Mathematical Modelling of the Spread of COVID-19 Cases in Sri Lanka

Silva KKTS*, Dharmawardane PMN

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka

*Email: ktshanika@gmail.com

Abstract – Sri Lanka confronted a substantial challenge in responding to the spread of the novel strain of coronavirus (SARS-CoV-2), which was declared as a pandemic by the World Health Organization. The government introduced multiple measures to contain the transmission of the virus within the country. This research provided information on COVID-19 and estimated the spread of the virus in Sri Lanka based on confirmed data from the national epidemiological reports for the period from October 01, 2020 to April 10, 2021. Stochastic modelling approaches were used to model the spread of the virus and its significance was also checked. The logistic growth model, where the number of infected cases are stochastic and the Susceptible-Infected-Recovered (SIR) model where the number of individuals who get disease, the number of deaths due to Coronavirus, and the number of recovered are stochastic, were employed under certain assumptions to identify the behavior of the spread of virus in a theoretical approach, during the second wave. The SIR model with a hyperparameter was used to improve the model fitting to actual data. Further, different infection and recovery rates were obtained using different time-variant scenarios. Results showed that the estimated cumulative cases were equal to 96845 within a certain period. It was also shown that both infection and recovery rates changed with time and followed different patterns. Due to changes in the COVID – 19 variants, the infected individuals also demonstrated different behaviors with increasing infection rate.

Keywords - COVID-19, Logistic growth model, Reproduction number, SIR model

1. INTRODUCTION

A novel strain of coronavirus (SARS-CoV-2) was declared as a pandemic on March 11, 2020, by the World Health Organization [8]. COVID-19 was first detected in Wuhan, a city in China's Hubei Province. The virus has now spread over two hundred nations and territories worldwide. Because of this pandemic, there are more than 135 million reported cases worldwide, including more than three million deaths (as of April 10, 2021).

On January 27, 2020, the first confirmed case of a Corona virus-infected person was reported in Sri Lanka Since then, there are more than 94 thousand reported cases, including 595 deaths (as of April 10, 2021). Social distancing, self-quarantine, and wearing a face mask have emerged as the most commonly used technique for minimizing the pandemic before finding a vaccine to control this pandemic. Also, Polymerase Chain Reaction (PCR) tests and Rapid Antigen Tests are done by the

ministry of health to confirm the infected and make them hospitalized to slow down the spread of the virus. As in other pandemics, researchers are interested in finding the start of COVID-19, its progress, and measures that can control this pandemic. In that case, mathematical models are used to estimate the quantitative behavior of the disease, such as transmission, recovery, deaths, and other significant parameters. Real incidence datasets from different countries which have different characteristics are used by researchers for developing mathematical models for COVID-19. These models are leads to consider the different patterns of COVID-19 spread in countries. Therefore, these models give information about how to control this kind of pandemic.

The second wave of COVID-19 in Sri Lanka, which ran from October 01, 2020 to April 10, 2021, was considered in this study. To identify the COVID-19 spread in Sri Lanka, two mathematical models were used: the logistic growth model and the

Susceptible-Infected-Recovered (SIR) model. This SIR modeling, approach requires less data with reduced computational requirements which makes it particularly more useful for an early-stage epidemic estimation [1]. Using these both models, predictions were obtained and compared them with original cases. As the parameters related to the spread of COVID-19, the number of susceptible, infected, and recovered individuals were considered and they were issued by the Epidemiology unit in Sri Lanka [6]. The model was fitted to the data using the nonlinear fittings, and it was used to find the spread of the virus in Sri Lanka.

2. MATERIALS AND METHOD

COVID-19 data from October 2020 to April 10, 2021 in Sri Lanka issued by the Epidemiological Unit, Sri Lanka were considered in this research [6].

2.1 Logistic growth model

A logistic growth model for finding the growth of an epidemic under the assumption that the rate of change in the number of new cases per capita linearly decreases with the number of cases [3]. The logistic model is given by,

$$\frac{dC}{dt} = rC\left(1 - \frac{C}{K}\right), \quad (1)$$

where C, r and K denote number of infected cases, infection rate and the final size, respectively. Given that the initial number of cases is $C(0) = C_0$, the number of infected cases (solution of (1)) is defined as,

$$C = \frac{K}{1 + \left(\frac{K - C_0}{C_0}\right)e^{-rt}} \quad (2)$$

Equation (2) is suitable to model the COVID -19 outbreak since there is an exponential growth in the beginning, but a decreasing growth at a later stage. When $\frac{d^2C}{dt^2} = 0$, the growth rate, $\frac{dC}{dt}$, reaches to its maximum. Therefore, the growth rate peaks at the time t_p is given by

$$t_p = \frac{\ln\left(\frac{K - C_0}{C_0}\right)}{r},$$

At time t_p, the number of cases is defined as

$$C_p = \frac{K}{2}.$$

At time t_p, growth rate is defined as,

$$\frac{dC}{dt_p} = \frac{rK}{4}.$$

To estimate parameters K and r, *NonlinearModelFit* function in *MATHEMATICA* software was used with total infect cases data [2]. *ParameterTable* function was used to check the nonlinear model validation with 95% confidence level. The function carried out the F-Test to check the significance of the model and hypothesis are,

- H$_0$: the coefficient of the parameters equals to 0
- H$_1$: the coefficient of the parameters not equals to 0

2.2 SIR model

When building up the SIR model, the population is divided into three classes labeled by S, I and R. Let $S(t)$ denote the number of individuals who are susceptible to the disease at time t (measured usually in days), $I(t)$ the number of infected individuals and $R(t)$ the number of individuals who have been recovered from the disease. To formulate a model, there are some assumptions, such as assuming infected are able to spread the disease by contact with susceptible and the total population (N) remains constant [7]. The total population size is the sum of the sizes of these three classes:

$$N = S(t) + I(t) + R(t).$$

The original SIR model is given by the following system of ODEs

$$S'(t) = -\beta IS$$
$$I'(t) = \beta IS - \alpha I,$$
$$R'(t) = \alpha I,$$

where α and β are recovery rate and infection rate, respectively. SIR model is particularly useful in estimating the maximum number of active cases and the peak time of attaining it.

There are some assumptions that were used in this study to check the behavior of COVID-19 spread in Sri Lanka. Main assumptions are recovers are not infected again with the virus, deaths are considered as recovers, total population is constant. The second wave of the virus is considered as the start of the spread of COVID-19 in Sri Lanka (*i.e.*, the data before the second wave is not considered for total infect cases).

The following SIR model is proposed for our study to describe the actual data:

$$\left.\begin{array}{l} S'(t) = -\beta IS - C, \\ \boldsymbol{I'(t) = \beta IS - \alpha I,} \\ \boldsymbol{R'(t) = \alpha I + C.} \end{array}\right]$$

where C is the hyperparameter.

Next we consider the basic reproduction number, $\mathcal{R}_0$, which plays a great role in a pandemic. It is important to know $\mathcal{R}_0$ as early as possible and it can be considered to find the solutions for the SIR model. This basic reproduction number gives the number of secondary cases one infectious individual will produce in a population consisting only of susceptible individuals. In other words, we assume that at early stages, $S(t) \approx S(0) \approx N$. Then $\mathcal{R}_0$ is given by,

$$\mathcal{R}_0 = \frac{\beta S(0)}{\alpha} \approx \frac{\beta N}{\alpha}$$

Here if infection rate (β) > recovery rate (α), then $\mathcal{R}_0 > 1$. This implies that the disease remains in the population. On the other hand, if $\beta < \alpha$ then $\mathcal{R}_0 < 1$. This implies that the disease disappears from the population.

To adjust the numerical solution to the proposed SIR model given in (3), the initial conditions are considered in the range [0, 1] and using a scaling factor they are scaled to fit with recorded data for visual test [4]. A constant value (hyperparameter) was also introduced to the SIR model to adjust the model graph function, which resulted in a significant fit with the actual data. The initial conditions for the three variables, *S, I* and *R,* were

chosen considering the most individuals are susceptible and few individuals are infected while no one recovered at the beginning. Consequently, the parameters α and β are constant values and scaling factors are adjusted manually to fit the model to recorded data as best as possible, based on a trial-and-error approach and visual inspections using the *MATHEMATICA* function *Manipulate.* In here, both α and β ranges are selected according to the infection period and reproduction number respectively, which was estimated by WHO [9].

3.   RESULTS AND DISCUSSION

The logistic model fits the confirmed cumulative cases in Sri Lanka from October 01, 2020 to April 10, 2021, and estimated parameters and their standard error and P-value are shown in Table 1. The result of the fitting model to actual data shows that R-Squared is 0.9988 which shows that 99.98% of the data fit to the model. P-values of the all the parameters are less than 0.05, and it indicates that all the parameters are statistically significant. The results of the regression of the logistic model and actual total cases are shown in Figure 1, and it shows a good agreement between the model and actual data. Estimated parameters in the logistic model show that the peak number of cases (K) as 96845 cases when time (t) is 105.44 days.
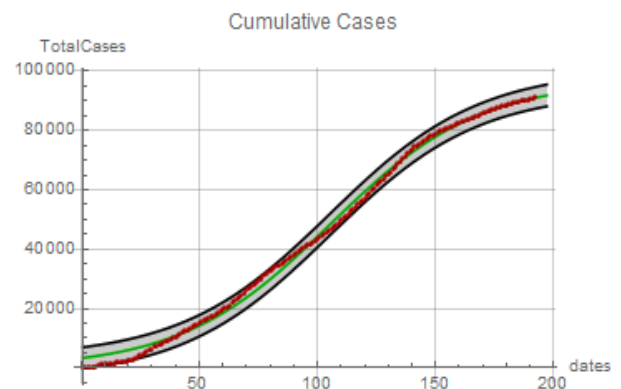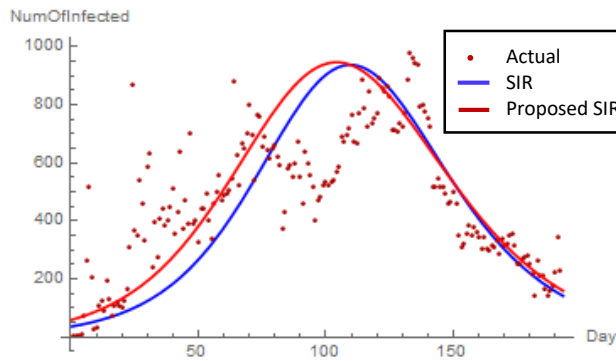

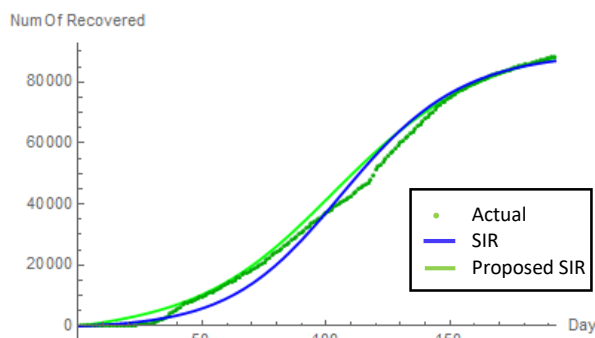**Figure 1: Confirmed and predicted total cases in Sri Lanka from October 01, 2020 to April 10, 2021.**

Root Mean Squared Error: 1.80742e+03.
R-Squared: 0.998875.

Table 1: Estimated results of the logistics model

	Estimate	Standard Error	t-Statistic	P-Value
K	96845.7399	658.0206	147.1774	6.4886e-197
r	0.0318	0.0004	76.1309	9.1972e-144
t_p	105.4457	0.6361	165.7780	1.3132e-206

Figure 4: Confirmed and predicted cumulative recoveries.

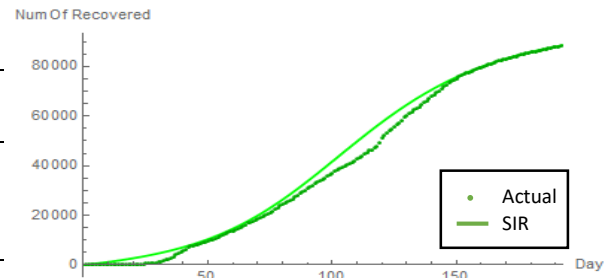
Figure 2: A comparison between the prediction curves of the original SIR model and proposed SIR model for daily infected.

Figure 5: Confirmed and predicted daily infected cases in Sri Lanka from October 01, 2020 to April 10, 2021.

The proposed SIR model for the second wave of COVID-19 in Sri Lanka with the actual data scenarios is shown in Figures 4 and 5, which are about cumulative recoveries and daily infected cases, respectively. This is used to illustrate the prediction of the cumulative recoveries and the rate of infection peaks. According to these figures, the infection and recovery rates are $\beta = 0.143$ and $\alpha = 0.1$ respectively, and it shows that reproduction number R_0 is equal to 1.43 during this time period. According to the previous studies, reproduction number for COVID-19 for early stage in Sri Lanka was vary between 0.93 and 1.23 while WHO estimated reproduction number was in the range 1.4 and 2.5 [5]. Therefore 1.43 can be considered as good reproduction number for second wave in Sri Lanka. From Figure 4, the total recoveries on the last day of this period are equal to 88341 while the actual value is 88226. When the hyperparameter C takes the value 0.0004, the best match of the model with actual data of daily infection is discovered, indicating that the model utilized here is quite close
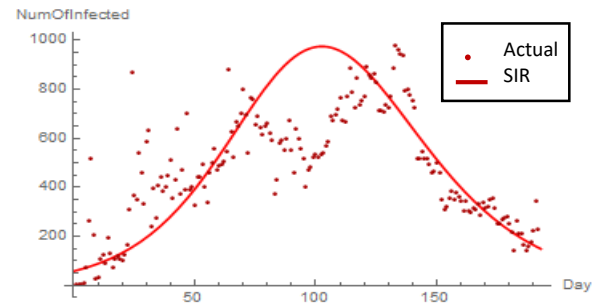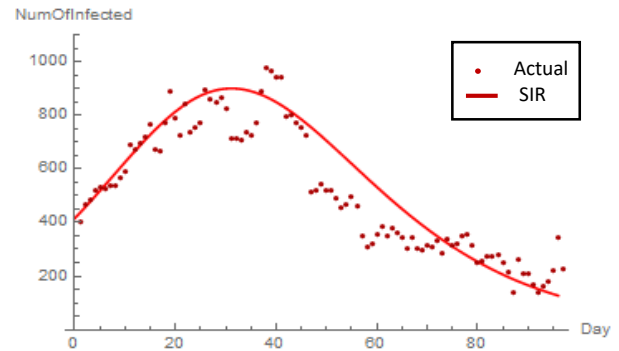

Figure 3: A comparison between the prediction curves of the original SIR model and proposed SIR model for cumulative recoveries

Figure 2 and 3 show the difference between original SIR model and proposed SIR model with actual data. It is evident that using a hyperparameter the variation between actual curves and prediction curves can be reduced. Therefore, the proposed SIR model is suitable for fitting actual data.

to the original SIR model. As a result of Figure 5 shows that the estimated maximum daily infected cases are equal to 980 while the actual maximum cases are 976. The estimated and actual observations of some peak points are shown in Table 2.
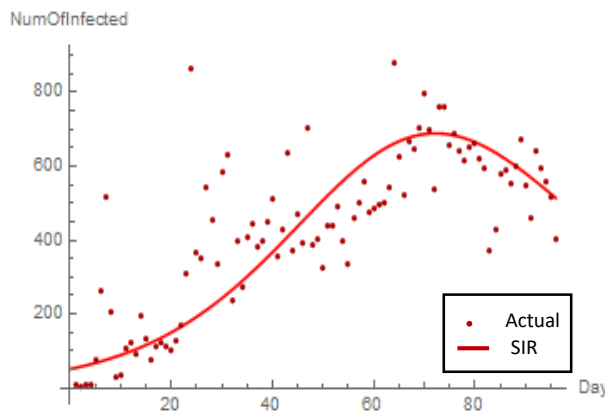
Table 2: Confirmed and predicted daily infected individuals.

Day	Confirmed Daily infected cases	Predicted Daily infected cases	Error Percentage
17	121	118	2.4
49	401	386	3.7
67	669	642	4.0
73	762	732	3.9
114	887	945	6.5
117	843	922	9.3
121	892	883	1.0
124	864	851	1.5
167	331	330	0.3
185	211	192	9.0

Figure 6: Confirmed and predicted daily infected cases in Sri Lanka from October 01, 2020 to January 04, 2021.

Figure 6 shows first curve which is from October 01, 2020 to January 04, 2021 time period daily infected, and the highest predicted infected cases are estimated to be 689 with $\beta = 0.15$ and $\alpha = 0.093$. The rest of the time

Figure 7: Confirmed and predicted daily infected cases in Sri Lanka from January 04 to April 10, 2021.

period which is follow second curve pattern in daily infected are shown in Figure 7, and it gives maximum infected cases as 900. In addition, the infection and recovery rates are $\beta = 0.16$ and $\alpha = 0.098$, respectively. This shows that with the time both infection and recovery rates are increased for the same hyperparameter 0.0004 in both cases.

4. CONCLUSION

In this study, stochastic approaches were carried out to identify the COVID-19 spread in Sri Lanka. It is based on the data from October 01, 2020 until April 10, 2021. This study shows that the spread of the virus using the SIR model, and the total recoveries and maximum daily infected were found to be 88341 and 980, respectively. The mean reproduction number was estimated to be 1.43, with the infection and recovery rates are $\beta = 0.143$ and $\alpha = 0.1$, respectively. But there are also some peak values of daily infected that varies differently from the common pattern that does not follow the mean infection and recovery rate.

The study further considered the epidemic peak values of daily infected with time. It can be observed that the peak values of infected individuals from January to March 2021 have

become higher than the time range from October to December 2020. Also, both infection and recovery rates are increased in this time range. The changes of the COVID-19 virus in different ways and created different variants are also affected to the changes of the infection rate and recovery rates.

Adding suitable hyperparameter to the SIR model it can be tuned up to a better model that describes the actual data. The peak number of cumulative cases in this period was estimated based on the logistic growth model and was found to be 96845 at the time 105(days). All the parameters in logistic growth model are statistically significant according to their P-values (<0.05) and this model shows 99.98% accuracy.

5. REFERENCES

1. A. M. Almeshal, A. I. Almazrouee and M. R. Alenizi*, Forecasting the Spread of COVID-19 in Kuwait Using Compartmental and Logistic Regression Models,* Appl. Sci.*, **10** (2020), 3402- 1-18

2. M. Bahrami, and B. Wood, *Logistic growth model for COVID-19*, Retreived from www.wolframecloud .com/ obj/covid-19/Published/ Logistic-Growth-Model-for-Covid-19.nb (April 28, 2021)

3. M. Batista, *Estimation of the final size of the COVID-19 epidemic*, 2020 *medRxiv:* 2020.02.16.20023606v5, DOI: https://doi.org /10.1101/ 2020.02.16.20023606

4. I. Cooper, A. Mondal and C. G. Antonopoulos, *A SIR model assumption for the spread of COVID-19 in different communities*, Chaos-Solitons-Fractals, **139** (2020) 110057. DOI: https://doi.org/ 10.1016/j.chaos.2020.110057

5. S. Dharmaratne, S. Sudaraka, I. Abeyagunawardena, K. Manchanayake, M. Kothalawala, and W. Gunathunga, *Estimation of the basic reproduction number (R_0) for the novel coronavirus disease in Sri Lanka*. Virol. J., **17**

(2020) no. 1, 1–7. DOI: https://doi.org/ 10.1186/s12985-020-01411-0

6. Epidemiology Unit Sri Lanka, *Corona Virus 2020/2021*. Epidemiology Unit Sri Lanka, 2020

7. M. Martcheva, *An Introduction to Mathematical Epidemiology*, Springer, 2020

8. World Health Organization, *Statement on the meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)*, 2020

9. World Health Organization, Coronavirus disease (COVID-19) pandemic, World Health Organization, 2021