Advances in Intelligent Systems Research, volume 130
6th International Conference on Mechatronics, Computer and Education Informationization (MCEI 2016)

Application of ARIMA Model in the Prediction of the Gross Domestic Product

Bing Yang[1, a*], Chenggang Li[1, 2, b], Min Li[1, c], Kang Pan[1, d], Di Wang[1, e]

[1]Faculty of Finance, Guizhou University of Finance and Economics, Guiyang, 550025, China;

[2]Faculty of Science, National University of Singapore, 117546, Singapore

[a]649747376@qq.com, [b]895977948@qq.com, [c]243165045@qq,com, [d]329223265@qq,com, [e]386954277@qq.com

*The corresponding author

Keywords: GDP; ARIMA model; Forecast; Residual

Abstract. GDP is an important index that is used to reflect the economy development and people's income. This paper chooses the annual data of Chinese GDP from 1978 to 2014 as the research object, and establishes ARIMA (2, 4, 2) model by applying the Eviews6.0 software. Then, the paper applies this model to forecast the GDP of the following five years, and compares the forecast values with the actual values. The result shows that this model is effective to forecast the GDP in a short term. In the end, the GDP of the following year is forecasted.

Introduction

Gross domestic product (GDP) refers to the market value of all the final products (goods and services) which is produced or provided by economic society (such as a country or a region) in a certain period of time. GDP is an important indicator to measure a country's wealth and economic strength. GDP is closely related with a country's employment level and living standard. So the prediction of GDP can be useful for the country in the future macroeconomic regulation and control, the calculation of the international balance of payments. Generally, there are two methods to calculate GDP, which are expenditure method and income method. The gross domestic product calculated by the expenditure method is equal to the sum of consumption, investment, government purchases, and net exports. It is often difficult to analyze and predict GDP by using the traditional structural method [1]. Therefore, the use of more suitable ARIMA model to predict the GDP can get a satisfactory result.

This paper uses the annual GDP data Chinese 1978 to 2014 as the research object, constructs the ARIMA (p,d,q) model, and analyzes the GDP in China with the help of EVIEWS 6.0 software[2]. Based on good model fitting effect, this paper forecasts the future trend of GDP.

ARIMA (p, d, q) Model

Setting Xt sequence is a random non-stationary time series. After the X sequence for D order difference [3], we can conclude:

$$Y_t = \nabla^d X_t \tag{1}$$

Among them, ∇ is the differential operation. d is the differential order. After the difference, sequence Yt has the following structure model:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{2}$$

Integrated auto-regressive average moving is referred as the ARIMA (p, d, q) model. Among them, p is auto-regressive order. d is the moving average. $\{\varepsilon_t\}$ is the white noise sequence. φ_i indicates the auto regressive coefficient, and θ_i is the moving average coefficient.

The auto-regressive moving average model, which is first proposed by the American statistician G.E.P.Box and G.M.Jenkins in 1970, is widely used in time series data analysis [4]. It is a short-term prediction method with high precision.

Application of ARIMA (p,d,q) Model in the GDP's Forecast

Data Selection and Processing. This paper selects the Gross Domestic Product (GDP) data from 1978 to 2015. All GDP data are from the official website of National Bureau of Statistics (http://www.stats.gov.cn/). All data processing and final data prediction are carried out using EViews 6.0 software.

Stationary Test of Time Series. Because the ARIMA model can only be applied to non-stationary time series analysis or can be transformed into stationary time series. Therefore, before the time series model analysis, we should first determine the stability of time series[5]. We do the stationary test of GDP by ADF test. The test results are shown in Table 1. It can be seen from the results of Table 1, the second order differential sequence of GDP is stable under the significant level of 1%, 5% and 10%.

Table 1 Stability test results

Sequence	T-statistic	Adjoint probability	Conclusion
Original sequence	6.365238	1.0000	Non stationary
First order differential sequence	-0.582695	0.8616	Non stationary
Second order differential sequence	-8.309458	0.0000	Stationary

There are three important parameters in the ARIMA(p,d,q) model, which are the auto-regressive order p, the difference order d and the moving average q. In the front, after second difference to determine the sequence stationary, we can get d=2. Next, it is necessary to determine the value of parameter p and parameter q in order to determine the final model ARIMA (p, d, q). By analyzing the auto-correlation and partial correlation graphs of time series GDP, the value of p and q can be determined[6]. The self-correlation and partial correlation test results of the sequence are shown in Fig. 1. From Fig. 1, according to the self-correlation and partial correlation sequence diagram, we can see that p=2, q=2. Therefore, the model ARIMA (2, 2, 2) is established.

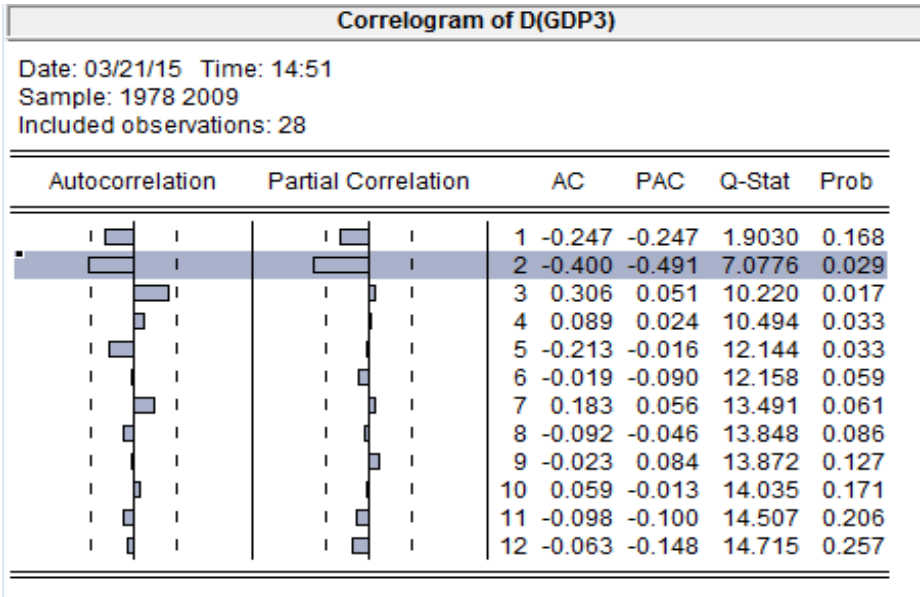
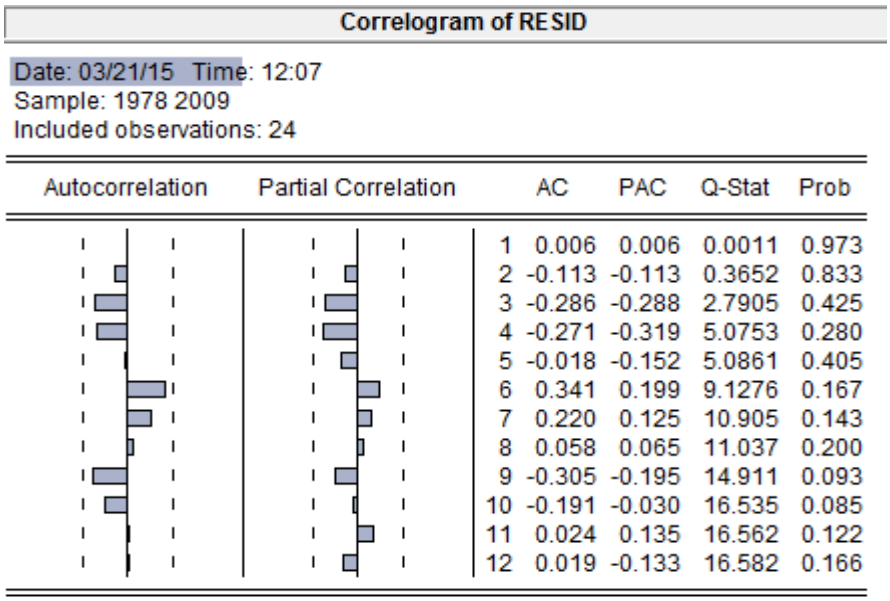


Figure 1. Self-correlation and partial correlation of GDP four order difference

Test of Model ARIMA (p,d,q). After the completion of the model construction, it is necessary

to carry out adaptive testing and to determine whether it is appropriate. For the ARIMA model, it is appropriate to test whether the residual sequence is white noise sequence to determine whether the model is appropriate [7]. Eviews6.0 software is used to test the white noise of the model residual series. The results are shown in Fig. 2.



	AC	PAC	Q-Stat	Prob
1	0.006	0.006	0.0011	0.973
2	-0.113	-0.113	0.3652	0.833
3	-0.286	-0.288	2.7905	0.425
4	-0.271	-0.319	5.0753	0.280
5	-0.018	-0.152	5.0861	0.405
6	0.341	0.199	9.1276	0.167
7	0.220	0.125	10.905	0.143
8	0.058	0.065	11.037	0.200
9	-0.305	-0.195	14.911	0.093
10	-0.191	-0.030	16.535	0.085
11	0.024	0.135	16.562	0.122
12	0.019	-0.133	16.582	0.166

Figure 2 Self-correlation and partial correlation coefficient test of residual

Fig. 2 shows that the auto-correlation coefficient of the residual sequence is in random interval, and there is no residual serial correlation [8]. They are close to zero in each order lag auto-correlation and partial auto-correlation values. All of the Q-statistic is not significant, and the lag of the P values of Q-statistic are significantly greater than 0.05. The residuals pass through white noise sequence test. Therefore, the model ARIMA (2, 2, 2) can be used for further analysis and prediction.

Model Estimation. From the previous calculation, the final model is determined as the ARIMA (2,2,2) model. Using views 7.2 software, we get the coefficients of the ARIMA (2,2,2) model. The results are shown in Table 2.

Table 2 Coefficient estimation of ARIMA(2,2,2) model

Variable	Coefficient	Std.Error	t-Statistic	Prob
C	-47563.08	36446.24	-1.305020	0.2009
AR(2)	1.202908	0.030922	38.90087	0.0000
MA(2)	0.395020	0.195986	2.015552	0.0521

From the above table, the mathematical specification of the model ARIMA(2,2,2) is

$$Y_t = -518.2927 + 1.336645Y_{t-2} + \varepsilon_t + 0.835119\varepsilon_{t-2} \tag{3}$$

Forecasting Analysis of GDP. We use the ARIMA(2,2,2) model to forecast the GDP in 2016 and 2017. Firstly, we use the GDP from 1978 to 2010 to forecast the GDP of 2011. The main calculation method is the deferred method. Then, we take the forecast GDP data of 2011 data back, combined the data from 1978 to 2012 to forecast the GDP of 2013. The deferred method can get more accurate results [9]. The predicated GDP from 2011 to 2017 are shown in Table 3.

Table 3 Prediction results of GDP from 2011 to 2017

Date	Predictive value	Actual value	Relative error	Average error
2011	406977.5	489300.6	-16.8%	6.78%
2012	506788.2	540367.4	-6.2%	
2013	615875.9	595244.4	3.5%	
2014	662066.6	643974	2.8%	
2015	708411.2	676708	4.6%	
2016	724077.6			
2017	773314.8			

The predicated errors are shown in Fig. 3. As can be seen from Fig. 3, the Thiel inequality coefficient is 0.376and its decomposition table shows that the error ratio is close to zero, and the variance is small. So, this shows that the model has better prediction effect.

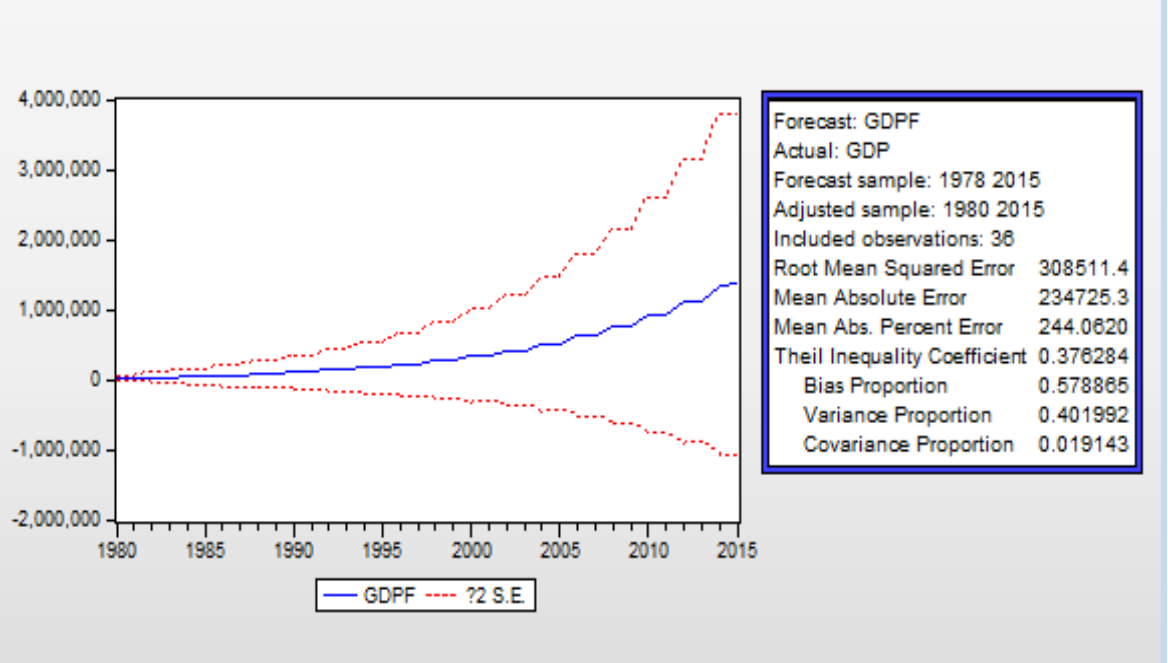


Figure 3. The predicated errors

Comparing the forecast value and the actual value, we can see that, the average error is 6.78%, which is less than 10%. Therefore, the ARIMA (2,2,2) model is an effective forecasting model. Seen from prediction results of Table 3, China's GDP will reach to 72,407.76 billion Yuan in 2016 and reach to 77,331.48 billion Yuan in 2017. So, China's GDP will increase in the next two years. Since 2012, China's GDP growth rate has remained at around 7%, which is also a great link with the national policy. Policy changes will seriously affect the trend of GDP. In recent years, China has changed from a high growth rate to a medium high growth rate, and has changed the economic growth strategy from the previous investment driven to innovation driven [10]. Therefore, in the future, China's GDP may decline.

By using MAIAR model, we can predict the later GDP. The table shows that with the help of ARIMA model, in the short term, the forecast accuracy is high and the average error is small. But with the extension of the forecast period, the prediction error may appear gradually increases, because this is only a theoretical data. In real life, there are a lot of factors influencing GDP. Therefore, there will be some deviations from the actual values.

Conclusions

This paper constructs ARIMA model to forecast China's gross domestic product. The forecast results show that, China's GDP will reach to 72,407.76 billion yuan in 2016 and reach to 77,331.48 billion Yuan in 2017. Thus, in the next two years, the gross domestic product (GDP) will keep rising trend. But the trend is relatively small rise compared to the previous few years. The reason is that China has changed the economic growth strategy from the previous investment driven to innovation driven. And there are a lot of factors influencing GDP. Therefore, the predicated values will be some deviations from the actual values.

References

[1] H. Gao, Western Economics (macro part), Fifth ed, Renmin University of China Press, Beijing, 2010.

[2] Z.N. Li, W.Q. Pan, Econometrics, Third ed, Higher Education Press, Beijing, 2010.

[3] Q. Wang. Time series analysis and application, Southwest Jiao-Tong University Press, Chengdu, 2008.

[4] J. Yu, Econometrics software-the use of Eviews, University of International Business and Economics Press, Beijing, 2006.

[5] Y. Wang, Applied time series analysis, China Renmin University Press, Beijing, 2005.

[6] Y. J. Xu , Y. Z.Wang, The application of ARIMA Model in forecasting of Henan's GDP and Realizing method in SAS. China Science and Technology Information, 10 (2012) 216-219.

[7] L.Q. Zhang, ARIMA model estimation for the GDP in Guangzhou area, Journal of Changchun University of Technology (Natural Science Edition). 31 (2013) 624-627.

[8] Z. Xiong. Research on RMB exchange rate forecasting model based on combining ARIMA with Neural Networks,The Journal of Quantitative & Technical Economics. 6(2011) 65-76.

[9] S.Shao.Forecast research on ARIMA model for China's health insurance, Journal of Baoji University of Arts and Sciences. 3(2016)22-26.

[10] L.Xiao. Forecast of resident's consumption level based on seasonal ARIMA model, Statistics and decision making.8 (2016)83-86.