Analysis of Urban Rail Transit Short-term Passenger Flow Forecast and Its Optimization

Ruxue Bai[1, *], Zhongrui Ren[2] and Xiongjie Tang[3]

[1]School of Information Engineering, Huaxia University of Technology, Wuhan, China

[2]Information Management and information systems, Neusoft University, Chengdu, China

[3]Internet of Things Engineering, Beijing University of Posts and Telecommunications, Beijing, China

* Corresponding Author Email: 21021320118@stu.nus.edu.cn

Abstract. With the growth of urban population and the increase of car ownership, the number of cars continues to increase, leading to traffic congestion becoming a major problem in many cities around the world. Traffic congestion not only wastes time, but also increases air pollution and energy consumption. To address these challenges, governments and transportation authorities around the world have begun to adopt advanced technical models and big data based approaches. By using the power of intelligent data analysis, this paper enables people to understand rail transit more deeply, establish an ARIMA model for short-term prediction, optimize traffic flow, and improve the overall performance of the system. Through the final experimental results, it is found that the trend of rail transit passenger volume in the next few days is increasing, but its growth trend is not as high as that of the previous few days. The ARIMA model infers that the peak period in the next few days may not be as high as that of the previous few days.

Keywords: Subway, ARIMA, short-time passenger flow.

1. Introduction

Along with the fast development of URT, URT has been the first choice for the majority of passengers, and URT is one of the most effective ways to ease the pressure of urban transportation. It is very important to predict the short term passenger flow of UMT, which affects the security, efficiency and operation cost of URT [1, 2]. In order to predict the main factors affecting subway traffic and optimize the service level of each period, it is necessary to accurately forecast the passenger flow of each period and each region. Finally, the paper proposes the forecast and optimization of the influencing factors of subway passenger flow [3].

The accurate forecast of UMT can alleviate the traffic congestion, and provide the city residents with quicker and better travel service. Passenger flow data with weak regularity and strong interference of random errors for short-term passenger flow prediction, ARIMA model can predict the delay of dependent variables, the lag value generating random errors and the current value [4]. To improve the forecast precision, this paper adopts the method of wavelet to denoise the passenger flow, and eliminates the tendency and season of the data with the different method. A comparison of the forecast value of the incoming passenger flow and the actual traffic flow shows that the forecast results are better and more applicable [5].

Passenger flow forecast is the foundation of URT planning, design and operation management. It is also an important component of URT construction [6]. In this paper, the DWT method is applied to analyze the daily traffic data of metro stations. Based on the decomposed wavelet coefficients of each layer and the AIC Akaike information standard, the ARIMA model is used to forecast the future passenger flow data. The results indicate that the proposed method can achieve a higher precision in predicting passenger flow in metro area [7].

Automatic fare collection based on urban rail transit (AFC) the historical data of inbound and outbound passenger flow collected by the system is used to construct the product difference autoregressive moving average (ARIMA) model, to achieve the accurate forecast of inbound and outbound passenger flow. Based on self-correlation and partial auto-correlation function, the effects

of trend and periodic characteristics are eliminated. Considering the product relationship between the periodicity and short-term correlation of the processed data, a product ARIMA passenger flow prediction model is constructed. The results indicate that this model is highly accurate and applicable [8].

High accuracy short time incoming passenger flow forecast plays an important role in UMT. It is more timely and advanced to use the results of passenger flow prediction to implement measures such as flow restriction and dredging in advance than to control afterwards. By collecting the data of subway inbound passenger flow at 15 min intervals, using the inbound volume of the same period last week, the inbound volume of the previous period of this day, and the parameters of peak and off-peak hours as input variables, short term forecast was carried out by means of weighted historical mean autoregression model, ARIMA model and WNN model. The results indicate that the combined forecast model can be used to forecast the incoming passenger flow in UMT [9, 10].

The passenger flow order of the metro is unstable and volatile, so it is hard to forecast it accurately with the conventional method. In order to improve the forecasting precision of metro passenger flow, the paper uses wavelet analysis to stabilize the passenger flow data, and then constructs the ARIMA to predict [11]. Taking into account the advantages of high passenger volume, safe environment and so on, the paper uses this method to forecast the passenger flow in Beijing metro station. Accurately analyze the forecast results, make reasonable arrangements for the service efficiency of each phase, provide reasonable support for the public facilities, and enhance the professional and refined level of public services.

2. Methods

2.1. Data Source

This paper uses the swipe card data of passengers of Beijing Line 1 from May 12 to 17, 2018, including the number of people entering, leaving and entering stations of all stations of Line 1.

2.2. Indicator Selection and Description

The swipe data of passengers can well record the information of subway passengers entering and leaving the station every day, and the accuracy is very high. According to the characteristics of the passenger card swipe data from May 12 to 17, the research data is more authoritative.

2.3. Introduction to the Method

2.3.1 ARIMA model

ARIMA Model is one of the most popular and widely used methods for forecasting time series. In ARIMA (p, d, q), AR is "autoregression", P is autoregression term, MA is "moving average", q is the number of terms in the moving average, and d is the number of variations (orders) that make it stationary.

ARIMA is used to make the model conform as closely as possible to the special form of time series data, so as to better predict passenger flow. In conclusion, ARIMA can be used as an effective method to forecast subway passenger flow.

2.3.2 Wavelet Neural Network

WNN is a kind of new ANN. It is a novel ANN model based on the theory of wavelet and wavelet. Instead of the normal non-linear sigmoid function, the signal is expressed by the linear superposition of the selected wavelet base. The method can avoid the blindness of BP neural network and the local optimal nonlinear optimization problem, so it is easier to train, and has a better ability to learn and generalize.

The results of computer simulation indicate that the combined model has better forecasting precision than ARIMA and WNN models. The combination model can improve the forecast precision of passenger flow, and it is an efficient way to predict passenger flow.

2.3.3 ARIMA-LSTM model

ARIMA-LSTM Model ARIMA and LSTM are both classical models for time series prediction. ARIMA is a statistical method based on differential and autoregressive moving average models that can be used to capture trends and seasonality in time series. LSTM is a neural network-based model that can make predictions by learning long-term dependencies of time series.

To sum up, combining ARIMA and LSTM can form an ARMI-LSTM hybrid model, which can make better use of the advantages of ARIMA and LSTM and improve the accuracy of time series prediction.

3. Results and Discussion

3.1. Descriptive Analysis

The short-term passenger flow statistics mainly come from the card data of Beijing Metro Line 1. Select a working day to exclude the special situation of excessive passenger flow on holidays. Taking the passenger swiping data of Tian'anmen East Station on May 17 as an example, the sampling interval of the data is 30 minutes, and the number of people entering the station, the number of people leaving the station and the number of people entering and leaving the station are calculated and plotted as shown in Fig 1.

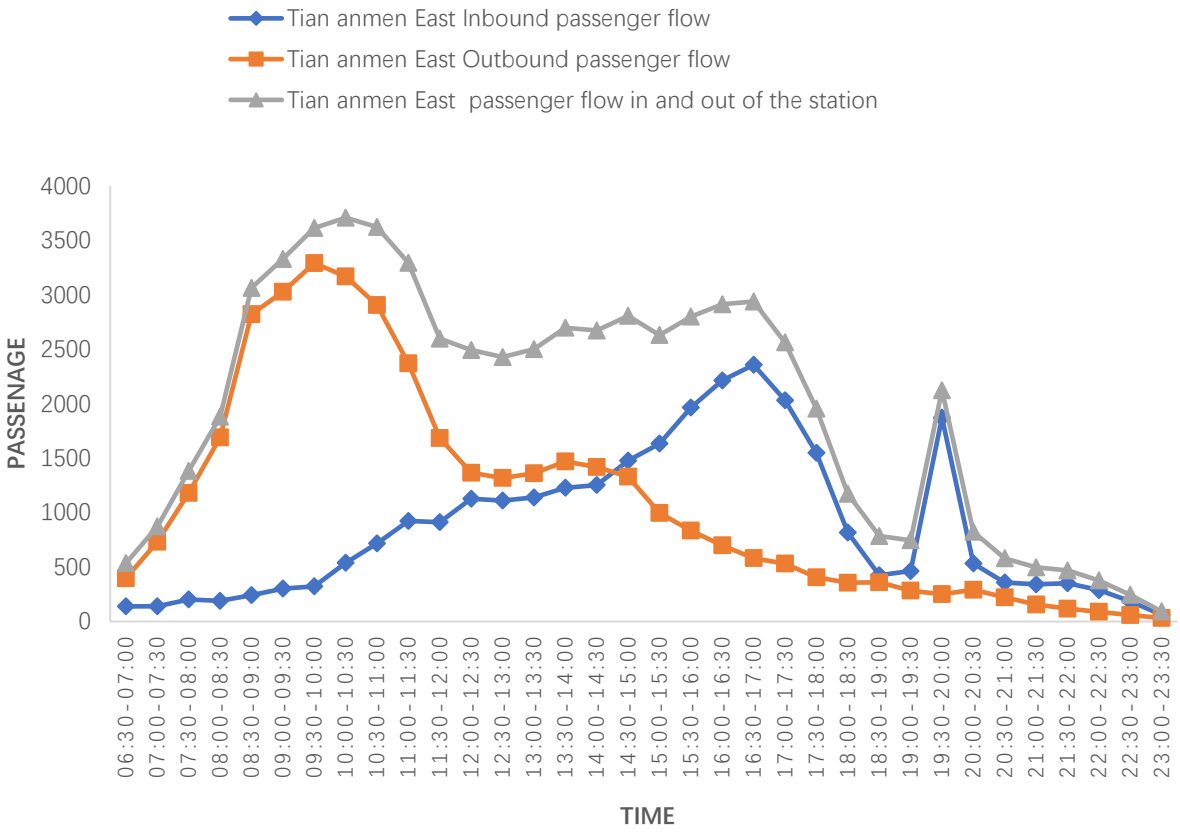


Fig. 1 Passenger flow distribution diagram of Tian'anmen East Station on weekdays.

The passenger flow of Tian'anmen East Station is calculated by using the credit card data of Tian'anmen East Station of Beijing Line 1. The results are shown in Fig 2. The height of the rectangle indicates the size of the passenger flow. From the graph, you can see that it gradually increases from 6:30 in the morning, peaks around 10 o'clock, and then starts to decrease. There are two peak passenger flows in the whole day, respectively at 10:00 and 17:00, which is consistent with the actual situation.

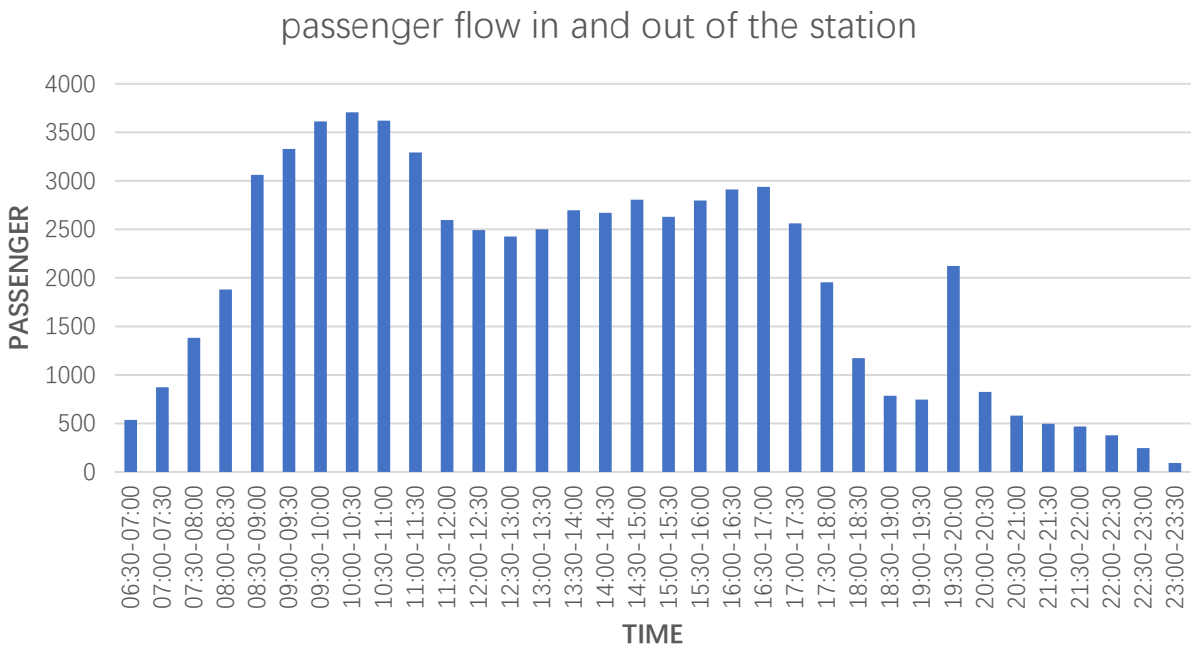


Fig. 2 Entry-exit Passenger Flow of East Tian'anmen Station in the Whole Day

The data of passenger flow in and out of the station from May 12 to 16 were selected for statistics and drawn as shown in Figure 3. After statistical analysis of the subway data on weekdays, it can be clearly observed from the figure that Monday is the day with the highest passenger flow in a week, which is very consistent with the actual situation.

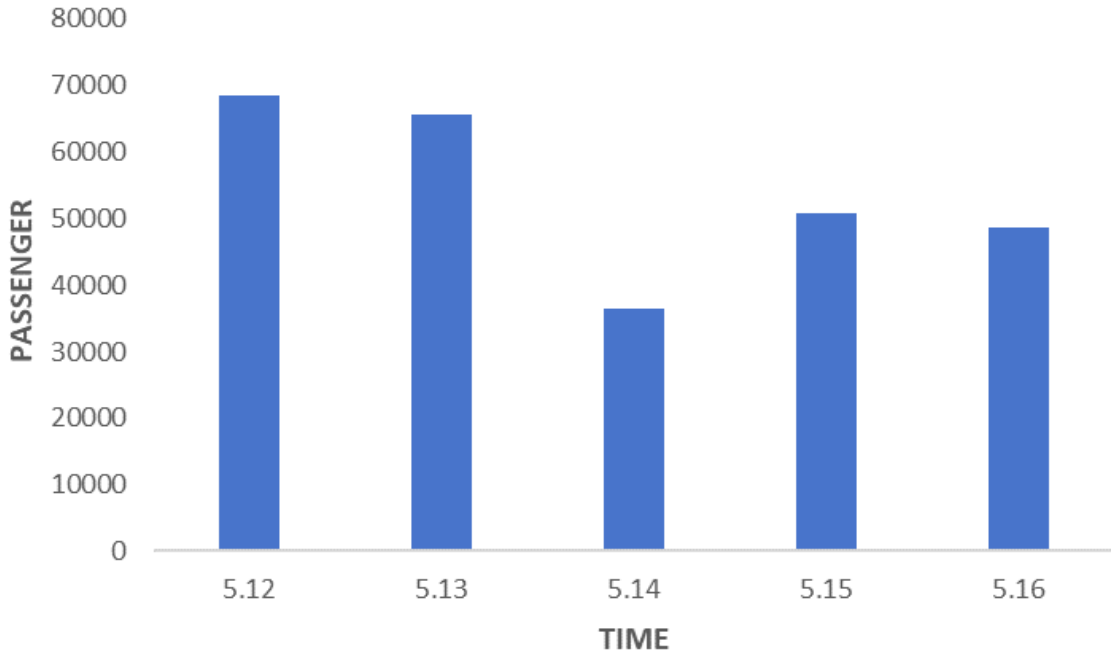


Fig. 3 Passenger Flow in Tian 'anmen East Station on weekdays

3.2. Inferential Analysis

As shown in Figure 4, the distribution of the flow of people at the entrance and exit of Tian 'anmen East Station on the timeline.

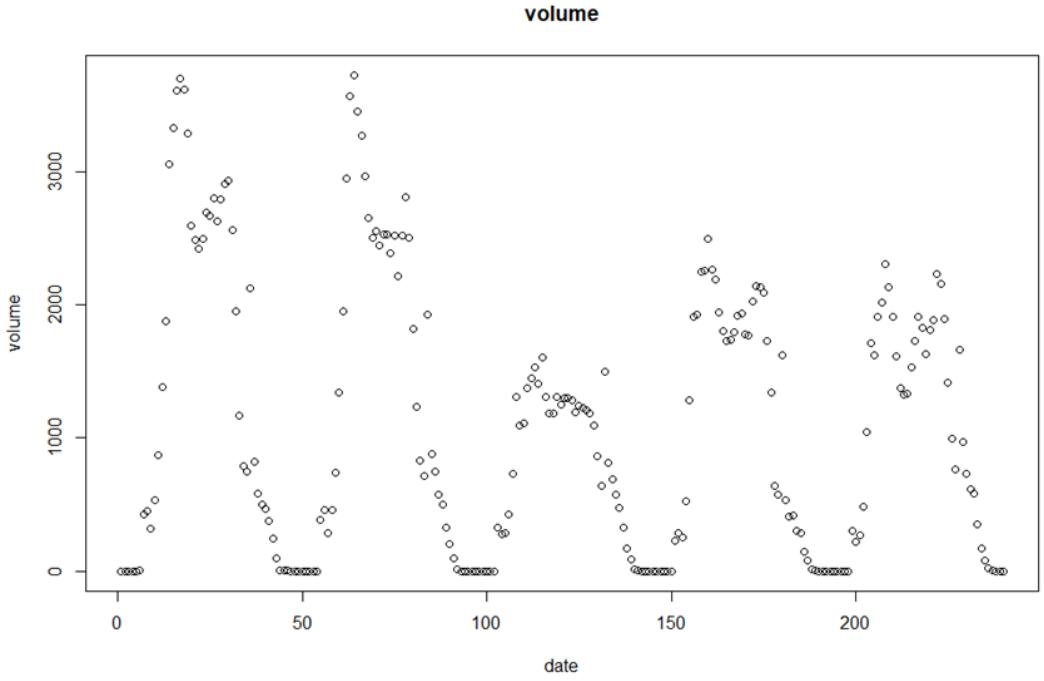


Fig. 4 Distribution of the import and export traffic of Tian 'anmen East Railway Station

Figure 5. shows the data and forecast data fitted according to the ARIMA(2,0,0) model. The shaded part is the forecast confidence interval, and the blue line is the forecast curve.

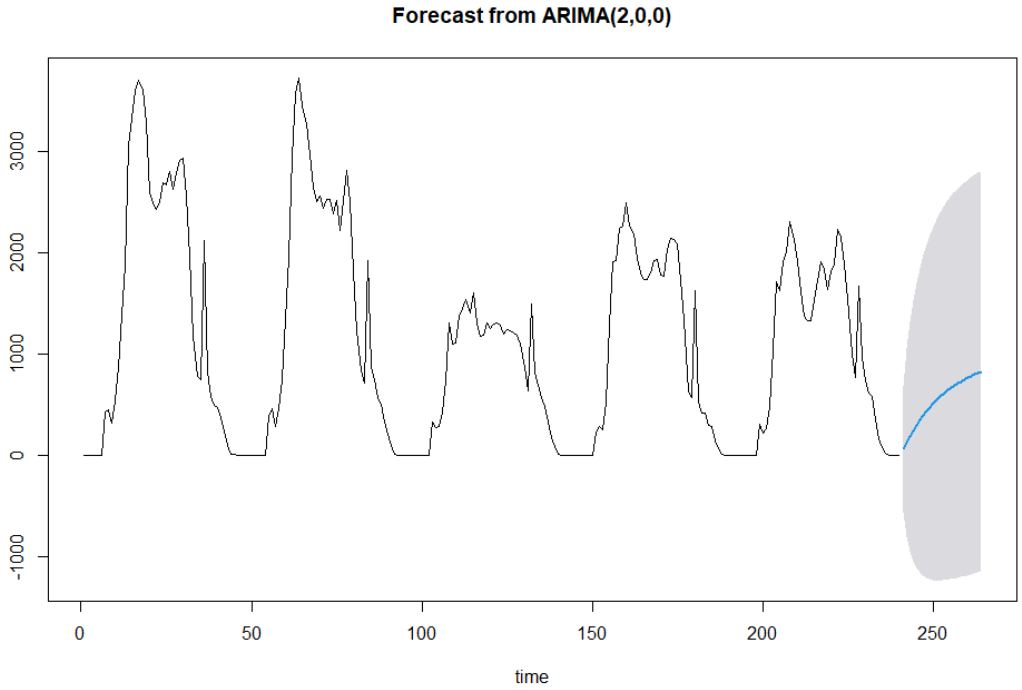


Fig. 5 Data fitted and predicted by the ARIMA (2,0,0) model

3.2.1 Autocorrelator

As shown in Figure 6, the model type and order q were determined according to the autocorrelation graph ACF. By observing the autocorrelation coefficient, we find that the autocorrelation coefficient diagram has a trailing condition, so it is determined that it follows the MA(0) process, that is, ARIMA (p, d, 0).

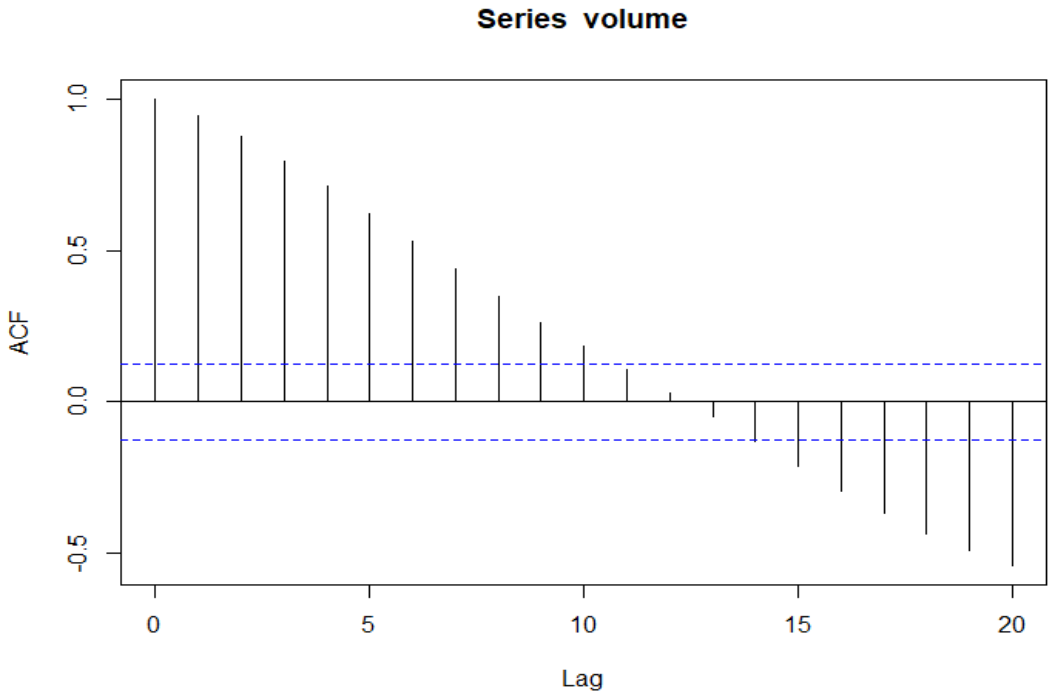


Fig. 6 Autocorrelator

3.2.2 Partial correlation diagram

As shown in Fig. 7, model type and order p are determined according to the partial correlator PACF. By observing the partial correlation coefficient, we found that the PD plot was out of the critical line in the 1st and 2nd cycle, and it was in the 3rd cycle. Therefore, we can certainly follow the procedure of AR (2), i.e., ARIMA (2, d, 0). Because of the stability of the original data, we used the original sequence, which is 0, so we concluded that d = 0. So now we know that the model parameter is ARIMA (2, 0).

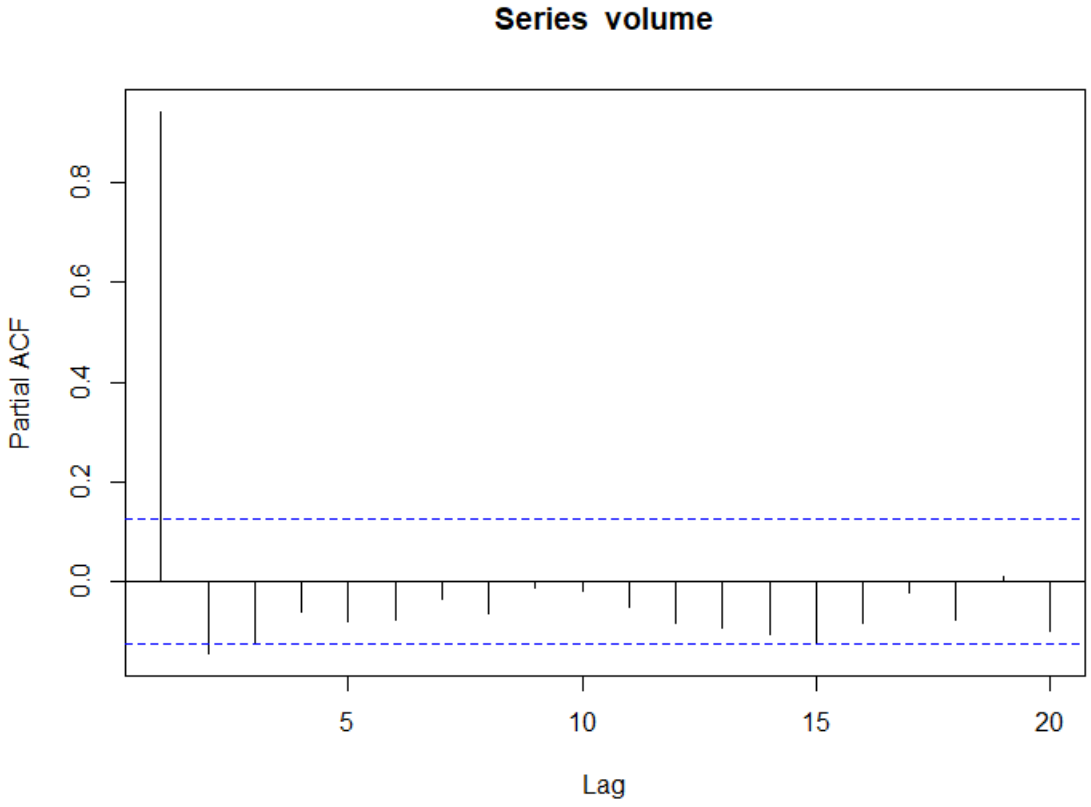


Fig. 7 Partial correlation Diagram

3.2.3 Model expression ARIMA(2, 0, 0)

After identifying the order of the model by autocorrelation graph (ACF) and partial correlation graph (PACF), ARIMA(2, 0, 0) was established, and the fitted model form was as follows:

$$y_t = 1005.0353 + 1.0949y_{t-1} - 0.1534y_{t-2} + \mu_t \quad (1)$$

4. Conclusion

ARIMA is a statistical method based on differential and autoregressive moving average models, which can be used to capture trends in time series. This paper constructed two graphs, PACF and PCF, using the model. Through the observation of partial correlation coefficient, it is found that the graph of PD is out of the critical line of standard deviation in the 1st and 2nd cycle. and fell within the critical line of standard deviation in the third period. Therefore, The data can be sure to follow the AR(2) process, that is, ARIMA (2, d, 0).

In this experiment, the ARIMA model was used for fitting prediction, but some results were not ideal enough. In subsequent experiments, other models should also be combined to improve prediction accuracy. The experiment results indicate that the precision of the single model should be increased, and the precision of the forecast can be increased by the combined model, such as the ARIMA-LSTM model. At the same time, in terms of data filtering, this paper used the dataset from May 12th to 17th, 2018, and processed it every 30 minutes, which is more complex in processing and analysis.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Long Xiaoqiang, Li Jie, Chen Yanru. Short-term Passenger Flow Prediction of urban rail Transit based on Deep Learning. Control and Decision, 2019, 34(8): 1589-1600.

[2] Hu Mingwei, Shi Xiaolong, Wu Wenlin, He Guoqing. Short-term Passenger Flow Prediction method of urban rail Transit Station based on machine learning. Journal of Shenzhen University: Science & TechnologyEdition, 2022, 39(5): 593-599.

[3] Yao Enjian, Cheng Xin, Liu Shasa, Zhang Rui. Prediction of existing inbound and outbound passenger flow of urban rail based on Accessibility. Journal of Railway Science, 2016, 38(1): 1-7.

[4] 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.

[5] Ni Jie, Yu Li, Jin Xiaonan. Passenger Flow Prediction and Research of Urban Rail Transit based on ARIMA Model. Intelligent Computer and Applications, 2021, 11(4): 135-138.

[6] Lu Zhiyi, Nie Weicong, Chen Lizhen. Urban rail transit passenger flow prediction based on ARIMA model. Henan Science, 2018, 36(5): 646-651.

[7] Shi Yuxin, Chen Lingyan, Liang Yingyi, Chen Kexin, Li Xiqin. Intercity railway passenger flow prediction based on wavelet decomposition and ARIMA model. Jiangsu Science and Technology Information, 2019, 36(29): 30-34.

[8] Cai Changjun, Yao Enjian, Wang Meiying, Zhang Yongsheng. Urban rail transit passenger flow prediction based on product ARIMA model. Journal of Beijing Jiaotong University, 2014, 38 (2): 135-140.

[9] LI Dewei, Yan Yixing, ZENG Xianfeng. Short-term combination forecasting Model for inbound passenger flow of Urban rail Transit. Urban Rapid Rail Transit, 2017, 30(1): 54-58.

[10] Cai Changjun, Yao Enjian, Wang Meiying, Zang Yongsheng. Urban rail transit passenger flow prediction based on product ARIMA model. Journal of Beijing Jiaotong University, 2014, 38(2): 135-140.

[11] Sun Yinshuo, Pan Simin, Geng Liyan, Yu Jianjian. Passenger flow prediction of Intercity high-speed rail based on Wavelet Analysis and ARIMA model. China Market, 2020, (19): 6-8.