Short-term Passenger Flow Prediction of Urban Rail Transit based on ARIMA Model

Junqing Chen*

Transport College, school of Chongqing Jiaotong University, Chongqing, 404100, China

* Corresponding Author Email: 632103090511@mails.cqjtu.edu.cn

Abstract. Urban rail transit is very important in modern urban transportation, which meets the needs of People's Daily travel. This paper aims to introduce the ARIMA model into the short-term passenger flow forecast of urban rail transit, analyze its principle and application method, and build ARIMA model combined with AFC data to forecast short-term passenger flow of urban rail transit. Through the research and improvement of ARIMA model, it is anticipated that it would give urban rail transport operators more precise and trustworthy short-term passenger flow forecasting capabilities, so as to realize more intelligent and efficient operation of urban rail transit system. In this study, the ARIMA model was established, and rigorous scientific methods such as residual test, autocorrelation function (ACF) and partial autocorrelation function (PACF) test, and comparison of predicted value and real value were adopted to conduct tests step by step according to the experimental process. Finally, it was confirmed that the excellent prediction accuracy of the ARIMA prediction model used in this paper, can provide accurate prediction results, and can be used in practical scenarios.

Keywords: Urban rail transit, ARIMA model, AFC data, short time forecast.

1. Introduction

In the past few years, urban rail transit has gradually become the travel choice of more and more citizens. Its advantages of green environmental protection, convenience and cheap travel cost have won passengers' favor [1]. Nowadays, the urban transportation system has included the urban rail transit as its indispensable component and is an indispensable "aorta" of the urban transportation system. Subway, light rail, tram, monorail and other means of transportation in recent years in the world's major cities have also been a huge number of popularity, the growing urban population is also contributing to the high number of passengers. Therefore, urban rail transit fills everyone's life.

At present, China's urbanization process is speeding up, making the contradiction between urban transport infrastructure and urbanization development become more and more obvious. With the sustained and steady development of Chinese society, the process of industrialization and urbanization has accelerated, and the level of development has been constantly improved, and people's living standards and quality have continued to improve, and there are increasingly high requirements for transportation travel [2]. A huge increase in the number of motor vehicles also leads to the contradiction between urban transportation facilities and urbanization. According to the statistics of the Traffic Management Bureau of the Ministry of Public Security [3]. A total of 79 cities in the country have more than one million vehicles, among which Chengdu, Chongqing and Beijing have more than 5 million. Therefore, the development, construction and good operation of urban rail transit is the "main weapon" to solve the above contradictions.

At present, China is already the fastest developing country of urban rail transit in the world [4]. Both major cities, such as Shanghai and Beijing, and other transportation hub cities, such as Xi 'an and Nanjing, have made strong support for China. In last several years, the pace of urban rail transit construction has continued to accelerate. Data show that five new operating cities have been added: Changzhou, Xuzhou, Wenzhou, Jinan and Hohhot. 968.77 kilometers of new routes were added, and the passenger volume exceeded 24 billion passengers, another record high [5]. The scale of rail transit operation is gradually increasing.

Short-term passenger flow is also often referred to as "real-time passenger flow", although there is no clear definition of the two, but the former generally refers to the passenger flow with a particle

size of 5 to 15 min, and the latter refers to the passenger flow with a particle size of 5 min [6], but the study believes that the expression of "real-term passenger flow" is less accurate than "short-time passenger flow". Generally speaking, passenger flow data with weak regularity will change greatly with the occurrence of uncertainty error [7]. Urban rail transit system has accurate network nodes and paths, but the number of inbound and outbound ridership is unknown. At present, there are many cases of using a single model to predict traffic flow: In 2007, Shekhar et al. proposed an adaptive parameter estimation method to predict traffic conditions by using three filtering technologies, namely Kalman filter, recursive least squares estimation and least mean square estimation [8]. In 2015, Kumar et al proposed a short-term traffic flow prediction method based on seasonal ARIMA model, which could predict the traffic flow in different seasons [9]. In terms of rail transit, Liu, Jiao et al. made predictions by establishing K-nearest neighbor model and other methods [10]. In terms of road transportation, Ma et al. used an interactive method based on hybrid models to predict short-term bus traffic demand.

Domestic and foreign scholars have developed many kinds of forecasting methods, and have drawn effective conclusions, but the research methods are very single. At present, the state is vigorously promoting urban construction, more and more data will appear, requiring deeper research. In this paper, several methods are integrated, using ARIMA model, and based on AFC data, from the aspects of "horizontal factors" and "vertical factors". This paper aims to introduce ARIMA model in short-term passenger flow forecasting of urban rail transit.

2. Method

2.1. Data acquisition channels and descriptions

This paper takes Chongqing as an example, so it adopts the information of Chongqing urban rail transit system. The experimental data comes from the data website of Chongqing Transportation Bureau, which contains more than 150 sample sizes and records the total volume for 23 weeks and 161 days from January 1, 2023 to June 11, 2023.

2.2. Indicator Selection and Description

Chongqing rail transit passenger flow is relatively dense, large flow of people, convenient for research. Therefore, the daily inbound traffic data of Chongqing rail transit is collected, and the operating period of rail transit is from 06:00 to 23:30, which can be obtained by analyzing the original data. Each data contains the information such as the start time, input flow, output flow and end time. Figure 1 shows the daily arrivals per week.

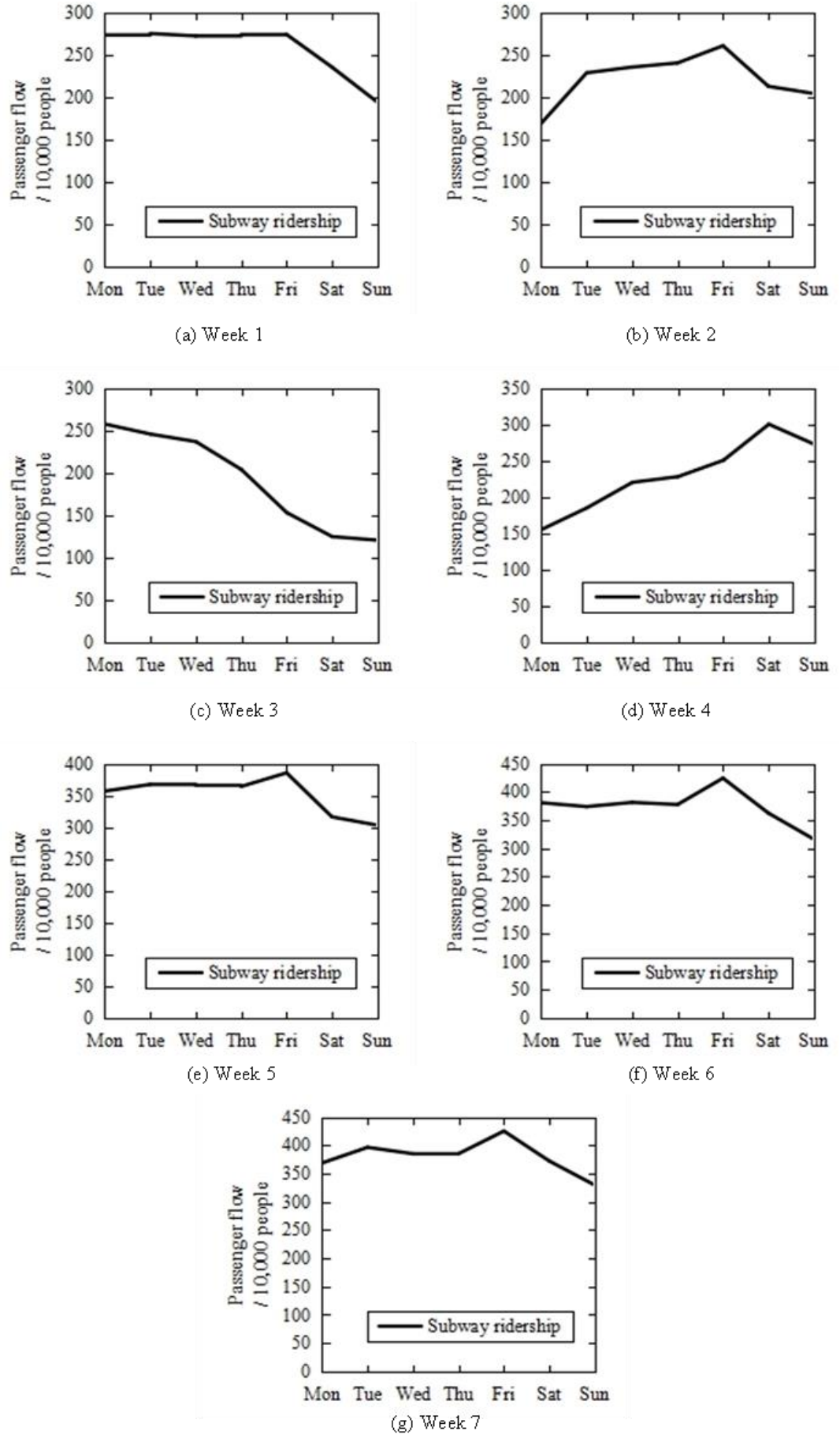


Fig. 1 Daily inbound passenger flow per week

2.3. Method Introduction

The main detection method is the stationarity detection of the model, which is usually to detect the early data values of the fitted model, in order to determine whether these data can present a stable state. The method of autocorrelation function test is used, because this method is based on the fact that the greater the distance between the data and the real data in the previous time breakpoint interval, the smaller the interference to the real value. With the gradual extension of time, the parameters of the autocorrelation function will gradually decrease under different time nodes and finally decrease to zero. Then it is proved that the data fitted by the model can present a stable state (Figure 2).

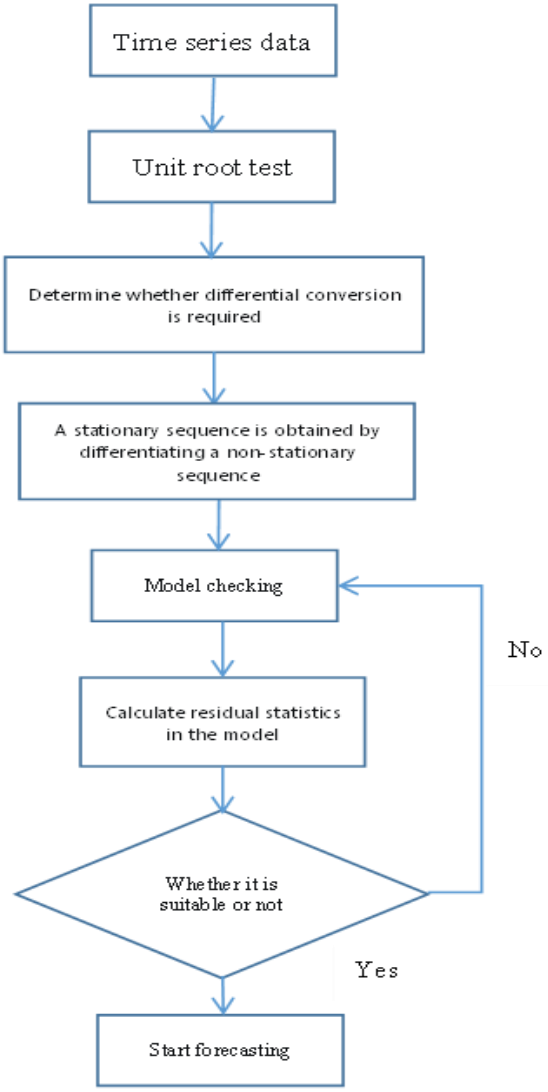


Fig. 2 Flow chart

3. Results and Discussion

3.1. The Stationarity Test of the Model

Before using the ARIMA algorithm, we need to test the stationarity of the training set. Because this algorithm is only suitable for time series with stable characteristics. The unit root test method is adopted, and the ADF diagram is used for the test. By observing the significance of the second part, we can judge whether the time series tends to be stable. When the value obtained is less than 0.5, it means that the unit root has a solution, and further. It indicates that the time series has stable characteristics. When the result exceeds the specified value, it means that the unit root is not solved, then it means that the time series does not have stability characteristics.

3.2. Predictive Model Test

After the model is constructed, it can be predicted. Before the prediction, it is also necessary to check whether the model can be predicted correctly and whether the residual symbol of the constructed model does not conform to the characteristics of normal distribution. The constructed model residuals obtained after bringing the data into the model are shown in the Figure 3.

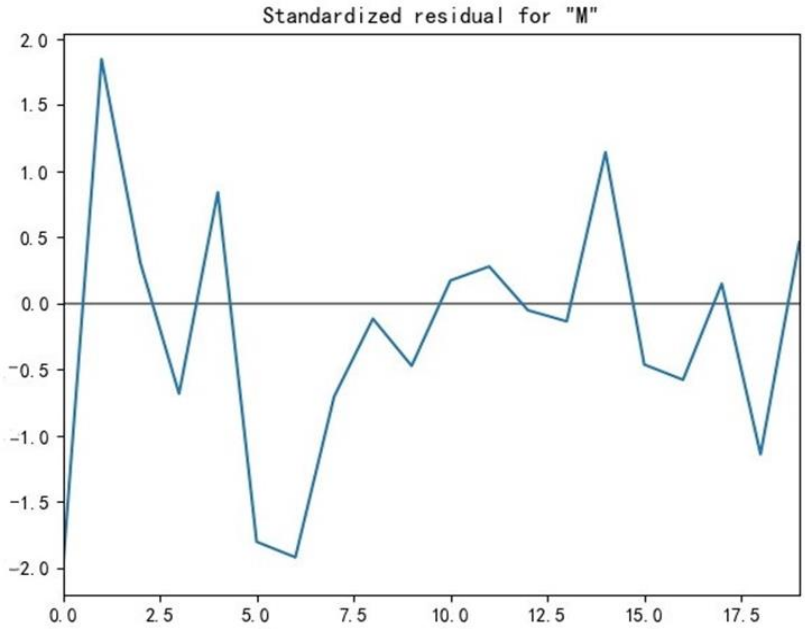


Fig. 3 Standardized residual for model

It can be seen from the chart that the residual is not correlated, which proves that the constructed model has effective characteristics and can fit the data for the following data prediction. Figure 4 shows the residual Q-Q distribution of the prediction model.

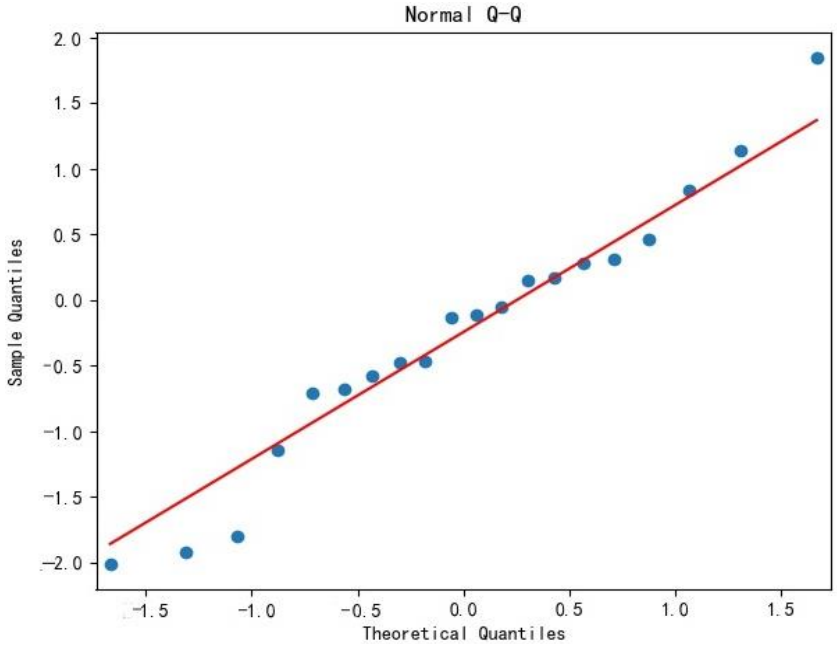


Fig. 4 Normal Q-Q

It can be clearly seen from the figure that the residual distribution basically conforms to the trend and characteristics of the normal distribution. The residual ACF and PACF of the generated prediction model are shown in the Figure 5.

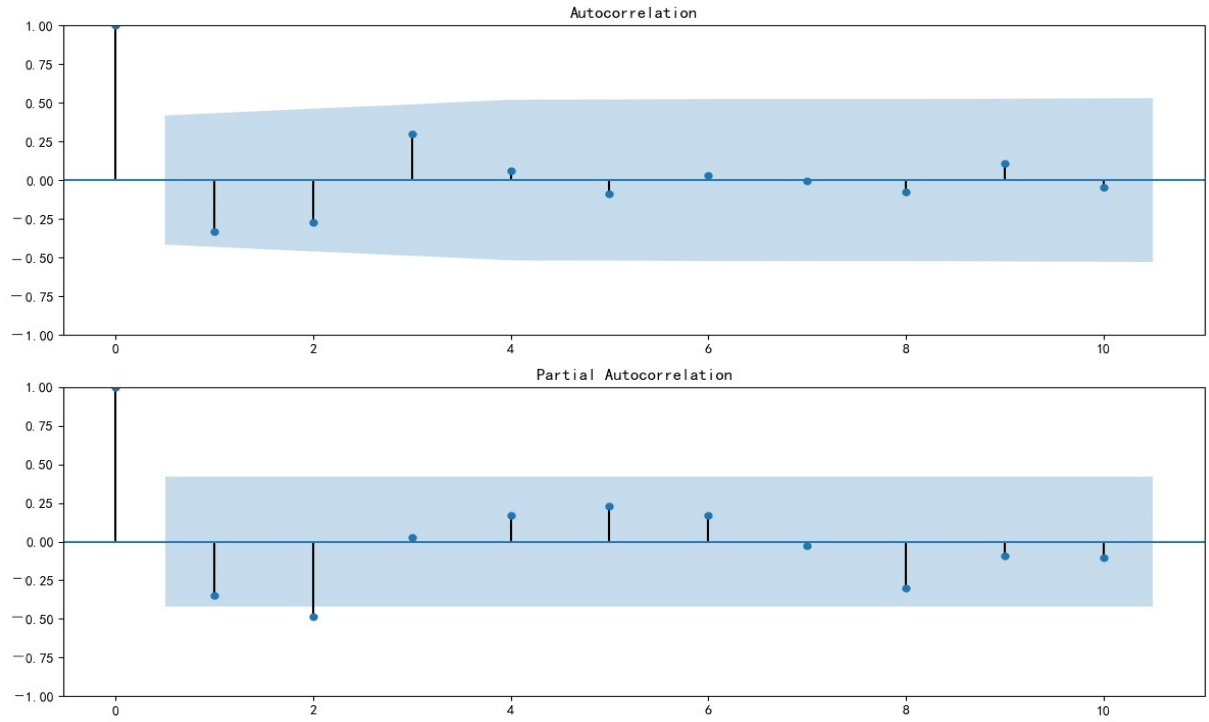


Fig. 5 Residuals ACF and PACF

After compared with the residual sequence's autocorrelation and partial autocorrelation functions on the prediction model, the two test results basically meet the requirements, so the model can make the next prediction.

3.3. Short-term Passenger Flow Prediction

When the model achieves the first prediction, the data in the test set following the time series will be imported into the training set. After the import, the training set data will be updated and the next model run will be started again. The next data prediction will be made based on the latest training set data and test set data. It is contrasted with actual passenger flow statistics. The model's predicted outcomes are displayed in Figure 6.

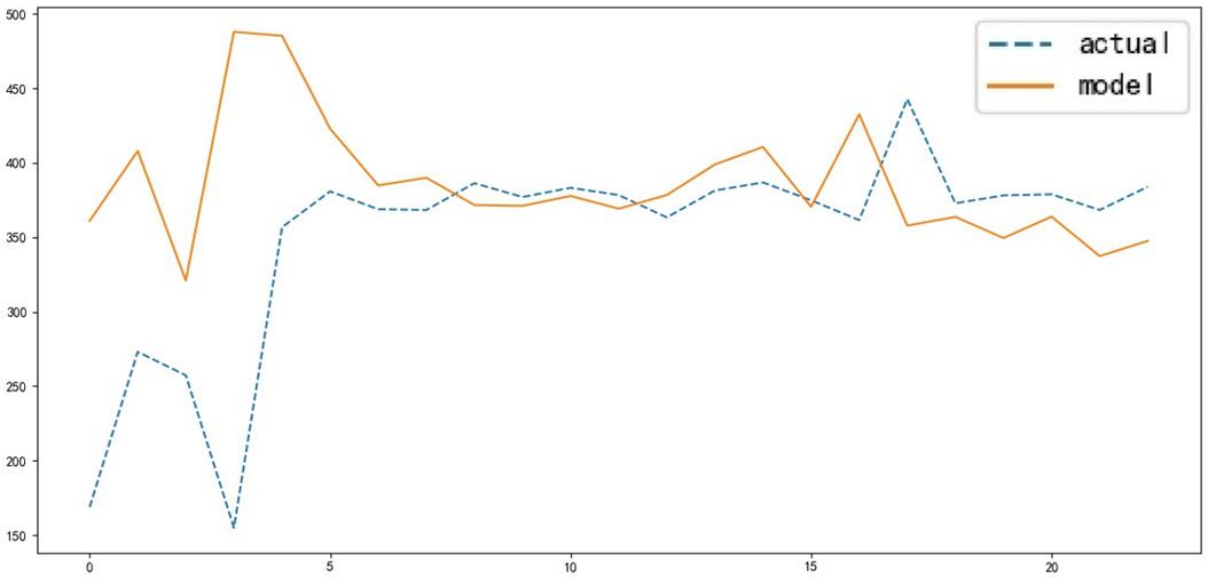


Fig. 6 The Predicted Value Compared with the True Value

According to the comparison chart, it can be intuitively seen that the deviation between the two values is very small, indicating that the model has strong prediction ability, which proves that the

method has certain practicability and applicability, and can be put into relevant research and application.

4. Conclusion

To sum up, urban rail transit's short-term passenger flow projection is crucial for maximizing the operation plan, improving service quality and meeting passenger demand. By modeling the trend, periodicity and seasonality of a huge number of passenger flow data, ARIMA model can accurately predict experimental objects and results. The advantage of this model is that it can use the existing data information to make predictions, and it is suitable for stable and linear relationships. Through applying the ARIMA model, urban rail transit operators can better understand the ridership in different periods, and formulate corresponding operation plans and resource allocation strategies.

This paper studies the problems existing in the short-term passenger flow prediction of urban rail transit system, and decides to establish ARIMA prediction model to discuss the problems after deep thinking. At each stage of the study, the operational analysis was carried out in strict accordance with precise steps, and different methods were tested for each study result. Finally, through case study and comparative analysis, the following conclusions are drawn: ARIMA prediction model has high prediction accuracy, can provide accurate prediction results, and can be used in practical scenarios.

References

[1] Teng Teng, Liu Zhengqi, Wang Xiaomin. Short-term passenger flow prediction of urban rail transit based on GA-SLSTM model. Railway Computer Applications, 2022, 31(8): 7-12.

[2] Liu Yaohong. Research on short-time Inbound passenger flow prediction of rail transit based on Combination model. Chongqing Jiaotong University, 2022.

[3] National Development and Reform Commission, PRC. Urban Rail Transit 2020 Annual Statistics and Analysis Report. Beijing: China Urban Rail Transit Association, 2019.

[4] Zeng Cheng, Wu Jiayuan, Luo Xia. Urban rail transit short-term prediction literature review. Journal of railway transportation and economy, Lancet, 2021, (8).

[5] Jin F. A study on network of domestic air passenger flow in China. Geographical Research, 2001, 20(1): 31-39.

[6] Zhang Guoyun, Jin Hui. Research on short-term passenger flow Prediction of Urban Rail Transit based on Improved ARIMA model. Computer Application and Software, 2022, 39(1): 339-344.

[7] Shekhar S, Williams B M. Adaptive seasonal time series models for forecasting short-term traffic flow. Transportation Research Record, 2007, 116-125.

[8] Kumar S V, Vanajakshi L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. European Transport Research Review, 2015, 7(3).

[9] Liu Meiqi, Jiao Pengpeng, Sun Tuo. Research on short-term prediction model of inbound passenger flow of urban rail transit. Urban rail transit research. 2015.

[10] Ma Z, Xing J, Mesbah M, et a1. Predicting short-term bus passenger demand using a paRem hybrid approach. Transportation Research Part C: Emerging Technologies, 2014, 39: 148-163.