ISSN 2616-5902 Vol. 5, Issue 21: 117-122, DOI: 10.25236/AJBM.2023.052117

# Analysis and Forecast of GDP in Shaoguan City Based on ARIMA Model

**Lianghui Zhao[a], Borun Chen[b,*]**

*School of Economics & Management, Wuyi University, Jiangmen, China*
*[a]Xingzhe7249@163.com, [b]948256678@qq.com*
**Corresponding author*

***Abstract:*** *GDP is a measure of a region's economic development, industrial structure, economic vitality, etc. It is of great significance to analyze the development of a region's GDP and predict its future development trend. The ARIMA model is an important model in time series analysis and forecast. In this paper, the GDP data of Shaoguan City from 1978 to 2019 are selected for empirical analysis using SPSS 25.0 software. After the smoothing test and processing the original data, the ARIMA (0, 2, 0) model is established through steps such as determining model parameters and model testing. Then, a comparison is made between the real GDP data and the data predicted by the ARIMA model from 2020 to 2022. It is found that the relative error values between the model prediction results and the real data are small, which indicates that the model fits well. Finally, the GDP data of Shaoguan City in the next three years are predicted to provide certain references and suggestions for relevant departments to plan for the future urban economic development of Shaoguan.*

***Keywords:*** *GDP, ARIMA model, Prediction*

## 1. Introduction

Gross Domestic Product (GDP) is a crucial economic indicator, reflecting relevant information such as the economic growth rate and industrial structure of the country or region. It is of great significance for policymakers, researchers, and investors. It is also closely related to our lives. With the rapid economic growth, people's income will continue to increase, and living standards and quality can be continuously promoted and improved.

Shaoguan is located in the northern part of Guangdong Province, serving as the North Gate of Guangdong Province. It has a superior geographical location. The gross domestic product of Shaoguan City has steadily increased since the reform and opening up. But in the past two years, it has been affected by various factors. The GDP growth of Shaoguan City has been relatively slow. Especially in 2022, compared with the previous year, GDP only increased by 997 million yuan. Therefore, it is necessary to actively and deeply explore how Shaoguan City should promote economic development. Based on this, it is of great practical significance to build a model to predict the development of GDP in Shaoguan City.

Currently, many scholars have used different models to predict GDP data. Wang Meina and Yang Xiaobin selected the GDP data of Guizhou Province from 2012 to 2018 based on the GM (1,1) grey prediction model to predict the development of GDP in the next three years, and analyzed the relationship between GDP and industrial structure in Guizhou Province through grey correlation analysis[1]. Li Nan selected the GDP data of Jiangxi Province from 2000 to 2016 and used the BP neural network model to predict the GDP of Jiangxi Province in the next two years[2]. Wu Boni selected the GDP data of Hangzhou from 1978 to 2017 as a sample based on the ARIMA model and predicted the GDP of Hangzhou in the next five years[3]. Xia Ruyu and Wang Ziqiao selected the GDP data of Chongqing from 2009 to 2020 and established the ARIMA model. They found that the model fitted well and predicted the GDP index of Chongqing in the next 12 years[4]. Wang E. and Zhang Ting used the GDP data of Hunan Province from 1978 to 2017 as samples to establish an ARIMA model. They compared the actual values of Hunan Province's GDP data from 2016 and 2017 with the predicted values fitted by the model and found that the relative error was relatively small. The ARIMA model constructed can accurately predict future GDP[5]. Through previous research by scholars, it has been found that the ARIMA model has a good fitting degree and predictive effect, making it suitable for predicting GDP. Therefore, in this paper, the GDP data of Shaoguan City from 1978 to 2019 are selected, and the ARIMA model is established using SPSS software. The historical data from 2020 to 2022 are used to test the model, and the GDP data of

ISSN 2616-5902 Vol. 5, Issue 21: 117-122, DOI: 10.25236/AJBM.2023.052117

Shaoguan City in the next three years are predicted by the established ARIMA model. The predicted results help relevant departments provide a reference for formulating more accurate economic policies in Shaoguan City that promote the development of Shaoguan's economy.

## 2. Theoretical basis

### *2.1. ARIMA Model introduction*

The ARIMA model, also known as the Autoregressive Integrated Moving Average Model, is a kind of statistical model for analyzing and forecasting time series data proposed by Box and Jenkins in the 1970s[6]. The ARIMA model is widely used in the analysis and prediction of time series data in fields such as production and economy. The general form of the ARIMA model is ARIMA (p, d, q), where p is the order of the autoregressive terms, q is the number of moving average terms, and d is the number of differences made when the original time series becomes stationary.

### *2.2. Steps to establish ARIMA model*

The specific steps for establishing a time series ARIMA model are as follows:

(1) Preprocessing of raw time series data. Firstly, it is necessary to determine whether there are missing values in the time series data and then create a time series for the original data in SPSS.

(2) Stability test and handling. If the original data don’t satisfy stationarity, the data are differentiated until they are stationary.

(3) Determination of parameters. Parameters p, q and d are determined based on the lag order of the autocorrelation function graph and the partial autocorrelation function graph for the stabilized data.

(4) Establishing and testing the model. The model is established according to the determined parameters. Through the results of statistical analysis, the model is tested, and the optimal model is selected.

(5) White noise test. Determine whether the residual sequence is a white noise sequence.

(6) Model prediction.

## 3. Empirical analysis

### *3.1. Data sources*

In this paper, the GDP data of Shaoguan City from 1978 to 2019 are selected for empirical analysis, and an ARIMA model is constructed. The data are sourced from the Statistical Yearbook of Shaoguan City. (GDP, unit: 100 million yuan)

### *3.2. Preliminary analysis and stability test*

In most of the time series data, the original data are non-stationary series. Firstly, we need to determine the stationarity of the original data by observing the original time series graph, autocorrelation function(ACF) graph, and partial autocorrelation function(PACF) graph. The original time series of Shaoguan's GDP from 1978 to 2019 is shown in Figure 1. It can be seen that the GDP of Shaoguan City shows an increasing trend year by year without showing seasonal fluctuations. Then the tests of the ACF graph and PACF graph are carried out, as shown in Figures 2 and 3.

ISSN 2616-5902 Vol. 5, Issue 21: 117-122, DOI: 10.25236/AJBM.2023.052117

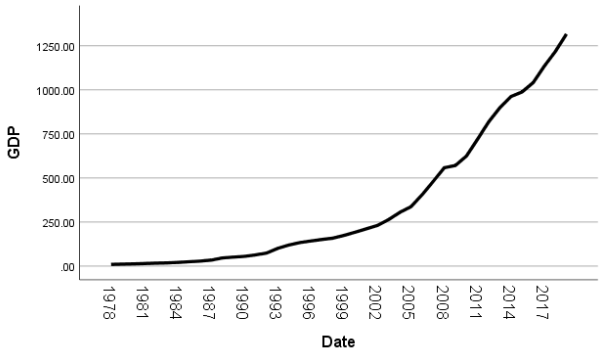


*Figure 1: Original time series of GDP in Shaoguan City from 1978 to 2019*

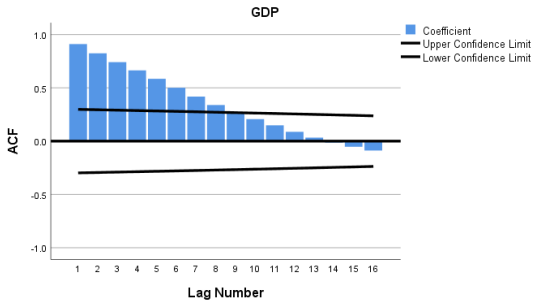


*Figure 2: Original time series ACF graph of GDP in Shaoguan City*

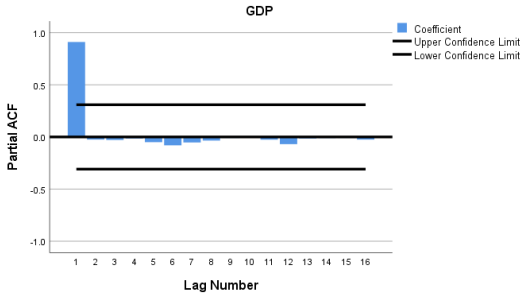


*Figure 3: Original time series PACF graph of GDP in Shaoguan City*

From the ACF and PACF graphs in Figures 2 and 3, it can be seen that the ACF graph exhibits a trailing trend, with the coefficients gradually decaying and then increasing. Most of the coefficients fall outside the 95% confidence interval. The PACF graph shows a sudden decrease in second-order coefficient, which manifests as a first-order truncation. Therefore, it can be judged that the original time series is non-stationary.

### *3.3. ARIMA model recognition*

Because the original time series is not stationary, it needs to be performed difference processing. Firstly, the data are made a first-order difference. The first-order difference sequence is shown in Figure 4. The sequence after the first-order difference shows a trend of slow growth. Then the ACF and PACF graphs are continued to be tested, and the results are shown in Figures 5 and 6. From Figures 5 and 6, it can be seen that the ACF graph also shows a trailing trend. The coefficients gradually decrease and then increase. A small number of coefficients fall outside the 95% confidence interval. The PACF graph is shown as a first-order truncation, with the coefficients fluctuating up and down the 95% confidence interval around the 0 axis. It indicates that the original sequence data have not yet reached the requirement of stationarity, and the data need to be processed by second-order difference.

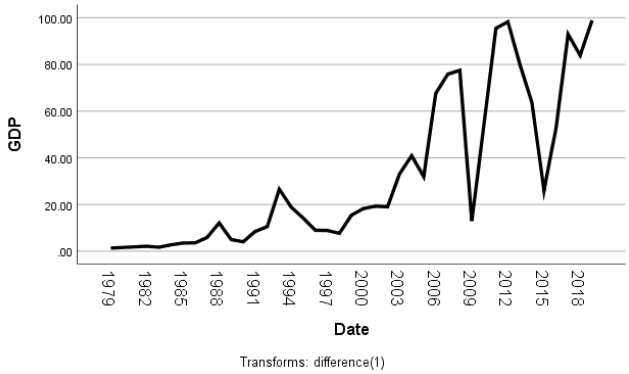


*Figure 4: First-order difference sequence of GDP in Shaoguan City*

ISSN 2616-5902 Vol. 5, Issue 21: 117-122, DOI: 10.25236/AJBM.2023.052117

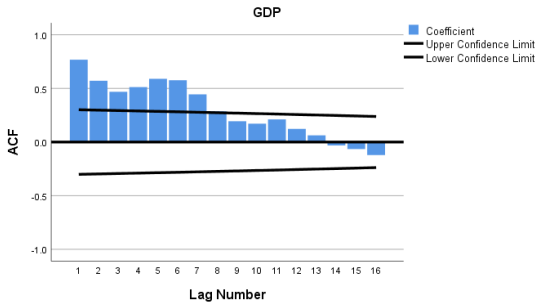


*Figure 5: First-order difference sequence ACF graph of GDP in Shaoguan City*

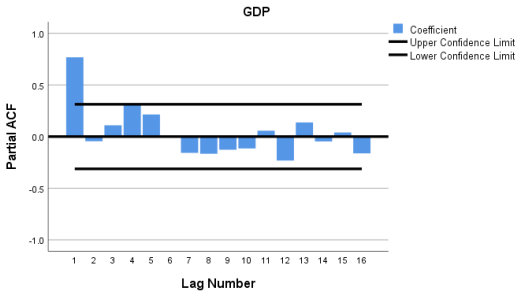


*Figure 6: First-order difference sequence PACF graph of GDP in Shaoguan City*

The sequence graph after the second-order difference is shown in Figure 7. It can be seen from the figure that the sequence graph after the second-order difference becomes a stationary time series, and the data fluctuate around a constant value. From the ACF and PACF graphs of the second-order difference sequence in Figures 8 and 9, it can be seen that both the ACF and PACF graphs are both of order 0 trailing, and the coefficients basically fall within the 95% confidence intervals. Therefore, the sequence data after the second-order difference are stationary sequence data. So it can be determined that the value of parameter p in the ARIMA model is 0, the value of d is 2, and the value of q is 0. According to the values of the three parameters, the ARIMA model in this paper is determined as ARIMA(0,2,0).

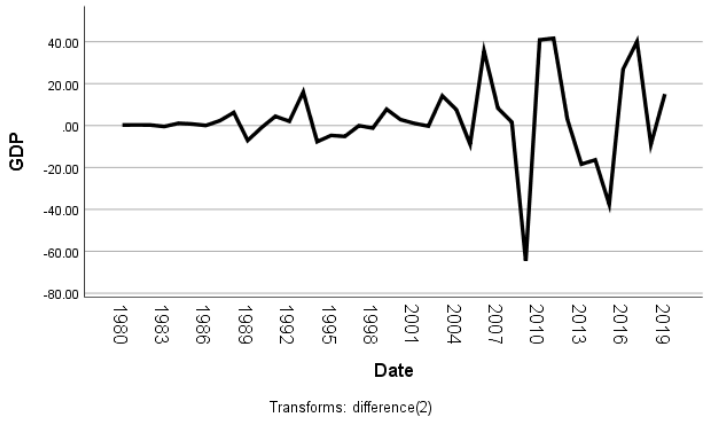


*Figure 7: Second-order difference sequence of GDP in Shaoguan City*

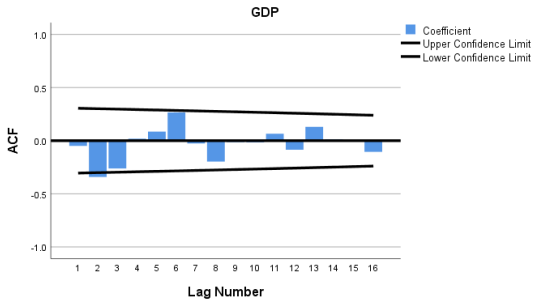


*Figure 8: Second-order difference sequence ACF graph of GDP in Shaoguan City*

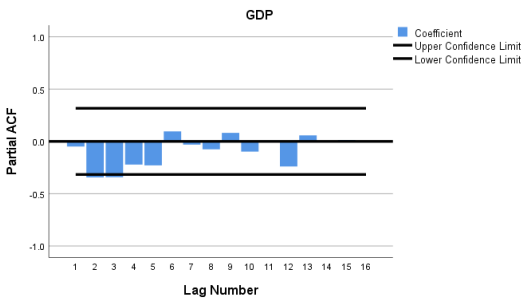


*Figure 9: Second-order difference sequence PACF graph of GDP in Shaoguan City*

### *3.4. Modeling and model test*

The ARIMA (0, 2, 0) is established in SPSS software according to the three parameters p, d, and q, and the option of automatic detection of outliers is selected. The results of the statistical analysis are shown in Figure 10 and Table 1. It can be seen that the model fits well, and the values of stationary R-squared and R-squared are relatively high. According to the statistic Ljung-Box Q (18), the significance is greater than 0.05, indicating that the model passes the test. Two outliers are detected.

ISSN 2616-5902 Vol. 5, Issue 21: 117-122, DOI: 10.25236/AJBM.2023.052117

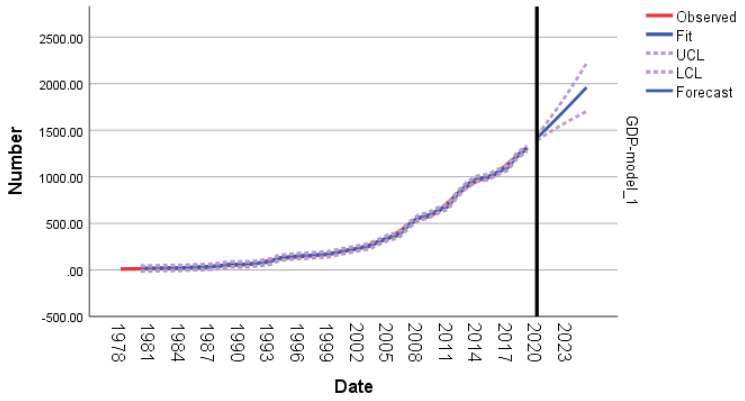


*Figure 10: Model fitting effect*

*Table 1: Model statistics*

| Model | Model Fit statistics | | Ljung-Box Q (18) | | | Number of outliers |
|---|---|---|---|---|---|---|
| | Stationary R-squared | R-squared | Statistics | DF | Sig. | |
| GDP-model_1 | .536 | .999 | 17.340 | 18 | .500 | 2 |

### *3.5. Residual white noise test*

After the model is established, it is also necessary to test whether the residual sequence of the model is a white noise sequence to ensure that there is no autocorrelation in the residual. The white noise sequence is tested by the ACF and PACF graphs, as shown in Figure 11. As can be seen from the figure, after second-order difference processing of the original sequence, all autocorrelation coefficients and partial autocorrelation coefficients of the residuals basically fall within the 95% confidence interval. The residual sequence is randomly distributed. Therefore, the residual sequence of the model is a white noise sequence, and the ARIMA (0, 2, 0) model passes the residual white noise test.

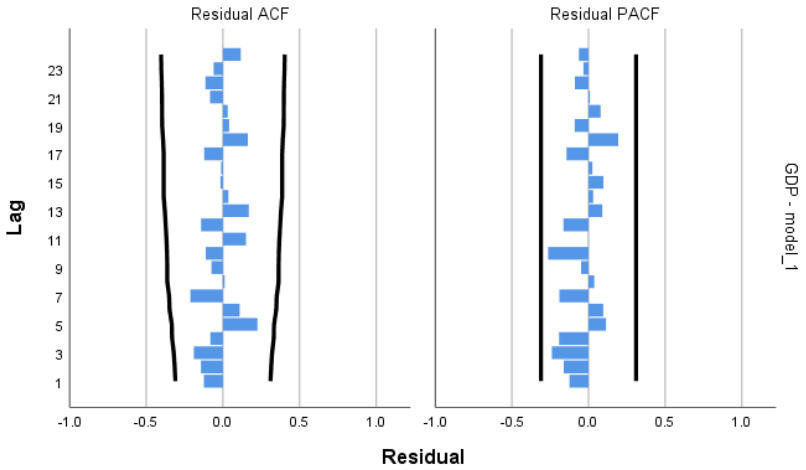


*Figure 11: ACF and PACF graphs of model fitting residuals*

### *3.6. Model prediction*

In order to further test the accuracy of model fitting and prediction, the ARIMA (0, 2, 0) model is used to forecast the GDP of Shaoguan City from 2020 to 2022. The actual GDP values from 2020 to 2022 are used as the observed values to test the accuracy of the model prediction, as shown in Table 2. As can be seen from the table, the relative error values between the predicted and actual values of GDP in Shaoguan City from 2020 to 2022 are small, and the relative error values are less than 5%. The fitting and prediction results of the model are good. The results indicate that the ARIMA (0,2,0) model constructed in this paper has high reliability. Finally, the model is further used to predict the GDP of Shaoguan City in the next three years.

ISSN 2616-5902 Vol. 5, Issue 21: 117-122, DOI: 10.25236/AJBM.2023.052117

*Table 2: Comparison of the predicted and actual GDP in Shaoguan City from 2020 to 2022*

| Year | Actual value | Predicted value | Relative error(%) |
|---|---|---|---|
| 2020 | 1375.16 | 1417.65 | 3.09 |
| 2021 | 1553.93 | 1521.40 | 2.09 |
| 2022 | 1563.90 | 1627.58 | 4.07 |

The prediction results of Shaoguan's GDP in the next three years are shown in Table 3. As can be seen from the table, the GDP of Shaoguan in the next three years will be 173.621 billion yuan, 184.727 billion yuan and 196.077 billion yuan respectively, indicating that the GDP of Shaoguan in the next three years will show a trend of steady growth.

*Table 3: Forecast results of GDP in Shaoguan City from 2023 to 2025*

| Year | Predicted value | UCL | LCL |
|---|---|---|---|
| 2023 | 1736.21 | 1884.42 | 1587.99 |
| 2024 | 1847.27 | 2047.95 | 1646.58 |
| 2025 | 1960.77 | 2218.90 | 1702.63 |

## 4. Conclusion

In this paper, the ARIMA (0, 2, 0) model is constructed by selecting the GDP data of Shaoguan City from 1978 to 2019 as the training set, and the GDP data of Shaoguan City from 2020 to 2022 are used as the test set to verify the fitting of the model. Finally, the model is used to predict the GDP of Shaoguan City in the next three years. It can be seen from the results that the relative error values of the actual and predicted values from 2020 to 2022 are no more than 0.05. The model has a good fitting effect, which can accurately predict the GDP of Shaoguan City in the future. From the predicted results, it can be seen that the GDP of Shaoguan City will maintain a stable growth trend in the next three years, providing a certain reference for decision-making institutions to formulate relevant policies for the future development of Shaoguan. At present, Shaoguan is promoting high-quality development of the economy through a series of measures such as attracting investment.

## References

*[1] Wang Meina and Yang Xiaobin. Based on GM(1,1) Forecast GDP of Guizhou Province and Grey Correlation Analysis of GDP and Industrial Structure [J]. Mathematics in Practice and Theory, 2021, 51(4): 180-188.(In Chinese)*

*[2] Li Nan. Prediction of Jiangxi Province GDP Based on BP Neural Network [J]. Science Mosaic, 2017(10): 47-49. (In Chinese)*

*[3] Wu Boni. Forecast of Hangzhou City GDP based on time series model [J]. Enterprise Reform and Management, 2019(23): 212-213. (In Chinese)*

*[4] Xia Ruyu and Wang Ziqiao. Forecast and analysis of Chongqing City GDP based on ARIMA model [J]. China Storage & Transport, 2022(8): 93-94. (In Chinese)*

*[5] Wang E and Zhang Ting. Application of Time Series in GDP Forecasting of Hunan Province——Based on ARIMA Model [J]. Journal of Qingdao University (Natural Science Edition), 2019, 32(3): 136-140. (In Chinese)*

*[6] Xiong Zheng and Che Wen-gang. Application of ARIMA-GARCH-M model in short-term stock forecasting [J]. Journal of Shanxi University of Technology: Natural Science Edition, 2022, 38(4): 69-74. (In Chinese)*