Passenger Flow Forecasting for Chinese High Speed Rail Network

Nuannuan Leng
School of traffic and transportation
Beijing Jiaotong University
Beijing, China
12120874@bjtu.edu.cn

Gencai Guo
School of traffic and transportation
Beijing Jiaotong University
Beijing, China
10114219@bjtu.edu.cn

Lei Nie
School of traffic and transportation
Beijing Jiaotong University
Beijing, China
lnie@bjtu.edu.cn

Xin Wu
School of traffic and transportation
Beijing Jiaotong University
Beijing, China
11114239@bjtu.edu.cn

Abstract—Forecasting passenger flow accurately, especially for the whole network, is extremely important for the holistic optimization of transportation organization. This paper is concerned with the entire high-speed rail network, not just focusing on one single channel, and proposes a complete forecasting process by considering the unique operating characteristics of Chinese high-speed rail network. The process consists of four steps to forecast the total passenger flow, the economy-based demand growth, the diverted demand and the induced demand. Meanwhile, three steps have been mainly improved. First, predict the future total passenger flow by analyzing its relation to the rail network and calculating the restriction of transport capacity. Second, cluster origin–destination (OD) flows by considering three factors (GDP, population and the distance) in order to get reasonable parameter. Third, adjust the induced demand according to the restriction of the total passenger flow. At last, the travel demand of the whole high-speed rail network of China in 2016 is forecasted.

Keywords- rail transport; passenger flow; forecasting; four-step method; high-speed rail network.

I. INTRODUCTION

Chinese HSR network structure has been changing rapidly with the construction of new lines almost every year since 2008. The passenger flow has been increased unconventionally for these years. Once a new line operates, large amount of passenger flow transfers to HSR from the conventional rail due to the restriction of capacity.

The common forecasting methods include time series analysis method [1], grey forecasting method [2], artificial neural network [3], etc. However, these forecasting methods have a better effect under the premise of complete historical data and stable passenger flow trend. In China, many scholars choose four-step method to forecast passenger flow for single high-speed railway, like Beijing-Shanghai high-speed railway [4-5], Wuhan-Guangzhou high-speed railway [6], etc. This paper also chooses four-step method whose forecasting accuracy is high [7]. But the date size is large for Chinese high-speed railway. It's hard to get all features of passenger flow of road, air and conventional railway.

This paper uses core idea of trip generation forecasting for reference and then forecasts total passenger flow by analyzing characteristics of the study object [8]. The OD passenger flow of high-speed railway can be divided into three parts: economy-based demand growth, diverted demand and induced demand [9]. The gravity model [10] is used to forecast economy-based demand growth and the Logit model [11] is used to forecast diverted demand.

II. FORECASTING PROCESS AND METHODS

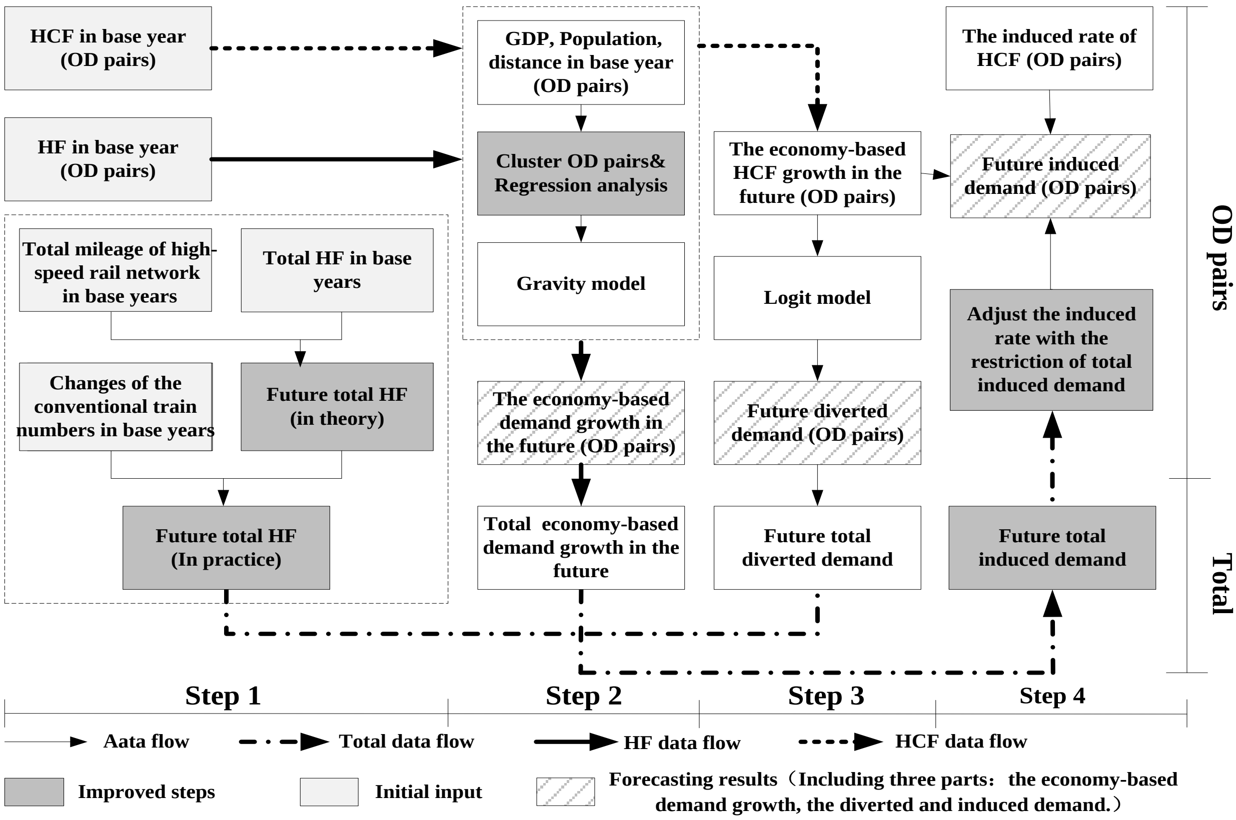
The complete forecasting process which is adaptable to the developing high-speed rail network is shown in Fig .1 According to the passengers' income level and the conventional train type they have chosen, the conventional rail passenger flow (CF) can be divided into two types: one is more likely to shift to the high-speed rail; the other prefers the conventional train. Thus, this paper primarily focuses on the former type and names this type as the high-end conventional rail flow (HCF). Correspondingly, the latter type is called low-end conventional rail passenger flow (LCF). Also, HF is short for the high-speed rail passenger flow.

A. Step1 Future Total HF Forecasting.

Try to find the reasonable relation between the total HF and the total mileage of high-speed rail network in base years by the method of data fitting. Then, assume this relation would stay the same in the future so as to calculate the total HF in the future. What should be noticed is that this result is just a theoretical value, without considering the actual rail operation situation. Hence, taking the changes of conventional train numbers into consideration, the theoretical result could be corrected in order to get the future total HF in practice.

B. Step2 Economy-Based Demand Growth Forecasting

In the condition of the quickly changing rail network structure, it is a better choice to use gravity model to



forecast the economy-based demand growth. The model is described as in (1):

$$q_{ij} = \theta \frac{(G_i \cdot G_j)^{\alpha}(P_i \cdot P_j)^{\beta}}{t_{ij}^{\gamma}} \tag{1}$$

Where q_{ij} is the passenger flow of the *ij* OD pair; G_i , P_i , G_j , P_j indicate the GDP and population of the origin *i* and the destination *j*, respectively; t_{ij} is the travel time of the *ij* OD pair; θ , α , β , γ are unknown parameters. Due to the large numbers of OD pairs, we need to cluster OD pairs which have similar characteristics in terms of GDP, population and the distance. Then, use the least square method to calibrate the parameters of each cluster. In the process of regression analysis, if the F or t inspection value is not extremely remarkable, we should cluster again until all clusters meet the inspection standards.

C. *Step3 Diverted Demand Forecasting*

The Logit model is described as in (2):

$$P_{ij}(k) = \frac{\exp(\alpha_k + \beta_1 \cdot t_{ij}(k) + \beta_2 \cdot mc_{ij}(k))}{\sum_{l=1}^{K} \exp(\alpha_l + \beta_1 \cdot t_{ij}(l) + \beta_2 \cdot mc_{ij}(l))} \tag{2}$$

Where $P_{ij}(k)$ is the share of the transport model *k*; t_{ij} is the travel time of the *ij* OD pair; mc_{ij} is the monetary travel cost of the *ij* OD pair; α_k , β_1 , β_2 are unknown parameters. After forecasting the economy-based HCF growth in the future by using (1), we use (2) to calculate the share of the high-speed rail which can be seen as the diverted percentage from the high-end conventional rail. Then, the diverted demand would be solved.

D. *Induced Demand Forecasting.*

The method to calculate induced rate is described as in (3):

$$R_{ij} = \left(\frac{t_{ij}(C)}{t_{ij}(H)}\right)^{\beta} \text{-1} \tag{3}$$

Where R_{ij} is the induced rate of the *ij* OD pair; $t_{ij}(C)$ is the travel time of the conventional rail; t_{ij}(H) is the travel time of the high-speed rail; β is the parameter in (1). In order to satisfy the total HF calculated in the first step, we need to adjust the induced rate of each OD pair with the restriction of total induced demand. The adjustment method is described as in (4):

$$I_{ij}^{'} = I_{ij} \times \frac{TI}{\sum_{i=1}^{N}\sum_{j=1}^{N} I_{ij}} = I_{ij} \times \frac{T - TG - TD}{\sum_{i=1}^{N}\sum_{j=1}^{N} I_{ij}} \tag{4}$$

Where I_{ij} is the induced demand of the *ij* OD pair; $I_{ij}^{'}$ is adjusted induced demand of the ij OD pair; *T* is total passenger demand; *TG* is total economy-based demand growth; *TD* is total diverted demand; *TI* is total induced demand.

III. APPLYING METHODS TO FORECAST PASSENGER FLOW

Fig .2 shows the scope of Chinese high-speed rail network in 2016, which can be divided into 437 traffic zones. The task is to forecast the passenger flow in 2016 based on the rail ticket data in 2011.

In the China Statistical Yearbook, there is only the data of total rail passenger flow, without distinguishing total HF from total CF. It is generally know that Chinese high-speed rail has been constructed since 2008, therefore, the total rail passenger flow is equal to total CF before 2008, while it is the sum of total CF and total HF after 2008.

Figure 2. Map of Chinese High-speed Railway Network in 2016

With the assumption that the growth trend of total CF is consistent regardless of the construction of the high-speed rail, the total CF in base years could be calculated by grey prediction. Then, the total rail passenger flow minus the total CF is the total HF. The data is shown in Table 1.

Years	2008	2009	2010	2011	2012	2013
Total rail flow	1.46	1.52	1.68	1.86	1.89	2.08
Total CF	1.33	1.38	1.44	1.49	1.55	1.61
Total HF	0.13	0.14	0.24	0.37	0.34	0.47
Total mileage	1.22	3.52	5.39	7.04	9.35	12.1

TABLE I. TOTAL PASSENGER FLOW (BILLION PEOPLE) AND TOTAL MILEAGE (THOUSAND KILOMETERS)

Fig .3 presents the relation between the total HF and mileage of high-speed rail from 2008 to 2013.

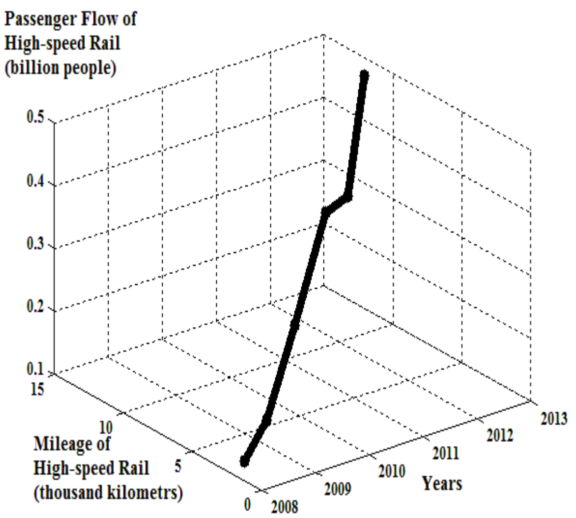


Figure 3. Relation between the Total HF and Mileage of High-speed Rail

The relation can be indicated by linear fitting as in (5):

$$f(x) = 0.03558x + 0.06523 \tag{5}$$

Where $f(x)$ is the total high-speed rail passenger flow; x is the total mileage of high-speed rail.

It is known that the total mileage of high-seep rail in 2016 is about 22.05 thousand kilometers, so the total HF in 2016 is forecasted to be about 0.85 billion people. The forecasting result is shown in Fig .4.

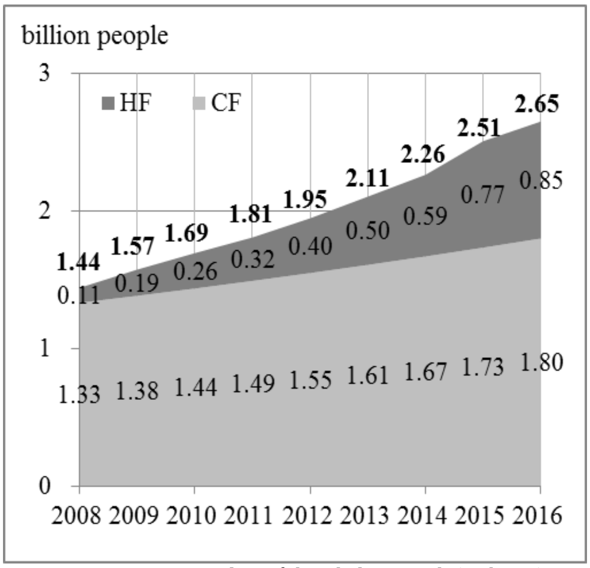


Figure 4. Passenger Flow of the whole network (in theory)

Fig .5 shows train numbers of conventional rail with different train types. Because of high-speed railway, the total number of conventional train remains unchanged. That means the number of passenger flow of conventional railways will keep unchanged in the future. As a result, the forecasting result in Fig .4 needs to be modified to be more realistic.

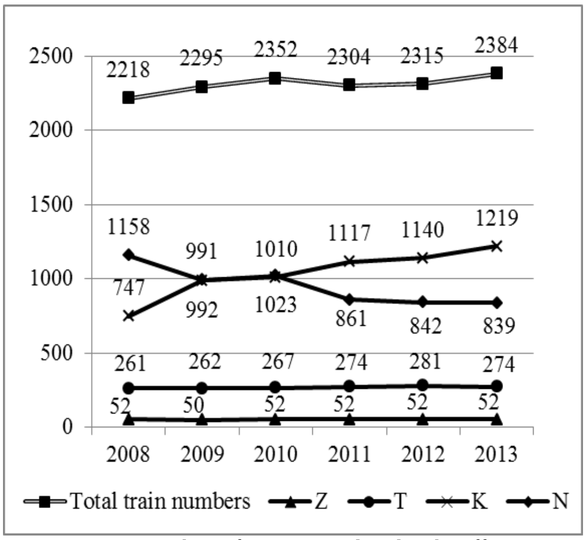


Figure 5. Train Numbers of Conventional Rail with Different Train Types

The final forecasting result of the whole rail network is 2.65 billion people. (shown in Fig .6). The passenger flow of high-speed rail is 1.21 billion people, that of high-end conventional rail is 0.63 billion people and low-end conventional rail is 0.81 billion people.

According to the division of traffic area, there are 6,915 ODs of high-speed passenger flow and there are 29,979 ODs of high-end conventional passenger flow. The high-speed passenger flow ODs are divided into 33 groups, while high-end conventional passenger flow ODs are divided into 96 groups by SPSS. The relationship between the dependent variable and independent variable integral is highly significant and the relationship between single independent variable and the dependent variable is also significant.

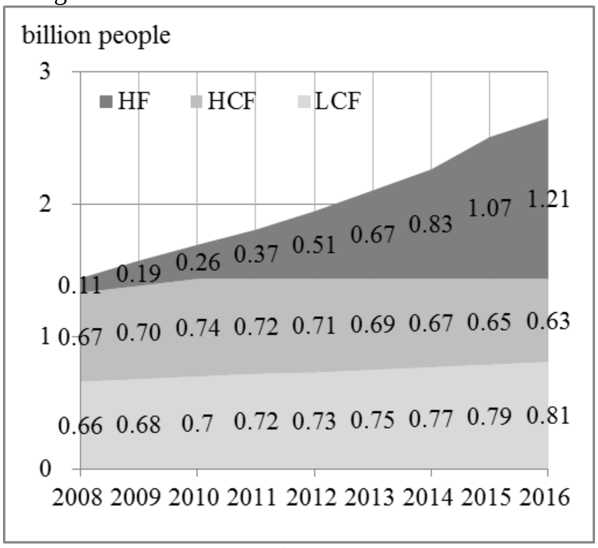


Figure 6. Passenger Flow of the whole network (in practice)

The economy-based demand growth of high-speed railway is 0.55 billion people, the diverted demand is 0.41 billion people and the induced demand is 0.25 billion people. Among which, the average diverted rate is 51.2% and the average induced rate is 61.1%.

IV. CONCLUSIONS

It is large scale work to forecast passenger flow for the whole High-speed railway. In order to ensure forecasting results more realistic, this paper proposes an entire improvement approach using traditional four-step method for reference. The core is to dig the relationship between passenger flow and network structure. By considering the influence of operation scheme to passenger flow, the whole forecasting progress is constrained by total passenger flow. At the same time, parameter calibration adopt the scientific method of clustering by considering different influence factors so that the parameters can really reflect the characteristics of different OD.

This paper has forecast passenger flow for the whole high-speed railway which will provide reasonable reference for passenger transport operation. The forecasting results will be helpful for train operation scheme.

ACKNOWLEDGMENT

This research was jointly supported by the National Natural Science Foundation of China (No. U1434207), the science and technology department of China Railway Corporation (No.2013X014-C, 2014X010-A).

REFERENCES

[1] WANG Hongye, L(u) Xiaoyan, ZHOU Liangjin,et al. Intelligent Seat Allotment Method for Railway Passenger Train Based on Passenger Flow Forecast[J]. China Railway Science, 2013, 34(3):128-132.

[2] YANG Jun, HOU Zhongsheng. A Wavelet Analysis Based LS-SVM Rail Transit Passenger Flow Prediction Method[J]. China Railway Science, 2013, 34(3): 122-127.

[3] M.G. Karlaftis, E.I. Vlahogianni. Statistical Methods Versus Neural Networks in Transportation Research: Differences, Similarities and some Insights [J].Transportation Research Part C: Emerging Technologies, 2011, 19(3):387-399.

[4] ZHAO Changjiang. A Passenger Demand Forecast for Great Railway Corridor Lines[J]. Journal of Transportation Systems Engineering and Information Technology, 2004, 4(3): 92-96.

[5] CHEN Shaokuan, MAO Baohua, DING Yong, CHU Qin. A Computer Aided System for Transportation Demand Prediction[J]. Journal of the China Railway Society,2004, 26(6):1-7.

[6] CAO Yaoqian, LI Xiamiao. Improved four-tage flow forecast method for Wuhan-Guangzhou passenger dedicated line[J]. Journal of Railway Science and Engineering, 2010, 7(3): 109-113.

[7] LI Xiamiao, HUANG Guizhang, TANG Jie. Passenger Flow Forecasting Based On OD-matrix Estimation Mode[J].Journal of the China Railway Society,2008,30(6):7-12.

[8] Shao Chunfu. Transportation planning principle[M]. China Railway Press,2004.

[9] Ennio Cascetta, Pierluigi Coppola. High Speed Rail (HSR) Induced Demand Models[J]. Procedia - Social and Behavioral Sciences, 2014, 111(5):147-156.

[10] CHU Qin, CHEN Shaokuan. On Calibration of the Gravity Model During Traffic Forecasts[J]. Journal of Transportation Systems Engineering and Information Technology, 2003,3(2):51-56.

[11] XIE Ruhe, QIU Zhuqiang, LI Qingyun,et al. Logit Model in the Application of Guangshen Railway Passenger Share Rate Estimation[J]. China Railway Science, 2006, 27(3):111-115.