo, R., Jara-Diaz, S.R., Chowdhury, S., Monzon, A., 2018. Transfer penalties in multimodal public transport networks. Transp. Res. Part A Policy Pract. 114, 52–66. https://doi.org/10.1016/j.tra.2018.01.016.

Guo, Z., Wilson, N.H.M., Rahbee, A., 2007. Impact of weather on transit ridership in Chicago. Illinois. Transp. Res. Rec. 3–10 https://doi.org/10.3141/2034-01.

Hadas, Y., Ranjitkar, P., 2012. Modeling public-transit connectivity with spatial quality-of-transfer measurements. J. Transp. Geogr. 22, 137–147. https://doi.org/10.1016/j.jtrangeo.2011.12.003.

Huang, Z., Xu, L., Lin, Y., Wu, P., Feng, B., 2019. Citywide metro-to-bus transfer behavior identification based on combined data from smart cards and GPS. Appl. Sci. 9 https://doi.org/10.3390/app9173597.

Iseki, H., Taylor, B.D., 2009. Not all transfers are created equal: towards a framework relating transfer connectivity to travel behaviour. Transp. Rev. 29, 777–800. https://doi.org/10.1080/01441640902811304.

Li, L., Wang, J., Song, Z., Dong, Z., Wu, B., 2015. Analysing the impact of weather on bus ridership using smart card data. IET Intell. Transp. Syst. 9, 221–229. https://doi.org/10.1049/iet-its.2014.0062.

Li, W., Chen, S., Dong, J., Wu, J., 2021. Exploring the spatial variations of transfer distances between dockless bike-sharing systems and metros. J. Transp. Geogr. 92, 103032 https://doi.org/10.1016/j.jtrangeo.2021.103032.

Ma, X., Zhang, J., Ding, C., Wang, Y., 2018. A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership. Comput. Environ. Urban. Syst. 70, 113–124. https://doi.org/10.1016/j.compenvurbsys.2018.03.001.

Miao, Q., Welch, E.W., Sriraj, P.S., 2019. Extreme weather, public transport ridership and moderating effect of bus stop shelters. J. Transp. Geogr. 74, 125–133. https://doi.org/10.1016/j.jtrangeo.2018.11.007.

Munizaga, M.A., Palma, C., 2012. Estimation of a disaggregate multimodal public transport origin – destination matrix from passive smartcard data. Transp. Res. Part C 24, 9–18. https://doi.org/10.1016/j.trc.2012.01.007.

Munizaga, M., Devillaine, F., Navarrete, C., Silva, D., 2014. Validating travel behavior estimated from smartcard data. Transp. Res. PART C 44, 70–79. https://doi.org/10.1016/j.trc.2014.03.008.

Navarrete, F.J., de Ortúzar, J.D., 2013. Subjective valuation of the transit transfer experience: the case of Santiago de Chile. Transp. Policy 25, 138–147. https://doi.org/10.1016/j.tranpol.2012.10.006.

Schakenbos, R., La Paix, L., Nijenstein, S., Geurs, K.T., 2016. Valuation of a transfer in a multimodal public transport trip. Transp. Policy 46, 72–81. https://doi.org/10.1016/j.tranpol.2015.11.008.

Seaborn, C., Attanucci, J., Wilson, N.H.M., 2009. Analyzing multimodal public transport journeys in London with smart card fare payment data. Transp. Res. Rec. 55–62 https://doi.org/10.3141/2121-06.

Singhal, A., Kamga, C., Yazici, A., 2014. Impact of weather on urban transit ridership. Transp. Res. Part A Policy Pract. 69, 379–391. https://doi.org/10.1016/j.tra.2014.09.008.

Taylor, D., Mahmassani, H., 1997. Analysis of stated preferences for intermodal bicycle-transit interfaces. Transp. Res. Rec. 86–95 https://doi.org/10.1177/0361198196155600111.

Taylor, B.D., Miller, D., Iseki, H., Fink, C., 2009. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transp. Res. Part A Policy Pract. 43, 60–77. https://doi.org/10.1016/j.tra.2008.06.007.

Tu, W., Cao, R., Yue, Y., Zhou, B., Li, Qiuping, Li, Qingquan, 2018. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. J. Transp. Geogr. 69, 45–57. https://doi.org/10.1016/j.jtrangeo.2018.04.013.

Wang, Z.-J., Liu, Y., Chen, F., 2018. Evaluation and improvement of the interchange from bus to metro using smart card data and GIS. J. Urban Plan. Dev. 144, 05018004. https://doi.org/10.1061/(asce)up.1943-5444.0000435.

Warf, B., 2014. Geographically weighted regression. Encycl. Geogr. 28 https://doi.org/10.4135/9781412939591.n478.

Wei, M., Liu, Y., Sigler, T., Liu, X., Corcoran, J., 2019. The influence of weather conditions on adult transit ridership in the sub-tropics. Transp. Res. Part A Policy Pract. 125, 106–118. https://doi.org/10.1016/j.tra.2019.05.003.

Wu, J., Liao, H., 2020. Weather, travel mode choice, and impacts on subway ridership in Beijing. Transp. Res. Part A Policy Pract. 135, 264–279. https://doi.org/10.1016/j.tra.2020.03.020.

Yan, X., Levine, J., Zhao, X., 2019. Integrating ridesourcing services with public transit: an evaluation of traveler responses combining revealed and stated preference data. Transp. Res. Part C Emerg. Technol. 105, 683–696. https://doi.org/10.1016/j.trc.2018.07.029.

Yang, M., Zhao, J., Wang, W., Liu, Z., Li, Z., 2015. Metro commuters' satisfaction in multi-type access and egress transferring groups. Transp. Res. Part D Transp. Environ. 34, 179–194. https://doi.org/10.1016/j.trd.2014.11.004.

Yang, X., Yue, X., Sun, H., Gao, Z., Wang, W., 2021. Impact of weather on freeway origin-destination volume in China. Transp. Res. Part A Policy Pract. 143, 30–47. https://doi.org/10.1016/j.tra.2020.11.007.

Yu, H., Peng, Z.R., 2019. Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression. J. Transp. Geogr. 75, 147–163. https://doi.org/10.1016/j.jtrangeo.2019.01.004.

Zhao, J., Deng, W., Song, Y., Zhu, Y., 2013. What influences metro station ridership in China? Insights from Nanjing. Cities 35, 114–124. https://doi.org/10.1016/j.cities.2013.07.002.

Zhao, D., Wang, W., Woodburn, A., Ryerson, M.S., 2017. Isolating high-priority metro and feeder bus transfers using smart card data. Transp. (Amst). 44, 1535–1554. https://doi.org/10.1007/s11116-016-9713-7.

Zhao, D., Wang, Wei, Li, C., Ji, Y., Hu, X., Wang, Wenfu, 2019. Recognizing metro-bus transfers from smart card data. Transp. Plan. Technol. 42, 70–83. https://doi.org/10.1080/03081060.2018.1541283.

Zheng, Z., 2021. Reasons, challenges and some tools for doing reproducible transportation research. Commu. Transp. Res. 1, 2772–4247. https://doi.org/10.1016/j.commtr.2021.100004.

Zhou, M., Wang, D., Li, Q., Yue, Y., Tu, W., Cao, R., 2017. Impacts of weather on public transport ridership: results from mining data from different sources. Transp. Res. Part C Emerg. Technol. 75, 17–29. https://doi.org/10.1016/j.trc.2016.12.001.