
Faculty of Science

Faculty Publications

Mathematical modeling approach to predict COVID-19 infected people in Sri Lanka

Premarathna, I. H. K., Srivastava, H. M., Juman, Z. A. M. S., AlArjani, A., Uddin, M. S., & Sana, S. S.

2022

© 2022 Premarathna, I. H. K., Srivastava, H. M., Juman, Z. A. M. S., AlArjani, A., Uddin, M. S., & Sana, S. S. This is an open access article distributed under the terms of the Creative Commons Attribution License. <https://creativecommons.org/licenses/by/4.0>

This article was originally published at:

<https://doi.org/10.3934/math.2022260>

Citation for this paper:

Premarathna, I. H. K., Srivastava, H. M., Juman, Z. A. M. S., AlArjani, A., Uddin, M. S., & Sana, S. S. (2022). "Mathematical modeling approach to predict COVID-19 infected people in Sri Lanka." *AIMS Mathematics*, 7(3), 4672-4699.

<https://doi.org/10.3934/math.2022260>



Research article

Mathematical modeling approach to predict COVID-19 infected people in Sri Lanka

I. H. K. Premarathna¹, H. M. Srivastava^{2,3,4,5}, Z. A. M. S. Juman⁶, Ali AlArjani⁷, Md Sharif Uddin^{7,8} and Shib Sankar Sana^{9,*}

¹ Department of General Studies, Academic Wing, Sri Lanka Military Academy, Diyatalawa, Sri Lanka

² Department of Mathematics and Statistics, University of Victoria, Victoria, BC V8W 3R4, Canada

³ Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan

⁴ Department of Mathematics and Informatics, Azerbaijan University, 71 Jeyhun Hajibeyli Street, AZ1007 Baku, Azerbaijan

⁵ Section of Mathematics, International Telematic University Uninettuno, I-00186 Rome, Italy

⁶ Department of Mathematics, Faculty of Science, University of Peradeniya, Peradeniya, 20400, Sri Lanka

⁷ Department of Industrial Engineering, Prince Sattam bin Abdulaziz University, AlKharj, KSA, 16273

⁸ Department of Mathematics, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh

⁹ Kishore Bharati Bhagini Nivedita College, Ramkrishna Sarani, Behala, Kolkata 700060, India

* **Correspondence:** Email: shib_sankar@yahoo.com.

Abstract: The novel corona virus (COVID-19) has badly affected many countries (more than 180 countries including China) in the world. More than 90% of the global COVID-19 cases are currently outside China. The large, unanticipated number of COVID-19 cases has interrupted the healthcare system in many countries and created shortages for bed space in hospitals. Consequently, better estimation of COVID-19 infected people in Sri Lanka is vital for government to take suitable action. This paper investigates predictions on both the number of the first and the second waves of COVID-19 cases in Sri Lanka. First, to estimate the number of first wave of future COVID-19 cases, we develop a stochastic forecasting model and present a solution technique for the model. Then, another solution method is proposed to the two existing models (SIR model and Logistic growth model) for the

prediction on the second wave of COVID-19 cases. Finally, the proposed model and solution approaches are validated by secondary data obtained from the Epidemiology Unit, Ministry of Health, Sri Lanka. A comparative assessment on actual values of COVID-19 cases shows promising performance of our developed stochastic model and proposed solution techniques. So, our new finding would definitely be benefited to practitioners, academics and decision makers, especially the government of Sri Lanka that deals with such type of decision making.

Keywords: COVID-19; stochastic model; model validation method; SIR and Logistic growth models
Mathematics Subject Classification: 92B05

1. Introduction

Better health leads to greater enjoyment of life. Nowadays, health has become an extremely vulnerable factor for all general public due to the COVID-19 outbreaks. The novel coronavirus disease (COVID-19) outbreaks is originated in December 2019 in Wuhan, China. The global growth of COVID-19-infected patients is gradually increasing and it is an outstanding rate compared with Sri Lanka. This situation is further exacerbated due to the mal practicing of health precautions. Better prediction of COVID-19 cases becomes a key factor for government's decisions.

1.1. Importance and research gap

The COVID-19 virus transmits from one individual to another quite rapidly in many communities. Due to its randomly changing behavior, it can be considered as dynamic system. These uncertain and unsteady patterns are difficult to understand. Numbers of infected patients have been observed to change unpredictably within a tiny amount of time. It can be considered a dynamical system in place of the Markov chain approach. This situation cannot be described using ordinary differential equations [1]. For that reason, stochastic differential equations can be used to describe the behavior of the virus. These stochastic models are produced by considering major factors such as the infected people and the cured ones from the COVID-19 virus. Random behavior of the virus as well as any random situation can be analyzed by using stochastic differential equations.

If government identifies the behavior of the virus, then following safety precautions can be arranged to minimize the death roll of the country: (1) Aggressive "social distancing" measures; (2) Issuing travel bans to other affected countries; (3) Island wide strict curfews; (4) Establishment of a COVID-19 virus Task Force, etc. In fact, preventive methods can be identified to reduce the factors which weaken the development of the country. These predictions can be used to maintain the economic growth rate of the country.

It is very important to understand the pattern of the variable to predict the future. Most of the share markets, global currency markets are fitted well with this stochastic one. Most of the investors are blind about the randomness. Because, small change will make a huge difference in the chaotic world. This research leads to identify the patterns of the chaotic situation like stock market by using stochastic differential equations. It benefits for investors to take necessary actions before they capitalize money.

It would be the most appropriate time to make decisions by the industries to function the company without failure in the pandemic situation. This prediction paves the way to the well-being of the employees of the particular companies and their families during the pandemic situation. Although conventional mathematical models are attracted by many researchers, work on this COVID-19 virus by stochastic model is scant. Besides, in practice, behaviors of any viruses are random. In this situation, none of the existing conventional mathematical models can be applicable. So, development of stochastic differential equations to describe the Covid-19 spread (especially to estimate the COVID-19 infected people in Sri Lanka) is desirable.

1.2. Contributions and organization

This research attempts to predict coronavirus cases in a community by considering stochastic differential equations. This study contributes for the Armed forces, health sector that battle against spreading of the virus. Moreover, epidemic diseases analysts and others who draw the future patterns of the virus are also benefited.

The proposed prediction model based on SDE, SIR and Logistic growth provides policy makers an evidence based scientific method for identifying suppression measures that can be relaxed at the most appropriate time.

Subsequently, the key contributions of the proposed article are of the following seven folds:

- Investigating first wave of COVID-19 outbreak and proposing a stochastic forecasting model to estimate the number of COVID-19-infected patients in the future.
- The accuracy of the model in forecasting is validated by existing methods in the literature.
- Achieving a solution technique to efficiently solve the proposed stochastic model.
- Investigating second wave of COVID-19 outbreak using two existing models (SIR model and Logistic growth model) for estimating the number of COVID-19-infected patients in the future.
- Developing another solution technique to efficiently solve two existing models.
- Solution techniques are validated via secondary data obtained from the Epidemiology Unit, Ministry of Health, and Sri Lanka.
- Finally, showing the merits and capabilities of the newly proposed solution methods in dealing with the novel COVID-19 scenario to make better decisions, especially for governments and scientific community, to prevent the spread of COVID-19 from the country.

The rest of the paper is organized as follows: Section 2 reviews the related work. Basic concepts of the Markov Chain, Martingale, Euler–Maruyama method and Logistic growth and Susceptible–Infected–Recovered (SIR) Models are described in Section 3, while Section 4 provides the methodology. The results and discussion are represented in Section 5. Finally, concluding remarks along with limitations and future research potential are drawn in section 6.

2. Related work

Over years, researchers have developed various types of mathematical models to recognize the pattern of random behavior of virus outbreaks. For a better understanding, we have the following subsections such as SIR models, SDE models, Virus behavior and other models.

2.1. SIR models

SIR model can be considered as a traditional mathematical model in the Mathematical and Statistical analysis. Rihan et al. [2] developed epidemic SIR model with long-range temporal memory. This model used a suitable Lyapunov functional to formulate the global stability of the steady states. According to the findings, the combination of fractional-order derivative and time delay in the model improve the dynamics and increase the complexity of the model. Sharma et al. [3] presented the SEIRD (Susceptible–Exposed–Infected–Recovered–Death) pandemic spread model with a time delay on the heterogeneous population. According to their work they have verified total infection cases in India considering the effect of nationwide lockdown at the onset of the pandemic. They have forecasted the total number of infection cases in two extreme situations of nationwide no lockdown and strict lockdown scenarios.

Guan et al. [4] identified 63 studies regarding modeling the transmission dynamics of COVID-19 epidemic. Epidemiological parameters estimation, trend prediction, and control measure evaluation were studied from this research. Din et al. [5] used an epidemic model composed on four compartments: susceptible, exposed, infected and recovered (SEIR), which were used to describes the dynamics of COVID-19 under convex incidence rate. They simulated the results by using nonstandard finite difference method (NSFDS) which is a powerful numerical tool. The new model on some random data and then by the available data of a particular regions of subcontinents were described.

Mello et al. [6] studied the key feature to the SIR model for various epidemics and given a brief conversation about its mathematical description. Numerous forms of mathematical models related to the containment, risk analysis, and features of COVID-19 were investigated by Adekola et al. [7]. Hajji et al. [8] considered a system of delay differential equations as a model for the dynamics of tumor immune system interaction. They carried out a stability analysis of the proposed model and showed that the system could have up to two steady states. They also determined an upper bound for the delay, such that stability was preserved. Numerical simulations of the system under different parameter values were performed.

2.2. SDE models

In practice, the spreading of virus follows a random process. Consequently, appropriate stochastic mathematical tools are urgent to study the random behavior of the virus. Most of the existing researches on predicting the virus' spread have used geometric Brownian motion with Ito's Lemma approach (refer to Allen [1]). Further, most of the existing works on the COVID-19 virus have been carried out using ODE or PDE (refer to Kucharski [9]).

Allen [10, 11] conducted a study to identify the factors and mathematical formulas which describe the stochastic differential equation (SDE) model for the wave of coins in circulation. Allen [12] found that the SDE model is suitable for environmental variability and mean-reverting processes. The mean-reverting processes possess several important features that better characterize environmental variability in biological systems than a linear function of white noise. Allen [1] investigated the topic of stochastic epidemic modeling from the perspective of an ordinary differential equations framework, with an emphasis on continuous-time Markov chains and stochastic differential equations. This work is more important to describe the COVID-19 behavior [13–15]. Bukiet and Ovens [16] found similarities in the cricket batting order of the ODI with the SDE model. While scoring strategies and

player performance in cricket had been studied, there were very few publications on the influence of batting order. This research is more helpful to understand the SDE applications.

2.3. *Virus behavior*

The organisms of all the viruses are similar in structure. But the effect of the different types of viruses on human being is different. Zika is a virus transmitted to humans through either the bites of infected mosquitoes or sexual transmission. Zika has been linked to congenital anomalies such as microcephaly. Spreading of this virus is simulated by using a stochastic method to identify the pattern (Agusto et al. [17]). Eltoukhy et al. [18] demonstrated data analytic approach to predict the spread of COVID-19 cases in top affected countries. Daniel [19] used digital tracing to consider variations in test sensitivities in a virus. Rao and Brandeau [20] developed analytical conditions showing the optimal allocation for four objectives on viruses. It is a vital research to comprehend the compartment of the viruses. Ciupe and Heffernan's [21] models had been used to gain new knowledge on the pathogenesis of in-host diseases (including characteristics of the pathogen lifecycle, information on medicine and public health, and aiding the development of effective drug therapies and vaccines); this was used to understand the spread of disease by differential equation.

2.4. *Other models*

Antonietta et al. [22] extended the algorithm for automatically simplifying systems of initial value ODEs. Khan et al. [23] investigated the existence of results and stability analysis for a nabla discrete ABC-fractional-order COVID-19 model. The existence and uniqueness theorems along with Hyers–Ulam stability were used to analyze the model. Sindhu et al. [24] developed the Gumbel type-II model to predict the percentage of patients from the pandemic. Lautaro et al. [25] studied the impact of a massive testing strategy with a threshold. Stefanie et al. [26] presented a hierarchy of mathematical models for spatiotemporal population dynamics using Galerkin projection and population scaling for model reduction. Ting-Yu et al. [27] used the Laplace-transformed likelihood method for multistate disease epidemiology, which estimates the parameters with it and dispenses with the individual time-stamped data. Singh et al. [28] investigated a dynamical systems of fractional-order corona model showing the efficiency of the model regarding CPU time. Srivastava et al. [29] presented a novel approach of fractional derivative to comprehensive study of transmission phenomena of dengue infection. Srivastava et al. [30] found out an effective power series solution of compartmental epidemiological models. Henry et al. [31] evaluated the discriminative ability of hematologic, biochemical and immunologic biomarkers in patients with and without the severe or fatal forms of COVID-19. The risk stratification models for predicting severe and fatal COVID-19 were identified. Bhattacharjee et al. [32] utilized a modified van der Pol symmetric oscillator equation for implementing the series connection of two same type conduction cells of the cardiovascular system that helped to describe the complete process of conduction.

3. Preliminaries

In this section, following the references [33, 34], the basic definitions of the Markov Chain and Martingale are described first. Then, the Euler–Maruyama method of finding the numerical solution

to the stochastic differential equations (SDEs) is presented. Thereafter, the two existing models, Logistic growth and Susceptible–Infected–Recovered (SIR), are provided.

3.1. Markov chain

Let $\{X_0, X_1, X_2, \dots\}$ be a sequence of discrete random variables. Then $\{X_0, X_1, X_2, \dots\}$ is a Markov chain if it satisfies the Markov property: $P(X_{t+1} = S | X_t = S_t, \dots, X_0 = S_0) = P(X_{t+1} = S | X_t = S_t) \forall t \in \mathbb{N}$ and for all states $S \in \{S_0, S_1, \dots, S_t\}$.

3.2. Martingale

Following [34], Discrete-time martingale is a discrete-time stochastic process (i.e., a sequence of random variables) $\{X_1, X_2, \dots, X_t\}$ that satisfies for any time t ,

$$E(|X_t|) < \infty \text{ and}$$

$$E(X_{t+1} : X_1, X_2, \dots, X_t) = X_t$$

Here, the conditional expected value of the next observation given all the past observations is equal to the most recent observation.

3.3. Euler–Maruyama method

The simplest computational method for the approximation of ordinary differential equations is Euler's method [35]. The Euler–Maruyama method [36] is an analogue of the Euler method for ordinary differential equations. To find the numerical solution to stochastic differential equations (SDEs), the Euler–Maruyama method is used. Although the Euler method for ordinary differential equations has order 1, the strong order for the Euler–Maruyama method for stochastic differential equations is $1/2$. This fact was proved in Gikhman and Skorokhod [36].

To develop an approximate solution on the given interval,

$$t_0 < t_1 < t_2 < \dots < t_n$$

Approximation of X

$$X_0 < X_1 < X_2 < \dots < X_n$$

will be determined at the respective t point,

$$dX(t) = \alpha(t, X)dt + \sigma(t, X)dW_t$$

where

$$X(t) = X_t$$

(One of the simplest numerical approximations for the SDE is the Euler-Maruyama method. If Ito's formula of the stochastic Taylor series is truncated after the first order terms, the Euler method or Euler-Maruyama method is obtained.)

Approximate solution of the above is as follows:

$$X_{t+1} = X_t + \mu X_{t-1} \Delta t + \sigma X_{t-1} \eta \sqrt{\Delta t};$$

$$\Delta t = t_{i+1} - t_i;$$

3.4. Logistic growth and Susceptible–Infected–Recovered (SIR) Models

3.4.1. The Logistic growth model

The Logistic growth model of the study indicates an exponential growth at the beginning of the epidemic, followed by a steady increase and finally ending with a declining growth rate. The Logistic growth model is a regression model that is widely used in epidemiology mathematical models to estimate the growth and decline rate of pathogens. The model assumes an exponential curve. The logistic model is presented by [37] as an equation,

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

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

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

and the maximum growth rate peaks at the time are

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

3.4.2. Susceptible–Infected–Recovered (SIR) model

The SIR model [38–41] has been widely used in literature and by policymakers to predict the spread of many diseases. The basic SIR model divides the population into three sections:

S : Susceptible healthy population who are at risk of getting infected;

I : Infected populations who have symptoms, whether mild or severe;

R : Recovered population who have recovered from the disease and gained immunity;

D : Death population.

Thus, the total population can be depicted as in Figure 1. i.e., $N = S + I + R + D$.

The model estimates the rate of change of the susceptible, infected, recovered and death population by differential equations, respectively, as follows:

$$\frac{dS}{dt} = -\beta I \frac{S}{N} ; \frac{dI}{dt} = \beta I \frac{S}{N} - \gamma I ; \frac{dR}{dt} = \gamma I ; \frac{dD}{dt} = \phi I$$

with the parameters defined as follows:

β : the infection rate, defined as the proportion of the infected population per day;

γ : the recovery rate, defined as the proportion of the recovered population per day;

\emptyset : death rate—defined as the proportion of the death population per day.

Using the infection and recovery rate parameters, the SIR model estimates the reproduction number of the disease R_0 as the ratio $R_0 = \frac{\beta}{\gamma}$.

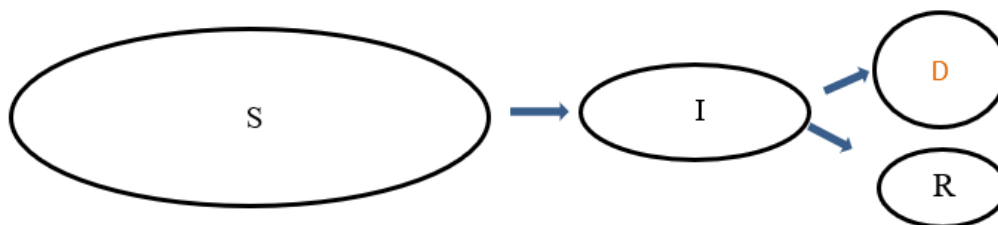


Figure 1. Schematic diagram of the model.

4. Methodology

This section reviews the methodology of the study for both the first and second waves of COVID-19 in Sri Lanka. It covers the hypothesis building, development of the stochastic differential equation (SDE) models, data collection (identification of the population for the study and sample) and data analysis and interpretation. Microsoft Excel (2013) and SPSS software are used to analyze and represent the outcome. This section presents assumptions and notations, the SDE–Virus spreading model formulations and two new solution approaches.

4.1. Assumptions and notation for the SDE–Virus spreading model formulations

Most of the predictions are based on assumptions that may be considered for a better understanding of the mathematical modeling. To develop the models for predicting numbers of the first and second waves of COVID-19 cases in Sri Lanka, we respectively adopt the notations of Tables 1 and 2 along with the following set of realistic assumptions.

Assumption. Large jumps in infected individuals caused by sudden major changes in the virus spreading environment are not considered in this model.

4.2. Development of the SDE model, its validation and the solution technique of the first wave of COVID-19

4.2.1. SDE model

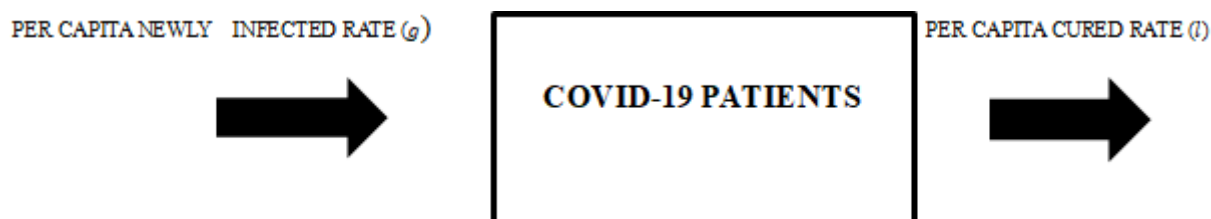
Following the notation in Table 1, new SDE model is developed first for predicting the first wave of COVID-19 cases in Sri Lanka. Thereafter, a solution method is presented to validate the developed SDE model.

Table 1. Notation of SDE model for the first wave of COVID-19.

Notations	Description
(a) Decision variables	
$X(t)$	Number of COVID-19 infected patients at time t
$W(t)$	Wiener process
t	Time
$P_{i,k}(t)$	Transition probability of a discrete stochastic process
(b) Input parameters	
N	Number of days
$i \in \{0, \dots, N\}$	Index of days
$j \in \{\dots, -2, -1, 0, 1, 2, \dots\}$	Index of patients
g	New infected rate per capita
l	Cured rate per capita
η	Random number generator
t_i	i^{th} day
Δt	Small time increment
σ	Volatility
μ	Drift

Prediction of the number of COVID-19 cases in any community is very difficult at this stage. The main reason for the virus spread is rapid contact between individuals [42]. The number of patients is presented at the National Epidemiology Center each day [43]. The secondary data is used to create the (SDE) model to identify the pattern of COVID-19 cases in the community. Following the Forward Kolmogorov equation (refer [1]) along with the notation of Table1, the stochastic differential equation (SDE) model is developed as follows:

According to the *Forward Kolmogorov equation*, let $t_i = i\Delta t$ for $i = 0, 1, \dots, N$ and let $x_j = j\delta$, for $j = \dots, -2, -1, 0, 1, 2, \dots$. Let x_0 be given. Consider the transition probability of a discrete stochastic process, where g and l are constant. If Δx changes in the stochastic process at time t , $X(t) = x_i$

**Figure 2.** Illustration of the simple virus infection procedure.

The main actions and reactions are indicated as newly infected and cured. According to Figure 2, the number of COVID-19 patients in the country is determined according to the infected action and cured action. Each infected person increases the number of patients and each cured person decreases the number of patients in the community. It is considered that ' g ' is the per capita newly infected rate and ' l ' is the per capita cured rate. So, we assume that ' g ' and ' l ' are constant according to the first case.

Assume that $g(t, x) = g$ and $l(t, x) = l$ are non-negative constants.

Define the transition probabilities of a discrete stochastic process by the following:

$$P_{i,k}(t) = \begin{cases} gx_i\Delta t & Fork = i + 1 \\ 1 - gx_i\Delta t - lx_i\Delta t & Fork = i \\ lx_i\Delta t & Fork = i - 1 \end{cases} \quad (1)$$

If $P_{i,i+1}(t)$ is considered, where $k = i + 1$, then $P_{(i,i+1)} = gx_i\Delta t$ by substituting

$$P_{(0,1)} = gx_0\Delta t, P_{(1,2)} = gx_1\Delta t, P_{(2,3)} = gx_2\Delta t \quad (2)$$

If $P_{i,i+1}(t)$ is considered, where $k = i + 1$, then $P_{(i,i+1)} = gx_i\Delta t$ by substituting

$$P_{(0,1)} = gx_0\Delta t, P_{(1,2)} = gx_1\Delta t, P_{(2,3)} = gx_2\Delta t \quad (3)$$

If $P_{i,i+1}(t)$ is considered, where $k = i + 1$, then $P_{(i,i+1)} = gx_i\Delta t$ by substituting

$$P_{(0,1)} = gx_0\Delta t, P_{(1,2)} = gx_1\Delta t, P_{(2,3)} = gx_2\Delta t \quad (4)$$

Let us consider Eq (1):

$$P_{i,k}(t) = \begin{bmatrix} P_{00} & P_{01} & \cdots & P_{0n} \\ P_{10} & P_{11} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ P_{m0} & \cdots & \cdots & P_{mn} \end{bmatrix}$$

Then, substituting Eqs (2)–(4) in the above matrix, we have

$$P_{i,k}(t) = \begin{bmatrix} 1 - gx_0\Delta t - lx_0\Delta t & gx_0\Delta t & 0 & \cdots \\ lx_1\Delta t & 1 - gx_1\Delta t - lx_1\Delta t & \cdots & \vdots \\ 0 & lx_2\Delta t & \ddots & \vdots \\ 0 & 0 & \cdots & 1 - gx_m\Delta t - lx_m\Delta t \\ \vdots & \vdots & & \end{bmatrix}$$

$E(\Delta x) = (g - f)\frac{\Delta t}{\delta}$; $Var(\Delta x) = (g + f)\Delta t$. It is assumed that $\Delta t/\delta^2$ is smallest so that $1 - b\Delta t/\delta^2 - d\Delta t/\delta^2$ is positive.

Let $P_k(t) = P(X(t) = x_k)$ be the probability distribution at time t . Then, $P_k(t + \Delta t)$ satisfies $P_k(t + \Delta t) = P_{k+1}(t)lx_{k+1}\Delta t + P_k(t)(1 - gx_k - lx_k)\Delta t + P_{k-1}(t)gx_{k-1}\Delta t$

Then, we derive $\frac{dP_k(t)}{dt}$,

$$\frac{P_k(t+\Delta t) - P_k(t)}{\Delta t} = P_{k+1}(t)lx_{k+1} - P_k(t)(gx_k + lx_k) + P_{k-1}(t)gx_{k-1} \quad (5)$$

It follows that the probability distribution in continuous time (letting $\Delta t \rightarrow 0$) satisfies the Forward Kolmogorov equations. As $\Delta t \rightarrow 0$, the discrete stochastic process is transformed to a continuous time process.

$$\therefore \Delta t \rightarrow 0 \frac{p_k(t+\Delta t) - p_k(t)}{\Delta t} = \frac{dp_k(t)}{dt} \quad (5.1)$$

From Eqs (5) and (5.1),

$$\begin{aligned} \frac{dP_k(t)}{dt} &= P_{k+1}(t) \left[\frac{lx_{k+1}}{2} + \frac{lx_{k+1}}{2} + \frac{gx_{k+1}}{2} - \frac{gx_{k+1}}{2} \right] + P_k(t) [-gx_k - lx_k] \\ &\quad + P_{k-1}(t) \left[\frac{gx_{k-1}}{2} + \frac{gx_{k-1}}{2} + \frac{lx_{k-1}}{2} - \frac{lx_{k-1}}{2} \right] \\ \frac{dP_k(t)}{dt} &= \frac{P_{k+1}(t)lx_{k+1}}{2} + \frac{P_{k+1}(t)lx_{k+1}}{2} + \frac{P_{k+1}(t)gx_{k+1}}{2} - \frac{P_{k+1}(t)gx_{k+1}}{2} - P_k(t)gx_k - P_k(t)lx_k \\ &\quad + \frac{P_{k-1}(t)gx_{k-1}}{2} + \frac{P_{k-1}(t)gx_{k-1}}{2} + \frac{P_{k-1}(t)lx_{k-1}}{2} - \frac{P_{k-1}(t)lx_{k-1}}{2} \\ \frac{dP_k(t)}{dt} &= -(g-l) \left[\frac{P_{k+1}(t)x_{k+1}}{2} - \frac{P_{k-1}(t)x_{k-1}}{2} \right] \\ &\quad + \left(\frac{g+l}{2} \right) [P_{k+1}(t)x_{k+1} - 2P_{k+1}(t)x_{k+1} + P_{k-1}(t)x_{k-1}] \end{aligned} \quad (5.2)$$

Assume that δ is small enough such that the stochastic process approaches a continuous valued process.

$$\text{Then, } \frac{F(x+\delta) - F(x-\delta)}{2\delta} = \left[\frac{P_{k+1}(t)x_{k+1} - P_{k-1}(t)x_{k-1}}{2\delta} \right]$$

$$\frac{F(x+\delta) - 2F(x) + F(x-\delta)}{\delta^2} = \frac{P_{k+1}(t)x_{k+1} - 2P_{k+1}(t)x_{k+1} + P_{k-1}(t)x_{k-1}}{\delta^2}$$

According to the above equations, it can be considered as central difference approximation for some values $\varepsilon_1, \varepsilon_2$ such that

$$x - \delta \leq \varepsilon_1, \varepsilon_2 \leq x + \delta$$

$$\frac{F(x+\delta) - F(x-\delta)}{2\delta} = F'(x) + \frac{\delta^2}{6} F'''(\varepsilon_1)$$

$$\frac{F(x+\delta) - 2F(x) + F(x-\delta)}{\delta^2} = F''(x) + \frac{\delta^2}{12} F''''(\varepsilon_2)$$

Note that we fixed $X(t) = x_i$ at time t , $E(\Delta X) = (g-l)x\Delta t$ and $Var(\Delta X) = (g+l)x\Delta t$ to order $(\Delta t)^2$. The above equations approximately satisfy the Fokker-Planck equation (refer to [1]). Then the above system of differential equation approximates the partial differential equation.

$$\frac{\partial p(t,x)}{\partial t} = -\frac{\partial}{\partial x}((g-l)xp(t,x)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}((g+l)^2xp(t,x)) \quad (5.3)$$

Consider the ‘ m ’ and ‘ n ’ functions and the stochastic differential equation,

$$dX(t) = m(t, X(t))dt + n(t, X(t))dW(t) \quad (5.4)$$

From (5.3) and (5.4),

$$dX(t) = (g-l)X(t)dt + \sqrt{(g+l)X(t)}dW(t) \quad (6)$$

Equation (6) is the SDE model used to predict the first wave of COVID-19 cases in Sri Lanka. The probability density of the solution of the stochastic differential equation satisfies the partial differential equation. Therefore, there exists a close relationship between the discrete stochastic process and the continuous process. In particular, for small Δt and δ , the probability distribution of the solutions to Forward Kolmogorov equation will be approximately the same as the probability distribution of solutions to the discrete stochastic process [1]. Then, an appropriate stochastic differential equation model is inferred from the above argument.

Finally, it is useful to note in the above argument that the coefficients at time t of the stochastic differential equation are related to the discrete stochastic model through the mean and variance in the change in the process ΔX over a short time interval Δt fixing $X(t) = x$.

4.2.2. SDE model validation

The stochastic equation in our SDE model can be validated by using the Mean absolute deviation (MAD) method, mean absolute percentage error (MAPE) method and Bias method. It is inevitable to ensure the data fittings of the curve drawn according to the stochastic equation. The predicted data deviate from the actual data and those changes ensure the validity of the derived equation.

The accuracy model equations and the scale of judgment of forecast accuracy (Table 2) are defined as follows [44]:

(1) Mean Absolute Percentage Error (MAPE) method

$$\varepsilon_{\text{MAPE}} = \frac{1}{n} \sum_{i=0}^n \left| \frac{A-P}{A} \right| \quad (7)$$

(2) Mean Absolute Deviation (MAD) method

$$\varepsilon_{\text{MAD}} = \frac{1}{n} \sum_{i=0}^n |A - P| \quad (8)$$

Where A and P represent the actual value and predicted value of the indices, respectively.

(3) Absolute Bias method

$$\varepsilon_{\text{Bias}} = |A - P| \quad (9)$$

Using SDE model in Eq (6) with the sample data (14 March to 17 April 2020), the MAPE, MAD and absolute bias are 2.63%, 0.247 and 8.6571, respectively. According to the error analysis, the new proposed model is highly accurate (less than 10%) with lowest MAPE error values (refer to Table 6).

Note that Table 2 represents the scale of judgment of forecast accuracy regarding to Error (MAPE) and Table 3 contains the notations of the proposed model. It clearly indicates that minimum values of MAPE result in greater accuracy for forecasting the future predictions [44].

Table 2. Model accuracy testing.

MAPE	Judgment of Forecast Accuracy
<10%	Highly Accurate
11% to 20%	Good Forecast
21% to 50%	Reasonable Forecast
>51%	Inaccurate Forecast

Table 3. Notation of SDE model for the second wave of COVID-19.

Notations	Description
(a) Decision variables	
C	C is the number of infected cases
K	Final epidemic size
t	Time
(b) Input parameters	
S	Susceptible healthy population that is at risk of getting infected.
I	Infected population showing symptoms, whether mild or severe
R	Recovered population who have recovered from the disease and gained immunity
β	Infection rate—defined as the proportion of the infected population per day.
γ	Recovery rate—defined as the proportion of the recovered population per day
\emptyset	Death rate—defined as the proportion of the infected population per day
$R_0 = 2.2$	Reproduction number of the disease
$r=0.126$	Rate of infection
N	Population

4.2.3. SDE model validation solution procedure

To obtain considerable valuation for forecasting the data set, a useful model requires more COVID-19 infection data. Thus, the major challenge of this study is to find a suitable data range for a particular community. The current study is carried out on the basis of secondary data, which are obtained from the Epidemiology Unit, Ministry of Health, Sri Lanka [43]. Daily trading data for one-month and four-day period, from 17 March 2020 to 14 April 2020, are extracted and tabulated for calculations. In this study, a new proposed stochastic model is widely applied to identify the random behavior of a virus with the combination of volatility and drift of the equation. To solve our new SDE model for predicting the first wave of COVID-19 patients, the following step-by-step procedure is presented.

Approach 1: Algorithm for solving the SDE model for the 1st wave of COVID-19

Step 1: First, during the past three months (14 of March 2020 to 8 of July 2020), infected patients in the country are taken as the population in this study. Data are collected through secondary data sources [43]. Then, the infected patients are selected during 14 March 2020 to 17 April 2020 as the sample using *random sampling technique*.

Step 2: The model is derived using the Kolmogorov equation. The per capita infected rate g and per capita cured rate l are taken to measure the volatility and drift of the stochastic differential equation.

Step 3: The Euler–Maruyama method is used to find the numerical solution to the presented stochastic differential equation–Virus spreading model in Eq (6). Let $X(t)$ be the dependent variable for stochastic differential equation:

$$X_{t+1} = X_t + \alpha X_{t-1} \Delta t + \sigma X_{t-1} \eta \sqrt{\Delta t} \quad (10)$$

η = Random number generator.

Step 4: First, calculate g and l by using secondary data. Then, Eq (10) is transformed into the following Microsoft Excel formula to find the data pattern.

We assume the following parameters in the Microsoft Excel (2013) sheet provided in Appendix A:

Time Δt : $1/365 = 0.0027397$

Drift μ : $\$C\$7 * \$C\$8 * E12$

Uncertainty σ : $\text{NORM.INV}(\text{RAND}(), 0, 1) * \text{SQRT}(\$C\$8) * \$C\$6 * \text{SQRT}(E12)$

Change: $C13 + B13$

Patients: $D14 + E13$

Random number generator is used to create the pattern and, by changing the values of the drift and volatility, many patterns are obtained until the most appropriate pattern is found. Simultaneously, the actual data are plotted in the same graph.

Step 5: The obtained data pattern in Step 4 is validated using the existing validation methods MAPE, MAD and bias to see whether this resulted data pattern is better fit with the actual data pattern.

4.3. Two well-known models and the solution technique for the second wave of COVID-19 outbreak

Here, two well-known models are presented for predicting the second wave of COVID-19 scenarios in Sri Lanka. Then, a solution method is proposed for the presented models.

4.3.1. Two well-known models for the second wave of COVID-19 outbreak

To predict the second wave of COVID-19 cases in Sri Lanka, two well-known Logistic growth and SIR models (refer to [37] and [38–41]) with the notations in Table 3 are obtained as follows:

The Logistic growth model [37]

$$\frac{1}{C} \frac{dC}{dt} = r \left(1 - \frac{C}{K} \right) \quad (11)$$

$$\frac{dC}{dt} = r \left(1 - \frac{C}{K} \right) C \quad (12)$$

Integrating (12),

$$\int \frac{dC}{C(1-\frac{C}{K})} = \int r dt \quad (13)$$

Then, by solving Eq (13), we can obtain

$$C = \frac{K}{1 + \left(\frac{K-C_0}{C_0}\right)e^{rt}} \quad (14)$$

Further, the maximum growth rate peaks at the time

$$t_p = \frac{\ln\left(\frac{K-C_0}{C_0}\right)}{r} \quad (15)$$

SIR Model [38– 41]

$$\frac{dS}{dt} = -\beta I \frac{S}{N} \quad (16)$$

$$\frac{dI}{dt} = \beta I \frac{S}{N} - \gamma I \quad (17)$$

$$\frac{dR}{dt} = \gamma I \quad (18)$$

$$\frac{dD}{dt} = \phi I \quad (19)$$

4.3.2. Solution technique for the two well-known models for predicting the second wave of COVID-19 outbreak

To solve the two model (Logistic growth equation and SIR model) for predicting the second wave of COVID-19 patient, the following step-by-step procedure is presented.

Approach 2: Algorithm for solving the SDE model for the second wave of COVID-19

Step 1: Define the population as the Sri Lankan population for this study.

Step 2: Apply the Logistic growth and SIR models.

Step 3: To plot the predicted pattern of the virus spreading, a MATLAB computational program (Appendix B) is developed for solving SIR model.

Step 4: To plot the predicted pattern of the Logistic growth model, Equations (14) and (15) are solved.

5. Results and discussion

5.1. Results of the first wave of COVID-19 patients in Sri Lanka

According to the solution technique (3.2.3), the first and second steps are completed in the previous sections. This section analyzes the data based on the predictive model [1].

According to Step 3, Eq (6) is approximated by using Euler–Maruyama method [36] in the given

time interval $t_0 < t_1 < t_2 < \dots < t_n$, approximating the $X_0 < X_1 < X_2 < \dots < X_n$ values. Determine the respective t point from (6)

$$dX(t) = (g - l)X(t)dt + \sqrt{(g + l)X(t)}dW(t) \quad (6)$$

Approximate solution from Eq (10) is as follows:

$$X_{t+1} = X_t + \mu X_{t-1}\Delta t + \sigma X_{t-1}\eta\sqrt{\Delta t} \quad (10)$$

$$\Delta t = t_{i+1} - t_i$$

where: η = Random number generator, σ = *Volatility* and μ = *Drift*.

According to Step 4, to find the predicted data, it is required to calculate the per capita infected rate (g) and per capita cured rate (l). Therefore, the number of sample data in Table 4 is used to determine these rates.

Table 4. Secondary data sources, published by the Epidemiology Unit, Ministry of Health, Sri Lanka [43].

Date_2020	Sum of infected	Active cases	Infected X_i	Cured Y_i
4-Mar	1	0	0	0
5-Mar	1	0	0	0
6-Mar	1	0	0	0
7-Mar	1	0	0	0
8-Mar	1	0	0	0
9-Mar	1	0	0	0
10-Mar	2	0	1	0
11-Mar	2	1	0	1
12-Mar	4	3	0	0
13-Mar	6	5	0	0
14-Mar	11	10	8	0
15-Mar	19	18	8	0
16-Mar	29	28	10	0
17-Mar	42	41	16	0
18-Mar	53	52	7	2
19-Mar	66	65	9	0
20-Mar	72	71	13	0
21-Mar	78	77	4	0
22-Mar	87	86	5	0
23-Mar	97	95	15	0
24-Mar	102	100	5	0
25-Mar	102	99	0	1
26-Mar	106	100	4	4
27-Mar	106	99	0	0

Continued on next page

Date_2020	Sum of infected	Active cases	Infected X_i	Cured Y_i
28-Mar	115	105	7	2
29-Mar	117	105	4	2
30-Mar	122	106	5	4
31-Mar	143	124	21	2
1-Apr	146	122	3	4
2-Apr	151	126	5	0
3-Apr	159	131	8	3
4-Apr	166	134	7	3
5-Apr	176	138	10	6
6-Apr	178	135	2	5
7-Apr	185	137	7	4
8-Apr	189	138	4	2
9-Apr	190	134	1	5
10-Apr	197	136	0	5
11-Apr	199	138	8	0
12-Apr	210	147	12	2
13-Apr	218	155	7	0
14-Apr	233	165	16	5
15-Apr	238	168	5	2
16-Apr	238	163	0	5
17-Apr	244	160	6	9

Following the infected values (X_i) and cured values (Y_i) in Table 4, g and l can be computed as follows:

$$g = \frac{\sum_{i=0}^N (X_{i+1} - X_i)}{N} = \frac{185}{30} = 6.11; l = \frac{\sum_{i=0}^N (Y_{i+1} - Y_i)}{N} = \frac{3}{30} = 0.1 \text{ thus, } (g - l) = 6.01; (g + l) = 6.21$$

Substituting Eq (6),

$$dX(t) = (g - l)X(t)dt + \sqrt{(g + l)X(t)}dW(t)$$

$$dX(t) = (6.01)X(t)dt + \sqrt{(6.21)X(t)}dW(t)$$

Table 5. Values of the parameters in our developed SDE model.

Parameters	Values
Initial number of patients	1
Volatility $\sigma = (g + l)$	6.21
Drift $\mu = (g - l)$	6.01
Time Step $\Delta t = 1/365$	0.002739726

Based on the above notion (Table 5), first, Eq (10) is converted using Microsoft Excel (2013).

Then, predicted patients are found in Appendix A. Finally, the predicted and actual data can be tabulated (Table 6) as follows:

Table 6. Difference between two values in consecutive days and related percentages.

Date	Actual	Predicted	Difference	Absolute bias
14-Mar	17	16	1	1
15-Mar	27	22	5	5
16-Mar	43	29	14	14
17-Mar	50	36	14	14
18-Mar	59	43	16	16
19-Mar	72	51	21	21
20-Mar	76	58	18	18
21-Mar	81	64	17	17
22-Mar	96	71	25	25
23-Mar	101	81	20	20
24-Mar	101	89	12	12
25-Mar	105	98	7	7
26-Mar	105	107	-2	2
27-Mar	112	117	-5	5
28-Mar	116	125	-9	9
29-Mar	121	131	-10	10
30-Mar	142	137	5	5
31-Mar	145	144	1	1
1-Apr	150	150	0	0
2-Apr	158	156	2	2
3-Apr	165	163	2	2
4-Apr	175	168	7	7
5-Apr	177	174	3	3
6-Apr	184	180	4	4
7-Apr	188	185	3	3
8-Apr	189	190	-1	1
9-Apr	189	194	-5	5
10-Apr	197	199	-2	2
11-Apr	209	205	4	4
12-Apr	216	211	5	5
13-Apr	232	217	15	15
14-Apr	237	222	15	15
15-Apr	237	227	10	10
16-Apr	243	233	10	10
17-Apr	253	240	13	13
MAPE			2.63%	
MAD			0.247	
Absolute bias			8.657	

In Figure 3, the predicted solution is plotted as a blue curve and actual data graph is plotted as red curve. The Euler–Maruyama approximations with drift and volatility parameters are set to 6.211 and 6.01, respectively.

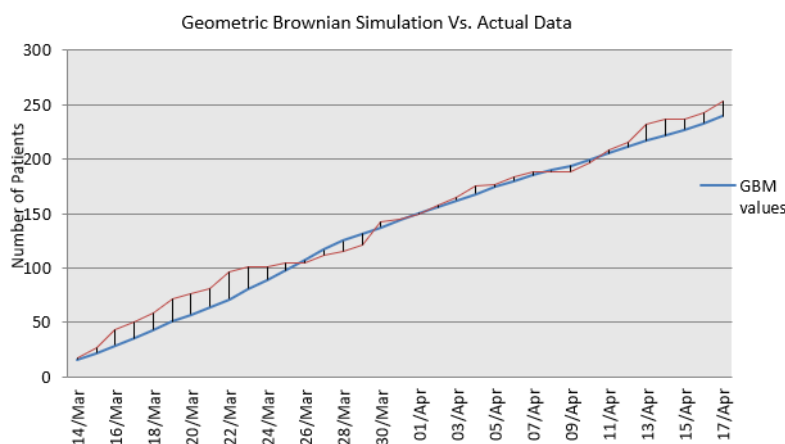


Figure 3. Solution to the virus infection changes from stochastic differential equation.

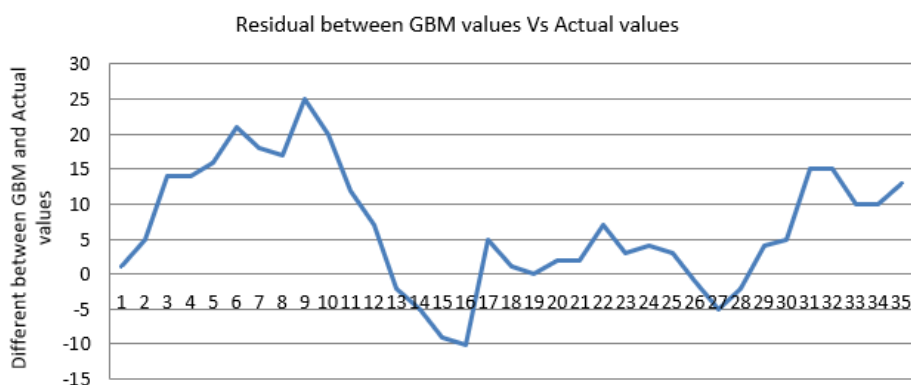


Figure 4. Residual values of geometric Brownian motion values vs. actual values.

In the above figure, the X axis shows the number of days and the Y axis shows the difference between GBM [45] and actual values. The corresponding results are summarized in Table 6 and Figures 3 and 4. The reported results in Table 6 represent the differences and similarities between forecasting results with respect to the actual data patterns. According to Table 3, the proposed model is highly accurate (less than 10%) with lowest MAPE error values.

The mean absolute bias is 8.6571 and Figure 4 depicts the residual values of GBM vs. actual values. Moreover, MAPE values suggest that our proposed stochastic differential equation (MAPE: 2.63%, MAD: 0.247) is more significant for forecasting short time predictions.

5.1.1. Discussion of the first wave of COVID-19 patients in Sri Lanka

In the current study, the proposed stochastic model approach is used to discuss our result. This newly developed model values and the actual data were plotted at the same graph (Figure 3). These

two graphs depicted similar pattern and the mean absolute percentage error (MAPE) results also reveal that (MAPE =2.63%) the proposed stochastic model is more accurate and gives the solution for predicting virus spread in a community. Generally, the COVID-19 virus spread is disordered and shows very quick transmission behaviors. Therefore, the accuracy of forecast might be enhanced by modeling the stochastic differential methods as well. It is more important for governments to take action to prevent the spread of COVID-19 all over the country.

5.2. Results and discussion of the second wave of COVID-19 patients in Sri Lanka

5.2.1. Logistic growth model

When analyzing the Daily Reported Cases, the country is yet to reach the spike rate of cases in the third week of December and the pandemic will reach the end in the fourth week of January 2021. Through firm preventive action, the number of cases, the peak date and the eradication is yet to come. According to the Logistic growth model, the normal behavior of the virus will affect 77,838 people.

Blue colored bars in Figure 5 indicate the daily actual infectious COVID-19 cases during the period of two months (from 5 October 2020 to 5 December 2020). The Logistic growth model prediction curve in Figure 5 depicts the daily predicted infectious COVID-19 cases from 05 October 2020 to 15 February 2021. It can be easily observed from Figure 5 that the highest number of infectious COVID-19 cases obtained on the 70th day (i.e., on 14 December 2020). These findings are also provided in Table 6.

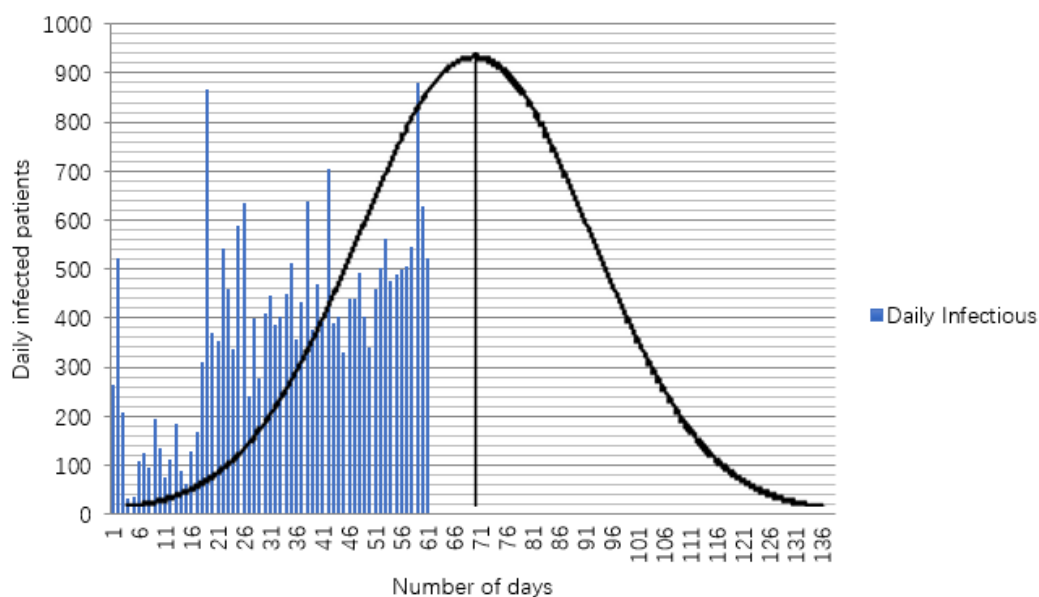


Figure 5. Logistic growth model prediction curve and actual situation of infectious individuals in Sri Lanka from 5 of October 2020 to 15 February 2021.

Table 7 indicates that the highest number of predicted infectious COVID-19 cases (estimated peak day cases) is 933, observed on 14 December 2020. However, the highest number of actual infectious COVID-19 cases (actual peak day cases) on this day is 701. Please note that a slight deviation of the

predicted cases (933) from the actual cases (701) can be observed. Similarly, on the 15 February 2021, the estimated and the actual epidemic sizes (cumulative) of the COVID-19 cases are 77,838 and 75,085 respectively. Note that a slight deviation of the estimated epidemic size (77,838) from the actual epidemic size (75,085) can be observed. Besides, the Logistic growth model can be extended by incorporating safety measures (taken by government) as the external factor (that affect the spread of the virus) to minimize the above stated deviations.

Table 7. Estimated simulation results of the logistics growth model.

Parameters	Values
Estimated epidemic rate (per day)	0.126
Estimated start of ending phase date	15/02/2021
Estimated epidemic size (Cumulative)	77,838 Cumulative cases
Actual epidemic size (Cumulative)	75,085 Cumulative cases
Estimated peak date	14/12/2020
Estimated peak day cases	933 cases
Actual peak day cases	701 cases

A comparison between actual cases and those without intervention is shown in Figure 6 below.

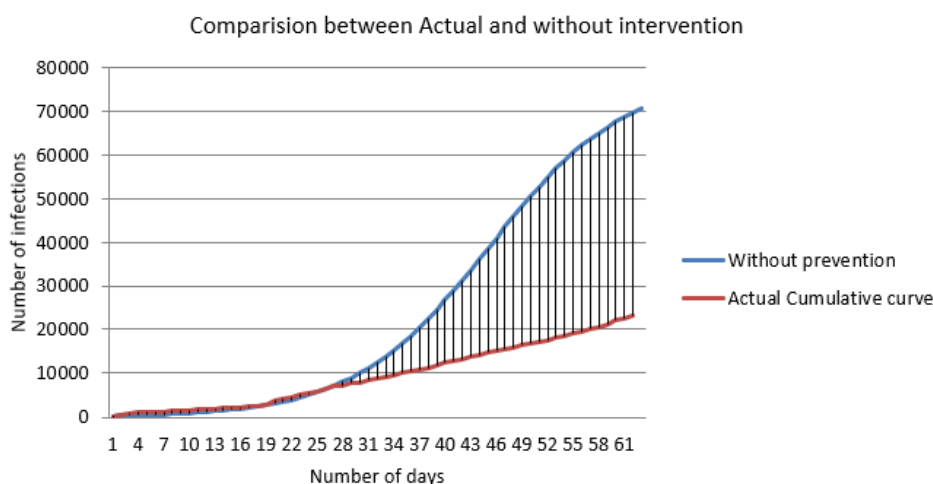


Figure 6. Logistic growth model prediction curves between no intervention of government and prevailing situation in Sri Lanka from 5 October 2020 to 5 December 2020.

When viewing the number of cases that should have been reported with no intervention, it will reach to approximately over 70,000 cases (refer to Figure 6). However, the government of Sri Lanka has been successfully mitigating the situation with less than 30,000 cases as of 10 December 2020. This shows the immense contribution of the government and its stakeholders in their decision making and preventive, fighting, managing and accountability theatres.

5.2.2. SIR model

When the existence of the virus is forecasted to an extended year from 5 October 2020 to 5

January 2021, in the Sri Lankan context, the people of the country are more susceptible to the virus when the government's involvement in the pandemic is absent. To solve the SIR model, a MATLAB computational program in Appendix B is utilized. Thus, the future prediction of COVID-19 in Sri Lanka (when the government's involvement of preventing the virus is absent) can be seen in Figure 7 below.

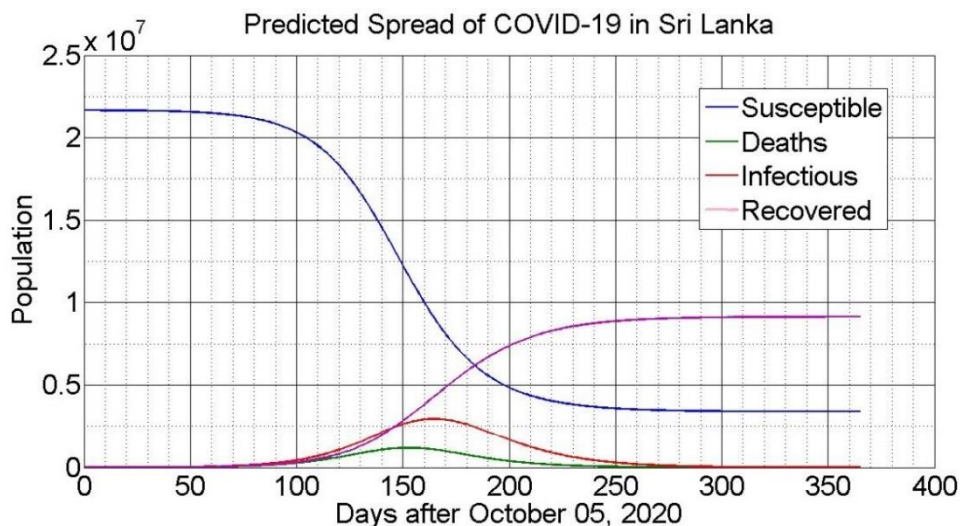


Figure 7. Predicted spread of COVID-19 in Sri Lanka if the government is not involved in prevention of the virus.

Suitable control over the infected people and reducing the risk of contact with the virus kept the country at an advantageous position. The stakeholders also deserve some credit for controlling the pandemic as the recovery rate remains comparatively higher than the death rate.

5.3. Comparison of the fatality rates and recovery rates of Sri Lanka vs. the Globe

Following [43], Sri Lankan fatality statistics are shown below, where the fatality rates of Sri Lanka are compared to other countries. Note that these fatality and recovery rate comparisons are obtained from [43] at the beginning of the second wave on 5 October 2020.

According to Figure 8, the fatality rate of Sri Lanka is the lowest compared with the global rate. By now, the world has a 2.29% fatality rate whereas Sri Lanka stands at 0.5%. This shows the success of Sri Lanka's COVID-19 response despite the fact that most of the developed and developing countries failed in this aspect, even those using the latest technology, health services and facilities.

When the main causes behind the COVID-19 deaths are analysed, it is apparent that people with NCDs (non communicable diseases) such as Cardiac arrest (17.7%), pneumonia (11.5%) and diabetes are the most vulnerable group. It is noticed in Figure 9 that Sri Lanka's recovery rate stands at a higher level compared with the global recovery rate.

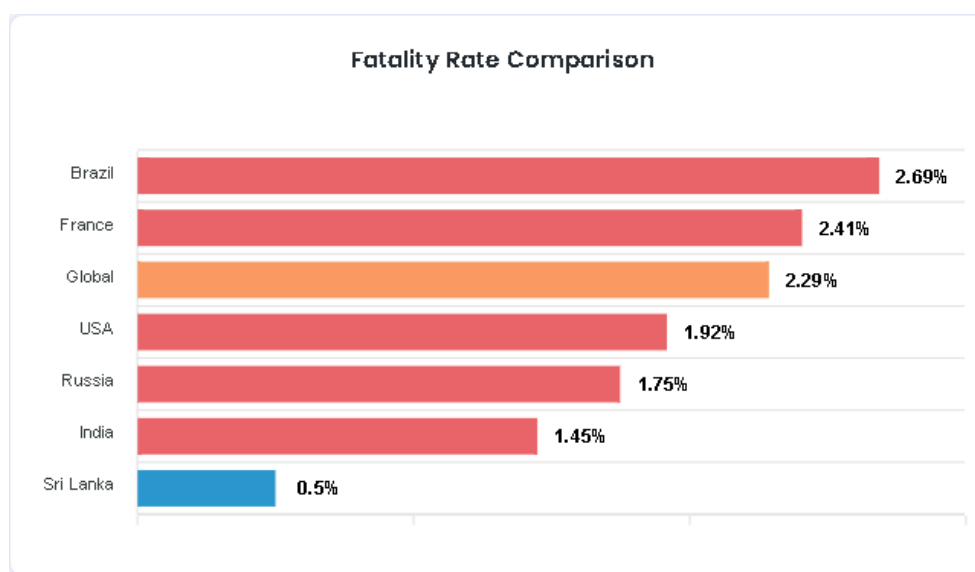


Figure 8. Comparison of fatality rates of many countries (COVID-19 tracker).

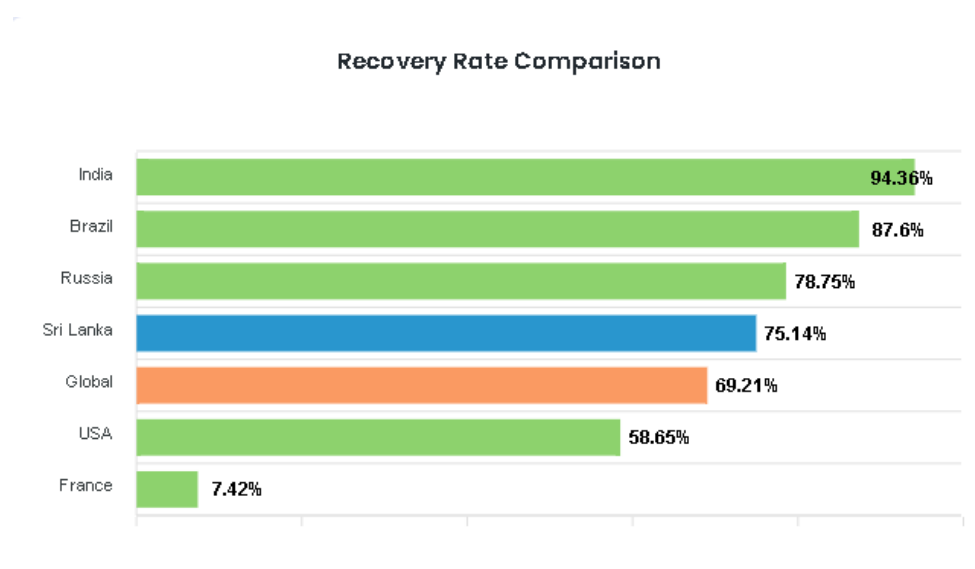


Figure 9. Comparison of recovery rates of many countries (COVID-19 tracker).

6. Conclusions

The novel coronavirus (COVID-19) has badly affected many countries in the world. Sri Lanka recorded the first confirmed case of COVID-19 on 27 January 2020, who was 44-year-old Chinese women from Hubei province in China. She had arrived as a tourist with another group of travelers and had been screened at the Bandaranayaka International Airport after having a high fever. She fully recovered and was later released on 19 February 2020. This incident made the health ministry and other relevant authorities more focused on the ongoing crisis around the globe. Predicting the number of COVID-19 cases has become vital for governments to take proper actions to prevent virus-spread. The first and second waves of COVID-19 cases in the country are investigated in this paper. First, a

new stochastic forecasting model is developed to predict the future COVID-19 outbreak in the country. The newly developed model is validated with the best available methods in the literature and a solution method is presented to solve the model efficiently. Thereafter, the solution method is presented to efficiently solve two existing models for predicting the number of second wave of COVID-19 outbreak. The presented models and solution approaches of this paper are validated with secondary data obtained from the Epidemiology Unit, Ministry of Health, Sri Lanka. Our new SDE model solutions approximately coincide with the actual data; thus, this forecasting model can predict future values with an accuracy of 92 percent. This percentage accuracy can be further improved by fine-tuning the SDE model parameters such as variance and expectation of the formulae. Moreover, the mean absolute percentage error (MAPE) results also reveal that (MAPE = 2.63%) the new proposed SDE is more accurate and gives better solution for predicting virus spread in a community. Besides, according to Figure 8, the fatality rate of Sri Lanka remains at the lowest level compared with the global rate and other countries. By now, the world suffers from a 2.29% fatality rate, whereas Sri Lanka stands at 0.5%. It is noticed from Figure 9 that Sri Lanka's recovery rate stands at a higher level compared with the global recovery rate. By utilizing our models, these fatality rates and recovery rates can be improved further. Hence, our findings (new proposed SDE model, two existing models, and the new solution techniques) are beneficial to the government of the country for making better decisions and to take better actions to prevent the spreading of the COVID-19 virus.

Although the proposed SDE model predicts the future values for the first wave of COVID-19 cases with a better accuracy (92%), a significant deviation of the estimated epidemic size (77,838) from the actual epidemic size (75,085) during the second wave of COVID-19 can be found with the Logistic growth model. So, future research can be directed to extend the Logistic growth model by incorporating safety measures (taken by government) as the external factor (that affects the spread of the virus) to minimize the deviation. Also, our proposed new SDE model could be further extended to study the behavior of new Covid variant, the Omicron.

Acknowledgments

The authors would like to express their sincerest thanks to the editors and reviewers for their constructive suggestions and correction to enhance the clarity of the article.

Conflict of interest

The authors declare that there are no potential conflicts of interest.

References

1. E. Allen, *Modeling with Itô stochastic differential equations*, Springer Science & Business Media, 2007. <https://doi.org/10.1007/978-1-4020-5953-7>
2. F. A. Rihan, Q. M. Al-Mdallal, H. J. Al-Sakaji, A. Hashish, A fractional-order epidemic model with time-delay and nonlinear incidence rate, *Chaos Soliton. Fract.*, **126** (2019), 97–105. <https://doi.org/10.1016/j.chaos.2019.05.039>

3. N. Sharma, A. K. Verma, A. K. Gupta, Spatial network based model forecasting transmission and control of COVID-19, *Physica A*, **581** (2021), 126223. <https://doi.org/10.1016/j.physa.2021.126223>
4. J. Guan, Y. Wei, Y. Zhao, F. Chen, Modeling the transmission dynamics of COVID-19 epidemic: a systematic review, *The Journal of Biomedical Research*, **34** (2020), 422–430. <https://doi.org/10.7555/JBR.34.20200119>
5. R. U. Din, A. R. Seadawy, K. Shah, A. Ullah, D. Baleanu, Study of global dynamics of COVID-19 via a new mathematical model, *Results Phys.*, **19** (2020), 103468. <https://doi.org/10.1016/j.rinp.2020.103468>
6. I. F. Mello, L. Squillante, G. O. Gomes, A. C. Seridonio, M. de Souza, Epidemics, the Ising-model and percolation theory: a comprehensive review focused on Covid-19, *Physica A*, **573** (2021) 125963. <https://doi.org/10.1016/j.physa.2021.125963>
7. H. A. Adekola, I. A. Adekunle, H. O. Egberongbe, S. A. Onitilo, I. N. Abdullahi, Mathematical modeling for infectious viral disease: The COVID-19 perspective, *J. Public Aff.*, **20** (2020), 2306. <https://doi.org/10.1002/pa.2306>
8. A. M. Hajji, Q. Al-Mdallal, Numerical simulations of a delay model for immune system-tumor interaction, *Sultan Qaboos University Journal for Science*, **23** (2018), 19–31. <https://doi.org/10.24200/squjs.vol23iss1pp19-31>
9. A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, et al., Early dynamics of transmission and control of COVID-19: a mathematical modeling study, *The Lancet Infectious Diseases*, **20** (2020), 553–558. [https://doi.org/10.1016/S1473-3099\(20\)30144-4](https://doi.org/10.1016/S1473-3099(20)30144-4)
10. E. Allen, Stochastic differential equation models for the wear of coins in circulation, *Tribol. Lett.*, **64** (2016), 45. <https://doi.org/10.1007/s11249-016-0780-x>
11. L. J. S. Allen, A primer on stochastic epidemic models: formulation, numerical simulation, and analysis, *Infectious Disease Modelling*, **2** (2017), 128–142. <https://doi.org/10.1016/j.idm.2017.03.001>
12. E. Allen, Environmental variability and mean reverting processes, *Discrete Cont. Dyn. Syst. B*, **21** (2016), 2073–2089. <https://doi.org/10.3934/dcdsb.2016037>
13. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, et al., Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia, *N. Engl. J. Med.*, **382** (2020), 1199–1207. <https://doi.org/10.1056/NEJMoa2001316>
14. C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *The Lancet*, **395** (2020), 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
15. S. Ahmad, A. Ullah, Q. M. Al-Mdallal, H. Khan, K. Shah, A. Khan, Fractional order mathematical modeling of COVID-19 transmission, *Chaos Soliton. Fract.*, **139** (2020), 110256. <https://doi.org/10.1016/j.chaos.2020.110256>
16. B. Bukiet, M. Ovens, A mathematical modeling approach to one-day cricket batting orders, *J. Sports Sci. Med.*, **5** (2006), 495–502.
17. F. B. Augusto, S. Bewick, W. F. Fagan, Mathematical model of Zika virus with vertical transmission, *Infectious Disease Modelling*, **2** (2017), 244–267. <https://doi.org/10.1016/j.idm.2017.05.003>
18. A. E. E. Eltoukhy, I. A. Shaban, F. T. S. Chan, M. A. M. Abdel-Aal, Data analytics for predicting COVID-19 cases in top affected countries: observations and recommendations, *Int. J. Environ. Res. Public Health*, **17** (2020), 7080. <https://doi.org/10.3390/ijerph17197080>

19. D. Xu, Modeling of network based digital contact tracing and testing strategies, including the pre-exposure notification system, for the COVID-19 pandemic, *Math. Biosci.*, **338** (2021), 108645. <https://doi.org/10.1016/j.mbs.2021.108645>
20. I. J. Rao, M. L. Brandeau, Optimal allocation of limited vaccine to control an infectious disease: simple analytical conditions, *Math. Biosci.*, **337** (2021), 108621. <https://doi.org/10.1016/j.mbs.2021.108621>
21. S. M. Ciupe, J. M. Heffernan, In-host modeling, *Infectious Disease Modelling*, **2** (2017), 188–202. <https://doi.org/10.1016/j.idm.2017.04.002>
22. A. Antonietta, J. P. Whiteley, Model reduction for initial value ODEs, *Math. Biosci.*, **337** (2021), 108618. <https://doi.org/10.1016/j.mbs.2021.108618>
23. A. Khan, H. M. Alshehri, T. Abdeljawad, Q. M. Al-Mdallal, H. Khan, Stability analysis of fractional nabla difference covid-19 model, *Results Phys.*, **22** (2021), 103888. <https://doi.org/10.1016/j.rinp.2021.103888>
24. T. N. Sindhu, A. Shafiq, Q. M. Al-Mdallal, On the analysis of number of deaths due to Covid-19 outbreak data using a new class of distributions, *Results Phys.*, **21** (2021), 103747. <https://doi.org/10.1016/j.rinp.2020.103747>
25. L. Vassallo, I. A. Perez, L. G. A. Zuzek, J. Amaya, M. F. Torres, L. D. Valdez, et al., An epidemic model for COVID-19 transmission in Argentina: Exploration of the alternating quarantine and massive testing strategies, *Math. Biosci.*, **2021**, 108664. <https://doi.org/10.1016/j.mbs.2021.108664>
26. S. Winkelmann, J. Zonkerab, C. Schütteab, N. D. Conrada, Mathematical modeling of spatio-temporal population dynamics and application to epidemic spreading, *Math. Biosci.*, **336** (2021), 108619. <https://doi.org/10.1016/j.mbs.2021.108619>
27. T. Y. Lin, A. Yen, S. Chen, Likelihood function for estimating parameters in multistate disease process with Laplace-transformation-based transition probabilities, *Math. Biosci.*, **335** (2021), 108586. <https://doi.org/10.1016/j.mbs.2021.108586>
28. H. Singh, H. M. Srivastava, Z. Hammouch, K. S. Nisar, Numerical simulation and stability analysis for the fractional-order dynamics of COVID-19, *Results Phys.*, **20** (2021), 103722. <https://doi.org/10.1016/j.rinp.2020.103722>
29. H. M. Srivastava, R. Jan, A. Jan, W. Deebani, M. Shutaywi, Fractional-calculus analysis of the transmission dynamics of the dengue infection, *Chaos*, **31** (2021), 53130. <https://doi.org/10.1063/5.0050452>
30. H. M. Srivastava, I. Area, J. J. Nieto, Power-series solution of compartmental epidemiological models, *Math. Biosci. Eng.*, **18** (2021), 3274–3290. <https://doi.org/10.3934/mbe.2021163>
31. B. M. Henry, M. H. S. De Oliveira, S. Benoit, M. Plebani, G. Lippi, Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): a meta-analysis, *Clin. Chem. Lab. Med.*, **58** (2020), 1021–1028. <https://doi.org/10.1515/cclm-2020-0369>
32. S. Bhattacharjee, A. Banerjee, A. Rakshit, S. Bhattacharyya, S. Chowdhuri, B. Sarkar, et al., Dynamics of cardiovascular muscle using a non-linear symmetric oscillator, *Symmetry*, **13** (2021), 151. <https://doi.org/10.3390/sym13010151>
33. Martingale Probability Theory, Wikipedia, The Free Encyclopedia. Available From: [https://en.wikipedia.org/wiki/Martingale_\(probability_theory\)](https://en.wikipedia.org/wiki/Martingale_(probability_theory)).

34. Markov Chains, Chapter-8, The University of Auckland, New Zealand. Available From: <https://www.stat.auckland.ac.nz/~fewster/325/notes/ch8.pdf>.
35. T. Sauer, Numerical solution of stochastic differential equations in finance, In: *Handbook of computational finance*, Berlin, Heidelberg: Springer, 2012, 529–550. https://doi.org/10.1007/978-3-642-17254-0_19
36. I. I. Gikhman, A. V. Skorokhod, Stochastic differential equations, In: *The theory of stochastic processes III*, Berlin, Heidelberg: Springer, 2007, 113–219. https://doi.org/10.1007/978-3-540-49941-1_2
37. A. M. Almeshal, A. I. Almazrouee, M. R. Alenziand, S. N. Alhajeri, Forecasting the spread of COVID-19 in Kuwait using compartmental and logistic regression models, *Appl. Sci.*, **10** (2020), 3402. <https://doi.org/10.3390/app10103402>
38. R. M. Anderson, Discussion: the Kermack-McKendrick epidemic threshold theorem, *Bull. Math. Biol.*, **53** (1991), 3–32. [https://doi.org/10.1016/S0092-8240\(05\)80039-4](https://doi.org/10.1016/S0092-8240(05)80039-4)
39. R. Singh, R. Adhikari, Age-structured impact of social distancing on the COVID-19 epidemic in India, 2020, arXiv:2003.12055.
40. L. Zhong, L. Mu, J. Li, J. Wang, Z. Yin, D. Liu, Early prediction of the 2019 novel coronavirus outbreak in the mainland china based on simple mathematical model, *IEEE Access*, **8** (2020), 51761–51769. <https://doi.org/10.1109/ACCESS.2020.2979599>
41. D. I. Vega, Lockdown, one, two, none, or smart. Modeling containing COVID-19 infection. A conceptual model, *Sci. Total Environ.*, **730** (2020), 138917. <https://doi.org/10.1016/j.scitotenv.2020.138917>
42. T. N. Sindhu, A. Shafiq, Q. M. Al-Mdallal, Exponentiated transformation of Gumbel Type-II distribution for modeling COVID-19 data, *Alex. Eng. J.*, **60** (2021), 671–689. <https://doi.org/10.1016/j.aej.2020.09.060>
43. Ministry of health, National Epidemiology report—Sri Lanka, cited April 19, 2020. Available from: <https://www.epid.gov.lk/web/>.
44. R. M. K. T. Rathnayaka, J. Wei, D. M. K. N. Seneviratna, Geometric Brownian motion with Ito's lemma approach to evaluate market fluctuations: A case study on Colombo Stock Exchange, In: *2014 International Conference on Behavioral, Economic, and Socio-Cultural Computing*, 2014, 1–6. <https://doi.org/10.1109/BESC.2014.7059517>
45. I. Karatzas, S. E. Shreve, Brownian motion, In: *Brownian motion and stochastic calculus*, New York: Springer, 1998, 47–127. https://doi.org/10.1007/978-1-4612-0949-2_2

Appendix A

Excel file is provided as supplementary data set.

Appendix B

MATLAB computational program for solving the S-I-R model

```
clearall; clc;
```

```
%% Parameters
```

```
Pre_infec = 5.2;
```

```

f = 1/Pre_infec;

Duration = 14;
r=1/Duration;

R_0 = 2.2; % A single infected person will infect about 2.2 others in a totally susceptible population

N = 21670000; % Population of Sri Lanka (2020)
beta = R_0/(N*Duration);

%% Differential equations
tspan = 0:1:365; % We will observe what happens over the next year
y0 = [N-3733, 262, 3733, 3259, 13]; % Coronavirus count of Sri Lanka (Mar 12, 2020)

[t,y]=ode45(@(t,y) ode_fun_simple(t,y,beta), tspan, y0);
% ode45 is a built in Numerical Differential Equations solver in MATLAB.
% There are packages for this solver in other programming languages.

%% plot
plot(t,y,'LineWidth', 1.5, 'MarkerSize', 18)
%legend('Susceptible','Deaths','Infectious','Recovered', 'Location', 'Best')
xlabel('Days after October 05, 2020')
ylabel('Population')
title('Predicted Spread of COVID - 19 in Sri Lanka')
gridon;
gridminor;
set(gca, 'FontSize', 36)
%saveas(gcf, 'corona_simple1.png')

```



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)