ATLANTIS PRESS

Advances in Economics, Business and Management Research, volume 76

4th International Conference on Financial Innovation and Economic Development (ICFIED 2019)

Analysis and Forecast of Resource-based City GDP Based on ARIMA Model

Huang Hao

Shengli College China University of Petroleum. department of Chinese Law and Economics Management,Dongying, Shandong,China

13805460321@139.com

*Huang Hao

Keywords: ARIMA model, GDP predicted, stable development, strategic adjustment

Abstract: As an important indicator to measure a country or region, GDP is often used to measure the level of economic development of a country or region. Doing a good GDP forecast is beneficial to the real-time and adjustment of government policies, and has important theoretical and guiding significance. This paper selects the GDP data of Dongying City from 1978 to 2016 as a research sample, and predicts the GDP of Dongying City from 2017 to 2020 by constructing a time series ARIMA model. The research results show that Dongying City's GDP has a first-order single-integration nature, and the current GDP will be affected by the interference of the last four periods of GDP and the last three disturbances. Due to the influence of uncertain factors such as the international situation, economic foundation and scientific and technological progress, the accuracy of model prediction decreases, and the actual GDP is underestimated. However, the future trend of the time series can be roughly judged according to the prediction of the ARIMA model, preparations can be made in advance to ensure the smooth development of the city's economy, and provide decision-making reference for resource-based city governments, relevant institutional policy formulation, and strategic adjustment.

1. Introduction

GDP is an important comprehensive statistical indicator in the accounting system. It is also a core indicator in China's national economic accounting system. It reflects the economic strength and market size of a country or region [1]. It is usually evaluated whether a country's economy is growing or declining, and it can be judged based on changes in GDP. The growth or decrease in GDP indicates whether the region's economy is expanding or entering a recession. GDP is often used as an important variable in the model. Its main purpose is to guide the development of the national economy and to have a very important significance for its prediction. Estimating GDP using various models and algorithms has been one of the research hotspots of scholars at home and abroad. This paper selects the GDP of the resource-based city Dongying City from 1978 to 2016 as sample data, explores the law of data change, establishes the ARIMA model, and uses this model to predict the development of the city's GDP in the next four years, and provides policies for the government to formulate economic development. support.

A time series is a sequence of numbers that are arranged by successive observations of the same phenomenon at different times. The time series prediction method reveals the law of the phenomenon changing with time through the historical data of the time series, and extends this law to the future to predict the future of the phenomenon [2]. The ARIMA model is called Autoressive Integrated Moving Average Model. A shorthand ARIMA is a famous time series prediction method proposed by Box and Jenkins. It is also called Box-Jenkins model. [3]. The AR in ARIMA(p,d,q) is autoregressive, MA is the moving average, P is the autoregressive term, g is the moving average term, and d is the number of differences made when the time series becomes stationary. The ARIMA model needs to transform the non-stationary time series into a stationary time series (Generally, the difference is obtained once.), and the dependent variable then regresses its lagged value and the present value of the random error term as well as the lagged value.

2.ARIMA Model theory basis

The ARMA(p,q) model is primarily an analytical model for stationary time series. In fact, most of the sequences in reality are non-stationary, so the analysis of non-stationary sequences is more common and more important. For non-stationary sequences, we usually use the summation autoregressive moving average model, the ARIMA (p, d, q) model for fitting. We refer to the model of the following structure as the ARIMA model:

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t \quad (1)$$

$$E(\varepsilon_t)=0, \quad Var(\varepsilon_t)=\sigma_\varepsilon^2, \quad E(\varepsilon_t,\varepsilon_s)=0, \quad t \neq s$$

$$E(X_s,\varepsilon_t)=0$$

$$\forall s < t$$

In the formula, $\Phi(B)=1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_p B^p$, which is an autoregressive coefficient polynomial of the stationary reversible $ARMA(p,q)$ model; $\Theta(B)=1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q$, which is a moving smoothing coefficient polynomial for the stationary reversible $ARMA(p,q)$ model.

3. Empirical analysis

This paper selects the 1978-2016 GDP data of the resource-based city Dongying City as a research sample, introduces the method of constructing the ARIMA model for time series, and predicts the future trend of Dongying GDP on this basis.

3.1 Descriptive statistical analysis

Before the prediction, the paper first descriptive statistical analysis of the sample data. From the timing chart, the GDP of Dongying City shows a fluctuating upward trend in the long run, which is consistent with the economic growth characteristics of resource-based cities. The steepness of the curve indicates the growth rate of GDP. In the initial stage of resource development, the GDP growth rate of Dongying City was small. After 1990, due to the development of resources, the economic growth rate accelerated. During the period of 2000-2007, the economic growth rate was relatively fast. Due to the impact of international oil prices, the GDP growth rate in 2010-2014 has cooled down. In 2015, the GDP of Dongying City has also entered a new normal, with the growth rate slowing down, and oil prices, The resource stock trend is consistent. The right distribution histogram and descriptive statistical analysis table show that during the sample period, the average GDP of Dongying is 90.988 billion yuan, and the standard deviation is 1178.001, which indicates that the GDP of Dongying has fluctuated greatly over the years. The JB statistic is 8.867, and the corresponding P value is 0.012 and less than 0.05, indicating that the Dongying GDP does not obey the normal distribution.

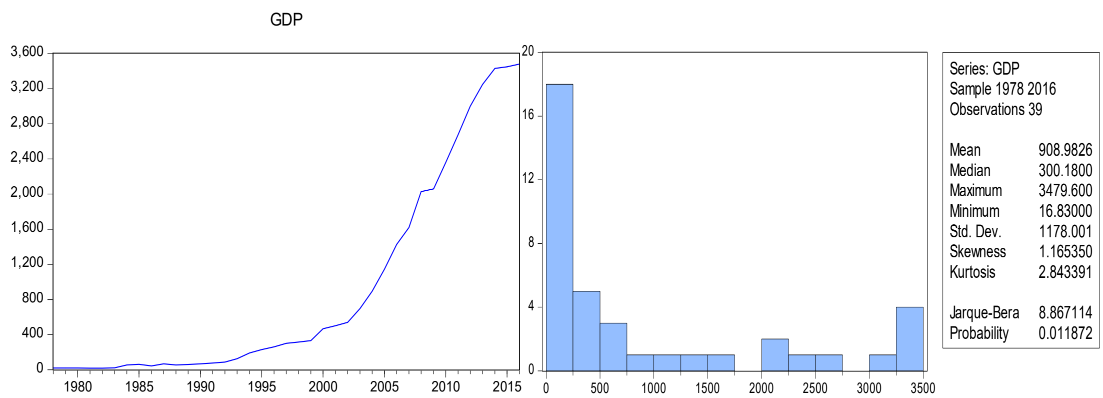


Fig.1 Sample data timing diagram

3.2 ARIMA Model prediction

This paper predicts the future trend of Dongying GDP by constructing the ARIMA model, which is often used for short-term forecasting of macroeconomic variables. Before the analysis, in order to eliminate the possible heteroscedasticity, the natural logarithm of GDP is processed, and then the ADF unit root method is used to test the stability of LNGDP. The specific results are shown in the following table:

Table 1 LNGDP stability test

Variable	(c,t,p)	ADF value	5% of the critical value	10% of the critical value	Stationarity
LNGDP	(c,t,3)	-3.014370	-3.544284	-3.204699	Non-stable
DLNGDP	(c,0,0)	-6.011601	-2.943427	-2.610263	Stable

Note: In (c, t, p), c denotes an intercept term, t denotes a linear trend, and p denotes a lag length.

It can be seen from the ADF test that the ADF value of LNGDP is -3.01 greater than the critical value of 0.05 significance level -3.54, so the null hypothesis that the unit root is present should be accepted, and the original sequence is considered to be non-stationary, while the first-order differential variable DLNGDP is its ADF. The value is less than 0.05 significance level test value, so the difference after the difference is considered to be stable. Therefore d=1 in ARIMA(p,d,q).

3.3 Autocorrelation analysis

In this paper, the order of the model is determined by the autocorrelation graph method. The model of this paper is tentatively defined as ARIMA(p,1,q). It can be seen from the autocorrelation graph that the partial autocorrelation coefficients appear truncated after the first order, and the autocorrelation coefficients exhibit the characteristics of tailing. Therefore, this paper can set the model to the ARIMA (1, 1, q) form. The final model is determined below based on the AIC and SC criteria.

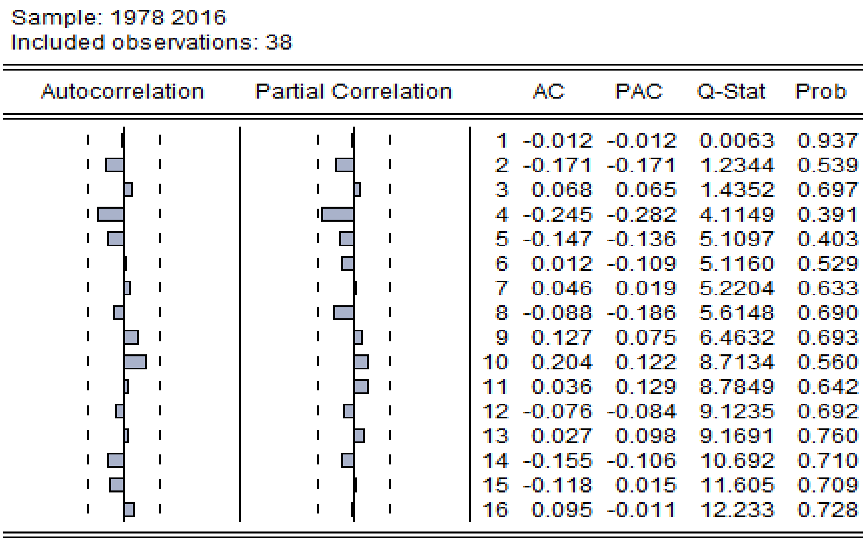


Fig.2 Autocorrelation diagram

After 4 steps of the autocorrelation function of DLNGDP, it oscillates back and forth to zero, and after 4 steps of the autocorrelation function, it slowly decays to zero. Therefore, we can initially judge that the values of p and q in ARIMA(p,1,q) are not more than 4.

3.4 Determination of model order

In the above, we roughly judge the range of p, q through the autocorrelation function and the partial autocorrelation function graph. The following is the optimal calculation based on the model AIC and SC criteria, the model goodness of fit and the significance of each variable. The model, the

ARIMA part of the model with different orders is summarized as follows:

Table 2 Summary table of ARIMA part models with different orders

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
					P value				
C	0.0000***	0.0001***	0.0000***	0.0000***	0.0012***	0.0000***	0.0011**8	0.0000***	0.0000***
AR(1)	0.6108	0.0356**	0.9937	0.7580		0.0406**		0.7252	0.0000***
AR(2)	0.8843	0.1061	0.3022	0.0319**	0.0072***	0.0013***	0.0557*	0.0332**	
AR(3)	0.2433	0.0159**	0.0128**		0.8037		0.0808*		
AR(4)	0.7259	0.5482			0.1806	0.0750*			
MA(1)	0.9609	0.0901*	0.5924	0.4716				0.1428	0.0000***
MA(2)	0.2374	0.0979*	0.2937	0.0000***	0.0000***	0.0000***	0.0043***	0.0000***	
MA(3)	0.0002***	0.0000***	0.0000***	0.8526	0.0000***	0.0201**	0.0000***		
MA(4)	0.2195		0.0144**						
Adj-R^2	0.5792	0.4264	0.4414	0.4724	0.6291	0.6417	0.4168	0.4970	0.4255
AIC	-0.9285	-0.6385	-0.6897	-0.7858	-1.0955	-1.1525	-0.7127	-0.8563	-0.7847
SC	-0.5244	-0.2793	-0.3342	-0.5219	-0.7813	-0.8831	-0.4905	-0.6364	-0.6541

Note: *, ** and *** denote 10%, 5%, and 1% significance levels, respectively.

It can be seen from the above table that the adjusted goodness of the model 6 is relatively maximum at 0.6417, and the corresponding AIC and SC values are the smallest. All the order variables pass the test of significance level. Therefore, the optimal model in this paper is model 6, we The estimated results of Model 6 can be obtained:

Table 3 Estimation results of Model 6

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.140125	0.023084	6.070329	0.0000
AR(1)	-0.214539	0.099947	-2.146530	0.0406
AR(2)	-0.489859	0.136955	-3.576791	0.0013
AR(4)	-0.195942	0.105973	-1.848978	0.0750
MA(2)	0.948167	0.016882	56.16312	0.0000
MA(3)	0.080575	0.032696	2.464370	0.0201

R-squared	0.695983	Mean dependent var	0.155796
Adjusted R-squared	0.641694	S.D. dependent var	0.209844
S.E. of regression	0.125610	Akaike info criterion	-1.152484
Sum squared resid	0.441781	Schwarz criterion	-0.883126
Log likelihood	25.59223	Hannan-Quinn criter.	-1.060625
F-statistic	12.82003	Durbin-Watson stat	1.701711
Prob(F-statistic)	0.000002		

Inverted AR Roots	.27-.55i	.27+.55i	-.38-.62i	-.38+.62i
Inverted MA Roots	.04+.98i	.04-.98i	-.08	

From the OLS estimation results, the equation of the model can be written as:

$$\nabla LNGDP_t = 0.140 - 0.215\nabla LNGDP_{t-1} - 0.490\nabla LNGDP_{t-2} - 0.196\nabla LNGDP_{t-4} + \varepsilon_t + 0.948\varepsilon_{t-2} + 0.081\varepsilon_{t-3} Var(\varepsilon_t) = 0.1256$$

3.5 Model diagnosis and testing

After constructing the model, it is also necessary to perform a white noise test on the residual sequence of the model. If the model satisfies the white noise sequence, the model fit is valid. Therefore, it is necessary to test the stationarity and autocorrelation of the residual sequence. The specific results are as follows:

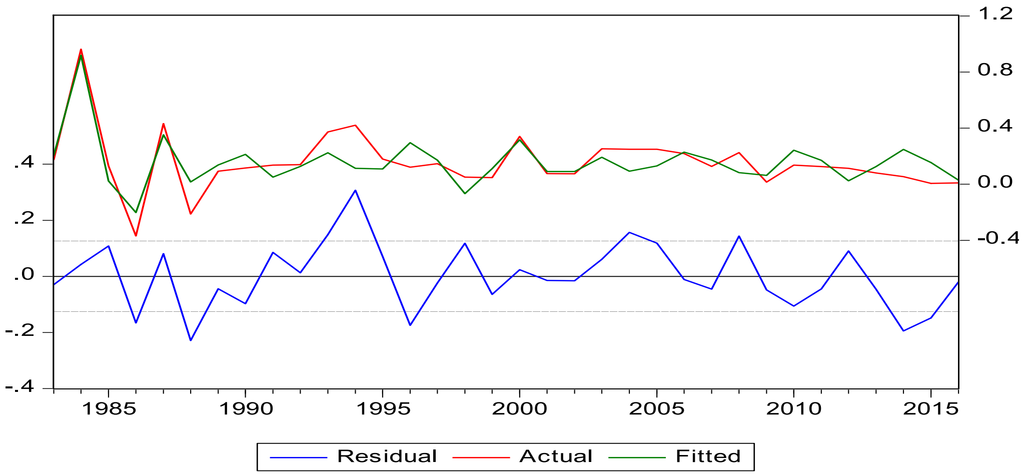


Figure 3 Stationarity and autocorrelation test of the residual sequence

It can be seen from the residual graph that the fitted value is consistent with the change of the actual value, and the residuals fluctuate around the zero mean. It can be judged that the residual sequence of the model is stable. For rigor, continue to perform ADF unit root test on the residual sequence, the results are as follows:

Table 4 Residual ADF unit root test

		T Statistics	P value
ADF Statistics		-3.818147	0.0004
	1% level	-2.639210	
	5% level	-1.951687	
.	10% level	-1.610579	

Note: The threshold here comes from MacKinnon (1996) one-sided p-values.

It can be seen from the ADF test that the residual sequence of the model is stable at a significance level of 0.01. The LM test autocorrelation is performed on the residual sequence below. The results are as follows:

Table 5 residual LM test

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.567723	Prob. F(2,26)	0.5737
Obs*R-squared	1.418813	Prob. Chi-Square(2)	0.4919

It can be seen from the LM test that the F statistic and the P value of Obs*R-squared are both greater than 0.05, so the assumption that the residual has no autocorrelation is accepted, so the residual sequence satisfies the white noise sequence.

3.6 Stability test

The stability of the model is very important for the prediction of the model. Therefore, the stability of the model needs to be tested before the prediction. The AR/MA root graph method is used for the test.

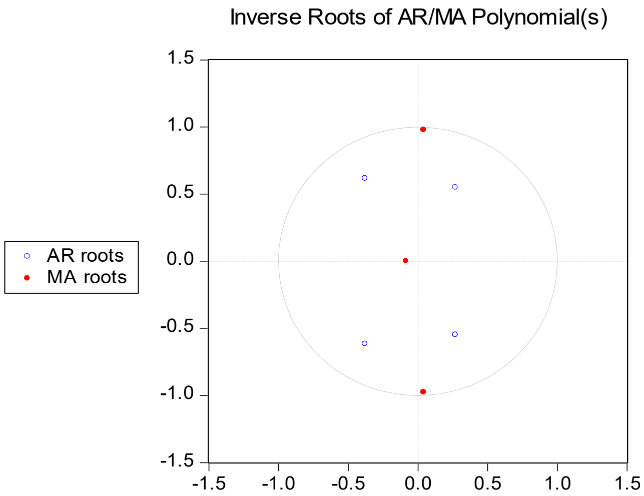


Fig.4 AR/MA root graphic test

It can be seen from the figure that the reciprocals of all the root modes fall within the unit circle, indicating that the constructed model is stable.

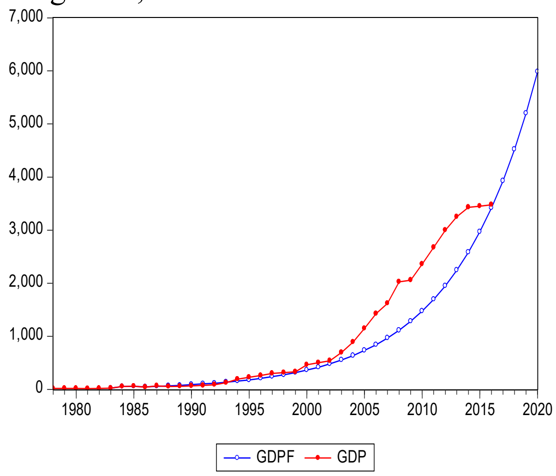
3.7 Model prediction

The prediction of the model is divided into static prediction and dynamic prediction. Static prediction is suitable for intra-sample prediction. Dynamic prediction is often used for extra-sample prediction. Except that the first prediction value is the actual value of the explanatory variable, the latter values are obtained by recursive method. . This paper mainly makes an off-sample forecast of Dongying's GDP. We can get the predicted value of Dongying GDP in 2017-2020 as follows:

Table 6 Forecast of Dongying GDP from 2017 to 2020

Year	GDP（Unit: billion yuan）
2017	392.9363
2018	452.0378
2019	520.0321
2020	598.256

And we can get the actual value and the predicted value trend chart and the accuracy of the model fitting table, as follows



Forecast: GDPF	
Actual: GDP	
Forecast sample: 1978 2020	
Adjusted sample: 1983 2020	
Included observations: 34	
Root Mean Squared Error	464.2629
Mean Absolute Error	285.6019
Mean Abs. Percent Error	23.11776
Theil Inequality Coefficient	0.167873
Bias Proportion	0.358544
Variance Proportion	0.371838
Covariance Proportion	0.269618

Fig.4 Trend graph of actual and predicted values and accuracy of model fitting

It can be seen from the timing diagram of the actual value and the fitted value that the model can better fit the trend of the Dongying GDP. From the perspective of the fitting effect, the actual GDP value of Dongying City in 1985-2004 is close to the predicted value. After 2005-2015, the fitting effect of the model is poor, and the fitting curve is below the actual curve during this period. The model underestimates the GDP of Dongying, but as time goes by, the distance between the two decreases gradually, indicating that the model can predict the future trend of Dongying GDP to a certain extent. The graph on the right is the prediction accuracy table. Although the Tyre inequality coefficient is smaller than 0.168, the mean square error percentage is 23.12%, and the percentage of deviation and the percentage of variance are larger than the percentage of covariance, indicating that the prediction accuracy of the model is not enough. This is because with the gradual maturity of China's accession to the WTO, GDP is affected by uncertain factors such as the international situation, economic foundation, scientific and technological progress, price changes, etc., but it is difficult to eliminate these factors in actual forecasting. In addition, we found that the model prediction is difficult to highlight the impact of major event shocks. For example, due to the financial crisis of 2007-2009, Dongying GDP showed a certain downward trend, but it was not reflected in the forecast curve.

4. Conclusion

Resource-based cities rely on natural resources as the leading industry, and the economic structure is relatively simple. Early resource-based cities relied on policy support and their strong resource advantages to develop rapidly, and GDP growth was relatively rapid. As the level of economic development continued to increase, the economic structure of single GDP growth was weak, and the problems of depletion of resources in the middle and late stages, especially The GDP growth has further slowed down. Especially in the development of regional economy, the development of resource-based cities is more and more lax. In general, the growth rate of GDP in resource-based cities will be expected to be closely related to the development cycle, experiencing rapid growth - slow whole-slow-slow growth phase, with a phased growth rate.

To a certain extent, the ARIMA model can use the analysis of past and present values from a quantitative perspective to predict the future development of the time series. However, the ARIMA model is only suitable for short-term prediction. As the prediction period is extended, the prediction error will increase accordingly, so the predicted data does not fully represent the actual data. However, we can still judge the future trend of the time series based on the prediction of the ARIMA model. On the one hand, it is beneficial to make measures in advance to ensure the stability and long-term nature of urban economic development; on the other hand, it can provide reference for resource-based cities economic development.

Acknowledgment

Fund of the project: The Social science planning research project of Shandong Province: “The research on the integration and innovation between Amoeba model and management accounting of petroleum equipment manufacturing industry" (NO.:17CKJJ04).

Reference

[1] Sun Hechao. Application of ARMA Algorithm in GDP Forecasting [J].Jin Tian,2013,06：389.

[2] Li hui,Tian Zichen,Li Bubei,Liu Miao. Prediction of GDP in Yili Based on Time Series Analysis Method [J].Mathematics In Pratice And Theory,2017,47(13):15-23.

[3] Gong Guoyong. Application of ARIMA Model in Shenzhen GDP Forecasting [J]. Mathematics In Pratice And Theory,2008,38(4):53-57.