

Spatio Temporal Modelling for Government Policy the COVID-19 Pandemic in East Java

Atiek Iriany¹, Novi Nur Aini¹, Agus Dwi Sulistyono²

¹ Department of Statistics Faculty of Mathematics and Natural Sciences, Brawijaya University, Indonesia
² Faculty of Fisheries and Marine Science, Brawijaya University, Indonesia

Email: atiekiriany@ub.ac.id

ABSTRACT

COVID-19 has cursorily spread globally. Just in four months, its status altered into a pandemic. In Indonesia, the virus epicenter is identified in Java. The first positive case was identified in West Java and later spread in all Java. The Large-scale Social Restrictions are seemingly inefficient as the SARS-CoV-2 transmission remains. As such, the government is struggling to find anticipatory policies and steps best to mitigate the transmission. In this particular article, we used a Spatio-temporal model method for the total COVID-19 cases in Java and forecasted the total cases for the next 14 days, allowing the stakeholders to make more effective policies. The data we were using was the daily data of the cumulative number of COVID-19 cases taken from www.covid19.go.id. Data modeling was conducted using a generalized Spatio-temporal autoregressive model. The model acquired to model the COVID-19 cases in Java was the GSTAR(1)(1,0,0) model.

Keywords: COVID-19; forecasting, pandemic; spatio-temporal

INTRODUCTION

As stipulated by WHO on 12 March 2020, COVID-19 had become a pandemic [1]. The virus, firstly identified in Wuhan in December 2019, rapidly spread throughout China and other 190 countries [2]. No research exactly explains how the SARS-CoV-2 was initially transmitted, but, in the meantime, it is believed that humans transmit this virus to humans. Later research reveals that symptomatic patients transmit SARS-CoV-2 through droplets or sneezes [3]. Moreover, another research mentions that SARS-CoV-2 can live in gas particles, e.g., air (generated through nebulizer) for approximately three hours [4]. Due to its relatively rapid transmission and mortality rate which cannot be overlooked and no definitive therapy found, COVID-19 is one of the diseases to which we should alert [5].

The Coronavirus epicenter in Indonesia is identified in Java. The first positive case was identified in West Java and later spread in all Java. It indicates that adjacent locations closely pertain to the SARS-CoV-2 transmission. In response to the virus, China's social distancing regulation is proven effective to stabilize the virus transmission, and hence the number declines [6]. Indonesia, similar to China, issues the same regulation, namely the Large-scale Social Restrictions (PSBB). Nevertheless, the regulation is seemingly inefficient as the SARS-CoV-2 transmission remains. As such, the government is struggling to find anticipatory policies and steps best to mitigate the transmission.

Many researchers, e.g., Jia *et al.* [7], Albana [8], and Fajar [9] have studied the COVID-19 transmission and aim to recommend some anticipatory efforts. Meanwhile, we made COVID-19 modeling using a Spatio-temporal approach due to the SARS-CoV-2 transmission, which is mostly influenced by interplay, and numerous positive cases. Several researchers used the Spatio-temporal model [10] and [11]. One of the methods used to handle data attributed to time and location was Generalized Space-time Autoregressive (GSTAR). Some researchers, such as Iriany [12], Ruchjana [13], and Prastyo [14], prefer this method. In this particular article, we used a Spatio-temporal model for the total COVID-19 cases in Java and forecasted the total cases for the next 14 days, allowing the stakeholders to make more effective policies.

METHODS

Data Source

The data we were using in this research were the daily data of the cumulative number of COVID-19 cases taken from www.covid19.go.id.

Data Stationarity

According to the stationary time series data, neither a sharp decrease nor an increase in data value nor fluctuated data was found around the constant mean value [15]. Stationary data had the mean $E(Z_t) = \mu$ and variance $Var(Z_t) = \sigma^2$. The mean value conditioned that data had to be stationary, so neither decrease nor an increase in data from time to time was allowed [16]. Furthermore, the characteristic of a stationary time series was endlessly constant average and variance. There were two types of time series stationarity, namely stationarity to variance and the mean.

a. Stationarity to Variance

Stationarity to variance was if $Var(Z_t) = Var(Z_{t-k})$ for all t and k, the variance was constant from time to time [17]. To observe whether or not the data was stationary to variance, we used a Box-Cox plot. Non-stationary data could be altered into stationary ones through transformation.

b. Stationarity to the Mean

Stationarity to the mean was if $E(Z_t) = E(Z_{t-k})$ for all t and k, the mean function remained constant from time to time. Stationarity to the mean was observed using the ACF (Autocorrelation Function) plot or the Dickey-Fuller test. Non-stationary data could be altered into stationary ones through differencing.

Generalized Space-Time Autoregressive Integrated (GSTAR)

The AR order was determined using the MPACF plot. Correlation between Z_t and Z_{t+k} , after a dependence relationship, was linear. The variables Z_{t+1} , Z_{t+2} , ..., and Z_{t+k-1} were thus negated. The formula of correlation partial matrix function is as follows:

$$\phi_{kk} = \frac{cov \left[(Z_t - \hat{Z}_t), (Z_{t+k} - \hat{Z}_{t+k}) \right]}{\sqrt{var(Z_t - \hat{Z}_t)} \sqrt{var(Z_{t+k} - \hat{Z}_{t+k})}}$$
(1)

Where

 ϕ_{kk} = Partial correlation matrix coefficient at lag k

- Z_t = Observation data at the time t
- \hat{Z}_t = Predictor for Z_t
- Z_{t+k} = Observation data at the time t + k
- \hat{Z}_{t+k} = Predictor for Z_{t+k}

The partial autoregression matrix at lag s became the last matrix coefficient when the data were leveraged for the vector autoregression process of the order s.

The best model was selected among some models considered feasible for MPACF testing. Model selection was conducted using AIC. The less the AIC value in a model, the better the model. The quantification of the AIC value was as follows:

$$AIC(i) = \ln(|S(p)| + \frac{2pb^2}{T})$$
 (2)

Where:

b = the number of predicted parameters in the model

T = the number of observations

S(p) = residual sum of squares

p = VAR model order

The GSTAR model was introduced by Borovkova, Lopuha, and Ruchjana in 2020 in Wutsqa *et al.* [18]. It was more flexible and generalized than the STAR model and did not require the same parameter values at all locations. The GSTAR model $(p, \lambda_1, ..., \lambda_l)$ is written as follows [19]:

$$Z_{t} = \sum_{k=1}^{p} [\Phi_{k0} + \Phi_{k1}W] Z_{t-p} + e_{t}$$
(3)

Where:

 $\Phi_{k0} = \text{diag}(\phi_{k0}^1, \dots, \phi_{k0}^n)$, diagonal matrix of the parameter space-time lag spatial 0 and the parameter autoregressive lag at the time kth

 $\Phi_{k1} = \text{diag}(\phi_{k1}^1, \dots, \phi_{k1}^n)$, diagonal matrix of the parameter space-time lag spatial 1 and the parameter autoregressive lag at the time kth

W = weighing matrix (N×N) selected as such that $W_{ii}^{(k)} = 0$ dan $\sum_{i \neq j} W_{ij}^{(k)} = 1$

e(t) = the white-nose vector in size of (N × 1)

Z(t) = the random vector in size of (N × 1) at the time t

Suhartono and Subanar [20] introduced a new method for determining weight using the result of cross-correlation normalization between locations at a congruent time lag.

$$\hat{\rho}_{ij}(k) = r_{ij}(k) = \frac{\sum_{k=1}^{n} [Z_i(t) - \overline{Z_l}] [[Z_j(t-k) - \overline{Z_j}]]}{\sqrt{(\sum_{t=1}^{n} [Z_i(t) - \overline{Z_l}]^2)(\sum_{t=1}^{n} [Z_j(t) - \overline{Z_j}]^2}}$$
(4)

The determination of location weight for the GSTAR model (1;p) is as follows:

$$w_{ij} = \frac{r_{ij}(1)}{\sum_{k \neq 1} |r_{ik}(1)|}$$
(5)

with i \neq j and the weight had fulfilled $\sum_{i \neq j} w_{ij} = 1$.

The weight of cross-correlation normalization represented the variance of correlation between locations occurring in the data.

RESULTS AND DISCUSSION

The COVID-19 cases in Indonesia were ever-increasing, and Java was regarded as the transmission epicenter. The increase in the COVID-19 cases is depicted in Figure 1.



Figure 1. The plot of the time series of COVID-19 cases in each province

Figure 1 indicates that as of 2 March-18 May 2020, the COVID-19 cases increased in all provinces in Java. On 18 May 2020, the highest number of cases, 5,555, was reportedly in Jakarta, whereas the lowest one, 185, was in Yogyakarta. Using the data of the total COVID-19 cases in six provinces in Java, we identified the correlation between provinces and the COVID-19 transmission in Java. Correlation between locations was identified using Pearson's correlation between provinces. The result of Pearson correlation quantification is presented in Table 1.

	Banten	Jakarta	West Java	Central Java	Yogyakarta	East Java
Banten	1	0.994	0.994	0.982	0.981	0.973
Jakarta	0.994	1	0.995	0.989	0.975	0.967
West Java	0.994	0.995	1	0.992	0.986	0.977
Central Java	0.982	0.989	0.992	1	0.983	0.980
Yogyakarta	0.981	0.975	0.986	0.983	1	0.996
East Java	0.973	0.967	0.977	0.980	0.996	1

 Table 1. The Correlation Value of the COVID-19 Cases between Provinces in Java

In Table 1, we can see that the data of the number of the COVID-19 cases in six provinces in Java had a high Pearson's correlation value which was higher than 0.9. It implies that the correlation of the COVID-19 cases between provinces in Java was strong.

Data Stationarity Test

Data stationarity testing was performed in two stages which were stationarity to variance and stationarity to the mean. Stationarity to variance was tested using the box-cox transformation. Data were regarded stationary if the lambda value was 1, signifying that $Var(Z_t) = Var(Z_{t-k})$. The result of the stationarity test to variance is shown in Table 2.

Table 2. The Result of Dox-Cox Transformation							
Location	2		Transfo	Final Transformation			
LUCATION	λ	Trans.	λ	Trans.	λ		
Banten	0.20	$Z_t^{0.20}$	1.00	-	-	$Zt^{0.20}$	
Jakarta	0.20	$Z_t^{0.20}$	1.00	-	-	$Z_{t}^{0.20}$	

Location	2		Transfo	Final Transformation		
Location	λ	Trans.	λ	Trans.	λ	
West Java	0.19	$Z_{t}^{0.19}$	1.00	-	-	$Z_{t}^{0.19}$
Central Java	0.00	Ln(Zt)	1.00	-	-	Ln(Zt)
Yogyakarta	0.00	Ln(Zt)	0.00	Ln(Zt)	1.00	$Ln(Ln(Z_t))$
East Java	0.00	Ln(Zt)	0.50	$Z_{t}^{0.50}$	1.00	$Ln(Z_t)^{0.50}$

As seen in Table 2, the initial data had not fulfilled the stationarity to variance yet. Several transformations were thus called for. After conducting the data stationarity test, we did the stationarity test to the mean. The test was conducted using an augmented Dickey-Fuller test. The result of the stationarity test to the mean is indicated in Table 3.

	Table 3. The Result of the Augmented Dickey-Fuller Test								
Location	Lag	0	1	2					
Danton	π	98.73	60.83	34.96					
Danten	p-value	0.001	0.001	0.001					
Iakarta	π	107.89	32.98	21.15					
Jakalta	p-value	0.001	0.001	0.001					
Westlere	π	100.74	40.22	28.78					
west java	p-value	0.001	0.001	0.001					
Control Iovo	π	122.94	55.84	29.66					
Central Java	p-value	0.001	0.001	0.001					
Vogualiarta	π	75.55	51.84	42.09					
TUgyakarta	p-value	0.001	0.001	0.001					
Fact Jawa	π	155.97	57.49	25.85					
Lasi java	p-value	0.001	0.001	0.001					

From the augmented Dickey-Fuller test, we acquired predicted values less than the real ones (0.05). It indicates that the data had fulfilled the stationarity to variance.

Interpretation of the GSTAR Model Parameter

Model identification was aimed to find the autoregressive GSTAR model order. The order was elicited by identification using AIC. The lag with the smallest AIC value was regarded as the autoregressive GSTAR model order. Table 4 lists the AIC values.

Table 4. The AIC Value in Model Order Selection										
Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5				
AR 0	34.0876	35.2909	35.5546	36.0924	36.8882	36.1828				
AR 1	31.9424	32.9755	33.3945	33.9363	33.9725	33.0619				
AR 2	32.0815	33.1868	33.0648	33.8555	34.487	34.3527				
AR 3	32.006	32.8989	33.096	35.128	35.9621	36.0125				
AR 4	32.9289	34.1187	32.882	35.2756	35.429	42.0827				
AR 5	34.6863	35.2259	35.61	35.7834	42.3544					

Table 4 shows the smallest AIC value at the lag AR(1) and MA(0), hence the GSTAR(1)(1,0,0) model.

Interpretation of the GSTAR Model Parameter

The GSTAR model was a particular form of VAR engaging spatial elements. Estimating the GSTAR(1) (1,0,0) spatial parameters with the Ordinary Least Square method using cross-correlation normalization weight generated the following parameters.

Table 5. The Parameters of the GSTAR(1)(1,0,0) Model							
Location	Parameter	Estimation					
Denter	$\phi_{10}^{(1)}$	1.015					
Banten	$\phi_{11}^{(1)}$	0.793					
Jakarta	$\phi_{10}^{(2)}$	0.915					
јакапа	$\phi_{11}^{(2)}$	0.984					
Most Isra	$\phi_{10}^{(3)}$	0.758					
west java	$\phi_{11}^{(3)}$	1.031					
Control Iour	$\phi_{10}^{-(4)}$	-0.003					
Central Java	$\phi_{11}^{(4)}$	0.256					
Varializatio	$\phi_{10}^{-(5)}$	0.118					
Yogyakarta	$\phi_{11}^{(5)}$	0.061					
East Java	$\phi_{10}^{-(6)}$	0.088					
	$\phi_{11}^{(6)}$	-0.013					

Referring to Table 5, we generated the matrix equation of the GSTAR(1)(1,0,0) model, which is as follows:

r7	(+)7							F7 (+ 1)
$ _{L_1}$	(l)	r1.01	5 0	0	0	0	0	$ _{21}^{(l-1)} $
Z_2	(t)	0	0.91	50	0	0	0	$ Z_2(t-1) $
Z_3	(t)	_ 0	0	0.758	B 0	0	0	$ Z_3(t-1) $
Z_4	(t)	- 0	0	0	-0.0	3 0	0	$ Z_4(t-1) $
Z_5	(t)	0	0	0	0	0.11	8 0	$ Z_5(t-1) $
Z_6	(t)	Γ0	0	0	0	0	0.088	$\left[\frac{1}{Z_{6}}(t-1) \right]$
0	r0.79	3 0	0	0	0	0	1	
	0	0.98	84 0	0	0	0		
	0	0	1.03	1 0	0	0		
Ŧ	0	0	0	0.25	6 0	0		
	0	0	0	0	0.06	61 0		
	L 0	0	0	0	0	-0.0)13	
Г	0	0.256	0.116	0.209	0.183	ן0.236	$Z_{1}(t-1)$	1)] $\begin{bmatrix} e_1(t) \end{bmatrix}$
0.2	205	0	0.187	0.217	0.160	0.231	$Z_{2}(t-1)$	1) $e_2(t)$
0.1	192	0.223	0	0.233	0.150	0.201	$Z_{3}(t-1)$	1) $ _{1} e_{3}(t) $
0.0)92	0.273	0.194	0	0.231	0.210	$Z_4(t-1)$	1) $ ^{\top} e_4(t) $
0.1	156	0.242	0.079	0.276	0	0.247	$Z_{5}(t-1)$	1) $ e_5(t) $
L0.1	134	0.241	0.180	0.124	0.321	0]	$Z_6(t-1)$	1) $\begin{bmatrix} e_6(t) \end{bmatrix}$

The following matrix equation was derived from the above equation.

$[Z_1]$	(t)	г1.015	0	0	0	0	ך 0	$[Z_1(t-1)]$
Z_2	(t)	0	0.91	5 0	0	0	0	$Z_2(t-1)$
Z_3	(t)	0	0	0.758	0	0	0	$Z_3(t-1)$
Z_4	(t)	0	0	0	-0.03	0	0	$Z_4(t-1)$
Z_5	(t)	0	0	0	0	0.118	0	$Z_5(t-1)$
LZ_6	(t)	L 0	0	0	0	0	0.088J	$Z_6(t-1)$
	г О	0.2	59	0.118	0.212	0.186	0.239-	$[Z_1(t-1)]$
	0.202	2 ()	0.184	0.213	0.157	0.227	$ Z_2(t-1) $
Т	0.198	3 0.2	29	0	0.240	0.155	0.207	$ Z_3(t-1) $
I	0.024	ł 0.0	69	0.049	0	0.059	0.054	$ Z_4(t-1) $
	0.009	0.0	15	0.005	0.017	0	0.015	$ Z_5(t-1) $
	L-0.00	2 -0.	004	-0.002	-0.002	-0.004	0 -	$\left\lfloor Z_{6}(t-1) \right\rfloor$
ſ	$e_1(t)$							
	$e_2(t)$							
	$e_3(t)$							
	$e_4(t)$							
	$e_5(t)$							
L	$e_6(t)$							

From the model generated, we made a comparison between the actual and predicted data, in which we acquired an RMSE and MAPE value of 0.005 and 1.43, respectively. The two gave us a hint that the model generated was good.

Prediction Result

From the equation, we forecasted the total cases for the next 14 days, namely 19 May-1 June 2020, the result of which is presented in Table 6.

	Table 6. The Predicted COVID-19 Cases on 19 May-1 June 2020										
	Banten	Jakarta	West Java	Central Java	Yogyakarta	East Java					
1	628	5662	1689	1214	195	2321					
2	651	5738	1746	1249	201	2438					
3	674	5802	1804	1283	207	2561					
4	699	5850	1862	1318	214	2689					
5	725	5881	1920	1352	220	2822					
6	754	5892	1978	1386	227	2961					
7	784	5880	2035	1419	233	3105					
8	817	5841	2092	1450	240	3255					
9	852	5773	2148	1481	247	3411					
10	891	5672	2202	1509	254	3573					
11	933	5533	2254	1535	260	3741					
12	979	5351	2305	1559	267	3916					
13	1030	5122	2352	1579	274	4097					
14	1086	4839	2396	1596	280	4284					

Table 6. The Predicted COVID-19 Cases on 19 May-1 June 2020

The prediction stated that the total cases in all provinces in Java would increase, except Jakarta, in which there would be a declined total number of cases. The prediction was based on the assumption that there was no change in social interaction in the community.

CONCLUSIONS

The model acquired in modeling the COVID-19 cases in Java was the GSTAR(1)(1,0,0) model. Our predicted COVID-19 case data was close to the actual number of COVID-19 cases in Java. A Spatio-temporal model could be used to predict the number of COVID-19 cases in Java. Human-to-human transmission likely had a cross-location impact due to an interaction between individuals. Our prediction indicates that all provinces in Java, but Jakarta, would likely have an increase in the total number of COVID-19 cases for the next 14 days.

ACKNOWLEDGMENTS

We would like to thank the University of Brawijaya University for funding and support of this research.

REFERENCES

- [1] W. H. Organization, "WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020," *Geneva, Switz.*, 2020.
- [2] W. H. Organization, "Novel Coronavirus (2019-nCoV), Situation report-70, as of March 30, 2020," URL https//www. who. int/docs/defaultsource/coronaviruse/situation-reports/20200330-sitrep-70-covid-19. pdf? sfvrsn= 7e0fe3f8_2.
- [3] Y. Han and H. Yang, "The transmission and diagnosis of 2019 novel coronavirus infection disease (COVID-19): A Chinese perspective," *J. Med. Virol.*, vol. 92, no. 6, pp. 639–644, 2020.
- [4] N. Van Doremalen, T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, and B. N. Williamson, "& Lloyd-Smith, JO (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1," *N. Engl. J. Med.*
- [5] A. Susilo *et al.*, "Coronavirus Disease 2019: Tinjauan Literatur Terkini," *J. Penyakit Dalam Indones.*, vol. 7, no. 1, pp. 45–67, 2020.
- [6] K. Roosa *et al.*, "Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020," *Infect. Dis. Model.*, vol. 5, pp. 256–263, 2020.
- [7] J. S. Jia, X. Lu, Y. Yuan, G. Xu, J. Jia, and N. A. Christakis, "Population flow drives Spatiotemporal distribution of COVID-19 in China," *Nature*, pp. 1–11, 2020.
- [8] A. S. Albana, "Optimasi Alokasi Pasien untuk Kasus COVID-19 Wilayah Surabaya," *J. TECNOSCIENZA*, vol. 4, no. 2, pp. 181–200, 2020.
- [9] M. Fajar, "ESTIMATION OF COVID-19 REPRODUCTIVE NUMBER CASE OF INDONESIA."
- [10] D. Giuliani, M. M. Dickson, G. Espa, and F. Santi, "Modelling and predicting the Spatio-temporal spread of Coronavirus disease 2019 (COVID-19) in Italy," *Available SSRN 3559569*, 2020.
- [11] D. Giuliani, M. M. Dickson, G. Espa, and F. Santi, "Modelling and predicting the spread of Coronavirus (COVID-19) infection in NUTS-3 Italian regions," *arXiv Prepr. arXiv2003.06664*, 2020.
- [12] A. Iriany and B. N. Suhariningsih, "Ruchjana and Setiawan," Prediction of Precipitation data at Batu Town using the GSTAR (1, p)-SUR Model,"," J. Basic Appl. Sci. Res., vol. 3, no. 6, pp. 860–865, 2013.
- [13] B. N. Ruchjana, S. A. Borovkova, and H. P. Lopuhaa, "Least-squares estimation of

Generalized Space-Time AutoRegressive (GSTAR) model and its properties," in *AIP Conference Proceedings*, 2012, vol. 1450, no. 1, pp. 61–64.

- [14] D. D. Prastyo, F. S. Nabila, M. H. Lee, N. Suhermi, and S.-F. Fam, "VAR and GSTARbased feature selection in support vector regression for multivariate Spatiotemporal forecasting," in *International Conference on Soft Computing in Data Science*, 2018, pp. 46–57.
- [15] W. W. S. Wei, "Time series analysis," in *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*, 2006.
- [16] S. Makridakis, S. C. Wheelwright, and V. E. McGee, "Metode dan aplikasi peramalan," *Jakarta: Erlangga*, 1999.
- [17] D. C. Jonathan and C. Kung-Sik, "Time series analysis with applications in R," *SpringerLink, Springer eBooks*, 2008.
- [18] D. U. Wutsqa, S. B. Suhartono, and B. Sutijo, "Generalized Space-Time Autoregressive Modeling," in *Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Applications (ICMSA2010)*, 2010.
- [19] S. Borovkova, H. P. Lopuhaä, and B. N. Ruchjana, "Consistency and asymptotic normality of least squares estimators in generalized STAR models," *Stat. Neerl.*, vol. 62, no. 4, pp. 482–508, 2008.
- [20] S. Suhartono and S. Subanar, "The Optimal Determination Of Space Weight in Gstar Model by Using Cross-Correlation Inference," *Quant. Methods*, vol. 2, no. 2, pp. 45– 53, 2006.