mathematics

Article

Analyzing the Passenger Flow of Urban Rail Transit Stations by Using Entropy Weight-Grey Correlation Model: A Case Study of Shanghai in China

Pei Yin [1], Jing Cheng [2,*] and Miaojuan Peng [1]

1 Department of Civil Engineering, School of Mechanics and Engineering Science, Shanghai University, Shanghai 200444, China
2 Department of Construction Management and Real Estate, College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518057, China
* Correspondence: jingcheng7-c@my.cityu.edu.hk or jingcheng@szu.edu.cn

Abstract: In this paper, the factors influencing the passenger flow of rail transit stations in Shanghai of China are studied by using the entropy weight-grey correlation model. The model assumptions and the corresponding variables are proposed, including traffic accessibility, built environment, regional characteristics of the district to which the rail transit station belongs, conditions of the station and spatial location, which affect the passenger flow of rail transit stations. Based on the assumptions and the variables, the entropy weight-grey correlation model for analyzing the passenger flow of urban rail transit stations is presented. By collecting the data of passenger flow of rail transit stations and corresponding influencing factors in Shanghai, the results of the entropy weight-grey correlation model are obtained. It is shown that the influencing factors, such as the distances from the rail transit station to the adjacent third-class hospital and the adjacent large commercial plazas, district committees, parking areas and the transaction price of important plots, and the gross output value of the tertiary industry, have significant impacts on the passenger flow of a subway station. Finally, some suggestions are proposed for the local governments to formulate improved policies for rail transit development. The conclusions can provide a reference for the development of rail transit in other large cities and countries.

Keywords: rail transit; station passenger flow; influencing factor; mathematical model; entropy weight-grey correlation model

MSC: 62H20

Citation: Yin, P.; Cheng, J.; Peng, M. Analyzing the Passenger Flow of Urban Rail Transit Stations by Using Entropy Weight-Grey Correlation Model: A Case Study of Shanghai in China. *Mathematics* **2022**, *10*, 3506. https://doi.org/10.3390/math10193506

Academic Editors: Roman Ivanovich Parovik, Junzo Watada and Sy-Ming Guu

Received: 18 July 2022
Accepted: 20 September 2022
Published: 26 September 2022

1. Introduction

Rail transit systems can save energy, reduce noise, lower carbon dioxide emissions and ease urban congestion, which can promote the sustainable development of large cities, and gradually become the core of urban public transit systems. Rail transit can be substituted effectively for cars when rail usage is supported by land development and other policies [1]. With the continuous expansion of the scale of rail transit networks in large and medium-sized cities, rail transit operations are faced with problems such as early warning of large passenger flows and improvement of operating arrangements. Shanghai, as one of the most prosperous megacities in China, had a subway operating mileage of 831 km at the end of December 2021. Figure 1 shows the distribution of the rail transit network in Shanghai in 2015. The Shanghai rail transit system has greatly relieved the pressure on the urban environment and passenger flow. Thus, it is crucial to investigate the factors that influence urban passenger flow in order for the Shanghai government to plan site selection and layout, optimize subway lines, and formulate more complete policies of rail transit development. This study can also provide a reference for the development of rail transit in other emerging large cities and countries.

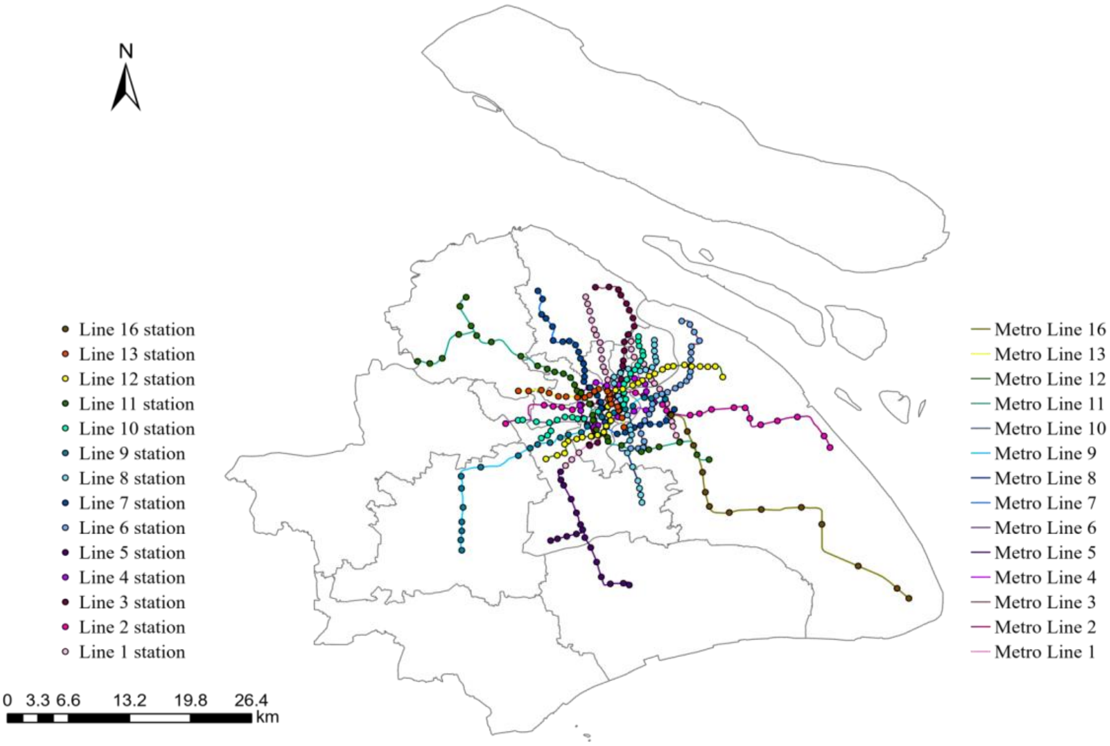

Figure 1. Rail transit network in Shanghai.

The conditions of the rail station and built environment around the station are regarded as the main factors affecting the passenger flow of the station [2,3]. The conditions of the station include the external connection between subway station and highway network, as well as the combined transportation with other modes of transportation. The development of stations also has a great impact on the attraction of rail transit passenger flow. If a station can be developed well, it can attract many passengers. The built environment is a significant factor that changes residents' travel behavior, and it is also a fundamental factor that affects the passenger flow of subways [4]. The built environment mainly embodies the principles of the three dimensions (3Ds) of building density, variety and design. Improvements in land accessibility around rail transit can promote activities of regional passenger flow and attract the passenger flow to nearby rail transit corridors [5]. In addition, the development and construction of the surrounding areas of rail transit is an important factor for the adjustment of passenger flow trends, which can also cause great changes in the intensity and nature of surrounding land development. Therefore, it is necessary to further explore the impact of land conditions around the station on its passenger flow. Some researchers found that the establishment of urban rail transit may bring 1/3 of the value added effect of the total rail investment to the nearby real estate [6–8]. Residential prices near subway stations may rise 32.1% after a light rail system is built [9]. The distance of the rail transit station has a significant impact on the nearby real estate prices, and with increase of the distance from a rail transit station, residential values near the rail transit decline [10]. Rail transit can promote the development and utilization of surrounding land, further accelerating and expanding urbanization [11,12].

Traffic accessibility is one of the aspects of subway stations affecting residential prices [13]. The development of urban rail transit consumes a large amount of government resource reserves. Much capital investment is needed during the planning, construction and operation of rail transit. It generally requires a long time to construct rail transit, and the operating income after completion is low. Thus, there is great pressure for the government to recover the construction costs. The main sources of funds to construct rail transit are the government investment, land transfer revenues, and loans based on government credit. The build-transfer (B-T) financing mode increases costs, which conceals the real debt situation of local governments [14]. In order to solve the problem of insufficient funds for the construction of the rail transit system, the land value-added along the subway line can be used to promote the construction of rail transit, which combines the capital demand with the promotion of rail transit construction, and is consistent with the concept of transit-oriented development (TOD). Especially in the urban areas that have not yet been developed, the government can acquire the land in the planned development area at a lower price in advance and introduce it into public transportation to gain a time advantage as regards the development of land. Then the government can sell the 'ripe land' with complete infrastructure, and will recover the investment in public transportation from the return of land appreciation to recoup the cost of the investment. Analyzing the influencing factors of passenger flow can predict the passenger flow status of newly developed rail transit stations, and provide policy suggestions for the land use around the station, leading to an optimal development strategy and design of the station.

Different mathematical models have been presented to study the factors influencing the passenger flow of rail transit stations in metropolitan areas. Sohn et al. applied multiple linear regression (MLR) to investigate the influencing factors of subway passenger flow demand in the metropolitan area of Seoul in Korea [2]. Blainey used the geographically weighted regression (GWR) method to predict the use of the Ebbw Vale branch line and the factors influencing passenger flow in South Wales of England [15]. Gutiérre et al. used a distance-decay weighted regression model to explore the influencing factors of the subway passenger flow in Madrid, Spain [16]. The ordinary least squares (OLS) method has been applied widely to analyze the influencing factors [17–19]. Sung and Oh used the OLS global spatial regression method to discuss the factors influencing passenger flow on weekdays, rest days and in the morning rush hour in Seoul, Korea [20]. Cardozo applied OLS and GWR methods to estimate the passenger flow of the Madrid subway station of Span, and compared the results obtained by these two methods [21]. Zhao et al. and Sun et al. proposed the direct passenger flow prediction model to investigate the factors influencing the passenger flow of subways in Nanjing and Beijing of China, respectively [3,22]. Sung et al. used the OLS global spatial regression method to study the impact of land use, railway service coverage and railway station accessibility on the passenger flow of rail transit stations in Seoul [23]. Jun et al. applied stepwise regression method and mixed geographically weighted regression (MGWR) method to investigate the factors influencing the passenger flow of subways in Seoul [24].

From the previous studies, it can be seen that most researchers have used the global linear regression method to analyze the index factors. The coefficients of the classical linear regression model (CLRM) are constant. Thus, the most of these studies are based on the premise of homogeneity, and spatial heterogeneity in the real environment is often ignored in the analysis process. Therefore, besides the built environment and the station's own conditions, the impacts of traffic accessibility and spatial location factors on the passenger flow of the station are considered in this paper to make up for the difference caused by the uneven spatial distribution of the passenger flow of the station. Moreover, the maturity level of the urban economy and infrastructure differs, and even within the same city, the policies, economic and regional conditions between districts and counties are vastly different. Besides considering the characteristics of land within the attraction range of the station, it is also necessary to analyze the regional characteristic of the district to which the station belongs.

The MLR model for analyzing the impact of passenger flow in rail transit considering spatial heterogeneity requires that sample data should satisfy typical probability distribution. In addition, if using a linear relationship between the characteristic and the factors, the factors should be independent. In particular, the current domestic statistical data is limited, while the grey-scale of data is relatively large, and the data have irregular discrete distributions. The related methods, such as dynamic modeling, system dynamics and discrete-event and agent-based modeling, are very good for multi-factor analysis, but they are difficult to combine with the traditional theory of degrees of grey correlation. The grey correlation model is an important method to quantitatively analyze the dynamic development process, examine whether the relationship between the various factors is close and identify the factors that affect the developments. Under the condition of fuzzy initial information, the decision-making methods of multi-factor analysis, which are based on measurement theory, mainly include TOPSIS, VIKOR and grey correlation analysis methods [25,26]. Grey correlation analysis has been successfully applied to multi-factor decision-making problems, such as interval number, linguistic information, intuitionistic fuzzy information, interval intuitionistic fuzzy information and hesitant fuzzy information [27–29]. This paper proposes the grey correlation model in grey system theory to analyze the factors influencing the passenger flow of Shanghai rail transit stations in 2015. The grey correlation model makes up for the shortcomings of using mathematical statistics methods for systematic analysis. The calculation process is efficient and simple, and quantitative results can be consistent with the results of qualitative analysis. Because of the disadvantage of low accuracy in the grey correlation model, the entropy is considered in the weight calculation to improve the accuracy of the solution.

Taking Shanghai as an example, this paper investigates the factors that affect the passenger flow of rail transit stations and how the passenger flow of rail transit stations is affected by the changes of these factors. The assumptions of factors affecting the passenger flow in Shanghai and the corresponding variables in the mathematical model are proposed, and the entropy weight-grey correlation model is used. Based on the grey correlation degree, it is shown that the traffic accessibility, built environment, regional characteristics of the district to which the rail transit station belongs, the conditions of the station and the spatial location will affect the passenger flow of the station. Finally, some policy suggestions for planning rail transit stations are proposed.

2. Variables in the Mathematical Model

To present the mathematical model of the factors influencing the passenger flow of rail transit stations, the factors including traffic accessibility, built environment, regional characteristics of the district to which the rail transit station belongs, station conditions and spatial location are proposed.

2.1. Traffic Accessibility of the Rail Tansit Station

The distances from the i-th rail transit station in Shanghai to the important transportation hub station (D_i), the district committee (DDC_i), the nearest large commercial plaza (DCP_i), the nearest university or key high school ($DUNI_i$), the nearest third-class hospital (DTH_i) and the nearest park ($DPAR_i$) are considered as traffic accessibility indicators in this paper. Lots of important transportation hubs, such as Shanghai Hongqiao Transportation Hub, are built in Shanghai [30]. Shanghai Hongqiao International Airport and Pudong International Airport are two of the three major gateway complex hubs in China. Shanghai Railway Station is the second largest railway station in Shanghai. This station, Shanghai Hongqiao Railway Station and Shanghai South Railway Station together constitute the largest railway hub in the eastern coastal area of China. In this study, Shanghai Hongqiao Railway Station, Shanghai Railway Station, Hongqiao International Airport and Pudong International Airport are considered as the important transportation hubs affecting the passenger flow of the rail transit stations.

The district committee is an organization that has overall leadership over the political, economic, cultural and social development of the district, and is a vital organization related to the policy of local rail transit development [31]. The distances of the rail transit station to the adjacent large commercial plazas, universities or key high schools, third-class hospitals and parks reflect the difficulty of taking the subway to the nearest core economic, education, medical and green area, which is an important indicator to measure the service capacity of the public transportation system. Figure 2 shows the distribution of traffic accessibility indicators in Shanghai.

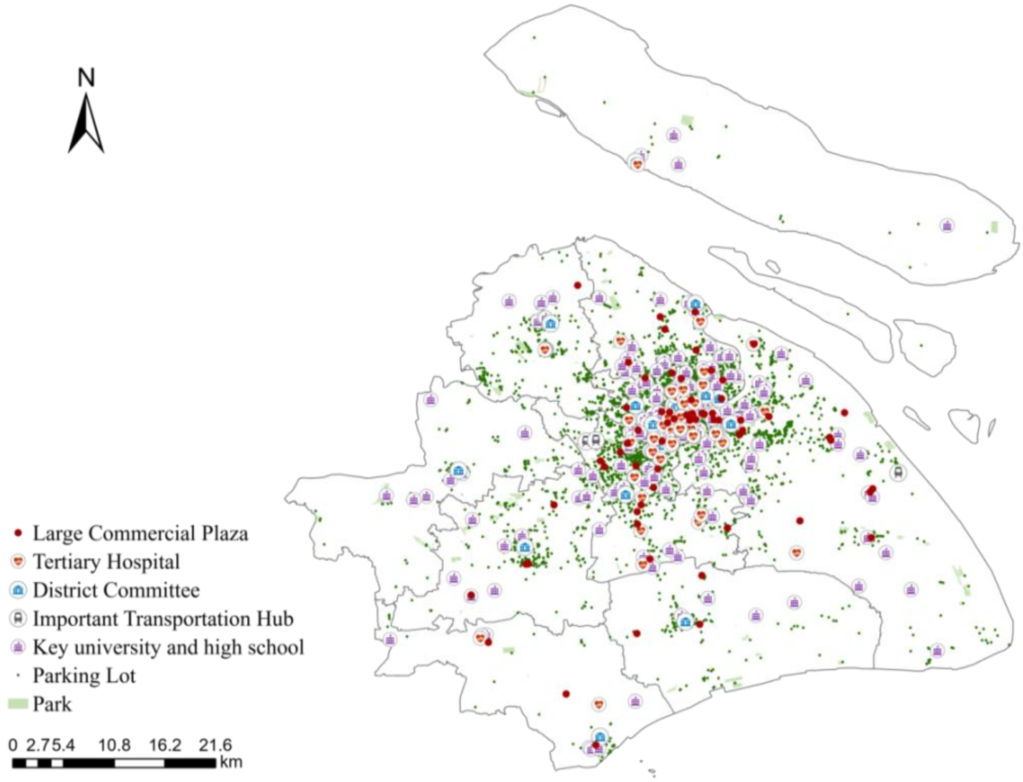

Figure 2. Distribution of traffic accessibility indicators in Shanghai.

2.2. Built Environment around the Rail Transit Station

The built environment around the rail transit station includes building density, diversity and design [32]. Indices, such as floor area ratio (IPR_i) and area (IPA_i) of important land within the attraction range of the *i*-th station, reflect the density. Four types of land use, including residential land, commercial or office land, industrial land and comprehensive land, are considered in this paper. Pedestrian friendliness around the station reflects the rational design of the station, and it is the average time spent by residents of nearby land walking to the station (IPT_i). The average travel time of residents is determined by the size of the area where passengers are located and the complexity of the road network. The better the road environment is, the smaller the personal cost to residents in travel time and distance. Figure 3 shows the distribution of important land within the attraction range of the station. Investment in urban rail transit projects can promote land appreciation, save travel time, and improve travel comfort [33]. Therefore, in this paper, the land transfer price (IPP_i) is also regarded as a variable in the built environment factors.

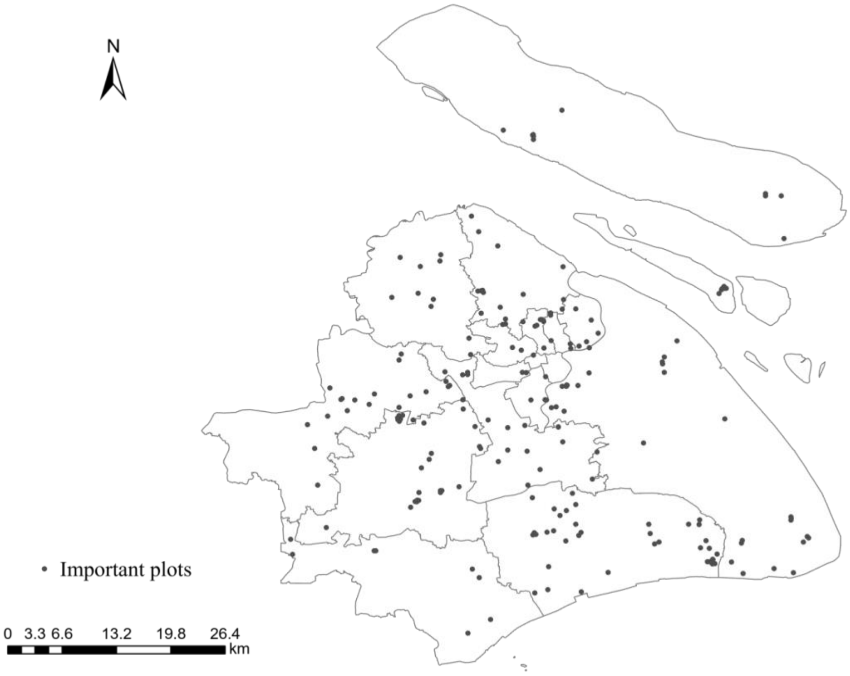

Figure 3. Distribution of important land around rail transit stations in Shanghai.

2.3. Regional Characteristics of the District to Which the i-th Station Belongs

The regional characteristic of the district to which the *i*-th station belongs include the land, economic and political characteristics of the district. The transportation system is very important to the industrial and commercial development of a city, which will affect land prices and house prices [34]. Correspondingly, the land grade (DIC_i), land area (DAA_i), number of permanent residents ($DPRN_i$), area of housing with more than eight floors (DHA_i), GDP ($DGDP_i$), average salary of employees ($DEAS_i$), the industrial output value ($DIEO_i$), secondary industry output value ($DSIO_i$), tertiary industry output value ($DTIO_i$) of the district, etc. will also affect the expansion and development of the transportation system, which in turn will impact the station passenger flow. In addition, government policies and intentions can have significant impacts on rail transit development, and the political promotion of local officials can affect their strategic decisions in rail transit [35], which will indirectly affect the changes in the passenger flow and its distribution of subway stations. Therefore, the political factors, including whether the district party secretary is re-elected ($DDSR_i$), whether the district head is re-elected ($DDMR_i$), and whether the district head is promoted to the district party secretary ($WMPS_i$), of the district where the *i*-th site belongs are considered in this paper.

2.4. Conditions of the Station

The contradiction between the limited urban land resources and the increasing traffic demand is becoming more and more large, and stations should be arranged and designed more rationally and scientifically to avoid waste of resources. For the factors influencing the conditions of the station, this paper focuses on the passenger service and evacuation capabilities of the station site [36]. Therefore, reasonable transfer stations and entrance and exit passages should be set up to avoid the phenomenon of passengers being stranded and crowded due to peak passenger flow. A transfer station is formed by the intersection of two or more rail transit lines. It is equipped with transfer facilities such as transfer passages and transfer station halls, and indicates the routes for passengers to transfer. The number of entrances and exits is determined by the topography of the station and the passenger flow during peak hours. Correspondingly, the number of entrances and exits determines the evacuation capacity of the station site, which has an obvious impact on the passenger flow of the station. Under normal circumstances, the number of entrances and exits of shallow-buried stations should not be less than 4, and the number of entrances and exits of small stations can be reduced according to the actual situation, but should not be less than 2 [37]. In this paper, whether the *i*-th station is a transfer station (WTS_i), the number of rail transit lines passing through the station ($NUML_i$), and the number of entrances and exits ($NUME_i$) are considered as factors.

2.5. Spatial Location of the Station

The spatial location factor considers the geographical location of the station, which is whether the *i*-th subway station is situated within the inner ring (WIR_i), middle of the middle and inner rings ($WIMR_i$), middle of the middle and outer rings ($WMOR_i$), middle of outer and suburban rings ($WOSR_i$) and outside suburban rings ($WOSR_i$) of Shanghai. The inner ring area, as the most central area in Shanghai, is surrounded by 47.7 km of elevated roads and covers an area of about 120 square km. A total of 89 rail transit stations are located in the inner ring area [38]. There are 3 dual-rail subway stations, which are Gucun Park Station, Gangcheng Road Station, and Yuqiao Station, in the area between the central and outer rings. The development of dual-rail transit has an effect on the surrounding real estate prices [39]. The rail transit network connects various areas between cities. Figure 4 shows the distribution of central districts, suburban districts, and a county in Shanghai. Figure 5 shows the inner, middle, outer and suburban rings in Shanghai.

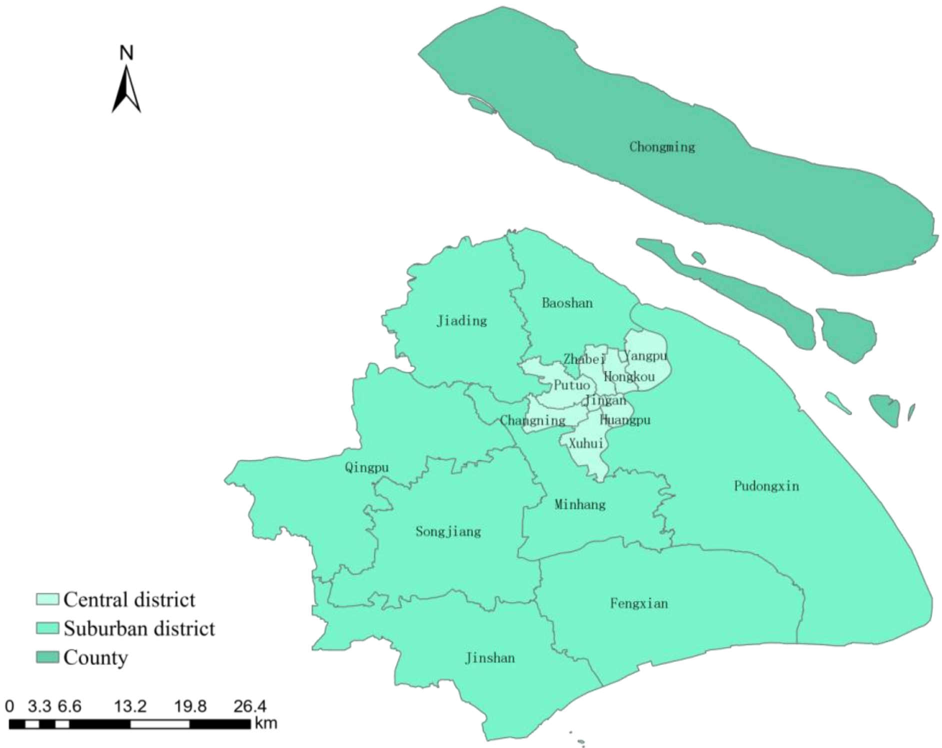

Figure 4. Distribution of central districts, suburban districts and a county in Shanghai.

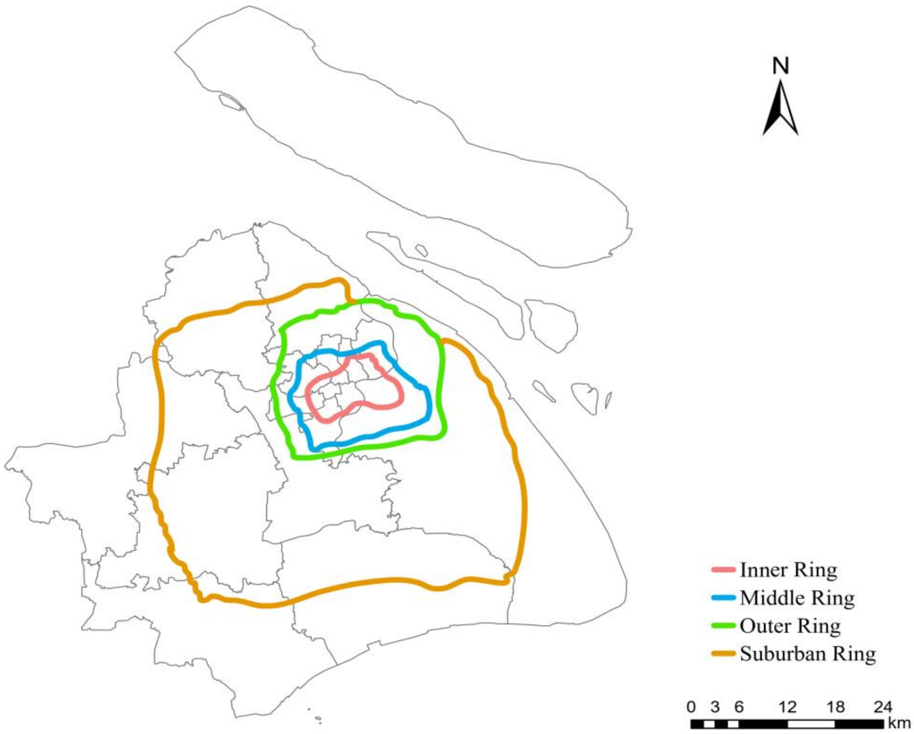

Figure 5. Inner, middle, outer and suburban rings in Shanghai.

3. Mathematical Model of Passenger Flow at Shanghai Rail Transit Stations

Combining the grey correlation model and entropy theory, this paper presents a mathematical model to analyze the factors affecting the passenger flow of urban rail transit stations. The grey correlation model is an important method to quantitatively analyze the dynamic development process and examine whether the relationship between the various factors is close. It is a measure of the relation degree between sequences, which is expressed as the similarity of magnitude changes and the development trends between sequences. The correlation degree is a reflection of the distance between two points of different sequences. Based on the reliability of the correlation coefficient, the entropy weight-grey correlation is established by weighting the grey correlation coefficient, which can control the volatility of the correlation coefficient effectively. Therefore, the entropy weight-grey correlation model can be used to analyze the influencing factors of urban passenger flow of rail transit stations directly and clearly.

The entropy weight coefficient method is used to analyze the subjective data. Based on the change of the value of the index, the weight coefficient in the model can be obtained from the utility value reflection of the information entropy of the data. The entropy represents the degree of disorder of the system according to information theory. The lower the entropy value of information, the lower the degree of disorder, and the higher the utility value. Contrarily, the increase of the entropy value of information will lead to the increase of the degree of disorder and the decrease of the utility value [40]. Based on the order degree and utility value of the information, the judgment matrix and information entropy are used to calculate the weight. This method eliminates the influence of human factors on the weight calculation, and makes up for the weaker aspects of the grey correlation model. The model is not easily affected by subjective factors when there are many indicators, and this can makes the evaluation results more realistic [41].

The factors influencing the passenger flow of Shanghai rail transit stations are analyzed in this paper, and the annual passenger flow of each station in Shanghai in 2015 is selected as the reference sequence. Five types of influencing factors, including traffic accessibility, built environment, regional characteristics of the district to which the *i*-th station belongs, the conditions of the station, and spatial location, of passenger flow of rail transit stations are used as a comparison sequence, which includes a total of 31 index factor of variables.

Suppose the reference sequence is Y, $Y = [Y^1, Y^2, \cdots, Y^m]$, the comparison sequence is X_k^i, $X_k^i = [X_k^1, X_k^2, \cdots, X_k^m]$, where m represents the *m-th* rail transit station, i denotes the i-th station, $i = 1, 2, \cdots, m$; n indicates that there are a total of n index factors, k ($k = 1, 2, \cdots, n$) is the k-th index factor.

The optimal index set F^* is

$$F^* = [X_1^*, X_2^*, \cdots, X_n^*] \tag{1}$$

where X_k^* is the ideal value of the k-th index.

For a certain type of index, the larger its value, and the more advantages it reflects, then the maximum value of this index in all stations should be taken; this type of index is called a positive index. Conversely, the smaller the value, the more fully the advantage can be exerted. The minimum value of this index in all stations should be selected, and this type of index is called a reverse index.

Combining with the reference and comparison sequences, the initial matrix is

$$M = \begin{bmatrix} Y^* & X_1^* & X_2^* & \cdots & X_n^* \\ Y^1 & X_1^1 & X_2^1 & \cdots & X_n^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Y^m & X_1^m & X_2^m & \cdots & X_n^m \end{bmatrix} \tag{2}$$

Suppose the change interval of the passenger flow Y value is $[\min Y, \max Y]$, where $\min Y$ is the minimum value of the passenger flow of Shanghai rail transit stations in 2015, and $\max Y$ is the maximum value among those of all stations. The variation interval of the k-th indicator is $\left[\min\limits_i X_{ik}, \max\limits_i X_{ik}\right]$, $\min\limits_i X_{ik}$ is the minimum value of the k-th index in all stations, and $\max\limits_i X_{ik}$ is the maximum value of the k-th index in all stations.

The normative formula is

$$V_y = \frac{y - \min y}{\max y - \min y} \tag{3}$$

$$V_{ik} = \frac{X_{ik} - \min\limits_i X_{ik}}{\max\limits_i X_{ik} - \min\limits_i X_{ik}} \tag{4}$$

After normalizing the original matrix, the matrix converted from M to V is obtained as

$$V = \begin{bmatrix} V_y & V_1^* & V_2^* & \cdots & V_n^* \\ V_y & V_1^1 & V_2^1 & \cdots & V_n^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ V_y & V_1^m & V_2^m & \cdots & V_n^m \end{bmatrix} \tag{5}$$

The correlation analysis method is used to obtain the correlation coefficient $\xi_i(k)$ between the passenger flow of the rail transit station and the optimal value of the k-th index in the i-th station, and

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k \left| V_y - V_k^i \right| + \max\limits_i \max\limits_k \left| V_y - V_k^i \right|}{\left| V_y - V_k^i \right| + \rho \max\limits_i \max\limits_k \left| V_y - V_k^i \right|} \tag{6}$$

where $\rho \in [0,1]$, and in general, $\rho = 0.5$. In the formula, $\left|Vy - V_k^i\right|$ is the absolute deviation between $\{Vy\}$ and $\{V_k^i\}$ of the k-th index at the i-th station. $\underset{i}{\min}\underset{k}{\min}\left|V_{y^-}, V_k^i\right|$ is the two-level minimum deviation, and $\underset{i}{\max}\underset{k}{\max}\left|V_{y^-}, V_k^i\right|$ is the two-level maximum devitation.

To understand these variables more easily, we need to introduce the concept of range. The range is the deviation between the largest one and the smallest one in a set of data. The range can reflect the distribution and dispersion range of data. The deviation between the standard values of any two units cannot exceed the range. The larger the range, the greater the degree of dispersion.

According to the definition of entropy, an analysis matrix $X = [x_{ik}]_{m\times n}$ $(i = 1, 2, \cdots, m;$ $k = 1, 2, \cdots, n)$ is established by a data table consisting of m stations and n items of influencing factors. The entropy of each index is defined as influencing factors.

The entropy of each index is defined as

$$H_k = -\frac{\sum\limits_{i=1}^{m} f_{ik} \ln f_{ik}}{\ln m} \tag{7}$$

where $f_{ik} = \frac{x_{ik}}{\sum\limits_{i=1}^{m} x_{ik}}$. When $f_{ik} = 0$, $\ln f_{ik}$ is meaningless, then let $f_{ik} \ln f_{ik} = 0$.

Further, the weight of the k-th index can be calculated as

$$w_k = \frac{1 - H_k}{\sum\limits_{k=1}^{n} (1 - H_k)} \tag{8}$$

where $0 < w_k < 1$, and $\sum\limits_{i=1}^{n} w_k = 1$.

From $\xi_i(k)$, the analysis matrix E of each index can be obtained as

$$E = \begin{bmatrix} \xi_1(1) & \xi_1(2) & \cdots & \xi_1(n) \\ \xi_2(1) & \xi_2(2) & \cdots & \xi_2(n) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_m(1) & \xi_m(2) & \cdots & \xi_m(n) \end{bmatrix} \tag{9}$$

The determined weights can be expressed as $w = [w_1, w_2, \cdots, w_n]^T$. Then the comprehensive evaluation result is

$$r_i = \sum_{k=1}^{n} w(k)\xi_i(k) \tag{10}$$

4. Data Collection

This paper collects data from 85 rail transit stations in Shanghai in 2015. The annual passenger flow data of each rail transit station is obtained by converting the monthly average card swiping data of each station, which was provided by Shanghai Shentong Metro Co., Ltd (Shanghai, China).

The data of GDP, average house price, paid-in foreign investment, fixed asset investment, total industrial output value, total industrial assets, secondary and tertiary industry output value of 19 districts in 2015, and important land, land area, land use type, floor area ratio, residential area and other data within the attraction range of the site are collected from the 'Land Planning Database' [42], based on the statistical yearbooks of the city and the districts of the city and the official website of the Municipal Bureau of Planning and Natural Resources. The Land Planning Database is a relational data management system that stores and manages all district-level data in big cities such as Shanghai. The map platform in this database was used to determine the spatial geographic location of each rail transit station in Shanghai.

For the data on traffic accessibility characteristics, this paper uses the network analysis provided by ArcGIS to calculate the distances from each rail transit station to important transportation hubs, district committees and the nearest large commercial plazas, universities or key high schools, third-class hospitals, parks and important land. The average walking speed of residents is assumed as 1.2 m/s [43], and thus the approximate time cost of residents walking from nearby plots to rail transit stations is calculated. Figure 6 shows the heat map of passenger flow distribution at various rail transit stations in Shanghai.

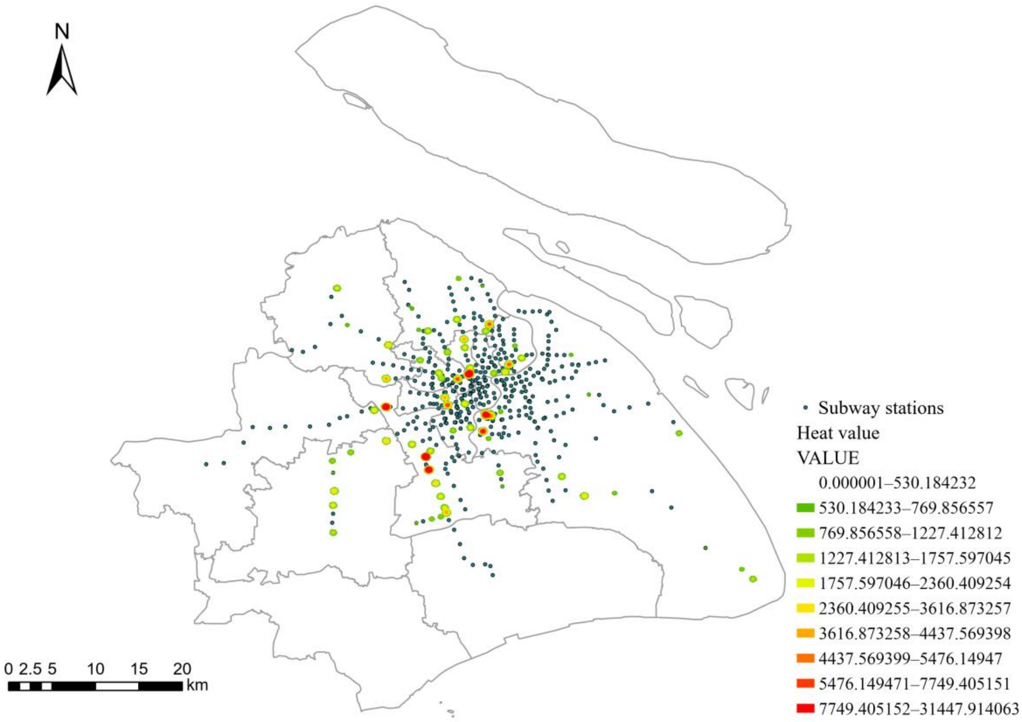

Figure 6. Heat map of passenger flow distribution of rail transit stations in Shanghai.

Google Earth (GE) is the virtual software of the earth developed by Google, and the global geomorphological imagery provides high-precision images with a resolution of about 1 m and 0.5 m for large cities and building areas, and the heights of viewing angle are about 500 m and 350 m, respectively. In this paper, GE high-resolution image software is used to obtain the data on the conditions of the rail transit stations. The high-resolution image with a spatial resolution of 0.5 m is selected, and through visual interpretation, the number of rail lines and the number of entry and exit channels for by each rail transit station can be identified. Table 1 summarizes the statistics of the passenger flow and its influencing factors.

Table 1. Summary statistics of passenger flow of stations and their influencing factors.

Variable	Observation	Mean	Standard Deviation	Minimum	Maximum
APS_i	211	705.6567	672.5237	63.2235	4679.2010
D_{1i}	211	233.5913	135.8824	4.7150	595.4320
D_{2i}	211	238.4747	153.6188	6.7170	686.0370
D_{3i}	211	233.8802	148.8693	22.1220	672.5820
D_{4i}	211	417.7905	147.1100	6.0840	737.9900
$DPAR_i$	211	13.9371	10.1799	0.9610	72.3710

Table 1. *Cont.*

Variable	Observation	Mean	Standard Deviation	Minimum	Maximum
$DUNI_i$	211	32.1301	22.7869	1.3720	99.2150
DCP_i	211	48.9716	47.5769	3.4030	246.6420
DDC_i	211	109.8904	108.9377	0.2160	451.4720
DPL_i	211	7.3663	7.2217	0.4330	34.1900
DTH_i	211	80.4863	75.0101	1.5400	302.5380
IPA_i	211	47,246.8800	38,565.9500	1658.0000	260,118.8000
IPR_i	211	1.9836	1.0436	0.2200	10.0000
IPP_i	211	79,180.5500	140,362.2000	186.0000	881,500.0000
IPT_i	211	104.6643	112.3101	1.7479	467.9843
DIC_i	211	5.7109	2.7493	1.0000	9.0000
DAA_i	211	625.2144	370.4273	23.4800	1210.4100
$DPRN_i$	211	196.2970	149.1328	69.1100	547.4900
DHA_i	211	2769.6540	2790.2910	135.0000	9233.0000
$DGDP_i$	202	2002.3940	2457.6720	291.2000	7898.3500
$DEAS_i$	177	80,474.4900	23,749.0800	60,285.0000	129,368.0000
$DIEO_i$	211	2926.2760	2784.0890	104.8800	9177.8000
$DSIO_i$	202	722.2753	646.8724	71.4100	2186.5200
$DTIO_i$	202	1228.5050	1844.4640	137.6000	5684.9100
$DDMR_i$	211	0.5687	0.4964	0.0000	1.0000
$DDSR_i$	211	0.4550	0.4992	0.0000	1.0000
$WMPS_i$	211	0.3649	0.4826	0.0000	1.0000
WTS_i	211	0.1327	0.3401	0.0000	1.0000
$NUML_i$	211	1.1801	0.5655	1.0000	5.0000
$NUME_i$	211	3.1138	1.4331	1.0000	8.0000
WIR_i	211	0.0427	0.2026	0.0000	1.0000
$WIMR_i$	211	0.1327	0.3401	0.0000	1.0000
$WMOR_i$	211	0.0900	0.2869	0.0000	1.0000
$WOSR_i$	211	0.5355	0.4999	0.0000	1.0000
WSR_i	211	0.1991	0.4002	0.0000	1.0000

At present, the domestic statistical data is limited and the grey-scale of data is relatively large. Among the statistical data, the panel data of Shanghai rail transit stations in 2015 is very complete, and the data information of the same stations in other years has many missing values. Besides, the theory of grey correlation degree can make up for the problem of large grey-scale of data, and will not cause distortion of the analysis results due to the smaller amount of data.

5. Results and Discussion

Based on the collected data and the mathematical model, the results of grey correlation between the passenger flow and the five influencing factors are calculated. The normalized processing of passenger flow and five types of influencing factors and the average value of their correlation coefficients are shown in Table 2. The entropy weight of influencing factors and the average value of comprehensive correlation degree are shown in Table 3. In this paper, the range method is applied to deal with the forward index and the reverse index, respectively. This method can remove the units involved in the physical quantity equation to simplify the calculation. Meanwhile, the indexes in the normalized matrix are all positive indexes. The maximum value of each column in the obtained normalized matrix is 1, and the minimum value is 0. There is no need to distinguish the positive and negative index factors in the subsequent correlation results. The calculated correlation degree should range from 0 to 1. The larger the value, the stronger the correlation between the index sequence and the reference sequence. If the result is less than 0, it does not conform to the calculation principle of grey correlation analysis, and should be used as 'Null'. In addition, the higher the weight, the greater the entropy, which means that the corresponding index system is more chaotic, the degree of variation is smaller, and the value information carried is less; on the contrary, the smaller the entropy weight, the more orderly the system is, and the more value information it carries. As shown in Table 3, most of the entropy weights in this study fluctuate between 0.025 and 0.3, indicating that the data of the entire system is relatively orderly, and the analysis of the results based on the collected data is more scientific and rigorous.

Table 2. Standardization and average correlation coefficients between passenger flow of stations and five types of influencing factors.

Variable	Mean of Normalized Values	Mean of Correlation Coefficient
APS_i	0.139175995	-
D_{1i}	0.612545015	0.538013464
D_{2i}	0.658838724	0.516097799
D_{3i}	0.674448592	0.509567495
D_{4i}	0.437487242	0.651135896
$DPAR_i$	0.818286843	0.436537605
$DUNI_i$	0.685637774	0.501344371
DCP_i	0.811458170	0.443615288
DDC_i	0.756957541	0.475163430
DPL_i	0.794611798	0.451009159
DTH_i	0.737718280	0.484190583
IPA_i	0.176386037	0.791569735
IPR_i	0.819669979	0.431951516
IPP_i	0.089632691	0.837244805
IPT_i	0.779261382	0.464259136
DIC_i	0.588862559	0.540888762
DAA_i	0.506967052	0.582531502
$DPRN_i$	0.265870154	0.767995587
DHA_i	0.289586072	0.771334204
$DGDP_i$	0.224945425	0.798307755
$DEAS_i$	0.292249699	0.744312855
$DIEO_i$	0.310968924	0.739375654
$DSIO_i$	0.307721714	0.758834759
$DTIO_i$	0.196154863	0.782534508
$DDMR_i$	0.568720379	0.545436215
$DDSR_i$	0.454976303	0.624641057
$WMPS_i$	0.364928910	0.640739983
WTS_i	0.132701422	0.763090882
$NUML_i$	0.045023697	0.814275221
$NUME_i$	0.301963415	0.753517275
WIR_i	0.042654028	0.794643939
$WIMR_i$	0.132701422	0.764587451
$WMOR_i$	0.090047393	0.763897732
$WOSR_i$	0.464454972	0.601352034
WSR_i	0.800947867	0.478019161

Table 3. The entropy weight of influencing factors and the average value of comprehensive correlation degree.

Variable	Mean of Entropy Weights	Mean of Comprehensive Correlation Degree
D_{1i}	0.064246000	0.034565220
D_{2i}	0.069125700	0.035675610
D_{3i}	0.066322000	0.033795550
D_{4i}	0.025575000	0.016652774
$DPAR_i$	0.095696900	0.041775317
$DUNI_i$	0.098295500	0.049279866
DCP_i	0.139211000	0.061756123
DDC_i	0.141756800	0.067357677
DPL_i	0.147649300	0.066591187
DTH_i	0.152121800	0.073655940
IPA_i	0.164913500	0.130540504
IPR_i	0.058285000	0.025176313
IPP_i	0.490088800	0.410324289
IPT_i	0.286712700	0.133108991
DIC_i	0.025456000	0.013768872
DAA_i	0.035685900	0.020788182
$DPRN_i$	0.040343300	0.030983443
DHA_i	0.072940100	0.056261209
$DGDP_i$	0.100004500	0.079834372
$DEAS_i$	0.037746600	0.028095270
$DIEO_i$	0.069242800	0.051200554
$DSIO_i$	0.065469000	0.049680169
$DTIO_i$	0.139005500	0.108776577
$DDMR_i$	0.099030300	0.054014714
$DDSR_i$	0.138185700	0.081616443
$WMPS_i$	0.176884700	0.113337136
WTS_i	0.919859200	0.701936198
$NUML_i$	0.035832000	0.029177141
$NUME_i$	0.044308700	0.033387387
WIR_i	0.321233800	0.255266524
$WIMR_i$	0.205659700	0.157244828
$WMOR_i$	0.245145600	0.187266177
$WOSR_i$	0.063589300	0.038239564
WSR_i	0.164371500	0.078572731

5.1. Influence of Traffic Accessibility on Passenger Flow of Rail Transit Stations

The influence of traffic accessibility is now discussed, drawing on the data in Table 3.

For the factors of traffic accessibility, the comprehensive correlation degree between the distance from the rail transit station to the adjacent third-class hospital (DTH_i) and the passenger flow of the station is 0.07365594, which is the maximum value in the comprehensive correlation degree column of this plate, indicating that the distance from the rail transit station to the neighboring third-class hospitals has the greatest influence on passenger flow among all traffic accessibility characteristics. With the acceleration of economic development and urbanization, a large number of people have poured into megacities such as Shanghai and Beijing, which has prompted the rapid growth of basic medical resources and related allocations. The third-class hospitals have strong diagnosis and treatment supply capacity, and the traffic flow generated by the patients gathered in the top three hospitals is not less than that of an important transportation hub [44]. At the same time, due to the imperfect tiered diagnosis and treatment system in China, grassroots hospitals have not been able to form effective patient shunting, resulting in large hospitals with absolute patient attraction, and patients from surrounding districts and counties also flocking to the city center, as well as the sources of disease being more extensive. It is also consistent with the conclusion drawn in this paper that the distance from the rail transit station to the adjacent top three hospitals has a great influence on the passenger flow of the station.

The comprehensive correlations between the distances (DCP_i, DDC_i and DPL_i) from the rail transit station to the adjacent large commercial plazas, district committees and parking areas and the station passenger flow are 0.061756123, 0.067357677 and 0.066591187, respectively, second only to the distance between the third-class hospitals and the station. The magnitude of the correlation degree indicates that the distances from the rail transit station to the adjacent large commercial plazas, district committees and parks also has greater impacts on the passenger flow of the station. Large-scale commercial plazas form a comprehensive high-end urban core area. Their developers choose valuable cities and regions for development and construction, and realize the concentrated urban economic value through comprehensive investment channels such as passenger flow and capital flow. Large commercial plazas are often located in the areas with high land value such as central business districts, financial centers, and transportation hubs. In order to alleviate the pressure on land resources, the government has set up some rail transit stations inside large commercial building complexes. The influence of the district committee on the passenger flow of rail transit stations is mainly due to the changes in the development and construction of rail transit stations in terms of political policies, which has an indirect impact on the passenger flow of the stations. In China, the district committee, as the center of the local district-level government, has overall leadership over the political, economic, cultural and social development of the district, and is responsible for organizing special forums or meetings to discuss development strategies with the rail transit companies, which plays an important role in the construction of rail transit. At present, the integrated construction of commercial buildings, underground parking lots and subways has become the mainstream development mode. The government vigorously advocates subway parking and transfer to non-motor vehicle parking lots, notably Biking & Riding (B-R) parking lots, to guide more people to use the green travelling mode of 'Bike to Subway'.

Secondly, the relationships between the distances from the rail transit station to the nearest park, university or key high school and the passenger flow of the station are also important. From the comprehensive correlation results, the correlations between the distances from the rail transit station to the nearest park, university or key high school ($DPAR_i$ and $DUNI_i$) and the station passenger flow are 0.041775317 and 0.049279866, respectively. As a part of the green space with the greatest social and ecological benefits, the accessibility of urban parks reflects the service capabilities of public facilities and the social distribution of public resources. There are clear requirements for the scale and service radius of parks and green spaces at all levels in the central districts in Shanghai [45]. At the same time, public rail transit provides a variety of travel options and convenient conditions for more people in urban space. In megacities, it is normal for people to choose the subway-walking composite transportation mode to reach parks and green spaces. The accessibility of educational resources in universities and key high schools in Shanghai is generally high, which can better meet the schooling needs of urban residents. In general, rail transit stations are named after their nearby landmark buildings or roads. In Shanghai, there are already some subway stations named after nearby universities, such as Tongji University Station, Jiaotong University Station, Shanghai University Station and Songjiang University Town Station, and the areas with high accessibility are mainly distributed in the areas with developed road networks and developed public transportation networks.

Finally, for the distances from the Shanghai Railway Station, Hongqiao Railway Station and Hongqiao International Airport to the rail transit station, the correlations between the distances (D_{1i}, D_{2i} and D_{3i}) of the stations and the passenger flow of the stations are roughly the same, which are 0.03456522, 0.03567561 and 0.03379555, respectively. The correlation of the distances (D_{4i}) from Pudong International Airport to the rail transit stations and the passenger flow of the stations is the smallest, at 0.016652774, indicating that Pudong International Airport has the weakest impact on the passenger flow of subway stations. The dense number of the stations on Metro Line 2 leads to a long running time. The average time it takes for passengers to arrive at Pudong Airport by rail transit is about 1.5 h, while that for passengers to arrive at Pudong Airport by car is only about 1 h [46]. In addition, the one-way time between Pudong International Airport and Hongqiao Hub by Metro Line 2 takes 2 h; thus, there is a lack of passenger facilities for rapid communication between the two airports [47]. According to the statistics, after the suspension of Metro Line 2 around 10:00 p.m., 8.8% of flights at Pudong International Airport were still arriving and 6.7% departing [48]. The premature shutdown of the subway causes a lack of service connection. Moreover, passengers can choose from various means of transportation to reach Pudong Airport. In addition to subways and taxis, they can also take maglev trains. It takes about 45 min from Longyang Metro Station to Pudong Airport, and only 8 min from the same starting station for maglev trains. At present, Pudong International Airport has two rail transits, including Metro Line 2 and Maglev Line, which can connect to the urban area. Although the subway is cheaper in price, its speed is slow, and multiple transfers may be required, which is inconvenience and uncomfortable.

5.2. Influence of Built Environment on Passenger Flow of Rail Transit Stations

The four index factors of the built environment include the area of important land around the rail transit station (IPA_i), the floor area ratio of the land (IPR_i), the land transfer price (IPP_i), and the time it takes residents to walk to the station (IPT_i). As shown in the correlation results in Table 3, the comprehensive correlation between APS_i and IPP_i is 0.410324289, which is the largest among the four indicators, and far exceeds other indicators of the same type. The land use types have a greater impact on passenger flow in cities, which are divided into four categories, including residential, commercial, office, and public buildings. The higher the land development intensity, the higher the economic benefits of land use and the corresponding increase in land prices. Meanwhile, it can cause increases in land area, so that more people can be accommodated, and thus the traffic demand will be greater. This also explains the impact of the surrounding land area (IPA_i) on the subway passenger flow. At the same time, the increase in land development intensity will also increase the share of public transportation, and the passenger flow of urban rail transit will also increase. The peak value of the cross-section passenger flow mainly occurs before and after the transfer station where the rail transit line enters the urban area in the peripheral area of the city and the station in the large office area. The main source of passenger flow is the commuting passenger flow generated by the residential area around the station along the line [48]. Most shops also have a competitive advantage because of the traffic created by the subway station. The rent of shops near the subway station is generally higher than that of other locations, so the land price will also be higher.

Next, the impact of the time it takes residents to walk to the station (IPT_i) on the passenger flow of the station is discussed. By investigating the passenger flow of Shanghai Metro Line 1, the average time spent by passengers walking from the starting point to the nearest rail transit station is 14 min, and the approximate distance walked is 2 km based on the average speed. The core of the user balance principle is that the users in the transportation network all choose the shortest path to travel, and the impedance of the final selected path is the smallest and equal. This principle reflects the behavioral criteria of road users for route selection; an individual with behavioral decision-making ability in any system always decides his behavioral decision by maximizing his own interests. In this study, residents of nearby land also tend to spend less time and cost in making travel decisions. The correlation between the floor area ratio (IPR_i) of the land and the passenger flow of the station is only 0.025176313, indicating that the impact of the floor area ratio on the passenger flow of the station is minimal. Generally speaking, the higher the floor area ratio, the lower the comfort level of residents. However, there is a certain relationship among the floor area ratio and the building density and the number of floors. The floor area ratio can be used to measure the level of land prices, and increasing the floor area ratio can also improve the efficiency of land use, so it cannot be ruled out that the floor area ratio has an indirect impact on the passenger flow.

5.3. Influence of Regional Characteristics of the District on Passenger Flow of Stations

There are 12 factors considered in the regional characteristics of the district to which the rail transit station belongs. The quantity of the influencing factors is large, and the set sum of the entropy weights of each factor under each station should be equal to 1, which leads to a very small proportion of the weight of each factor, and the calculated comprehensive correlation is also smaller than that of other categories of influencing factors. Thus, the results are reasonable and analyzable.

Among the first nine influencing factors in Table 3, the output value of tertiary industry has the greatest impact on the passenger flow of rail transit stations, and its correlation value is 0.108776577. In China, the tertiary industry is divided into two major departments, including circulation and service, and four levels, which are the circulation department, the department serving production and life, the department serving the improvement of scientific and cultural level, and the department serving the improvement of residents' quality of life [49]. In the definition and statistics of the tertiary industry in China, the fourth level is categorized as belonging to the tertiary industry division, but its added value is not included in the output value of tertiary industry or the gross national product. It can be seen that the tertiary industry in China is still in the traditional service sector. The traditional service industry used traditional consumption methods before the emergence of large-scale industries, mainly including commerce and transportation. The transportation industry is very important in the tertiary industry, and the optimization of the industrial structure will inevitably lead to the evolution and development of passenger flow and transportation to adapt to the new industrial structure. In comparison, the gross output value of the secondary industry has a much smaller impact on the rail passenger flow, and the correlation between them is only 0.049680169. From the GDP proportion of the three industrial structures, the industrial structures of the United States and Japan are relatively similar. In these two countries, the proportion of the added value of the primary and secondary industries continued to decline, and the added value of the tertiary industry kept rising until 2015. The industrial structure in China is different from other countries. The most obvious difference is that the proportion of added value of secondary industry in China has not declined, remaining stable at around 45%, which is higher than the proportion of added value of the primary and tertiary industries. Because the development of tertiary industry started late in China, although it has developed rapidly, it still lags behind the United States and Japan, and there is still much space for development. Moreover, rail transit is mostly located in the downtown area, and suburban rail transit is not well developed. In recent years, Shanghai has vigorously built and developed suburban rail transit. For example, the Disney Station of Metro Line 11 and Metro Line 15, 17 and 18 opened for operation, Metro Line 5 extends south from Dongchuan Road Station to Fengxian New City Station, and the Phase II of Metro Line 10 extends east from the New Jiangwan City Station to Keelung Road Station. In addition, industries related to secondary industry are often located in the suburbs of cities, so the passenger flow of stations has little correlation with the output of the secondary industry.

Secondly, GDP of the district has a greater impact on the passenger flow of the rail station, and its correlation value is 0.079834372. Some studies have found that in Chengdu, China, before 2010, subway passenger flow increased by 1% while the per capita GDP increased by 0.33%. After 2010, that is, after the construction of the subway in Chengdu, the coefficient of subway passenger flow increased to 0.40, indicating that the impact of passenger flow after the construction of the subway on per capita GDP was 0.07% more than the impact on per capita GDP before 2010. It is shown that rail transit has positively influenced economic construction and development in Chengdu [50]. The contribution of urban rail transit to the increased value of real estate along the line is obvious, and even the price of commercial housing under planning has soared. When the supply is fixed and the demand goes up, the land price will also see a gradual rise. Correspondingly, the land grade (DIC_i) will also increase with the growth of land price. The correlation between the land grade and the passenger flow of rail transit stations is 0.013768872. The higher the land grade is, the stronger is the service capacity of the district, which attracts passengers to gather and benefit from the better carrying capacity and service level of the infrastructure. Districts with lower land grades also have less complete rail transit, and the development intensity of land along the line guided by rail transit is about 30–100% higher than that of the same type of land in general districts [51]. Under the premise that the total amount of land supply remains unchanged, the local government can appropriately increase the supply of residential land by adjusting the land supply structure, and can obtain land revenue from the land transfer to reduce the investment pressure on the construction of rail transit. The influences of the district area (DAA_i) and the number of permanent residents ($DPRN_i$) on the passenger flow of rail transit stations are mainly reflected in the station layout. In commercial and residential areas with large passenger flow, smaller station spacing should be considered in order to meet the needs for convenient travel of large passenger flows; in the peripheral areas of the city, the number of trips by the population is large, so larger distance between stations can be considered. In practice, the distance between stations increases in the peripheral areas of the city. In China, houses with more than eight floors are generally considered as high-rise residences. In the case of relatively poor land resources, high-rise residential buildings, which occupy a small area and accommodate many residents, become the necessity of urban development. Shanghai Planning and Land Resources Administration has investigated the high-rise residential buildings, and found that there are about 240 new high-rise residential buildings per year on average, which is equivalent to that a high-rise residence can be built in about one and a half days in Shanghai. If the establishment of rail transit stations can facilitate the travel of residents in a nearby high-rise residential community, most residents will choose to take the subway, and the influx of a large number of nearby residential residents will also affect the passenger flow. The larger the population density is, the greater the passenger flow of rail transit. Similarly, the closer the house is to rail transit, the higher the housing price. The accessibility of rail transit thus is very important for residents' work and living needs.

In most previous studies, the district-level political factors of each district are often ignored. From the results in Table 3, the re-election of district chiefs in each district ($DDMR_i$), the re-election of district party secretaries ($DDSR_i$), and whether the district chief is promoted to district party secretary ($WMPS_i$) exert great influence on the passenger flow of stations, and their correlation degrees are 0.054014714, 0.081616443 and 0.113337136, respectively. In China, the financial decision-making of local governments shows a clear tenure effect. For consideration of promotion, local officials tend to adopt different resource allocation strategies at different times during their tenure, which may bring fluctuations in financial expenditures for rail transit construction. From Table 3, the re-election of the local district head, the re-election of the district party secretary, and the promotion of the local district head to district party secretary have positive impacts on the passenger flow growth of stations. At present, the existing researches mainly focus on the analysis of the impact of official replacement on total fiscal expenditure, productive fiscal expenditure, etc., and the research on rail transit development and expenditure on construction is sparse.

5.4. Influence of Conditions of the Station on Passenger Flow of Stations

From Table 2, it can be seen that there are small differences in the correlation coefficients according to whether the station is a transfer (WTS_i), the number of rail transit lines passed by the station ($NUML_i$), and the number of station entrances and exits ($NUME_i$), which are 0.763090882, 0.814275221 and 0.753517275, respectively. The weight ratio of stations as transfer stations is 0.9198592, which indicates that the data of this indicator value show little change or very little difference. The data of transfer stations show that only 29 items are transfer stations (=1), and 182 items are non-transfer stations taking the station (=0), which is consistent with the above discussion. However, whether it is the correlation coefficient or the comprehensive correlation value, it can still be seen that the transfer station has a more influential impact on the passenger flow of the rail transit station [52]. Under the same external conditions, travelers are more inclined to choose the low-cost rail transit mode. However, when the low-cost mode of transportation cannot meet the passengers' time requirements, the economic factor is no longer the primary factor affecting the transfer, and passengers often gain time at the expense of price. For subway lines, passengers hope to reach the station quickly from the psychological expectation, and are willing to reduce transfer as much as possible.

Currently, among all rail transit stations in Shanghai, the fourth-line transfer station is only Century Avenue Station, and there are 16 third-line transfer stations including Hanzhong Road Station, Shanghai South Railway Station, Xujiahui Station, People's Square Station, South Shaanxi Road Station, Shanghai Railway Station, Hongqiao Railway Station, and Nanjing West Road Station. There are 50 s-line transfer stations including Xinzhuang Station, Changshu Road Station, Caobao Road Station, Shanghai Stadium Station, Nanjing East Road Station, Loushanguan Road Station, Jiangsu Road Station, and Jing'an Temple Station. The number of rail transit lines ($NUML_i$) passing through the station also has an obvious influence on the passenger flow of the rail transit station, and its correlation coefficient is 0.814275221, which is the highest among other indicators of conditions of the station. As the only four-line transfer station in Shanghai's urban rail transit network, Century Avenue Station has the highest average daily passenger transfer flow in the road network, and its transfer passenger flow during the morning rush hour is close to its carrying capacity.

During the station layout planning, the station entrances and exits are the only passages for passengers to enter and exit the station, and must meet the requirements of urban planning and traffic and facilitate the entry and exit of passengers. The number of entrance and exit passages ($NUME_i$) should be determined by the number of passengers getting on and off at the station, and the size of entrance and exit space should consider passenger demand, station facilities, safety of waiting passengers and related specifications of construction design.

5.5. Influence of Spatial Location of the Stations on Passenger Flow of Stations

From the comprehensive correlation degree results in Table 3, the correlation degree between rail transit stations located within the inner ring and station passenger flow is 0.255266524, which is much higher than that of the other four factors of spatial location. It is indicated that the rail transit stations being located within the inner ring of the city has a huge impact on the passenger flow of the station. The location of the inner ring road in Shanghai is completely within the city center. From the heat map shown in Figure 6, the central area of the city is the most dynamic area of the passenger flow in Shanghai, and the streets with high vitality show the same format characteristics. Whether a street combines commercial streets with residential areas, or mainly combines catering, shopping and other formats, the surrounding public transportation facilities are complete and the rail transit stations are densely set up, which is convenient for passengers to travel. These areas also bring great passenger flow to the subway stations in the inner ring. The central ring road is basically located at the edge of the downtown area, which acts as a traffic barrier for the central area, and can quickly divert the transit traffic from the periphery into the urban area. The expressway of the middle ring buffers the traffic between the radial entrance and exit and the downtown area, and shares the passenger flow pressure of the subway line to a great extent. Therefore, it can also be found that the rail transit stations situated between the inner and middle rings ($WIMR_i$) have even less impact on the passenger flow than the stations situated between the middle and outer rings ($WMOR_i$).

The stations located between the outer ring and the suburban ring and outside the suburban ring have little impact on the passenger flow of the station, and their correlation is only 0.038239564 and 0.078572731. From the city center to the suburbs, the distance between rail transit stations shows an increasing correlation. The outer layer of the outer ring line leads to the suburban ring and the expressway that diverges radially in all directions. The Pudong section has fewer expressways connected to the outside than the Puxi section, but it connects the Pudong International Airport, ports and other vital traffic hubs. For the

Metro Line 2 to Pudong International Airport, the average time cost of passengers is about an hour, and more passengers are willing to choose the expressway in the outer ring of the city and the suburban ring to their destination. In addition, the outer ring also connects four main roads to the urban area, including Shangnan Road, Yanggao Road, Shenjiang Road and Jinhai Road; the Pudong section of the outer ring road is also closely connected with the roads in the downtown area.

6. Conclusions and Suggestions

The factors influencing the passenger flow of rail transit stations is explored here based on the mathematical model of entropy weight-grey correlation in this paper. The model assumptions and corresponding variables are proposed, including traffic accessibility, built environment, the regional characteristics of the district to which the rail transit station belongs, conditions of the station and spatial location, affecting the passenger flow of rail transit stations. The mathematical model of entropy weight-grey correlation is established, and the results for factors influencing the passenger flow of rail transit stations are analyzed.

To sum up, the traffic accessibility, built environment, the regional characteristics of the district to which the rail transit station belongs, conditions of the station and spatial location will affect the passenger flow of the station. For traffic accessibility, the distance from the rail transit station to the adjacent third-class hospitals has the greatest impact on the passenger flow of the station, the distance to the Pudong International Airport has the least impact on the passenger flow, and the distance to the adjacent large commercial plazas, district committees and parking areas also have greater impacts on station traffic. Among the built environment factors, the transaction price of important land around the station has the greatest impact on the passenger flow of the station, and the direct impact of the floor area ratio on the rail transit station is very small, but its indirect effect in improving land efficiency cannot be ignored. For the regional characteristics of the district, the ones that exert greater influence on the passenger flow of stations are the total output value of the tertiary industry, the GDP of the district, and whether the head of the district is promoted to the secretary of the district party committee. Among the conditions of the station, whether the station is a transfer station and the number of track lines passed through the station have greater influences on the subway passenger flow. For the spatial location, rail transit stations situated between the inner ring and the middle and outer rings have the greatest influence on the passenger flow of stations, followed by those situated in the middle and inner rings, while the stations outside the outer ring and suburban rings have the least impacts.

The results show that the passenger flow of rail transit stations in Shanghai is closely related to the quality of people's livelihood and economic conditions, and most factors with a greater level of influence are related to people's clothing, food, housing, transportation, employment and medical care and other life matters. The public service of urban rail transit is correlated to the daily travel of people's basic life; this service is public-facing and is an important part of people's life security. At the same time, the construction of rail transit stations will increase the value of the land along the line. The government can make up for the investment in the pre-construction of rail transit by later transferring these lands in exchange for high profits, and the construction of rail transit will also bring new vitality to the regional economy. The results of this paper are consistent with those of recent related studies. Based on the mathematical model and district data of Shanghai from 2007 to 2015, influencing factors such as land area, GDP, tenure of the district chief, and distance from the land to the nearest subway station affect the land transfer behavior of government [35,53,54]. There is an association between land use around a metro station and metro passenger flow, and the composition of land use around the metro station or along the metro line affects the passenger flow generation and the prediction accuracy [55]. Plot potential, development intensity, land ownership, residential appreciation, and mix-value of land use are closely related to the passenger flow [56]. In this paper, the method of entropy weight-grey correlation is used to analyze the influencing factors of the passenger flow of rail transit stations. Under the condition of fuzzy initial information, this model has higher integrity and intuitiveness, which makes it easier to understand the influence level of each factor on station passenger flow.

Furthermore, the regional characteristics of the district to which the station belongs have various and profound impacts on the passenger flow, which has been ignored in the previous analysis of passenger flow. A good urban rail transit planning should be based on the overall urban planning. It is suggested that local governments should study the distribution characteristics of population, land use, employment and tertiary industry at the district level of the city, as well as the distribution characteristics of large passenger flow distribution areas, in order to determine the road network structure and station location plan of rail transit, and promote the development of subway stations and sustainable urban development. In addition, the rail transit stations in the Pudong section have not achieved the expected affordability of passenger flow. When the government studies the new round of station layout planning, factors such as the ring expressway, passenger travel time, and transfer lines in the Pudong section should be considered as having impacts on passenger flow.

This study will help the local governments to formulate rail transit development policies more effectively. The conclusions can provide a reference for the development of rail transit in other large cities and countries.

Author Contributions: Conceptualization, M.P.; methodology, P.Y.; software, P.Y.; validation, P.Y.; formal analysis, P.Y; investigation, M.P.; resources, J.C.; data curation, J.C.; writing—original draft preparation, P.Y.; writing—review and editing, J.C.; visualization, P.Y.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Start-up Funds for Scientific Research of Shenzhen University, grant number 000002112313.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data, models or code generated or used during the study are available from the authors upon request.

Conflicts of Interest: The authors declare that there is no conflict of interest in the publication of this paper.

References

1. Bowes, D.R.; Ihlanfeldt, K.R. Identifying the impacts of rail transit stations on residential property values. *J. Urban Econ.* **2001**, *50*, 1–25. [CrossRef]
2. Sohn, K.; Shim, H. Factors generating boardings at Metro stations in the Seoul metropolitan area. *Cities* **2010**, *27*, 358–368. [CrossRef]
3. Zhao, J.; Deng, W.; Song, Y.; Zhu, Y. What influences Metro station ridership in China? Insights from Nanjing. *Cities* **2013**, *35*, 114–124. [CrossRef]
4. Lemp, J.; Zhou, B.; Kockelman, K.M.; Parmenter, B. Visioning versus modeling: Analyzing the land use-transportation futures of urban regions. *J. Urban Plan. Dev.* **2008**, *134*, 97–109. [CrossRef]
5. Ling, X.; Huang, Z.; Wang, C.; Zhang, F.; Wang, P. Predicting subway passenger flows under different traffic conditions. *PLoS ONE* **2018**, *13*, e0202707. [CrossRef] [PubMed]
6. Tang, S.; Lo, H.K. The impact of public transport policy on the viability and sustainability of mass railway transit – The Hong Kong experience. *Transp. Res. Part A Policy Pr.* **2008**, *42*, 563–576. [CrossRef]
7. Pior, M.-Y.; Shimizu, E. GIS-aided evaluation system for infrastructure improvements: Focusing on simple hedonic and Rosen's two-step approaches. *Comput. Environ. Urban Syst.* **2001**, *25*, 223–246. [CrossRef]
8. Knaap, G.J.; Ding, C.; Hopkins, L.D. Do plans matter? The effects of light rail plans on land values in station areas. *J. Plan. Educ. Res.* **2001**, *21*, 32–39. [CrossRef]
9. So, H.; Tse, R.; Ganesan, S. Estimating the influence of transport on house prices: Evidence from Hong Kong. *J. Prop. Valuat. Invest.* **1997**, *15*, 40–47. [CrossRef]
10. Kim, S.; Ulfarsson, G.F.; Hennessy, J.T. Analysis of light rail rider travel behavior: Impacts of individual, built environment, and crime characteristics on transit access. *Transp. Res. Part A Policy Pr.* **2007**, *41*, 511–522. [CrossRef]

11. Pan, H.; Zhang, M. Rail transit impacts on land use: Evidence from Shanghai, China. *Trans. Res. Record* **2008**, *2048*, 16–25. [CrossRef]
12. Zhang, M.; Wang, L. The impacts of mass transit on land development in China: The case of Beijing. *Res. Transp. Econ.* **2013**, *40*, 124–133. [CrossRef]
13. Zhang, W.; Li, H.; Duan, X. The impacts of rail transit on property values-The case of No. 1 Line in Beijing. *Econ. Geogr.* **2012**, *32*, 46–51.
14. Fu-Zhou, L.; Hong-Yuan, G. The risk assessment model of BT construction engineering project financing. *Syst. Eng. Procedia* **2011**, *1*, 169–173. [CrossRef]
15. Blainey, S. Trip end models of local rail demand in England and Wales. *J. Transp. Geogr.* **2010**, *18*, 153–165. [CrossRef]
16. Gutiérrez, J.; Cardozo, O.D.; García-Palomares, J.C. Transit ridership forecasting at station level: An approach based on distance-decay weighted regression. *J. Transp. Geogr.* **2011**, *19*, 1081–1092. [CrossRef]
17. Cheng, J. Mathematical models and data analysis of residential land leasing behavior of district governments of Beijing in China. *Mathematics* **2021**, *9*, 2314. [CrossRef]
18. Cheng, J. Residential land leasing and price under public land ownership. *J. Urban Plan. Dev.* **2021**, *147*, 05021009. [CrossRef]
19. Cheng, J.; Luo, X. Analyzing the land leasing behavior of the government of Beijing, China, via the multinomial logit model. *Land* **2022**, *11*, 376. [CrossRef]
20. Sung, H.; Oh, J.-T. Transit-oriented development in a high-density city: Identifying its association with transit ridership in Seoul, Korea. *Cities* **2011**, *28*, 70–82. [CrossRef]
21. Cardozo, O.D.; García-Palomares, J.C.; Gutiérrez, J. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Appl. Geogr.* **2012**, *34*, 548–558. [CrossRef]
22. Sun, L.-S.; Wang, S.-W.; Yao, L.-Y.; Rong, J.; Ma, J.-M. Estimation of transit ridership based on spatial analysis and precise land use data. *Transp. Lett.* **2016**, *8*, 140–147. [CrossRef]
23. Sung, H.; Choi, K.; Lee, S.; Cheon, S. 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. Jun, M.-J.; Choi, K.; Jeong, J.-E.; Kwon, K.-H.; Kim, H.-J. Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *J. Transp. Geogr.* **2015**, *48*, 30–40. [CrossRef]
25. Keçeci, B.; Ic, Y.T.; Eraslan, E. Development of a spreadsheet DSS for multi-response taguchi parameter optimization problems using the TOPSIS, VIKOR, and GRA methods. *Int. J. Inf. Technol. Decis. Mak.* **2019**, *18*, 1501–1531. [CrossRef]
26. Orazbayev, B.; Zhumadillayeva, A.; Orazbayeva, K.; Iskakova, S.; Utenova, B.; Gazizov, F.; Ilyashenko, S.; Afanaseva, O. The system of models and optimization of operating modes of a catalytic reforming unit using initial fuzzy information. *Energies* **2022**, *15*, 1573. [CrossRef]
27. Wei, G.-W. Gray relational analysis method for intuitionistic fuzzy multiple attribute decision making. *Expert Syst. Appl.* **2011**, *38*, 11671–11677. [CrossRef]
28. Wei, G.-W. GRA method for multiple attribute decision making with incomplete weight information in intuitionistic fuzzy setting. *Knowl. Based Syst.* **2010**, *23*, 243–247. [CrossRef]
29. Sunaga, T. Theory of an interval algebra and its application to numerical analysis. *RAAG Memoirs* **2009**, *26*, 125–143. [CrossRef]
30. Cao, J.; Liu, X.C.; Wang, Y.; Li, Q. Accessibility impacts of China's high-speed rail network. *J. Transp. Geogr.* **2013**, *28*, 12–21. [CrossRef]
31. Edin, M. State Capacity and local agent control in China: CCP cadre management from a township perspective. *China Q.* **2003**, *173*, 35–52. [CrossRef]
32. Ewing, R.; Cervero, R. Travel and the built environment: A synthesis. *Transp. Res. Rec. J. Transp. Res. Board* **2001**, *1780*, 87–114. [CrossRef]
33. Tang, W.; Cui, Q.; Zhang, F.; Yan, H. Evaluation of the land value-added benefit brought by urban rail transit: The case in Changsha, China. *J. Transp. Land Use* **2021**, *14*, 563–582. [CrossRef]
34. Li, L. The dynamics of the Shanghai land market–An intra city analysis. *Cities* **2011**, *28*, 372–380. [CrossRef]
35. Cheng, J. Analyzing the factors influencing the choice of the government on leasing different types of land uses: Evidence from Shanghai of China. *Land Use Policy* **2019**, *90*, 104303. [CrossRef]
36. Xue, H.; Yang, P.; Zhang, H.; Jing, E. Study on the control strategy of urban rail transit passenger flow under the condition of large passenger flow. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *234*, 012001. [CrossRef]
37. *GB 50157-2003*; Code for Design of Metro. China National Standard: Beijing, China, 2003.
38. Dong, Y.-H.; Peng, F.-L.; Guo, T.-F. Quantitative assessment method on urban vitality of metro-led underground space based on multi-source data: A case study of Shanghai Inner Ring area. *Tunn. Undergr. Space Technol.* **2021**, *116*, 104108. [CrossRef]
39. Zhang, C. Money, housing, and inflation in China. *J. Policy Model.* **2013**, *35*, 75–87. [CrossRef]
40. Fitzpatrick, K.; Brewer, M.A.; Turner, S. Another look at pedestrian walking speed. *Trans. Res. Record* **2006**, *1982*, 21–29. [CrossRef]
41. Liu, H.; Dong, Y.; Wang, F. Gas outburst prediction model using improved entropy weight grey correlation analysis and IPSO-LSSVM. *Math. Probl. Eng.* **2020**, *2020*, 1–10. [CrossRef]
42. Yin, P.; Cheng, J. The software system of urban land planning database of Shanghai in China based on MySQL. *CMES-Comput. Model. Eng. Sci.*. In press.

43. Sun, C.C. Combining grey relation analysis and entropy model for evaluating the operational performance: An empirical study. *Qual. Quant.* **2014**, *48*, 1589–1600. [CrossRef]
44. Jabareen, Y.R. Sustainable urban forms: Their typologies, models, and concepts. *J Plann. Educ. Res.* **2006**, *26*, 38–52. [CrossRef]
45. Wang, X.-J. Analysis of problems in urban green space system planning in China. *J. For. Res.* **2009**, *20*, 79–82. [CrossRef]
46. Givoni, M.; Chen, X. Airline and railway disintegration in China: The case of Shanghai Hongqiao Integrated Transport Hub. *Transp. Lett.* **2017**, *9*, 202–214. [CrossRef]
47. Chang, Y.C. Factors affecting airport access mode choice for elderly air passengers. *Trans. Res. Part Logistics Trans. Rev.* **2013**, *57*, 105–112. [CrossRef]
48. Crowley, D.F.; Shalaby, A.S.; Zarei, H. Access Walking Distance, Transit Use, and Transit-Oriented Development in North York City Center, Toronto, Canada. *Transp. Res. Rec. J. Transp. Res. Board* **2009**, *2110*, 96–105. [CrossRef]
49. Wang, Q.; Liu, L.; Wang, S.; Wang, J.-Z.; Liu, M. Predicting Beijing's tertiary industry with an improved grey model. *Appl. Soft Comput.* **2017**, *57*, 482–494. [CrossRef]
50. Wei, X.; Zhang, W.K.; Wang, C.; Xu, G.J. The Impact of Urban Rail Transit on Surrounding Residential Prices–Line 1 of Chengdu Metro as an Example. *Mod. Appl. Sci.* **2012**, *6*, 58. [CrossRef]
51. Ma, X.; Wu, Y.-J.; Wang, Y.; Chen, F.; Liu, J. Mining smart card data for transit riders' travel patterns. *Transp. Res. Part C Emerg. Technol.* **2013**, *36*, 1–12. [CrossRef]
52. Sun, Y.; Zhang, G.; Yin, H. Passenger Flow Prediction of Subway Transfer Stations Based on Nonparametric Regression Model. *Discret. Dyn. Nat. Soc.* **2014**, *2014*, 397154. [CrossRef]
53. Cheng, J. Data Analysis of the Factors Influencing the Industrial Land Leasing in Shanghai Based on Mathematical Models. *Math. Probl. Eng.* **2020**, *2020*, 9346863. [CrossRef]
54. Cheng, J. Analysis of commercial land leasing of the district governments of Beijing in China. *Land Use Policy* **2021**, *100*, 104881. [CrossRef]
55. Lin, C.; Wang, K.; Wu, D.; Gong, B. Passenger Flow Prediction Based on Land Use around Metro Stations: A Case Study. *Sustainability* **2020**, *12*, 6844. [CrossRef]
56. Jin, S.; Wang, D.; Su, B. Functional orientation and spatial siting of subway stations based on land potential and mixed land use. *Ain Shams Eng. J.* **2022**, *13*, 101733. [CrossRef]