Information Systems and Economics (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/infse.2024.050112
ISSN 2523-6407 Vol. 5 Num. 1

Analysis and Forecasting of GDP Using the ARIMA Model

Yao Ma

Haojing College, Shaanxi University of Science and Technology, Xi'an, Shaanxi, 712046, China
1194351167@qq.com

Keywords: GDP, ARIMA Model, Time Series Analysis, Trend Forecasting

Abstract: Gross domestic product (GDP) is an important indicator to measure the development of the national economy, which is important for promoting economic growth and assisting relevant departments in making economic decisions. In this paper, an ARIMA time series model is used to model China's GDP from 1978 to 2022 for empirical analysis. The results show that the predicted GDP values are in good agreement with the actual values, i.e. the ARIMA (0, 2, 0) model has high prediction accuracy. Based on the established ARIMA (0, 2, 0) model, China's GDP is predicted in order from 2023 to 2027. From the prediction results, it can be seen that China's GDP will still maintain steady growth. In order to promote China's economic growth, the following suggestions are made: (1) attract high-tech talents and complex talents; (2) optimise and upgrade the industrial structure; (3) adhere to the innovation drive; (4) strengthen the deepening of foreign economic cooperation.

1. Introduction

Gross domestic product (GDP) is an important macroeconomic indicator that reflects the level of economic development of a country or region and the living conditions of the people, which can reflect the scale of economic development of a country or region and judge the overall strength of its economy and the speed of its economic development; It can also be used to carry out the analysis of economic structure and is an important basis for macroeconomic decision-making; therefore, an empirical model analysis of GDP is conducive to the real-time adjustment of government policies and has important theoretical and guiding significance for the formulation of economic development policies.

Researchers have proposed a large number of models to forecast economic time series data such as VAR (Vector Autoregressive)[1,2], ARIMA (Autoregressive Integrated Moving Average)[3,4], LSTM (Long Short-Term Memory)[5-7], etc. for forecasting GDP of different countries. However, among all these forecasting models, ARIMA is the most popular[8,9]. Besides, it has been widely used by many researchers in forecasting economic indicators (e.g. GDP and CPI (consumer price index)), which shows that the ARIMA model has a good predictive effect on GDP[10-12]. Hongye Cai et al[13] selected the GDP data of Shenzhen City for the period 1980-2020 and built an ARIMA (2, 2, 3) model to predict the GDP of Shenzhen City for the next five years and the results showed that the relative error was only 2.9%, which indicates that the model has a high degree of accuracy in its prediction results. Uddin K et al[14] used an ARIMA (1, 2, 1) model to predict the GDP of Bangladesh for the years 2019-2025, and the results showed that the trend of GDP in Bangladesh has been steadily increasing in recent years and will continue to increase in the future. Abdullah Ghazo et al[15] modelled and predicted the GDP and CPI of Jordan using the ARIMA model. GDP data from 1976 to 2019 were selected for the model. The results show that ARIMA (3, 1, 1) is the best model for predicting GDP and it predicts that GDP will decrease in 2020.

Therefore, the purpose of this paper is to construct a suitable ARIMA model to predict the trend of China's GDP. First, the ARIMA model is constructed and empirically analysed using Python program on the annual data of China's GDP from 1978 to 2022, and the trend of China's GDP from 2023 to 2027 is reasonably predicted and analysed, which has certain practical significance and reference value, and it is expected to provide a scientific reference and basis for the future development of China's GDP.

2. Introduction to the ARIMA Model

The basic time series models[16] are Autoregressive (AR) model, Moving Average (MA) model, Autoregressive Moving Average (ARMA) model and Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model is one of the most widely used forecasting methods for time series data, first proposed in 1970 by two American statisticians, G.E.P. Box and G.M. Jenkins, and also known as the Box-Jenkins method. The ARIMA model is generally favored for its flexibility and forecasting accuracy for all types of time series data.

The ARIMA model aims to determine the most appropriate values of the p, d and q parameters and apply them to the forecasting and analysis of time series data. For a given time series data, the number of differences, d, is determined based on the smoothness of the series, and then the values of p and q for the AR(p) and MA(q) terms are determined based on the ACF and PCAF plots. The general form of the ARIMA model is

$$y_t' = c + \phi_1 y_{t-1}' + \cdots + \phi_p y_{t-p}' + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{1}$$

where p is the autoregressive coefficient, q is the moving average coefficient and d is the number of differences made to make the non-stationary time series stationary. ε_t denotes a normal distribution with a mean of 0 and a variance of σ^2.

The ARIMA model is mainly used for forecasting smooth time series, and the steps include plotting the time series, data pre-processing (including smoothness and white noise tests), model identification and ordering, and model forecasting.

3. Empirical Analysis Using an ARIMA Model

3.1 Pre-Analysis of Data

The annual GDP data selected for this paper come from the National Bureau of Statistics of China. (http://data.stats.gov.cn/)

As can be seen from the GDP time series graph in Figure 1, China's GDP data from 1978 to 2022 shows an exponential growth trend and significant non-stationarity characteristics. This is because the implementation of the reform and opening-up policy marked the entry of the Chinese economy into a new stage of development. In this stage, the Chinese government has vigorously promoted market-oriented reforms, gradually deregulated economic activities, introduced the competition mechanism and stimulated the vitality and innovation of the market; China's reform, opening-up and upgrading has entered a new stage of vigorous development, with unprecedented motivation in all regions and sectors, and a comprehensive acceleration of the pace of reform and opening-up, which has vigorously promoted the high-speed growth of the national economy and achieved significant achievements in all respects. Second, China's economic growth model has shifted from investment-driven to consumption-driven. In this process, the rapid development of the service industry, especially the rise of finance, information technology and tourism, has provided a powerful impetus to economic growth.

Since the 18th Party Congress, under the background of the economy entering a new normal and high-quality development becoming a theme, China's GDP has been growing faster and faster, bringing China's GDP into a new stage of rapid development. By 2022, China's GDP will be worth 1,210,207.2 billion yuan, about 329 times that of 1978, indicating that China's economic construction has made tremendous development since the reform and opening up, and the country's comprehensive national power, international influence and people's living standards have been steadily improving, highlighting the fact that China's development is on a more solid foundation, with better quality of development and more abundant development force.

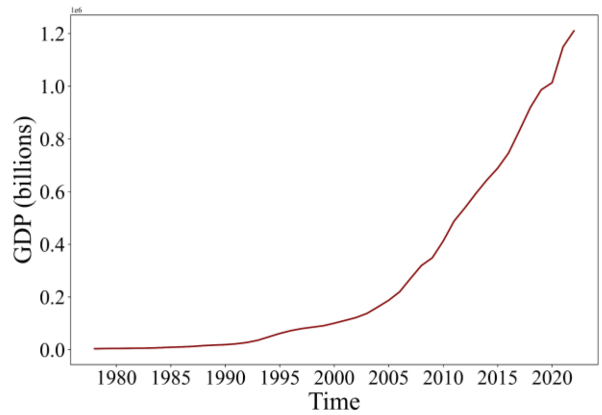

Figure 1: GDP time series

First, the log function is used to logarithmise the GDP series, denoted as INGDP. Figure 2 shows the time series plot of INGDP. From Figure 2 it can be seen that the INGDP series from 1978 to 2022 still shows obvious non-stationarity characteristics. Therefore, the INGDP series was differenced to eliminate the trend characteristics of the GDP series.

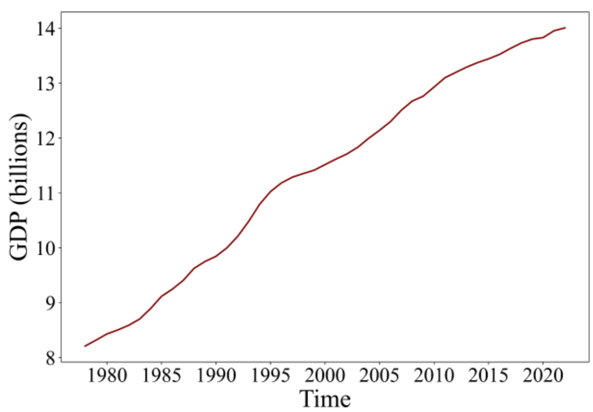

Figure 2: INGDP time series plot

3.2 Identification and Establishment of Models

3.2.1 Smoothness Test

From the preliminary analysis of the data, it can be seen that the Chinese GDP series is obviously a non-stationary series with strong exponential trend characteristics. Therefore, the Chinese GDP series is first logarithmically transformed into the INGDP series, as shown in Figure 2. For further modelling analysis, the INGDP series is first differenced to smooth the series and obtain the new time series DINGDP. To further test the smoothness of the DINGDP series, a unit root (ADF) test is performed using the t-statistic to compare the critical values of the t-statistic at the 1%, 5% and 10% test levels for different samples. If the value of the t-statistic is less than the critical value, the time series is smooth; otherwise, the time series is not smooth. As can be seen from Table 1, the t-value of the INGDP series is greater than the critical value at both the 5% and 10% test levels and the p-value is greater than 0.05, making it a non-stationary series. Therefore, the INGDP series must be treated as a D2INGDP series by second order differencing. Fortunately, the t-values of the D2INGDP series are all less than the critical values at the 1%, 5% and 10% test levels, and the p-values are all less than 0.05, rejecting the initial hypothesis and indicating a smooth series.

Table 1: ADF test results for INGDP, DINGDP and D2INGDP time series

sequences	t-value	p-value	1% level	5% level	10% level
INGDP	-2.60	0.093	-3.606	-2.937	-2.607
DINGDP	-2.74	0.067	-3.593	-2.932	-2.604
D2INGDP	-6.19	6.19e-8	-3.601	-2.935	-2.606

3.2.2 Model Identification and Ordering

The identification and ordering of the ARIMA (p, d, q) model is a crucial step in model construction, and according to the smoothness test, it is known that the ARIMA (p, d, q) model is selected to model the Chinese INGDP series. To determine the parameters of the ARIMA (p, d, q) model, the ACF and PACF plots of the D2INGDP series are plotted as shown in Figure 3. From Figure 3, it can be seen that the ACF plot of the Chinese D2INGDP series is truncated at order 0, and the PACF plot is also truncated at order 0. Therefore, the parameters of the ARIMA model in this paper are: $p = 0$ or 1, $q = 0$, $d = 2$.

Based on the ARIMA model parameter selection, the candidate ARIMA models in this paper are the ARIMA (0, 2, 0) and ARIMA (1, 2, 0) models. Secondly, parametric significance tests were performed on the ARIMA (0, 2, 0) and ARIMA (1, 2, 0) models, and it was found that only the ARIMA (0, 2, 0) model passed the parametric significance test. Therefore, the ARIMA (0, 2, 0) model is selected for the empirical modelling analysis of China's GDP in this paper.

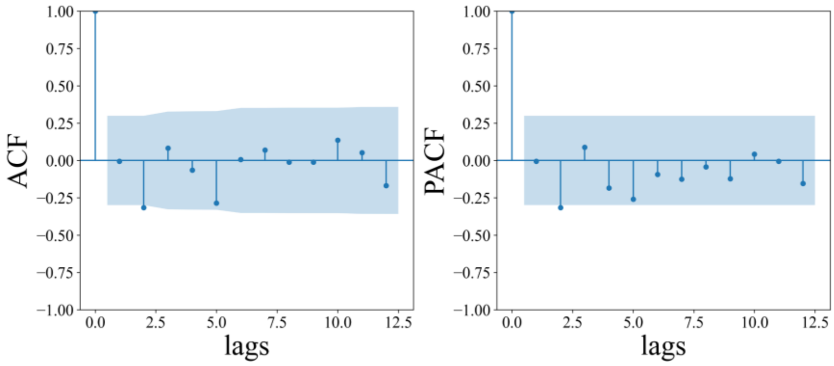

Figure 3: ACF plot and PACF plot of D2INGDP sequence

3.3 Model Testing

Model significance test, i.e. testing whether the residual series of the fitted model is a white noise series. If the residual series is a white noise series, it means that the useful information of the original GDP series has been sufficiently extracted and the model is significantly valid, otherwise the model needs to be reselected.

Figure 4 shows the standardised residual plot, the histogram estimated density plot, the Q-Q plot and the residual autocorrelation plot of the residual series of the ARIMA (0, 2, 0) model. It is clear from Figure 4 that the standardised residual plot, histogram estimated density plot, Q-Q plot and residual autocorrelation plot show that the residuals of the model satisfy the assumption of normality and there is no significant correlation. Therefore, the model has fully extracted all the information contained in the Chinese GDP series, indicating that the model passes the significance test.

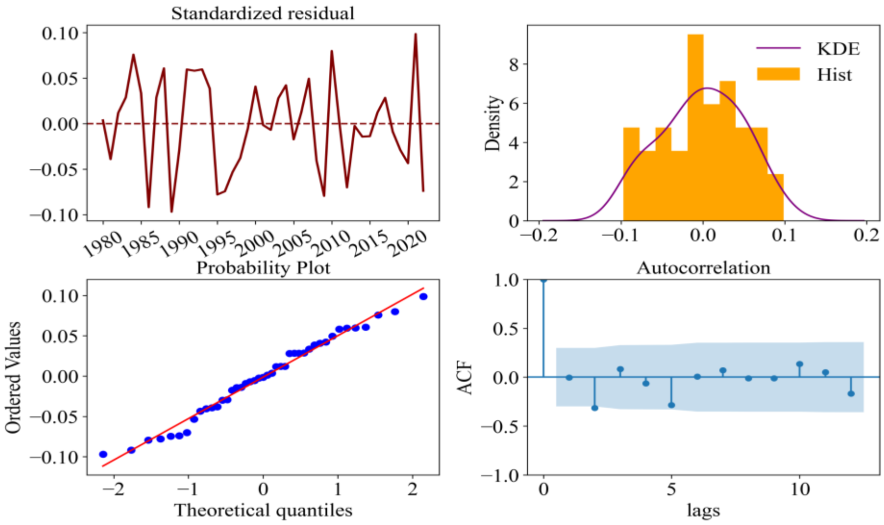

Figure 4: Residual series test plot

3.4 Modelling and Analysis

In this paper, the sample static forecast is chosen to forecast and analyse China's GDP from 2023 to 2027. Figure 5 shows the comparison between the predicted and actual values of China's GDP, where the solid line represents the actual value and the circle represents the predicted value. From the figure it can be seen that the GDP values predicted by the model are closer to the actual values and capture the trend of China's GDP. It can also be seen that China's GDP will grow steadily over the next five years.

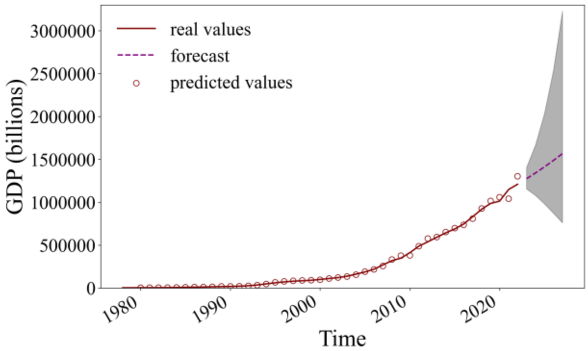

Figure 5: Comparison of actual and forecast values of China's GDP curve

China's GDP from 2023 to 2027 can be predicted by the prediction function of Python programme, and China's GDP in the next five years will be 1,274,412 billion, 1,342,023 billion, 1,413,221 billion, 1,488,196 billion and 1,567,149 billion, in the order of the next five years, as shown in Table 2. As can be seen from Table 2, China's GDP will still maintain a high growth rate in the next five years.

Table 2: GDP projections for Shaanxi Province and their 95% confidence intervals, 2023-2027

Years	Forecast (billion)	LCL	UCL
2023	1274412	1156086	1404848
2024	1342023	1079271	1668743
2025	1413221	981442.5	2034958
2026	1488196	872703.1	2537780
2027	1567149	760781.9	3228201

4. Conclusion

This paper uses the ARIMA model to empirically model and analyse China's annual GDP data from 1978 to 2022, and to forecast and analyse the GDP data from 2023 to 2027, and establishes an ARIMA (0, 2, 0) model with high forecasting accuracy. The relative errors of the model predictions are all less than 1%, and after testing, the errors generated by the forecasting model are within a reasonable range, then the model is used to forecast China's GDP from 2023 to 2027 with a certain degree of accuracy and reference value.

Based on the empirical analysis in this paper, the following recommendations are made to further promote China's economic development:

Firstly, attract and retain talent. China has huge labour resources, which provide a rich reserve of manpower for China's economic development. At the same time, with the improvement of education level, the quality of labour force will continue to improve, injecting new vitality into the development of China's economy, so it is necessary to pay attention to human capital investment and improve the talent security system. In addition, there should be a firm push to reform population policy, optimise fertility policy, actively explore ways to deal with the problems of population ageing and paedophilia, and stabilise the regional labour market.

Secondly, we will continue to optimise and upgrade our industrial structure to attract investment. As China's economy enters a new stage of development, the proportion of service industries continues to increase, especially in finance, information technology, tourism and other industries with great development potential; the optimisation of industrial structure has greatly improved the efficiency and effectiveness of economic operations; therefore, in order to optimise the industrial structure, improve policy initiatives and enhance the motivation of social investment, local governments should implement effective industrial guidance and actively attract investment.

Thirdly, we should adhere to the innovation drive and improve innovation capability. We will encourage and support enterprises to increase investment in research and development, establish a sound incentive mechanism for innovation, promote the effective integration and fusion of capital and technology, realise the economic impact of innovation, promote the effective application of technology, and raise the level of innovation in the country.

Fourthly, we need to launch an innovation drive and improve innovation capacity. This study encourages and supports enterprises to increase R&D investment, establish and improve innovation incentive mechanism, promote the effective integration and integration of capital and technology, achieve the economic effect of innovation, promote the effective application of technology, and improve the national innovation level.

However, we should also see that there are still some problems and challenges in China's economic development, such as environmental pollution, resource constraints and uneven income distribution. These problems require our due attention and solutions in future development.

References

[1] Narayan P K, Narayan S, Prasad A. A structural VAR analysis of electricity consumption and real GDP: Evidence from the G7 countries. Energy Policy, 2008, 36(7): 2765-2769.

[2] Salisu A A, Gupta R, Olaniran A. The effect of oil uncertainty shock on real GDP of 33 countries: a global VAR approach. Applied Economics Letters, 2023, 30(3): 269-274.

[3] Wabomba M S, Mutwiri M, Fredrick M. Modeling and forecasting Kenyan GDP using autoregressive integrated moving average (ARIMA) models. Science Journal of Applied Mathematics and Statistics, 2016, 4(2): 64-73.

[4] Kontopoulou V I, Panagopoulos A D, Kakkos I, et al. A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. Future Internet, 2023, 15(8): 255.

[5] Saadah S, Wibowo M S. Prediction of gross domestic product (GDP) in Indonesia using deep learning algorithm. 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE, 2020: 32-36.

[6] Ouaadi I, Ibourk A. The Contribution of Deep Learning Models: Application of LSTM to Predict the Moroccan GDP Growth Using Drought Indexes. International Conference on Advanced Intelligent Systems for Sustainable Development. Springer, 2022: 284-294.

[7] Zhang Y, Shang W, Zhang N, et al. Quarterly GDP Forecast Based on Coupled Economic and Energy Feature WA-LSTM Model. Frontiers in Energy Research, 2023, 11: 1329376.

[8] Agrawal V. GDP modelling and forecasting using ARIMA: an empirical study from India. Central European University, 2018.

[9] Muma B, Karoki A. Modeling GDP Using Autoregressive Integrated Moving Average (ARIMA) Model: A Systematic Review. Open Access Library Journal, 2022, 9(4): 1-8.

[10] Rana S B. Forecasting GDP Movements in Nepal Using Autoregressive Integrated Moving Average (ARIMA) Modelling Process. Journal of Business and Social Sciences Research, 2019, 4(2): 1-20.

[11] Polintan S N, Cabauatan A L L, NEPOMUCENO J P, et al. Forecasting Gross Domestic Product in the Philippines Using Autoregressive Integrated Moving Average (ARIMA) Model. European Journal of Computer Science and Information Technology, 2023, 11(2): 100-124.

[12] Hassan S M. Modelling and Forecasting Somalia GDP Using Autoregressive Integrated Moving Average (ARIMA) Models. East African Journal of Business and Economics, 2023, 6(1): 255-264.

[13] Cai H, Qiu W. Prediction Analysis of Shenzhen GDP Based on ARIMA Model and Implementation in R Language. Academic Journal of Computing & Information Science, 5(10): 28-34.

[14] Uddin K, Tanzim N. Forecasting GDP of Bangladesh Using ARIMA Model. International Journal of Business and Management, 2021, 16(6): 56-65.

[15] Ghazo A. Applying the ARIMA Model to the Process of Forecasting GDP and CPI in the Jordanian Economy. International Journal of Financial Research, 2021, 12(3): 70.

[16] Box G E, Jenkins G M, Reinsel G C, et al. Time series analysis: forecasting and control John Wiley & Sons. Hoboken, NJ, 2008.