

The Current Situation and Prediction of Urbanization in China Based on ARIMA Model

Yue Zhou

School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China

small_potato0303@163.com

*Yue Zhou

Keywords: Urbanization, Prediction, ARIMA Model.

Abstract. This document introduces the current situation of urbanization in our country, and analyzes the development trend and the existing problems in the process of urbanization in our country. Then using the basic theory of ARIMA model in sequence theory and combining with the data of urbanization level in our country, this article use E-views and SPSS software to model and predict the urbanization rate data of China from 1982 to 2016. The results show that using ARIMA (0,2,1) model to predict the level of urbanization in China is more accurate in short-term data, which indicates that the urbanization level in our country develops more rapidly and shows a gradual upward trend.

1. Introduction

With the further deepening of the reform of the economic system, China's economy has entered a phase of rapid growth. At the same time, China's urbanization has also been further strengthened and the level of urbanization has been greatly increased, showing growth trend year by year. Is the growth trend steady and sustainable? With all sorts of questions, I measured the level of urbanization in our country. The measure of urbanization refers to applying scientific and standardized methods and according to the principle of comprehensively and truly reflecting the status and level of urbanization in the region. The measure of urbanization refers to applying scientific and standardized methods and according to the principle of comprehensively and truly reflecting the status and level of urbanization in the region.

2. Literature references

At present, it is very extensive for researchers to use ARIMA model to analyze and predict China's macroeconomic situation. Wang Longbing, Chen Xizhen (2012) made a short-term forecast of China's GDP using ARIMA (1,1,1) model, showing high accuracy [3]. Jiali Wen and Xu Deyi (2014) used ARIMA model to predict the demand of China's iron ore [4]. Zhu Yanke (2013) used ARIMA (2,2,2) model to predict the energy consumption in Guangdong province [8]. The results show that the fitting effect is better.

In addition to the above macroeconomic forecasts, there are many literature on the use of time series to study the level of urbanization, but the use of specific models is not the same. Shi Lei, Shi Wanqun (2010); Wang Xiaoxin and Yang Chen (2013) used time series models to analyze the urbanization level in different regions. The former used OLS estimation to establish the time series model of urbanization rate in Chongqing. The model carries out parameter estimation and model test [9]. The latter used VAR model to make an empirical study on the relationship between changes in expenditure levels of social security and social security funds, and changes in urbanization level [5]. Ma Jun (1999) predicts the level of urbanization in China with a 10-year span using logistic curve models, but does not predict the relevant data for each specific year in 10 years [11], while Shima Hamidi, Reid Ewing(2014) predict changes in urban through a longitudinal aspect [2]. Using the ARIMA (4,2,2) model, Wang Hao (2012) analyzed the farmer's income and choices and predicted

that rural household income will continue to grow, but the growth rate will decline slightly [13]. Among them, there are a lot of articles devoted to forecasting the level of urbanization in China. For example, Chen Fukai (2014) found that it is more accurate to use ARIMA (0, 1, 5) to predict the short-term data of the level of urbanization in China, and our urbanization level has enjoyed a good momentum of development [12].

3. The present situation of urbanization in China

3.1 The development of urbanization in China

Urbanization is a natural and historical process accompanied by industrialized development, the gathering of non-agricultural industries and the concentration of rural population towards to cities and towns. After the implementation of the reform and opening up in 1978, China's urbanization has entered a new stage of development. Reform and opening up policy has become an important driving force for the development of urbanization. With the profound changes in the development of the environment inside and outside our country, urbanization has entered a new stage which focus on the promotion of quality. In the new stage, the key to the development of urbanization in our country lies in taking people as the core.

3.1.1 The stage of promoting rural reform

This stage was mainly the stage of urbanization promoted by rural economic system reform and rural industrialization. Since the early 1950s, in order to tie in with the country's strategy of giving priority to the development of heavy industry, China has begun to limit the flow of rural population to cities step by step. At this stage, the urbanization of our country was in the stage of restoration and development. The level of urbanization increased from 17.92% in 1978 to 23.01% in 1984.

3.1.2 The stage of promoting urban reform

At this stage, there were two driving forces which promoted the development of urbanization--the reform of cities and township enterprises. In 1984 and 1986, the state successively liberalized the standards of established towns and cities and the number of established cities increased substantially. In 1992, the number of established cities in the country reached 517, an increase of 217 over 1984. The town has been built from 9,140 to 14,539. The level of urbanization rose from 23.01% to 27.46%, an increase of 4.45 percentage points from 1984 and an annual increase of 0.56 percentage points.

3.1.3 The stage of promoting market economy system reform

This stage was the fastest stage of urbanization in China's history. Compared with 1992, the number of cities under construction increased from 517 to 660 in 2002, and the number of established towns reached 20,226. The urbanization rate increased from 27.46% to 40.53%, an average annual increase of 1.19 percentage points.

3.1.4 The stage of harmonious development of urban-rural area

At this stage, the urban development started to shift from quantity expansion to quality improvement. Urbanization in our country continues the trend of rapid development. However, the rate has been adjusted with the urbanization rate rising by 1.07 percentage points on average each year, with an average annual increase of 2.3%. The number of established cities has been reduced, but the size of individual towns has rapidly expanded. As of the end of 2013, the urban resident population was 73,111,000, an increase of 207,350,000 over 2003, representing an average annual increase of 20,735,000. The urbanization rate in China reached 53.7%, an increase of 13.17 percentage points from 2003.

3.1.5 The stage of urbanization transformation and development based on human beings

The specific background of China's urbanization has experienced or is about to undergo profound changes, which are mainly reflected in the highly integrated growth, transformation and reform in China in the next five to 10 years. Therefore, in the past, large-scale urbanization aimed at increasing the total amount of GDP was difficult to survive. Urbanization must enter a new stage of

transformation and development with emphasis on improving quality. At the core of human beings, urbanization became the key to the development of China's urbanization in the new stage.

Urbanization must enter a new stage of transformation and development with emphasis on quality improvement. The traditional model of industrialized urbanization with high input, high consumption and high emission is unsustainable. Depending on the cheap supply of labor force, extensive depletion of resources such as land, the non-equalization of basic public services to reduce costs and promote the rapid urbanization is unsustainable. The development of China's urbanization from speed to quality is imperative. Taking people as the core is the key to the development of urbanization in our country.

3.2 The achievements in China's urbanization

Since the reform and opening up, the urbanization in our country has maintained a sustained and rapid development. The city has become the core carrier of the national economic and social development. Since the 16th CPC National Congress in particular, in accordance with the principle of overall planning of urban and rural areas, rational distribution, land conservation and functional improvement, we have to stick to the road of urbanization with Chinese characteristics and promote the coordinated development of cities and towns. China's urbanization has achieved Great achievements.

The level of urbanization continued to increase rapidly and the urban system continued to improve. By the end of 2008, China's urban population reached 607 million with an urbanization rate of 45.68%, an increase of 35 percentage points over 10.64% of the initial liberation while a 25.6 percentage point increase from 21.1% in 1982, an average annual increase of 0.95 percentage point. The contact between cities and towns become closer, and urban areas gradually formed. Municipal utilities service capacity and attack capability increased while the investment scale gradually expanded. By the end of 2008, the urban water penetration rate reached 94.73%. Gas penetration rate reached 89.55%. Sewage treatment rate reached 70.16%. Domestic garbage harmless treatment rate reached 66.76%. Urban built-up area green coverage rate reached 37.37%. Per capital park The green area amounts to 9.71 square meters, with 11.31 public transport vehicles per 10,000 people, with a per capital road area of 12.21 square meters. Urban and rural planning work continuously improved and strengthened, and played a guiding role in urbanization and urban development.

3.3 The problems in the development of urbanization in China

In the rapid development of urbanization in our country, there are some problems which have emerged and need to be given high priority and solved. "Land Urbanization" is faster than "Population Urbanization." In recent years, in order to expand urban area, some places emerged such situations as the urban framework has been blindly drawn up, over-occupation of cultivated land, indiscriminate development zones. Spatial planning and urban planning exits issues. The structure of some cities is irrational and the coordination between neighboring cities is not coordinated enough so that it is not conducive to forming a town system and spatial pattern with a clear division of labor and a scientific layout. In addition, the jobs of most migrant workers is unstable. Their social security problem is very prominent. Their medical insurance work has just started while their pension insurance has not yet begun. At present, in the process of urbanization in our country, different government departments manage the process of urbanization from the perspective of their respective responsibilities and their own interests. It is very common for many government departments and long-term management to emerge. This is not conducive to the formulation and implementation of a unified urbanization policy and affected the process of urbanization to develop healthily and efficiently.

4. Empirical analysis

4.1 The stationary test of time series

The Stationary Stability (ADF) test of the urbanization rate from 1982 to 2016 shows the results which are shown in Figure 1. As shown in the figure, the sample ACF discovers the sine-cosine fluctuation of the damping, while the sample PACF censures after the delay of 1, which means the data is not stable. It can be seen from Table 1 that the ADF value is 0.9960, which is larger than the value of the 10% test level. It shows a serious sequence instability. Besides, the auto-correlation coefficient is truncated after the seventh order, and the partial auto-correlation coefficient is truncated after the first order.

Table 1. The ADF test results of Urbanization rate¹.

		T statistics
ADF statistics		2.500651
Significance level	1% level	-2.639210
	5% level	-1.951687
	10% level	-1.610579

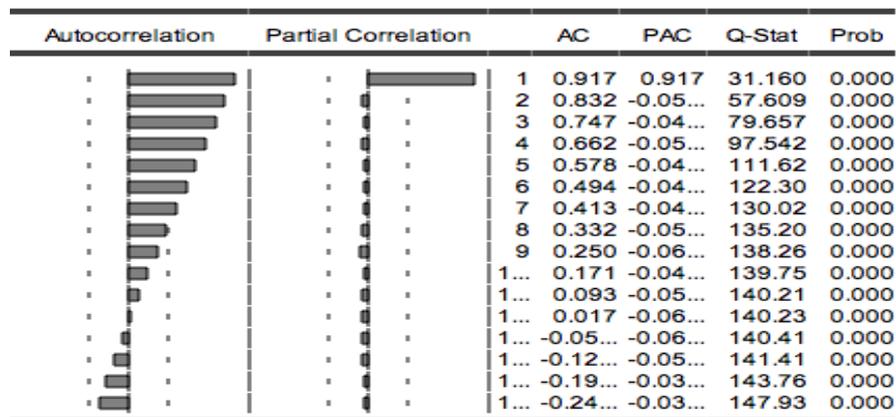


Fig. 1. Autocorrelation and partial autocorrelation coefficients in horizontal conditions.

Deal with data into the first-order difference, and the fluctuations are shown in Figure 2. It shows that the first-order difference shows a steady trend when the data are in the vicinity of the average uniform distribution. Through the figure, we could realize the ADF values gradually approach the 10% significance level after the first difference. But the coefficient is still greater than -2.62, so I think the data is still not stable. In addition, as is shown to us, the autocorrelation coefficient is truncated after the 4th order, and the partial autocorrelation coefficient is truncated after the 1st order. So we could conclude that the data has gradually stabilized after the first-order difference.

Table 2. ADF test results after first difference of urbanization rate².

		T statistics
ADF statistics		-2.356474
Significance level	1% level	-3.653730
	5% level	-2.957110
	10% level	-2.617434

Therefore, take the second-order difference into consideration. According to the unit root test results, as is shown in Table 3, we can know that the data becomes stable and passes the 1% significance level through the second-order difference. Therefore, we consider that the sequence is stationary under the second-order difference. It means the value of “d” is assumed to be 2.

¹ Source: State Statistics Bureau in 2017.

² Source: State Statistics Bureau in 2017.

Table 3. ADF test results after second difference of urbanization rate³

		T statistics	Prob,*
ADF statistics		-8.522062	0.0000
Significance level	1% level	-4.284580	
	5% level	-3.562882	
	10% level	-3.215267	

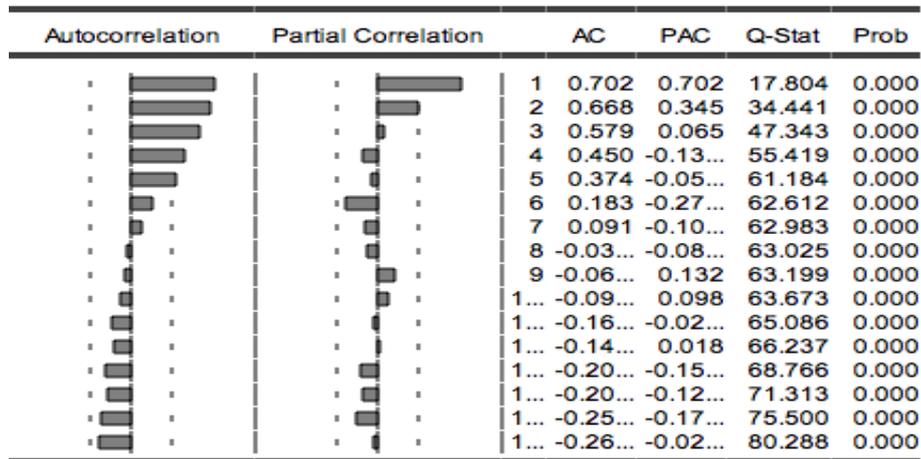


Fig. 2. Autocorrelation and partial autocorrelation coefficients in first order conditions.

Select the lag 32, and then the autocorrelation and partial autocorrelation coefficient diagram after second-order difference are shown in Figure 3. As is shown to us, they are all trailing. But the sequence is significant and gradually decreasing from the first order after the second order difference, so we set the value of “q” is 1 or 0. Similarly, partial autocorrelation functions are significant from the first order, so we pre-set the value of “p” to 0 or 1. There are four models that we could establish: ARIMA(1, 2, 1); ARIMA(0, 2, 0); ARIMA(1, 2, 0); ARIMA(0, 2, 1).

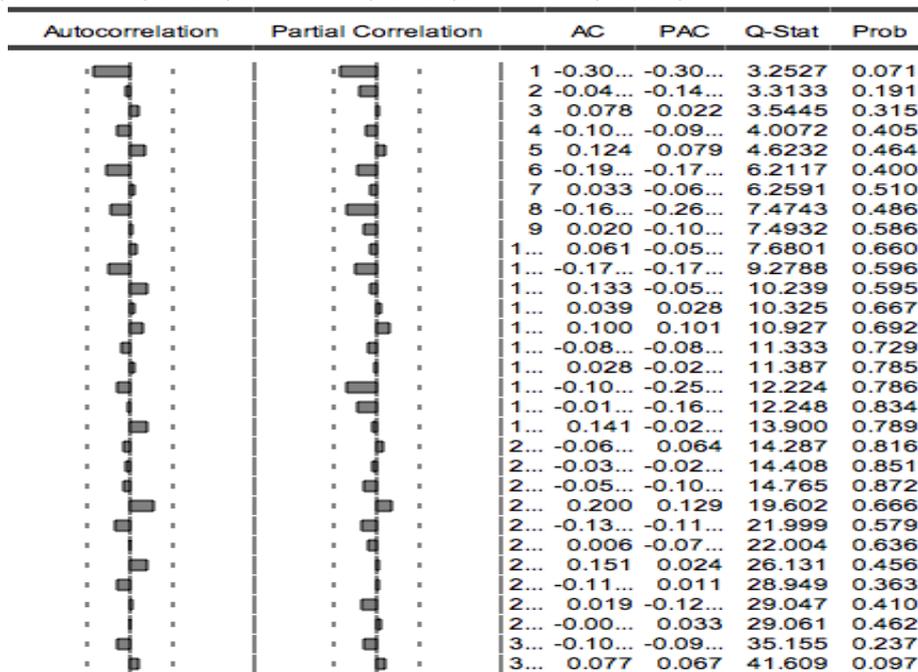


Fig. 3. Autocorrelation and partial autocorrelation coefficients in second order conditions.

³ Source: State Statistics Bureau in 2017.

4.2 The estimation and diagnosis of model

Table 4. AIC values comparison of four models⁴.

ARIMA (p, d, q)	(1, 2, 1)	(0, 2, 0)	(1, 2, 0)	(0, 2, 1)
AIC	0.574	0.531	0.521	0.516

Comparing the AIC values of the four cases, we should choose the model with the smallest AIC value. So we choose the group of parameters “p = 0, q = 1, d = 2”. Therefore, the resulting model is ARIMA (0,2,1). According to the selected parameters, we could make ARIMA model which is shown below.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.016413	0.043565	0.376745	0.7091
MA(1)	-0.395502	0.179610	-2.202002	0.0358
SIGMASQ	0.080856	0.017849	4.530014	0.0001
R-squared	0.135274	Mean dependent var		0.026250
Adjusted R-squared	0.075638	S.D. dependent var		0.310678
S.E. of regression	0.298698	Akaike info criterion		0.515606
Sum squared resid	2.587389	Schwarz criterion		0.653019
Log likelihood	-5.249699	Hannan-Quinn criter.		0.561155
F-statistic	2.268323	Durbin-Watson stat		1.656947
Prob(F-statistic)	0.121545			
Inverted MA Roots	.40			

Fig. 4. The regression results of ARIMA (0,2,1).

```

Estimation Command:
=====
LS(OPTMETHOD=OPG) Y C MA(1)

Estimation Equation:
=====
Y = C(1) + [MA(1)=C(2), UNCOND, ESTSMPL="3 34"]

Substituted Coefficients:
=====
Y = 0.016412922776 + [MA(1)=-0.395502426347, UNCOND, ESTSMPL="3 34"]
    
```

Fig. 5. The regression equation of model.

4.3 The prediction of the model

There are two ways to predict: one is Dynamic and the other is Static. The dynamic one is based on the selected estimation interval, and the multi-step forward prediction is performed. While the static one is rolling-only one-step prediction. For each prediction, the real value is used instead of the predicted value, which is included in the estimation interval and the prediction is made in the previous step.

In order to certify the result, I make a comparison table to show some predicted value and the actual value of data, which is shown in Table 5.

⁴ Source: State Statistics Bureau in 2017.

Table 5. The comparison between predicted value and actual value of the model⁵.

Year ¹	actual value [%]	predicted value [%]	Percentage	
			Relative error value [%]	absolute value of Relative error [%]
2010	49.95	49.68	0.27	0.54
2011	51.27	51.44	-0.17	0.33
2012	52.57	52.66	-0.09	0.17
2013	53.73	53.94	-0.21	0.39
2014	54.77	55.03	-0.26	0.47
2015	56.10	56.01	0.09	0.16
2016	57.35	57.23	0.12	0.23

5. Summary

Based on the AIC minimum criterion, this article selected ARIMA (0,2,1) because it gives us the prediction of urbanization level in China, which is better. However, there are many criteria for judging the model, and the optimal choice of the model is unilateral. In addition, the prediction of the model is only accurate in the short run. The long-term prediction error needs to be factually verified. The level of urbanization in our country will also be affected by the international and domestic environmental influences, including government policies, social concepts, changes in economic level, and sudden disasters. However, the urbanization level in China is still growing. Using ARIMA model to forecast the short-term level of urbanization in our country has certain reference value to the government's urban planning.

References

- [1] Mingxing Chen, Weidong Liu, and Xiaoli Tao, Evolution and assessment on China's urbanization 1960–2010: Under-urbanization or over-urbanization?, *Habitat International*, vol.09, pp. 25-33, 2013.
- [2] Shima Hamidi and Reid Ewing, A longitudinal study of changes in urban sprawl between 2000 and 2010 in the United States, *Landscape and Urban Planning*, vol.04, pp. 72-82, 2014.
- [3] Longbing Wang and Xizhen Chen, Short term forecast of China 's GDP based on ARIMA model, *Science Technology and Engineering*, vol.08, pp. 1671-1815, 2012.
- [4] Liwen Jia and Deyi Xu, Multi-model comparison of iron ore demand analysis and prediction capability--panel model and gray model, coefficient model, ARIMA model, *Resources Science*, vol. 07, pp. 1382-1391, 2014.
- [5] Xiaoxin Wang and Chen Yang, Social security expenditure and urbanization: an empirical analysis based on time series, *Journal of Chongqing Institute of Technology*, vol. 12, pp. 49-55, 2013.
- [6] Yanpeng Gao and Jiadong Yuan, The comprehensive measurement and evolution of population urbanization in northeast China since 2003, *Economic Geography*, vol. 12, pp. 44-50, 2016.

⁵ Source: State Statistics Bureau in 2017.

- [7] Zhibin Xiong, Forecast of GDP time series based on integration of ARIMA and neural network, *Journal of Applied Statistics and Management*, vol. 02, pp. 306-314, 2011.
- [8] Yanke Zhu, Predictive analysis of energy consumption in Guangdong Province based on ARIMA model, *Mathematics in Practice and Theory*, vol. 03, pp. 14-18, 2012.
- [9] Lei Shi and Wanqu Shi, Research on urbanization in Chongqing based on time series model, *Journal of Chongqing College*, vol. 05, pp. 163-165, 2010.
- [10] Jun Ma, Measurement, evaluation and prediction of urbanization level, *ZheJiang statistics*, vol.02, pp. 18-21, 1999.
- [11] Fukai Chen and Letian Xia, Prediction of urbanization level based on APdMA model in China, *Journal of Chongqing Institute of Technology*, vol.04, pp. 133-137, 2014.
- [12] Wang Hao, Prediction of farmers' income and selection of Model ARIMA, *Asian Agricultural Resereh*, vol. 11, pp. 37-41, 2012.