Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 13(2) 2020: 170-181
http://ejournal.undip.ac.id/index.php/media_statistika

SUSCEPTIBLE INFECTED RECOVERED (SIR) MODEL FOR ESTIMATING COVID-19 REPRODUCTION NUMBER IN EAST KALIMANTAN AND SAMARINDA

Sifriyani Sifriyani[1], Dedi Rosadi[2]
[1] Study Program of Statistics, Department of Mathematics, Mulawarman University
[2] Department of Mathematics, Gadjah Mada University

e-mail: sifriyani@fmipa.unmul.ac.id

DOI: 10.14710/medstat.13.2.170-181

Article Info:
Received: 13 June 2020
Accepted: 1 October 2020
Available Online: 28 December 2020

Keywords:
COVID-19, Estimate, SIR, Simulation, Reproduction
.

Abstract: Modeling and analysis of Covid-19 data, especially on the modeling the spread and the prediction of the total number of cases for Indonesian data, has been conducted by several researchers. However, to the best of our knowledge, it has not been studied specifically for East Kalimantan Province data. The study of the data on the level of provincial and District/City level could help the government in making policies. In this study, we estimate the Covid-19 reproduction number, calculate the rate of recovery, the rate of infection, and the rate of death of East Kalimantan Province and Samarinda City. We also provide a prediction of the peak of the infection cases and forecast the total incidence of Covid-19 cases until the end of 2020. The model used in this research is the Susceptible Infected Recovered (SIR) model and the data used in the study was obtained from the East Kalimantan Public Health Office.

1. INTRODUCTION

Corona Virus Disease 2019 or Covid-19 is an infectious disease caused by a type of corona virus SARS-CoV-2. The Covid-19 virus infection was first discovered in the city of Wuhan, China, at the end of December 2019. This virus spreads very quickly and has spread to almost all countries, including Indonesia. This has led several countries to implement policies to impose lockdowns in order to prevent the spread of the Corona virus. Several cities in Indonesia have implemented a Large-Scale Social Restriction or Pembatasan Sosial Berskala Besar (PSBB) policy to suppress the spread of the virus.

The Covid-19 has spread to all provinces in Indonesia. According to data submitted by Covid-19 Response Acceleration Task Force on June 10, 2020, there were 1,241 confirmed positive cases recorded, therefore the accumulated number of confirmed positive cases in Indonesia was 34,316. The highest addition of cases for confirmed data occurred from March to June 6, 2020. The government also noted that there were an additional 715 recovered patients, thus a total of 12,129 patients have recovered.

Meanwhile, there were an additional 36 death, thus a total of 1,959 people have died (Gugus Tugas Percepatan Penanganan Covid -19, 2020).

The spread of Covid-19 in East Kalimantan, according to data from the East Kalimantan Public Health Office, on June 10, 2020, recorded a cumulative number of confirmed positive cases of 362 cases. A total of 221 people recovered and 3 died (Dinas Kesehatan Provinsi Kalimantan Timur, 2020). The growth rate of Covid-19 in East Kalimantan is shown in Figure 1.

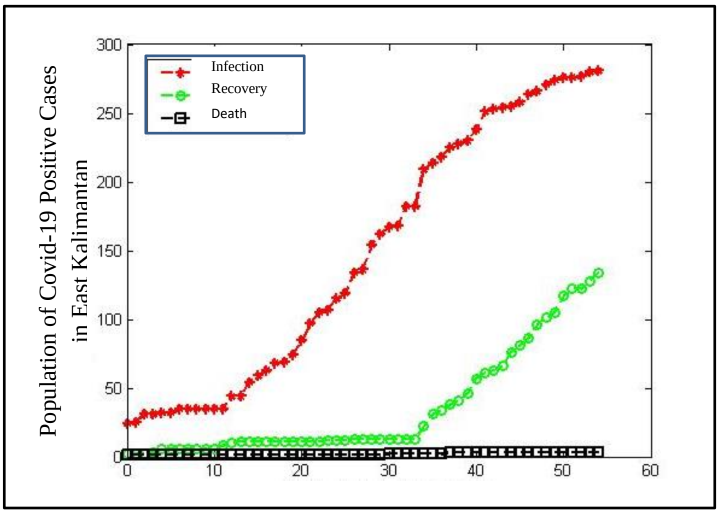

Figure 1. Time Series Graph of Accumulative Data for Covid-19 in East Kalimantan

Figure 1 shows the accumulative data of patients who test positive for Covid-19 continues to increase and the curve has not sloped or cases continue to increase every day. Patients recovered had an increase in contrast to the number of patients who died which is still at the same rate.

Several researchers have conducted Covid-19 modeling to predict the spread and the peak of Covid-19 cases, both in Indonesia and in several other provinces. This study uses Covid-19 data for East Kalimantan Province and Samarinda City using disease distribution modeling. The data used was obtained from the East Kalimantan Public Health Office on the https://dinkes.kaltimprov.go.id/ and the East Kalimantan Communication and Informatics Office on the https://covid19.kaltimprov.go.id/.

This study uses mathematical model in epidemiology called Susceptible Infected Recovered (SIR) to estimate the reproduction number, to predict the peak of the infection cases, and to forecast the total incidence of Covid-19 cases until the end of 2020 in East Kalimantan Province and Samarinda City. We use three indicators: the rate of infection, the rate of recovery, and the rate of death based on daily time.

2. MATHEMATICAL MODELING IN EPIDEMIOLOGY

Susceptible Infected Recovered (SIR) model is one of mathematical models in infectious disease epidemiology, this model was introduced by W.O. Kermack and A.G. McKendrick in 1927. The model consists of three compartments: ***S*** for *Susceptible*, ***I*** for *Infectious* and ***R*** for *Recovered*. SIR model has been used to estimate infectious diseases which are transmitted from human to human and where recovery confers lasting resistance (Hethcote, 2000), (Diekmann, Heesterbeek and Britton, 2012).

2.1. Susceptible Infected Recovered (SIR) Model

SIR model is used to predict how a disease spreads, the total number infected, the duration of an epidemic, and to estimate various epidemiological parameters such as the reproduction number. Such models can show how different public health interventions may affect the outcome of the epidemic. SIR model can be shown as a flow diagram in which the boxes represent the different compartments and the arrows the transition between compartments, as in Figure 2.

Figure 2. Flow Diagram of SIR Model

The arrows should be labeled with the transition rates between compartments. The transition rate between S dan I is given by equation (1).

$$\frac{d(S/N)}{dt} = \beta SI/N^2 \tag{1}$$

SIR model by Kermack-McKendrick can be expressed by the following set of ordinary differential equations (2) (Hethcote, 2000).

$$\frac{dS}{dt} = -\frac{\beta IS}{N} \tag{2}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

SIR model on equation (2) have parameter β which is the rate of transmission of a disease through contact, parameter γ is the rate of recovery from disease infection, S is the stock of susceptible population, I is the stock of infected, R is the stock of recovered population and N is the sum of these three.

2.2. Basic Reproduction Number

Basic reproduction number R_0 is the average number of infections caused by a single infectious subject in a wholly susceptible population (Diekmann, Heesterbeek and Britton, 2012). Therefore, R_0 contains information on the spread of infectious disease outbreaks. This value shows the index of the occurrence of the outbreak. The followings are the explanation of R_0.

1. When $R_0 < 1$, it indicates that an infectious disease will stop being epidemic and even disappear by itself. The non-endemic equilibrium point is globally asymptotically stable, that is, the number of individuals infected initially will fade out on its own over time. If $R_0 < 1$, the area can be a New Normal life recommendation.
2. When $R_0 > 1$, it indicates that the number of infected will continue to increase until it reaches its equilibrium point.
3. When $R_0 = 1$, it indicates the infection will be endemic and persist in the population.

2.3. Effective Reproduction Number

Effective reproduction number (R) is the average number of secondary cases per infectious case in a population composed of both susceptible and non-susceptible hosts. We refer to R as an effective reproduction number when there is some immunity or some intervention measures are in place. When individuals are homogeneous and mix uniformly, R is defined as the mean number of infections generated during the infectious period of an infected individual. When $R > 1$, the number of cases will increase, as in the beginning of the epidemic, when $R_0 = 1$, it indicates the infection will be endemic and persist in the population, whereas $R < 1$, the number of cases will decrease. The effective reproduction number in this study uses several estimations; one of which is the reproduction number by involving the compartment **S** for susceptible.

$$R_t = R_0 \left(\frac{S(t)}{N(t)} \right) \tag{3}$$

where R_t is the effective reproduction number, R_0 is the basic reproduction number and $N(t) = S(t) + I(t) + R(t) + D(t)$. $D(t)$ is the number of death. Another method of reproduction number using the Time Dependent method was introduced by Jacco Wallinga in a study in the journal of epidemiology (Wallinga, 2004). The Time Dependent method calculates the reproduction number by finding the mean and probability value of the growth dynamics of Covid-19 every day. Equation (4) consists of individual case *i* and *j* at time t_i and t_j respectively, the probability of infection transmission from case *i* to case *j*:

$$P_{ij} = \frac{(t_i - t_j)}{\sum_{i \neq k} (t_i - t_k)} \tag{4}$$

The effective reproduction number R_t for all cases is the number of probability (4) from all Covid-19 cases.

$$P_{ij} = \frac{(t_i - t_j)}{\sum_{i \neq k} (t_i - t_k)} \tag{5}$$

We used equation (5) to calculate the reproduction number of Covid-19 in East Kalimantan Province for confirmed cases from March 19, 2020 to June 19, 2020.

2.4. Vaccination Coverage

We calculated the vaccination coverage by using basic reproduction number R_0 and effective reproduction number R_t (Grohskopf *et al.*, 2011).

$$V_R = 1 - \frac{1}{R} \tag{6}$$

$$V_R = 1 - \frac{1}{R} \tag{7}$$

where V_{R_0} and V_R are the proportion of individuals that must be vaccinated.

3. SIR Model of Covid-19 in East Kalimantan

3.1. The Growth Rate of Covid-19 Cases

The East Kalimantan Public Health Office has published Covid-19 daily data including confirmed positive cases, treated patients, PDP (en: patient under supervision), ODP (en: person under supervision), OTG (en: asymptomatic person), and other data related to Covid-19 cases. Figure 3 shows the statistical description of the daily growth rate of confirmed cases.

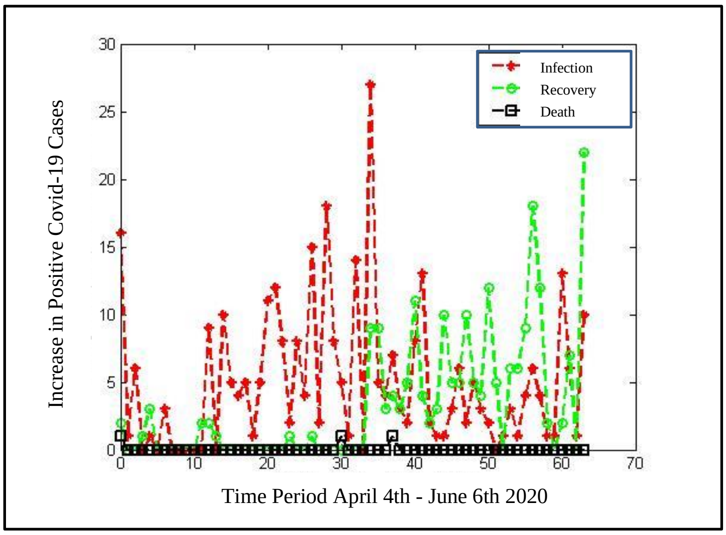

Figure 3. The Daily Growth Rate of Covid-19 in East Kalimantan

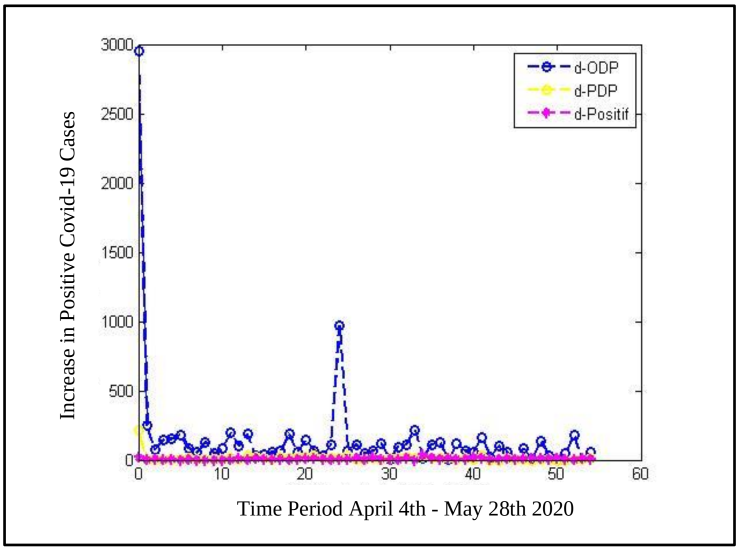

Figure 4. The Daily Growth Rate of ODP, PDP and Positive Cases in East Kalimantan

Each point in Figure 3 and 4 defines the daily data changes which is the difference between today's events (t) and yesterday's events (t-1). The daily growth rate and the highest number of cases from April 20 to May 28 was on May 7 with an increased by 27 positive cases and the average positive cases was increased by 4 cases per day.

3.2. Estimation of Covid-19 Reproduction Number in East Kalimantan

We modelled the data using the modification of SIR model (2) with several parameters such β as the rate of infection, γ as the rate of recovery and μ as and the rate of death.

Equation (2) was modified by adding the rate of death to obtain the derivative of the infected population which is obtained from the difference between the susceptible population and the number of recovered and dead patients (Yulida and Karim, 2020), the modification is shown on equation (8). We used compartment S as the stock of susceptible population at intervals *t*, we obtained S by accumulating the number of ODP, PDP and OTG in East Kalimantan Province. I is the stock of infected at intervals *t*. R is the stock of recovered population at intervals *t*.

$$\frac{dS}{dt} = -\frac{\beta IS}{S+I+R} = -\frac{\beta IS}{N}$$
$$\frac{dI}{dt} = \frac{\beta IS}{S+I+R} - (\gamma + \mu)I \qquad (8)$$
$$\frac{dR}{dt} = \gamma I$$

Equation (8) defines parameter β as the rate of infection, parameter γ as the rate of recovery, and parameter μ as the rate of death. We estimated parameters β, γ, μ in SIR model by using nonlinear least square method and Runge-Kutta model and we applied them to Covid-19 data in East Kalimantan Province. The objective function of SIR model (Yuni *et al.* 2020) is formed as follows:

$$\min_{\beta,\gamma,\mu} \|f(\beta,\gamma,\mu)\|^2 = \min_{\beta,\gamma,\mu} \frac{1}{3(S+I+R)}(Z) \qquad (9)$$
$$Z = \sum_{t=0}^{N-1}\left(S(t) - S_{data}(t)\right) + \sum_{t=0}^{N-1}\left(I(t) - I_{data}(t)\right) + \sum_{t=0}^{N-1}\left(R(t) - R_{data}(t)\right)$$

where f is a function in β, γ, μ and parameters β, γ, μ are estimated at intervals $[0,1]$ respectively.

The numerical method used in estimating SIR in this study is Runge-Kutta method (Karim *et al.*, 2018), (Karim and Gunawan, 2020). Meanwhile, the optimization method used is nonlinear least square method (Howell, Bates and Watts, 1990). The computational program used in this research is MATLAB (Keviczky *et al.*, 2019), (Duffy, 2016). Figure 5 shows the SIR estimation for confirmed cases at the beginning of Covid-19 from March 19 to April 2020, Figure 6 shows the SIR estimation for confirmed cases since April 2020, during that period several places and schools were closed, until the end of May 2020, and Figure 7 for confirmed cases from April 4 to June 6, 2020.

Figure 5 shows three curves; the red curve with dotted points represents the infection cases data (*I*) and another red curve represents the solution to infection cases data, circle marks on the green dotted curve represent recovery cases data (*R*) and another green curve represents the solution to recovery cases data, while the black curve represents the data and solution for the death cases (*D*). According to the SIR model estimation, in Figure 6 the rate of infection $\beta = 0{,}07$ interprets that 70 out of 1000 suspected individuals were infected, furthermore, the rate of recovery $\gamma = 0{,}019$ interprets that 19 out of 1000 infected individuals have recovered and the rate of death $\mu = 0{,}001$ interprets that 1 out of 1000 infected individuals have died due to Covid-19.

According to SIR analysis on data from April 4 to May 28, we obtained $R_0 = 3$ with MAPE 18%. The value of R_0 defines every 1 infected individual (*I*) could lead to 3 newly infected individuals. The value of $R_0 = 3 > 1$ indicates that the number of infected individuals would continue to increase until it reaches its equilibrium point. MAPE criteria with value $10\% \leq \text{MAPE} \leq 20\%$ states that the estimation is good and the results of the forecast predictions are acceptable (Gaspersz, 2014).

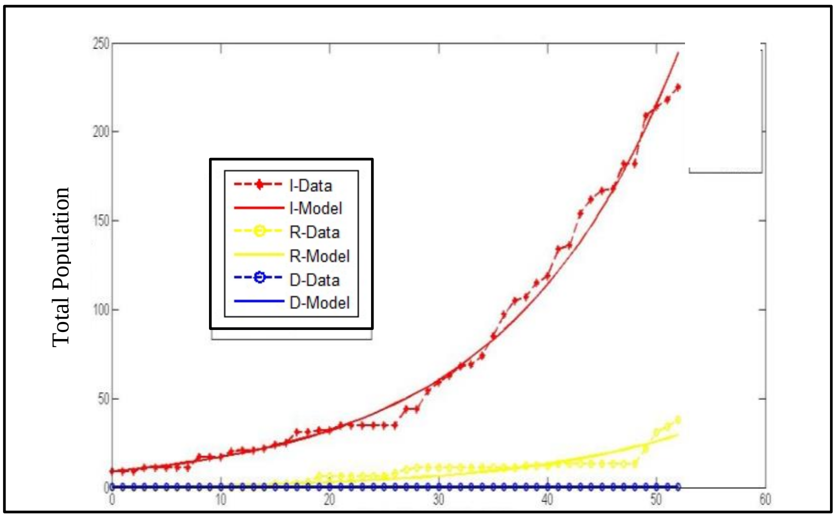

Figure 5. Estimation of Covid-19 Cases on March 19 - April 2020 in East Kalimantan Province

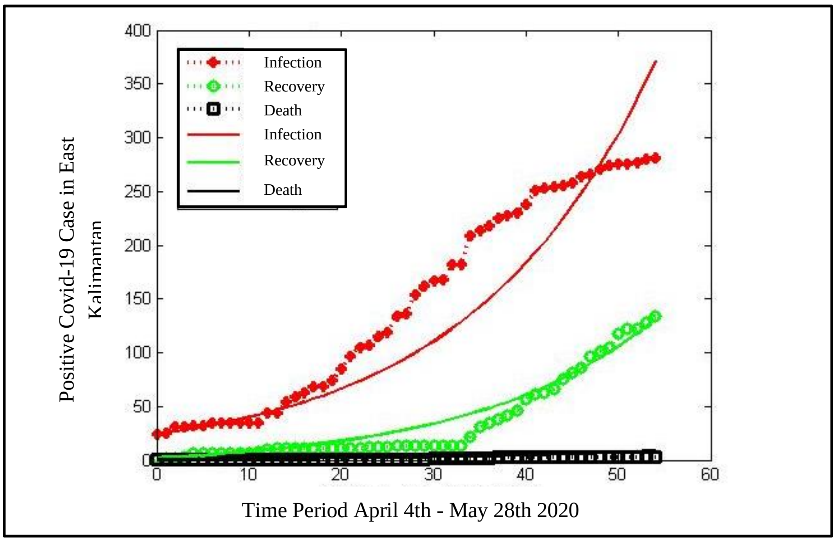

Figure 6. Estimation of Covid-19 Cases on April - May 2020 in East Kalimantan Province

In Figure 7, the rate of infection $\beta = 0{,}0707$ interprets that 71 out of 1000 suspected individuals were infected, the rate of recovery $\gamma = 0{,}0238$ interprets that 24 out of 1000 infected individuals have recovered, and the rate of death $\mu = 0{,}0013$ interprets that 1 out of 1000 infected individual have died due to Covid-19. We summarized the SIR estimation in Table 1.

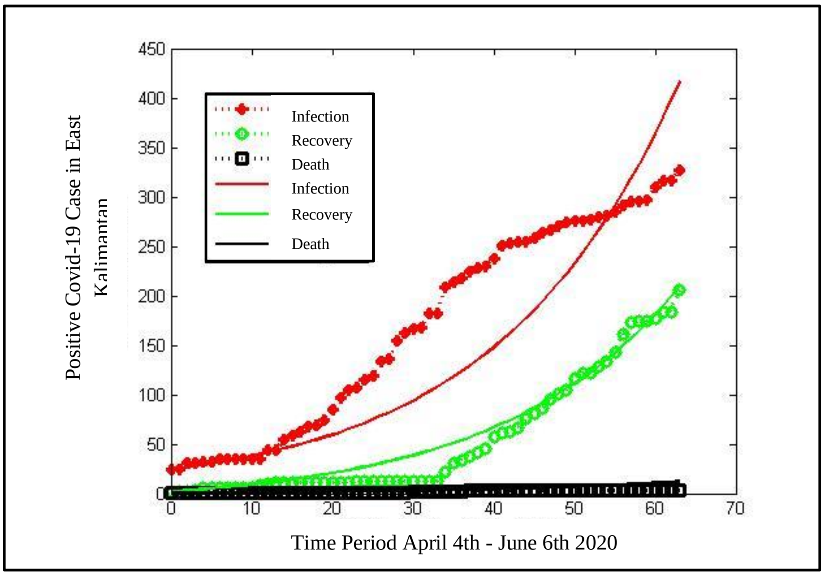

Figure 7. Estimation of Covid-19 Cases on April - June 2020 in East Kalimantan Province

Table 1. SIR Model Numerical Solution

Parameter	Measurement	Covid-19 Cases		
		March–April 2020	April–May 2020	April-June 2020
β	Rate of infection	0,0725	0,07	0,0707
γ	Rate of recovery	0,0079	0,019	0,0238
μ	Rate of death	0,001	0,001	0,0013
R_0	Reproduction number	8,0053	3	2,8179
R_t	Effective reproduction number (3)	8,004115	2,998	2,81
R_t	Time independent	1,0658	1,0397	1,03189
MAPE	Modeling Accuracy	0,0779	0,018	0,02341
V_{R_0}	Vaccination coverage of the value R_0	0,87	0,67	0,65
V_R	Vaccination coverage of the value R_t	0,87	0,66	0,64

We obtained R_0 = 8,0053 for confirmed data from March to April 2020, it shows that 1 infected individual could lead to 8 other individuals to be infected. Furthermore, if 8 individuals were infected, then 64 other individuals would be infected or $(R_0)^n$ as the rate of increase in cases. The implementation of government policies in East Kalimantan, such as suspending activities in schools, universities and certain places, resulted in a decline in reproduction number. This condition was shown in the estimated data from April to June 6, 2020. In the period from April to the end of May 2020, we obtained the reproduction number $R_0 = 3$ interpreting 1 infected individual could lead to 3 other individuals to be infected, if 3 individuals were infected, then 9 other individuals were susceptible to infection. For confirmed cases up to June 6, 2020, we obtained the reproduction number $R_0 = 2{,}8179$. The policies given by the East Kalimantan government affected the reproduction number of the spread of covid-19.

We calculated the vaccination coverage in Table 1 using equations (6) and (7). The simulation results for confirmed Covid-19 cases until June 6, 2020 produced vaccination coverage value V_{R_0} = 0,65 or 65% individuals need to be vaccinated (if the experts find a Covid-19 vaccine).

3.3. Estimation of Covid-19 Reproduction Number in East Kalimantan

According to equations (8) and (9) for Covid-19 data in Samarinda City, we obtained the rate of infection $\beta = 0{,}051$ interpreting that 51 out of 1000 suspected individuals were infected and the rate of recovery $\gamma = 0{,}038$ interprets that 38 out of 1000 infected individuals have recovered. The basic reproduction number $R_0 = 1{,}35$ $(R_0 > 1)$ with MAPE 2% interprets that the number of infected individuals would continue to increase until it reaches its equilibrium point.

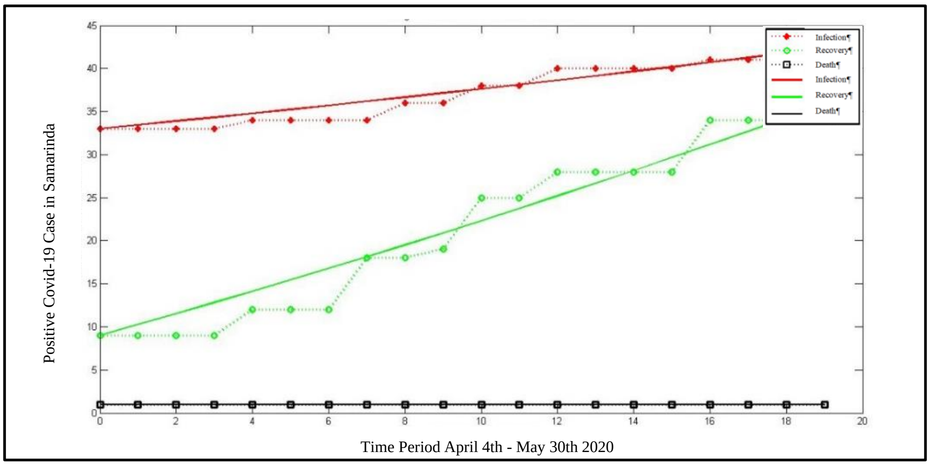

Figure 8. Estimation of Covid-19 Cases on March - May 30, 2020 in Samarinda City

3.4. Prediction of The Peak of Covid-19 Infection Cases in East Kalimantan and Samarinda City

The estimation of SIR model for confirmed Covid-19 cases from April to June 6, 2020 shows that the prediction of the peak of the infected cases (I) is on September 27 until October 11, 2020 with an estimated 14,000 infected individuals or 28% per 50,000 population in East Kalimantan Province. We show the prediction in Figure 9.

The prediction for Covid-19 cases in the next 2 weeks is shown in Figure 10, the rates of infection, recovery, and death are increasing. Further research for estimation and prediction results is continued with spatial mapping (Sifriyani *et al.*, 2018), (Sifriyani, Ruslan and Susanty, 2019) based on the number of infections in each district and city in East Kalimantan.

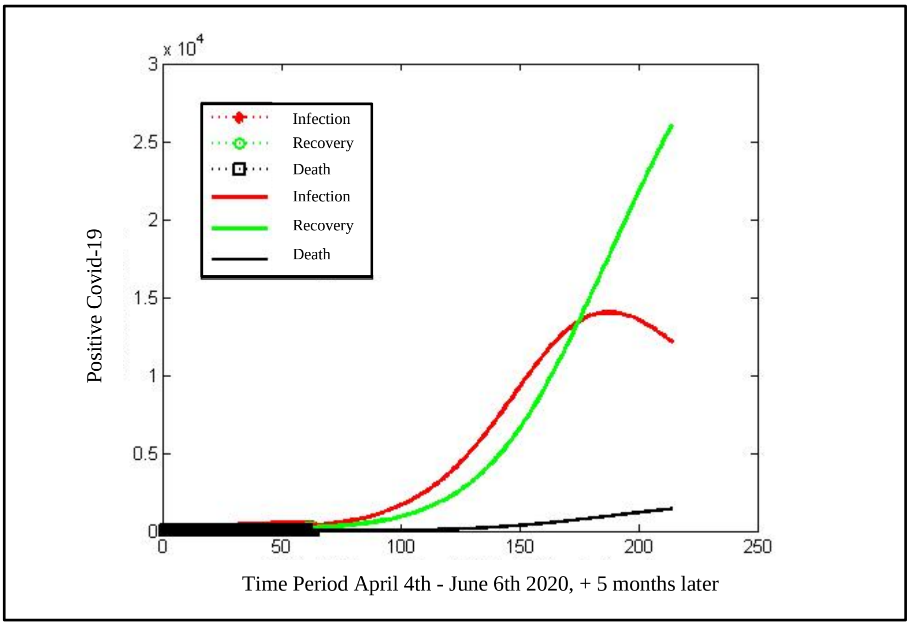

Figure 9. Prediction of The Peak of Infected Cases for Data April 4 - June 6, 2020 in East Kalimantan

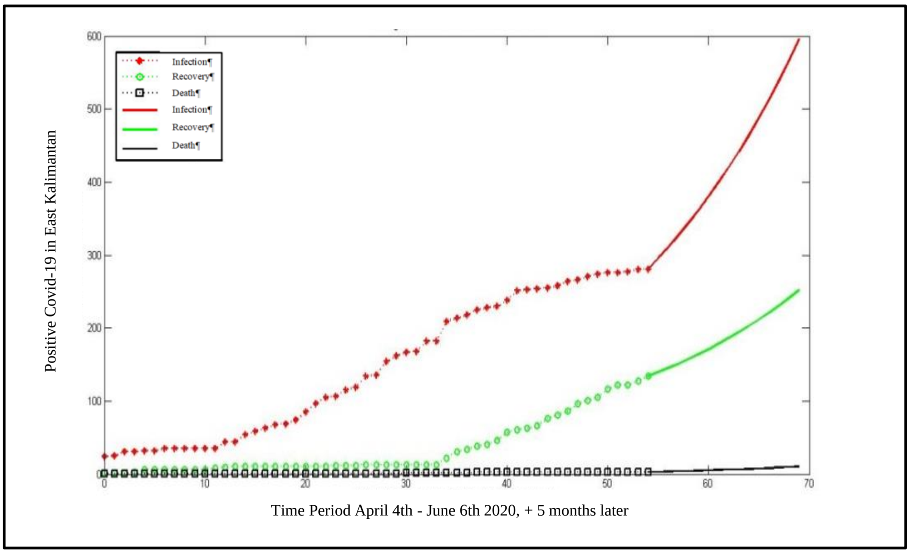

Figure 10. Prediction Curve in The Next 2 Weeks

4. CONCLUSION

SIR model produced the basic reproduction number $R_0 = 2{,}8179$ interpreting that 1 individual infected of Covid-19 in East Kalimantan could lead to 3 newly infected individuals. The rate of infection $\beta = 0{,}0707$ interprets that 71 out of 1000 suspected individuals were infected. Furthermore, the rate of recovery $\gamma = 0{,}0238$ interprets that 24 out of 1000 infected individuals have recovered.

The existence of a regional government policy for East Kalimantan Province had a good impact on the Covid-19 reproduction number. We obtained the reproduction number $R_0 = 8{,}0053$ according to cases from March to April 2020. The implementation of lockdowns in East Kalimantan decreased the reproduction number. This condition is shown in the estimated data from April to the end of May 2020, the reproduction number $R_0 = 3$ and for confirmed data up to June 6, 2020, $R_0 = 2{,}8179$. The policies given by the East Kalimantan government affected the reproduction number of the spread of Covid-19.

The peak date of infected cases (I) is predicted to occur on September 27 until October 11, 2020 with an estimated 14,000 infected individuals or 28% per 50,000 population in East Kalimantan Province.

The rate of infection $\beta = 0{,}051$ interprets that 51 out of 1000 suspected individuals were infected. Furthermore, the rate of recovery $\gamma = 0{,}038$ interprets that 38 out of 1000 infected individuals have recovered. The basic reproduction number $R_0 = 1{,}35$ with MAPE 2% interprets that every individual infected of Covid-19 in Samarinda City could lead to a newly infected individual and the number of infected individuals would continue to increase until it reaches its equilibrium point.

REFERENCES

Diekmann, O., Heesterbeek, H. & Britton, T. (2012). *Mathematical Tools for Understanding Infectious Disease Dynamics*. doi: 10.23943/princeton/9780691155395.001.0001.

Duffy, D. G. (2016). *Advanced Engineering Mathematics with Matlab*. *BMC Bioinformatics*.

Dinas Kesehatan Provinsi Kalimantan Timur. Accessed May 30, 2020. *Data Kasus Covid-19 yang Terkonfirmasi*. URL: https://dinkes.kaltimprov.go.id/.

Gaspersz, V. (2014). ISO 9001: 2000 and Continual Quality Improvement. Jakarta: *Gramedia Pustaka Utama*.

Gugus Tugas Percepatan Penanganan Covid -19. Accessed June 11, 2020. *Kasus Positif Covid-19 Melesat Naik 1.241, Total 34.316*. URL: https://covid19.go.id/p/berita/kasus-positif-Covid-19-melesat-naik-1241-total-34316.

Grohskopf, L. *et al.* (2011). Prevention and Control of Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices (ACIP), 2011. *American Journal of Transplantation*, 11(10). doi: 10.1111/j.1600-6143.2011.03793.x.

Hethcote, H. W. (2000). Mathematics of Infectious Diseases. *SIAM Review*, 42(4), pp. 599–653. doi: 10.1137/S0036144500371907.

Howell, R. D., Bates, D. M. & Watts, D. G. (1990). Nonlinear Regression Analysis & Its Application. *Journal of Marketing Research*, 27(1). doi: 10.2307/3172558.

Karim, M. A. *et al.* (2018). Solving a Parameter Estimation Problem of Goodwin Model With Fuzzy Initial Values. *Far East Journal of Mathematical Sciences (FJMS)*, 107(2). doi: 10.17654/ms107020321.

Karim, M. A. & Gunawan, A. Y. (2020). Parameter Estimations of Fuzzy Forced Duffing Equation: Numerical Performances by the Extended Runge-Kutta Method. *Abstract and Applied Analysis*, 2020, pp. 1–9. doi: 10.1155/2020/6179591.

Keviczky, L. *et al.* (2019). Introduction to MATLAB. *Advanced Textbooks in Control and Signal Processing*. doi: 10.1007/978-981-10-8321-1_1.

Sifriyani *et al.* (2018). Development of Nonparametric Geographically Weighted

Regression using Truncated Spline Approach. *Songklanakarin Journal of Science and Technology*, 40(4). doi: 10.14456/sjst-psu.2018.98.

Sifriyani, R. & Susanty, F. H. (2019). Evaluation of Forest Productivity and Governance on the Preservation of Tropical Rain Forests in Kalimantan using the NGWR-TS Nonparametric Geospatial Method', *EurAsian Journal of BioSciences*, 13(2).

Wallinga, J. (2004). Wallinga and Teunis Respond to "Real-Time Tracking of Infection Control Measures". *American Journal of Epidemiology*, 160(6). doi: 10.1093/aje/kwh257.

Yulida, Y. and Karim, M. A. (2020). Pemodelan Matematika Penyebaran COVID-19 di Provinsi Kalimantan Selatan. *Jurnal Binawakya*, 14(10).