The Spatial-SIR Model for COVID-19 Transmission by Resampling Method to Selecting Parameter Study Case: Java Island, Indonesia

Abdullah Sonhaji
Program Study of Mathematics
Institut Teknologi Bandung
Bandung, 40132, INDONESIA
a.sonhaji@gmail.com

Udjianna S. Pasaribu
Program Study of Mathematics
Institut Teknologi Bandung
Bandung, 40132, INDONESIA
udjianna@math.itb.ac.id

Kurnia Novita Sari
Program Study of Mathematics
Institut Teknologi Bandung
Bandung, 40132, INDONESIA
kurnia@math.itb.ac.id

Hennie Husniah
Department of Industrial Engineering
Langlangbuana University
Bandung, 40116, INDONESIA
hennie.husniah@gmail.com

Nisa Fadlilah Fathul Ilmi
Program Study of Mathematics
Institut Teknologi Bandung
Bandung, 40132, INDONESIA
nisafadlilah.fi@gmail.com

Abstract

The current outbreak of coronavirus disease (COVID-19) has become a global issue to its quick and widespread over the world, including in Indonesia. More than 60% of positive cases came from Java island, therefore the proposed model focused on six provinces in this area. We developed a discrete-time stochastic epidemic model, such as Spatial-SIRD model, associated with the mobility of people by public transportation (air and land). Model parameters were estimated by fitting the data of October 22nd – 28th, 2020 and November 8th-November 14th, 2020 with the model. At the beginning of the estimation process, we used the coefficient of regression from the observation to estimate the range of parameters. Afterward, the order statistics method was carried out to determine the best parameters so we could forecast the number of infectious of each province. The SIR model was created by applying the regression rate

of infection parameters before and after the long holiday from October 28 to November 1, 2020. The effect of this long holiday was that it could increase the number of cases so that there was a difference in the rate of infection.

Keywords
Contact rate, COVID-19, Origin-Destination Matrix, Spatial-SIR model, regression.

1. Introduction

The coronavirus disease 19 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-COV-2). Some symptoms arising from COVID-19 is fever, cough, fatigue, shortness of breath, pneumonia, and other respiratory tract symptoms, and in many cases progresses to death (Bi et al, 2020). On March 11, 2020, COVID-19 had spread to more than 100 countries and 4292 deaths, so WHO characterized as a pandemic. The first case of COVID-19 in Indonesia is confirmed on March 2, 2020, which infected two women in Depok, West Java. Based on data as of June 9, 2020, more than 60% of the total infected people came from Java island.

This virus can spread by contact with an infected person through droplets when a person coughs or sneezes, or through saliva droplets (He et.al, 2020). Thus, the first thing to do is not go to the place where people gather to reduce the rate of contact between infected people and not. Preventive steps taken by the Indonesian government are social distancing that took effect since March 16. Since then, office activities in postpone, shopping centers had been closed, and school or college activities had also carried out online. Displacement or mobilization from the population also dramatically influences the transmission rate of COVID-19 because the transmission occurs among humans.

Petropoulos and Makridakis (2020) had applied time series forecasting approaches to predict the number of infected in China. They produced forecasts using the exponential smoothing family model without involving factors that influence it. Other researchers, who had adapted the time series technique are Sarkodie and Owusu (2020). In addition to the time series model, their model also were based on data panel settings. For the case in Indonesia, Nuraini (2020) applied Richard's curve model to predict the number of infected by matching observational curves in Indonesia to those in several countries, namely: China, Italy, Iran, South Korea, and the United States.

The dynamic model of SEIR (Susceptible, Exposed, Infected, Removed) is prevalent in modeling this COVID-19 case. Most of the parameters estimated from the model are R_0 (basic reproduction rate). Wu et al. (2020) applied this model to predict the potential domestic and international spread of the COVID-19 outbreak in Wuhan. The Monte Carlo Markov Chain method with Gibbs sampling and non-informative flat prior was used to estimate R_0. In addition to determining cases in China, Kucharski et al. (2020) also estimated the number of internationally exported cases from Wuhan through the model. In another study, He et al. used the stochastic model discrete by adding hospitalized and quarantine compartments to the SEIR model scheme (He at al., 2020). The modified SEIR model was also applied to predict many cases in Indonesia by Rustan and Handayani (2020).

This article contributed a spatial analysis method of COVID-19 transmission that has not involved in the above studies, such as the Spatial-SIR model. This model adapted from the ITB Report of COVID-19 Modeling by involving community mobilization behavior between provinces by land and air. This model concerned COVID-19 transmission in 6 provinces in Java Island as an epicenter. The resampling method and principle in order statistics were applied to estimate the parameters of the model. Hopefully, the spatial COVID-19 transmission approach can be used for preventive action against COVID-19 in each province of Java.

2. Methods

Basically, SIR model consists of compartments: Susceptible (S), Infected (I), and Removed (R). Let *N* denote the total population size which the following differential equation

$$\frac{dS(t)}{dt} = -\beta \frac{I(t)}{N} S(t)$$
$$\frac{dI(t)}{dt} = \beta \frac{I(t)}{N} S(t) - \gamma I(t)$$
$$\frac{dR(t)}{dt} = \gamma I(t)$$

with initial conditions $S(0), I(0)$, and $R(0)$. The parameters used are β for infection rate and γ^{-1} for infection period. Discrete time-point series $t = 1, 2, 3, \ldots$ are considered as the development of disease time (day as unit of time). In this research, SIR model is modified so that the removed compartment is divided into two, namely recovery and death.

Modified SIR models by entering the mobility factor for discrete times can be written as in this equation.

$$S_i(t+1) = S_i(t)\left(1 - \frac{\beta_i I_i(t)}{N_i} - \alpha_i\left(\beta_i \frac{\sum_{j=1}^{M} m_{j,i} I_i(t)}{N_i + \sum_{j=1}^{M} m_{j,i}}\right)\right) + \eta_i R_i(t)$$

$$I_i(t+1) = I_i(t) + S_i(t)\left(\frac{\beta_i I_i(t)}{N_i} + \alpha_i\left(\beta_i \frac{\sum_{j=1}^{M} m_{j,i} I_i(t)}{N_i + \sum_{j=1}^{M} m_{j,i}}\right)\right) - \left(\frac{\gamma_i + d_i}{2}\right) I_i(t)$$

$$R_i(t+1) = R_i(t) - R_i^*(t)\eta_i + \left(\frac{\gamma_i + d_i}{2}\right) I_i(t)$$

with

i	:	number of provinces, $i = 1, 2, 3, \ldots, 6$
t	:	day, $t = 1, 2, 3, \ldots$
$S_i(t)$	:	the number of susceptible individuals in province i who can be infected at day t
$I_i(t)$	:	the number of individuals in province i who has been infected (from susceptible) and able to transmit the virus at day t
$R_i(t)$	:	number of individuals in province i who has been recovered and death (from infected) at day t.
$R_i^*(t)$	:	number of individuals in province i who has been death (from infected) at day t
β_i	:	the individual contact rate in district i
δ_i	:	the death rate in province i
γ_i	:	the recovery rate in province i
α_i	:	the percentage of incoming migrant using public transportation in province i
$m_{i,l}$	:	the number of individuals migrate from province l to province i.
$\mathcal{T}_i$	:	the neighboring association from province i
N_i	:	the population in province i

In this model, the initial number $I_i(0)$ and $R_i(0)$ are dependent on the first day of the data (fixed). Specially for $S_i(0)$, initial number is obtained from the number of total population in i^{th} province multiplied by a constant based on the incubation period and the estimated number of meeting people for the population in Indonesia. The maximum value for $S_i(0)$ is the total population in each province i. The Spatial-SIR model scheme in this study can be seen in Figure 1.

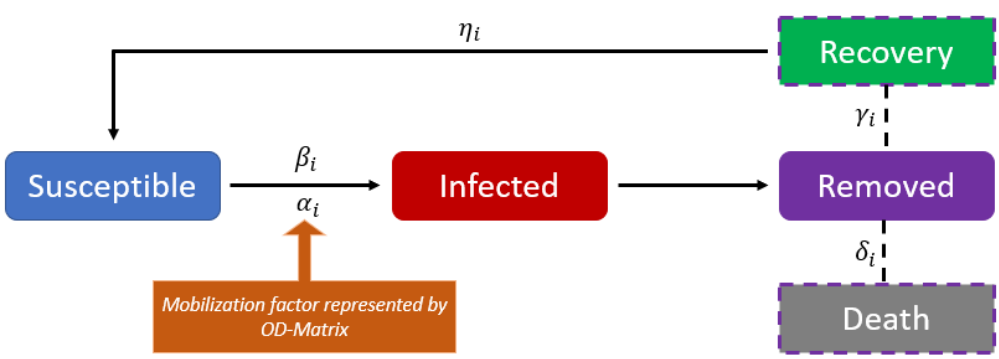


Figure 1. Scheme of Spatial-SIR model including mobilization factor for susceptible become infected.

The estimation of each parameter is carried out through the resampling method of distribution. Information on the distribution of each parameter is unknown, so it is assumed that parameters follow uniform distribution as in Table 2. The interval value of each parameter is determined based on the previous research or the accepted values in the corresponding field. The distribution interval of the contact rate β_i is obtained from the slope or coefficient of the data of the last seven days of infected cases of the six provinces. The recovery interval, the rate of death and the rate of re-infection are determined based on existing data. The parameter estimation algorithm is explained in Figure 2.

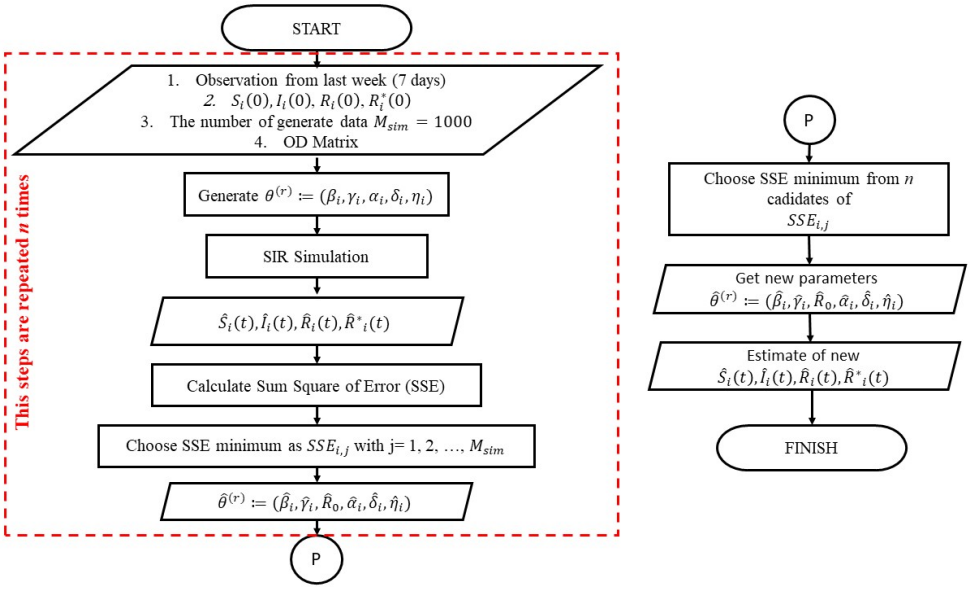


Figure 2. The parameter estimation algorithm in Spatial-SIR by resampling method.

3. Result and Discussion

In this paper, we used data on COVID-19's positive cases on October 22nd – 28th, 2020 (before the long weekend) and November 8th-14th, 2020 (after a long weekend). There was a suspicion that the mobility during holidays would affect the increasing cases 2 weeks thereafter (because of the incubation period lasts for approximately 2 weeks). In addition, due to this incident, the spread rate of COVID-19 in several or every provinces could also change.

Table 1. The Origin-Destination (OD) matrix for population mobilization (in Java island) through land and air.

O \ D	**Banten**	**DI Yogyakarta**	**DKI Jakarta**	**Jawa Barat**	**Jawa Tengah**	**Jawa Timur**
Banten	0 (0)	17908 (3341)	23693 (0)	20763 (1)	18486 (4621)	17209 (8472)
DI Yogyakarta	606 (3912)	0 (0)	627 (1577)	643 (197)	818 (0)	713 (468)
DKI Jakarta	21559 (0)	17005 (1718)	0 (0)	19786 (99)	17340 (1567)	15930 (2256)
Jawa Barat	68185 (0)	64417 (184)	71185 (57)	0 (0)	64141 (190)	59827 (1001)
Jawa Tengah	9121 (5515)	11946 (1)	9493 (1517)	9606 (189)	0 (0)	10125 (736)
Jawa Timur	8898 (6845)	10846 (347)	9148 (3417)	9107 (866)	10472 (525)	0 (0)

Other data that we used in this model was the data on the number of population mobilization of incoming from and outgoing to each province (in Java island) through land and air. The information was obtained from BPS, which had been built into an OD (origin-destination) matrix with units of people per day. The OD matrix is shown in Table 1. Moreover, the distributions applied to the Spatial-SIR parameters are shown in Table 2.

Table 2. The distributions of Spatial-SIR model parameters of each province-*i*.

Parameters	Definition for each province	Distribution	Detail/source

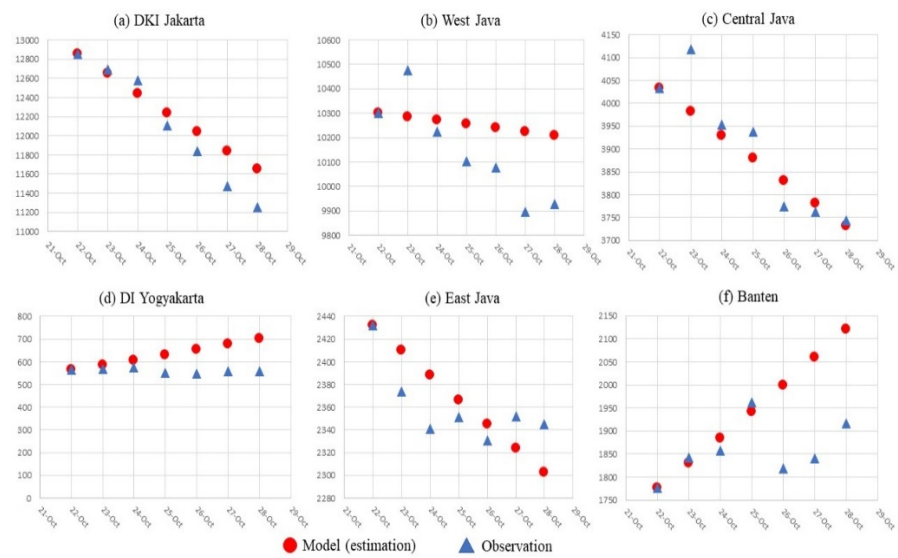
β_i	Contact rate	$\beta_i \sim U\,[0.09, 0.2]$	
γ_i	Recovery rate	$\gamma_i \sim U\,[\frac{1}{21}, \frac{1}{14}]$	
R_{0_i}	Basic Reproduction Rate	$R_{0_i} = \frac{\beta_i}{\gamma_i}$	
α_i	Percentage of mobility	$\alpha_i \sim U\,[0.4, 0.7]$	
δ_i	Death rate	$\delta_i \sim U\,[0.04, 0.09]$	
η_i	Re-infected rate	$\eta_i \sim U\,[\frac{1}{1100}, \frac{1}{1000}]$	
$S_i(0)$	Initial susceptible population	Fixed	Population data per province in 2018
$I_i(0)$	Initial infected population	Fixed	Ministry of Health's data
$R_i(0)$	Initial total removed population	Fixed	Ministry of Health's data
$R_i^*(0)$	Initial removed (dead) population	Fixed	Ministry of Health's data

3.1 Prediction Cases Before Long Weekend

Figure 3 shows the comparison plots of the infected's number of the observation data on October 22nd – 28th, 2020 which was estimated by using Spatial-SIR model. Moreover, the percentage of errors of each province is also shown. It was calculated using the following equation.

$$\%\,Error_i = \frac{I_i(t) - \hat{I}_i(t)}{I_i(t)} \times 100\% \qquad (1)$$

During this period, almost all positive cases trend in every province decreased, except for Banten province. The best fitting result is shown in Figure 3.(c) i.e. of Central Java province. This can also be seen from the smallest average error, that is 1.09%.

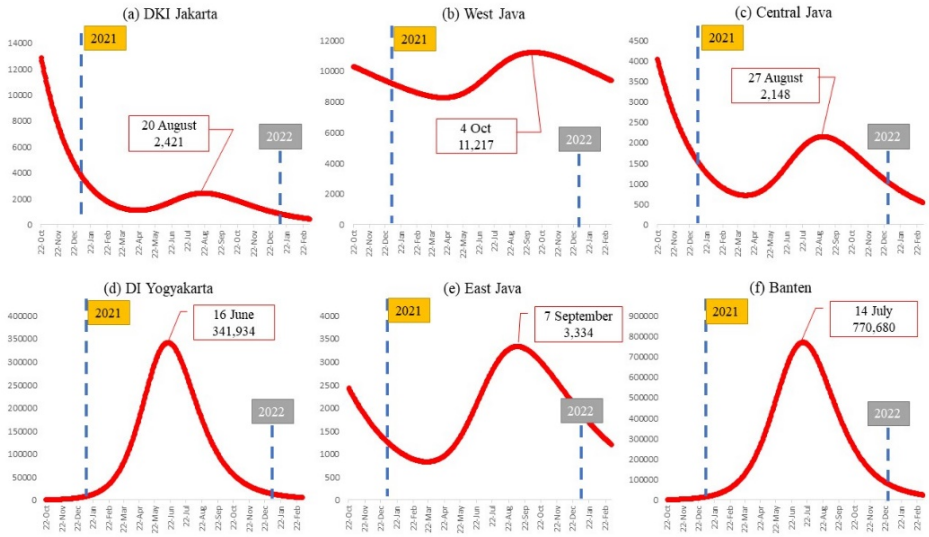


% Err	(a)	(b)	(c)
22	0	0	0
23	0.38	1.82	3.35
24	1.12	0.45	0.58
25	1.09	1.51	1.48
26	1.70	1.62	1.45
27	3.22	3.31	0.50
28	3.51	2.84	0.29
Mean	1.57	1.65	1.09
% Err	(d)	(e)	(f)
22	0	0	0
23	3.18	1.51	0.69
24	5.52	2.01	1.46
25	14.17	0.65	1.04
26	19.02	0.59	9.99
27	21.61	1.21	11.86
28	25.61	1.81	10.69
Mean	12.73	1.11	5.10

Figure 3. Data estimation and percentage of error fitting of the Spatial-SIR model for infected cases of the six provinces before long weekend data (22 – 28 October 2020).

Based on the parameter in Figure 4, the number of infected people for the next 500 days can be predicted. The outbreak would occur faster in DI Yogyakarta along with a quite high rate of transmission, 0.088. The behavior of the society in DKI Jakarta, West Java, Central Java, and East Java were similar, such as the number of positive case would decrease until the first quartile in 2021.

Par	(a)	(b)	(c)

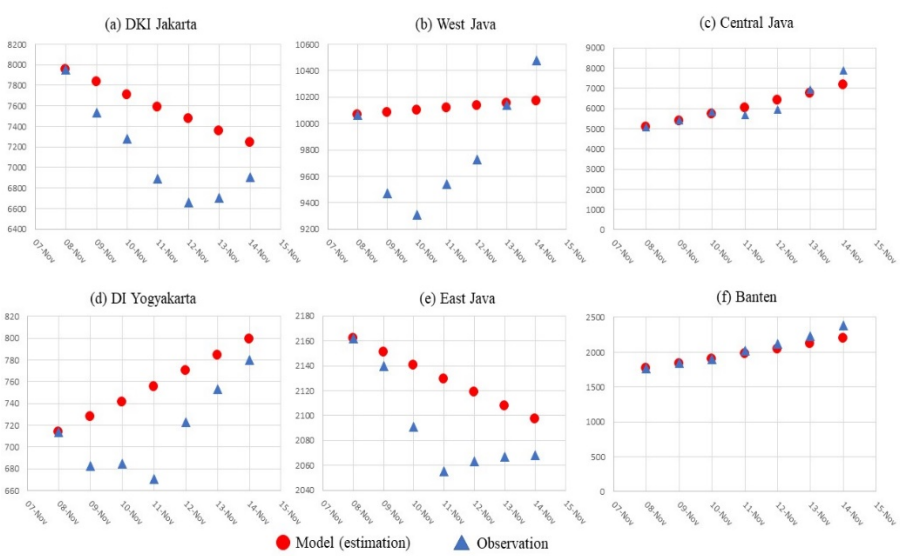


Par	(a)	(b)	(c)
β	0.050	0.061	0.057
γ	1/20	1/20	1/19
R_0	0.992	1.229	1.088
α	0.638	0.582	0457
δ	0.081	0.076	0.088
η	$\frac{1}{1053}$	$\frac{1}{1020}$	$\frac{1}{1043}$
Par	(d)	(e)	(f)
β	0.088	0.057	0.091
γ	1/15	1/20	1/18
R_0	1.324	1.140	1.646
α	0.542	0.524	0.400
δ	0.042	0.083	0.069
η	$\frac{1}{1092}$	$\frac{1}{1092}$	$\frac{1}{1087}$

Figure 4. Prediction and parameter estimation of Spatial-SIR model for the infected cases of the six provinces before long weekend data.

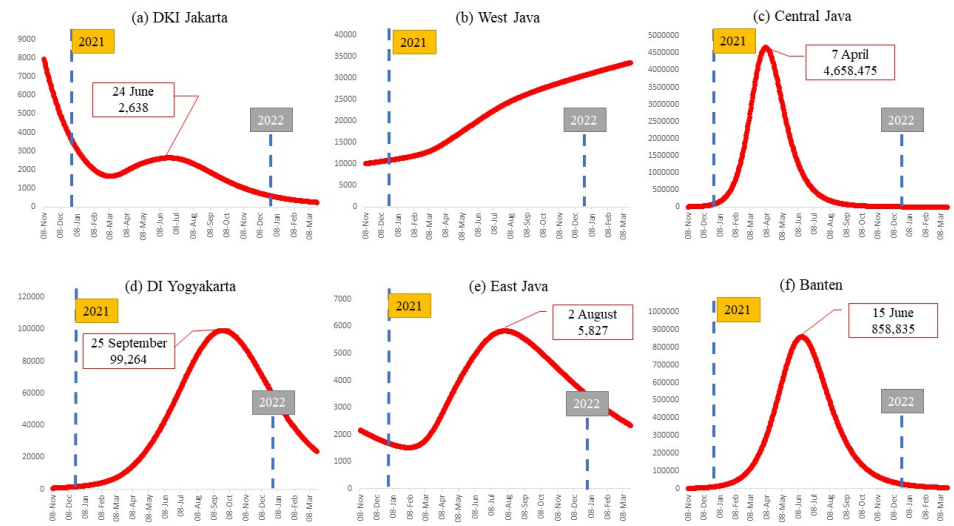
3.2 Prediction Cases After Long Weekeend

The percentage of error in Figure 5 is calculated using Equation (1). The impact of the long holiday on October 28th - November 1st began in this period. The trend of positive cases in each province began to rise again since November 10th. The best fitting results are shown in Figure 3.(c) and (f) for East Java and Banten provinces, with an average of percentage error 1.79% and 2.73%. The unfavorable estimation results shown in Figures 5(a), 5(b), and 5(d) are presumed that there was a significant change in trends during that period.



%Err	(a)	(b)	(c)
8	0	0	0
9	3.93	6.43	0.25
10	5.87	8.47	2.25
11	10.07	6.04	6.39
12	12.15	4.17	7.72
13	9.74	0.09	2.51
14	4.85	2.95	9.26
Mean	6.66	4.02	4.05
%Err	(d)	(e)	(f)
8	0	0	0
9	6,53	0,52	0,40
10	8,24	2,35	0,25
11	12,59	3,61	2,42
12	6,48	2,69	3,41
13	4,17	1,97	4,84
14	2,46	1,40	7,80
Mean	5.78	1.79	2.73

Figure 5. Data estimation and percentage of error fitting of the Spatial-SIR model for infected cases of the six provinces after long weekend data.



Par	(a)	(b)	(c)
β	0.053	0.061	0.125
γ	1/19	1/16	1/15
R_0	1.013	0.975	1.876
α	0.432	0.515	0.657
δ	0.084	0.056	0.067
η	$\frac{1}{1073}$	$\frac{1}{1093}$	$\frac{1}{1041}$
Par	**(d)**	**(e)**	**(f)**
β	0.086	0.054	0.106
γ	1/17	1/18	1/14
R_0	1.459	0.970	1.490
α	0.457	0.494	0.453
δ	0.078	0.063	0.070
η	$\frac{1}{1086}$	$\frac{1}{1052}$	$\frac{1}{1099}$

Figure 6. Prediction and parameter estimation of Spatial-SIR model for the infected cases of the six provinces after long weekend data.

The prediction's results based on data during this period show that the behavior of many infected cases in West Java will continue to rise until 2022 if current conditions are maintained. In Central Java and Banten provinces, the outbreak is hitting in the first and second quartiles of 2021. This is because of the contact rate in these provinces is the highest at more than 10%.

The comparison of parameter estimation before and after the holiday is shown in Table 3. After the long holiday, the changes in population behavior were seen to be very significant in Central Java Province. All parameters were greater than before holiday, except the death rate. This has led to an explosion of cases (Figure 6c) so that conditions of mobility and interaction in this province need to be limited.

Table 3. Comparison of parameter estimation of the six provinces between before (B) and after (A) long weekend.

Province	DKI Jakarta		West Java		Central Java		DI Yogyakarta		East Java		Banten	
Par	B	A	B	A	B	A	B	A	B	A	B	A
β	0,050	**0,053**	0,061	0,061	0,057	**0,125**	0,088	0,086	0,057	0,054	0,091	**0,106**
γ	1/20	1/19	1/20	**1/16**	1/19	**1/15**	1/15	1/17	1/20	**1/18**	1/18	**1/14**
R_0	0,992	**1,013**	1,229	0,975	1,088	**1,876**	1,324	1,459	1,140	0,970	1,646	1,490
α	0,638	0,432	0,582	0,515	0,457	**0,657**	0,542	0,457	0,524	0,494	0,400	**0,453**
δ	0,081	**0,084**	0,076	0,056	0,088	0,067	0,042	**0,078**	0,083	0,063	0,069	0,070
η	$\frac{1}{1053}$	$\frac{1}{1073}$	$\frac{1}{1020}$	$\frac{1}{1093}$	$\frac{1}{1043}$	$\frac{1}{\mathbf{1041}}$	$\frac{1}{1092}$	$\frac{1}{\mathbf{1086}}$	$\frac{1}{1092}$	$\frac{1}{\mathbf{1052}}$	$\frac{1}{1087}$	$\frac{1}{1099}$

4. Conclusion

This model aims to obtain an overview of the condition of each province in the case of Covid-19 desease spread in the future. One of the factors that can influence the condition of the case is the long vacation which causes an increase in the intensity of population mobilization and interaction. This could obstruct the government efforts to break the chain of COVID-19 spreading. Through this simulation, a comparison of the parameters before and after the long weekend on October 28th - November 1st, 2020 is obtained. The result of the simulation - with M_{sim} = 1000, n = 10 repetitions, and the distribution of the parameters depicted in Table 2 - shows that there has been a society behavior change during these two periods. The average of percentage error of the estimation results of the two periods is only 0.3% difference. Therefore, the resampling method in the Spatial-SIR model is sufficient to provide the desired description.

References

Petropoulos, F. and Makridakis, S., Forecasting the Novel Coronavirus COVID-19, *PLOS ONE*, 2020.

Sarkodie, S.A. and Owusu, P.A., Investigating the Cases of Novel Coronavirus Disease (COVID-19) in China using dynamic statistical techniques, *Heliyon*, 2020.

Nuraini, N., Khairudin, K., and Apri, M., Modeling Simulation of COVID-19 in Indonesia based on Early Endemic Data, *Commun. Biomath. Sci.*, Vol. 3, No. **1**, PP. 1-8, 2020

Wu., J.T., Leung, K., and Leung, G.M., Nowcasting and Forecasting the Potential Domestic and International Spread of the 2019-nCov outbreak originating in Wuhan, China: a Modelling Study, *The Lancet*: 395: 689-97, 2020.

Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., On behalf of the Centre for Mathematical Modelling of Infectious Diseases COVID-19 working group, Early Dynamics of Transmission and Control of COVID-19: A Mathematical Modelling Study, *Lancet Infect Dis* 20:553-58, 2020.

He, S., Tang, S., and Rong, L., A Discrate Stochastic Model of COVID-19 Outbreak: Forecast and Control, *Mathematical Biosciences and Engineering* 17(**4**): 2792-2804, 2020.

Rustan and Handayani, L., The Outbreak's Modeling Of Coronavirus (Covid-19) Using The Modified Seir Model In Indonesia, SPEKTRA: *Jurnal Fisika dan Aplikasinya*, 5(**1**): 61-68, 2020.

Biographies

Abdullah Sonhaji is a Doctoroal student in the Program Study of Mathematics, Institut Teknologi Bandung, Indonesia. His research interest is in Statistics Modeling specially in Genetics and Optimization Algorithm. His educational background includes the following: Master (Informatics Engineering, ITB, 2005); Bachelor (Mathematics, ITB, 1999).

Udjianna S. Pasaribu is an associate Professor in the Program Study of Mathematics, Institut Teknologi Bandung, Indonesia. Her primary research interests are in Hidden markov model, warranty, survival model, point process analysis, Customer lifetime values, Biostatistics, statistical process control, multivariate analysis, and sampling technique. Her educational background includes the following: Doctoral (Biostatistics, University of Wales Swansea, The United Kingdom, 1993); Master (Biostatistics, University of Wales Swansea, The United Kingdom, 1987); Bachelor (Mathematics, ITB, 1985). Numerous papers have been published in several scientific journals, including International Journal of Applied Mathematics and Statistics, Journal of Information, and many international proceedings published by IEEE, AIP, IOP, etc.

Kurnia Novita Sari is a Lecturer at the Program Study of Mathematics, Institut Teknologi Bandung, Indonesia. Her primary research interests are in Geostatistics and Actuarial Model. Her educational background includes the following: Doctoral (Mathematics, ITB, Indonesia, 2016); Master (Mathematics, ITB, 2009); Bachelor (Mathematics, ITB, 2004).

Hennie Husniah is a Lecturer at Department of Industrial Engineering, Langlangbuana University. She received her PhD from Department of Industrial, Engineering, Institut Teknologi Bandung, Indonesia. Her research interests include product warranties, analysis of warranty data, maintenance and reliability models. Her works have been published in the following journals: International Journal of Collaborative Enterprise, Quality Technology and Quantitative Management, Computer and Industrial Engineering, Journal of the Indonesian Mathematical Society, International Journal of Applied Mathematics and Statistics, Journal of Information, Jurnal Teknik Industri, Jurnal Teknologi, Computer & Indusrtrial Engineering Journal and many international proceedings published by IEEE, AIP, IOP, etc.

Nisa Fadlillah F.I. is a research assistant in the Program Study of Mathematics, Institut Teknologi Bandung, Indonesia. Her research interest is in multivariate analysis. Her educational background includes the following: Master (Mathematics, ITB, 2018); Bachelor (Mathematics, ITB, 2014).