land

MDPI

Article

The Impact of Land Use on Time-Varying Passenger Flow Based on Site Classification

Kexin Lei [1], Quanhua Hou [1], Weijia Li [2], Meng Zhao [1], Jizhe Zhou [1], Lingda Zhang [1], Shihan Chen [1] and Yaqiong Duan [1,*]

[1] School of Architecture, Chang'an University, Xi'an 710061, China
[2] The Engineering Design Academy of Chang'an University Co., Ltd., Xi'an 710064, China
* Correspondence: duanyaqiong@chd.edu.cn

Abstract: During the different periods of a day, the imbalanced distribution of inbound ridership, that is related to land use, results in extreme flow, which makes metro management challenging. The causes of imbalanced passenger flow from the perspective of land use in metro station areas are studied in this paper. More specifically, based on site classification, the impact of land use, including the floor area ratio and gross floor area on passenger flow, was explored by using a multiple linear regression model. The results first indicate that the impact intensities of the floor area ratio on peak hourly flow were 0.41, 0.21, and 0.20 around business, residential, and mixed sites, respectively. Second, for the abovementioned sites, the types with the greatest impact intensities of gross floor area on peak hourly flow were commercial and business facilities (B), residential (R), as well as administration and public services (A), which were 0.73, 0.32, and 0.87, respectively. Finally, for the land-development-control schemes for business, residential, and mixed sites, the maximum values of the floor area ratio were roughly 7.2, 5.3, and 8.2, respectively. The results presented in this study provide guidance for land development in metro station areas and contribute to avoiding the emergence of extreme passenger flow.

Keywords: land use; site classification; metro station areas; time-varying passenger flow; Xi'an

check for updates

Citation: Lei, K.; Hou, Q.; Li, W.; Zhao, M.; Zhou, J.; Zhang, L.; Chen, S.; Duan, Y. The Impact of Land Use on Time-Varying Passenger Flow Based on Site Classification. *Land* **2022**, *11*, 2189. https://doi.org/10.3390/land11122189

Academic Editors: Sheng Zheng, Yuzhe Wu and Ramesh P. Singh

Received: 20 October 2022
Accepted: 29 November 2022
Published: 2 December 2022

1. Introduction

With the rapid construction of metros, due to its characteristics of high efficiency and rapid speed, it has become the main means of public transport in domestic and foreign cities [1–3]. At the same time, metros also cause a series of urban safety management problems as a result of an increase in passenger flow on weekdays [4]. As the first city in Northwest China to open a metro, the explosive growth in construction in Xi'an in recent years has impacted upon the concentrated development of land in station areas. In turn, this has stimulated extreme flow, which has resulted in safety management issues, such as platform congestion and flow blockage [5,6]. Furthermore, for 80% of stations in Xi'an, the inbound passenger flow during off-peak periods is only around one-third of that during peak periods, and the passenger flow during peak periods exceeds the designed carrying capacity of the metro. Transportation demand is shaped by urban land use, and irrational land development stimulates extreme passenger flow [7,8]; therefore, identifying the impact of the floor area ratio and gross floor area on time-varying passenger flow is essential to controlling the high volume of metro users [9]. This study reviewed the literature with these research questions in mind. A growing body of work has explored the "source-flow correlation" following the proposed integration of transportation and land that [10,11]. This included the following two aspects: theoretical and methodical analysis. The first research aspect identified high-density land use as a key factor stimulating urban rail flow [12]. Kuby et al. proposed that land use is the dominant factor in metro ridership [13]. Related studies showed that the built environment in various station areas significantly impacts the travel behavior of residents, and the characteristics of passenger flow distribution at a metro station (i.e., the time-varying passenger flow) may differ between different periods [14]. Voulgaris et al. applied a variable set with 20 indicators concerning built environmental characteristics in order to investigate the relationship between the built environment and residents' travel activity [15]. Previous studies examined the mechanism of how the built environment impacts this flow [16–19]. It was suggested that the development of rail transit station land can yield passenger flow data [20]. The core factors pertaining to land in the station area that affect the time-varying passenger flow include form, intensity, and mixedness [21–24]. For instance, Cynthia et al. and Ding et al. adopted empirical location data and showed that intensity is one of the most important factors affecting passenger flow. More specifically, intensity exerts a positive effect on passenger flow [25,26]. Hence, the control of land use intensity in the station area has been proposed as a key measure to address this issue [27–29]. In a case study of the Shenzhen metro, both temporal and spatial features of passenger travel behaviors were examined using land use information to determine local land use development patterns [30].

The second aspect of this study established a regression model to explain the correlation between land and flow of travel. This model could identify a particular linear causal correlation between multiple independent factors and the dependent variable of flow [31]. Su et al. adopted this regression model to explore the correlation between passenger flow and various land use indicators, including density, diversity, and other elements, within an impact radius of 500 m around each Shanghai station [32,33]. With a similar methodology, by drawing on cases from Shenzhen and Xi'an, Gu and Duan et al. applied chart and linear regressions as essential approaches [34,35]. Moniruzzaman and Páez analyzed the relationship between railway service demand and land use in Hamilton, Canada, using a regression model that considered spatial auto correlation [36]. In addition, site classification is also a critical factor influencing station-level transit ridership [37,38].

From the abovementioned literature, it can be observed that the relationship between flow and land use has been explored in depth; however, the research regarding the imbalanced distribution of time-varying passenger flow caused by irrational land use in the station area has not been explored to the same extent. More specifically, previous studies examined the total daily passenger flow, which tends to obscure the imbalanced distribution characteristics of passenger flow at different times of the day. This paper examined the impact that the mechanism of land use around stations has on time-varying passenger flow. The aim of this study is to clarify the root causes of extreme flow, and to ensure the maximum values of land use in order to provide control schemes for local land development practices.

In comparison to previous studies, this paper emphasized the classification of site types, discussed the distribution characteristics of time-varying passenger flow, and proposed land development control schemes that are tailored to each type of metro station. On the one hand, from the perspective of the time-varying distribution of passenger flow within a day, this study focused on the problems arising from extreme flow at stations. It then explored the logical relationships behind the inbound passenger flow that is generated by land use around different types of stations; this can enrich the theoretical and methodological systems of land development in metro station areas. On the other hand, in many areas, the metro is still in the construction stage. Whether in newer districts or older districts, the areas around the station are faced with major tasks such as land development and renewal; therefore, in terms of solving practical problems, this study proposed control schemes for different types of stations to more accurately and effectively avoid extreme passenger flow and to ensure public safety.

2. Data and Methods

2.1. Data

As the metro line with the highest regional passenger volume, Xi'an metro line 2 runs through the Chang'an, Yanta, Beilin, Lianhu, and Weiyang districts, with an average daily passenger volume of 1.25 million [39]. In addition, the development of the surrounding land around 21 stations tends to be advanced (Figure 1).

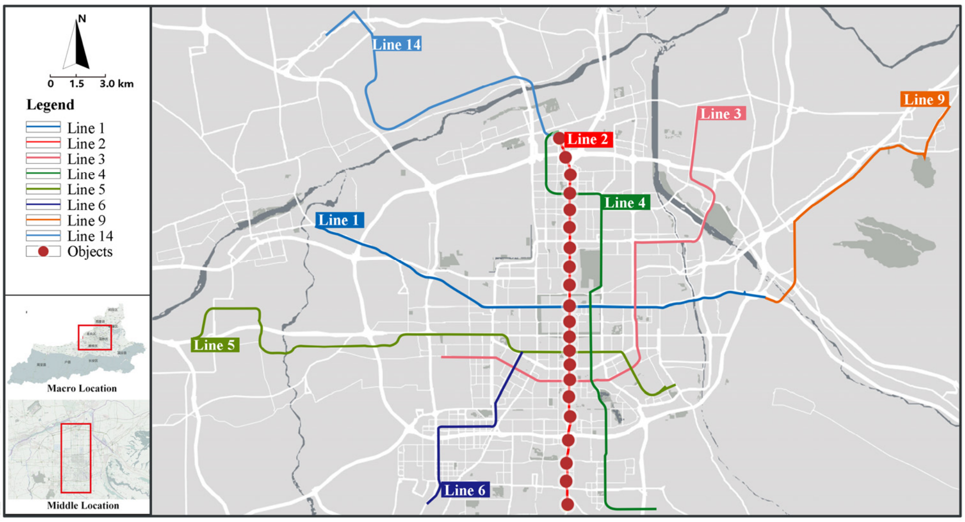

Figure 1. Station location schematic.

After selecting the abovementioned stations as the research object, in order to delineate the metro station areas, two main aspects were relied upon: pedestrian scale and site spacing. In particular, this study first considered the spatial distance that can be reached by pedestrians leaving the metro station on foot within 10 min, and then, the station radius was set within the range of 500–800 m due to the variation in walking speed. Second, for different countries, regions, and cities, the spacing of metro stations should be considered in accordance with local conditions. Generally, the radius of the station areas is set to half that of the station spacing in order to prevent the scope from overlapping [40,41]. In this study, based on a spacing distance of approximately 1000 m between metro stations in Xi'an, the buffer zone radius was determined to be 500 m. Inside the buffer zone (Figure 2), the data included land use and passenger flow (Table 1). First, the land use data aided the design by utilizing the Autodesk computer from the station area; the data were processed using high-definition satellite images (https://maps.google.com/ (accessed on 23 August 2021)), the GIS database, and field research conducted in 2021 which concerned the total floor area ratio (FAR) and gross floor area of different types of land (GFAs) (one unit: 10,000 m^2). The floor area ratio is the relationship between the total amount of usable floor area that a building has, or has been permitted to have, and the total area of the lot upon which the building stands. The gross floor area represents the sum of the gross floor area of all floors of all buildings on a site; it is measured by the exterior faces of the exterior walls or from the centerlines of walls separating two buildings. Second, the hourly passenger flow data were collected from the station's automatic fare collection (AFC) system from 23 to 27 August 2021.

Table 1. Data categories and sources.

Data Categories	Index	Significance	Sources	Unit
Land use	Floor area ratio (FAR)	Land develop-ment intensity	maps.google.com GIS database field research	/
	Gross floor area (GFA)	Land development function	maps.google.com GIS database field research	m^2
Passenger flow	Hourly flow	Passenger flow per hour	AFC system	person/hour

Figure 2. Processed land use data showing the scope of the research.

Each small graphic inside Figure 2 depicts the scope of the research; there are 21 discrete graphs that map the surroundings of the stations of line 2, and 11 land use types are colored differently, which refer to the cutlines.

According to the code for classifying urban land use, and the planning standards for developing land in China [42], the types of land in the station areas are divided into 11 categories, and the area of each type of land is measured. As the areas that contain streets and transportation services (S), logistical services and warehouses (W), railways (H21), special services (H4), structures under development (E9), municipal utilities (U), and industrial services (M) account for less than 5% each, the GFAs were calculated with the other four land types in mind. These four land types are: residential (R), administration and public services (A), commercial and business facilities, (B) and green spaces and squares (G) (Figure 3a). The GFAs of the four types are symbolized as R-GFA, A-GFA, B-GFA, and G-GFA. The ratios of the GFAs (R-GFA, A-GFA, B-GFA, and G-GFA) amount to the buffer area, which is the FAR, and they are plotted in Figure 3b, thus providing an essential indicator for determining the amount of construction land used [43].

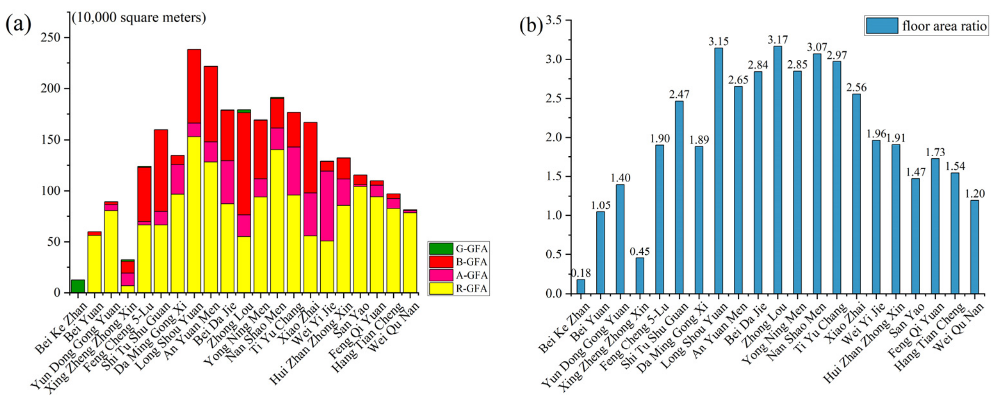

Figure 3. *Cont.*

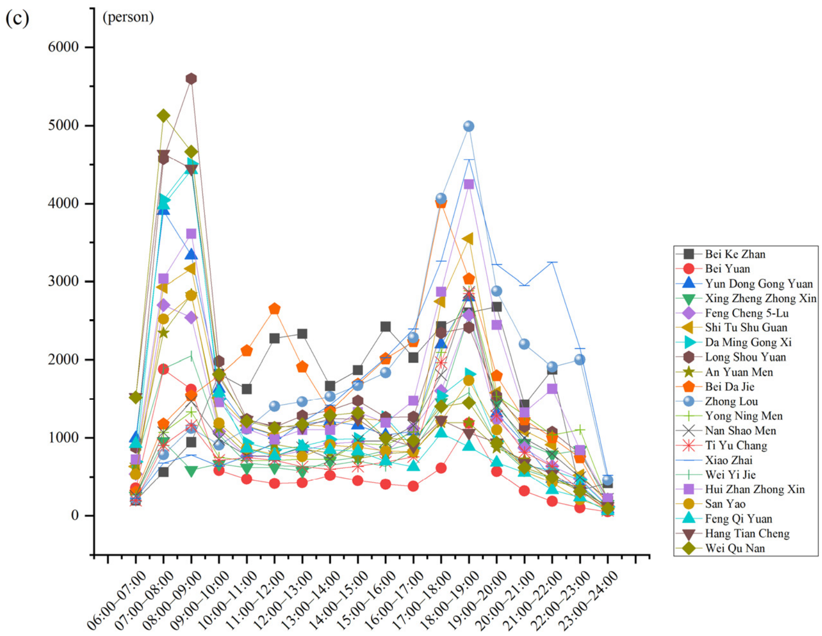

Figure 3. (**a**) Gross floor area of different types of land (GFAs); (**b**) total floor area ratio (FAR); (**c**) hourly passenger flow.

For the passenger flow, because the volume fluctuation of the total flow in different hours is slight, the time-varying distribution characteristics (adopting the total flow in and out of the station) may diminish. Compared with the outbound flow occurring on weekdays, the inbound flow has a higher degree of agglomeration, thus imposing greater pressure on the station's operation and management [32]; hence, the hourly inbound flow was adopted as the research index, and the data were collected from 06:00 to 24:00 in the week of 23 to 27 August 2021, as shown in Figure 3c.

2.2. Methods

In this study, a "data collection characteristics analysis model and construction index control" research path was proposed (Figure 4), and specific research content was applied.

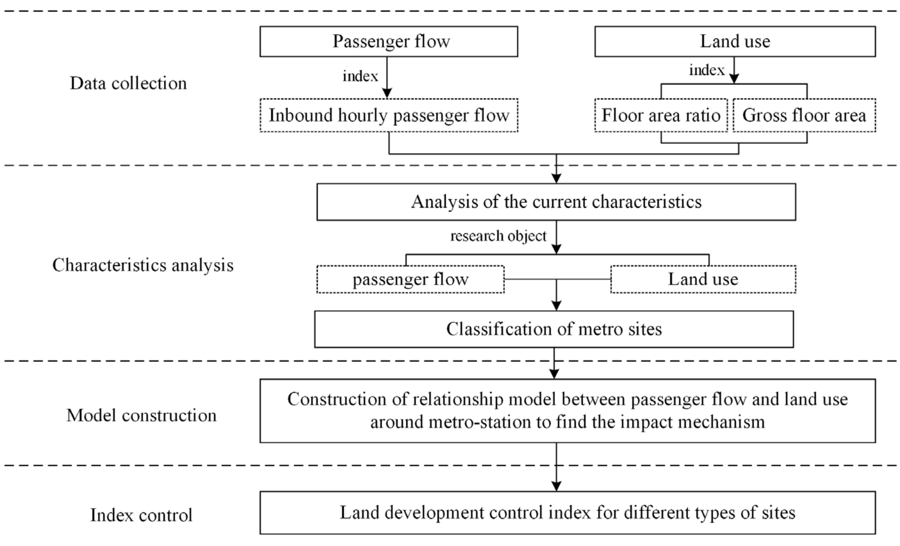

Figure 4. Research path.

- Classification of metro sites. According to the characteristics of land use around the station at present, and the time-varying law of the hourly coefficient on weekdays, the metro sites were classified into several types. Based on the site classification, a specific period division to vary the time periods was proposed.
- Impact mechanism of land use on hourly flow at various stations. Based on the independent variable types determined after applying the Pearson parameter dimensionality reduction, a linear regression model connecting the hourly flow with different land use indicators was constructed for different periods.
- Land use control for different types of sites. According to the fitting equation through the multiple linear regression model, the maximum values of the floor area ratio and gross floor area were proposed.

The methods were further refined for the above content. The probes on categorizing sites were conducted by using local land use mixtures [44] and the hourly coefficient for each station [45]. Then, the correlation of land use and the time-varying flow was analyzed with the Pearson correlation index. The multiple linear regression model was used to determine the separate impact of the factors. Finally, the reliability of the results was corroborated by the use of the mean absolute percentage error (MAPE).

1. Land use mixture

Land use mixture quantifies the abundance and complexity of land use types in the station area. Normally, it is abbreviated H with the general form:

$$H = -\sum_{k=1}^{n} P_k ln(P_k) \tag{1}$$

where P_k is the proportion of the total area of the site of the type in k, and n is the type of land use.

2. Hourly coefficient

The hourly coefficient normalizes the unit passenger volume, which can be used to characterize the volume variations in the time-varying distribution. Its general form is:

$$Y = \frac{a}{b} \times 100\% \tag{2}$$

where a is the passenger volume per hour, and b is the total passenger volume over the entire day.

3. Pearson correlation coefficient

The Pearson correlation coefficient can help to screen for indicators that are less correlated with dependent variables, while avoiding the covariance between independent variables. The value is bounded by -1 and 1 and it embodies the correlation between land use indicators and hourly flow. Its general form is:

$$r = \frac{\sum_{i=1}^{n}(X_i - X)(Y_i - Y)}{\sqrt{\sum_{i=1}^{n}(X_i - X)^2}\sqrt{\sum_{i=1}^{n}(Y_i - Y)^2}} \tag{3}$$

where r is the Pearson correlation coefficient, X represents the GFAs of four area types (i.e., residential, administration and public services, commercial and business facilities, as well as green spaces and squares), and Y represents the hourly flow that corresponds with different periods.

4. Multiple linear regression model

The multiple linear regression model is more accurate than using one variable to estimate the dependent variable; this can reflect the actual situation in a realistic and effective way. Moreover, multiple linear regression is suitable for simulations using multiple scenarios and the model is relatively simple and easy to manipulate. In this study, a multiple regression analysis by IBM SPSS Statistics Version 22 software was employed to quantify the impact that different land use indicators have on the hourly flow in the station area. Equation (4) is the general equation for the regression model(s), comprising one dependent variable, Y, and k independent variables, X_1, X_2, ..., X_k.

$$Y = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \beta_0 \quad (4)$$

where the dependent variable Y represents the average hourly flow in a certain period, the independent variable X symbolizes the FAR and GFAs, β_0 is a constant, and β_1, β_2, ..., β_k are unstandardized regression coefficients.

5. Mean absolute percentage error
6. To test the reliability of the fitting equation, the mean absolute percentage error (MAPE) was calculated using Equation (5). This formula was used in this study to measure the difference between the actual and fitted flow. If the value is less than 30%, the fitting equation is valid.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \times 100\% \quad (5)$$

where Y_i represents the actual value, $\hat{Y}_i$ represents the fitted value, and n represents the sample size.

3. Results and Discussion

3.1. Site Classification

There were four categories for all sites that were distinguishable with the hourly coefficient: residential, business, mixed, and traffic type (Table 2). The proportional variation on the hourly inbound volume, displayed in the broken-line graphs of Figure 5, embodied the time-varying flow distribution for each station category. Remarkably, the result of the site classifications through land use mixture was convergence with an hourly flow percentage (Figure 6). This indicates that certain correlations existed between land use and hourly flow in the station area.

Table 2. Site classification.

Type	Station(s)	Number
Residential [1]	Bei Yuan, Yun Dong Gong Yuan, Da Ming, Gong Xi, Long Shou Yuan, An Yuan Men, Wei Yi Jie, San Yao, Feng Qi Yuan, Hang Tian Cheng, and Wei Qu Nan	10
Business [2]	Xing Zheng Zhong Xin, Bei Da Jie, Zhong Lou, Yong Ning Men, Nan Shao Men, Ti Yu Chang, and Xiao Zhai	7
Mixed [3]	Feng Cheng 5-Lu, Shi Tu Shu Guan, and Hui Zhan Zhong Xin	3
Traffic [4]	Bei Ke Zhan	1

[1] For stations at 07:00–09:00, the coefficient is 10–25%; [2] for stations at 17:00–19:00, the coefficient is 10–25%; [3] for stations at 07:00–09:00 and 17:00–19:00, the coefficient is 5–15%; [4] and for stations at 06:00–24:00, the coefficient is 1–10% [46].

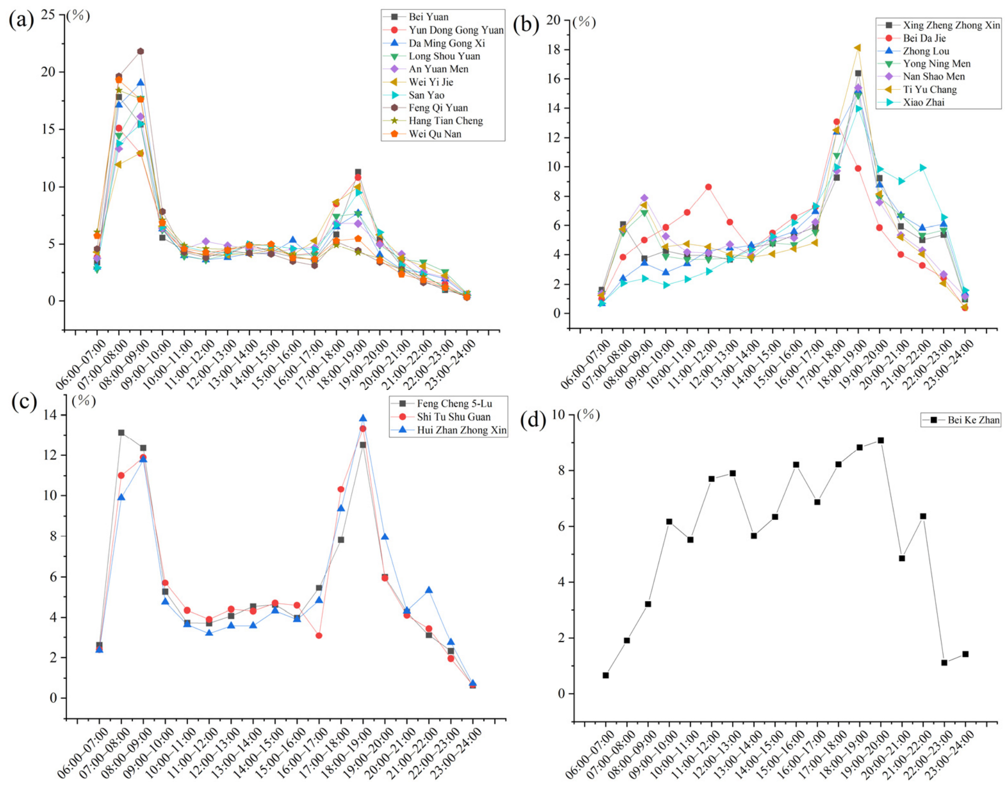

Figure 5. (**a**) Residential; (**b**) business; (**c**) mixed; and (**d**) traffic sites.

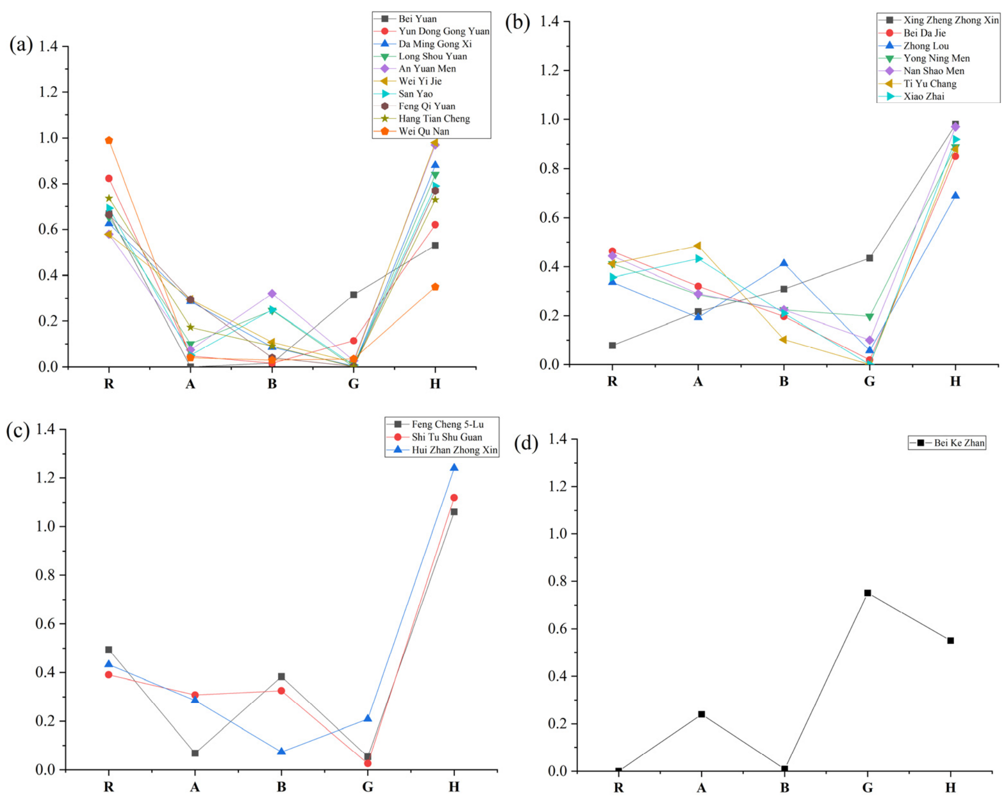

Figure 6. (**a**) Residential; (**b**) business; (**c**) mixed; and (**d**) traffic sites.

As shown in Figure 5a–d, with the varying times for flow distribution at all stations, different peak and off-peak periods existed at different sites, in accordance with the hourly coefficient. At the residential sites, 07:00–09:00 was set as the peak period because significant tidal phenomena can be observed in those stations, and the off-peak periods referred to the two periods, 06:00–07:00 and 09:00–24:00. At the business sites, 17:00–19:00 was set as the peak period, with two off-peak periods at 06:00–17:00 and 19:00–24:00. At the mixed sites, the peak periods were 07:00–09:00 and 17:00–19:00, which displayed a double peak feature as the volume character and the off-peak periods were 06:00–07:00, 09:00–17:00, and 19:00–24:00. Conversely, as shown in Figure 5d, the passenger volume had no particular distribution in terms of time for the traffic site, with primarily flexible travelers using the service (non-commuting).

According to the land characteristics at present, R-type land use occupied a major proportion of the residential sites; it accommodated a high number of residents, and the mixture was less than 1.0 (Figure 6a). Most inbound passengers during 07:00–09:00 on weekdays originated from R-type land and commuted to work. Around the business sites, the main land use categories were generally identified as B-types with a large, employed population; the mixture was less than 1.0 (Figure 6b). Most inbound passengers during 17:00–19:00 on workdays originated from B-type land and commuted to their households. For mixed sites, the mixture of land use types was relatively higher, exceeding 1.0 (Figure 6c). Lastly, the land surrounding the traffic site, where H was less than 1.0, contained a lower residential population and fewer business premises (Figure 6d), thus resulting in a lower overall volume and a balanced distribution.

Additionally, the passenger volume variation was largely determined by the railway operation schedule in the only railway–metro hinge on line 2, called Bei Ke Zhan; therefore, the pressure on station management as a result of extreme flow was relatively low, which shows that a traffic-type station is not an interesting subject for research.

3.2. Impact of Floor Area Ratio on Hourly Flow

Multiple linear regression was used to explore the impact of FAR (X) on the hourly flow (Y), where three independent variables, X_1, X_2, and X_3 represent the FAR of the residential, business, and mixed sites, respectively, and the dependent variables, Y_1 and Y_2, represent the hourly flow during peak and off-peak periods, respectively (Table 3). The standardization coefficient (k) represents the impact intensity of FAR on the hourly flow. The impact intensities of FAR on the hourly flow change in residential, business, and mixed sites were 0.21, 0.41, and 0.20 during peak periods, respectively, and 0.42, 0.33, and 0.23 during off-peak periods, respectively. In Table 3, a represents the unstandardized coefficient and k represents the standardized coefficient.

Table 3. Corresponding coefficients.

	Residential			**Business**			**Mixed**		
	a	**Constant**	**k**	**a**	**Constant**	**k**	**a**	**Constant**	**k**
Y_1	403.46	2841.24	0.21	422.36	1961.39	0.41	331.99	2284.76	0.20
Y_2	167.40	602.56	0.42	136.18	755.26	0.33	15.14	976.63	0.23

The results indicate a positive impact between FAR and hourly flow, which is consistent with the results of previous studies [47–49]. Total development intensity addresses the increase in the local population which imposes an increase in the hourly flow. In short, a reduction in FAR contributes to reasonable restrictions on local developments, which indirectly and effectively controls the flow rate for station management.

3.3. Impact of Gross Floor Area on Hourly Flow

The Pearson correlation was adopted to analyze the correlation between GFAs and passenger hourly flow in order to filter the types of GFAs; then, linear regression was used to explore the impact of this correlation and filtration on the hourly flow.

3.3.1. Pearson Correlation

The Pearson correlation was used for correlating the GFAs of four area types (i.e., residential, administration and public services, commercial and business facilities, as well as green spaces and squares) and the peak (PHF: person/hour) or off-peak hourly flow (OPHF: person/hour). Figures 7–9 show the matrix results, in which the panels marked with *a* represent peak periods and those marked with *b* represent off-peak periods.

7. Residential Sites

During the peak periods, PHF was positively correlated with R-GFA and B-GFA (the coefficients were 0.45 and 0.07, respectively), and negatively correlated with G-GFA and A-GFA (the coefficients were −0.43 and −0.30, respectively). During the off-peak periods, the OPHF was positively correlated with R-GFA and B-GFA (the coefficients were 0.53 and 0.30, respectively), and negatively correlated with G-GFA and A-GFA (the coefficients were −0.12 and −0.11, respectively); therefore, the positive correlation between R-GFA and the hourly flow was the strongest, as the R-GFA had a greater proportion in the station area, which may result in higher passenger flow rates.

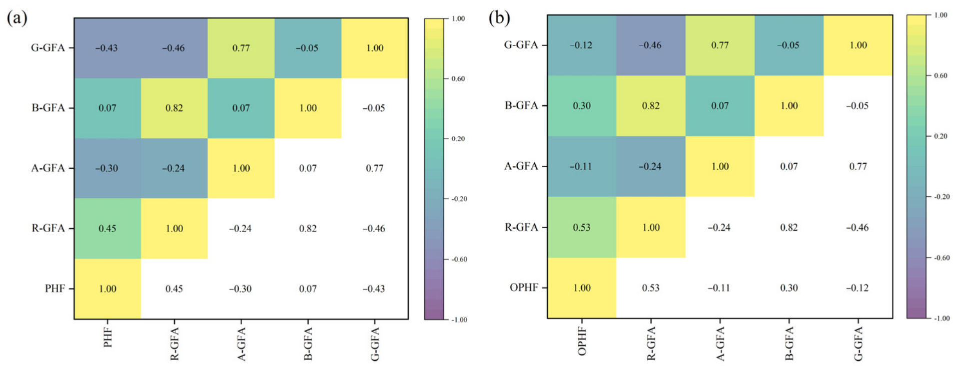

Figure 7. Correlations at residential sites, (**a**) peak periods; (**b**) off-peak periods.

8. Business Sites

For the business sites during the peak periods, a positive correlation was observed between PHF and B-GFA, A-GFA, and G-GFA, as indicated by the positive coefficients (0.91, 0.29, and 0.29, respectively). Nevertheless, the PHF was negatively correlated with R-GFA, as indicated by a coefficient of −0.17. During the off-peak periods, the OPHF was positively correlated with the floor areas of B, A, and G (the coefficients were 0.77, 0.33, and 0.12, respectively). Conversely, a negative correlation was observed for R-GFA (the coefficient was −0.15). A positive correlation was observed between the B-type GFA with an hourly flow, which addressed a large proportion of the employed population.

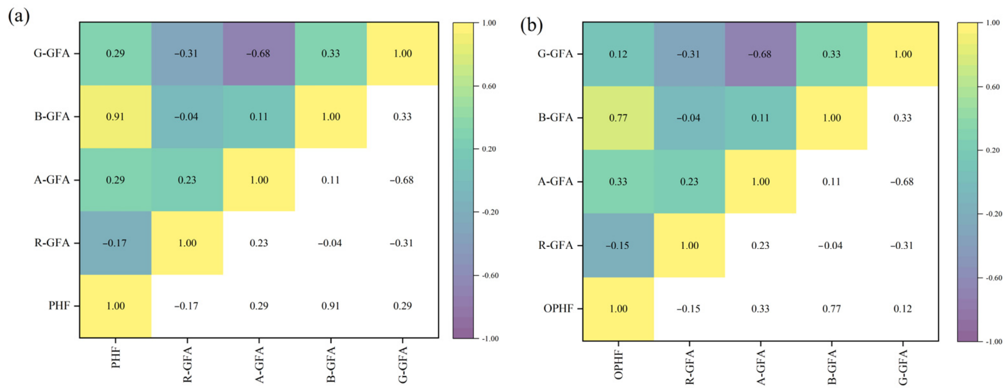

Figure 8. Correlations at business sites, (**a**) peak periods; (**b**) off-peak periods.

9. Mixed sites

The floor area types of A and R and the hourly flow at mixed sites were all positively correlated during the peak (coefficients of 0.96 and 0.75, respectively) and off-peak (coefficients of 0.99 and 0.85, respectively) periods. Conversely, the hourly flow was negatively correlated with G-GFA and B-GFA during the peak (coefficients of −0.84 and −0.37, respectively) and off-peak (coefficients of −0.73 and −0.52, respectively) periods. Consequently, A-GFA was identified as the most relevant factor that affected the local inbound flow.

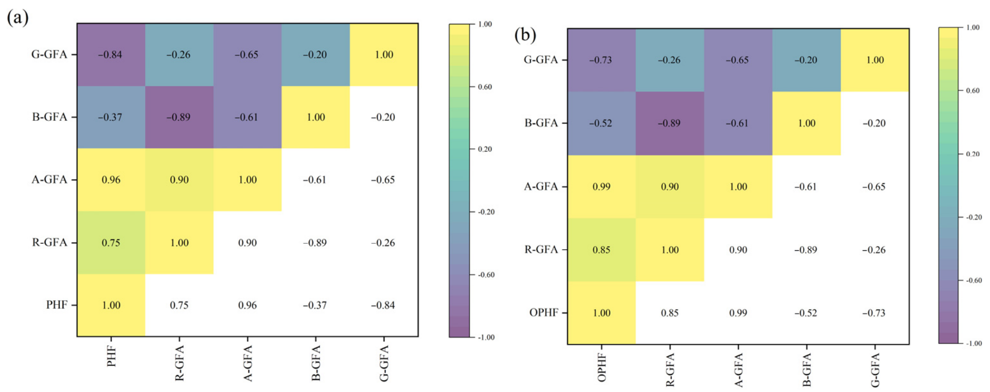

Figure 9. Correlations at mixed sites, (**a**) peak periods; (**b**) off-peak periods.

3.3.2. Impact Intensity

According to Yao and Dong [50], a threshold value greater than 0.2 and lower than 0.8 was proven as a valid coefficient to screen out the types of GFAs (Table 4). R-GFA, A-GFA, and G-GFA were obtained and used to compute the respective impact intensity on the PHF in residential sites. R-GFA and B-GFA were sorted for intensity computations with OPHF. For the business sites, A-GFA, B-GFA, and G-GFA were applied for computing the impact intensity on PHF, and A-GFA and B-GFA were adopted for OPHF. At the mixed sites, the land use factors A-GFA, B-GFA, and G-GFA were linked to different periods. Linear regression was employed in order to identify the impact of GFAs on flow using IBM SPSS Statistics Version 22 software. The correlation coefficients shown in the following tables, R and R^2, of the regression model both exceed 0.4; thus, the fitting equation is reliable.

Table 4. Types of gross floor area after screening.

Types of Sites	Residential	Business	Mixed
Peak hourly flow	R, A, G	A, B, G	A, B, G
Off-Peak hourly flow	R, B	A, B	A, B, G

10. Peak periods

Tables 5 and 6 shows the multi-regression results of PHF and R-GFA, A-GFA, and G-GFA at the residential sites. The following model-fitting equation was constructed:

$$Y = 12.82R - 0.99A - 2512.61G + 2613.14 \quad (6)$$

Table 5. Summary of models for residential sites.

R	R^2	Adjusted R^2	Standard Error	Sig.
0.54	0.51	0.48	1286.51	0.03

Table 6. Corresponding coefficients.

Model	Non-Standard Coefficient B	Non-Standard Coefficient Standard Error	Standard Coefficient Beta	T	Sig.
	2613.14	593.96		4.40	0.02
R-GFA	12.82	1.25	0.32	10.29	0.00
A-GFA	−0.99	0.17	−0.02	−5.82	0.01
G-GFA	−2512.61	592.98	−0.27	−4.32	0.02

From the standardized coefficient, it can be observed that the impact intensity of R-GFA on PHF was highest at 0.32.

Tables 7 and 8 shows the multi-regression results for the PHF and A-GFA, B-GFA, and G-GFA at the business sites. The following model-fitting equation was constructed:

$$Y = 31.13A + 23.88B + 360.54G + 609.90 \tag{7}$$

Table 7. Summary of models for business sites.

R	R^2	Adjusted R^2	Standard Error	Sig.
0.95	0.91	0.81	413.65	0.04

Table 8. Corresponding coefficients.

Model	Non-standard Coefficient B	Non-standard Coefficient Standard Error	Standard Coefficient Beta	T	Sig.
	609.90	143.51		4.25	0.03
A-GFA	31.13	6.90	0.47	4.51	0.02
B-GFA	23.88	7.26	0.73	3.29	0.04
G-GFA	360.54	62.70	0.37	5.75	0.01

From the standardized coefficient, it can be observed that the impact intensity of the B-GFA on PHF was highest at 0.73.

For the mixed sites, Tables 9 and 10 shows the multi-regression results for PHF and A-GFA, B-GFA, and G-GFA. The following model-fitting equation was constructed:

$$Y = 20.15A - 10.39B - 1244.61G + 4077.70 \tag{8}$$

Table 9. Summary of models for mixed sites.

R	R^2	Adjusted R^2	Standard Error	Sig.
0.73	0.62	0.56	413.65	0.02

Table 10. Corresponding coefficients.

Model	Non-standard Coefficient B	Non-standard Coefficient Standard Error	Standard Coefficient Beta	T	Sig.
	4077.70	959.50		4.24	0.03
A-GFA	20.15	1.78	0.87	11.34	0.00
B-GFA	−10.39	3.05	−0.41	−3.41	0.04
G-GFA	−1244.61	216.93	−0.33	−5.74	0.01

From the standardized coefficient, it can be observed that the impact intensity of A-GFA on PHF was highest at 0.87.

11. The results presented above indicate that the inbound flow originates from regions with different land use types. During the peak periods, R-type land around residential sites provided the majority of the flow due to the high number of people traveling to work, and the peak times for mixed sites was the result of people commuting on weekdays. B-type land around the business sites provided the majority of the flow as individuals were returning to their accommodation [51,52]. Notably, the lower R and R^2 of the PHF and GFAs that are fit for residential sites may have been due to the fact that during weekday peak hours, other factors have a greater impact on the flow because of the requirements for compressed commuting times; for instance, residents may consider whether to transfer, bus stop density, and so on. These factors create a strong interference with the dependent variable, thus resulting in a slightly lower fitting effect of R-GFA, A-GFA, and G-GFA on passenger flow.
12. Off-peak periods

Tables 11 and 12 shows the results of the multi-regression analysis between the OPHF of the residential sites and the R-GFA and B-GFA. The following model-fitting equation was constructed:

$$Y = 7.43R - 3.99B + 289.60 \quad (9)$$

Table 11. Summary of models for residential sites.

R	R^2	Adjusted R^2	Standard Error	Sig.
0.59	0.45	0.43	232.17	0.03

Table 12. Corresponding coefficients.

	Non-Standard Coefficient		Standard Coefficient	T	Sig.
Model	B	Standard Error	Beta		
	289.60	45.67		6.34	0.02
R-GFA	7.43	0.82	0.90	9.06	0.01
B-GFA	−3.99	0.78	−0.45	−5.12	0.03

From the standardized coefficient, the impact intensity of R-GFA on OPHF was highest at 0.90.

Tables 13 and 14 shows the results of the multi-regression analysis between the OPHF, A-GFA, and B-GFA at the business sites. The following model-fitting equation was constructed:

$$Y = 6.63A + 9.83B + 417.66 \quad (10)$$

Table 13. Summary of models for business sites.

R	R^2	Adjusted R^2	Standard Error	Sig.
0.80	0.65	0.47	283.13	0.02

Table 14. Corresponding coefficients.

	Non-Standard Coefficient		Standard Coefficient	T	Sig.
Model	B	Standard Error	Beta		
	417.66	42.10		9.92	0.01
A-GFA	6.63	0.31	0.24	21.39	0.00
B-GFA	9.83	2.05	0.74	4.79	0.04

From the standardized coefficient, the impact intensity of B-GFA on OPHF was highest at 0.74.

For the mixed sites, Tables 15 and 16 shows the results of the multi-regression analysis between OPHF and A-GFA, B-GFA, and G-GFA. The following model-fitting equation was constructed:

$$Y = 19.25A - 4.81B - 426.41G + 1453.02 \quad (11)$$

Table 15. Summary of models for mixed sites.

R	R^2	Adjusted R^2	Standard Error	Sig.
0.95	0.89	0.88	416.22	0.03

Table 16. Corresponding coefficients.

	Non-Standard Coefficient		Standard Coefficient	T	Sig.
Model	B	Standard Error	Beta		
	1453.02	429.50		3.38	0.04
A-GFA	19.25	3.23	0.84	5.96	0.01
B-GFA	−4.81	1.09	−0.53	−4.42	0.03
G-GFA	−426.41	96.93	−0.29	−4.40	0.03

From the standardized coefficient, the impact intensity of A-GFA on OPHF was highest at 0.84.

The results presented above show that during the off-peak periods, for the sites with different types, the inbound flow of people primarily originates from residential, commercial, and business facilities, as well as administration and public services land. This is generally consistent with the findings of previous studies [53,54]. These types of land are the primary sources of station space which generate a steady flow of people into the station. Moreover, cumulative passenger flow during off-peak periods is lower than during the peak periods at all stations, which is due to the reduced travel demand.

The actual and fitted passenger flow for the peak and off-peak periods of each station is obtained using Equations (6)–(11), as shown in Figure 10a–f, respectively. The trend of the value was basically similar, and the average MAPE values of the three types of sites during peak and off-peak periods were all less than 30%, within an acceptable range (Table 17); therefore, all fitting equations were reliable for ascertaining the correlations between the GFAs and hourly flow. Additionally, they can be employed for forecasting the trend when the flow changes; however, other factors, beyond those pertaining to land, can also impact the hourly flow, and thus, the fitting equations need to be perfected in subsequent research.

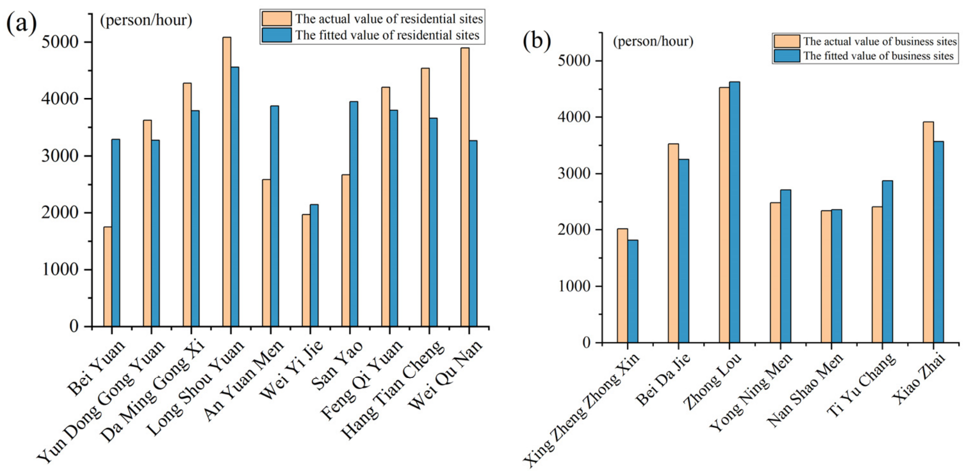

Figure 10. *Cont.*

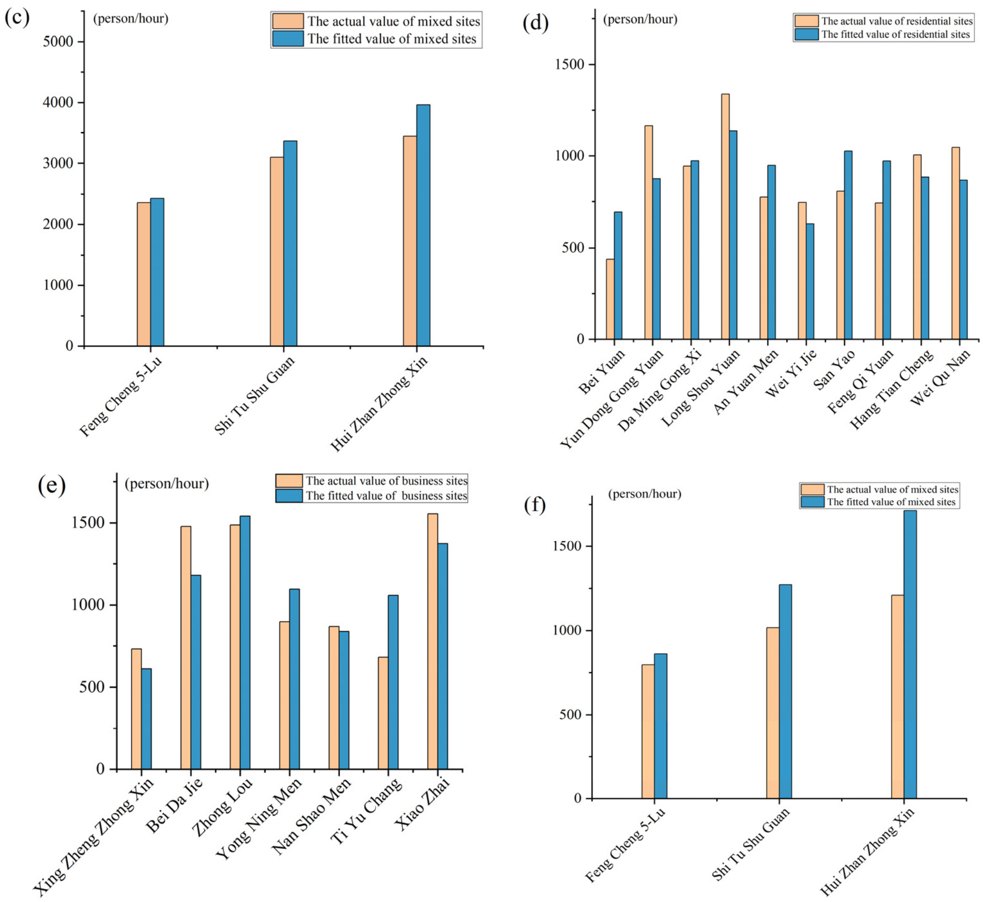

Figure 10. (**a**–**c**) the actual and fitted passenger flow values during peak periods; (**d**–**f**) the actual and fitted passenger flow values during off-peak periods.

Table 17. Mean absolute percentage error (MAPE).

Periods	Residential		Business		Mixed	
	Peak	**Off-Peak**	**Peak**	**Off-Peak**	**Peak**	**Off-Peak**
MAPE (%)	28	22	8	19	8	25

3.4. Land Development Control

The rational development of the area surrounding the station ensures that the appropriate population type provides a reasonable passenger flow for public transportation. According to the empirical analysis of Xi'an metro line 2, the flow was positively correlated with the FAR and GFAs in the station area. The three types of land (R, B, and A) had the greatest impact intensities on the PHF of the metro. As a result, these types of land use are more important than others to curb the appearance of extreme flow during peak periods. According to previous research, the constraint values can be determined by combining the long-term peak hourly flow [55,56]. From the above research results presented in Sections 3.2 and 3.3, it can be observed that the control suggestions mainly target the orbit of the FAR and GFAs for these three types of different sites (Table 18).

Table 18. Maximum development values for control.

Site Type	FAR	GFA (One Unit: 10,000 m^2)
Residential sites	5.3	204.68(R-GFA)
Business sites	7.2	132.29(B-GFA)
Mixed sites	8.2	101.19(A-GFA)

- Residential sites

The proportion of residential land in the station area is close to 50% with a simple land structure, and less than 10% of land is dedicated to office, commercial, and municipal utilities. As indicated by the fitting formula, the maximum development values of the FAR and R-GFA of the station area should be approximately 5.3 and 204.68, respectively (one unit: 10,000 m^2).

- Business sites

The proportion of land is mainly used for offices and residences, the total proportion of which exceeds 50%, and there are few commercial or municipal utilities, or green spaces. B-GFA presented the greatest number of impact intensities on the flow during peak periods. As indicated by the fitting formula, the maximum development values of the FAR and B-GFA of the station area should be approximately 7.2 and 132.29, respectively (one unit: 10,000 m^2).

- Mixed sites

In terms of land use, spatial distribution, and passenger flow characteristics, the mixed sites area lies between residential and business areas, as it has a high land use mix and the most balanced flow distribution. Overall, mixed sites present outstanding conditions in all aspects. As indicated by the fitting formula, the maximum values for the development of the FAR and A-GFA of the station area should be approximately 8.2 and 101.19, respectively (one unit: 10,000 m^2).

In this study, the impact of land use on the hourly flow was identified by adopting a multiple linear regression model, including two aspects concerning the FAR and GFAs. As shown in Figure 11, these two aspects representing land use intensity and function impact the increase in hourly flow in residential, business, and mixed sites. Moreover, the intensity control of land use development can be guided by assuming the maximum flow as the threshold.

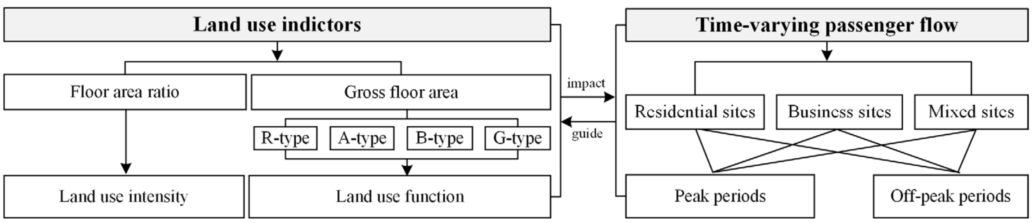

Figure 11. Results framework.

4. Conclusions

In recent years, we have witnessed the rapid construction of metros. Identifying the impact of land use around stations on passenger flow is of great significance for obtaining a scientific understanding of the time-varying distribution of flow, and it provides a theoretical basis for avoiding extreme flow. According to the results framework presented above (Figure 11), the following conclusions can be summarized.

Based on the distribution characteristics of the hourly flow during the day, all stations were classified as residential, business, mixed, and traffic, and each type of site had a different division between peak and off-peak periods. The impact mechanism of land use on hourly flow obtained by multiple linear regression at different stations may differ.

In terms of the methodological application, the study also followed the linear regression method adopted by Chang et al. and Su et al. [31,32]; however, it is worth noting that the Pearson correlation coefficient was used to filter out the independent variable indicators beforehand, which can improve the fitting degree of the gross floor area to passenger flow.

For the theoretical study, the surrounding land use was the main factor influencing the passenger flow and there was an objective "source–flow correlation" between the two factors. The floor area ratio and different types of gross floor area were core indicators that represented land use development intensity and function, which is similar to the results obtained by Yoshiharu et al. and Ding et al. [21,26]. In particular, the types of gross floor area that had the greatest impact on the hourly flow during peak periods were residential areas (R), commercial and business facilities (B), and administration and public services (A). At the same time, this paper was based on site classification, and therefore, development control values also differed. At the residential, business, and mixed sites, these indicators could be subject to the strongest control factors, with values of 204.68, 132.29, and 101.19 (one unit: 10,000 m^2), and the maximum values of the total FAR should be approximately 5.3, 7.2, and 8.2, respectively. In the future development of new districts and the renewal of old cities, the results can be used to guide site classification and land development strategies.

- The emergence of extreme passenger flow at metro stations was influenced by a number of complex factors, including the built environment of the station area; therefore, there were some limitations to this study. In particular, spatial factors other than land use were not considered, and the sample size was small, as only line 2 was considered for the study. Finally, the presented study failed to consider the spatial variability of station distribution; therefore, subsequent studies should be improved in the following three ways:
- More consideration should be given to the factors concerning the built environment that impact passenger flow at the station, including the distribution of public stations, the amount of parking, and the quality of the pedestrian environment, so that extreme passenger flow can be more accurately and effectively managed to ensure public safety.
- The diversity of the study sites can further improve the accuracy of the results; therefore, the number and type of sample sites needs to be expanded in future studies. For example, all 116 sites within the main city of Xi'an could be selected for the study.
- The GWR and GTWR models consider the spatial variability of different sites, and therefore, they have the advantage of being more accurate. Further research can introduce spatial parameters through these models to propose more reasonable diversified development strategies for different types of sites.

Author Contributions: Conceptualization, Q.H., L.Z. and Y.D.; Data curation, K.L., W.L. and S.C.; Formal analysis, K.L. and S.C.; Funding acquisition, K.L. and J.Z.; Investigation, W.L.; Methodology, K.L., J.Z. and L.Z.; Project administration, M.Z.; Resources, Q.H. and W.L.; Software, M.Z. and L.Z.; Supervision, Q.H. and Y.D.; Validation, J.Z. and Y.D.; Visualization, S.C.; Writing—original draft, K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Basic Research Plan in Shaanxi Province of China, grant number 2021JQ-285, the Fundamental Research Funds for the Central Universities, CHD, grant number 300102411106, and the Fundamental Research Funds for the Central Universities, CHD, grant number 300102411608.

Data Availability Statement: Not applicable.

Acknowledgments: We are thankful to the anonymous reviewers for their valuable comments.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Atkinson-Palombo, C.; Kuby, M.J. The geography of advance transit-oriented development in metropolitan Phoenix, Arizona, 2000–2007. *J. Transp. Geogr.* **2011**, *19*, 189–199. [CrossRef]
2. Zegras, C. Land Use and Transport in China. *J. Transp. Land Use* **2010**, *3*, 1–3. [CrossRef]
3. Choi, J.; Lee, Y.; Kim, T.; Sohn, K. An analysis of metro ridership at the station-to station level in Seoul. *Transportation* **2012**, *39*, 705–722. [CrossRef]
4. Shi, J.G.; Yang, J.; Yang, L.X. Safety-oriented cooperative passenger flow control model in peak hours for a metro line. *J. Transp. Syst. Eng. Inf. Technol.* **2019**, *19*, 125–131.
5. Truong, R.; Gkountouna, O.; Pfoser, D.; Züfle, A. Towards a Better Understanding of Public Transportation Traffic: A Case Study of the Washington, DC Metro. *Urban Sci.* **2018**, *2*, 65. [CrossRef]
6. Wang, K.; Zhang, J.J.; Zhang, D.; Wu, X. A priority in land supply for sustainable transportation of Chinese cities: An empirical study from perception, discrimination, linkage to decision. *Land* **2022**, *11*, 78. [CrossRef]
7. Cao, X.S.; Yang, F.; Xiao-Pei, Y. Study on the urban transport and land-use of Guangzhou. *Chin. Geogr. Sci.* **2000**, *2*, 144–150. [CrossRef]
8. Hou, Q.H.; Duan, Y.Q.; Ma, R.G. Collaborative optimization of land use intensity and traffic capacity under hierarchical control rules of regulatory detailed planning. *J. Chang'an Univ.* **2015**, *35*, 114–121.
9. Moon, Y.I.; Rho, J.H. An empirical analysis on public transportation demand and TOD design factors in Seoul subway adjacent area. *Int. J. Highw. Eng.* **2011**, *13*, 211–220. [CrossRef]
10. Knight, R.L.; Trygg, L.L. Land use impacts of rapid transit: Implications of recent experience. *Energy Plan. Policy Econ.* **1977**. [CrossRef]
11. Werner, C. Integrated land-use and transport modeling-Decision chains and hierarchies—Delabarra, T. *Environ. Plan. B Plan. Des.* **1990**, *19*, 487–490.
12. Zhao, L.; Shen, L. The impacts of rail transit on future urban land use development: A case study in Wuhan, China. *Transp. Policy* **2018**, *81*, 396–405. [CrossRef]
13. Kuby, M.; Barranda, A.; Upchurch, C. Factors influencing light-rail station boardings in the United States. *Transp. Res. Part A Policy Pract.* **2004**, *38*, 223–247. [CrossRef]
14. Cervero, R.; Kockelman, K. Travel demand and the 3ds: Density, diversity, and design. *Transp. Res. Part D Transp. Environ.* **1997**, *2*, 199–219. [CrossRef]
15. Voulgaris, C.T.; Blumenberg, E.; Taylor, B.D.; Brown, A.; Ralph, K.M. Neighborhood Character and Travel Behavior: Comprehensive Analysis of the United States in the 2000s. In Proceedings of the Transportation Research Board 95th Annual Meeting, Washington, DC, USA, 8–12 January 2016.
16. Teng, S.H. Analysis of the relationship between land use, development, and construction near a rail traffic station and passenger flow volume: The case of Tianjin metro line 1. *Appl. Mech. Mater.* **2013**, *357–360*, 1856–1862. [CrossRef]
17. Wu, H.; Gao, L. Functional analysis of passenger flow at rail transit in Beijing based on the usage of land. *J. Beijing Univ. Civ. Eng. Archit.* **2015**, *31*, 28–35.
18. Zhang, Y.; Cao, K.; He, Y.W.; Zhou, W. Discussion on spatial match between land use and rail transit: A case study of Shenzhen subway line 2. *City Plan. Rev.* **2017**, *41*, 107–115.
19. Lewis-Workman, S.; Brod, D. Measuring the neighborhood benefits of rail transit accessibility. *Transp. Res. Rec. J. Transp. Res. Board* **1997**, *1*, 147–153. [CrossRef]
20. Zhuang, Y.; Yuan, M. An analysis on the influence of passenger flow on development intensity in metro station area in the core area of downtown shanghai. *Archit. J.* **2017**, *2*, 22–26.
21. Yoshiharu, S. Factors in expansion of low-use and unused land in core areas of regional centers: Comparative study between Matsuyama city and Takamatsu city. *Geogr. Sci.* **2007**, *62*, 65–78.
22. Chen, X. *Driving Factors Analysis on Urban Vibrancy: A Case Study of Chongqing Main Area*; Springer: Singapore, 2021; pp. 1137–1147.
23. Sung, H.; Choi, K.; Lee, S.; Cheon, S.H. Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership. *J. Transp. Geogr.* **2014**, *36*, 134–140. [CrossRef]
24. Duan, Y.Q.; Hou, Q.H.; Zhang, S.Y. Review of intensive land use in built-up area based on low-carbon travel. *Planners* **2019**, *35*, 5–11.
25. Cynthia, C.; Jason, C.; James, B. Diurnal pattern of transit ridership: A case study of the new york city subway system. *J. Transp. Geogr.* **2009**, *17*, 176–186.
26. Ding, C.; Cao, X.Y.; Liu, C. How does the station-area Build environment influence metro rail ridership? Using gradient boosting decision trees to identify non-linear threshold. *J. Transp. Geogr.* **2019**, *77*, 70–78. [CrossRef]
27. Mo, Y.K.; Deng, J.; Wang, J.Y. A land use model for urban rail station area planning based on Tod strategy. *J. Civ. Archit. Environ. Eng.* **2009**, *31*, 116–120.
28. Wang, J.Y.; Zheng, X.; Yikui, M.O. Establishment of density zoning and determination of plot ratio along rail transit line based on TOD: A case study on rail transit line 3 in Shenzhen. *City Plan. Rev.* **2011**, *4*, 30–35.
29. Zhang, W.; Long, C.P. Determination of plot ratio surrounding rail transit station—A case of YuanYang station on Chongqing light rail transit. *Urban Mass Transit.* **2017**, *01*, 91–95.

30. Tang, X.L.; Huang, Z.D.; Lin, M. The characteristics of urban rail transit and its relationship with land use in Shenzhen. *Urban Rural Plan.* **2018**, *03*, 98–104.
31. Chang, L.Y.; Lin, D.J. Analysis of international air passenger flows between two countries in the APEC region using non-parametric regression tree models. *Lect. Notes Eng. Comput. Sci.* **2010**, *2180*, 415–420.
32. Su, H.L.; Wang, Y.; Wang, X.; Zhou, R. Transit-oriented development correlates of rail transit ridership in shanghai municipality. *J. Tongji Univ.* **2014**, *42*, 71–77.
33. Yuankun, L.; Xiafei, Y. The relationship between the scale of urban rail transit network and population and post density. *Urban Rail Transit Res.* **2017**, *20*, 10–14.
34. Gu, J.C. Relationship between ridership characteristics at railway stations and surrounding land-use in Shenzhen. *Traffic Transp.* **2019**, *35*, 47–50.
35. Duan, Y.; Lei, K.; Tong, H.; Li, B.; Hou, Q.H. Land use characteristics of Xi'an residential blocks based on pedestrian traffic system. *Alex. Eng. J.* **2021**, *60*, 15–24. [CrossRef]
36. Moniruzzaman, M.; Páez, A. Accessibility to transit, by transit, and mode share: Application of a logistic model with spatial filters. *J. Transp. Geogr.* **2012**, *24*, 198–205. [CrossRef]
37. Loo, B.P.; Chen, C.; Chan, E.T. Rail-based transit-oriented development: Lessons from New York City and Hong Kong. *Land Sc. Urban Plan.* **2010**, *97*, 202–212. [CrossRef]
38. Cervero, R. Alternative approaches to modeling the travel-demand impacts of smart growth. *J. Am. Plan. Assoc.* **2006**, *72*, 285–295. [CrossRef]
39. Wei, Z.Y.; Su, H.M.; Huang, R.J. The influence of the construction of Xi'an Metro Line 2 on land use change. *Chin. J. Ecol.* **2018**, *37*, 2474–2482.
40. Kim, H.; Nam, J. The size of the station influence area in Seoul, Korea: Based on the survey of users of seven stations. *Int. J. Urban Sci.* **2013**, *17*, 331–349. [CrossRef]
41. Jin, D.; Meng, X. Characteristics of subway station ridership with surrounding land use: A case study in Beijing. In Proceedings of the 2015 International Conference on Transportation Information and Safety (ICTIS), Wuhan, China, 25–28 June 2015; pp. 330–336.
42. *GB/T 50137-2011*; Code for Classification of Urban Land Use and Planning Standards of Development Land. Ministry of Housing and Urban-Rural Development of the People's Republic of China: Beijing, China, 2012.
43. Tang, T. Development control, planning incentive and urban redevelopment: Evaluation of a two-tier plot ratio system in Hong Kong. *Land Use Policy* **1999**, *16*, 33–43. [CrossRef]
44. Feng, H.; Zhang, S. Study on travel efficiency and land use mixing degree based on meta fractal dimension theory. *J. Xi'an Univ. Archit. Technol.* **2014**, *46*, 882–887.
45. Ye, X.F.; Ming, R.L.; Li, R.X. Rail transit in Tokyo and Seoul: Development trends & characteristics of passenger flow. *Urban Transp. China* **2008**, *6*, 16–20.
46. Jin, Y. The influence of l and use on diurnal pattern of urban mass transit station boardings: A case study of shanghai city. *Mod. Urban Res.* **2015**, *6*, 13–19.
47. Li, J.B.; Cui, Y. Research on relationship between passenger flow and attraction of surrounding land on Tianjin metro line 1 based on cell phone signaling data. *Urban Passeng. Transp.* **2019**, *6*, 50–56.
48. Xia, F.S.; Chen, H.; Yuan, J.M.; Zhang, K.; Zhang, K. Research on automatic detection method of building in aerial stereo image based on elevation constraints. *Geogr. Geo-Inf. Sci.* **2019**, *35*, 46–50.
49. Chen, W.G. The Discussion of Rational Exploitation Strength in the Plot around the Metro Station: Starting with the Detailed Plan of Rail Transport in Second Phase in Shenzhen. *Mod. Urban Res.* **2006**, *08*, 44–50.
50. Yao, J.T.; Dong, Z.P. Correlativity of random variables and structural reliability. *J. Xi'an Univ. Archit. Technol.* **2000**, *32*, 114–118.
51. Cui, Y.C.; Guan, H.Z.; Yang, S.I.; Qin, Z.T. Residents' travel characteristics based on order data of on-line car-hailing: A case study of Beijing. *Transp. Res.* **2018**, *4*, 20–28.
52. Zhuang, C.T.; Wu, C.C.; Gu, T.Q.; Li, J. Analysis on characteristics of rail transit and conventional bus transfer in Suzhou. *Traffic Transp.* **2021**, *37*, 51–55.
53. Zheng, W.H. Discussion on land use layout of residential rail transit station areas. *Planners* **2009**, *025*, 58–62.
54. Cui, X.; Yu, B.J.; Liang, P.P.; Wang, L.; Zhang, L.F. Urban rail transit station type identification and job-housing spatial rebalancing based on passenger flow and land use: Taking Chengdu as an example. *Mod. Urban Res.* **2021**, *36*, 68–79.
55. Kittelson, Associates Inc. *TRB's Transit Cooperative Research Program (TCRP) Report 100: Transit Capacity and Quality of Service Manual*, 2nd ed.; Transportation Research Board | National Academies: Washington, DC, USA, 2003.
56. *GB/T 50157-2013*; Metro Design Specification. China Architecture and Building Press: Beijing, China, 2013.