Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia
Vol. 4, 2021, page 81-91

Simulation of COVID-19 for The Third Wave in Sabah by Using SEIR Model

[1]Nur Aishah Athirah Mohd Salleh and [2]Anati Ali

[1,2]Department of Mathematical Sciences
Faculty of Science, Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia.

e-mail: [1]nuraishahathirah@graduate.utm.my, [2]anati@utm.my

Abstract Mathematical model in epidemiology is widely used to predict the number of individuals for infectious diseases. As Malaysia confronts with a new wave of COVID-19 which is more challenging compared to the previous ones, Sabah is one of the states that sparked the third wave of COVID-19 in Malaysia. Susceptible, Exposed, Infectious and Removed (SEIR) model is a nonlinear ordinary differential equation with four compartments, which is used to simulate of active cases for several situations with different assumption of R-naught values.

Keywords SEIR model; Epidemiology; Nonlinear Ordinary Differential Equations; R-naught; COVID-19; Sabah.

1 Introduction

Coronavirus disease (COVID-19) is an infectious disease, which the first case is being reported in Wuhan, China in the late of 2019 [1]. As the number of the cases rapidly increases in China, the COVID-19 has been spreading to the other countries as well. Ever since then, the whole world is not the same as it used to live. Malaysia has recorded the first few cases of COVID-19 on January 25, 2020 with three China tourists who entered Malaysia from Singapore on January 23, 2020 were tested positive COVID-19 [2]. This marks Malaysia hit the first wave of COVID-19 but all confirmed cases have been fully discharged on February 16, 2020 [3]. Later, the second wave was started on February 27, 2020 after 11 days of zero confirmed cases and lasted until June 30, 2020 [4]. Malaysia recorded the first ever no local transmission cases being reported on the July 1, 2020, [5] and indicated the ending of the second wave of COVID-19 in Malaysia.

After a few months, Malaysia encounters the third wave of COVID-19 that began on September 20, 2020, which is much worser than the previous waves. Sabah had only 417 cases up until August 31, 2020 with no new cases are recorded within seven days [6]. Unfortunately, Sabah starts to record new confirmed cases from September 1, 2020 and keep rising until it became the beginning of the third wave in Malaysia. The reason Sabah started recording new confirmed cases is due to a cluster at detention center in Lahad Datu, Benteng LD cluster and has spread and impacted the state election [7].

Mathematical model in epidemiology is used to predict the number of cases infectious diseases. Kermack-McKendrick model is among the simplest of SIR model for describing the endemic disease [8]. There are many suggested models that have been used to predict the number of COVID-19 cases around the world such as SIR and SEIR model. SEIR model was considered for simulation studies and data fitting of Ebola epidemic spreading in West Africa [9]. Therefore, SEIR model can be used for COVID-19 too. There are four compartments, which each of these compartments represents the number of individuals in SEIR model, which are *S*, *E*, *I* and *R*. *S* is the number of individuals those are susceptible to the infectious disease but not infected at time *t*. *E* is the number of individuals that are exposed to the disease but not infectious to others. *I* is the number of infected individuals that are infectious to the population. *R* is the number of removed individuals either they are recovered from the disease or removed by death due to the disease.

2 Methodology

The classical infectious disease model is used to model the spread of disease. By manipulating the basic reproduction number, R_0, there are some different types of simulation can be made. From the simulated curves, the simulations can be compared with the actual data cases that were reported. Simulation is made by using R Studios with SEIR R-package [10] and used R-app with Shiny package [11].

2.1 Mathematical Formulation in SEIR Model

The mathematical model used in this research is SEIR model that can be described in the Figure 1. The figure is describing the progress of an individual in sequence from the susceptible (*S*), exposed (*E*), infectious (*I*) and removed (*R*) when encounter with the contagious diseases.

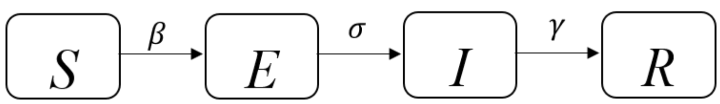

Figure 1 The Compartment of SEIR Model.

The SEIR model describes the change in number of individuals at each compartment in terms of three parameters, β, σ and γ. The transmission rate, β is given by the product of basic reproduction number, R_0 with the recovery rate, γ. σ is the rate of infection and calculated as the inverse of incubation while the recovery rate, γ is the inverse of the infectious period. The governing equations describing the SEIR model can be described by a set of ordinary differential equations as given [9]:

$$\frac{dS}{dt} = \frac{-\beta SI}{N} \tag{3.1}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E \tag{3.2}$$

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{3.3}$$

$$\frac{dR}{dt} = \gamma I \tag{3.4}$$

$$\beta = R_0 \gamma \tag{3.5}$$

2.2 Data Analysis Indicators

For the data of confirmed cases of COVID-19 in Sabah from September 1,2020 until December 31, 2020, are taken from the Jabatan Kesihatan Negeri Sabah [6]. The size of population in Sabah is estimated to be 3.9 million for the initial susceptible, $S(0)$ [12]. The initial infected, $I(0)$ is 18, which is the number of active cases reported until September 1, 2020. Next, the initial exposed per infectious, $E(0)$ is $20I$ according to Read et al.'s estimation [13] and initial removed $R(0)$ is set to be zero [14]. The rate of infectious, σ, which is the inverse of incubation, is set to be 5.2 days according to Li et al. [15]. As for the recovery rate, γ, it is the inverse of the infectious period, 14.5 days is taken which is the total average recovery time [16].

3 Results

There are several simulations of active cases and total cases for some situations by using different basic reproduction number for the result. Also, the graph of actual data in Sabah is included in the result.

3.1 Estimation Active Cases Assuming That Insufficient Prevention and Control Measures

The simulations are made by several values of R_0 to observe different worst situations assuming the prevention and control measures taken during the spread of COVID-19 are insufficient. As for these simulations, all parameters and initial condition values are same except the basic reproduction number. The R-naught values of 3.7, 2.2 and 1.5 are taken to predict the number of infectious and total number of cases for 120 days by using SEIR model [17, 18].

By assuming the $R_0 = 3.7$, the number of COVID-19 active cases is predicted will reach the peak on 100th day with 1.47 million active cases. It is the worst scenario if the basic reproduction number is not reduced throughout the outbreak. From the Figure 2, the total cumulative number is observed and it might reach 3.5 million of cases on 120 days which is almost all population in Sabah will be infected if there is insufficient prevention and control measures taken.

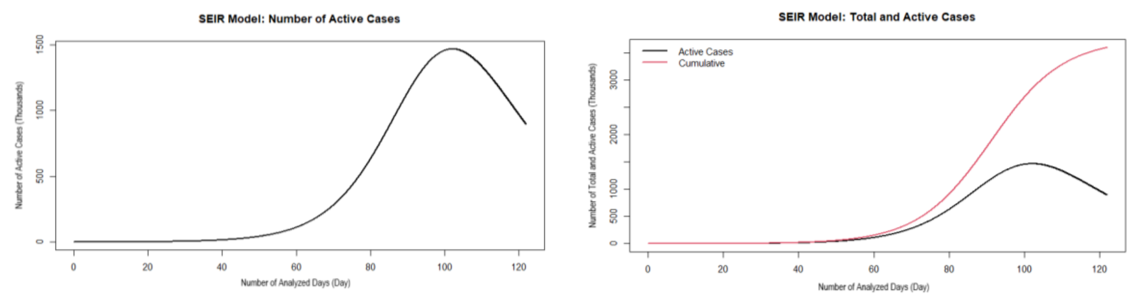

Figure 2 The Number of Active and Total Cases with $R_0 = 3.7$.

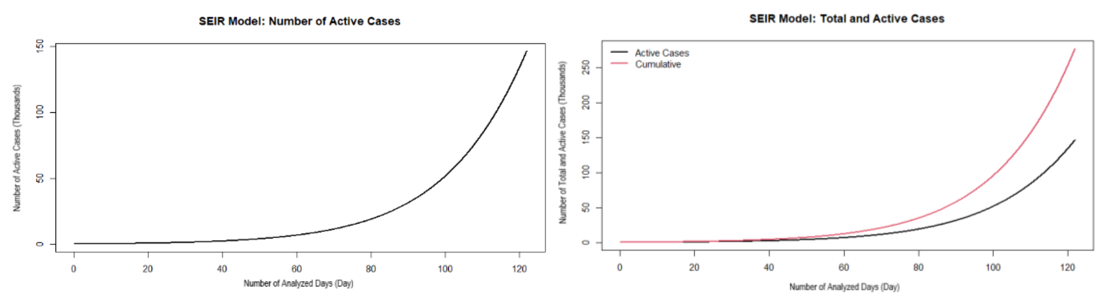

Figure 3 The Number of Active and Total Cases with $R_0 = 2.2$.

Next, R_0 is assumed to be 2.2, which is the basic reproduction number on the early third wave in Malaysia. As shown in the Figure 2, the simulation does not reach the peak after 120 days compared to the R_0 of 3.7. Also, there is around 50 000 active cases estimated on the 100th day while the cumulative active cases exceed 250 000 cases after 120 days as shown in Figure 3. When the R_0 is 1.5, the number of infectious cases is estimated to be less than 4000 which is approximately around 3 200 cases on the 100th day. Besides that, the total number of cases will be exceeding 15 000 cases after 120 days as observed in Figure 4.

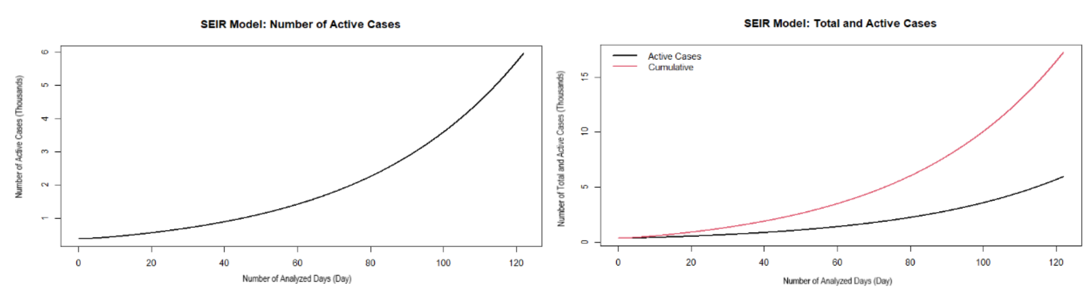

Figure 4 The Number of Active and Total Cases with $R_0 = 1.5$.

3.2 Estimation Active Cases Assuming That Sufficient Prevention and Control Measures

On the other hand, the number of active cases is simulated, which such prevention and control measures are sufficient in these scenarios. The value of S, E, I, R, σ and γ are same as the previous estimation. There are sufficient prevention and control measures taken assumed during the outbreak. So, the basic reproduction number for the scenarios will be reduced for a certain time gradually. For this estimation, the simulation is divided into 5 stages. For the first stage, from day 1 to day 14, the R_0 has been set to 3.7 due to the re-emerge of Covid-19 active cases in Sabah.

The third wave happened unexpectedly during the time of recovery movement control order where the strict prevention and control measures previously have been loosening.

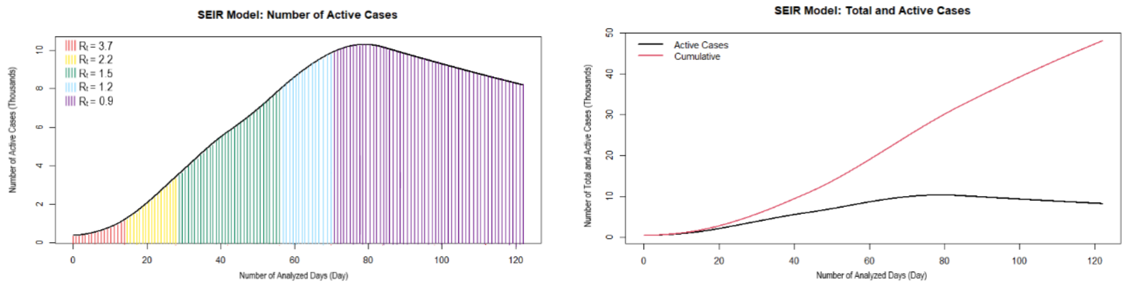

Figure 5 Estimation of COVID-19 Active Cases with Prevention and Control Measures ($R_0 = 3.7$, $R_t = 2.2, 1.5, 1.2, 0.9$).

At the end of the first stage, there are 1 200 infectious people of Covid-19. The second stage is from day 15 to day 28 where the government has taken action to take stricter control measures to break the chain by implementing the targeted enhanced movement control order (TEMCO) on the specific high-risk area. The result shows that 3 400 was infectious at the end of this phase with R_0 reduces to 2.2. On the third stage, Condition Movement Control Order (CMCO) was introduced to all district in Sabah which is from day 29 to day 56. With the CMCO, the number of infectious is decreasing to 7 900 and the R_0 keep on reducing to 1.5. Looking on the fourth stage, when the R_0 reduced to 1.2 from day 57 to day 70, the number of infectious cases is 9 900. The R_0 reduced due to many medical health staff from peninsular have been sent to Sabah since the third stage and an armed field hospital has been deployed Sabah. On the fifth stage, from the effect of the control measures taken for the past days, it has showed positive outcome.

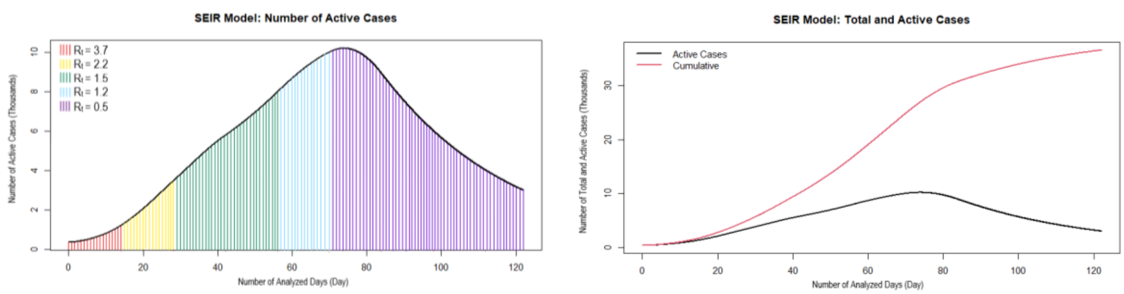

Figure 6 Estimation of COVID-19 Active Cases with Prevention and Control Measures ($R_0 = 3.7$, $R_t = 2.2, 1.5, 1.2, 0.5$).

By looking at the graph in Figure 5 and 6, R_0 gradually decreased to 0.9 or 0.5, with the number of actives cases reached the peak at day 79 or 73 with 10 300 cases and 10200 cases respectively. On day 84, there are 10 200 cases when the R-naught is 0.9 while 9 000 cases when the R-naught is 0.5. By using the R_t of 0.9 and 0.5 on the fifth stage, there is difference in the total number of cases. There are about 30 000 cases for R_t of 0.9 and 28 000 cases for R_t of 0.5 on the 80th day on Figure 5 and 6 respectively which both of the numbers are not really far apart. But, on the 120th day, the total number of infectious cases is approaching 50 000 cases with R_t of 0.9 while when R_t is 0.5, the estimated total number of cases is around 35 000 cases.

3.3 Actual Data of Active Cases in Sabah

The actual data is collected from the Jabatan Kesihatan Negeri Sabah from 1 September 2020 until 31 December 2020 [9]. The graph of the active cases and graph of total and active case are illustrated below in Figure 7. The peak of the number of active cases at day 67 where 8886 people were infectious while on 80th day, the total of active cases is 24929 cases. On the 120th day, the gradient of the total active cases reduces and the number becomes 36461 cases.

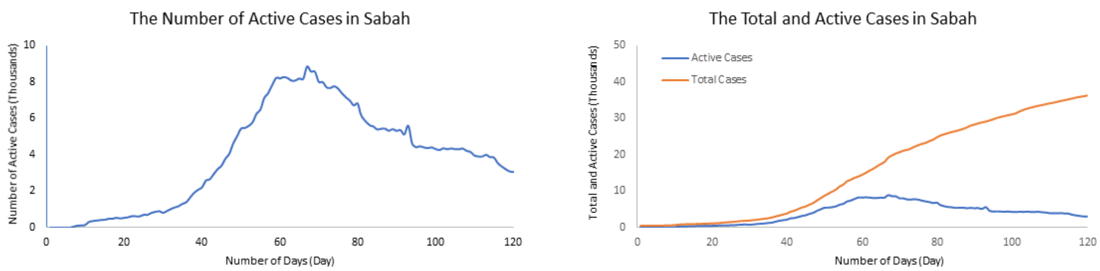

Figure 7 The Actual Number of Active and Total Cases in Sabah.

4 Discussion

The comparison number of cases between two different simulations based on result is discussed. Next, observation on the two different values of R-naught that being used in these simulations and comparison between the actual data in Sabah with one of the simulation results to see on the reliability of the simulation model are also included in the discussion. Besides that, there are simulation on the number of days taken to have the decline in number of COVID-19 cases and simulation that contains two waves inspired by the real number of cases.

4.1 Comparison Number of Cases Between Two Different Simulations

The number of days for the simulation is 121 days which is the total number of days between 1 September 2020 until 31 December 2020. It has been broken down into eight parts with their own R-naught values. Each part of the R-naught values contains 14 days of simulation. R-naught values of 3.7, 2.2 and 1.5 have been selected for the simulation without sufficient prevention and safety measures. The number of active cases increases exponentially up until to a peak of 1.47 million at day 100 when R-naught is 3.7 as shown in Figure 2. The cumulative number of active cases reaches 3.5 million at day 120. While for the second simulation in Figure 3, which the R-naught is estimated to be 2.2, the number of active cases is 50 000 at day 100. At day 120, the number of active cases becomes 148 000 and the total of active cases that being estimated is 250 000. The number is decreasing and reduced by 29.4 times compared to number of active cases at day 100 and 14 times compared to the total number of active cases at day 120 between R-naught of 3.7 and 2.2.

For the next R-naught, it is assumed to be 1.5 as shown in Figure 4. The number of active cases at day 100 is 3 200 while at day 120, it is estimated to be 6 000. The sum of active cases recorded at day 120 is 18000. Based on observation, the number decreases 15.6 times compared to the number of active cases at day 100 and 13.9 times compared to the total of number active cases at day 120 between R-naught 2.2 and 1.5.

Table 1 The comparison of simulated active cases and cumulative number of COVID-19 with different R_0 on day 100 and day 120.

	Day 100			Day 120		
R_0	3.7	2.2	1.5	3.7	2.2	1.5
Simulated Number of Active Cases	1 470 000	50 000	3200	950 000	148 000	6 000
Simulated Cumulative Number of Active Cases	2 900 000	100 000	9800	3 500 000	270 000	18 000

Besides that, the simulations are made to differentiate few based on assumption when there is sufficient prevention and control measures implemented such as wearing mask, social distancing, and lockdown in the region. R-naught of 3.7, 2.2, 1.5 and 1.2 are chosen for break 1 until 5 but on 6^{th} break until 8^{th} break, 2 different R-naught are assigned, 0.9 and 0.5 by assuming the decreasing trend with respectively assigned R-naught until the end of the days.

Table 2 The differences in active cases on the simulation between R_t of 0.9 and 0.5.

Break	**Day**	**Number of Active Cases**		**Difference**
		R_t = 0.9	R_t = 0.5	
6	84	10200	9000	1200
7	98	9400	6100	3300
8	112	8800	4100	4700

From the Figure 5 and Figure 6, the number of active cases for both simulation from break 1 until 5 are same due to the same R-naught applied on the breaks. But the difference is on when two different R-naught are assigned for the rest of breaks. The number of active cases on day 84 is 10 200 cases when the R-naught is 0.9 while 9 000 active cases when the R-naught is 0.5. The difference of active cases between both simulation on day 84 is 1200. Next, there are 9 400 cases and 6 100 cases when the R-naught is 0.9 and 0.5 with difference of 3 300 on day 98. On the last break of assigned R-naught, the number of active cases is 8 800 cases with R-naught of 0.9 and 4100 active cases with R-naught of 0.5. The difference on day 112 is 4 700 cases between these R-naught which more than half of the active cases with R-naught of 0.9.

4.2 Comparison of Active Cases Between Simulation and Actual Data

Based on the observation in Figure 5.1, there are difference in the number of active cases on certain days. There is a huge gap between the simulation and the actual, which made error of the simulation is high. The actual data is exponentially increasing while the simulation is gradually increasing before reaching the peak of the graph. But, the trend of the graph is relatively the same. The number of active in the simulation is a bit higher than the actual data, which gives some disadvantages on the accuracy of the simulation made. On the bright side, it is better to expect a higher number or margins, so that government or authorities can prepare to plan for the worst. As it might be seen in the graph, the peak of the simulation result is around 10 200 on day 79 while the peak of the actual data is 8 886 on day 67. The actual peak is 12 days earlier than the simulation peak.

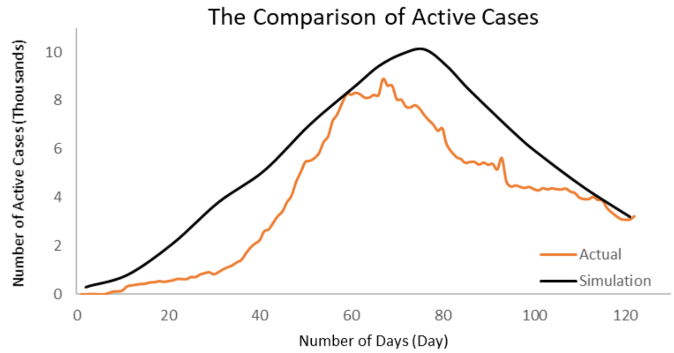

Figure 8 Graph of Simulation and Actual Active Cases.

The SEIR model has been used by the Ministry of Health Malaysia in order to predict COVID-19 cases. Based on the prediction, the government has made several preventions to anticipate the number of predicted COVID-19 cases. That might be one of the reasons why the actual data is less than the predicted data.

4.3 Estimation on the Number of Days to Reach Zero Active Cases

Figure 5 is the simulation when the number of analyzed day is lengthen until the number of active cases reaching zero by using the R-naught value, 0.9 for the rest of the analyzed day, it takes a lot time to reach zero cases which is around at day 900 on February 17, 2023. Other than that, the total number of cases after 1 000 days is estimated to be 110 000.

Next, the number of analyzed day is lengthen based on the situation in Figure 6. It is shown that the estimated day when the active cases become zero is on day 370. This is when the R-naught of 0.5 is continue being applied after day 112 until it reached zero. On this case, the total number of active cases is below than 40 thousand cases which is approximately 39 000 until there is no new cases.

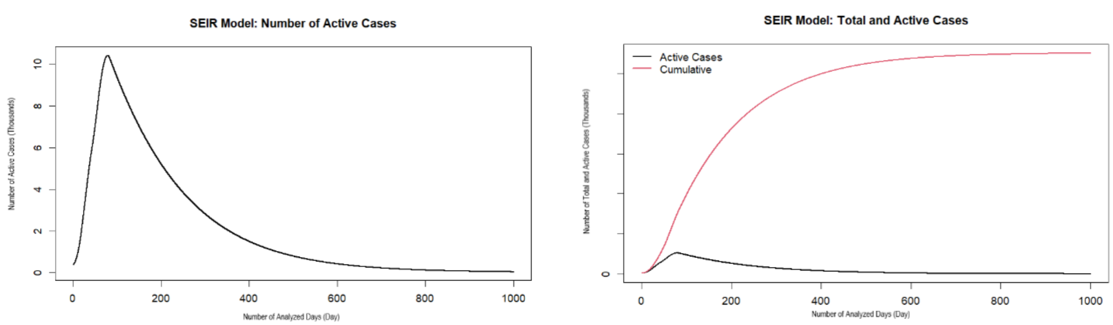

Figure 9 Estimation of Analyzed Days to Reach Zero Active Cases with $R_t = 0.9$.

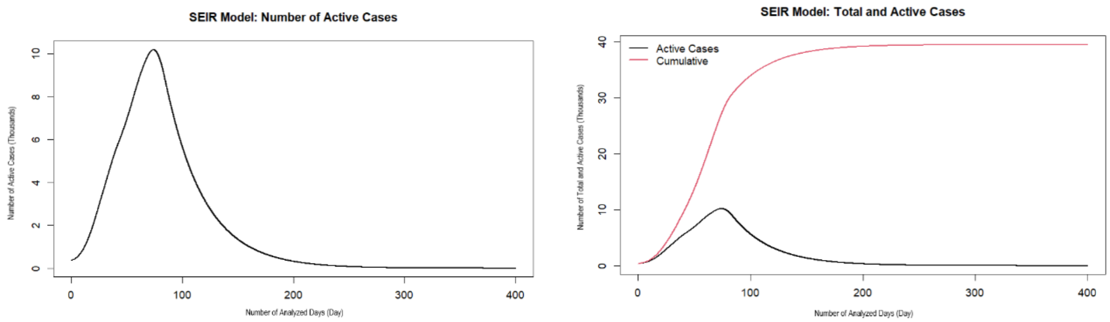

Figure 10 Estimation of Analyzed Days to Reach Zero Active Cases with $R_t = 0.5$.

From these two situations, when the R-naught value is 0.9, it takes 900 days while 370 days for R-naught of 0.5. Even with 0.5, it might take a year to have a decline in number of active cases. This is only considered when consistent and effective sufficient prevention and safety measures taken such as wearing mask, washing hands regularly, practicing social distancing and in such situation, lockdown in a certain area.

4.4 Estimation of COVID-19 Cases with A New Spike

There is decrease in active cases at the end of analyzed day within 1 September 2020 until 31 December 2020. Safety measures are also have been lighten by allowing some economic sectors to operate. The active cases have been increasing since January 2021 and led to a new spike. There are two waves in the simulation based on the actual active cases from 1 September 2020 until 31 March 2021 in Figure 11. The second peak happened due to the withdrawal of interstate and inter-district travel ban on December 2020. The reopening of tourism sector also contributes to the second peak of COVID-19 in Sabah. To make it worse, it is found that there is a new identified of COVID-19 mutation that similar to the strain found in South Africa, Australia and Netherlands, which is more infectious in Sabah. The peak of second wave is slightly lower than the peak of first wave, with 7502 cases on day 146 compared to 8886 cases on day 67.

The second wave is simulated by increasing the value of the R-naught on break 9. The same R-naught values as the previous spike with a little modification, which are 2.2, 1.2, 0.9 and 0.5. This creates a second peak in the graph. In this simulation, the first peak is estimated on day 79 with 10 200 cases while the peak of second wave is estimated on day 149 with approximately 8 000 cases. From Figure 12, due to the second wave between the analysed day, there is an increase in number in cumulative number of cases too. At the end of the graph, a total of 60 thousand cases is estimated to be infected with COVID-19.

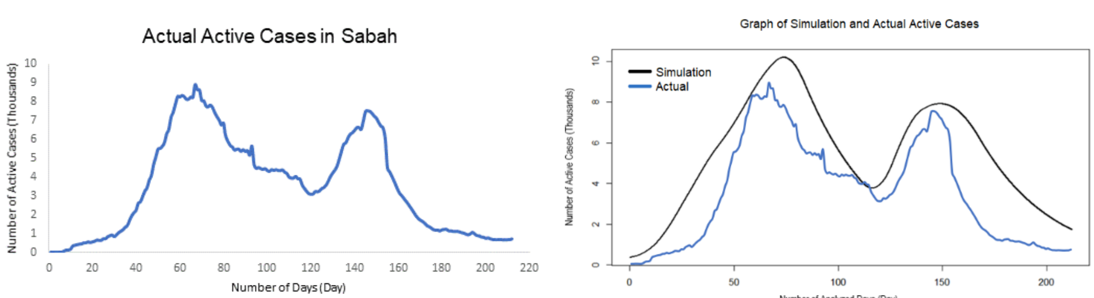

Figure 11 The Actual Active Cases in Sabah and Simulation with A New Spike.

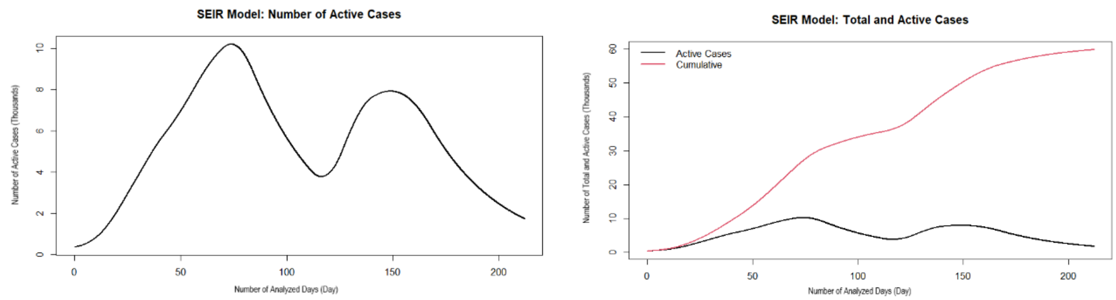

Figure 12 Number of Active and Total Cases with Two Spike.

5 Conclusion

SEIR model is one of the mathematical models in the field of epidemiology that helps to predict the upcoming infectious cases such as the current pandemic of COVID-19. Previous assumption of infectious diseases on the parameters and initial values does affect the result of the simulation. As George Box (1976) said, “All models are wrong, but some are useful” [19], which means that all the models are just prediction. For the limitations of this research, several coefficients of parameters that are collected from another country which might had some differences and affecting the results. This research does not take into account the mutation of COVID-19. Currently, there are several new variants of COVID-19 have been identified during this research which can manipulate the infection rate for the SEIR model.

Therefore, there are some suggestions on future research by considering the coefficients of parameters that are collected in Malaysia and to include the mutation of COVID-19. Next, SEIR model can be improvised by optimizing the parameters of the model. An optimization algorithm should be considered to estimate the best values for the parameters to estimate the R-naught for the infectious diseases. Other than SEIR model, the prediction of COVID-19 cases can be simulated by using machine learning method.

Even though there is significant error between the number active cases in simulation and the actual active cases, but the trend of the simulations is likely similar to the actual data of active cases. It is believed that the reliability of the mathematical modeling in SEIR model can help us to predict the forthcoming COVID-19 cases. Therefore, a swift action can be taken in order to anticipate and triumph over the COVID-19 pandemic and live freely as per old normal again.

References

[1] World Health Organization (WHO), 2020. Coronavirus. *World Health Organization*. Available at: https://www.who.int/health-topics/coronavirus#tab=tab_1

[2] Minister of Health Malaysia, 2020. Kenyataan Akhbar Ketua Pengarah Kesihatan Malaysia Pada 24 Januari 2020 -Tindakan Kkm Bagi Pengesanan Kontak (Contact Tracing) Kepada Kes Pertama Positif Novel Coronavirus di Negara Singapura Pada 24 Januari 2020. Available at:https://www.moh.gov.my/index.php/database_stores/store_view_page/21/1300

[3] Ministry of Health Malaysia, 2020. Updates on the Coronavirus Disease 2019 (COVID-19) Situation in Malaysia. *Press Statement Ministry of Health Malaysia Updates*. Available at: http://www.moh.gov.my/moh/modules_resources/database_stores/96/337_1380.pdf
[4] Rampal, L, and B S Liew. "Malaysia's third COVID-19 wave - a paradigm shift required." *The Medical journal of Malaysia* vol. 76,1 (2021): 1-4.
[5] The Straits Times, 2020. Coronavirus: Malaysia records zero locally transmitted infections for first time in 4 months. *The Straits Times*. Available at: https://www.straitstimes.com/asia/se-asia/coronavirus-malaysia-records-zero-locally-transmitted-infections-for-first-time-in-4
[6] Jabatan Kesihatan Negeri Sabah
[7] Danial, M. et al., 2020. Mitigation of COVID-19 clusters in Malaysia. *Journal of Global Health*, 10(2).
[8] Weisstein, E., SIR Model. *from Wolfram MathWorld*. Available at: https://mathworld.wolfram.com/SIRModel.html
[9] Mamo, D.K. & Koya, P.R., 2015. Mathematical Modeling and Simulation Study of SEIR disease and Data Fitting of Ebola Epidemic spreading in West Africa. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, 2(1), pp.106–114.
[10] SEIR R Package. Welcome – R-Forge. *R*. Available at: https://r-forge.r-project.org/
[11] Chausse, P., 2020. *The Susceptible, Exposed, Infectious, Removed (SEIR) Model*. Available at: http://www.arts.uwaterloo.ca/~pchausse/seir [Accessed January 10, 2021].
[12] Department of Statistical Malaysia.
[13] Read, J.M. et al., 2020. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions.
[14] Wang, H. et al., 2020. Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. *Cell Discovery*, 6(1).
[15] Li, Q. et al., 2020. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine*, 382(13), pp.1199–1207.
[16] Mahmud, A. & Lim, P.Y., 2020. Applying the SEIR Model in Forecasting the COVID-19 Trend in Malaysia: A Preliminary Study.
[17] Minister of Health Malaysia, Perkembangan Nilai Rt ATAU R NAUGHT (R0) Dalam Negara. *COVID-19 Malaysia*. Available at: http://covid-19.moh.gov.my/sorotan/092020/perkembangan-nilai-rt-atau-r-naught-r0-dalam-negara
[18] Minister of Health Malaysia, 2020. Penurunan Nilai R di Malaysia. *COVID-19 Malaysia*. Available at: http://covid-19.moh.gov.my/sorotan/102020/penurunan-nilai-r-di-malaysia-29102020
[19] Wasserstein, R., 2010. George Box: a model statistician. *Significance*, 7(3), pp.134–135.