Forecasting with ARIMA Model in Anticipating Open Unemployment Rates in South Sulawesi

Didiharyono D., Muhammad Syukri

Abstract. The purpose this study is to forecasting with using ARIMA model in anticipating open unemployment rates in South Sulawesi. Research method used is applied research with quantitative secondary data obtained at Central Statistics Agency (BPS) South Sulawesi. Research procedures include identification models; analyze of autocorrelation function (ACF) and partial autocorrelation function (PACF); differences data; estimation parameter in model; and forecasting. Based on results of study obtained the best time series model used for forecasting is ARIMA (1,2,1) with the small Mean Square value is 2,0474. General form equation model of ARIMA (1,2,1) is $Z_t = 1{,}9267Z_{t-1} - 0.8534Z_{t-2} - 0{,}0733Z_{t-3} + a_t - 1{,}0504a_{t-1}$.

Index Terms. *Forcasting, ARIMA, and Open Unemployment Rates*.

1. Introduction

Forecasting is used to predict something that will happen in the future so that the right action can be taken. Forecasting through statistical and non-statistical approaches has the goal to estimate the expected results close to the actual data (Bisgaard & Kulahci, 2011). In principle, forecasting is based on time series data according to a certain period (Bakhtiar & Didiharyono, 2018). Forecasting activities are an important thing in human life to anticipate things that happen in the future. One problem predicted is the unemployment rate in South Sulawesi in the future using time series data. Unemployment is a serious problem faced by governments in various countries of the world including Indonesia. Unemployment has a direct effect on poverty, crime and other social and political problems. There are many factors caused by unemployment in Indonesia, apart from low investment, education is also blamed as a factor causing increasing unemployment problems (Dewita Hia, 2017).

Unemployment is an important macroeconomic variable, defines the condition of a country's economic balance and is a barrier to social development (Dritsakis & Klazoglou, 2018). Unemployment is a fundamental economic problem with significant negative social impacts that cause people to become poor. High unemployment has both direct and indirect impacts on poverty, crime and socio-political problems that are also increasing (Nikolaos, Stergios, Tasos, & Ioannis, 2016). With a large enough workforce, a steady flow of migration, and the impact of the prolonged economic crisis to date, makes the labor problem very large and complex (Dewita Hia, 2017; Pertiwi, 2015). If such a complicated unemployment problem is allowed to drag on, it is very likely to encourage a social crisis.

The beginning of the unemployment explosion occurred around the end of 1997 or early 1998. When a severe monetary crisis hit Asia, especially Southeast Asia, it encouraged the creation of tight liquidity as a reaction to monetary turmoil in Indonesia. In 2006, the overall unemployment rate in Indonesia reached 10.27% with a total unemployment of

- *Author: Didiharyono, Senior Lecture in Mathemetics at Andi Djemma University, Indonesia. E-mail: muh.didih@gmail.com*
- *Co-Author: Muhammad Syukri, economics lecturer in Andi Djemma University. E-mail: syukrimuhammad302@yahoo.co.id*
- *Correnspondent* : *muh.didih@gmail.com*

10,932,000 people. Although Indonesia's unemployment rate has decreased in 2017 by 7.01%, in 2018 it was 6.87% and it is possible that the following year will again increase (BPS, 2018). As it is, the open unemployment rate in South Sulawesi February 2017 unemployment was at 4.77%, increasing in August 2018 to 5.34% (Indonesia, 2018)(Sulsel, 2018). This figure will always change depending on the political economy policy of the government.

The debate about the unemployment rate is the main topic of discussion of political economy in many countries including in developing countries. One factor that has led to an increase in unemployment in Indonesia is that too many workers are directed to the formal sector so that when they lose their jobs in the formal sector, they are frustrated and cannot try to create their own jobs in the informal sector (Dewita Hia, 2017). Projections of the unemployment rate in the future are very important for economic policy in detecting, planning and stopping any continuous increase in the unemployment rate in a country including in South Sulawesi.

The best forecasting methods used to forecast open unemployment rate South Sulawesi in the future is Autoregressive Integrated Moving Average (ARIMA) model or Box-Jenkins method (Dobre & Alexandru, 2008). This model is considered very accurate because ARIMA is one of forecasting methods developed based on time series data. This method uses an iterative approach in identifying the most appropriate model or the best model of all possible models. The advantages of this method can handle almost all time series data (Baziar & Kavousi-fard, 2015) (Didiharyono & Bakhtiar, 2018).

Common equations that represent non-stationary time series is ARIMA (*p*,*d*,*q*) model where *p*,*d*,*q* namely orde to autoregressive process, differencing and moving average. ARIMA (*p*,*d*,*q*) model can be written in equations is (Dobre & Alexandru, 2008).

$$\left(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p\right)(1-B)^d \tilde{z}_t=(1-\theta_1 B-\cdots-\theta_q B^q)a_t \qquad (1.1)$$

where, $\phi_1 \cdots \phi_p$ is parameter AR model; $(1\text{-B})^d$ is *differencing* level; Z_t is forecast value period *t;* $\theta_1 \cdots \theta_q$ is parameter MA model; and a_t is residual value period *t.*

Research using the ARIMA model is widely used by researchers, such as which examines forecasting unemployment rates in the US (Dritsakis & Klazoglou, 2018), unemployment rates in Greece (Nikolaos et al., 2016), unemployment rates in Nigeria (Nkwatoh, 2012), and who examined the unemployment rate across countries (Barnichon & Garda, 2016). Later, continued to use the ARIMA model, but the application was in different cases (Baziar & Kavousi-fard, 2015) (Didiharyono & Bakhtiar, 2018) (Ohyver & Pudjihastuti, 2018) (Abdullah, 2012). Based on the explanation, the research objective is to forecast using the ARIMA model in anticipation of the open unemployment rate in South Sulawesi.

2. Research Methods

Research method used is applied research with quantitative secondary data obtained at Central Statistics Agency (BPS) South Sulawesi from 1986 until 2018. Research procedures include identification model; analyze of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF); differences data; estimation parameter in model; and forecasting with soft computers, Minitab (Bisgaard & Kulahci, 2011) (Didiharyono & Bakhtiar, 2018) (Nkwatoh, 2012)(Abdullah, 2012).

3. RESULTS AND DISCUSSION

3.1. Identification Model

Open Unemployment Rates is an indicator that can be used to level of labor supply not absorbed in the labor market. Open Unemployment Rates data South Sulawesi Province from 1986 until 2004 presented in years. While, from 2005 until 2018 presented every six months in February and August. So that, total periods is 47 data. Based on the research data it's shown in form graph below this,

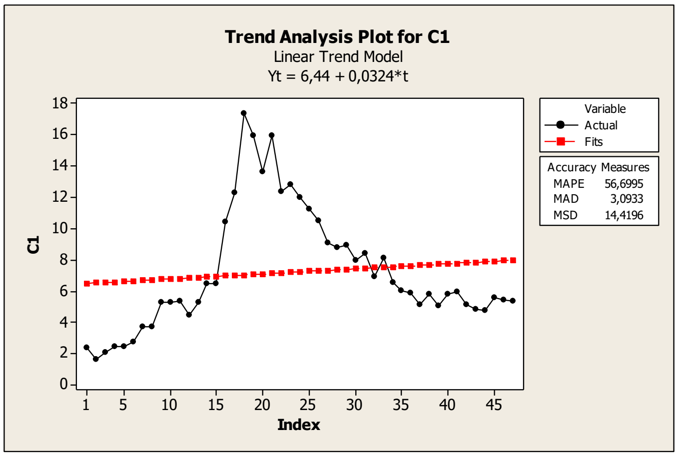


Fig. 1. Trend Chart Open Unemployment Rates

Base on plot data trend chart open unemployment rates at Fig. 1 can be known that open unemployment rate has increased and decreased over time, so the graph includes time series models that are not stationary in the average.

3.2. Autocorrelation Function and Partial Autocorrelation Function

The ARIMA forecasting model was identified by observing the Autocorrelation Function (ACF) model and Partial Autocorrelation Function (PACF) model.

Autocorrelation Function: C1

Lag	ACF	T	LBQ
1	0,912576	6,26	41,69
2	0,827472	3,47	76,74
3	0,709185	2,42	103,06
4	0,572993	1,75	120,64
5	0,458980	1,32	132,20
6	0,321045	0,89	137,99
7	0,232418	0,63	141,10
8	0,121913	0,33	141,97
9	0,018291	0,05	141,99
10	-0,073327	-0,20	142,33
11	-0,179295	-0,48	144,38
12	-0,267716	-0,72	149,10

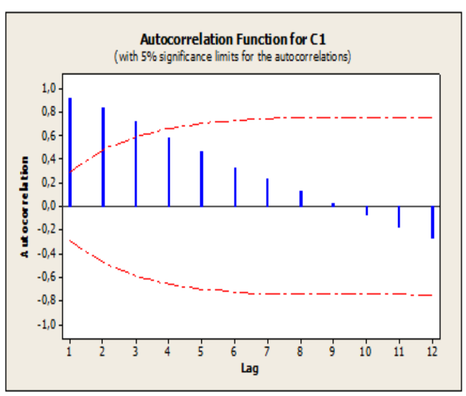


Fig. 2. ACF Value and ACF Chart

From the graph in Fig. 2, the autocorrelation value decreases slowly and is exponential so that the data is not stationary so that the best model is not formed. Next will be determined the following PACF value and PACF chart,

Partial Autocorrelation Function: C1

Lag	PACF	T
1	0,912576	6,26
2	-0,031830	-0,22
3	-0,245065	-1,68
4	-0,192527	-1,32
5	0,082112	0,56
6	-0,182684	-1,25
7	0,174329	1,20
8	-0,206031	-1,41
9	-0,104943	-0,72
10	-0,074567	-0,51
11	-0,093671	-0,64
12	-0,125260	-0,86

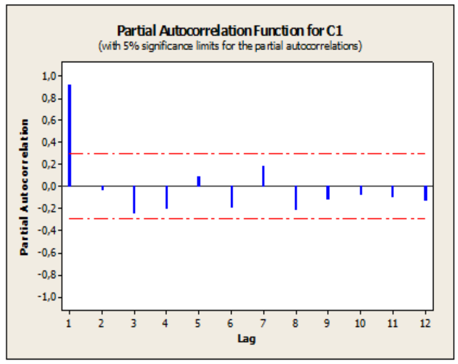
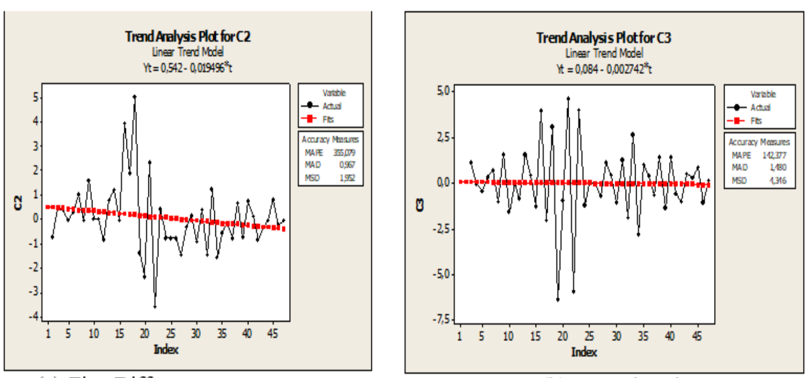


Fig. 3. PACF Value and PACF chart

The PACF graph in Fig. 3 is also not stationary, because the PACF autocorrelation values still appear to dies down and cut off at lag-1. In Fig. 2 and Fig. 3 shows 12 lags as in (Didiharyono & Bakhtiar, 2018) (Iriawan & Astuti, 2008), shows that the number of observations is 47 so 47/4 = 11.75 or 12 lag. In the identification of the data in Fig. 1 the best model has not yet been formed so that data differences are needed to stationer the data.

3.3. Differences Data

Based on Fig. 1 shows the data is not stationary, then differences are made so that the graph is as follows.



(a) First Differences (b) Second Kedua

Fig. 4. Data Analysis Graph after differences

Based on the analysis plot of the first data differences in Fig. 4 (a) shows a graph that has not been stationary. So the second data differences need to be done as shown in Fig. 4 (b). After the second data differences, it can be seen that the data is stationary. Because the average amount of production does not move freely in a given time and has quite a small variance and the actual value is approaching a linear line. The following will be determined the ACF and PACF values and graphs, namely:

Autocorrelation Function: C3

Lag	ACF	T	LBQ
1	-0,651892	-4,37	20,43
2	0,219374	1,08	22,79
3	0,076833	0,37	23,09
4	-0,314911	-1,51	28,21
5	0,410133	1,87	37,10
6	-0,453106	-1,93	48,24
7	0,288346	1,14	52,86
8	-0,030196	-0,12	52,91
9	-0,122296	-0,47	53,79
10	0,195426	0,74	56,10
11	-0,252252	-0,95	60,06

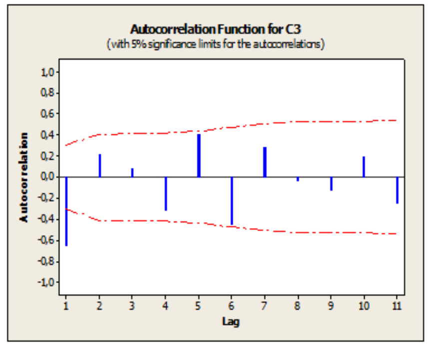


Fig. 5. ACF values and graphs after the second differences

Partial Autocorrelation Function: C3

Lag	PACF	T
1	-0,651892	-4,37
2	-0,357522	-2,40
3	0,075576	0,51
4	-0,272660	-1,83
5	0,069022	0,46
6	-0,278869	-1,87
7	-0,184483	-1,24
8	-0,029640	-0,20
9	0,020463	0,14
10	-0,063366	-0,43
11	-0,156429	-1,05

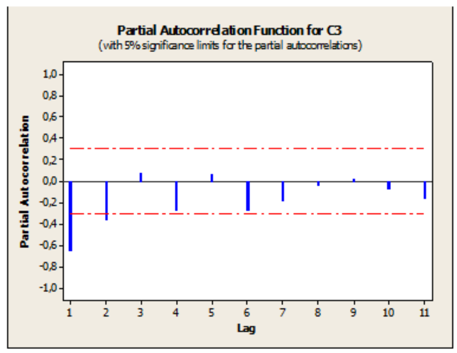


Fig. 6. PACF values and graph after the second differences

From the ACF graph in Fig. 5 the data is stationary because the graph is dies down so that the model can be directly estimated and ρ_k value cut off of the white noise line in lag-1 or lag-q for the highest moving average level. Whereas, from the PACF graph in Fig. 6 it is seen that it follows the sine graph and the graph dies down exponentially. By looking at the two graphs above, you can see ρ_k value cut off the white noise line at lag-1, while, ϕ_{kk} value cut white noise at lag-1. So the estimated model is ARIMA (1,2,1) or AR (1) because the ACF graph dies down exponentially and PACF cut off at lag-1, differences two times, MA (1) because the ACF graph is cut off at lag-1 and PACF dies down exponentially (sinusoidal), so the general form can be written:

$$Z_t = (2 + \phi_1)Z_{t-1} - (1 + 2\phi_1)Z_{t-2} + \phi_1 Z_{t-3} + a_t - \theta_1 a_{t-1} \tag{1.2}$$

And it is estimated that the approaching model is ARIMA (2,2,1) or AR (2) because the ACF graph dies down exponentially and PACF is cut off at lag-2, differences two times, MA (1) because the ACF graph is cut off at lag-1 and PACF dies down exponentially. As a comparison, the ARIMA (1,2,2), ARIMA (2,2,2) and ARIMA (3,2,1) forms are used to see the smallest Mean Square value.

3.4. Parameter Estimation in the Model

After identifying the data, the next step is to estimate the parameters. The output results obtained by using the Minitab Software are obtained as follows.

Table 1.
Sum Square and Mean Square values

No	Model	Sum Square (SS) Value	Mean Square (MS) Value
1	ARIMA (1,2,1)	85,9890	2,0474
2	ARIMA (2,2,1)	90,8799	2,2166
3	ARIMA (1,2,2)	89,8512	2,1915
4	ARIMA (2,2,2)	89,6013	2,2400
5	ARIMA (3,2,1)	90,3677	2,2592

This stage aims to examine the best model by finding the smallest Mean Square (MS) value based on the estimated parameter results. From the five ARIMA models it can be concluded that the ARIMA (1,2,1) which has the smallest MS value. Thus this model is appropriate to be used to predict the unemployment rate in the future.

Based on the analysis results obtained by ARIMA (1,2,1) which has parameters of ϕ_1 = -0,0733 and θ_1 = 1,0504. By using Equation 1.1, the model equation is obtained:
$Z_t = (2+\phi_1)Z_{t-1} - (1+2\phi_1)Z_{t-2} + \phi_1 Z_{t-3} + a_t - \theta_1 a_{t-1}$
$Z_t = (2-0{,}0733)Z_{t-1} - \big(1+2(-0{,}0733)\big)Z_{t-2} - 0{,}0733Z_{t-3} + a_t - 1{,}0504a_{t-1}$
$Z_t = 1{,}9267Z_{t-1} - 0.8534Z_{t-2} - 0{,}0733Z_{t-3} + a_t - 1{,}0504a_{t-1}$ (1.3)
Equation (1.3) is used to forecast future unemployment rates.

3.5. Forecasting

Model of ARIMA (1,2,1) is the most suitable model for modeling time series data and forecasting future periods. The forecasting results are shown in Table 2.

Table 2.
Forecasting Results in the Future Period

Forecasts from period 47

Period	Forecast	95% Limits Lower	95% Limits Upper	Actual
48	5,6799	2,8749	8,4849	
49	5,9797	2,2500	9,7093	
50	6,2707	1,8666	10,6748	
51	6,5507	1,6234	11,4779	
52	6,8198	1,4723	12,1672	
53	7,0780	1,3873	12,7686	
54	7,3252	1,3523	13,2981	
55	7,5616	1,3561	13,7671	
56	7,7871	1,3907	14,1835	
57	8,0017	1,4496	14,5538	
58	8,2054	1,5279	14,8828	
59	8,3981	1,6212	15,1750	

Table 2 shows that after 47 periods (t) in the coming 12 periods there was a continuous increase with 95% data at the lower end and the upper limit for the actual data. Forecasting the increase in the number of unemployed every time gives a warning to the government to formulate economic strategies and policies in reducing the number of unemployed each year.

4. Conclusion

Based on the results of the study obtained the best time series model used in doing forcasting the unemployment rate in the future is ARIMA (1,2,1). That model is the best model with smallest Mean Square value compared to other models. The general equation of the ARIMA (1,2,1) is $Z_t = 1{,}9267Z_{t-1} - 0.8534Z_{t-2} - 0{,}0733Z_{t-3} + a_t - 1{,}0504a_{t-1}$. Forcasting results obtained for the next 12 periods are 5,6799; 5,9797; 6,2707; 6,5507; 6,8198; 7,0780; 7,3252; 7,5616; 7,7871; 8,0017; 8,2054; and 8,3981.

5. Recommendation

Forcasting results show the unemployment rate has increased from year to year. So, it is hoped that the government, especially the provincial government of South Sulawesi, can formulate strategies and policies that are ideal in dealing with the unemployment problem. For example, government policies in increasing knowledge about entrepreneurship to Micro, Small and Medium Enterprises (UMKM) and capital assistance in increasing the entrepreneurial spirit of the community. Not only that, the government can also develop the tourism and cultural sectors, develop marine and agricultural potential in South Sulawesi so as to create an increase in productive employment opportunities in improving the welfare of the community.

Acknowledgment

The authors wish to thank research institutions and community service (LP2M) Andi Djemma University for your help and support.

References

Abdullah, L. (2012). ARIMA Model for Gold Bullion Coin Selling Prices Forecasting. *International Journal of Advances in Applied Sciences (IJAAS)*, *1*(4), 153–158.

Bakhtiar, B., & Didiharyono, D. (2018). Perkiraan Lonjakan Jumlah Kunjungan Wisatawan di Kabupaten Toraja Utara Tahun 2021 dengan Penerapan Metode ARIMA. *Jurnal Sainsmat*, *VII*(2), 153–162.

Barnichon, R., & Garda, P. (2016). Forecasting unemployment across countries: The ins and outs. *European Economic Review*, *84*, 165–183. https://doi.org/10.1016/j.euroecorev.2015.10.006

Baziar, A., & Kavousi-fard, A. (2015). Short Term Load Forecasting Using A Hybrid Model Based On Support Vector Regression. *International Journal of Scientific & Technology Research*, *4*(05), 189–195.

Bisgaard, S., & Kulahci, M. (2011). *Time Series Analysis and Forecasting by Example* (1st editio). English: Wiley.

BPS. (2018). Keadaan Ketenagakerjaan Indonesia Februari 2018. *Badan Pusat Statistik*, (42), 1–16.

https://doi.org/No. 74/11/35/Th.XVI, 5 November 2018

Dewita Hia, Y. (2017). Strategi Dan Kebijakan Pemerintah Dalam Menanggulangi Pengangguran. *Economica*, *1*(2), 208–213. https://doi.org/10.22202/economica.2013.v1.i2.121

Didiharyono, D., & Bakhtiar, B. (2018). Forecasting Model With Box-Jenkins Method To Predict Tourists Who Visit Tourism Place In Toraja. *Jurnal of Economic, Management and Accounting*, *1*(1), 34–41. Retrieved from http://ojs.unanda.ac.id/index.php/jemma/article/download/75/65

Dobre, I., & Alexandru, A. A. (2008). Modelling unemployment rate using Box-jenkins procedure. *Journal of Applied Quantitative Methods*, *3*(Quantitative Models), 156–166.

Dritsakis, N., & Klazoglou, P. (2018). Forecasting Unemployment Rates in USA Using Box-Jenkins Methodology. *International Journal of Economics and Financial Issues*, *8*(1), 9–20.

Indonesia, S. Tingkat Pengangguran Terbuka (TPT) Menurut Provinsi, 1986-2018 (2018). Retrieved from https://www.bps.go.id/statictable/2014/09/15/981/tingkat-pengangguran-terbuka-tpt-menurut-provinsi-1986---2018.html

Iriawan, N., & Astuti, S. P. (2008). *Buku Mengolah Data Statistik Dengan Mudah Menggunakan Minitab 14* (1st ed.). Yokyakarta: Penerbit Andi.

Nikolaos, D., Stergios, A., Tasos, S., & Ioannis, S. (2016). Forecasting Unemployment Rates in Greece. *International Journal of Sciences: Basic and Applied Research*, *4531*, 43–55.

Nkwatoh, L. (2012). Forecasting Unemployment Rates in Nigeria Using Univariate time series Models. *International Journal of Business and Commerce*, *1*(12), 33–46. Retrieved from http://www.ijbcnet.com/1-12/IJBC-12-11202.pdf

Ohyver, M., & Pudjihastuti, H. (2018). Arima Model for Forecasting the Price of Medium Quality Rice to Anticipate Price Fluctuations. *Procedia Computer Science*, *135*, 707–711. https://doi.org/10.1016/j.procs.2018.08.215

Pertiwi, S. Y. (2015). Model Kebijakan Pemerintah Daerah dalam Penanggulangan Kemiskinan. *Pandecta: Research Law Journal*, *9*(2), 216. https://doi.org/10.15294/pandecta.v9i2.3446

Sulsel, B. (2018). Keadaan Ketenagakerjaan Sulawesi Selatan Agustus 2018. *Badan Pusat Statistik Sulsel*.