Analysis of Civil Aviation's Passenger Turnover Based on X11-ARIMA Model

Lixin Zhang and Cuifang Yang

College of Information Engineering, Tarim University, Alar Xinjiang843300, China

Abstract—**The analysis of civil aviation's passenger turnover is of great significance to civil aviation department. According to the data of civil aviation's passenger turnover in the last 12 years, using the X11 method, the long-term trend was separated, and the results showed that the civil aviation's passenger turnover was linearly increasing and there was a seasonal effect of one year cycle. Using the difference method to eliminate the influence of long-term trend and seasonal effect, ARIMA model was proposed to fit the development of the sequence. Taking the data from January to September in 2017 as the test sample, the model predicted the data very precisely.**

Keywords- turnover; long-term trend; sequence

I. INTRODUCTION

Under the influence of the new normal economic, China's civil aviation has stepped into a more rational and steady development path. Data from China Civil Aviation Administration shows that during the 12th FYP period, CAAC delivered a total of 1.8 billion passengers safely. Civil aviation's passenger turnover in the integrated transport system has reached 22.8%.Traveling by air has become one of the popular modes of transportation[1].Under this background, correct analysis of civil aviation's passenger turnover is of great significance for the operation management of airport, the reform of charging mechanism and other related work. Hongyao QIN, Yong TANG, Dan HU, Yan FANG, Xiaoqiang LI implied that there are significant correlations between public festivals and seasonal indexes of civil aviation's passenger turnover[2].Ting PENG, yaping ZHANG,, Siqi HAO, Chuanyun FU pointed out that when proposing a forecasting model of passenger flow suitable for newly constructed airports and improving the forecast accuracy of the traditional models, it should not only analyze the influences of social economy on passenger throughput of civil aviation, but also consider the factors, such as the structure of aviation network etc[3].Jianguo KONG and Xiao LI based on the wavelet analysis theory to predict the amount of civil passenger, but the selection of the wavelet base increases the complexity of the problem[4].Jing GUAN introduced a grey Markov model to predict the amount of passengers handled at airports in order to make forecasts more accurate and reliable, but the division of states has a strong subjectivity[5]. Qian ZHU and Jinghong LI combined the function of wavelet analysis and neural network to establish the prediction model, but the difficulty in selecting variables to a certain extent, affect the practicality of the model[6-7].Based on the characteristics of gray forecasting and support vector machine, Jing GUAN proposed a gray support vector model which combined the two forecasting methods. However, the choice of model parameters for SVM has a greater impact on the accuracy of the prediction results[8].There are also some combined models to predict the volume of civil aviation's passenger, but there is no uniform standard for the selection of the combination weight[9-10]. Lihua CHEN set up a database using SPSS software to predict quantitatively some economic indexes such as the quantity of all the passengers, but did not consider the seasonal impact on the volume of civil aviation's passenger[11].Moreover, the above studies mainly focus on the volume of civil aviation's passenger, the research on the civil aviation's passenger turnover is relatively few.

The rapid development of time series theory in recent years has provided a brand new way for people to analyze and study various problems in the fields of natural sciences and social sciences. The X11 method was a time-based factorization process based on computer automatic calculations commissioned by the National Bureau of Census after World War II. Unfortunately,the conclusion that using X11 theory to analyze civil aviation's passenger turnover is seldom seen. This paper attempts to use X11-ARIMA model to analyze the trend of civil aviation's passenger turnover, and uses the data from January to September in 2017 as the sample to test the accuracy of the model.

II. MATERIALS AND METHOD

A. Data Material

From the website of the National Bureau of Statistics, we collected the monthly data on civil aviation's passenger turnover from January in 2005 to September in 2017.

B. Research Methods

1) X11 model

The X11 model assumes that the time series consists of four elements: long-term trends, seasonal variations, cyclical fluctuations, and random fluctuations. The following three moving average methods are used in the X11 program to achieve accurate decomposition of various sequences: simple center moving average, Henderson weighted moving average, Musgrave asymmetric moving average.

2) ARIMA model

Assuming that x_t is the t-th observation of a time series, the model $ARIMA(p,d,q)$ is:

$$\nabla^d x_t = \frac{\Theta(B)}{\Phi(B)}\varepsilon_t \tag{1}$$

The t here represents time, B stands for delay operator, then $\nabla = 1 - B$,

$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ represents the regression polynomial,

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ represents the moving average polynomial.

The multiple seasonal model of order $(p,d,q) \times (P,D,Q)_S$ can be expressed as

$$\nabla^d \nabla_S^D x_t = \frac{\Theta(B)\Theta_S(B)}{\Phi(B)\Phi_S(B)} \varepsilon_t \qquad (2)$$

The S here represents the seasonal cycle, then $\nabla_S = 1 - B^S$,

$\Phi_S(B) = 1 - \varphi_1 B^S - \varphi_2 B^{2S} - \cdots - \varphi_P B^{PS}$ represents the seasonal autoregressive polynomial.

$\Theta_S(B) = 1 - \theta_1 B^S - \theta_2 B^{2S} - \cdots - \theta_Q B^{QS}$ represents the seasonal moving average polynomial.

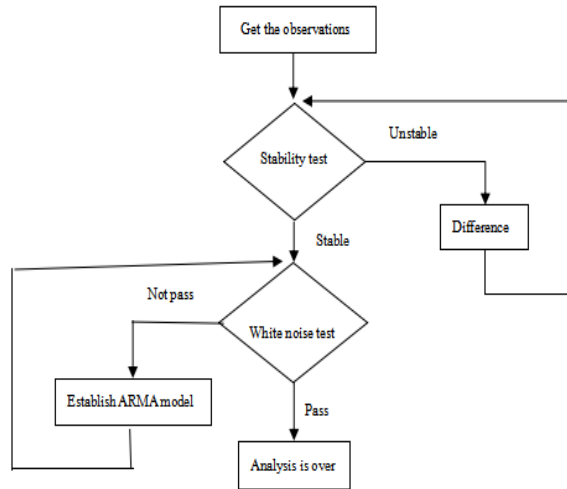The modeling process of the ARIMA model is shown in figure 1:

FIGURE I. MODELING PROCESS

III. EMPIRICAL ANALAYSIS

A. Long-Term Trend Analysis

We select the data of civil aviation's passenger turnover from January in 2005 to December in 2016, in order to show the trend of changes in the sequence, draw the timing diagram, as shown in figure 2.
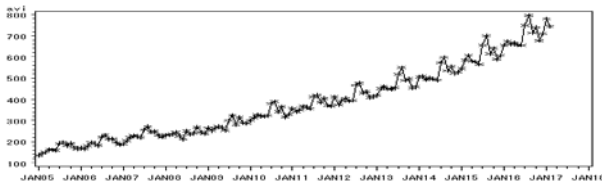

FIGURE II. ORIGINAL SEQUENCE DIAGRAM

From the sequence diagram as a whole we can see the linear increasing trend, seasonal effect for one year period and the influence of random fluctuations, using X11 method to decompose the long-term trend, and draw the long-term trend chart, as shown in Figure 3.
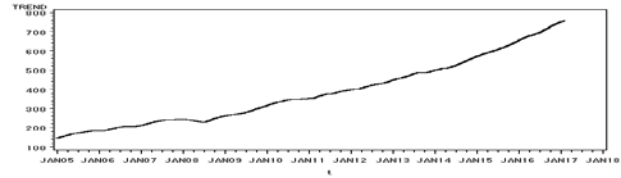

FIGURE III. LONG-TERM TREND CHART

From the long-term trend chart, it can be clearly seen that the civil aviation's passenger turnover has shown a linear growth trend in the past 12 years. The reason is that the main factor affecting the civil aviation's passenger turnover has been growing rapidly in recent years, such as GDP. Under this influence, civil aviation's passenger turnover has been growing rapidly.

B. Seasonal Effect Analysis

According to the timing chart, it is found that the civil aviation's passenger turnover from 2005 to 2016 has a seasonal effect of one year. Assuming that there is no seasonal effect, the civil aviation's passenger turnover should fluctuate around a certain average value. The existence of seasonal effect makes passenger turnover take a year as a cycle, showing similar characteristics in the same months of different years. In order to obtain quantitative seasonal information, the seasonal effect was separated by X11 method and the average seasonal index of each month was calculated by the simple average method, as shown in table 1.

TABLE I. AVERAGE SEASONAL INDEX

Month	Average Seasonal Index（%）
January	97.25701801
February	98.10717606
March	98.59940332
April	100.5423884
May	97.02770434
June	94.22283459
July	109.2572942
August	112.9657346
September	100.624833
October	103.7877948
November	94.6104037
December	92.99741501

As can be seen, the seasonal indies of July and August are larger than the other months, the reason is that the summer holidays in July and August, which are the peak season for airlines. Followed by October, because on the National Day, people go out to travel, then forming the peak season. The seasonal indies of September and April near 100, because in September, all over the country has beautiful scenery, so a tourist season. The traditional Qingming Festival in April coincides with the spring season, so it is suitable for outbound travel in April. Coupled with the low discount sales of air tickets, it also increases civil aviation's passenger turnover. The lower is in December, affected by the Spring Festival, people generally go home, so it is relatively off-season. This is

consistent with Hongyao QIN's findings, as in [2].

C. ARIMA Model Analysis

1) Stability test

Since the sequence has a very clear linear increasing trend and the seasonal effect of one year cycle, we eliminate the linear trend by first order difference, then make a 12-step difference to eliminate the seasonal effect. The sequence after differences is recorded as y_t , and the stability test is carried out.

The test results are shown in table 2, where P-value is far less than the critical value of 0.05. Therefore, the differential sequence is stable.

TABLE II. STABILITY TEST

Type	Delay Order	τ -Statistic	*P*-Value
Type1	0	-15.69	<.0001
	1	-14.12	<.0001
Type2	0	-15.64	<.0001
	1	-14.10	<.0001
Type3	0	-15.57	<.0001
	1	-14.06	<.0001

2) White noise test

The white noise test of the sequence after the differences is shown in table 3.

TABLE III. WHITE NOISE TEST

Delay Order	χ^2 -Statistic	DF	P-Value
6	27.59	6	0.0001
12	67.25	12	<.0001
18	78.69	18	<.0001
24	82.77	24	<.0001

P-Value is far less than the critical value of 0.05, so the sequence is stable and is not a white noise sequence, there is a certain degree of relationship between the sequence, which need to establish a model to fit the changing trend of this sequence.

3) Model establishment

The order of the model is determined according to the characteristics of the autocorrelation coefficient and the partial autocorrelation coefficient of the post-differential sequence y_t .

Autocorrelation coefficient and partial autocorrelation coefficient are shown in table 4:

TABLE IV. AUTOCORRELATION COEFFICIENT AND PARTIAL AUTOCORRELATION COEFFICIENT

Delay Order	Autocorrelation Coefficient	Partial Autocorrelation Coefficient	Standard Deviation
0	1.00000	--	0
1	-0.31688	-0.31688	0.086711
2	-0.25401	-0.39399	0.099502
3	0.16398	-0.09615	0.099995
4	0.02315	-0.05251	0.101997
5	-0.07837	-0.05231	0.102037
6	0.05543	0.02069	0.102488
7	-0.00236	-0.00196	0.102713
8	-0.01541	0.01477	0.102714
9	-0.12713	-0.17445	0.102731
10	0.07843	-0.07059	0.103907
11	0.26081	0.25505	0.104353
12	-0.42163	-0.25895	0.109142
13	0.08283	-0.03912	0.120770
14	0.15277	-0.08099	0.121196
15	-0.11284	-0.05885	0.122636
16	-0.07422	-0.12356	0.123414
17	0.14908	-0.00567	0.123749
18	-0.06555	-0.02842	0.125092
19	0.04218	0.09467	0.125350
20	-0.04463	-0.01015	0.125457
21	0.06296	0.00833	0.125576
22	0.0188	0.01034	0.125813
23	-0.12706	0.03847	0.125834
24	0.03081	-0.17874	0.126795

Both the autocorrelation coefficient and the partial autocorrelation coefficient are not truncated, so we try to use the $ARMA(p,q)$ model to extract short-term related information. We select $ARMA(12,1)$ model after repeated attempts, combined with the previous differential operation, the final model is $ARMA(0,1,1) \times (1,1,0)_{12}$,that is

$$(1-B^{12})(1-B)x_t = \frac{1-\theta_1 B}{1-\varphi_{12}B^{12}} \varepsilon_t \tag{3}$$

The conditional least square method was used to estimate the parameters. The fitted model is:

$$(1-B^{12})(1-B)x_t = \frac{1-0.57791B}{1+0.45409B^{12}} \varepsilon_t \tag{4}$$

The test results show that the model successfully passes the residual's white noise test and the significance test of the parameters, which show that the model fully extracts the information of the data. The test results are shown in table 5 and table 6.

TABLE V. RESIDUAL'S WHITE NOISE TEST

Delay Order	χ^2 -Statistic	DF	P-Value
6	3.33	4	0.5038
12	8.92	10	0.5393
18	16.63	16	0.4097
24	20.24	22	0.5680

TABLE VI. PARAMETER TEST

Parameter	Delay Order	Estimated Value	*t*-Value	P-Value
MA1，1	1	0.57791	8.02	<.0001
AR1，1	12	-0.45409	-5.29	<.0001

In order to display the effect of the model fitting more intuitively, the fitting value and the observation value are drawn in figure 4, where the curve represents the fitting value and the asterisk represents the observation. It can be intuitively seen from the figure that the model fits well with the data.
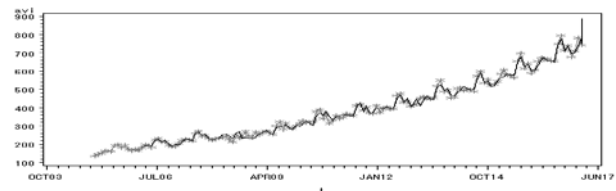

FIGURE IV. FITTING DIAGRAM

4) Model prediction

According to (4), the civil aviation's passenger turnover from January to December in 2017 is predicted. The data from January to September are taken as the test samples. The test results are shown in table 7.

TABLE VII. PREDICTED RESULTS

Month	Actual Value (100millionkm)	Predictive Value (100millionkm)	Relative Error(%)
January	781.61	736.7981	6.0820
February	743.42	789.0771	-5.7861
March	761.68	770.3587	-0.0113
April	754.65	757.9022	-0.0043
May	766.95	749.0057	0.0240
June	753.32	750.6305	0.0036
July	840.52	843.8066	-0.0039
August	869.84	888.2518	-0.0207
September	800.92	797.7830	0.0039
October		825.028	
November		767.272	
December		794.2412	

The value of MAPE is

$$MAPE = \frac{1}{9}\sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| = 1.3266\%$$

So the model predicts the data very accurately.

IV. CONCLUSION

According to the changing tendency of civil aviation's passenger turnover, we separate the long-term trend by X11 method. The result shows that civil aviation's passenger turnover tends to increase linearly. The reason is that the main factor affecting the civil aviation's passenger turnover has been growing rapidly in recent years, such as GDP. In this environment, civil aviation's passenger turnover has been growing rapidly. Using the difference method to eliminate the effects of long-term trends and seasonal effects, after repeated attempts, we use $ARMA(0,1,1)\times(1,1,0)_{12}$ model to fit the development of the sequence and forecast civil aviation's passenger turnover from January to December in 2017. The predicted value from January to September show that the model predicts the data very accurately, which can provide reference for the planning and construction of civil aviation departments.

REFERENCES

[1] Jinping FENG .2016 year worth looking forward to and concerned about the development of China's civil aviation hot [J]. air business, 2016,02:10-20.

[2] Hongyao QIN, Yong TANG,Dan HU, Yan FANG, Xiaoqiang LI. Investigating Seasonality of Air Passengers in China[J]. aeronautical computing technology, 2016,46 (05): 95-98.

[3] Ting PENG, Yaping ZHANG,Siqi HAO,Chuanyun FU. A Forecasting Model for Estimating Passenger Throughput of civil Aviation at Navigable City[J]. Journal of Transportation Information and Safety, 2016,34 (01): 37-44 + 63.

[4] Jianguo KONG, Xiao LI . Research on the Application of Wavelet Analysis Based on MATLAB in Predicting the Passenger Volume of Civil Aviation [J]. Sci-Tech Information of China, 2014, (17): 40-41.

[5] Jing GUAN. Apply Grey Markov SCGM（1，1） Model to Predict Airports Passenger Traffic in China[J] .Journal of Civil Aviation University of China, 2010,28 (01): 45-47.

[6] Qian ZHU. The Research of Airline Passenger Flow Prediction Based on Wavelet Neural Network [D]. Beijing Jiaotong University, 2014.

[7] Jinghong LI. Research on the Master Budgeting Management of Chinese Airlines Based on the Data Mining [D]. Civil Aviation University of China, 2009.

[8] Jing GUAN. Prediction of Passenger Throughput Based on Gray Support Vector Machine [J] .Journal of Dalian Jiaotong University, 2013,34 (03): 41-43.

[9] Yao YAO, Jing TAO , Yi LI. Prediction of Civil Aviation Passenger Transport Volume Based on ARIMA-BP Combined Model [J]. Computer Technology and Development, 2015,25 (12): 147-151.

[10] Yubao CHEN, Gang ZENG. Civil Aviation Passenger Throughput Forecasting Research Based on Combination Forecasting Method—Capital Airport[J]. Journal of Civil Aviation University of China, 2014,32 (02): 59-64.

[11] Lihua CHEN. Analysis and Forecasting of Civil Aviation Transportation Market[J]. Shanghai Jiaotong University, 2003,04: 623