Epidemic Modeling of COVID-19 in the ASEAN countries using a Genetic Partial Fitting Algorithm with the Presence of a Second Wave

Venansius R. Tjahjono[1], Hengky Kurniawan[1], Amirul Hakam[1], Endah R.M. Putri[1*], and Hadi Susanto[2,3]

[1] *Department of Mathematics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Indonesia*
[2] *Department of Mathematics, Khalifa University, United Arab Emirates*
[3] *Department of Mathematical Sciences, University of Essex, United Kingdom*
[*] *Corresponding author. E-mail: endahrmp@matematika.its.ac.id*

Received: Feb. 22, 2021; Accepted: Apr. 27, 2021

We present an SIRD epidemic modelling for COVID-19 outbreak in the ASEAN member countries. The occurrence of a second wave in the region adds complexity to the parameter estimation of the SIRD model. In this case, a standard genetic algorithm cannot fully capture the dynamic transmission of the pandemic. We therefore introduce a genetic partial fitting algorithm (GPFA) of seven-day intervals. We show that our method outperforms the standard algorithm with a significant reduction in the Root Mean Square Error (RMSE) value. We also extend our study to produce a real-time estimation of the effective reproduction number with a confidence interval to incorporate uncertainties in the model.

Keywords: Genetic algorithm, Epidemic model, Partial fitting, Second wave

http://dx.doi.org/10.6180/jase.202112_24(6).0011

1. Introduction

The global pandemic of novel coronavirus (COVID-19), starting in Wuhan city, China, at the beginning of December 2019, has affected countries all over the world. The most affected ones, in terms of positive cases, were the US, Spain, Italy, Germany, France, China, Iran, and UK, with the US currently having the highest total death toll [1]. There are several different symptoms of COVID-19, with the most common ones are fever, cough, and shortness of breath. Severe symptoms include respiratory and organ failures, which might end up in death due to the seriousness of failures and limited capacity of health facilities [2, 3].

The pandemic has been followed by dramatic policies such as lockdown and quarantines, which have been applied to reduce the spread of the epidemic. Globally, the number of infected people has reached more than one million in less than six months, and the positive cases are still growing daily. The rapid growth made this novel coronavirus outbreak immediately receive serious attention from researchers in different fields worldwide.

Mathematical and statistical modeling tools provide additional resources in facing the epidemic [4]. Emerging studies related to the outbreak have arisen with various models to predict the spread of the virus. Some of them are heuristic due to the limited infection data available, i.e., exponential [5] or logistic [6] model. However, the model parameters do not have an adequate physical interpretation and are not used for a long-term prediction. Epidemic compartmental models have been developed based on physical parameters, where the population is divided into smaller groups based on the infection status, such as susceptible (S), infected (I), recovered (R), and death (D). The mathematical model representing the population's dynamic is then called a SIRD model [7].

A comparative assessment of COVID-19 on the dynamic of the epidemic spreading in mainland China, Italy, and

France, has been done by Fanelli and Piazza [8]. It can be used to observe reported deaths that happen severely in the COVID-19 pandemic and other features such as the number of infected and recovery during the pandemic period. Fanelli and Piazza pointed out that the initial values of the reported and infected cases tend to be underestimated, so that they need to be floating and optimized.

In this paper, the SIRD model is revisited for the COVID-19 epidemic and its forecasting. The study will involve data set from a well-known regional group of nations, the Association of Southeast Asian Nations (ASEAN). Due to the uncertainty of the COVID-19 epidemic and its irregularity in some different countries, obtaining a more reliable understanding of the outbreak's spread using the SIRD model becomes a challenge [9]. Initial values of the SIRD model might be estimated using an artificial intelligence approach, i.e., a genetic algorithm. There are several advantages of the genetic algorithm to conventional optimization algorithms. Two of the most notable ones are its capacity to deal with complicated problems and its parallelism [10]. A careful fitting should be done in the COVID-19 SIRD model so that the government could commence a COVID-19 policy-making properly. Parameter estimation using a standard genetic algorithm can be performed precisely if the data set is commonly assumed to be under the SIRD curve.

However, the occurrence of a second wave makes the SIRD model may not fit the data precisely due to the presence of the resurgence (second wave) in the data [11]. Hence, the current modeling method of the SIRD curve produces a limitation, and the model's prediction could not represent the actual data [12]. The misfitting of the SIRD model can possibly lead policymakers to a wrong settlement. Xie et al. improved a Genetic Algorithm reconstruction on a time series model by partitioning the observation into independent piecewise intervals for accuracy boosting of the prediction [13]. Other works from Baragona et al. to reconstruct the Genetic Algorithm on time series model also produced satisfactory results [14], where they divided the auto-regressive model into N-independent parts (regimes) by setting up a certain threshold using the Genetic Algorithm. Yet, the existing modification of the Genetic Algorithm cannot be implemented directly to the SIRD model due to its independent properties.

According to the WHO, the incubation phase of COVID-19, which is the interval between susceptibility to the virus and symptom onset, is on average 5-6 days but can be as long as 14 days [15]. So, in the incubation period, the data from yesterday certainly perform an effect, which is dependent on today's data. Therefore, in this study, we introduce our novel approach, the genetic partial fitting algorithm (GPFA), on the COVID-19 data set fitting process. We divided our actual data into k overlapping sub-intervals to fit our model in the data set. The overlapping sub-intervals will produce a 4-tuple piecewise parameter of the SIRD model that is useful to measure the reproduction number.

There are two advantages to using GPFA as a fitting algorithm. Firstly, we can offer a more fitted result to the COVID-19 data set containing second wave appearances. Secondly, some measures related to the SIRD model, such as the case fatality rate and reproduction number, can be estimated in real-time.

This paper is organized as follows. In Sections 2 and 3, we discuss the governing SIRD model and parameter estimations for the model using a genetic algorithm. Data set description and some measures are provided in Section 4 and 5, respectively. Simulations, analysis, and further discussions of the epidemic in the aforementioned countries are presented in Section 6. Finally, we give the conclusion of our work in Section 7.

2. The SIRD Model

In this paper, the disease transmission is modeled by the following set of ordinary differential equations:

$$\begin{aligned} \frac{dS(t)}{dt} &= -rS(t)I(t), \\ \frac{dI(t)}{dt} &= rS(t)I(t) - (a+d)I(t), \\ \frac{dR(t)}{dt} &= aI(t), \\ \frac{dD(t)}{dt} &= dI(t), \end{aligned} \tag{1}$$

under an assumption that the population is grouped into four compartments given by susceptible individuals (S) which are healthy but not immune to the disease, infected individuals (I) that can infect others, recovered individuals (R) after being infected, and death individuals due to the viral infection (D). The system in Eq. 1 is complemented with initial conditions $[S(t_0), I(t_0), R(t_0), D(t_0)] = [S_0, I_0, R_0, D_0]$. An initial time t_0 is equivalent to the first day when the virus began to spread in the country. Parameters r, a, and d in the system represent transmission, recovery, and death rates, respectively. The parameters r, a, and d will be estimated using our proposed method, a genetic partial fitting algorithm (GPFA) based on the available COVID-19 data from [16]. The value of S_0 is also optimized since there is no information about how many healthy people could be infected initially. We extend our analysis of the SIRD model into case fatality rate and reproduction number in Section 5 and 6.

3. Genetic Partial Fitting Algorithm as Parameter Estimation Techniques

A genetic algorithm is a meta-heuristic searching that is inspired by Charles Darwin's ideas of natural evolution. This algorithm exhibits rules of natural selection where the most eligible individuals are chosen for breeding the offspring of the next generation. Genetic algorithms have four basic steps: an initial population generation, a selection process, a crossover mechanism, and a mutation scheme [17], as depicted in Fig. 1.

The process begins with a generated set of values, called an initial population with N individuals (or chromosomes), $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N)$ where

$$\mathbf{X_i} = (S_0, r, a, d)_i, \quad i = 1, 2, ...N. \tag{2}$$

Each individual has four parameters (S_0, r, a, d) that can be considered as one strain of chromosome. We generated 250 $(N = 250)$ types of chromosomes for each country as our initial population. Before entering the selection process, we should determine our fitness value. We used the Root Mean Square Error (RMSE) measurement to find the fitness value in the genetic algorithm. Henceforth, our objective is to minimize the RMSE by systematically choosing one strain of chromosomes from within the set of the initial population.

The RMSE is calculated using Eq. 3,

$$\mathbf{RMSE}(\mathbf{A}, \mathbf{P}) = \min_{\{S_0, r, a, d\}} \sqrt{\frac{1}{n} \sum_{i}^{n} (A_i - P_i)^2} \tag{3}$$

where n is the number of observed data, $A_i \in \{I(t), R(t), D(t)\}$ is the observed data, and $P_i \in \{\hat{I}(t), \hat{R}(t), \hat{D}(t)\}$ is the predicted (simulated) data using the SIRD model. To make our simulation easier, we prefer to maximize $\mathbf{RMSE}(\mathbf{A}, \mathbf{P})^{-1}$ instead of minimizing $\mathbf{RMSE}(\mathbf{A}, \mathbf{P})$. In the second step, the selection process, we used the roulette system to determine the chromosome that corresponds with the smallest fitness value. The selection process by roulette algorithm is described in Table 1. The roulette system will provide a high probability of selecting the chromosome that has the smallest fitness value. Thus, the roulette system will decide the fittest chromosome for the next generation.

In the crossover mechanism, we design a weighted arithmetic summation to create the offspring of the next generation $\mathbf{X}_{child}$ as follows,

$$\mathbf{X}_{child} = (1 - \alpha)\mathbf{X}_{parent1} + (\alpha)\mathbf{X}_{parent2}$$

where α is a constant crossover rate, and $\mathbf{X}_{parent1}$ and $\mathbf{X}_{parent2}$ are chosen randomly by conditioning α on $U[0, 1]$.

Table 1. Roulette Algorithm.

No	Roulette Algorithm
1.	Calculate $\mathbf{RMSE_i}(\mathbf{A}, \mathbf{P})^{-1}$ for every $\mathbf{X_i} \in \mathbf{X}$.
2.	Define $w_i = \dfrac{\mathbf{RMSE_i}(\mathbf{A}, \mathbf{P})^{-1}}{\sum_i \mathbf{RMSE_i}(\mathbf{A}, \mathbf{P})^{-1}}$, resulting $0 < w_i < 1$.
3.	Create roulette's partition, π_i, based on w_i.
4.	Generate $s \sim U[0, 1]$.
5.	If s is located in π_i, for certain i, save the value of X_i.
6.	Repeat step 4 and 5 up to N times.

In the last step, we design the mutation scheme $(\mathbf{X}_i)_{mutated}$ as an addition or subtraction of the value in a chromosome $\mathbf{X}_i$ in Eq. 2 with a normal distributed random noise $\mathbf{z}$. This scheme can be described in a mathematical formula as follows,

$$(\mathbf{X_i})_{mutated} = \mathbf{X_i} + \mathbf{z} \tag{4}$$

where $\mathbf{z} \sim N(0, \overline{\mathbf{X}})$. The chromosomes will be mutated fortuitously (see Eq. 4) based on a mutation rate, θ, depending on $U[0, 1]$. The genetic algorithm terminates if the maximum iteration is reached or the fitness value is already converged.

In this paper, we propose a modification of the standard genetic algorithm into a genetic partial fitting algorithm (GPFA) that employs piecewise fitting every seven days. Thus, our objective function evolves to a modified RMSE based on the one in Eq. 3 and is defined as

$$\mathbf{RMSE_m}(\mathbf{A}, \mathbf{P}) = \min_{\{S_0(k), r(k), a(k), d(k)\}} \sqrt{\frac{1}{7} \sum_{i=k}^{k+6} (A_i(k) - P_i(k))^2} \tag{5}$$

where $k = 0, 1, 2, \ldots, N - 6$ denotes the k-th piecewise data set containing seven reported COVID-19 data, and N denotes the total entry of data. Seven points of data (i.e., seven days) represent the average incubation period according to World Health Organization (WHO) [1].

4. Some Measures on Epidemic Model

We use two measures to describe the severity of the epidemic: case fatality rate (CFR) and effective reproduction number (). Each of the measures is explained in the following. Case fatality rate (CFR) is obtained by dividing the last data of death individual with a total infected individual. The value is calculated as [18]

$$CFR = \frac{1}{c} \frac{D(n)}{I(n)}, \tag{6}$$

where n and c indicate the index for the latest data and the correction factor. We chose the value of c in Eq. 6 under the

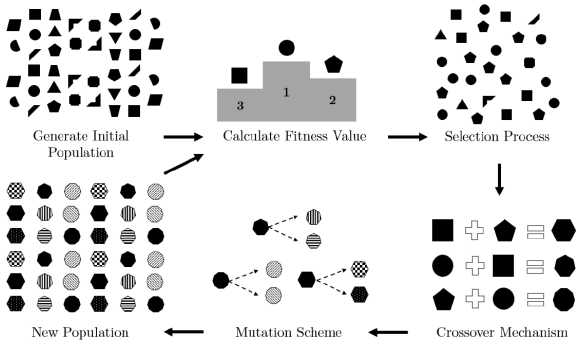


Fig. 1. General Scheme of Genetic Algorithm.

assumption of under-reported cases are minimized. Thus, the value of c is 1.

The time-varying effective reproduction number is a measure of the outbreak's transmission describing the continuation of the transmission after control Non-Pharmaceutical Interventions (NPI) [19]. A quantitative evaluation of efficacy of the measure by using , is applied by controlling and reducing to less than one. Based on [20] and the SIRD model in Eq. 1, where a and d refer to the recovery rate and death rate, respectively. The rate of infection can be obtained as follows,

$$\frac{dI(t)}{dt} = rS(t)I(t) - (a+d)I(t) \tag{7}$$

$$\frac{dI(t)}{dt} = (a+d)(-1)I(t), \tag{8}$$

where

$$\mathcal{R}_t = \frac{S(t)}{\mathcal{N}}\mathcal{R}_0, \tag{9}$$

and

$$\mathcal{R}_0 = S_0\left(\frac{r}{a+d}\right). \tag{10}$$

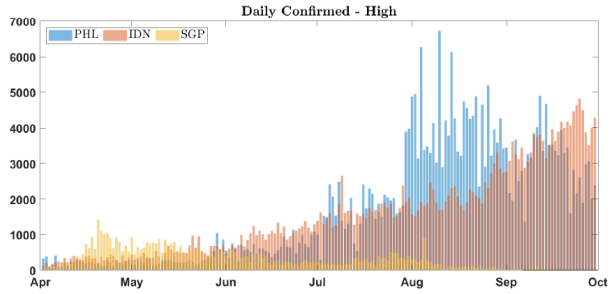
$\mathcal{R}_0$ is a basic reproduction number of the outbreak. It is defined as the average number of increasing cases caused by an infected individual during their infectious period in a susceptible and uninfected population. It should be noted that $\mathcal{R}_0$ is estimated before the control intervention while is estimated after. Comparing the two measures, $\mathcal{R}_0$ and , the effectiveness of control intervention is determined [21].

5. ASEAN COVID-19 Data Set

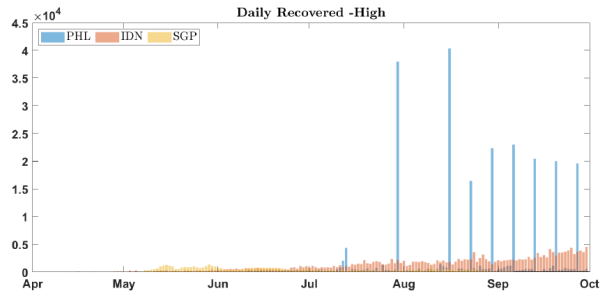
The data set used in this paper is obtained from COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [16], which is considered as a leading data source [22], from 1 April 2020 to 30 September 2020. The data set consists of daily confirmed cases, recovered cases, and death cases. We collect the data from ten countries of the ASEAN (The Association of Southeast Asian Nations) minus Timor Leste.

The countries are classified as high, mid, and low categories based on the total number of confirmed cases on 30 September 2020 [16]. The high-cases category is the countries with the three highest number of confirmed cases. The next four countries are mid-cases, while the last three countries are low-cases. The daily confirmed cases and the total number of death are depicted in Figs. 2a, 3a, 4a, and Fig. 5 for all three clusters respectively. In high-cases clusters (see Fig. 2), both the Philippines and Indonesia, show monotone increasing numbers of daily confirmed cases. The first waves of infection in both countries seem not yet finished, although after a quite long period. Singapore is one of the countries that has successfully dealt with the outbreak and has passed its first wave of infection. The number of cases in Singapore has increased rapidly from April to June due to the massive tests conducted by Singapore. In July and afterward, Singapore succeeded in suppressing the increase of daily confirmed cases, and accordingly, the daily recovered cases follow the trend.

Myanmar, Malaysia, Thailand, and Vietnam, are clustered as mid cases. In Fig. 3, confirmed cases in Myanmar increased significantly in April and recorded the first recovered case since impacted by Coronavirus on 8 April. Then, Myanmar survived nearly one month without any local transmissions from mid-July to mid-August. Even with that achievement, Myanmar had to suffer the second wave that started on 16 August, which causes dramatic increases in the number of daily confirmed cases in Myanmar. In Malaysia, the Coronavirus transmission grew insignificantly, thanks to the Movement Control Order (MCO) extension by the government on 10 April [23]. Malaysia was able to manage the spread of Coronavirus effectively. However, at the end of September 2020, Malaysia is facing the new challenge of the incoming second wave of infection without a significant increase in Coronavirus infection. In Thailand, the government suspended all commercial inter-

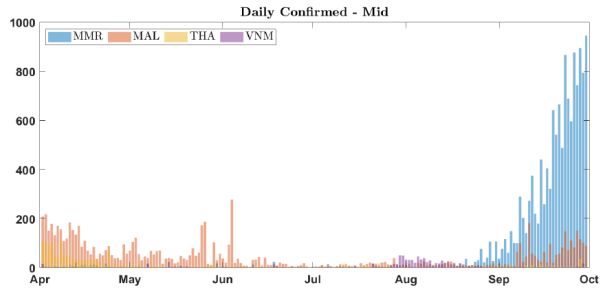


(a) Daily confirmed-high

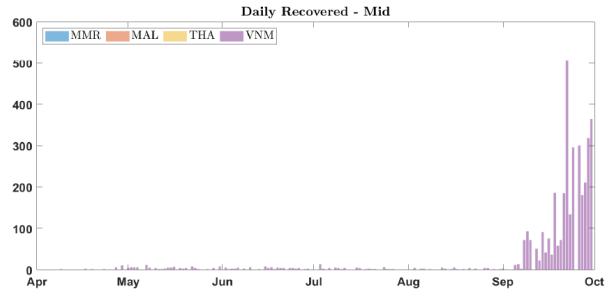


(b) Daily recovered-high

Fig. 2. High-Cases Data Set
PHL = Philippines, IDN = Indonesia, SGP = Singapore.

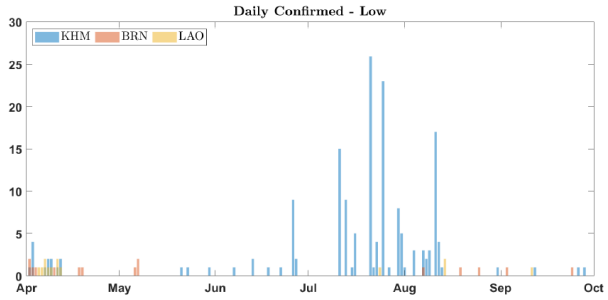


(a) Daily confirmed-mid

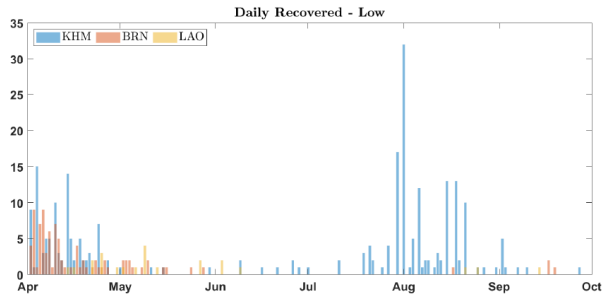


(b) Daily recovered-mid

Fig. 3. Mid-Cases Data Set
MMR = Myanmar, MAL = Malaysia, THA = Thailand, VNM = Vietnam.



(a) Daily confirmed-low



(b) Daily recovered-low

Fig. 4. Low-Cases Data Set
KHM = Cambodia, BRN = Brunei Darussalam, LAO = Laos.

national flights and implemented lockdown measures from 4 April. The decision successfully diminished the increase in infection in Thailand. Even in mid-May, the infection number had fallen close enough to zero. But recognizing a sharp rise in infection at the end of September, Malaysia should prepare for a second wave possibility in a short time. In Vietnam, the infection rate from April to June is very low. But then, at the end of July, Vietnam was impacted by the second wave of infection. It happened due to a new Coronavirus strain that led to high infection rates.

Next, Cambodia, Brunei Darussalam, and Laos clustered as low-cases (see Fig. 4). Despite limited health infrastructure, Coronavirus in Cambodia is most certainly considered as undercount [24]. The keys to pandemic management in Cambodia lie in strong surveillance mechanisms, including border controls, real-time databases and risk assessments, rapid response teams, massive epidemic training, enhancing laboratory capacity, and effective communications [25]. Brunei Darussalam has been positioned as a country that successfully managed the spread of the Coronavirus together with Singapore and Taiwan, despite its lack of experience in handling such a pandemic. While taking advantage of its wealth, a high human development index, and good health facilities, the country has faced the possibility of viral transmission. Consequently, the country

has sustainable programs of pandemic management: huge funding support, massive surveillance including border control, tracing and PCR testing, and transparent and responsive public communication on a daily basis [26]. Like Cambodia and Brunei, Laos also has the typical policies in handling the Coronavirus spread so that the country less suffered from the disease [27].

The cumulative number of death caused by the Coronavirus for the three countries are shown by Fig. 5. Only Singapore has no increase in the cumulative number of death for the high-cases cluster. Unlike Malaysia, Thailand, and Vietnam in the mid-cases cluster, Myanmar faced an increasing number of deaths as the country is still unable to control the spread of the disease and lack effective actions implemented. The last cluster, low-cases, consists of countries with almost countless and even zero death reported. The countries are Cambodia with zero death reported, Brunei Darussalam with a low constant number of death, and last Laos with very low death cases.

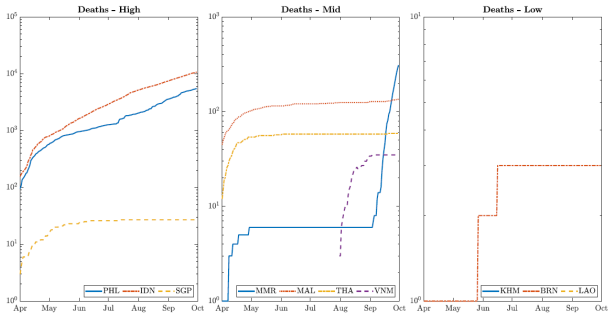


Fig. 5. ASEAN COVID-19 Cumulative Death Cases

A measure to determine the hazard level of the pandemic based on the death class of an individual is the case fatality rate (CFR). Table 2 shows the CFR of all countries observed in this paper. The countries considered to be the most affected by Coronavirus, in term of the number of infected individuals, Indonesia and the Philippines, have the CFR less than 2%-4% by September, 30th, 2020. This value is much less than the early period which was around 6% as reported in [28]. Myanmar has a higher CFR value and becomes one of the most affected by the outbreak. Cambodia and Laos have zero death since an early stage of the pandemic, so that the CFR is also zero. Due to the high number of infections but low level of death, Singapore has a very low value of CFR by September, 30th, 2020. Therefore, Singapore has been considered a successful country for controlling the outbreak. Other ASEAN countries listed in Table 2 have CFR around 1%-3% and the countries are considered to have a low level of CFR.

Table 2. ASEAN Countries COVID-19 Prevention Schemes
FW = First Wave, SW = Second Wave

Nation	Status	CFR at 30/09/2020
Philippines	FW Ongoing [29]	0.02125
Indonesia	FW Ongoing [30]	0.04759
Singapore	FW Passed [31]	0.00047
Myanmar	FW Ongoing [32]	0.07626
Malaysia	SW Ongoing [33]	0.01346
Thailand	SW Possibility [34]	0.01716
Vietnam	SW Passed [35]	0.03349
Cambodia	SW Passed [36]	0
Brunei	FW Passed [37]	0.02069
Laos	FW Passed [38]	0

6. Simulation Results and Discussion

Data set of each country have been presented and elaborated in the previous section. In this section, based on the data set observed, we present the dynamic of the outbreak spread using our novel approach, general partial fitting algorithm (GPFA), especially dealing with the presence of second wave phenomena.

6.1. Genetic Partial Fitting Algorithm (GPFA) on ASEAN COVID-19 Data Set

The genetic partial fitting algorithm (GPFA) is implemented to the data set following the steps presented in Section 3. The COVID-19 data set is fitted every seven overlapped days with daily increments. To manage the overlapped fit results, calculations of the mean of the data are applied to the same timestamp. Resulting, the GPFA can follow the trend of COVID-19's data points. The GPFA can well-detect turning points on the data set fluctuation using the partial fitting. The ability of GPFA in dealing with the fluctuations is considered as an advantage of the algorithm in fitting the second wave phenomena. In our finding, the genetic partial fitting algorithm (GPFA) shows excellent simulation outcomes for both, the first wave and second wave phenomena.

Fig. 6 provides a visualization concerning the GPFA on the two most COVID-19 cases in ASEAN countries, Indonesia and the Philippines, according to the Center for Strategic and International Studies (CSIS), [39]. The dot-dashed lines can follow the turning points perfectly compared to the solid line representing the standard genetic algorithm (GA).

Complete visualizations for each country are provided in Appendix A. We present the Root Mean Square Error (RMSE) of the GPFA in Table 3 to analyze and compare

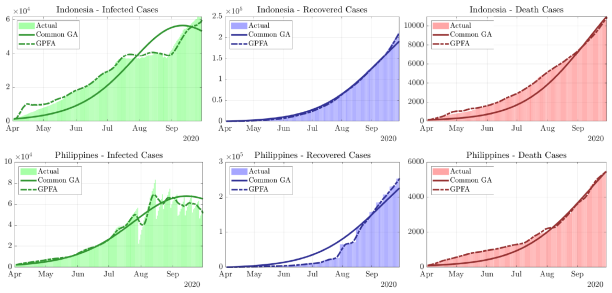


Fig. 6. The GPFA vs Common GA on Indonesia and Philippines COVID-19 Dataset

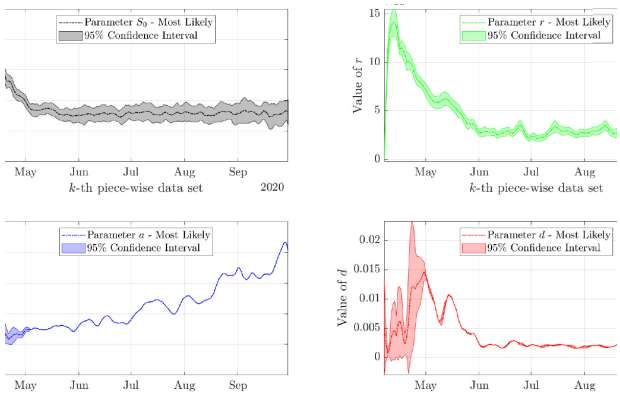


Fig. 7. Parameter Plot on Indonesia Data Set

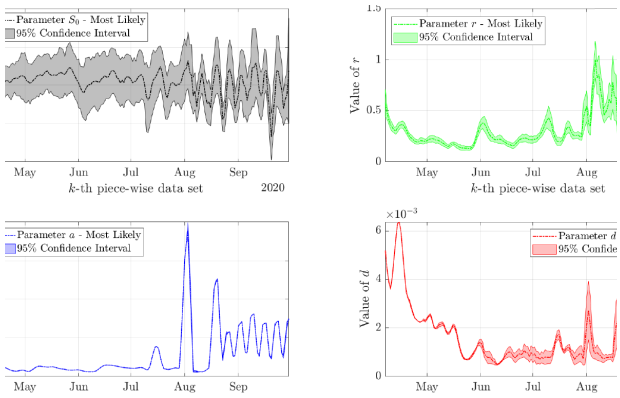


Fig. 8. Parameter Plot on Philippines Data Set

the significance of GPFA and standard GA. Note, a higher RMSE in value does not always mean that our algorithm worked poorly. However, it can also explain how big the value in the data set we chose. On the other hand, smaller RMSE implies that the fitting process has efficiently operated. Table 3 shows that the GPFA can significantly reduce the RMSE for more than 90%. The piecewise approach in GPFA can be more beneficial if the data set is more volatile and complex and is set up for short-term forecasting in each piecewise, for example, only for a 7-day piecewise or less. A higher accuracy will be obtained as a shorter period is used.

6.2. 4-Tuple Parameter Series Plot

The capability of our GPFA to capture the movement of COVID-19 reported cases has been shown in the previous section. Better fitting results using GPFA compared to the standard GA are depicted in Fig. 6. The figure shows the fluctuation of COVID-19 reported cases from the end of April/July 2020 to September 2020. In this section, the parameters, namely a 4-tuple parameter, which correspond to GPFA for each k-th piecewise, are observed (see Eq. 5). The value of the 4-tuple parameter in each piecewise, $(S_0(k), r(k), a(k), d(k))$ will be used to estimate the value of basic reproduction number in the next section.

The plots of parameters are represented by two countries, Indonesia and the Philippines, and are depicted in Figs. 7 and 8. Complete results are presented in Appendix B. In every piecewise data, the parameter S_0 (i.e., initial susceptible population) provides a wider confidence interval than the others. It illustrates the sensitiveness of S_0 amongst the other parameters. A small change in the data set can significantly change the estimated value of S_0. Other parameters (r,a,d) also move according to the fluctuation of the ASEAN COVID-19 data set.

6.3. Real Time Estimation of the Effective Reproduction Number

In this part, we performed the GPFA method to investigate viral transmission in the ASEAN data set. Based on the 4-tuple parameter obtained using the GPFA method, those are transmission rate $r(k)$, recovery rate $a(k)$, and death rate $d(k)$, are interpreted the previous section. The value of effective reproduction number $\mathcal{R}_t$, as the measure of viral transmission, is determined using Eq. 9. The value of parameters $(S_0(k), r(k), a(k), d(k))$ are used to estimate the value of basic reproduction number as described in Eq. 10 of k-th piecewise,

$$(k) = \frac{S(t)}{N}\mathcal{R}_0(k) \tag{11}$$

$$\mathcal{R}_0(k) = S_0(k)\left(\frac{r(k)}{a(k)+d(k)}\right). \tag{12}$$

$S_0(k)$ is the value of an estimated susceptible of the k-th interval, which is updated by a one-day shifting of the next seven-day overlapping interval. The transmission rate $r(k)$, recovery rate $a(k)$, and death rate $d(k)$, are interpreted the rates for each k-th piecewise interval. The rates change over

Table 3. Comparison of RMSE value between Common Genetic Algorithm and GPFA

Abbreviation	Nation	RMSE	$RMSE_m$	Reduction (%)
PHL	Philippines	12137.88	2794.30	76.98
IDN	Indonesia	4795.61	1436.56	70.04
SGP	Singapore	4803.15	93.25	98.06
MMR	Myanmar	636.94	23.35	96.33
MAL	Malaysia	269.57	20.89	92.25
THA	Thailand	120.01	9.41	92.16
VNM	Vietnam	130.39	5.02	96.15
KHM	Cambodia	52.89	2.19	95.85
BRN	Brunei	1.20	0.09	92.30
LAO	Laos	2.09	0.06	97.27

time and need to be readjusted in each k-th overlapping interval.

We visualized our simulations in graphic forms in Fig. 9 and 10. Complete simulation results for other ASEAN countries can be seen in Appendix C. The red-colored bar charts in both, Fig. 9 and 10, represent the reported active (infected) cases. Simultaneously, the black-colored line represents the estimated value of $\mathcal{R}_t$. In our analysis, the movement of $\mathcal{R}_t$ follows the change in the reported active cases. Resulting, we can produce a real-time estimation of $\mathcal{R}_t$ anytime when the new data is already recorded in the system. The spiking value in the reported cases also makes the $\mathcal{R}_t$ becomes higher and produces a wider gap in the confidence interval. The confidence interval (CI) expresses the uncertainty in any estimation results and describes an interval of values in which we can be reasonably convinced that the true value prevails. So, wider CI implies less knowledge about the future COVID-19 occurrence, but it can be used as a rough guideline.

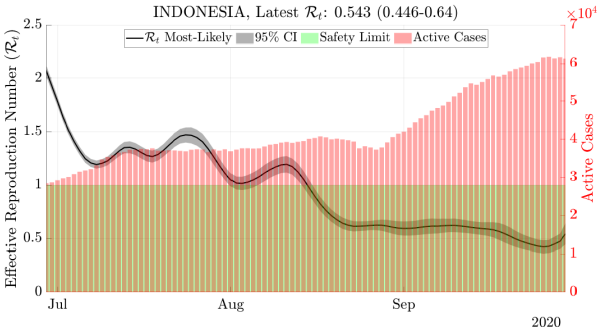


Fig. 9. The Estimated $\mathcal{R}_t$ Value on Indonesia Data Set

Based on our results, we have found that the changes mentioned above at active cases trigger a significant spike in the value of $\mathcal{R}_t$. For example, in Fig. 10, there were major value decrements in active cases at the end of July. Resulting, the value of $\mathcal{R}_t$ dropped significantly from around 2 into 1. This finding proposes a relationship between the percentage of increase in the infected case and the $\mathcal{R}_t$ value. Also, we infer that $\mathcal{R}_t$ is sensitive to the change of infected cases. Thus, it can be useful for identifying the severity transmission of COVID-19 when new cases are reported.

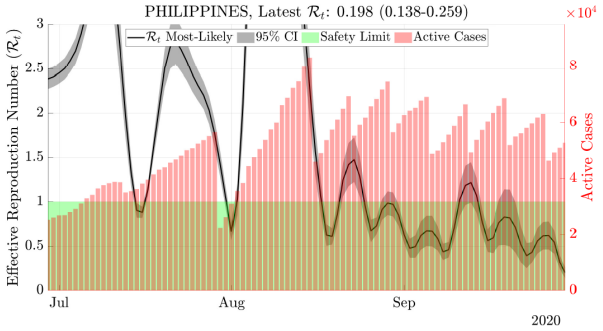


Fig. 10. The Estimated $\mathcal{R}_t$ Value on Philippines Data Set

Table 4. The GPFA Estimation Result of $\mathcal{R}_t$

Nation	Estimated $\mathcal{R}_t$ at 30 September 2020	95% CI
Philippines	0.198	0.138-0.259
Indonesia	0.553	0.446-0.640
Singapore	0.249	0.224-0.275
Myanmar	3.217	3.166-3.268
Malaysia	1.894	1.849-1.938
Thailand	2.625	1.694-1.721
Vietnam	1.299	1.290-1.308
Cambodia	1.774	1.760-1.787
Brunei	0.364	0.125-0.603
Laos	1.846	1.823-1.869

In addition, it has been found that estimating based on the SIRD model might not always work out. As shown in Fig. 29 for Brunei, the values of increase sharply in some

time steps, although the number of infected, recovered, and deceased individuals are low and the growth is nearly zero. As the spread is practically slowing down, the value of should be low. However, based on Eq. 9 obtained by the SIRD model, the value of may increase sharply with relatively low growth. Therefore, we may consider that the formula based on the SIRD model has failed in measuring the spread of the outbreak.

7. Conclusion

This study provided the value of the case fatality rate (CFR) to describe the severity of the pandemic based on the number of deaths, for some ASEAN countries. Indonesia and the Philippines, are still considered to be the most affected countries by the COVID-19 spread and the highest CFR in ASEAN. Singapore, which has the lowest CFR with a high number of infections, showing its success in handling the outbreak. We have employed the parameter estimation of the SIRD differential equation using the genetic partial fitting algorithm (GPFA) based on the data.

We presented the performance of the GPFA that outperforms the standard genetic algorithm. We also show that our method can be implemented on the COVID-19 data set that contains second wave phenomena. It produces smaller RMSE values than the standard genetic algorithm (GA), with a reduction of approximately 90%. Using 4-tuple parameters obtained from GPFA, we have observed the behavior of the effective reproduction number, , that illustrates the outbreak's transmission of COVID-19. We found that the spiking value in the data increases the value of and provides a wider confidence interval. It is a loose interpretation but is practical to prepare for the possibility of the incoming viral transmission.

A. Visualization of GPFA on COVID-19 ASEAN Data Set

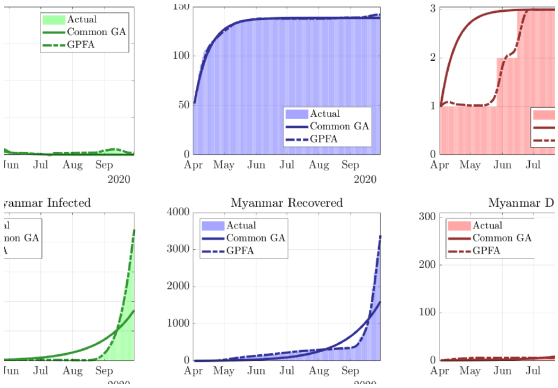


Fig. 11. The GPFA vs Common GA on Brunei and Myanmar COVID-19 Data Set

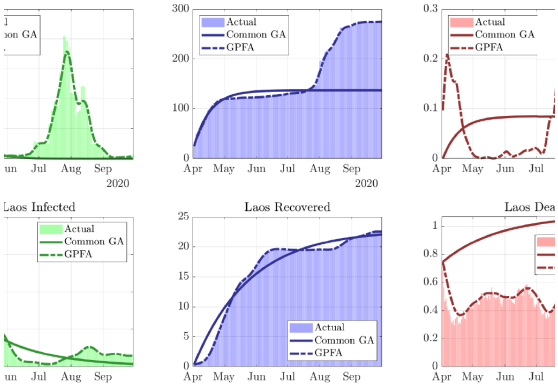


Fig. 12. The GPFA vs Common GA on Cambodia and Laos COVID-19 Data Set

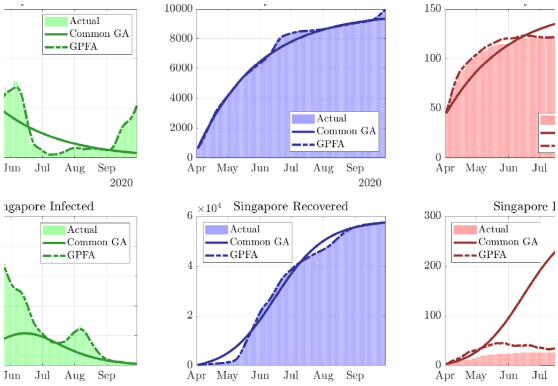


Fig. 13. The GPFA vs Common GA on Malaysia and Singapore COVID-19 Data Set

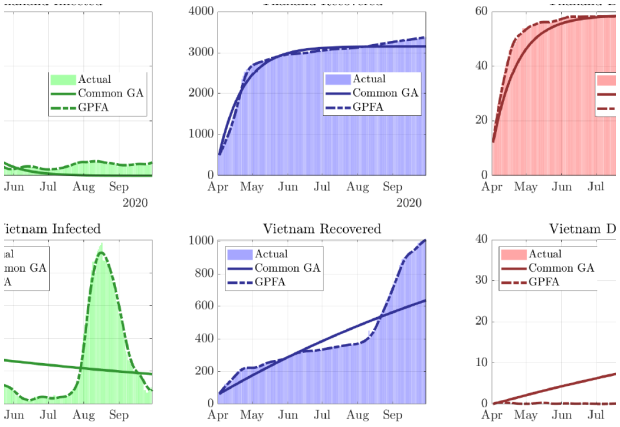


Fig. 14. The GPFA vs Common GA on Thailand and Vietnam COVID-19 Data Set

B. Parameter Plot of GPFA on COVID-19 ASEAN Data Set

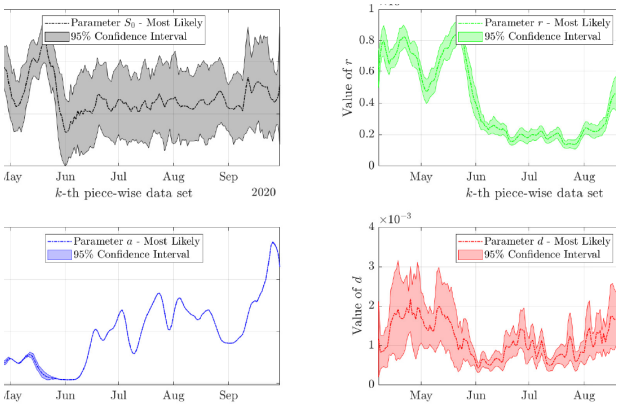


Fig. 15. Parameter Plot on Singapore Data Set

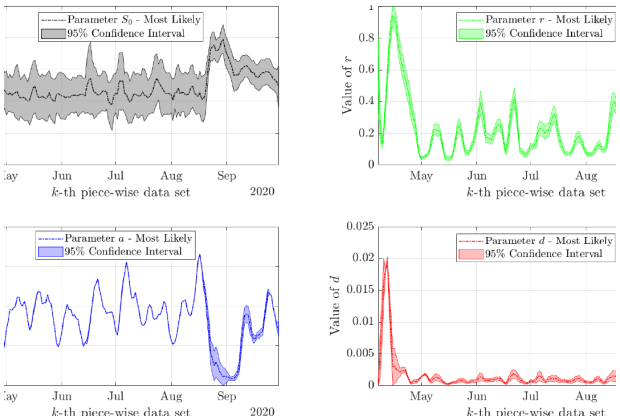


Fig. 16. Parameter Plot on Myanmar Data Set

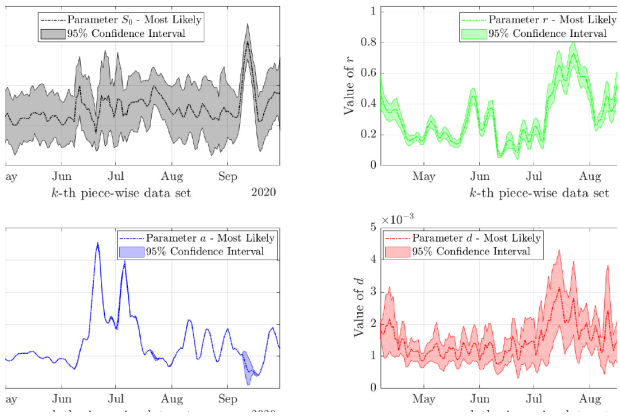


Fig. 17. Parameter Plot on Malaysia Data Set

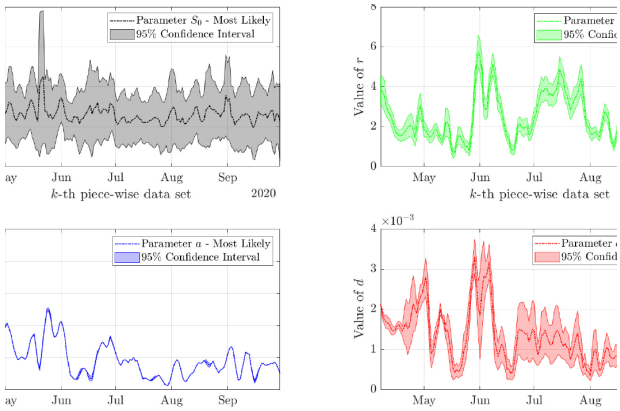


Fig. 18. Parameter Plot on Thailand Data Set

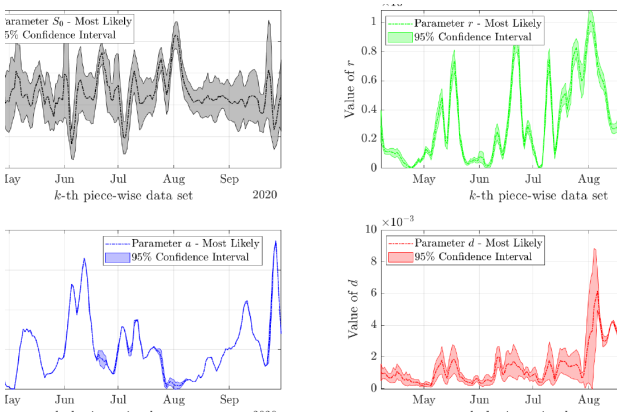


Fig. 19. Parameter Plot on Vietnam Data Set

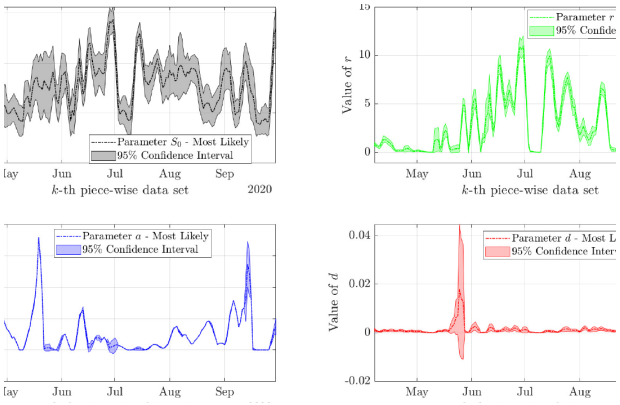


Fig. 20. Parameter Plot on Cambodia Data Set

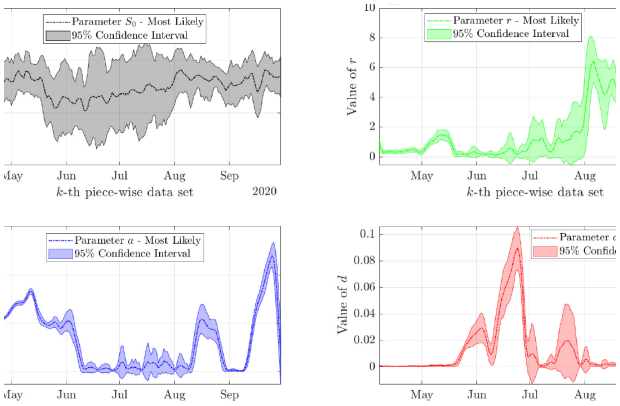


Fig. 21. Parameter Plot on Brunei Data Set

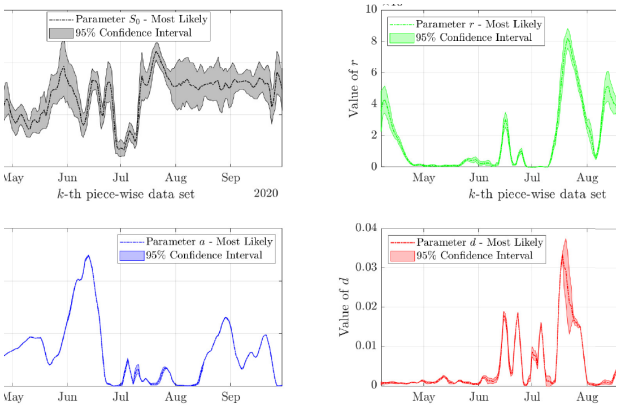


Fig. 22. Parameter Plot on Laos Data Set

C. The Estimated Value of the Effective Reproduction Number on COVID-19 ASEAN Data Set

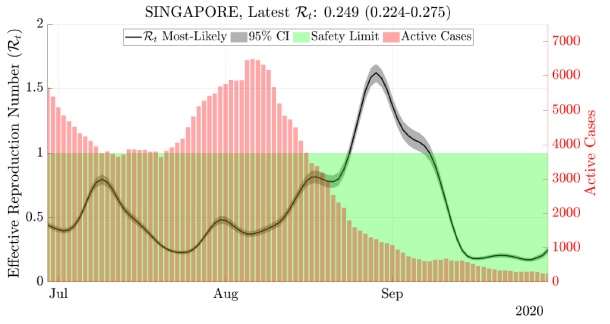


Fig. 23. The Estimated Value of $\mathcal{R}_t$ on Singapore Data Set

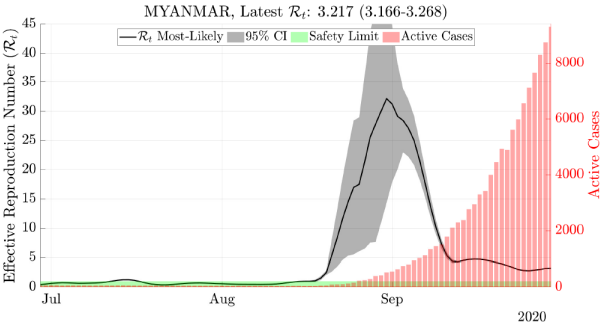


Fig. 24. The Estimated Value of $\mathcal{R}_t$ on Myanmar Data Set

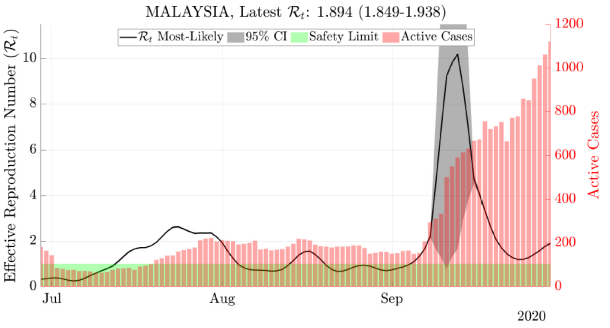


Fig. 25. The Estimated Value of $\mathcal{R}_t$ on Malaysia Data Set

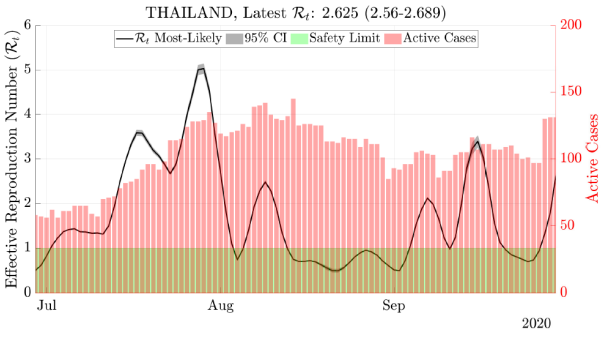


Fig. 26. The Estimated Value of $\mathcal{R}_t$ on Thailand Data Set

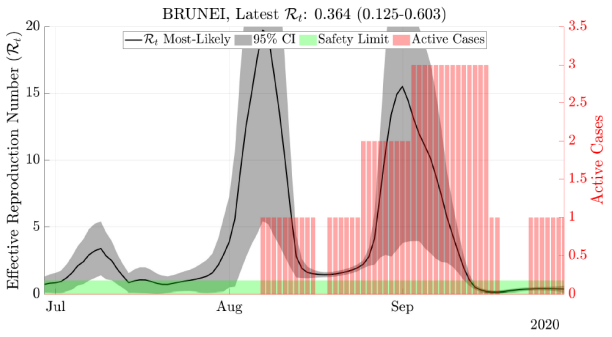


Fig. 29. The Estimated Value of $\mathcal{R}_t$ on Brunei Data Set

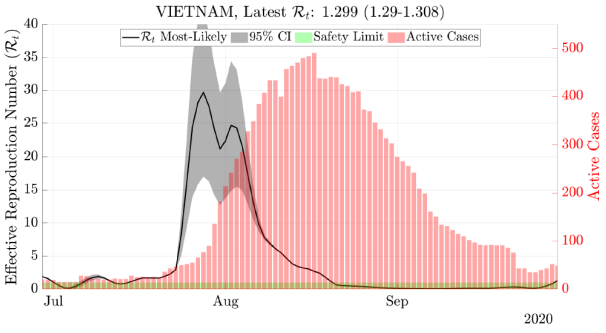


Fig. 27. The Estimated Value of $\mathcal{R}_t$ on Vietnam Data Set

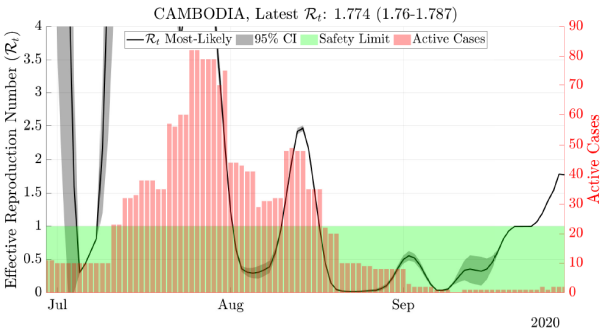


Fig. 28. The Estimated Value of $\mathcal{R}_t$ on Cambodia Data Set

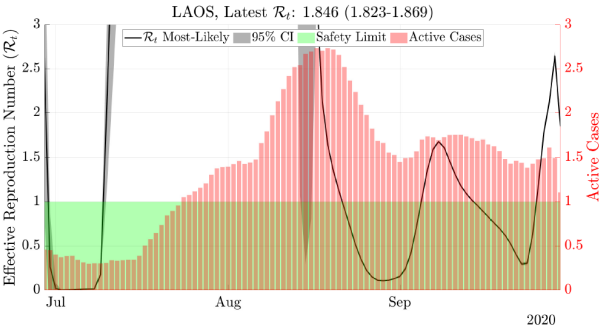


Fig. 30. The Estimated Value of $\mathcal{R}_t$ on Laos Data Set

References

[1] World Health Organization Regional Office for Europe. About the virus, 2020. http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov.

[2] Borame L Dickens, Joel R Koo, Annelies Wilder-Smith, and Alex R Cook. Institutional, not home-based, isolation could contain the covid-19 outbreak. *The Lancet*, 395(10236):1541–1542, 2020.

[3] Marie Gousseff, Pauline Penot, Laure Gallay, Dominique Batisse, Nicolas Benech, Kevin Bouiller, Rocco Collarino, Anne Conrad, Dorsaf Slama, Cédric Joseph, Adrien Lemaignen, François-Xavier Lescure, Bruno Levy, Matthieu Mahevas, Bruno Pozzetto, Nicolas Vignier, Benjamin Wyplosz, Dominique Salmon, Francois Goehringer, and Elisabeth Botelho-Nevers. Clinical recurrences of covid-19 symptoms after recovery: Viral relapse, reinfection or inflammatory rebound? *Journal of Infection*, 81(5):816 – 846, 2020.

[4] Liangrong Peng, Wuyue Yang, Dongyan Zhang, Changjing Zhuge, and Liu Hong. Epidemic analysis of covid-19 in china by dynamical modeling. *medRxiv*, 2020.

[5] Andrea Remuzzi and Giuseppe Remuzzi. Covid-19

and italy: what next? *The Lancet*, 395(10231):1225 – 1228, 2020.

[6] Milan Batista. Estimation of the final size of the second phase of the coronavirus covid 19 epidemic by the logistic model. *medRxiv*, 2020.

[7] Jesse Pietz, Scott McCoy, and Joseph H. Wilck. Chasing john snow: data analytics in the covid-19 era. *European Journal of Information Systems*, 29(4):388–404, 2020.

[8] Duccio Fanelli and Francesco Piazza. Analysis and forecast of covid-19 spreading in china, italy and france. *Chaos, Solitons & Fractals*, 134:109761, 2020.

[9] Saptarshi Chatterjee, Apurba Sarkar, Swarnajit Chatterjee, Mintu Karmakar, and Raja Paul. Studying the progress of covid-19 outbreak in india using sird model. *medRxiv*, 2020.

[10] Xin-She Yang. *Nature-inspired optimization algorithms*. Elsevier, 2014.

[11] Francois A. Engelbrecht and Robert J. Scholes. Test for covid-19 seasonality and the risk of second waves. *One Health*, 12:100202, 2021.

[12] Kamil Kozioł, Rafał Stanisławski, and Grzegorz Bialic. Fractional-order sir epidemic model for transmission prediction of covid-19 disease. *Applied Sciences*, 10(23), 2020.

[13] Nan Xie and Henry Leung. Reconstruction of piecewise chaotic dynamic using a genetic algorithm multiple model approach. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(6):1210–1222, 2004.

[14] R. Baragona, F. Battaglia, and D. Cucina. Fitting piecewise linear threshold autoregressive models by means of genetic algorithms. *Computational Statistics & Data Analysis*, 47(2):277–295, sep 2004.

[15] World Health Organization Regional Office for Europe. Transmission of sars-cov-2: implications for infection prevention precautions, 2020. https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions.

[16] Johns Hopkins University Center for Systems Science and Engineering. Covid-19 data repository, 2020. https://www.github.com/CSSEGISandData/COVID-19.

[17] Pratibha Bajpai and M. Kumar. Genetic algorithm - an approach to solve global optimization problems. *Indian Journal of Computer Science and Engineering*, 1(3):199–206, 2010.

[18] Agus Hasan, Endah R. M. Putri, Hadi Susanto, and Nuning Nuraini. Data-driven modeling and forecasting of covid-19 outbreak for public policy making. *medRxiv*, 2020.

[19] Agus Hasan, Hadi Susanto, V. R. Tjahjono, R. Kusdiantara, E. R. M. Putri, P. Hadisoemarto, and N. Nuraini. A new estimation method for covid-19 time-varying reproduction number using active cases. *medRxiv*, 2020.

[20] H. Susanto, V. R. Tjahjono, A. Hasan, M. F. Kasim, N. Nuraini, E. R. M. Putri, R. Kusdiantara, and H. Kurniawan. How many can you infect? simple (and naive) methods of estimating the reproduction number. *Communication in Biomathematical Sciences*, 3(1), 2020.

[21] Hyuk-Jun Chang. Estimation of basic reproduction number of the middle east respiratory syndrome coronavirus (mers-cov) during the outbreak in south korea, 2015. *Biomedical engineering online*, 16(1):79, 2017.

[22] Bert George, Bram Verschuere, Ellen Wayenberg, and Bishoy Louis Zaki. A guide to benchmarking covid-19 performance data. *Public Administration Review*, 80(4):696–700, 2020.

[23] Kuok Ho Daniel Tang. Movement control as an effective measure against covid-19 spread in malaysia: an overview. *Journal of Public Health*, pages 1–4, 2020.

[24] Voice of America. Luck, culture helped cambodia contain coronavirus, 2020.

[25] World Bank Blogs. What explains cambodia's effective emergency health response to covid-19 (coronavirus)?, 2020. https://blogs.worldbank.org/health/what-explains-cambodias-effective-emergency-health-response-covid-19-coronavirus.

[26] Justin Wong, Wee Chian Koh, Mohammad Fathi Alikhan, Anita B. Z. Abdul Aziz, and Lin Naing. Responding to covid-19 in brunei darussalam: Lessons for small countries. *Journal of Global Health*, 10(1), 2020.

[27] World Health Organization. Covid-19 situation report for lao pdr#19: 06 - 25 november 2020, 2020. https://www.who.int/laos/internal-publications-detail/covid-19-situation-report-for-lao-pdr-19.

[28] George R Puno, Rena Christina C Puno, and Ida V Maghuyop. Covid-19 case fatality rates across southeast asian countries (sea): a preliminary estimate using a simple linear regression model. *Journal of Health Research*, 2021.

[29] The Jakarta Post. Broke and helpless: Philippines virus lockdown maroons dozens at airport, 2020.

[30] Detikcom. Terpopuler sepekan: Ini yang tak boleh dilakukan saat dki psbb 14 september, 2020.

[31] Gov.sg. Moving into phase 2: What activities can resume, 2020.

[32] The Straits Times. Myanmar puts yangon in covid-19 lockdown for 2 weeks as elections loom, 2020.

[33] CNA. Malaysia to allow interstate travel from monday, some covid-19 restrictions remain for kl, selangor and sabah, 2020.

[34] Bloomberg. Thailand to impose broad lockdown to fight novel coronavirus, 2020.

[35] VnExpress. Ha noi, hai durong tim nguoi den nha hang the gioi bo tuoi, 2020.

[36] Fresh News. Cambodian government closes public and private schools in phnom penh capital and siem reap city to fight covid-19, 2020.

[37] BruDirect.com. All mosques, suraus and religious halls in the country will be temporarily close for a week-period, 2020.

[38] TTR Weekly. Laos goes into total lockdown, 2020. https://www.ttrweekly.com/site/2020/04/laos-goes-into-total-lockdown/.

[39] The Center for Strategic and International Studies. Southeast asia covid-19 tracker, 2020.