Study on Railway Locomotive Demand Prediction Model

XIAO Xiang [1], WANG Laixiang [2], XIE Hongwen [1], YE Yujie [4]
1. School of Economics and Management, Beijing Jiaotong University, China, 100044
2. China Railway Corporation, China, 100038
xxiao@bjtu.edu.cn

Abstract: Locomotive is one type of transport carrier, however, its procurement still has some problems, leading to high locomotive idle rate and low utilization rate. In other words, we now do not have a scientific demand evaluation system for locomotive procurement. In this article, the author first analyzes how different locomotive types are related to passenger and freight transport volume, their running speed, and transport time, and then gets the demand prediction formulae for passenger trains, freight trains, locomotives, and China Railway High Speed (CRH) respectively. Then the author uses the data from 1978 to 2012 to build unary and binary linear regression models for locomotives and improves the model for evaluating the procurement size of locomotives, hoping this research achievement can be an important reference for China Railway Corporation, road bureaus and companies during their decision-making for locomotive procurement.
Keywords: Railway locomotive, Demand analysis model, Transport organization, Linear regression model

1 Introduction

In China, railway is an important national infrastructure and also the backbone of the transportation system. It is the transport means that accords with our economic and geographic conditions as well as the income levels in China. As the transport carrier, railway locomotive is the material base of railway transport, however, its procurement still has some problems, the idle rate is high and utilization ratio is low. The evaluation model for locomotive procurement is quite special, and there is no normative evaluation model yet. In March 2013, the national railway begins to separate administrative functions from enterprises and China Railways Corporation is established to take the corporate functions. In this context, the Corporation is in urgent need for a way to properly predict the procurement size for railway locomotives, so that it can reasonably allocate funds and improve its economic performance. This article analyzes the prediction model for different locomotives based on transport organization relations and also uses linear regression analysis method to assist the prediction.

2 Transport-Organization-Based Prediction Model

According to the factor correlation principle of the transport organization, this article analyzes how different transport vehicles are related to passenger and freight volume and their transport time, and then gets the demand analysis models for major railway locomotive types, including passenger trains, freight trains, locomotives and CRHs. These models are elaborated in the following sections.

2.1 Passenger trains

2.1.1 Prediction based on number of opening passenger train lines

Passenger carriage = (2 * Moving distance/Travel speed + stay time in outside of passenger carriage) * Traffic Density/24 (1)

2.1.2 Prediction based on passenger traffic turnover

Based on the ratio between railway passenger traffic turnover over the years and actually owned train number, this article performs a recursive analysis and gets the proper passenger train number to be owned on short and long terms. The calculation formula is as follows:

P =Z/Q (2)
P — Passenger train number owned on short and long terms;
Z — Passenger traffic turnover on short and long terms;
Q — Ratio of passenger traffic turnover completed by each passenger train to the total turnover on short and long terms.

2.2 Freight trains
Number of freight trains to be purchased = Number of owned freight trains on short and long terms - Number of currently owned freight trains + Number of freight trains scrapped in current stage.
2.2.1 Number of freight trains owned on short and long terms

$$Number\ of\ owned\ freight\ trains = \frac{Annual\ traffic\ volume \times Turnover\ days\ of\ freight\ trains}{365 \times Static\ train\ load} \times \frac{1}{1-Repair\ rate} \times \frac{1}{1-Reserve\ rate} \times Fluctuation\ coefficient \quad (3)$$

Where:
Annual traffic volume - Freight volume prediction on short and long terms;
Static train load — Goods tonnage every freight train loaded on average in a certain period (one year);
Repair rate — Rate of trains under repair to the total number of owned trains;
Reserve rate — Rate of reserved trains to the total number of owned trains;
Fluctuation coefficient — Coefficient set to compensate the uncertainty of traffic volume prediction, static load of freight trains and turnover days of freight trains, and is set to 1.1.
2.2.2 Number of owned freight trains in current stage
This data can be acquired from statistics.
2.2.3 Number of scrapped freight trains

$$Number\ of\ scrapped\ freight\ trains = \sum_{ij} Q_{ij} x_{ij} \quad (4)$$

Q_{ij} —Number of i type freight trains owned in j year;
X_{ij} —Ratio of i type of freight trains passing service life in j year to the total number of i freight trains owned in that year.

2.3 Locomotives
2.3.1 Daily production method
Based on the ratio between all-lines transport turnover and the daily transport volume of the locomotive, this article calculates the proper locomotive demand of the entire railway lines.
Freight locomotive = (All-lines freight turnover volume/(Daily transport volume of freight locomotives * 365)) * A (5)
A -- Ratio of gross freight weight (ton-kilometer) to freight turnover
Passenger transport locomotive = All-lines passenger transport turnover/(Daily transport volume of passenger transport locomotives * 365) (6)
Daily transport volume of passenger transport locomotives = Passenger transport turnover * Average car-kilometers per car-day / Total running kilometers of passenger transport locomotives
2.3.2 Prediction based on turnover volume
Based on the ratio between all-lines converted turnover and actually owned train number, this article performs a recursive analysis and gets the proper locomotive number for the plan year.
Number of owned locomotives = Converted turnover/Q (7)
Q = Ratio of converted turnover to number of owned locomotives, in other words, it is the average converted turnover completed by every owned locomotive.

2.4 CRHs

Number of CRHs demanded = Total turnover time / Travel time of CRHs * (1 + Repair rate) * (1 + Chonglian ratio) (8)

Where:

Total turnover time of the locomotives refers to the time that the locomotives take to complete one turnover (exclude the idle time)

Total turnover time = Operating mileage/Travel speed * Initially number of opening lines * 2 + go and back time * (2 * Initially number of opening lines - 1) (9)

With these formulae, we can directly predict the number of locomotives to be purchased based on the actual conditions, however, the data used to deduct these formulae are historical data or the average data of the whole industry, the calculation may be not applicable to specific condition of individual enterprises. In this case, other methods shall be used to assist the prediction. According to historical data analysis, there is a significant linear relation between number of owned railway locomotives and some other factors. Therefore, a linear regression model can be used to assist the prediction.

3 Study on Linear Regression Model

The operating history of high-speed rail and passenger transport line is quite short, the statistics about CRHs is quite limited, therefore, in this article, the author uses the data from 1978 to 2012 of railways in china to perform a regression analysis for passenger trains, freight trains and locomotives.

3.1 Linear regression model for passenger trains

After the literature studies, the author selects several independent variables, including operating mileage, passenger transport volume, passenger traffic turnover, population, and consumer price index (CPI). The analysis with spss software shows that there is a collinearity relation among these variables. After we remove the independent variables with insignificant linearity, only two are left, which are operating mileage and passenger traffic turnover. The scatter diagrams are as follows:

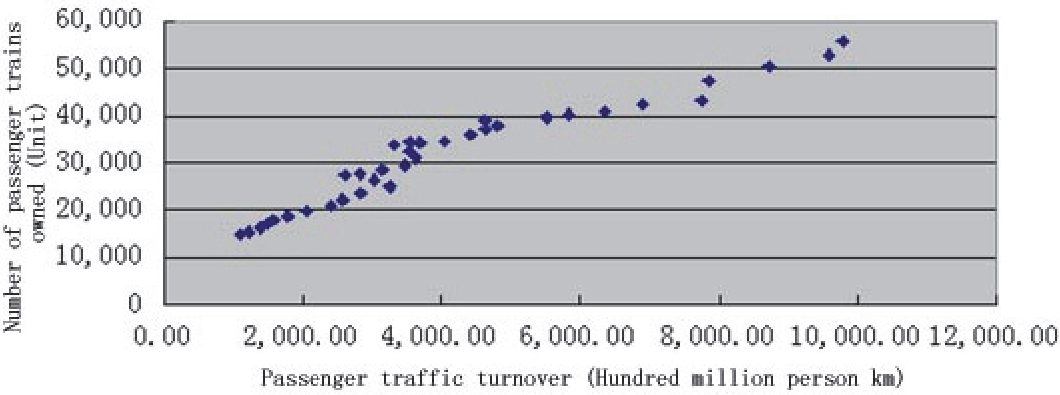

Figure 1 Scatter diagram between number of owned passenger trains and passenger traffic turnover

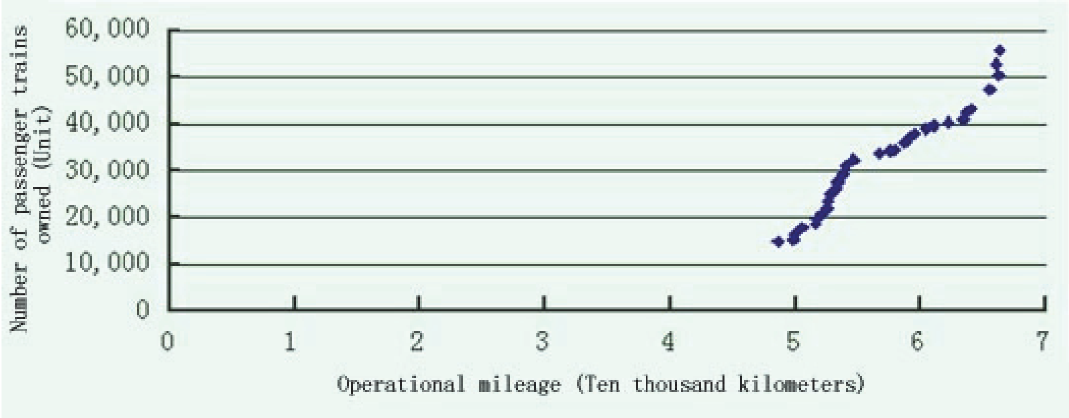

Figure 2 Scatter diagram between numbers of owned passenger trains and operating mileage

As shown in Figure 1 and Figure 2, the number of owned passenger trains, passenger traffic turnover, and operating mileage are linearly related, and the scatter diagram shown in figure one (between number of owned passenger trains and passenger traffic turnover) is smoother.

Take number of owned passenger trains as dependent variable and passenger traffic turnover as independent variable, we can get a unary regression equation:

Number of owned passenger trains = 13 228.399 + 4.48 * Passenger traffic turnover (10)

Adjusted R square = 0.925, F = 418.169 > F (1, 32), pass the F test. t = 20.449 > $t_{0.975}$, pass the T test, and there is a significant linearity relation.

Take number of owned passenger trains as dependent variable, take passenger traffic turnover and operating mileage as independent variables, and use spss software to get a binary regression model:

Number of owned passenger trains = -53 804.106 + 14 092.842 * Operating mileage + 1.383 * Passenger traffic turnover (11)

Adjusted R square = 0.957, F = 378.896 > F (2, 31), pass the F test. Passenger traffic turnover t = 2.187 > $t_{0.975}$, Operating mileage t = 5.076, pass the T test, and there is a significant linearity relation.

Compare the actual application data with prediction data, the error rate remains less than 3.5%.

3.2 Linear regression model for freight trains

The author also uses SPSS software to analyze the railway data from 1978 to 2012 in China and gets a scatter diagram. As shown in the diagram, there is a linear relation between the number of owned freight trains and operating mileage as well as between the number of owned freight trains and freight turnover.

Take number of owned freight trains as dependent variable and freight turnover as independent variable, we can get a unary linear regression equation, as shown below:

Number of owned freight trains = 171 582.355 + 18.313 * Freight turnover (12)

Adjusted R square = 0.975, F = 1 302.285 > F (1, 32), pass the F test. t = 36.087 > $t_{0.975}$, pass the T test, and there is a significant linearity relation.

Take number of owned freight trains as dependent variable and take freight turnover and operating mileage as independent variables, we can get a binary regression model:

Number of owned freight trains = -261 211.189 + 94 897.498 * Operating mileage + 10.625 * Freight turnover (13)

Adjusted R square = 0.983, F = 990.513 > F (2, 31), pass the F test. Freight turnover t = 5.679 > $t_{0.975}$, Operating mileage t = 4.213, pass the T test, and there is a significant linearity relation.

The numbers of owned freight trains predicted by unitary and binary regression prediction formulae are similar. Compare the actual application data with prediction data, the error rate remains less than 3%.

3.3 Linear regression model for locomotives

As we learn from SPSS scatter diagram, the number of locomotives, freight trains, and passenger trains are significantly linearly related. Therefore, take the number of owned locomotives as dependent variable and take number of owned passenger trains and number of owned freight trains as independent variables, and then we can get the linear regression formula:

Number of owned locomotives = 6 102.170 + 0.154 * Number of owned passenger trains + 0.008 * Number of owned freight trains (14)

Adjusted R square = 0.971, F = 569.803 > F (2, 31), number of owned passenger trains, Number of owned freight trains t=3.111 and t = 2.753, > $t_{0.975}$, pass the T test, and there is a significant linearity relation. As from the comparison between the predicted results for the last four years and the actually owned number, the error rate remains less than 5%. The average error rate remains less than 2%.

3.4 Linear regression model evaluation

Advantage: With multiple-variable linear regression model, we can directly and speedily get the linear relation among the variables by inputting multiple groups of data. Regression analysis can accurately measure the fitting and correlation degrees among different elements, and can help to improve the accuracy of the prediction result.

Disadvantage: The analysis may neglect the causation and nonlinear relations, and the prediction result may be influenced by sample selection. During the regression analysis, the selection of factors and the factor expressions are only determined by prediction, which restricts the factor diversity. Besides, some factors are immeasurable, and both have restricted the application of regression analysis in certain conditions.

From the linearity significance degree and the prediction result for recent years, we conclude that the prediction result of linear regression model is reasonable.

4 Conclusion

In practice, we can use the formula method to perform prediction from multiple aspects. The test shows that linear regression prediction model has significant linear feature, prediction precision is high, and error rate remains less than 5%, therefore, linear regression prediction can be an important auxiliary prediction method. And if we take average or the weighted average of multiple prediction results, we can get a proper procurement number, helping us to maximize the usage of capitals, to develop railway transport in a more efficient and scientific way, to avoid waste of social resources, and to protect legal rights and interests of taxpayers.

Of course, this article only predicts the total number of locomotives, passenger trains and freight trains due to lack of specific statistic data. The future researches can collect the statistic data of specific locomotive types, and they can also include the influences of technology and policy to further improve prediction accuracy and practicability.

Acknowledgements:

We thank the financial supports from the National Nature Science Foundation of China (71372012) and the National Nature Science Foundation of China (71272055).

References

[1]. D·R·Burns. Demand Prediction for Diesel Locomotive and Electric Locomotive [J]. Foreign Diesel Locomotive. 1994 (09)

[2]. James Abbott, David Burns, Hu Guangzhi. Prediction Report for Locomotive Market (the Third Part) — Locomotive Demand Prediction [J]. Foreign Diesel Locomotive. 1999 (05)

[3]. Lyonso, Jain J, Holley D. The Use of Travel Time by Rail Passengers in Great Britain [J]. Trans

Portation Research Part A: Poliey and Praetiee. 2007, 41 (l)
[4]. Dieter Wolfgram* (German). Satisfying Market Demand for High-Performance Locomotives [J]. Foreign Diesel Locomotive. 2004 (06)
[5]. Bai Yaoyao. Prediction of Bus/Coach Market Based on Second Exponential Smoothing Forecasting Method [J]. Bus & Coach Technology and Research. 2013, 3 (in Chinese)
[6]. Qin Lu. Fuzzy Evaluation of Purchasing Risks of Passenger Transport Technical Equipment in Railway: Great Leap Forward Development [J]. Logistics Technology. 2011, 2 (in Chinese)