Passengers Flow Forecasting Model of Urban Rail Transit Based on the Macro-Factors

Bin Shang [1, a], Xiaoning Zhang [2, b]

[1]Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China

[2] School of Economics and Management, Tongji University, Shanghai 200092, China

[a] 2010shangbin@ tongji.edu.cn, [b] *Corresponding author:* cexzhang@ tongji.edu.cn

Keyword: Local finance; Gross domestic product; Urban population; Urban rail transit; Passengers flow forecasting

Abstract. In China, many cities are planning urban rail transit system, but a comprehensive passenger flow estimation model is still lacking. The total passenger flow of urban rail transit in a city depends on many factors, such as urban population, total length of rail lines, gross domestic production of the city etc. To estimate the total passenger flow of urban rail transit, a linear regression model with multiple variables is established in the paper, based on the real data collected in many cities with urban rail transit operating. The comparison of the estimated flow and the real flow in many cities shows that the model is very accurate in passenger flow forecasting.

Introduction

Urban rail transit (URT) has developed rapidly in China which has important influence on exerting city function efficiently, improving city environments and promoting the economic and social development. Before the year of 2003, some cities had built the projects of URT blindly without considering their revenue. There were so many projects to build which did not be state-sanctioned or with high standard and causing the huge wasting of investment or causing the financial loss after operating because of the lack of capital. In order to improve the construction management, promote UTR developing healthfully and confirm the building of UTR adapting the economic development of city, General Office of the State Council of the People's Republic of China (GOSCPR) put forward a notice about strengthening the building and management of UTR.

The notice of GOSCPR put forward the basic conditions about application for construction of metro and light rail. If a city wants to construct metro, it must be reaching the basic conditions such that the general budgetary revenue of local finance (GBRLF) is over the 10 billion yuan, gross domestic production (GDP) is over the 100 billion yuan, the number of urban population(NUP) are over 300 million and the one-way planning passenger is over 30 thousand. If a city wants to construct light rail, it must be reaching the basic conditions such that the GBRLF is over the 6 billion yuan, GDP is over the 60 billion yuan, the urban population is over 150 million and the one-way planning passenger is over 10 thousand [1]. The study of passenger forecasting model is generally base on the four-step method of traffic planning which is not deeply analyzing the macro-factors impact on passenger flow [2-4].

This paper builds the passengers flow forecasting model of UTR based on the macro factors through analyzing the basic conditions which have influences on the passengers flow. The model not only builds the relationship between the macro-factors and passengers, but also makes the calculated value closing to the actual value. The model is simple in form and can be application easily.

Analyzing of macro-factors

There are thirteen cities in Chinese Mainland ending 2011 that have operated the urban rail transit such as Beijing, Shanghai, Tianjin, Chongqing, Guangzhou, Shenzhen, Nanjing, Chengdu, Shenyang, Xi'an, Dalian, Changchun and Wuhan, more than 50 lines, including metro and light rail. The overall

length of total lines is up to 1600 km. UTR in some cities including Beijing, Shanghai, Shenzhen, Guangzhou, have been network operation, and the traffic circumstance in these cities have been improved obviously. Table1 gives the data of these cities about the average daily passenger flow(ADPF), GBRLF , GDP, NUP who live in the main city and the length of UTR (LURT) and so on [5].

The paper mainly uses macro-factors including ADPF, GBRLF, GDP and LURT to make regression analysis. Following analyzing the impact of macro factors on passenger flow of UTR.

Table1　Data statistics of urban rail transit in the main city [6]

Num.	Cities	ADPF [million]	GBRLF of 2011 [billion]	GDP of 2011 [billion]	NUP [million]	LURT [Km]
1	Beijing	5.60	300.63	1600.04	16.21	336.610
2	Shanghai	6.735	342.983	1919.569	18.21	506.246
3	Chongqing	0.40	148.825	1001.113	5.23	35.650
4	Tianjin	0.133	145.487	1119.099	9.93	26.200
5	Guangzhou	4.00	97.947	1230.312	8.51	243.365
6	Shenzhen	2.10	133.959	1150.206	8.91	178.989
7	Nanjing	1.12	63.5	614.552	6.69	84.750
8	Chengdu	0.1515	68.07	685.46	5.76	18.500
9	Xi'an	0.189	31.855	386.421	5.99	20.500
10	Shenyang	0.12	62.01	591.49	5.62	27.800
11	Wuhan	0.20	67.326	675.62	5.63	28.500
12	Changchun	0.055	28.3	404.0	3.59	32.000
13	Dalian	0.145	65.1	615.01	3.78	63.500

General budgetary revenue of local finance(GBRLF). The investment of UTR is usually more than 4 billion yuan because that the investment of URT is about 300 million yuan per Kilo-meters and the length of UTR is about 15 Km. The operating cost of UTR which is a variety of system including vehicle system, condition system and aerator system and so on is higher than other kind of traffic mode. So the project of UTR has the disadvantages of more initial investment, high operating cost. The notice of GOSCPR put forward the basic conditions which ensure the metro and light rail system can be constructed successfully such that If a city wants to construct metro system, GBRLF is over the 10 billion yuan; If a city wants to construct light rail, GBRLF is over the 6 billion yuan.

Table 2　Regression equation and parameters statistics

Method	Regression equation	R^2
Linear	y = 0.2037x - 82.701	0.740
logarithmic	y = 234.67ln(x) - 1438.8	0.580
Exponential	$y = 10.962e^{0.0013x}$	0.552
Power	$y = 0.0007x^{1.6403}$	0.555

At first, we analyze the relationship between GBRLF and passengers flow of URT. Table2 gives the regression equation including linear regression, logarithmic regression, exponential regression, Power regression, correspondingly correlation coefficient is listed in table 2. From table2 we can see that correlation coefficient of linear is higher than 0.740 and the others kind of correlation coefficient is less than 0.6. the linear regression is not only simpler, but also better than other kind of regression. So using the linear function to illustrate the relationship between GBRLF and passengers flow of URT is reasonable.

Gross domestic product(GDP). The price of UTR ticket is higher than other kind of public transport because that the project of UTR has the disadvantages of more initial investment and high operating cost. The price of tickets is related with the residents travel cost. The economic

development level of city can be seen from the GDP and the ratio of GDP to the number of urban population can reflect the average consumption level of city. This shows that the GDP is the sufficient conditions for ensuring the metro and light rail system can be operating successfully.

The notice of GOSCPR put forward the basic conditions for ensuring the metro and light rail system can be operating successfully such that If a city wants to construct metro system, GDP is over the 100 billion yuan; If a city wants to construct light rail, GDP is over the 600 billion yuan.

We analyze the relationship between GDP and passengers flow of URT. Table3 gives the regression equation including linear regression, logarithmic regression, exponential regression, Power regression, correspondingly correlation coefficient is also listed in table 3. From table3 we can see that correlation coefficient of linear is higher than 0.796 and the others kind of correlation coefficient is less than 0.7. the linear regression is not only simpler, but also better than other kind of regression. So using the linear function to illustrate the relationship between GDP and passengers flow of URT is reasonable.

Table 3   Regression equation and parameters statistics

Method	Regression equation	R^2
Linear	y = 0.0444x - 248.92	0.796
logarithmic	y = 374.42ln(x) - 3214.2	0.647
Exponential	$y = 3.2667e^{0.0003x}$	0.684
Power	$y = 2E-09x^{2.6634}$	0.642

Number of urban population (NUP). The notice of GOSCPR put forward the basic conditions for ensuring the metro and light rail system can be operating successfully such that If a city wants to construct metro system, NUP is over the 300 million; If a city wants to construct light rail, NUP is over the 100 million. This shows that the NUP is the other sufficient conditions for ensuring the metro and light rail system can be operating successfully besides the GDP.

We analyze the relationship between NUP and passengers flow of URT. Table4 gives the regression equation including linear regression, logarithmic regression, exponential regression, Power regression, correspondingly correlation coefficient is listed in table4. From table4 we can see that correlation coefficient of linear is higher than 0.821 and the others kind of correlation coefficient is less than 0.75. The linear regression is not only simpler, but also better than other kind of regression. So using the linear function to illustrate the relationship between GDP and passengers flow of URT is reasonable.

Table 4   Regression equation and parameters statistics

Method	Regression equation	R^2
Linear	y = 0.468x - 213.53	0.821
logarithmic	y = 405.15ln(x) - 2498.5	0.735
Exponential	$y = 4.7877e^{0.0029x}$	0.624
Power	$y = 7E-07x^{2.7427}$	0.660

Length of urban rail transit (LURT). The project of UTR is usually first built one lines in the central of city or along the major economic zone and then add the other lines through considering the economic development of city. With the length of UTR becoming longer, the passengers of UTR is increasing quickly; and following the network operation of UTR, the passenger will increase substantially.

From the column of ratio of passengers flow to the length of UTR in table5, the UTR in the Chinese main land can be divided to two categories, one for Beijing, Shanghai, Guangzhou, Shenzhen, Nanjing, Chongqing which the ratio is greater than 1. Through the column of LURT in table5, we know that the length of UTR is longer than 80 Km except Chongqing city. The passenger flow is so big because of network operation which is consisted of 3 lines at least. Chongqing is an mountain city and the people in Chongqing is distributed along the mountain which is formed an

narrow economic zone. Though Chongqing only has one line of UTR which is along the narrow economic zone, the passenger flow is big. The other category is including Tianjin, Chengdu, Xi'an, Shenyang, Wuhan, Changchun, Dalian which the ratio of passengers flow to the length of UTR is less than 1. Through the column of LURT in table5, we know that UTR in these cities are not formed network. And then reachability of UTR system is not enough to attract the passenger. Based on the above analysis, this paper analyzes impact of the total length of UTR on passenger.

Table5   The ratio of average passengers flow and the length of urban rail transit

Num.	Cities	ADPF（million）	LURT（Km）	The ratio of ADPF to LURT(10^4/km)
1	Beijing	5.60	336.610	1.664
2	Shanghai	6.735	506.246	1.330
3	Chongqing	0.40	35.650	1.122
4	Tianjin	0.133	26.200	0.508
5	Guangzhou	4.00	243.365	1.644
6	Shenzhen	2.10	178.989	1.173
7	Nanjing	1.12	84.750	1.322
8	Chengdu	0.1515	18.500	0.819
9	Xi'an	0.189	20.500	0.922
10	Shenyang	0.12	27.800	0.432
11	Wuhan	0.20	28.500	0.702
12	Changchun	0.055	32.000	0.172
13	Dalian	0.145	63.500	0.228

We analyze the relationship between LURT and passengers flow of URT. Table6 gives the regression equation including linear regression, logarithmic regression, exponential regression, Power regression, correspondingly correlation coefficient is listed in table6. From table6, we can see that correlation coefficient of linear is higher than 0.971 and the others kind of correlation coefficient is less than 0.9. the linear regression is not only simpler, but also better than other kind of regression. So using the linear function to illustrate the relationship between GDP and passengers flow of URT is reasonable.

Table 6   Regression equation and parameters statistics

Method	Regression equation	R^2
Linear	y = 1.5012x - 23.938	0.971
logarithmic	y = 188.54ln(x) - 625.48	0.8625
Exponential	$y = 15.13e^{0.0096x}$	0.7767
Power	$y = 0.1774x^{1.349}$	0.866

The model

Through above analysis between passenger flow and macro factors such as GBRLF, GDP, NUP and LURT, we can see that linear function is better than others and linear regression is close to actual data. From above thinking and using the linear function, the paper builds the passenger forecasting model of UTR. Thus:

$$Y = b + a_1X_1 + a_2X_2 + a_3X_3 - a_4X_4 \qquad (1)$$

Where   : X_1—General budgetary revenue of local finance(GBRLF);

X_2—Gross domestic product (GDP);

X_3—Number of urban population(NUP);

X_4—Length of UTR (LURT);

b——Constant;

a_1, a_2, a_3, a_4, ——correlation coefficient;

Application of model and analysis of results

Through comparing the predictive value with actual data, it is verified that model is valid which is using the statistic software of SPSS to make a multiple linear regression. Actual data and predictive value is listed in the table7. Regression function is:

$$Y = -0.0396X_1 + 0.0055X_2 + 0.0904X_3 + 1.3316X_4 - 78.878 \tag{2}$$

Table 7   The actual and predictive value of urban rail transit

| Num.	Cities	Actual data	Statistic of SPSS	
			Forecasting data	Standardized residuals
1	Beijing	560.00	485.0616	1.6622
2	Shanghai	673.50	729.9356	-1.2518
3	Chongqing	40.00	12.1815	0.6170
4	Tianjin	13.30	49.9389	-0.8127
5	Guangzhou	400.00	351.2142	1.0821
6	Shenzhen	210.00	250.4438	-0.8971
7	Nanjing	112.00	103.2102	0.1950
8	Chengdu	15.15	8.6945	0.1432
9	Xi'an	18.90	11.2709	0.1692
10	Shenyang	12.00	17.0249	-0.1115
11	Wuhan	20.00	20.5956	-0.0132
12	Changchun	5.50	7.2718	-0.0393
13	Dalian	14.50	48.0064	-0.7432
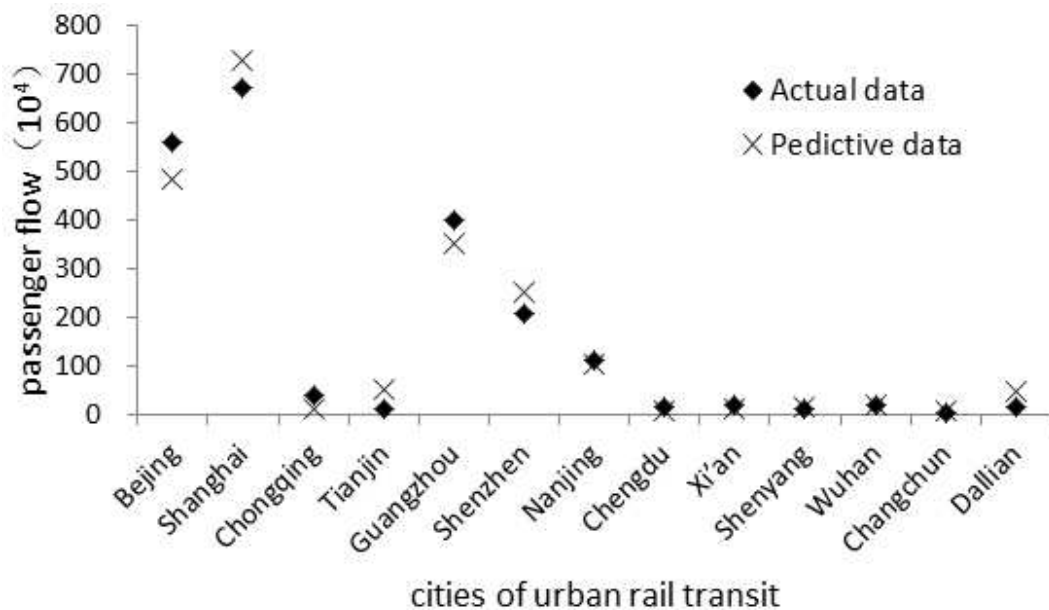
Fig.1   Relationship between actual and predictive passengers flow value

From the fig.1 and table7, we can see that the value calculated by model is very close to the practice. Standard residuals in table7 show that the model is reasonable and valid. Standard residuals subject to the standard normal distribution N(0，1). From the column of standardized residuals in table7, the data do not exist outside of the interval (-2,2), so this can explain that the probability of standardized residual data falls outside the interval (-2,2) is less than 0.05. If the data points do not fall outside the (-2,2) interval, 95% confidence level for determining abnormal data as the data does not exist. Through the analysis of standardized residual data, the model is verified to be valid and effective.

Conclusions

This paper use the actual data from the URT city to analyze the relationship between passenger flow of UTR and macro-factors such as GBRLF, GDP, NUP and LURT which are listed in the notice about strengthening the building and management of UTR which is informed by General Office of the State Council of the People's Republic of China.

Sufficient GBRLF guarantees the metro and light rail system to be constructed successfully. GDP reflects the economic development level of city and per capita consumption level. If per capita consumption level is increasing, per capita expenditure on transport will be also increasing which will drive the passenger flow increasing. Sufficient NUP which can provide enough passenger flow ensures that the UTR system can be operating successfully.

We also found that the relationship between passenger flow and LURT is so strong that the passengers of UTR are increasing quickly with the length of UTR becoming longer; The passenger data of cities such as Beijing, Shanghai, Guangzhou and Shenzhen shows that passenger flow increase substantially following the network operation of UTR. The result shows that linear regression is better than other kind of regression and macro factors have a greater effects on the passenger flow of UTR.

Finally,, using the data of UTR in main-land to make the regression and the result shows that predictive value is closing to the actual data. The model has the value of theory and engineering applications which can be used in the passenger flow forecasting.

Acknowledgments

References

[1] General Office of the State Council of the People's Republic of China. A notice about strengthening the building and management of urban rail transit. On

[2] http://www.gov.cn/gongbao/content/2003/content_62476.htm

[3] M. MAO, J. ZHANG, H.W. ZHOU. Passenger Volume Forecasting of Guangzhou-Zhuhai Intercity Rapid Mass Transit, Journal of Southwest JiaoTong University, 39(2004): 195-198.

[4] R.H. Xu, Q. Luo, P. Gao. Passenger flow distribution model and algorithm for urban rail transit network based on multi-route choice. Journal of the China Railway Society. 31(2009): 110-114.

[5] D. Chen, W.Z. XIAO, X.H. Li et al. Forecast and analysis of passenger flow of urban rail transit under iterative feedback constraints. Journal of South China University of Technology: Natural Science Edition. 39(2011): 99-103.

[6] B. Shang, X.N. Zhang. Current Situation and Thinking of Urban Rail Transit in China, Traffic and Transportation. 2011 90-92

[7] Annual Statistical Bulletin of 2011 on http://www.city-net.cn/p!feature.do?cid=2