Forecasting USA GDP Base on ARIMA Model

Yanyi Peng[1, *]

[1] Chang Zhou NO.1 High School, Chang Zhou, 213000, China

*Corresponding author: 15020440311@xs.hnit.edu.cn

Abstract. Since the international tendency is more serious and complex today, GDP as the standard of evaluating economic final output has been paid more attention. GDP is the market value of all final goods and services produced within a country in a given period of time. In this paper, Box-Jenkins method is used to build ARIMA model for forecasting American GDP from 2000 to 2010. The most appropriate model for USA GDP is ARIMA (1, 1, 2). At last, this paper uses the model to forecast American GDP for next a few years. According to the forecasting plot, the trend decreases slowly, which means that GDP will decrease in several future years. In the case, America can take action like increasing job opportunities, encouraging consumption immediately to prevent the decreasing trend. The significant meaning of forecasting GDP is to let countries notice the possible future trend of the whole country's economy and change policies in time to make the situation better instead of being unprepared to face problems.

Keywords: American, GDP, ARIMA model, Box-Jenkins method.

1. Introduction

Gross domestic product (GDP) is a monetary measure of the market value of all the final goods and services produced in a specific time period by countries and it not only reflects the national economies, but also represents the mean standard of living [1]. If GDP increases positively, it means that the country develops quickly and people's income and consumption will increase. In this case, the living standard will be up to next level and the national happiness index will increase at the same time. Due to its important role played in both country and people, GDP will affect policies the government announces and the trend of economy. Taking fiscal policy as an example, if the government spends more on social security, GDP will increase according to the formula that GDP is equal to consumption plus investment plus government plus next export. What's more, GDP interacts closely with macroeconomy. There are three economic indexes influencing macroeconomy, including economic growth rate, inflation ratio and unemployment rate, all of which are connected with GDP. Firstly, economic growth can be reflected obviously by GDP using specific number. Secondly, as GDP increases, the demand will increase too so that the price will increase, which means GDP will cause inflation. Thirdly, according to Okun's law, employment rate will increase 0.05% as employment rate increases 1%.

The paper will use statistical techniques in time series to forecast GDP in America from 2000 to 2010. An Autoregressive Integrated Moving Average (ARIMA) model is an effective and popular technique among all of these. First, the paper preprocesses the time series by using stationarity test and white noise test and then we use the first differencing to make it stationary. Consequently, the research use Box-Jenkins approach to build ARIMA model to forecast American GDP. Finally, the paper uses diagnostic tests to check the normality and the stationarity of the residuals from the model. GDP may be used to gauge a nation or region's economic strength overall and the rate at which it is developing economically. It may also be utilized for economic structure analysis, which forms a critical foundation for macroeconomic decision-making. Examples of this study include industry structure, demand structure, and regional structure. Other significant indicators can be calculated using GDP in addition to pertinent indicators. For instance, combining GDP with population indicators may be used to calculate GDP per capita, and combining GDP with energy consumption can be used to determine GDP energy consumption indicators. Additionally, the GDP reduction index8—which accounts for changes in a nation's total price level—can be computed using current GDP and constant GDP (or region). The GDP indicator is a crucial one [2]. There are many studies

used an ARIMA model for studying the GDP in different countries and cities, such as Abonazel, M.R., Abd-Elftah, A.I.(2019) Forecasting Egyptian ARIMA Models and Xia, R.Y., Wang Z.J (2022) Prediction and analysis of Chongqing's GDP based on ARIMA Models

2. Methodology

2.1 Data Collection

The research uses the American GDP as the target in this paper to forecast. American GDP is the market value of all final goods and services produced in America in a given period of time [1]. It represents the strength of the US economy and the American mean standard of living. Although it is not accurate to estimate the comprehensive strength of a country, if the country's GDP is much lower than the average in the national situation, it cannot develop quickly in the future because only after having a high GDP is the country rich enough to afford many expenses on military, education and technology. America takes the lead in many fields like finance, high technology, astronomy and military and its GDP always ranks among top list in the world. In this case, according to GDP, it is easy to learn about American economic strength.

2.2 Data Selection

The paper uses the data of American GDP from 1980 to 2000 to forecast American GDP from 2000 to 2008 because America suffers a serious financial crisis in the early twentieth. Stock fell so fast that investors could not sell it immediately. America lost 5 hundred billion just on one day. Consequently, American regressed considerably to 1913. Many companies went out of the business and the unemployment rate rose sharply. The great collapse of the stock market led to financial regression for four years [3]. In the case, GDP hardly increased so that the increasing trend was not obvious at that time. This is the reason of choosing the data from 1980-2000.

2.3 Statistic Background

2.3.1 Moving average (MA) model

TheMoving average refers to a series of averages obtained by arithmetic averages of sev eral items in the time series using a step-by-item progressive approach. If the average num ber of data items is N, it is called the movement average of N period (item). Predicting ba sed on the moving average is the moving average prediction.Some of the movement averag e predictions only need to make a movement average, while others need to calculate the se condary moving average based on the characteristics of the time series [4].

2.3.2 Autoregressive (AR) model

It is to use the time sequence of a variable as the sequence of dependent variables, use the same variable to move several periods into the sequence of independent variable sequences, analyze the correlation between a series of dependent variables and one or more other independent variable sequences, and establish regression equations for prediction [5].

2.3.3 ARIMA

The ARIMA model as one of the most common and effective way to forecast stationary time series, which is not a white noise, uses a linear model to predict future values and describe the autocorrelations in the data. Because time series are fully random if the data is white noise, the trend and seasonality will change the value of the time series at various periods, therefore the time series is not worthwhile to analyze if the trend and seasonality are uncorrelated with the prior data [6]. If a time serious is not stationary, differencing which can make non-stationary time series stationary can solve the problem. ARIMA (p,d,q) is knowns as a non-seasonal model and p is the lag order, d is the degree of differencing, and q is the order of moving-average [7].

2.3.4 Box-Jenkins Approach

The Box-Jenkins approach is named after statistics George and Gwilym and it aims to use ARIMA model to find the most suitable model to forecast a time series based on past values [6].

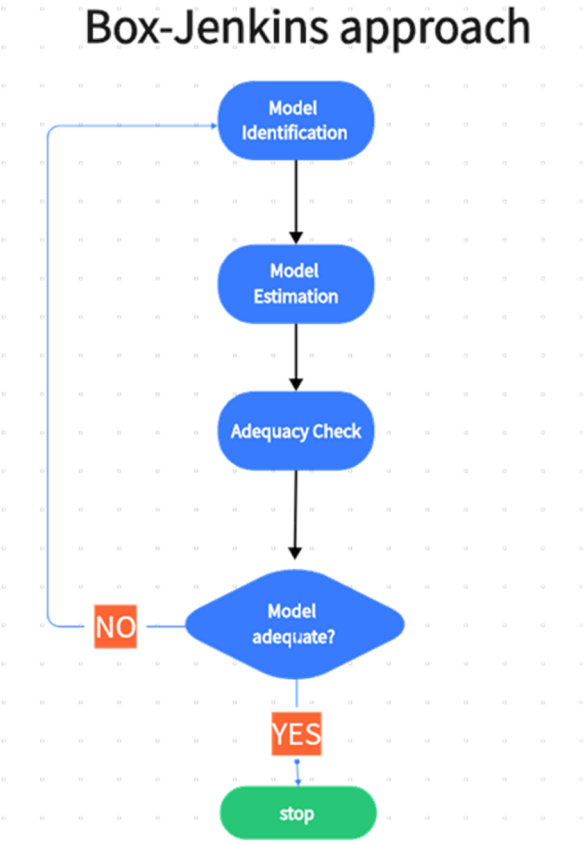

Fig. 1 Box-Jenkins approach（Photo credit: Original ）

Box-Jenkins approach is summarized in figure 1. Model Identification: It is necessary to make sure that the time series is not non-stationary and white noise so that Partial Auto-correlation Function (PACF) and Auto-Correlation Function (ACF) can identify whether the time series is stationary or white noise. If it is non-stationary, differencing can solve it completely. The PACF image can determine the specific order of AR(P) and ACF can be used to estimate MA (q). To be specific, if the PACF has truncation and close to 0 when lag >p, the order is p, which means that the time series is fit for building autoregressive model. If lag > q is close to 0 in ACF image, the order is q, indicating that the time series is suitable for moving average model [8]. In general, several models will be built and figure out the best fit one by comparing AIC.

Adequacy check: When building the most suitable model to forecast a time series, the next step which also plays significant role in Box-Jenkins approach is to test whether the estimated model conforms to the specifications of a stationary univariate process. Firstly, estimating residuals of ACF and PACF from the model are helpful to test. If they are white noise, the model works well. Otherwise, there must be something wrong in the process. Secondly, if many of autocorrelations are not within the limits, it means the model is not relatively suitable. After being sure about that the chosen model is adequate, the research can use it to forecast the time series.

3. Estimation Results

The preliminary analysis such as stationarity test and white noise test of the time series was done by using time plots in figure 2-3. If the time series is stationary, it will not be predicted in the long-term and time plot will be roughly horizontal with constant variance. Obviously, the series has

positive trends so that it is non-stationary in figure 2. ACF and PACF also show the result in p-value which is less than 0.05.

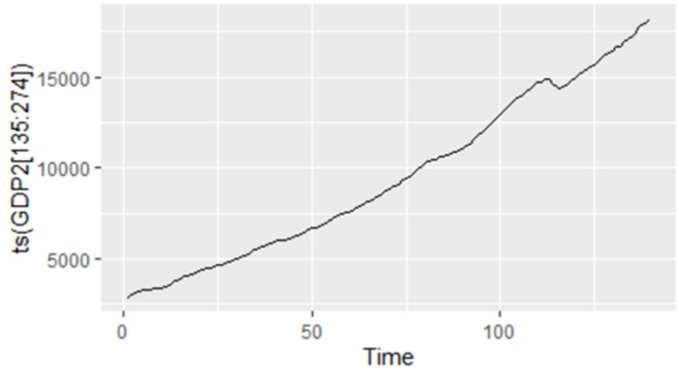

Fig. 2 Time plot of American GDP

What's more, figure 3 confirms that there is not white noise, which means it is worth to researching, because if it is, the ACF will not drop to zero quickly and almost none of the autocorrelations will lie outside the limits [6].

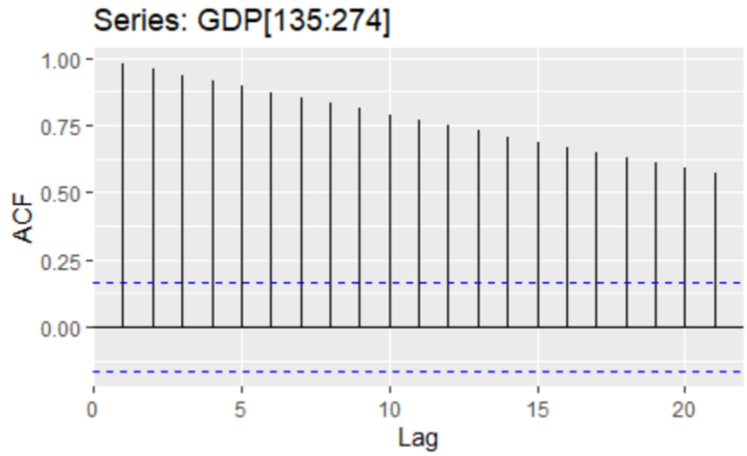

Fig. 3 ACF of American GDP

In the case, differencing can solve the problem of non-stationary time series in figure 4 by making non-stationary time series more stationary. The series which looks stationary is transformed by the first differencing, and so it is no need to consider further differences.

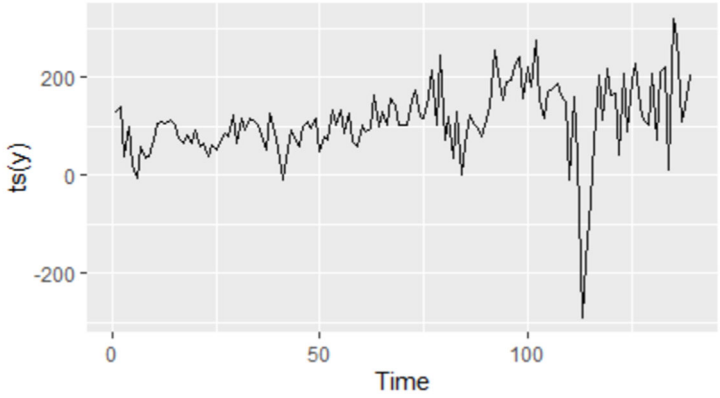

Fig. 4 The first differencing of American GDP

After the first transformation, the series become stationary and do not have trend so that d=1 in ARIMA model. Usually, ACF and PACF plots will be used to identify the other two, i.e. q and d and there are four results of q and d. Firstly, if the ACF plot is truncated at q order, and the PACF is rendered with a smear, so the best suitable model for the time series is moving average model instead of ARIMA model. If the PACF is truncated at the p order and the ACF shows a smear, however,

autoregressive model is fit for ARIMA model. Thirdly, the time series is appropriate to ARIMA model if both of plots have clearly smears, and it is significant to determine p and q value by ACF and PACF plots. Finally, if both of plots are truncated, there is white noise and certainly some problems to use ARIMA model. Obviously, according to ACF and PACF plots in figure 5 and figure 6, they are truncated so that ARIMA model is suitable for the series. The truncation is in the second order in ACF so that q=2 and it is in the first order in PACF so that p=1 in figure 5-6.

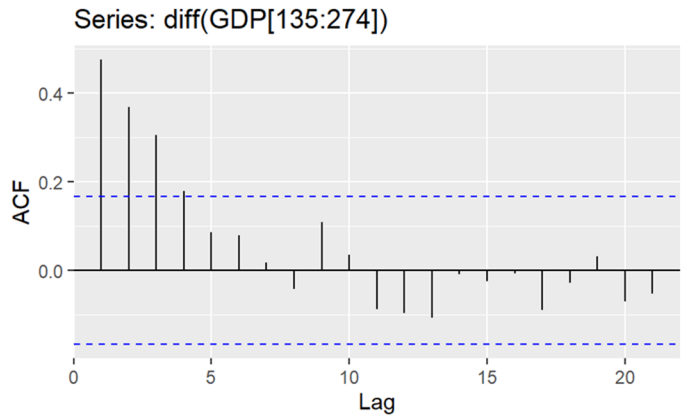

Fig. 5 ACF of the first differencing

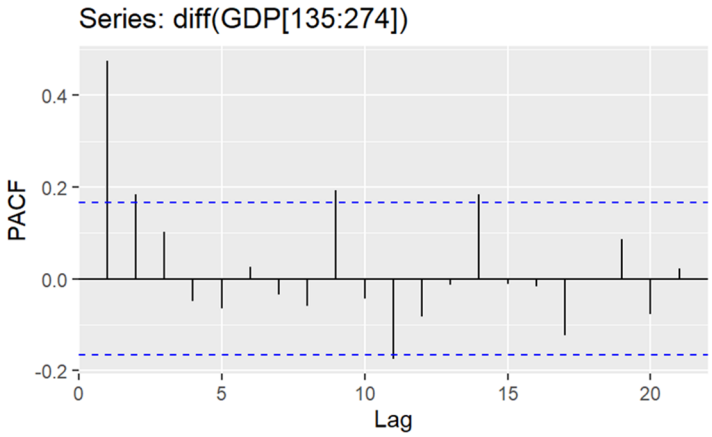

Fig. 6 PACF of the first differencing

In order to find the fit model for the time series to forecast, it is necessary to compare several different ARIMA models to find the best one by comparing AIC in figure 7. In the case, the best ARIMA model order is ARIMA (1,1,2).

Table 1. Summary of AIC results under different models

ARIMA model	AIC
ARIMA (0,1,0)	1598.161
ARIMA (1,1,0)	1576.731
ARIMA (0,1,1)	1566.033
ARIMA (0,1,0)	1596.168
ARIMA (0,1,2)	1566.142
ARIMA (0,1,1)	1564.084
ARIMA (1,1,1)	1560.41
ARIMA (1,1,0)	1574.739
ARIMA (2,1,1)	1558.813
ARIMA (2,1,0)	1569.045
ARIMA (3,1,1)	1559.739
ARIMA (2,1,2)	1560.229
ARIMA (1,1,2)	1558.345
AIRMA (0,1,2)	1564.573
ARIMA (1,1,3)	1560.155
ARIMA (0,1,3)	1564.564
ARIMA (2,1,3)	1562.287

After figuring out the model, it is vital to check whether the chosen one is adequate by residuals from the model in figure 7 regarding Box-Jenkins approach. An appropriate forecasting method will yield uncorrelated residuals so that the point is to test whether the residuals behave like white noise. The residuals' ACF plot demonstrates that the residuals behave like white noise since all autocorrelations fall inside the crucial value. The portmanteau test also indicates the presence of white noise by returning a larger p-value than the significant level p-value of 0.05 in figure 7.

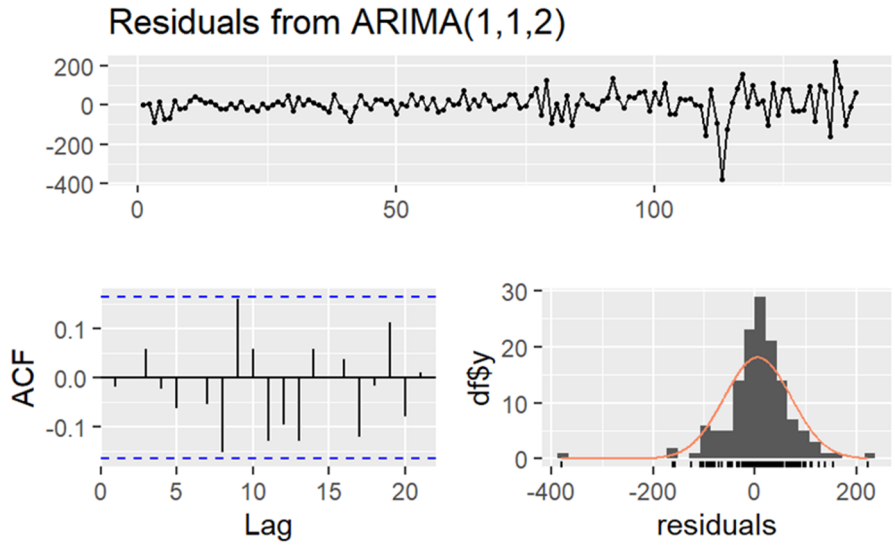

Fig. 7 residuals of the first differencing

Obviously, the ARIMA (1,1,2) residuals' ACF plot reveals that all autocorrelations are within acceptable bounds, which suggests that the residuals are acting like white noise.

Table 2. Ljung-Box test

Ljung-Box test		
Q*=9.5625	df=7	p-value=0.2148
Model df:3	Total lags used:10	

The Ljung-Box test returns a large p-value which is greater than the significant level p-value 0.05, also suggesting that the residuals are white noise. In the case, the ARIMA (1,1,2) is adequate for forecasting the time series since it passes the test. As a result, the research uses the chosen model to forecast American GDP in figure 8. The plot indicates that the USA GDP will have a decreasing trend in the future.

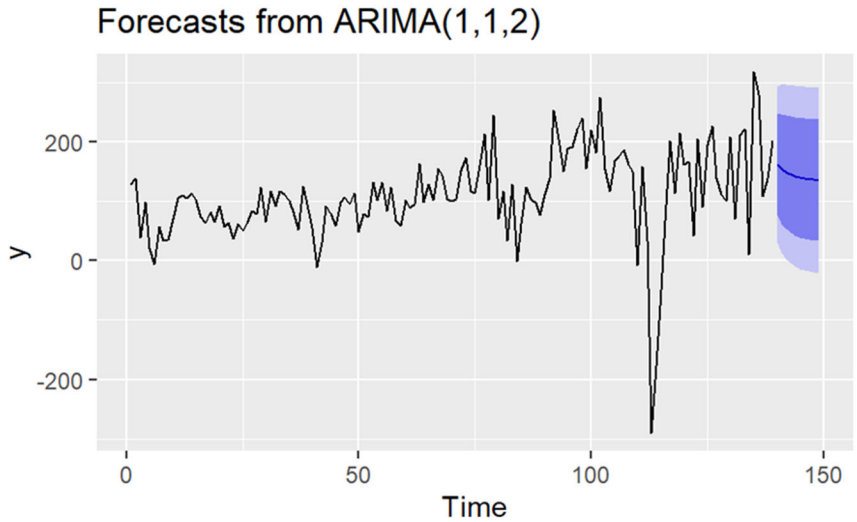

Fig. 8 Forecasting plot

According to Box-Jenkins approach, this article builds the best fit ARIMA (1,1,2), which is used for forecasting American GDP from 2000-2010, and check its adequacy. Finally, the plot shows that USA GDP will have a slowly decreasing trend in next 10 years late. Based on the forecasting result, America should take action to prevent GDP decreasing in future 10 years, which means announce policies to increase GDP. In practice, GDP can almost represent a country's economy so that there are some suggestions for economy growth.

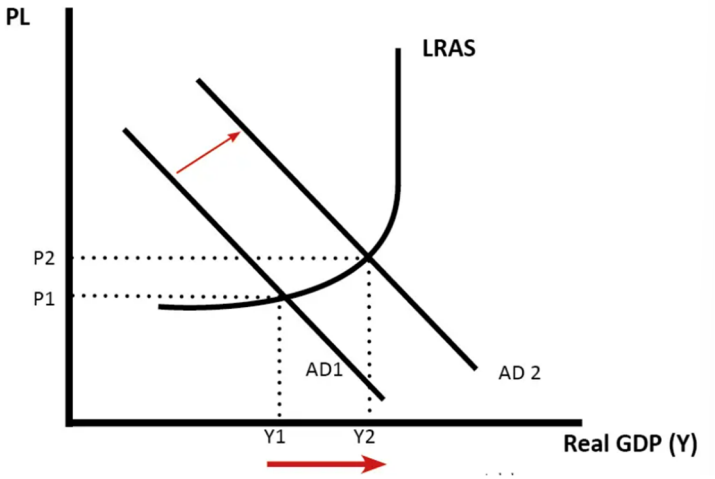

Fig. 9 Rise in aggregate demand

One of methods is to increase aggregate demand. There are various reasons causing the result in figure 9. The first is to lower interest rates, which can increase consumer spending and investment. The second is to increase real wages. If people have more wages to consume, consumption will increase [9].

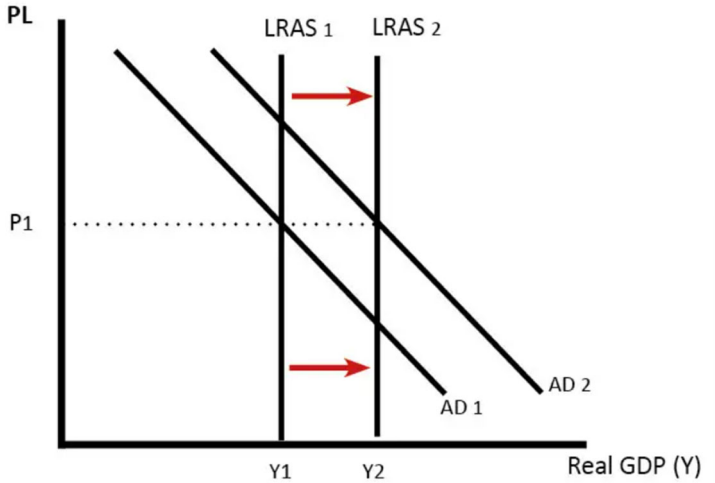

Fig. 10 Growth in productivity

Another method is to increase productivity. The first approach to grow productivity is to develop new technology such as tide energy, solar energy and wind energy, which is helpful to increase productivity. The second approach is to introduce new management techniques because better industrial relations help workers become more productive [9].

4. Conclusion

ARIMA model have two rules to follow and the most significant is stationarity. The data from countries' GDP is usually non-stationary so that it is necessary to preprocess a time series like differencing to transform it to a more stationary series. Then using PACF and ACF to estimate approximate values of p and q and building the most suitable ARIMA model to forecast by comparing the value of AIC. On top of that, regarding Box-Jenkins approach, adequacy check plays irreplaceable

role in the process of forecasting. The ACF of residuals from the chosen model behaves like white noise, which means the model is fit for American GDP. Finally, the plot of forecasting shows that there will be a small decrease in USA GDP in a few next years.

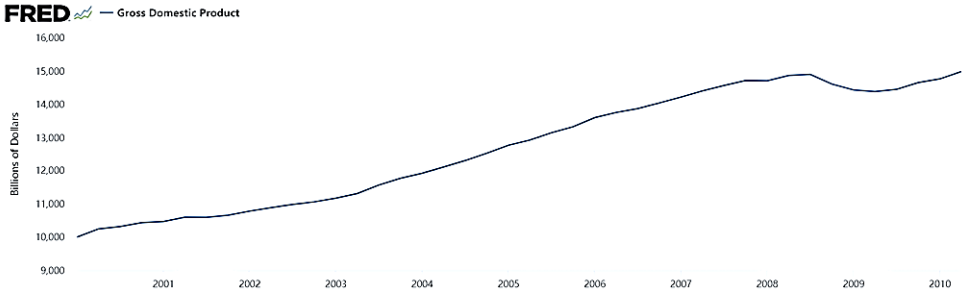

Fig. 11 American GDP from 2000 to 2010

According to the forecasting plot in figure 11, American GDP will decrease in next 10 years [10]. But although GDP fluctuates and has decreasing trend from 2008 to 2009, the overall trend is upward in reality. In the case, the forecasting result from the research is not consistent with data in real world. This paper uses Box-Jenkins approach, which is an economic forecasting method. Although it is effective and popular, there are some limits. This method assumes that the time series is stationary and will not change whenever in the past or the future, which is inconsistent with reality. In the case, it is impossible to use the origin model to forecast, which is inconvenient to change frequently, and the method is also inappropriate to forecast in long term like twenty or thirty years so that the paper can only forecast what the series will be like in future ten years. What's more, ARIMA have to use large amount of previous data to forecast in order to make the result more accurate so that the forecasting method only is fit for long-term time series. Also, it can only be used for stationary series and if the series is non-stationary, it is necessary to use some methods like differencing to transform it to a more stationary series. GDP cannot represent a country's economy totally. GDP excludes values of activities performed outside the market since it utilizes market pricing to modify goods and labor force. The value of the products and services generated in homes is specifically excluded from GDP. Additionally, neither income nor its distribution is included in GDP. GDP per capita merely depicts average living standards; nevertheless, average living standards do not apply to everyone. There is a huge difference between people.

References

[1] Wikipedia, 2022, Gross domestic product. https://en.wikipedia.org/, last accessed 2022/9/22.

[2] Xia. R.Y., Wang, Z.J. Prediction and analysis of Chongqing's GDP based on ARIMA model. Ln: National Knowledge Infrastructure. China. 2022, 93-94.

[3] Baidu, 2022, Financial crisis. https://baike.baidu.com/, last accessed 2022/9/22.

[4] Chen Min. Strong compatibility of an order-determining method for MA model. Journal of Shanxi University: Natural Science Edition, 1992.

[5] Xu Huijuan. Research on Global Least Squares Analysis of Autoregressive AR Model. Donghua University of Science and Technology, 2012.

[6] Hyndman, R.J., & Athanasopoulos, G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2, 2018.

[7] Abonazel, M.R., Abd-Elftah, A.I. Forecasting Egyptian GDP Using ARIMA Models. Economics and Finance, 2019, 5(1), 35-47.

[8] CSDN. 2019, Time Series Analysis Tutorial(6): Box-Jenkins approach (part 2). https://blog.csdn.net/, last acessed 2022/9/22.

[9] Pettinger, T. 2018, How to increase economic growth. https://www.economicshelp.org/, last accessed 2022/9/22.

[10] U.S. Bureau of Economic Analysis, Gross Domestic Product, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GDP, 2022.