Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 14(1) 2021: 44-55
http://ejournal.undip.ac.id/index.php/media_statistika

AUTOREGRESSIVE FRACTIONAL INTEGRATED MOVING AVERAGE (ARFIMA) MODEL TO PREDICT COVID-19 PANDEMIC CASES IN INDONESIA

Puspita Kartikasari, Hasbi Yasin, Di Asih I Maruddani
Department of Statistics, Faculty of Science and Mathematics, Diponegoro University

e-mail: puspitakartikasari@live.undip.ac.id

DOI: 10.14710/medstat.14.1.44-55

Article Info:
Received: 8 September 2020
Accepted: 25 June 2021
Available Online: 30 June 2021

Keywords:
ARFIMA; Prediction; COVID-19
.

Abstract: Currently the emergence of the novel coronavirus (Sars-Cov-2), which causes the COVID-19 pandemic and has become a serious health problem because of the high risk causes of death. Therefore, fast and appropriate action is needed to reduce the spread of the COVID-19 pandemic. One of the way is to build a prediction model so that it can be a reference in taking steps to overcome them. Because of the nature of transmission of this disease which is so fast and massive cause extreme data fluctuations and between objects whose observational distances are far enough correlated with each other (long memory). The result of this determination is the best ARFIMA model obtained to predict additional of recovering cases of COVID-19 is (1,0,489.0) with an SMAPE value of 12,44%, while the case of death is (1.0.429.0) with SMAPE value of 13,52%. This shows that the ARFIMA model can accommodate well the long memory effect, resulting in a small bias. Also in estimating model parameters, it is also simpler. For cases of recovery and death, the number is increasing even though the case of death is still very high compared to cases of recovery.

1. INTRODUCTION

The novel coronavirus (Sars-Cov-2) which causes the COVID-19 pandemic has become a global health problem. Even the World Health Organization (WHO) has set it to be a global pandemic because of the fast and massive nature of transmission, as well as the health impacts it causes, such as pneumonia, multi-organ failure, and even a high risk causes of death (WHO, 2020; Sifriyani & Rosadi, 2020). Due to the fast and massive nature of transmission, data on the spread of the COVID-19 pandemic, both the addition of cases of recovery and death fluctuate in extreme ways and between objects whose observation distances are quite far correlated (long memory effect). Therefore, a quick and precise action is needed to minimize the number of patients exposed (Shi et al., 2020; Chen et al., 2020)

Building a predictive model requires information from previous events, but often the information from these events are extremely fluctuating and random. Like data on the cases of the spread of the COVID-19 pandemic which is currently sweeping the world, including Indonesia, it also fluctuates in extreme, random ways and has a long memory effect. There are many methods in statistics that can be used to build a predictive model, ranging from parametric and nonparametric statistical methods such as regression methods to time series

based methods, Autoregressive Integrated Moving Average (ARIMA). However, there are limitations to the formation of a prediction model for the COVID-19 pandemic case through the regression approach or ARIMA, which is unable to build a model with high accuracy due to the nature of the data that fluctuates in extreme, random and has a long memory effect. In addition, this method must also fulfill the assumptions that have been determined in the analysis so that it is not well used in the interpretation of the parameters and the confidence interval of the model being built. Therefore, the prediction model with the Autoregressive Fractional Integrated Moving Average (ARFIMA) approach is used to accommodate this so that a robust and accurate prediction model is obtained.

2. LITERATURE REVIEW

2.1. Long Memory Model Identification

Long memory effects can be identified using several estimation methods, including Geweke and Porter Hudak (GPH), Nonlinear Least Square (NLS), Exact Maximum Likelihood (EML) and Modified Profile Likelihood (MPL) (Doornik & Ooms, 1999). However, in this study the GPH Estimator is used because the parameter d can be estimated without identifying the order p and q first (Geweke & Porter-Hudak, 1983).

$$\hat{d}_{GPH} = \frac{-0{,}5\sum_{j=1}^{m}\left(\tilde{X}_j - \bar{\tilde{X}}\right)\log I_j}{\sum_{j=1}^{m}\left(\tilde{X}_j - \bar{\tilde{X}}\right)^2} \quad (1)$$

2.2. ARFIMA Model

Time series method is one of the most popular prediction methods in statistics. This is because the method is simple but able to solve complex problems, if the case under study is influenced by the time (Zheng & Zhong, 2011). Besides that, the prediction model built is also accurate. In the case prediction of the transmission of COVID-19, which is require a model that can accommodate extreme fluctuations due to the nature of transmission and effects (long memory). From several previous studies, the ARFIMA model build predictive values with better accuracy than other models because it has flexibility, robustness and able to accommodate long memory effects, including (Kartikasari, 2020) show that the ARFIMA model show better result than the LSTAR and FILSTAR models. Aye et al., (2014) examines the effect of long memory daily stock returns from Brazil, Russia, India, China, and South Africa, also explain the efficacy of the ARFIMA model. (Baillie & Morana, 2012) examines simple adaptive modification of the basic ARFIMA model using the flexible Fourier form which is enable for a variety of interceptions

In the long memory case the stationary process causes the autocorrelation function to slow down slowly. This causes overdifferencing to have an unexpected impact on parameter estimation and prediction (Bhardwaj & Swanson, 2006). In the prediction there is a condition that the process is stationary and convergent at the average value. Thus, the prediction of long memory processes should converge to the average value of the process, even though it will converge slowly.

The ARFIMA model is a time series model that generalize and simplify parameter estimates in the ARIMA model with the value of the differencing 0 <d <1 (Chuang & Wei, 1991). The ARFIMA model for the Y_t time series is given by:

$$\emptyset_p(B)(1-B)^d Y_t = \theta_q(B)a_t \quad (2)$$

with $\emptyset_p(B) = \sum_{i=0}^{p} -\emptyset_i B^i$, $\theta(B) = \sum_{j=0}^{q} -\emptyset_j B^j$ and $\emptyset_0 = \theta_0 = -1$. In addition, B^k is a backshift operator or lag of Y_t then $B^k Y_t = Y_{t-k}$

From the Equation (2), we can see that in the case of long memory being modeled using the ARIMA model, many parameters will be estimated so that it does not adhere to the parsimony principle. However, the ARFIMA model provides simpler parameter estimates. More in-depth research on parameter estimation of the ARFIMA model was carried out by Dueker & Startz (1998), Reisen & Abraham (2001), Lopes, et al. (2004), Mayoral (2003) and for the fractional parameter d have been proposed in the literature of Hassler (1993), Lobato & Robinson (1996), Hurvich & Deo, (1999), TAQQU et al., (1995), Velasco, (1999). In general, the estimator d can be categorized into parametric methods as proposed by Fox & Taqqu (1991), Sowell (1992) and the semi-parametric methods Geweke & Porter-Hudak (1983), Robinson (1995) With parametric methods, the focus is on using Maximum Likelihood (MLE) Estimates. The MLE technique is used because the estimation method has several advantages over other estimation techniques, all the information contained in the data is used. In addition, this method is able to generate information from a large sample under common conditions. So that the results of the resulting prediction model are robust.

2.3. Best Model Selection

The selection of the best ARFIMA model is seen based on the symmetric mean absolute percentage error (SMAPE), which is presented in the formula:

$SMAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|Y_t - \hat{Y}_t|}{\left[\frac{1}{2}(Y_t - \hat{Y}_t)\right]}$ for out sample data, where:

Y_t : actual data

$\hat{Y}_t$: forecast data

n : observation size

2.4. Diagnostic Assumption Test

There are two diagnostic assumption tests in the ARFIMA model, the white noise test using the L-Jung Box test and normal distribution using the Kolmogorov Smirnov test, presented in the following formula.

Ljung-Box Test : $Q = n(n+2)\sum_{k=1}^{K} \frac{\hat{\rho}_k^2}{n-k}$

Kolmogorov Smirnov Test : $D = Sup|S(x) - F_0(x)|$

3. METHODOLOGY OF RESEARCH

3.1. Data Source and Variable

The Data and variable used are daily data for the additional of recovered and death cases cause of the COVID-19 in Indonesia. The data comes from the "Daily Technical Report" issued by the Ministry of Health of the Republic of Indonesia (https://covid19.go.id) (https://covid19.kemkes.go.id/). The data was taken from March 2, 2020 to June 14, 2020. The data is divided into two parts, in-sample data and out-sample data. March 2, 2020 to June 4, 2020 is used as the in-sample data, while the out-sample data is from June 5, 2020 to June 14, 2020. The data that has been collected will be analyzed using R software (Venables et al., 2021; Torres-Reyna, 2013).

3.2. Analysis Step

The analysis steps of the ARFIMA model are as follows:

1. Descriptive Statistics
   Descriptive statistics stage is examining how the feature / characteristics of the COVID-19 data, including the additional of recovered cases and those who died from March 2, 2020 to June 4, 2020.
2. Long Memory Testing
   This stage is carried out to defect the presence or absence of a long memory effect on the data using of the GPH Estimator method. If the value of d obtained from the GPH Estimator method is between 0 to 1, then there is a Long Memory effecr of the data.
3. ARFIMA Model Formation
   a. Stationary data in variants.
   b. Select one or more ARFIMA models according to the ACF and PACF plots.
   c. Estimating the parameters of the ARFIMA model obtained.
   d. Choose the best ARFIMA model based on SMAPE.
   e. Test the assumption requirement of the residual and handling the data if the residual assumptions are violated.
4. Predict the next 12 periods, then calculate the SMAPE value from the forecast data.

4. RESULTS AND DISCUSSION

The results from the data processing in cases of recovery and death due to COVID-19 are presented as follows.

1. Descriptive Statistics

Table 1. The Additional of Recovered Cases of COVID-19

Variable	Week	Mean	Minimum	Maximum	Range
The Additional of Recovered Cases	1	0	0	0	0
	2	1.143	0	6	6
	3	1.714	0	4	4
	4	6.100	0	13	13
	5	17.57	9	28	19
	6	31.71	4	73	69
	7	59.40	20	102	82
	8	75.60	40	137	97
	9	122.70	64	243	179
	10	162.40	91	231	140
	11	188.60	108	285	177
	12	247.00	72	523	451
	13	396.00	298	486	188

Based on table 1 above, from week 1 to week 3, the additional of recovered cases of COVID-19 have not shown a significant increase. However, in week 4 to week 13 there was a significant increase. This shows that there were an extreme fluctuation in the data for the additional of recovered cases. We can see in week 12 where there was the highest weekly increase in cases which there were 523 cases recovered and in week 13 there were a minimum of 298 cases recovered in a day. This does not mean that there will be progress in dealing with this pandemic problem because the minimum value and resistance have actually

decreased. The steps to anticipate the implementation of health protocols have not been effective, this is reflected in the inability to increase the minimum value and range for additional cases of recovery.

Table 2. The Additional of Death Cases of COVID-19

Variable	Week	Mean	Minimum	Maximum	Range
The Additional of Death Cases	1	0	0	0	0
	2	0.714	0	3	3
	3	4.710	0	14	14
	4	9.800	1	20	19
	5	12.14	7	21	14
	6	34.00	19	60	41
	7	22.43	8	47	39
	8	19.56	8	42	34
	9	19.71	8	35	27
	10	16.14	13	21	8
	11	33.57	13	59	46
	12	33.50	19	55	36
	13	27.00	22	35	13

Table 2 shows that from week 1 to week 6 the incidence of death cases show an extreme increase. Similar to the additional of recovered cases, the data on the additional of death cases also fluctuate in an extreme manner from day to day. We can see that at the beginning of this case it was identified on March 2, 2020, that the highest death cases occurred in week 6, or the second month with an average of 34 cases and a minimum of 19 people per day. Even on April 14, 2020 this week there were also 60 cases of death in a day, this is the highest case. At the end of the 3rd month or 9th week there was a decline, although not too significant, the number of positive confirmed cases increased. This phenomenon is due to the fact that in the 2nd month the handling of this pademic has not been fully implemented, there is no enforcement of health protocol rules such as PSBB, physical distancing, social distancing and so on and this is also the early days of the pandemic entering Indonesia. Meanwhile, at the end of week 13, social restrictions began to be loosened and coincided with the entry of Ramadan and Eid al-Fitr, where the culture of the Indonesian people held homecoming so that community mobilization was getting higher, therefore there was an increase in cases.

2. Long Memory Detection

In this step, long memory detection with GPH estimator is first carried out for the additional data of Covid-19 cases, both recovered and death cases. On the basis that if the value of d is obtained between 0 and 1, it indicates that there is a long memory effect, whereas if the value of d is less than 0 or more than 1, there is no long memory effect. The following are the results of the d value obtained from the GPH estimator method.

Table 3. d Value with the GPH Estimator Method

Variable	d Value
The Additional of Recovered Cases	0.7189
The Additional of Death Cases	0.5231

Based on table 3, the estimation results of the d parameter in the data for the additional of recovered cases of COVID-19 is 0.7189, while the results of the d parameter estimation in the data for the additional of death cases of COVID-19 is 0.5231. This prove

that the data has a long memory effect. From long memory detection, data of the additional cases of recovery and death of Covid-19 must be modeled using the ARFIMA model.

3. ARFIMA Model Formation
   a. Stationary data in Variance

Before forming the ARFIMA model, the first thing to do is check whether the data is stationary in variants or not by estimating the rounded value or λ using the Box-Cox method. If the value of λ = 1 means that the data is stationary in variants. It should be noted that in the formation of the ARFIMA model the data only needs to be stationary in variants, while the mean does not need to fulfill the stationarity assumption.

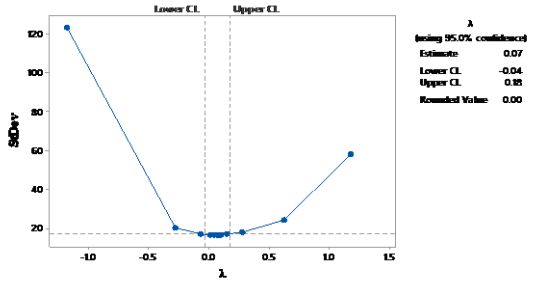

Figure 1. Box Cox of The Additional of Recovered Cases Data

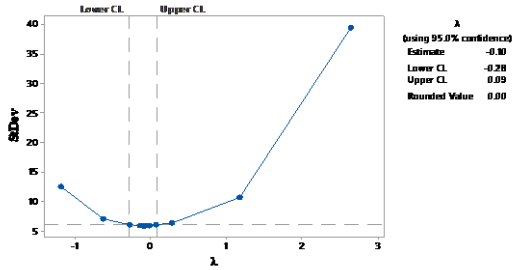

Figure 2. Box Cox of The Additional of Death Cases Data

Based on Figure 1 and Figure 2 shows the data pattern is not stationary in variants. This can be seen x from the value of λ = 0, therefore it is necessary to transform using the Box-Cox method. Box-Cox with a single parameterized transformation class, λ which is raised to the data, where λ is the expected parameter. The basis for selecting the transformation model is that if the observational data Y1, Y2, Y3, ... Yn are not stationary in variants with a value of λ = 0 then the observational data will be transformed using natural logarithms (Ln) and if the value λ = 0 then use root transformation. The results of the Ln transformation are as follows.

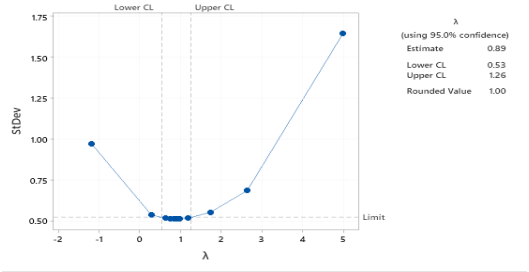

Figure 3. Transformation of Ln using Box-Cox Method for The Additional of Recovered Cases

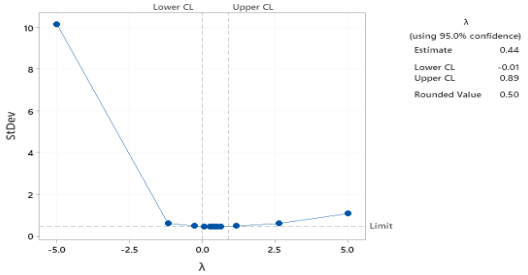

Figure 4. Transformation of Ln using Box-Cox Method for The Additional of Death Cases

In Figure 3 and Figure 4 shows the value of λ = 1 for the additional of recovered cases while for the additional of death cases the value of λ = 0.50. From these results it can be concluded that in the data, the additional of recovered cases is stationary in variants, but for the additional of death cases, it is not stationary. Because the data has been transformed using Ln needs to be transformed again using root transformation to stationary the data for the additional of death cases. After the transformation, the root data on death cases shows the following results.

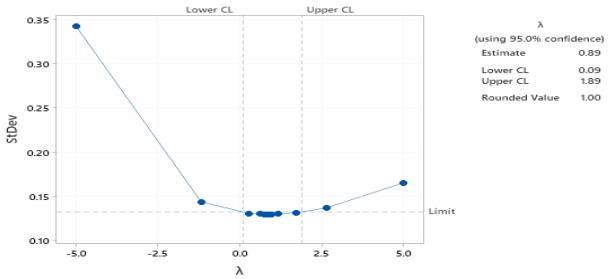

Figure 5. Root Transformation using Box-Cox Methods for The Additional of Death Cases

In Figure 5, it can be seen that the value of λ = 1, then both the additional of recovered and death cases, have fulfilled the stationarity assumption in variants. After the descriptive statistics stage, long memory identification, and stationarity assumptions in the next analysis variant, the determination of the appropriate ARFIMA model is based on the ACF and PACF plots generated from the transformed data.

b. ARFIMA Model Selection

To select the ARFIMA model (p, d, q) it is necessary to select each value. For p model related parameters, Autoregressive (AR), q parameter Moving Avarage (MA) while for parameter d which the value between 0 to 1 is called Fractional Integrated (FI). The choice of this model is based on the results of the ACF and PACF plots where the AR model is based on the cut off lag in the PACF plot, while the MA model is based on the cut off lag in the ACF plot.

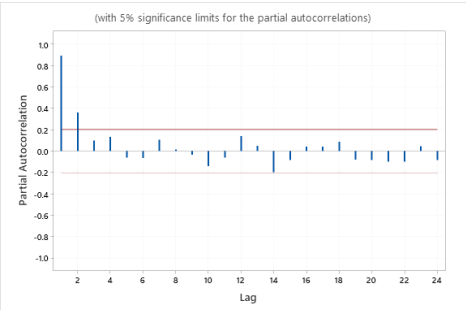

Figure 6. PACF Plot of Ln Transformation Result for The Additional of Recovered

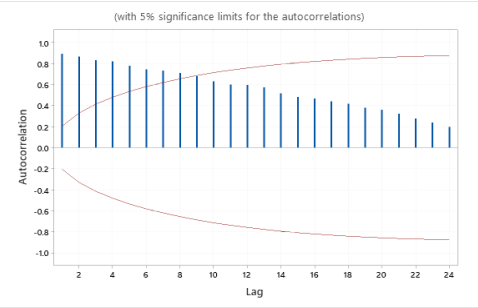

Figure 7. ACF Plot of Ln Transformation Result for The Additional of Recovered

From Figures 6 and 7, the PACF plot indicates the model that will be used to predict the additional of recovered cases are AR 1 or AR 2. This is based on the PACF plot which is cut off (significant) at lag 2 and the ACF plot decreases slowly resulting in MA model 0 or not indicated. Likewise in Figures 8 and 9 the model will be used to predict the additional of death cases, AR 1 or AR 2 because the PACF plot is cut off

(significant) at the second lag and the ACF plot drops slowly resulting in the MA model 0 or not being indicated. After knowing the model that will be used as a predictive model, the next step is to estimate the p and d values in the selected model, AR (1) and AR (2), both for the additional recovered and death cases.

c. Estimating parameters of the ARFIMA model

After carrying out several experiments by including significant lags, the estimated model obtained and which is suitable for data on recovery and death cases are presented in Table 4 and Table 5.

Table 4. Estimation of ARFIMA Model Parameters for The Additional of Recovered Cases

	Estimate	Std. Error	z value	Pr(>|z|)
d	0,489	0,008	63,48	<2e-16
AR (1)	-0,265	0,102	-2,60	0,00934

Table 5. Estimation of ARFIMA Model Parameters for The Additional of Death Cases

	Estimate	Std. Error	z value	Pr(>|z|)
d	0,429	0,009	46,805	<2e-16
AR (1)	-0,235	0,099	-2,353	0,0186

Table 4 shows that the parameter estimate for AR (1) is -0.265 with an estimated value of d is 0.489. The overall coefficient of these parameters is significant to the model because the p-value is <α, so the ARFIMA model that is formed is ARFIMA (1,0.489.0), so the model formed is as follows.

$A_1 \nabla^{0,489} Z_t = A_2 a_t$, with

$A_1 = 1 + 0{,}265B$

$A_2 = 1$,

$(1 + 0{,}265B)(1 - B)^{0,489} Y_t = a_t$

While table 5 shows that the parameter estimate for AR (1) is -0.235 with an estimated value of d amounting to 0.429. The overall coefficient of these parameters is significant to the model because the p-value is <α, so the ARFIMA model that is formed is ARFIMA (1,0.429.0, so the model formed is as follows.

$\boldsymbol{A_1 \nabla^{0,429} Z_t = A_2 a_t}$,with

$A_1 = 1 + 0{,}235B$

$\boldsymbol{A_2 = 1}$

$(1 + 0{,}235B)(1 - B)^{0,429} Y_t = a_t$

d. Residual Assumption Testing

After obtaining the best ARFIMA model parameters, the next thing to do is a diagnostic check by testing the residuals to see whether the residuals meet the white noise assumption and are normally distributed or not. The results are presented in the following table.

Table 6. Residual White Noise Test of ARFIMA Model (1,0.489,0)

X-Squared	df	P-Value
49,78	48	0,4023

Table 7. Residual White Noise Test of ARFIMA Model (1,0.429,0)

X-Squared	df	P-Value
54,84	48	0,2312

Table 8. Residual Normality Test of ARFIMA Model (1,0.489,0)

D	P-Value
0,125	0,7235

Table 9. Residual Normality Test of ARFIMA Model (1,0.429,0)

D	P-Value
0,115	0,7122

In Table 6 and Table 7, by using the L-Jung Box test statistic, the P-value is obtained of 0.4023 in the ARFIMA model (1.0.489.0) and 0.1927 in the ARFIMA model (1,0.429.0) where the value is greater than $\alpha = 0.05$. Therefore, we can conclude that in the residual assumption test, the ARFIMA model (1,0.489.0) for the addition of cured cases and the ARFIMA model (1,0.429.0) for the death of cases has fulfilled the residual white noise assumption. Meanwhile, for the residual normality test, the test statistic used is the Kolmogorov Smirnov test, where the results are presented in table 8 and table 9.Based on the results of the residual normality assumption test using the Kolmogorov Smirnov test statistic, the residual results are normally distributed, because the P-value obtained in ARFIMA (1,0.489.0) and (1,0.429.0) models of 0, 7235 and 0.7122 where the value is greater than $\alpha = 0.05$. After all the residual assumptions are fulfilled, the final step is to make predictions for additional cases of recovered and death cause of COVID-19 using the ARFIMA model that has been obtained.

4. Prediction

After all the assumptions are fulfilled, both stationary and residual assumptions, the next step is to use the ARFIMA model selected for prediction. In this study we will predict the next 12 periods. The prediction results are presented in the following table.

Table 10. Forecast Result ARFIMA Model (1,0.489,0)

Data	The Patient Recovered	Forecast
June 5, 2020	551	486
June 6, 2020	464	411
June 7, 2020	591	504
June 8, 2020	406	333
June 9, 2020	510	442
June 10, 2020	715	651
June 11, 2020	507	452
June 12, 2020	577	524
June 13, 2020	563	503
June 14, 2020	755	696
	SMAPE	**12,44%**

Tabel 11. Forecast Result of ARFIMA Model (1,0.429,0)

Date	The Patient Death	Forecast
June 5, 2020	49	43
June 6, 2020	31	26
June 7, 2020	50	45
June 8, 2020	32	28
June 9, 2020	40	35
June 10, 2020	36	32
June 11, 2020	41	37
June 12, 2020	48	39
June 13, 2020	43	37
June 14, 2020	43	39
SMAPE		**13.52%**

Based on the prediction results for the next 12 periods in table 10 and table 11, it can be seen that the forecast results have good values of the SMAPE model of 12.44%, and 13.52%. This means that the ARFIMA (1,0.489.0) and (1,0.429.0) models can be said to be accurate to be used to predict the addition of recovered and death cause of COVID-19 cases. The predicted data presented in tables 10 and 11 are data that have been returned to their original form. This is done because at the time of modeling the prediction the data must be transformed so that the assumption of stationarity in the variants is fulfilled.

5. CONCLUSION AND SUGGESTION

The best ARFIMA model obtained for predicting COVID-19 cases in Indonesia, both the additional of recovered and death cases every day is $(1 + 0{,}265B)(1 - B)^{0{,}489}Y_t = a_t$ with the SMAPE value 12,44% for the number of recovered cases, while the number of death cases $(1 + 0{,}235B)(1 - B)^{0{,}429}Y_t = a_t$ with the SMAPE value 13,52%. This is because the ARFIMA model is able to accommodate the long memory effect well, resulting in a small bias, besides that the estimation of model parameter is also simpler. The ARFIMA model development in this study goes through several stages, one of which is the transformation of the data so that the assumption of stationarity in variant is fulfilled. Therefore, after obtaining the prediction results from the model built, it is necessary to return to the original data. The prediction results obtained for additional cases of recovery and death, The number of cases are increasing even though the death cases are lower than the recovered cases. The ARFIMA model that has been built provides an illustration of the addition of cases of recovery and death due to COVID-19 so that anticipatory steps can be taken and decisions that need to be made.

REFERENCES

Aye, G. C., Balcilar, M., Gupta, R., Kilimani, N., Nakumuryango, A., & Redford, S. (2014). Predicting BRICS stock returns using ARFIMA models. *Applied Financial Economics*, *24*(17), 1159–1166. https://doi.org/10.1080/09603107.2014.924297

Baillie, R. T., & Morana, C. (2012). Adaptive ARFIMA models with applications to inflation. *Economic Modelling*, *29*(6), 2451–2459. https://doi.org/10.1016/j.econmod.2012.07.011

Bhardwaj, G., & Swanson, N. R. (2006). An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series. *Journal of*

Econometrics, *131*(1–2), 539–578. https://doi.org/10.1016/j.jeconom.2005.01.016

Chen, Q., Liang, M., Li, Y., Guo, J., Fei, D., Wang, L., He, L., Sheng, C., Cai, Y., Li, X., Wang, J., & Zhang, Z. (2020). Mental health care for medical staff in China during the COVID-19 outbreak. *The Lancet Psychiatry*, *7*(4), e15–e16. https://doi.org/10.1016/S2215-0366(20)30078-X

Chuang, A., & Wei, W. W. S. (1991). Time Series Analysis: Univariate and Multivariate Methods. In *Technometrics* (Vol. 33, Issue 1, p. 108). https://doi.org/10.2307/1269015

Doornik, J. a, & Ooms, M. (1999). A Package for Estimating , Forecasting and Simulating Arfima Models : Arfima package 1 . 0 for Ox. *Erasmus*, 1–31.

Dueker, M., & Startz, R. (1998). Maximum-likelihood estimation of fractional cointegration with an application to U.S. and Canadian bond rates. *Review of Economics and Statistics*, *80*(3), 420–426. https://doi.org/10.1162/003465398557654

Fox, R., & Taqqu, M. S. (1991). Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to The Annals of Statistics. ® www.jstor.org. *Annals of Statistics*, *19*(3), 1403–1433.

Geweke, J., & Porter-Hudak, S. (1983). the Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, *4*(4), 221–238. https://doi.org/10.1111/j.1467-9892.1983.tb00371.x

Hassler, U. (1993). Regression of Spectral Estimators With Fractionally Integrated Time Series. *Journal of Time Series Analysis*, *14*(4), 369–380. https://doi.org/10.1111/j.1467-9892.1993.tb00151.x

Hurvich, C. M., & Deo, R. S. (1999). Regression Estimates of the Memory Parameter of a Long-. *Journal of Time Series Analysis*, *1*.

Kartikasari, P. (2020). *Prediksi Harga Saham Pt . Bank Negara Indonesia Dengan Menggunakan Model Autoregressive Fractional Integrated Moving Average (Arfima)*. *8*(1), 1–7.

Lobato, I., & Robinson, P. M. (1996). Averaged periodogram estimation of long memory. *Journal of Econometrics*, *73*(1), 303–324. https://doi.org/10.1016/0304-4076(95)01742-9

Mayoral, L. (2003). *Centre de Referència en Economia Analítica Barcelona Economics Working Paper Series Working Paper n^{o} 55 Inequality*. *October 2001*.

Reisen, V., & Abraham, B. (2001). Estimation of parameters in arfima processes: A simulation study. *Communications in Statistics Part B: Simulation and Computation*, *30*(4), 787–803. https://doi.org/10.1081/SAC-100107781

Robinson, P. M. (1995). Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to The Annals of Statistics. ® www.jstor.org. *Annals of Statistics*, *19*(3), 1403–1433.

S. Lopes., B. Olbermann., V. R. (2004). *A Comparison of Estimation Methods in Non-Stationary ARFIMA Processes Lopes , S . R . C* . 1–14.

Shi, Y., Wang, J., Yang, Y., Wang, Z., Wang, G., Hashimoto, K., Zhang, K., & Liu, H. (2020). Knowledge and Attitudes of Medical Staff in Chinese Psychiatric Hospitals Regarding COVID-19. *Brain, Behavior, & Immunity - Health*, *4*(March), 100064.

https://doi.org/10.1016/j.bbih.2020.100064

Sifriyani, S., & Rosadi, D. (2020). Susceptible Infected Recovered (SIR) Model for Estimating Covid-19 Reproduction Number in East Kalimantan and Samarinda. *Media Statistika*, *13*(2), 170–181. https://doi.org/10.14710/medstat.13.2.170-181

Sowell, F. (1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. *Journal of Econometrics*, *53*(1–3), 165–188. https://doi.org/10.1016/0304-4076(92)90084-5

TAQQU, M. S., TEVEROVSKY, V., & WILLINGER, W. (1995). Estimators for Long-Range Dependence: an Empirical Study. *Fractals*, *03*(04), 785–798. https://doi.org/10.1142/s0218348x95000692

Torres-Reyna, O. (2013). PPT:Introduction to RStudio. *Princeton University*, *August*, 1–16.

Velasco, C. (1999). Gaussian Semiparametric Estimation of Non-Stationary Time Series. *Journal of Time Series Analysis*, *20*(1), 87–127. https://doi.org/10.1111/1467-9892.00127

Venables, W. N., Smith, D. M., & R Core Team. (2021). An Introduction to R. *Quantitative Geography: The Basics*, *0*, 250–286. https://doi.org/10.4135/9781473920446.n12

Who. (2020). A Coordinated Global Research Roadmap: 2019 Novel Coronavirus. In *Who* (Issue March). http://dx.doi.org/10.1038/s41591-020-0935-z

Zheng, F., & Zhong, S. (2011). Time Series Forecasting Using A Hybrid RBF Neural Network and AR Model Based on Binomial Smoothing. *World Academy of Science, Engineering and Technology*, *51*(3), 1464–1468. https://doi.org/10.5281/zenodo.1072611