UTM UNIVERSITI TEKNOLOGI MALAYSIA
INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

Fitting Time-varying Coefficients SEIRD Model to COVID-19 Cases in Malaysia

Norsyahidah Zulkarnain [1], Muhammad Salihi Abdul Hadi [2], Nurul Farahain Mohammad [3], Ibrahim Shogar[4]
Department of Computational and Theoretical Sciences, Kulliyyah of Science
International Islamic University Malaysia
25200, Kuantan, Pahang, Malaysia
Email: norsyahidah.z@live.iium.edu.my[1], salihi@iium.edu.my[2], farahain@iium.edu.my[3], shogar@iium.edu.my[4]

Submitted: 13/2/2023. Revised edition: 2/5/2023. Accepted: 2/5/2023. Published online: 30/5/2023
DOI: https://doi.org/10.11113/ijic.v13n1.397

Abstract—**This paper proposes a compartmental Susceptible-Exposed-Infected-Recovered-Death (SEIRD) model for COVID-19 cases in Malaysia. This extended model is more relevant to describe the disease transmission than the SIRD model since the exposed (*E*) compartment represents individuals in the disease's incubation period. The mathematical model is a system of ordinary differential equations (ODEs) with time-varying coefficients as opposed to the conventional model with constant coefficients. This time dependency treatment is necessary as the epidemiological parameters such as infection rate β, recovery rate γ, and death rate μ usually change over time. However, this feature leads to an increasing number of unknowns needed to be solved to fit the model with the actual data. Several optimization algorithms under Python's LMfit package, such as Levenberg-Marquardt, Nelder-Mead, Trust-Region Reflective and Sequential Linear Squares Programming; are employed to estimate the related parameters, in such that the numerical solution of the ODEs will fit the data with the slightest error. Nelder-Mead outperforms the other optimization algorithm with the least error. Qualitatively, the result shows that the proportion of the quarantine rule-abiding population should be maintained up to 90% to ensure Malaysia successfully reaches disease-free or endemic equilibrium.**

Keywords—**SEIRD model, COVID-19, simulation, optimization, Malaysia**

I. INTRODUCTION

The coronavirus disease 2019 (COVID-19) has been battled by people worldwide ever since it was first discovered on 31 December 2019 in Wuhan, China. It is an infectious disease caused by SARS-CoV-2, a coronavirus with the possibility of causing fatal respiratory infections in people [1] and transmitting through droplets and airborne particles [2]. Three COVID-19 waves have so far had an impact on Malaysia. Since then, epidemiological modeling of COVID-19 spread via compartmental models is crucial and becoming a powerful tool for outbreak prevention and control.

Epidemiological modeling is a mathematical description of how an infectious disease will spread in a population [3]. Such models may forecast the future growth of COVID-19 transmission dynamics via the prediction of the daily number of infected (*I*) active cases, peak infected (*I*) cases and duration of an epidemic for a particular wave [4]. Various fundamental traditional compartmental models of infectious disease with constant coefficients, including the SIR model, the SIRD model, and the SEIR model, were created. Such models have their roots in research by Kermack & McKendrick [5], who looked at the prevalence and distribution of infectious disease cases as they spread across a population over time.

The SIR model is one of the simplest mathematical models. It divides the population into three groups which are susceptible (*S*), infected (*I*), and removed (*R*). The SIR model is applied by Beckley *et al*. [6] to predict the transmission trend of COVID-19 and Abuhasel *et al*. [7] has studied the effectiveness of government intervention. SIRD and SEIR are the extended models of the fundamental classical SIR model. The SIRD model is an extension of SIR that divides the removed compartment into two compartments: one for recovery (*R*) and the other for death (*D*). This model enables the government to know the exact predicted number of each case. Besides, a new compartment *E* representing the exposed (*E*) individuals in the incubation period is added to the SEIR model, which is a more advanced model to examine the transmission of infectious diseases precisely which has been proposed by Mahmud & Lim [8] and Gupta *et al*. [9]. However, the *D* compartment is absent from this model.

Infectious diseases commonly have an incubation phase once the individuals get exposed to the virus. The incubation period is between a person's initial exposure to a pathogen and the onset of the infection's first symptoms [10]. This delay is crucial to be incorporated within the epidemiological model by adding an exposed (E) compartment which may mimic the actual progression of infected (I) individuals, who progress in sequence from S to E, then E to I.

The SEIRD model, which is the extended model of SEIR and SIRD, is more reliable to be applied in modelling COVID-19 transmission dynamics as studied by Muka & Sannyal [11], Piccolomiini & Zama [12] and Maugeri *et al.* [13] in their studies. These studies adopted the conventional SEIRD model with constant coefficients as they all assume stationarity of epidemiological parameters throughout the simulation. The constant coefficient may be employed during the early stages of an outbreak and with simulations that take place over a shorter time window, but not for the current scenario, which values longer predictions when analyzing the COVID-19 transmission trend. Time-varying coefficients must be employed in the mathematical model as the epidemiological parameters such as infection rate β, recovery rate γ, and death rate μ usually change over time. However, this time dependency treatment leads to an increasing number of unknowns needed to be solved to fit the model with the actual data of COVID-19 cases.

Thus, this study's primary purpose is to fit the time-varying coefficients SEIRD model to actual COVID-19 cases in Malaysia using optimization algorithms. In this paper, we first (1) present our proposed SEIRD model formulation under the Methodology section. It is subsequently followed by the Findings and Discussion section (2), which is divided into two subsections i) optimization algorithm performance of four different algorithms (Levenberg-Marquardt, Nelder-Mead, Trust-Region Reflective and Sequential Linear Squares Programming) (ii) simulation of COVID-19 cases with optimized parameters which shows the effects of population behavior by the value of r, proportion of quarantine rule-abiding population. We summarized our findings in this paper's Conclusion part (3). These findings may give insight into the effectiveness of Malaysia's government health policy in countering the outbreak in 2020 and raise public awareness of the importance of following the rules and standard operating procedures (SOPs) in preventing COVID-19 spread.

II. METHODOLOGY

In this section, we propose our modified SEIRD model. This section will be divided into four subsections i) study area and epidemic data, ii) mathematical model formulation, iii) research design, and iv) input data and initial values. The subsections are presented accordingly.

A. Study Area and Epidemic Data

Since this study is a simulation of general COVID-19 cases in Malaysia, daily data of COVID-19 infected (I), recovered (R), and death (D) cases from the first wave to the beginning of the third wave (from 25 January 2020 to 18 September 2020) are gathered from the Ministry of Health Malaysia (MOH) official data [14]. The COVID-19 data were available to the public on the Ministry's GitHub page. Before importing the data into Pandas DataFrame for the purpose of Python simulation, the data is gathered and sorted in a spreadsheet file. Data on the daily active infected (I), total recovered (R), and total death (D) are utilized to simulate the SEIRD model.

B. Mathematical Model Formulation

Several assumptions have been made to model the reliable transmission dynamics of COVID-19 due to the limited availability of several data. The assumptions are as follows:

1. Malaysian population was a closed population due to the implementation of the international travel restriction enforced on 25 January 2020, which limits the movement of traveling foreigners into Malaysia.
2. There has yet to be an immunization for SARS-CoV-2 since the period of this study is before the immunization program starts. Thus, the Malaysian population is susceptible to COVID-19.
3. It is assumed that the Malaysian population was constant during the relatively short time span of the model simulation. The natural death and newborn were not counted in, which is not realistic, however considering that a person's lifetime is much longer than the disease's outstanding phase, the number is insignificant.
4. There was a chance for the recovered (R) population to be reinfected again.
5. The sporadic case is now considered where we assume the disease can also be probably spread by an exposed (E) individual.
6. The epidemiological parameters except the incubation rate are non-stationary.

The schematic design in Fig. 1 shows how the populations in each compartment progress in sequence with a particular transition rate that accommodates reinfection incidences while considering the assumptions mentioned before. Based on the figure, the *chocolate dashed-dotted curve* represents the interaction or contact between everyone in the indicated compartments. In contrast, the *black line arrow* and *black arrow curved down* to show how the population progress in sequence from one compartment to another subsequent compartment with specific probabilities.

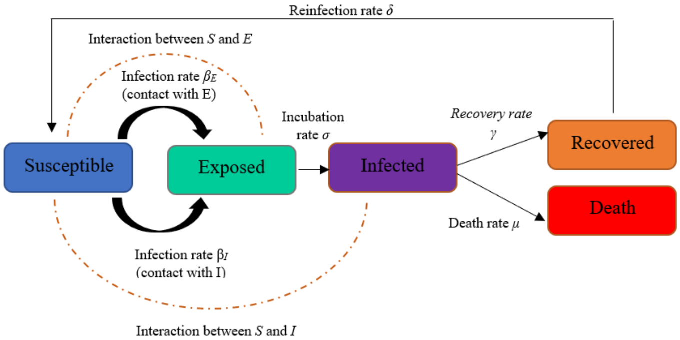

Fig. 1. Schematic diagram of disease progression

By following the assumptions, the modified SEIRD model proposed by Jamil *et al.* [15] that considers reinfection cases and the modified SEIRD model proposed by Muka & Sannyal [11] that considers sporadic cases which highlighted two types of infection rates can be used for constructing a new form of a modified model in this study. This paper proposes a modified SEIRD model with some modifications considering sporadic cases, reinfection cases, and the non-stationarity of epidemiological parameters simultaneously. As depicted in Fig. 1, this non-spatial model consists of five population compartments which is mathematically translated into a system of non-linear first order ordinary differential equations (ODEs) given by Eq. (1) – Eq. (5). Additionally, Eq. (6) represents the assumption that the total population is a constant, denoted by *N*.

$$\frac{dS}{dt} = -\frac{\beta_I(t)SI}{N} - \frac{\beta_E(t)SE}{N} + \delta(t)R \quad (1)$$

$$\frac{dE}{dt} = \frac{\beta_I(t)SI}{N} + \frac{\beta_E(t)SE}{N} - \sigma E \quad (2)$$

$$\frac{dI}{dt} = \sigma E - \gamma(t)I - \mu(t)I \quad (3)$$

$$\frac{dR}{dt} = \gamma(t)I - \delta(t)R \quad (4)$$

$$\frac{dD}{dt} = \mu(t)I \quad (5)$$

$$N = S(\mathrm{t}) + E(t) + I(t) + R(t) + D(t) \quad (6)$$

This study adopts time-varying coefficients of the SEIRD model in predicting COVID-19 cases in Malaysia. In our proposed SEIRD model, we highlight six critical epidemiological parameters. It is because they may provide insights into the actual phenomenon that make it easier for policymakers to make informed decisions on the approach to contain the COVID-19 pandemic. The parameters include infection rate β_I (contact with *I*), infection rate β_E (contact with *E*), incubation rate σ, recovery rate γ, death rate μ, and reinfection rate δ.

In general, those parameters are the rates of individual changes in their state from one to another. Infection rate β_I is the probability of transmitting disease between susceptible (*S*) and infected (*I*) individuals, which then will progress the susceptible (*S*) individuals to the exposed (*E*) compartment. Generally, $\beta = R_0\gamma$, where R_0 is the basic reproduction number [16][17]. Meanwhile, infection rate β_E is the probability of transmitting disease between the susceptible (*S*) and infected (*I*) compartment. Incubation rate σ is the probability of latent individuals who have close contact with the infected (*I*) or exposed (*E*) population becoming infectious in the incubation period, which then will progress the exposed (*E*) population to the infected (*I*) population once the incubation period ends. Traditionally, $\sigma = 1/\mathrm{T_c}$, where $\mathrm{T_c}$ is the average latency duration [18].

In addition, recovery rate γ is the probability of the infected (*I*) population becoming resistant and recovering from COVID-19, which will progress the population to the recovered (R) compartment, where traditionally $\gamma = 1/(\mathrm{T_i})$ with $\mathrm{T_i}$ as the average of recovery duration [1] which is also known as infectious period [19]. Traditionally, death rate μ may be determined by using the formula of $\mu = D/N$ over a certain period, where *D* refers to the cumulative death up to a certain date, and *N* refers to the total population in the country. Moreover, reinfection rate δ can be easily calculated using the simple formula of $\delta = Re/I$, where *Re* is the number of reinfected individuals. Malhotra [20] highlighted that fully vaccinated healthcare workers in India had a lower risk of reinfection than unvaccinated and partially vaccinated with a percentage of 1.6%, 12.7% and 11%, respectively. Since our epidemic data are from the first wave to the early third wave, only before reinfection cases and immunization emergence, the value of reinfection rate δ throughout the simulation of COVID-19 cases in Malaysia is indeed zero.

Since in our model, the epidemiological parameters, especially infection rate β, recovery rate γ and death rate μ are functions of time, so now it is important to determine what type of functions should be employed. Inspired by SIRD model of Jamil *et al.* [15], time-varying infection rate $\beta_I(t)$, recovery rate $\gamma(t)$ and death rate $\mu(t)$ are formulated as piecewise functions, as in Eq. (7) – Eq. (10). While for the infection rate $\beta_E(t)$, it is based on the SEIRD model of Muka & Sannyal [11], where they considered $\beta_E = 5\beta_I$, based on COVID-19 cases in China. In general, one can assume $\beta_E(t) = p\beta_I(t)$ where $p \geq 1$ and therefore $\beta_E(t)$ is defined as Eq. (8). Time interval of these piecewise function was divided into three phases as follows:

- Phase I: Before Movement Control Order (MCO), $t < t_{\mathrm{lock}}$ **(27/2/20-17/3/20)**
- Phase II: During Movement Control Order (MCO) and Conditional Movement Control Order (CMCO), $t_{\mathrm{lock}} \leq t < t_{\mathrm{lift}}$ **(18/3/20-9/6/20)**
- Phase III: During Recovery Movement Control Order (RMCO), $t \geq t_{\mathrm{lift}}$ **(10/6/20-23/2/21)**

$$\beta_I(t) = \begin{cases} \beta_1 t + \beta_2, & t < t_{\mathrm{lock}} \\ \beta_0 e^{-((t - t_{\mathrm{lock}})/(\tau_\beta))}, & t_{\mathrm{lock}} \leq t < t_{\mathrm{lift}} \\ (1 - r)\left(\beta_1 (t - t_{\mathrm{lift}}) + \beta_2\right), & t \geq t_{\mathrm{lift}} \end{cases} \quad (7)$$

$$\beta_E(t) = p\,\beta_I(t) \quad (8)$$

$$\gamma(t) = \begin{cases} \gamma_2(t) + \gamma_3, & t < t_{\mathrm{lock}} \\ \gamma_0 + \dfrac{\gamma_1}{1 + e^{(-t + t_{lock} + \tau_\gamma)}}, & t \geq t_{\mathrm{lock}} \end{cases} \quad (9)$$

$$\mu(t) = \begin{cases} \mu_2(t) + \mu_3, & t < t_{\mathrm{lock}} \\ \mu_0 e^{-((t - t_{\mathrm{lock}})/(\tau_\mu))} + \mu_1, & t_{\mathrm{lock}} \leq t < t_{\mathrm{lift}} \\ \mu_2\,(t - t_{\mathrm{lift}}) + \mu_3, & t \geq t_{\mathrm{lift}} \end{cases} \quad (10)$$

$$\delta(t) = 0 \quad (11)$$

To understand the formulated time-varying infection rate β_I, firstly, at the beginning of the outbreak, which is Phase I: before Movement Control Order (MCO), people had high mobility and were free to move. Thus, the infection rate $\beta(t)$ is assumed to be a linear function, $\beta_1(t) + \beta_2$. For Phase II, when Movement Control Order (MCO) was introduced in Malaysia, the infection rate decayed due to the isolation of people and physical distancing, and this behavior was described by an exponential function, $\beta_0 e^{-((t - t_{\mathrm{lock}})/(\tau_\beta))}$ with β_0 as the initial value of infection rate during that phase and $1/\tau_\beta$ refers to the decay rate. Conditional Movement Control Order (CMCO) will still be regarded as a lockdown or restricted movement even though more economical and social activities were allowed due to limited operating hours and stringent standard operating procedures (SOP). Lastly, for Phase III, when the lockdown was lifted whereby more economic and social activities were allowed, for instance, interstate travel, tourism business, unessential premises or even schools, the infection rate was assumed to follow the trend at the beginning of the outbreak, which is increasing when no lockdown or stringent movement order was implemented.

Following the work of Jamil *et al*. [15], a fraction of compliance to the SOP was included in the infection rate as it affects or contributes to the new values of infection rate and based on the experience in facing the pandemic $(1-r)(\beta_1 (t-t_{lift}) + \beta_2)$, where r was the percentage of Malaysians who followed the SOPs and practiced the 3Ws, even after the government had lifted the lockdown. The numeric value of r was between 0 and 1, which will be regarded as a percentage value. $(1-r)$ here refers to the percentage of Malaysians who are not following the SOPs that lead to the increment of the values of infection rate β. Meanwhile, the rest of it is the linear function.

Time-varying recovery rate $\gamma(t)$ for Phase I was formulated based on linear function as it increases linearly before MCO, and logistic growth function for Phase II as the value of recovery rate is believed to display in a sigmoidal curve starting the first day of MCO and reach equilibrium state at the end of the simulation period. While time-varying death rate $\mu(t)$, it was formulated based on linear function for Phase I and III and using exponential decay function during Phase II.

C. Research Design

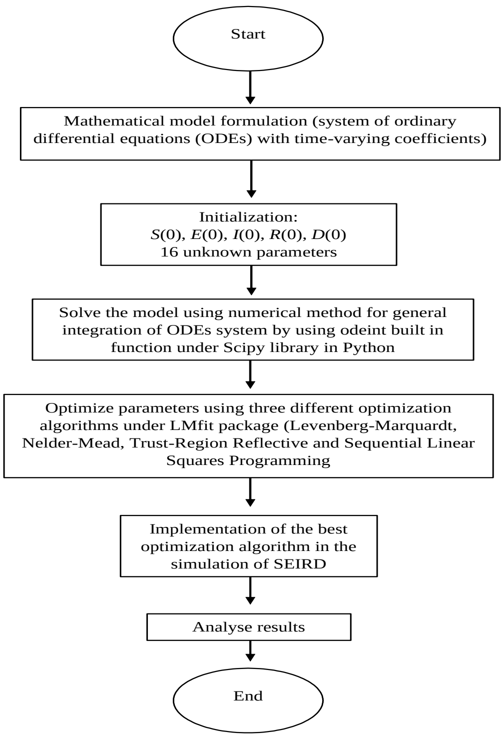

Fig. 2. Research flowchart

The flowchart for the computational work of this study is depicted in Fig. 2 where Python via Jupyter Notebook is the programming platform. SciPy *odeint* function is used to solved the ODEs [21]. The ODEs require initial values in order to obtain the unique solutions, in this case, the initial values for each of population states are given in Table I. In order to fit the model to the actual cases in Malaysia, there are totally 16 parameters need to be optimized by virtue of Least-Squares minimization, using several optimization algorithms under the LMfit package, which are Trust-Region Reflective (TRR), Levenberg-Marquardt (LM), Nelder-Mead (NM), and Sequential Linear Squares Programming (SLSQP). Since all the algorithms are iterative methods, they require initial guesses to begin with. It will be better if the guesses can be set closer to the actual optimized solutions as it will speed up the convergence. The initial guesses for all 16 parameters are summarized in Table II.

D. Input Data and Initial Values

According to Department of Statistics Malaysia [22], the population of Malaysia as of early April 2020 is 32,653902 and this is the value of N in Eq. (6). The initial values of infected $I(0)$, recovered $R(0)$ and death $D(0)$ on 25 January 2020 are 3, 0 and 0, respectively [14]. In reality, exposed (E) data are not available, thus by looking at small value of $I(0)$, the initial exposed $E(0)$ can be assumed 0. Since the total population is considered unchanged over the simulation time interval, we set the constant N = 32,653902 and therefore, from Eq. (6) the initial value of susceptibility $S(0)$ is 32,653899.

TABLE I. INITIAL VALUES FOR SOLVING THE ODEs OF SEIRD MODEL

Initial States	**Values**	**Sources**
Susceptible, $S(0)$	32,653899	Eq. (6)
Exposed, $E(0)$	0	Assumption
Infected, $I(0)$	3	[14]
Recovered, $R(0)$	0	[14]
Death, $D(0)$	0	[14]

There are totally sixteen unknown parameters that required initial guesses but fifhteen of them are already been investigated in References [15] and [18]. These values are considered the closest to the solutions by the iterative optimization algorithms. Thus, only one parameter i.e. the proportionality constant p left where it is assumed to be close to unity rather than five as suggested in Reference [11], the study on China's COVID-19 spread. This is because the Malaysian population is lower compared to China, and Malaysia government has stringently enacted a mandatory 7-days quarantine for everyone with a closed contact history for the whole simulation period, which a bit reduce the risk of the SARS_CoV2 transmission between exposed (*E*) and susceptible (*S*), giving *p* a value of unity. This unity implies that the infection rate β_I, which is based on ordinary transmission is equal to the infection rate β_E, which signifies transmission in sporadic cases that involving asymptomatic individuals infecting susceptible (*S*).

TABLE II. THE INITIAL PARAMETER INPUT FOR THE OPTIMIZATION PROCEDURES

Parameters	Values	Sources
Infection rate, $\beta(t)$	$\beta_0 = 0.16100732$ $\beta_1 = 0.00142347$ $\beta_2 = 0.07373335$	[15]
Incubation rate, σ	0.15	[18]
Recovery rate, $\gamma(t)$	$\gamma_0 = 0.02590983$ $\gamma_1 = 0.0267004$ $\gamma_2 = 8.3000e\text{-}05$ $\gamma_3 = 0.00606688$	[15]
Mortality rate $\mu(t)$	$\mu_0 = 0.00151062$ $\mu_1 = 1.5316e\text{-}04$ $\mu_2 = 8.0126e\text{-}05$ $\mu_3 = 2.5064e\text{-}04$	[15]
Characteristic time of transmission, τ_β	21.7321554	[15]
Characteristic time of recovery, τ_γ	12.3593006	[15]
Characteristic time of death, τ_μ	26.3593227	[15]
Proportion β_E over β_I, *p*	1.002	Assumption

III. RESULT AND DISCUSSION

In this section, we will compare the four optimization algorithms' performances to identify which one has the least error. Then, by using the optimized parameters obtained from the best algorithm, we will predict the possible infected (*I*), recovered (*R*) and death (*D*) cases of COVID-19 by changing the parameter that corresponds to the variation in the Malaysian isolation abiding behaviors.

A. Optimization Algorithm Performance

Trust-Region Reflective (TRR), Levenberg-Marquart (LM), Nelder-Mead (NM) and Sequential Linear Squares Programming (SLSQP) are different built-in optimization algorithms under the LMfit Python package, provided for non-linear optimization curve fitting problems. Both TRR and LM techniques are based on the knowledge on derivative of the objective (or cost) function. While TRR adopts the process of solving a system of equations with a gradient from an objective function, which constitutes the first-order optimality condition for a bound-constrained minimization [23], LM combines gradient descent and the Gauss-Newton as potential possibilities for the algorithm's direction at each iteration [24]. NM is a directive free simplex-based direct search technique for multidimensional unconstrained optimization [25] and SLSQP is a technique that transforms an optimization problem into a successive solution of quadratic programming problems [26].

These four optimization algorithm performances in fitting the time-varying coefficient SEIRD model can be compared based on the obtained Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). These measurements give us information on the exact amount of deviation from the actual values. The optimization technique with the lowest error of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) therefore is the best fitting approach [27].

The fitting of the SEIRD model by the four different optimization algorithms, for *I*, *R*, and *D* cases, are presented in Fig. 3 – Fig. 5, respectively and the error measurements of TRR, LM, NM and SLSQP are listed in Table III. Table III consists of MAE, RMSE, average of each accuracy metric for infected (*I*), recovered (*R*) and death (*D*), and average of the accuracy metrics. Finding the average of the accuracy metrics helps in selecting the best optimization algorithm. Nelder-Mead (NM) algorithm resulted the least average value of accuracy metrics which is 224 showing the best overall performance compared to the other algorithms. It reduces the measurement error of MAE and RMSE in fitting the time-varying coefficients SEIRD model with the least MAE of 246 and RMSE of 347 for infected (*I*) while for recovered (*R*), as low as 270 for MAE and 407 for RMSE. Further, for death (*D*), 20 of MAE and 39 of RMSE.

Nelder Mead (NM) optimization algorithm has significantly reduced the RMSE as much as 66.63%, 85.88% and 72.78% from the early simulation using the initial value of parameters (initial guesses) compared to the other optimization algorithms for infected (*I*), recovered (*R*) and death (*D)* cases, respectively. NM algorithm shows the best performance in reducing the errors. The best result showed by Nelder-Mead (NM) algorithm is in line with what Jamil *et al*. [15] has observed while fitting the 14 unknown parameters of the SIRD model to actual COVID-19 cases in Malaysia.

By using Nelder-Mead (NM) as an optimization algorithm for our predictive modified SEIRD model, we obtained the result of a modified SEIRD model simulation that has a better fit to the actual data of COVID-19 infected (*I*), recovered (*R*) and death (*D*) cases as in Fig. 3–Fig. 5.

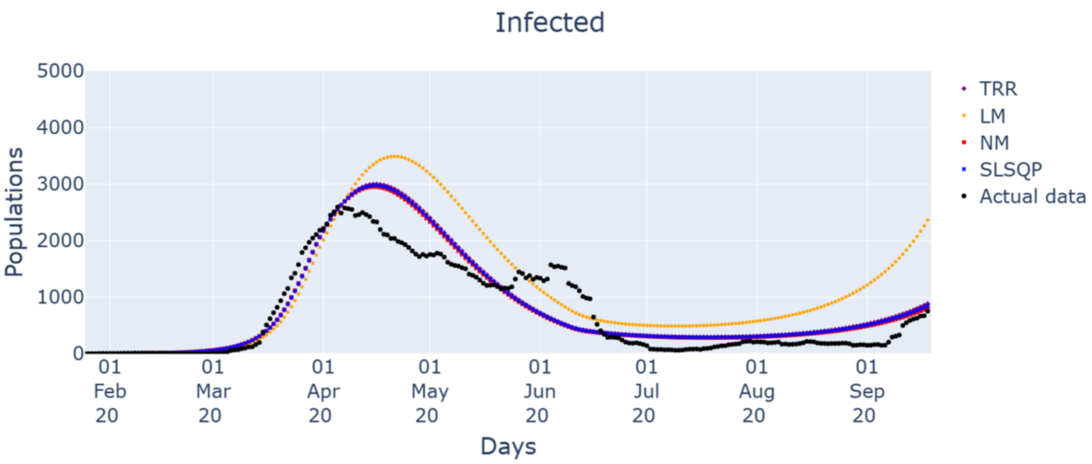

Fig. 3. Fitted infected (*I*) curves by the optimization methods

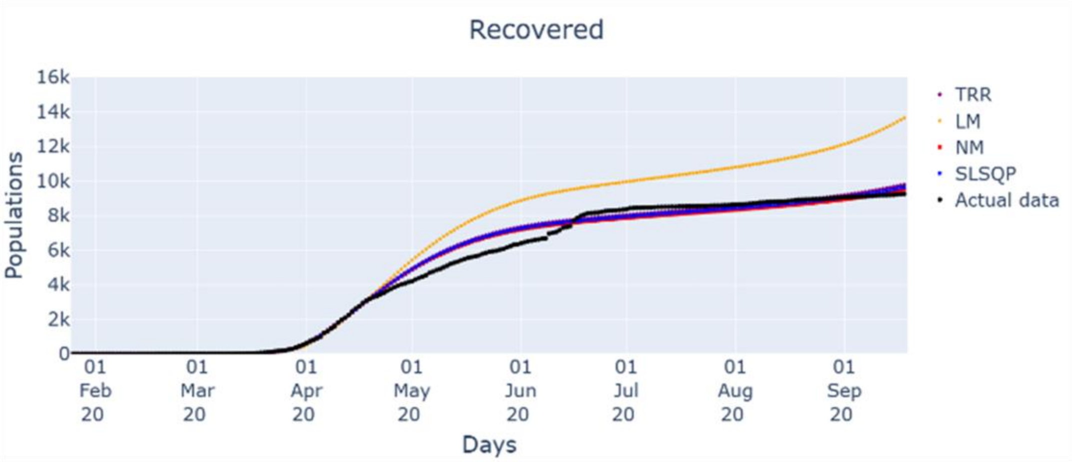

Fig. 4. Fitted recovered (*R*) by the optimization methods

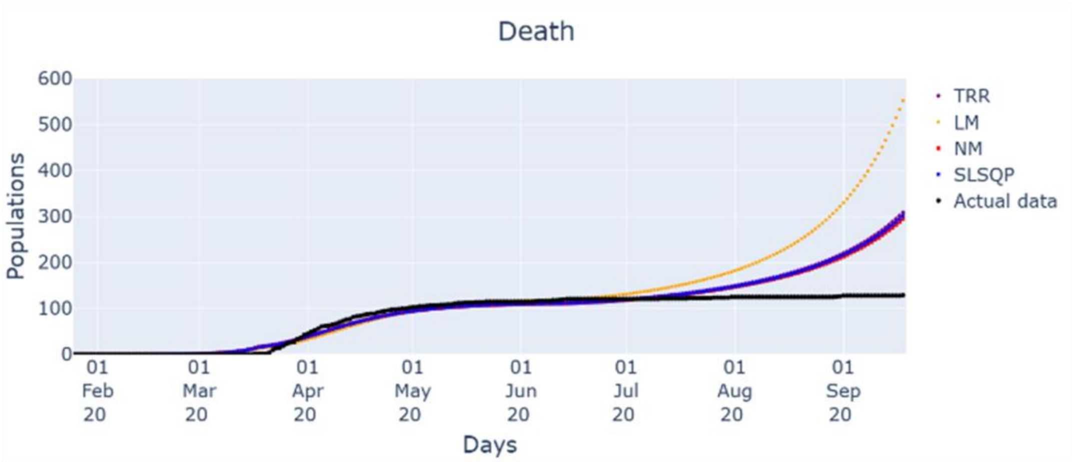

Fig. 5. Fitted death (*D*) by the optimization methods

TABLE III. PERFORMANCE MEASUREMENT COMPARISON FOR FOUR OPTIMIZATION ALGORITHMS

Optimization Algorithms	Cases	MAE	RMSE
Trust-Region Reflective (TRR)	*I*	264.421	364.691
	R	270.931	424.759
	D	22.029	42.9426
	Average of each accuracy metric		
		185.794	277.4642
	Average of accuracy metrics		
		231.629	
Levenberg-Marquardt (LM)	*I*	524.918	700.379
	R	1391.19	1836.63
	D	44.7798	95.8968
	Average of each accuracy metric		
		653.629	877.632
	Average of accuracy metrics		
		765.631	
Nelder-Mead (NM)	*I*	246.249	347.739
	R	288.001	407.077
	D	20.6546	39.601
	Average of each accuracy metric		
		184.968	264.806
	Average of accuracy metrics		
		224.887	
Sequential Linear Squares Programming (SLSQP)	*I*	255.055	355.759
	R	275.624	409.632
	D	21.3168	41.2157
	Average of each accuracy metric		
		183.999	268.869
	Average of accuracy metrics		
		226.434	

Fig. 3. highlights the simulation of infected (*I*) cases which shows how our proposed SEIRD model fit the actual data of the infected (*I*) cases. In this study, infected (*I*) cases are the data of COVID-19 active cases resulting in the bell-shaped form for the infected (*I*) curve. Our proposed modified SEIRD model proved to be better than Anastassopoulou et al. [28], who adopted the compartmental SIRD model without considering the exposed (*E*) compartment. Table IV presents the comparison of Root Mean Squared Error (RMSE) between our proposed model and other published study.

TABLE IV. COMPARISON OF ROOT MEAN SQUARED ERROR (RMSE) FOR EPIDEMIOLOGICAL MODEL

No.	Source	Model	RMSE of infected *(I)*
1.	Anastassopoulou *et al*. [28]	SIRD	6.128918×10^2
2.	Our proposed model	SEIRD	3.47739×10^2
3.	Jamil *et al*. [15]	SIRD	0.80078×10^2

Table V shows the 16 optimized parameters of our proposed modified SEIRD model. Table VI summarizes the date of peak infection, plateau and new wave that has been predicted by our proposed SEIRD model. Our proposed model can predict the peak of infected (*I*) cases during the second wave of COVID-19, which is almost similar to the actual peak of infected (*I*) which is in April 2020. Besides, the predicted date of the infected (*I*) curve to be plateaued, and the plateau breakthrough (new wave) is from mid to end of July 2020. However, based on Fig. 3. there is a slight unfit at the infected (*I*) curve during Jun 2020 due to the sudden spike of infected cases reported among immigrants in immigration detention centers near Kuala Lumpur, which started on 22 May with 35 reported new infected cases. Then, the number had jumped to 410 new infected cases across four sites by the end of May 2020, which portrays the second highest peak of infected (*I*) cases during the second wave of COVID-19 in Malaysia [29].

TABLE V. VALUE OF OPTIMIZED PARAMETERS FOR SEIRD MODEL ESTIMATED BY NELDER-MEAD ALGORITHM

Parameter	Value	Parameter	Value
β_0	0.22160899	μ_0	0.07271124
β_1	0.06311759	μ_1	0.01222001
β_2	0.18181217	μ_2	0.34447642
σ	0.50	μ_3	0.02449339
γ_0	0.00205247	p	1.42616385
γ_1	0.06448763	τ_β	11.9729414
γ_2	0.04059854	τ_γ	19.0321813
γ_3	0.07289225	τ_μ	5.31669064

TABLE VI. COMPARISON OF INFECTED (*I*) CASES BETWEEN ACTUAL AND SIMULATED DATA

Infected Cases	Actual Data	Simulated Data
Peak infection of second wave	2597 (5th April 2020)	2954 (15th April 2020)
Plateau infection of second wave	77 (4th July 2020)	284 (15th July 2020)
New wave	123 (20th July 2020)	290 (28th July 2020)

B. Simulations of COVID-19 Cases with the Effects of Population Behaviour

In combating the pandemic of COVID-19, there are two things that need to be considered and studied: virus behavior and population behavior. SARS-CoV-2 behavior is very complex to control as it mutates itself. Fortunately, we have the capacity to control population behavior. Good population behavior might be the best exit to this life-threatening pandemic as it helps break the chain of COVID-19 transmission.

Based on Fig. 3, we observed that the third wave of COVID-19 in Malaysia started earlier, on 20 July 2020, due to a consistent increment in the daily new infected (*I*) cases. By 10 June 2020, the Malaysian government has implemented Recovery Movement Control Order (RMCO) which relaxes a few stringent movements control orders. To simulate our proposed modified SEIRD model, we regard the first date of Recovery Movement Control Order (RMCO) as the reopening date from Movement Control Order (MCO) on 10 June 2020. Fig. 6. shows an exciting and paramount significant result: the effects of population behavior on the transmission dynamics of COVID-19 in Malaysia which portrays the projection of infected (*I*) cases of COVID-19 in Malaysia after Movement Control Order (MCO) has been lifted-up during early of the third wave.

The percentage of the quarantine rules-abiding population may indicate population behavior. It is essential to be monitored in curbing the spread of COVID-19 as this value affects the occurrence of a new wave of the pandemic and the number of infected (*I*) cases. Rules that must be followed include keeping a safe distance from others, donning face masks, avoiding large gatherings, and checking the temperature. To study the effect of population behavior on transmission dynamics of COVID-19 in Malaysia during the early third wave, we set five different percentage values of populations who abide by the rules (*r*) to simulate the effects of population behaviors in Malaysia which are 90%, 60%, 55%, 50% and 30%. 90% is categorized as strong compliance. 60%, 55% and 50% are categorized as moderate compliance, while 30% are weak. Fig. 6 highlights that the higher the percentage of the quarantine rules-abiding population, the lesser the number of peak infected (*I*), peak recovered (*R*) and peak death (*D*) cases, which are good for the country.

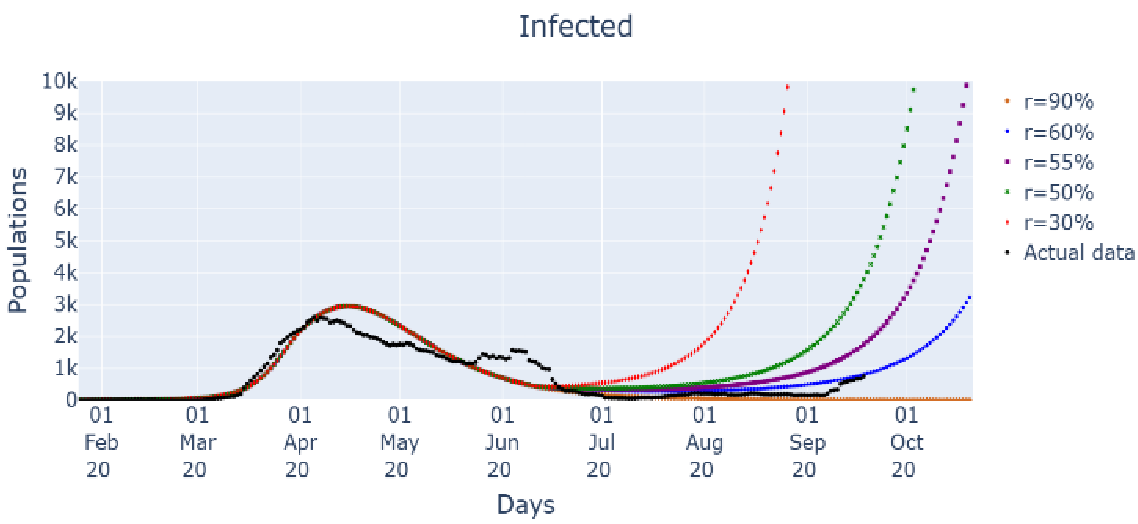

Fig.6. Forecast of COVID-19 cases in Malaysia after the reopening date, 10th June 2020 with the effects of population behavior

Based on Fig. 6, with the higher value of *r*, 90%, which is indicated by the chocolate diamond marker curve, may subside the second wave of COVID-19 in Malaysia, resulting in the endemic equilibrium on 17 October 2020, where the infected (*I*) cases remain approximately constant to 2 infected (*I*) cases for long period of time, while with the lower value of *r*, 30%, this may lead to the occurrence of new wave incidence within a short term where it may infect up to 10,000 population in Malaysia on 28 July 2020 and becoming a catastrophe to every sector of Malaysia includes the essential sector, healthcare system of Malaysia. From this figure, we also observed that the blue curve almost fits the actual data of COVID-19. This means the percentage of the quarantine rules-abiding population after the reopening date, 10 June 2020, is about 60% indicating moderate compliance. This moderate compliance might be referred to the enormous amount of 67,542 compound notices reported by the Malaysia government from March 2020 until the end of the year.

To summarize this section, the population behavior effects upon reopening date of Movement Control Order (MCO) can be studied via our proposed SEIRD model, which considers sporadic cases using different values of *r*, the percentage of the quarantine rules-abiding population. With the above-presented results, the Malaysian government needs to regulate the population behavior as it may be the best strategy to exit the pandemic of COVID-19.

IV. CONCLUSION

A more extended compartmental epidemiological model leads to increasing unknown parameters that must be solved. This study proposed a time-varying coefficient SEIRD model that considers sporadic cases. Intending to fit the time-varying coefficient SEIRD model to the actual data of COVID-19 cases, this study has observed the comparison between four optimization algorithms (Trust-Region Reflective, Levenberg-Marquart, Nelder-Mead, Sequential Linear Squares Programming). The evaluation using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) indicates that Nelder-Mead Algorithm outperforms the other optimization algorithm. Thus, this study proposed Nelder-Mead as the best optimization algorithm among Trust-Region Reflective (TRR), Levenberg-Marquardt (LM) and Sequential Linear Squares Programming (SLSQP) in fitting time-varying coefficient SEIRD model with more extended time-window simulation. The SEIRD model simulations presented in Section 3 were done by using Python programming language via Jupyter Notebook, adopting Python built-in functions which are *odein*t from *Scipy.integrate* package and *minimize* from *Lmfit* package to solve the system of ordinary differential equations (ODEs) and reduce the errors for optimization, respectively. We found that our proposed SEIRD model, which an extended SIRD model is better in predicting COVID-19 cases in Malaysia for longer time-window simulation compared to other study [25], which adopted SIRD model in predicting the COVID-19 spread as we managed to obtained less value of Root Mean Squared Error (RMSE), as low as 347 infected (I) for 238 days simulations. With that, remarks on the importance of considering the exposed (E) compartment in modelling transmission dynamics of COVID-19 as it helps in structuring and formulating a more relevant model. Qualitatively, this study suggests that the percentage of quarantine rules-abiding population, r needs to be remained at a high level, as high as 90%, to ensure Malaysia successfully reaches disease-free or endemic equilibrium so it may halt a leading emergence of any new wave of this infectious disease. Everyone is responsible for keeping themselves and everyone safe by following preventive measures against the transmission of COVID-19, for instance, being aware of social distancing while having a close conversation, avoiding mass gatherings in confined spaces or crowded places and observe on self-hygiene by regularly washing hands.

ACKNOWLEDGMENT

The authors acknowledge the Ministry of Higher Education Malaysia (MOHE) for the financial support. This work is a part of a research project, entitled "Systemic Approach on Book of Life: Methods, Theories and Implications" (Reference Code: FRGS/1/2018/SSI04/UIAM/02/1). They also acknowledge the authors who their thoughtful ideas are benefited in this work.

REFERENCES

[1] Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, *368*(6490): 493-497. https://doi.org/10.1126/science.abb4218.

[2] EPA. (2023, February 1). Indoor Air and Coronavirus (COVID-19). Retrieved from United States Environmental Protection Agency:https://www.epa.gov/coronavirus/indoor-air-and-coronavirus-covid-19x.

[3] Anderson, R. M., May, R.M. (1992). *Infectious diseases of humans: Dynamics and control.* Oxford, UK: Oxford University Press.13. doi 10.1017/s0950268800059896

[4] Salim, N., Chan, W. H., Mansor, S., Bazin, N. E. N., Amaran, S., Faudzi, A. A. M., Zainal, A., H, S. H., Hooi, E. K. J., & Shithil, S. M. (2020). COVID-19 epidemic in Malaysia: Impact of lock-down on infection dynamics. *MedRxiv.* Doi:10.1101/2020.04.08.20057463

[5] Kermack, W. O., McKendrick, A. G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A. 115*(772): 700-721. Bibcode:1927RSPSA.1I15..700K. https://doi.org/10.1098/rspa.1927.0118.

[6] Beckley, R., Weatherspoon, C., Alexander, Michael., Chandler, Marissa., Johnson, A., & Bhatt G. S (2013). Modeling epidemics with differential equations. 1-9. Group Project SIR.

[7] Abuhasel, K. A., Khadr, M., & Alquraish, M. M. (2020). Analyzing and forecasting COVID-19 pandemic in the Kingdom of Saudi Arabia using ARIMA and SIR models. *Computational Intelligence, 2020,* 1-14. https://doi.org/10.1111/coin.12407.

[8] Mahmud, A., Lim, Y. P. (2020) Applying the SEIR model in forecasting the COVID-19 trend in Malaysia: A preliminary study. *Medrxiv.* https://doi.org/10.1101/2020.04.14.20065607.

[9] Gupta, R., Pandey, G., Chaudhary, P., & Pal, S. (2020). Seir and regression model based COVID-19 outbreak predictions in India. https://doi.org/10.1101/2020.04.01.20049825.

[10] Osborn, D. (2020, November 2). *Difference between incubation period and infectious period.* Difference between similar terms and objects. http://www.differencebetween.net/science/difference-between-incubation-period-and-infectious-period/.

[11] Muka, A., & Sannyal, M. (2020). Transmission dynamics of COVID-19 in Bangladesh-A compartmental modeling approach. *Journal of Mathematics in Industry*, https://doi.org/10.21203/rs.3.rs-31194/v1.

[12] Piccolomiini, E. L., & Zama, F. (2020). Monitoring Italian COVID-19 spread by an adaptive SEIRD model. *Medrxiv.* https://doi.org/10.1101/2020.04.03.20049734.

[13] Maugeri, A., Barchitta, M., Battiato, S., & Agodi, A. (2020). Modeling the novel coronavirus (SARS-CoV-2) outbreak in Sicily, Italy. Preprints. https://doi.org/10.20944/preprints202004.0267.v1.

[14] COVID-19 Database in Github MOH [https://github.com/MoH-Malaysia/covid19public/blob/main/epidemic/cases_malaysia.csv].

[15] Jamil, N. M., Rosli, N., & Muhammad, N. (2021). Simulation of COVID-19 outbreaks via Graphical User Interface (GUI). *Journal of Public Health Research.* https://doi.org/10.4081/jphr.2021.2130

[16] Gotz, T., & Heidrich, P. (2020). Early stage COVID-19 disease dynamics in Germany: models and parameter identification. *Journal of Mathematics in Industry, 10*(1), 1-13. https://doi.org/10.1186/s13362-020-00088-y.

[17] Korolev, I. (2020). Identification and estimation of the SEIRD Epidemic Model for COVID-19. *Journal of Econometrics*, *220*(1), 63-85. Doi: 10.1016/j.jeconom.2020.07.038.

[18] Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S. M., Lau, E. H. Y., Wong, J. Y., Xing, X., Xiang,

N., Wu, Y., Li, C., Chen, Q., & Fheng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*. Doi: 10.1056/NEJMoa2001316

[19] Read, J. M., Bridgen, J. R., Cummings, D. A., Ho, A., & Jewell, C. P. (2020). Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. *medRxiv.*

[20] Malhotra, S., Mani, K., Lodha, R., Bakhshi, S., Mathur, V. P., Gupta, P., ... & COVID Reinfection AIIMS Consortium. (2022). SARS-CoV-2 reinfection rate and estimated effectiveness of the inactivated whole virion vaccine BBV152 against reinfection among health care workers in New Delhi, India. *JAMA Network Ope*n, *5*(1), e2142210-e2142210.

[21] Scipy. Integration (scipy.integrate) — SciPy v1.6.0 Reference Guide. (2008). SciPy.Org. https://docs.scipy.org/doc/scipy/reference/tutorial/integrate.html#mol.

[22] Department of Statistics Malaysia (2020, July 15). Current Population Estimates, Malaysia, 2020. Department of Statistics Malaysia Official Portal. Retrieved June 15, 2022, from https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=155&bul_id=OVByWjg5YkQ3MWFZRTN5bDJiaEVhZz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09.

[23] Branch, B., A., Coleman, T., F., & Li, Y. (1999). A subspace, Interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal on Scientific Computing*, 21(1), 1-23.

[24] Gavin, H., P. (2019). The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. Unpublished.

[25] Lagarias, J., Reeds, J., & Wright, M., & Wright, P. (1997). Convergence properties of the Nelder-Mead Simplex Algorithm in low dimensions.

[26] Jorge, N., & Stephen, J., W. (2006). *Numerical optimization.* Springer.

[27] Saleh, C., Dzakiyullah, N. R., & Nugroho, J. B. (2016). Carbon dioxide emission prediction using support vector machine. *IOP Conference Series: Materials Science and Engineering*, *114*(1). https://doi.org/10.1088/1757-899X/114/1/012148.

[28] Anastassopoulou, C., Russo, L., Tsakris, A. & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PloS One, 2020,* 15:e0230405

[29] Fishbein, E., & Hkawng, J. T. (2020, June 2). Immigration detention centres become Malaysia coronavirus hotspot. Human Rights News | Al Jazeera. Retrieved February 3, 2023, from https://www.aljazeera.com/news/2020/6/2/immigration-detention-centres-become-malaysia-coronavirus-hotspot.