Trends in Computational and Applied Mathematics, **24**, N. 3 (2023), 459-473
Sociedade Brasileira de Matemática Aplicada e Computacional
Online version ISSN 2676-0029
www.scielo.br/tcam
doi: 10.5540/tcam.2023.024.03.00459

Stochastic Modeling of a Measles Outbreak in Brazil

M. LAU[1*] and Z. G. ARENAS[2]

Received on December 22, 2021 / Accepted on November 26, 2022

ABSTRACT. Development of mathematical models and its numerical implementations are essential tools in epidemiological modeling. Susceptible-Infected-Recovered (SIR) compartmental model, proposed by Kermack and McKendrick in 1927, is a widely used deterministic model which serves as a basis for more involved mathematical models. In this work, we consider two stochastic versions of the SIR model for analysing a measles outbreak in Ilha Grande, Rio de Janeiro, in 1976; Continuous Time Markov Chain and Stochastic Differential Equations. The SIR Continuous Time Markov Chain model is used to extract specific information from the measles outbreak. The outbreak probability, final size distribution and expected duration of the epidemic were computed, obtaining results in excellent agreement with the reported epidemic values. Numerical simulations are performed in Python.

Keywords: stochastic epidemiological models, SIR model, measles outbreak.

1 INTRODUCTION

Measles virus, *Measles morbillivirus*, have affected human populations since many centuries ago, even before the Common Era [8]. Measles is an airborne, highly contagious disease, which may cause mortality, mainly on childhood and in developing countries. In reference [16], a detailed study is carried out, reporting on incidence of measles in global level, vaccine coverage, and risk factors from 1990 to 2019. Vaccination has been a fundamental strategy for measles control. Even so, in 2018, there were more than 140.000 deaths worldwide, mainly children under 5 years [19]. Since that time, Brazil has faced some outbreaks, losing its certification of measles eradication, previoulsy obtained in 2016 [13]. A careful epidemiological surveillance is maintained by the Brazilian Health Ministry and its associated institutions.

For studying infectious diseases and keeping an epidemiological controlled situation, mathematical modeling appears as a major importance tool. Among several models, compartmental models

*Corresponding author: Michelle Lau – E-mail: miiichellelau@gmail.com
[1]Programa de Pós-graduação em Ciências Computacionais, IME, Universidade do Estado do Rio de Janeiro, R. São Francisco Xavier, 524, 20550-900, Rio de Janeiro, RJ, Brazil – E-mail: michelle.lau@ime.uerj.br https://orcid.org/0000-0002-0607-7426
[2]Departamento de Matemática Aplicada, IME, Universidade do Estado do Rio de Janeiro, R. São Francisco Xavier, 524, 20550-900, Rio de Janeiro, RJ, Brazil – E-mail: zochil@ime.uerj.br https://orcid.org/0000-0001-8377-2687

have been extensively used [2, 5]. In the compartmental modeling, population is divided into compartments and some relations among these compartments are assumed. SIR model is one of the basics compartmental models, useful for studying diseases that confer inmunity. It has been widely applied to chilhood diseases [1], as it is measles. Its name refers to the considered compartments: susceptible, infected and recovered (or removed) classes. Use of compartmental models for studying disease transmission was primarily done by W.O. Kermack and A.G. McKendrick, in early 20th century [11]. The SIR compartmental model is used as a basis for the construction of several more complex models. In addition to the compartmental SIR model for analyzing measles outbreaks [7], there are also more complex models and approaches that use, for example, birth and death rates, vaccination rates, probabilistic cellular automata based models, among others [4, 9, 15, 20]. For the purposes of this study, it is enough to consider a basic SIR model. Simulation results will be compared with an outbreak results in Brazil, in an isolated island and with a poor vaccination coverage.

It is commonly considered a deterministic dynamics for basic compartmental models, built upon differential equations. In this case, the evolution of the system is completely determined for given initial conditions. On the other hand, stochastic models [1] can be taken into account by using probabilistic concepts. In contrast with a deterministic approach, stochastic models allow obtaining a distribution of possible populations behaviours, included the possibility of having an asymptotic convergence to a limit state different from its deterministic counterpart expected result. Stochastic epidemic models own other properties that are unique for this approach, like computing the probability of an outbreak, the expected duration of the epidemic and its final size distribution.

In this work, we consider two stochastic versions of the SIR model for analysing a measles outbreak in Ilha Grande, Rio de Janeiro, in 1976 [3]. We use a Continuous Time Markov Chain [2], which allows a detailed follow-up of the epidemic dynamics. Deterministic SIR model and its stochastic differential equation version were also applied for comparison purposes. We are interested in computing specific properties from stochastic modeling for this particular outbreak, as its expected duration and final size distribution. Through numerical simulations implemented in Python, we compute all the specific properties of stochastic modeling and it was possible to estimate a rate of spread for the epidemic outbreak in Ilha Grande. All the computed values through SIR CTMC model for this specific measles outbreak, were in good agreement with the reported results [3], showing a better behaviour than the deterministic SIR model. From the simulation, it was possible to estimate a rate of spread for the epidemic outbreak, in good agreement with the expected value for measles disease. All the specific properties of stochastic modeling were computed. Numerical simulations were implemented in Python.

In the next section, deterministic SIR model is introduced, as well as the considered stochastic versions. In Section 3, results of simulations are discussed and, finally, conclusions are presented in Section 4.

2 THE SIR MODEL

As it was introduced above, SIR model is used for modeling diseases which have no re-infection of the individuals who were infected. In consequence, the passage of individuals among the compartments is a flux, from the susceptible class S, passing through the infected class I, to the removed compartment R. In the basic SIR model, the total number of individuals in the population is assumed to be constant, say N, which means that there are no births or deaths during the evolution period of the disease. It will also be considered that there is no latent period, a period in which those infected do not yet transmit the disease.

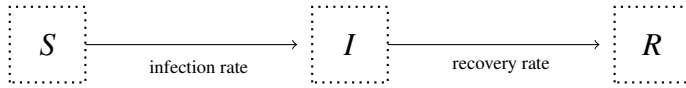


Figure 1: SIR epidemic model diagram.

Figure 1 represents the diagram of the model's operation. The susceptible individuals class communicates with the infected individuals class through an infection rate and the infected individuals class communicates with the recovered individuals through a recovery rate.

The concept of the basic reproduction number, $\mathscr{R}_0$, plays a fundamental role in epidemiological modeling. It is used to measure the transmission potential of the disease and represents the average number of secondary cases that an infectious individual can produce, considering a fully susceptible population. It is defined as

$$\mathscr{R}_0 = \frac{\text{infection rate}}{\text{recovery rate}}$$

and allows to estimate when an epidemic will occur, being a parameter of high social value. If $\mathscr{R}_0 > 1$, then there is an outbreak of the disease in the population; if $\mathscr{R}_0 < 1$, the spread of the disease is controlled and there is no epidemic; finally, for $\mathscr{R}_0 = 1$, the disease is endemic, that is, there is an endemic equilibrium.

2.1 Deterministic SIR Model

The deterministic SIR epidemic model [11] can be stated through the following differential equations,

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N}SI \\ \frac{dI}{dt} &= \frac{\beta}{N}SI - \gamma I \\ \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{2.1}$$

with initial values $S(0), I(0) > 0, R(0) \geq 0$, such that $S(0) + I(0) + R(0) = N$. The parameter β is defined as the contact or infection rate and γ, as the recovery rate of an infected individual.

Combining the first two equations in the system (2.1) and assuming a single infected individual and no recovered ones at the beginning, initial conditions are $I(0)=1, S(0)=N-1$ and $R(0)=0$, and the following solution for the class of infected individuals at time t is obtained [12],

$$I(t) = -S(t) + N + \frac{\gamma N}{\beta} ln\left(\frac{S(t)}{N-1}\right) . \tag{2.2}$$

As time evolves, it is known that the number of infected individuals will decrease and so the epidemic will end. Consequently, for $t \to \infty$, $I(\infty) = 0$ and from equation (2.2), it is obtained

$$S(\infty) = N + \frac{\gamma N}{\beta} ln\left(\frac{S(\infty)}{N-1}\right). \tag{2.3}$$

Equation (2.3) gives a long-term generic solution for the class of susceptible individuals. A detailed deduction of this equation can be found in reference [6]. Similar expressions in literature can be seen in [17]. The final size of the epidemic is given by $R(\infty) = N - S(\infty)$, which represents all individuals who became infected and recovered (or eventually died).

2.2 Stochastic Modeling

There were considered two different stochastic versions of the SIR model, Continuous Time Markov Chain (CTMC) and Stochastic Differential Equations (SDE). We will refer to CTMC SIR and SDE SIR model in what follows. For stochastic modeling, compartments are considered as stochastic processes.

2.2.1 CTMC SIR Model

For implementing the CTMC SIR model [1], we consider the compartments S, I and R as being discrete random variables, $S(t), I(t)$ and $R(t)$, called states, for any time instant t in a given interval (continuous time process). The *chain* term indicates that the random variables are discrete. We will refer to the states of the system as pairs (S,I), indicating the number of individuals in compartment S and in compartment I, for each state. Summarizing, for this model, $S(t), I(t), R(t) = \{0,1,2,\ldots,N\}$ and $S(t)+I(t)+R(t) = N$, for each $t \in [0,\infty)$. Assuming that at the onset of the epidemic period there is no recovered individual, the initial distribution of the model is defined as $(S(0), I(0)) = (s_0, i_0)$, where $i_0 > 0$ and $s_0 + i_0 = N$.

From the system of equations (2.1), we have that only two of the random variables in the stochastic model are independent. Considering $S(t)$ and $I(t)$ as independent variables, we have a bivariate process $\{(S(t), I(t))\}_{t=0}^{\infty}$ and we will use its joint probability function, $p_{(s,i)}(t) = Prob\{S(t) = s, I(t) = i\}$, to determine the transition probability from one state to another. In this way, the model is defined through the transition probabilities indicated in Table 1. Each equation models the variation of only one individual in each compartment. The first equation illustrates the instant in which a susceptible individual becomes infected and the second one models the passage of an infected individual to the recovered compartment. The last equation ensures the normalization of

the transition probability. In the Table 1, $\Delta S = S(t+\Delta t) - S(t)$ and $\Delta I = I(t+\Delta t) - I(t)$ are the state variation for a small time Δt.

As previously mentioned, by means of the stochastic modeling, particular and important properties can be obtained for analyzing the spread of the disease. One of these properties is the probability of occurrence of a disease outbreak in the population, computed at the start of the epidemic, for a small value $I(0) = i$. It can be computed [18] as

$$\text{outbreak probability} = \begin{cases} 0 & ,\mathscr{R}_0 \leq 1 \\ 1 - \left(\dfrac{1}{\mathscr{R}_0}\right)^i & ,\mathscr{R}_0 > 1 \end{cases} \tag{2.4}$$

where $\mathscr{R}_0$ is the basic reproduction number.

As another particular property, stochastic modeling provides a distribution for the final size of the epidemic. The final size represents the strength of the epidemic, that is, how many individuals were infected in total. For obtaining this distribution, it will be necessary to consider the transition probabilities between all the possible states (s,i). To this end, all possible states will be ordered as follows:

$$(N,0),(N-1,0),(N-2,0),\ldots,(0,0),(N-1,1),(N-2,1),\ldots,(0,1),\ldots,(0,N).$$

So, there are $(N+1)(N+2)/2$ possible states. The vector of the probabilities for each state of the system can be defined by $p(t) = (p_{(N,0)}, p_{(N-1,0)}, \ldots, p_{(0,N)})^T$ and the transition rates between states can be determined.

Considering $(S(t),I(t)) = (s,i)$, the transition matrix elements of the Embedded Markov Chain P_Y [10] are determined from the transition probabilities between states. When the transition from state (s,i) to state $(s,i-1)$ occurs, an infected individual is recovered and its probability is

$$\tilde{p}_s = \frac{\gamma i}{\gamma i + (\beta/N) si} = \frac{\gamma}{\gamma + (\beta/N) s}.$$

Table 1: Transition probabilities between states for the SIR CTMC model.

$(\Delta S, \Delta I)$	Probability
$(-1,1)$	$\dfrac{\beta}{N} S(t) I(t) \Delta t + o(\Delta t)$
$(0,-1)$	$\gamma I(t) \Delta t + o(\Delta(t)$
$(0,0)$	$1 - \left[\dfrac{\beta}{N} S(t) I(t) + \gamma I(t)\right] \Delta t + o(\Delta t)$

On the other hand, when the transition is from the state (s,i) to the state $(s-1,i+1)$, a susceptible individual has been infected and its probability is

$$1-\tilde{p}_s = 1-\frac{\gamma}{\gamma+(\beta/N)s} = \frac{(\beta/N)s}{\gamma+(\beta/N)s}.$$

The transition matrix P_Y is very useful for calculating the final size of the epidemic. The epidemic ends when the infected compartment becomes empty, *i.e.*, $I(t)=0$. In general, for any population of size N, starting with an infected individual, which means $p_{(N-1,1)}(0)=1$, the maximum number of transitions until there are no more infected individuals, is $2N-1$. The situation of having no infected individuals is known as absorption. Once the system reachs an absorving state, it is no more possible to leave it.

Assuming that initially the number of infected individuals is 1 and there are no recovered individuals, that is, $S(0)=N-1$, $I(0)=1$ and $R(0)=0$, the probability associated with the final size of the epidemic can be obtained by calculating the absorption probabilities,

$$\lim_{t\to\infty}\sum_{s=0}^{N-1} p_{(s,0)}(t)=1.$$

If there are s susceptible individuals when the number of infected individuals reaches zero, the final size of the epidemic is $N-s$. Therefore, it is possible to find the absorption probabilities using the transition matrix P_Y. In particular, $\lim_{t\to\infty} p(t) = p(2N-1) = (P_Y)^{2N-1}p(0)$.

Finally, another specific property which stochastic modeling allows accessing is the expected duration of an epidemic. In the model, the duration of an epidemic corresponds to the time it takes to reach absorption, that is, the time T such that $I(T)=0$. For the SIR CTMC model, it can be calculated using the first passage time method. Time to absorption may be too short or too long depending on the modeling conditions, as the size of the population N and the value of $\mathscr{R}_0$, and the initial number of infected i [1]. Denoting by $\tau_{s,i}$ the expected duration of the epidemic with initial condition (s,i), this value is calculated from the system of equations

$$(\gamma i+(\beta/N)si)\left(\tilde{p}_s\tau_{(s,i-1)}-\tau_{(s,i)}+(1-\tilde{p}_s)\tau_{(s-1,i+1)}\right)=-1. \tag{2.5}$$

To perform stochastic simulations in a CTMC, we need to know the time distribution between successive events. That is, the time it takes to go from one state to another. This time is called the time between events, denoted by T_E, and it is defined as a random variable. In the SIR CTMC model, it can be calculated as

$$T_E = -\frac{\ln U}{\frac{\beta}{N}S(t)I(t)+\gamma I(t)},$$

where U is a uniform random variable defined in the range $[0,1]$.

2.2.2 SDE SIR Model

For the SIR SDE epidemic model [2], $S(t)$ and $I(t)$ denote again continuous random variables for susceptible and infected compartments and let denote $X=(S,I)^T$, $\Delta X=(\Delta S,\Delta I)^T$, in vector

notation. As it was seen in Section 2.1, infected individuals recover at a rate γI and susceptible individuals get infected at a rate $\beta SI/N$. Table 2 lists the probabilities for the two possible changes in the SIR epidemic model for a small time interval Δt. These changes refer to the possibility of a susceptible individual gets infected, $(\Delta S = -1, \Delta I = 1)$ or an infected individual gets recovered, $(\Delta S = 0, \Delta I = -1)$.

Table 2: Probabilities associated with changes between states for the SIR SDE model.

i	$(\Delta X)_i$	Probability (p_i)
1	$(-1,1)$	$\frac{\beta}{N} S(t) I(t) \Delta t$
2	$(0,-1)$	$\gamma I(t) \Delta t$

The expectation $\mathbb{E}(\Delta X)$ and covariance $Cov(\Delta X)$ for the change in the two populations are computed from the probabilities defined in Table 2. The expectation can be expressed as

$$\begin{aligned}\mathbb{E}(\Delta X) &= \Delta X \cdot \text{probability} \\ &= \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \beta SI/N \\ \gamma I \end{pmatrix} \Delta t \\ &= \begin{pmatrix} -\beta SI/N \\ \beta SI/N + \gamma I \end{pmatrix} \Delta t\end{aligned}$$

The lines of matrix ΔX represent the transitions that may occur in each compartment. The first line represents transitions that occur in the susceptible compartment and the second line, transitions that occur in the infected compartment. In addition, the covariance matrix associated with these changes is a 2×2 matrix such that

$$\begin{aligned}Cov(\Delta X) &= \mathbb{E} \begin{bmatrix} (\Delta S)^2 & (\Delta S)(\Delta I) \\ (\Delta S)(\Delta I) & (\Delta I)^2 \end{bmatrix} \\ &= \begin{pmatrix} \beta SI/N & -\beta SI/N \\ -\beta SI/N & \beta SI/N + \gamma I \end{pmatrix} \Delta t\end{aligned}$$

Diffusion matrix for standard Itô SDE equations involve the computation of the square root of the covariance matrix. Instead of this, another matrix B is shown to have the property that $BB^T \Delta t = Cov(\Delta X)$ and can be used for defining an equivalent Itô SDE model for the SIR epidemic process as $dX = \mu dt + BdW^*$, where $W^* = (W_1^*, W_2^*)$ is a vector of two independent Wiener processes. Let denote the i-th change in Table 2 $(\Delta X)_i$, for $i = 1,2$, as $(\Delta_{1i}, \Delta_{2i})^T$. Each component represents the amount and direction (sign) of change in the variables Δ_{1i} and Δ_{2i}. The (i,j) entry in matrix B is defined as

$$B_{ij} = \Delta_{ij} \sqrt{p_j / \Delta t}. \tag{2.6}$$

Now, the 2×2 matrix B can be computed from the entries in Table 2 and the equation (2.6), obtaining

$$B = \begin{pmatrix} -\sqrt{\beta SI/N} & 0 \\ \sqrt{\beta SI/N} & -\sqrt{\gamma I} \end{pmatrix}.$$

Therefore, the explicit form for the SDE epidemic model, in this case, is

$$\begin{aligned} dS &= -\frac{\beta}{N} SI dt - \sqrt{\beta SI/N} dW_1^*, \\ dI &= \left(\frac{\beta}{N} SI - \gamma I\right) dt + \sqrt{\beta SI/N} dW_1^* - \sqrt{\gamma I} dW_2^*. \end{aligned} \tag{2.7}$$

3 NUMERICAL RESULTS

The parameters values used for performing numerical simulations were taken from a report on an epidemic measles outbreak in a fishing village in the Ilha Grande region, in the state of Rio de Janeiro, Brazil, in 1976 [3]. The total population of residents in the village was 453 individuals. The disease spreads from one infected person who traveled to the location. Thus, we consider $I(0) = 1$ as the initial condition.

Table 3: Incidence of measles according to age. Information extracted from [3].

Age group	Total cases	Resident population
< 1	-	11
1 a 5	16	66
6 a 10	20	60
11 a 15	11	53
16 a 19	3	41
> 19	-	222

In Table 3, the values of the population size by age group are listed. It can also be observed that the incidence of the disease affects more individuals between 1 and 19 years, that is, children and teenagers. In the analysis of a measles outbreak, there are some age groups that do not get infected. It is the situation for individuals over 19 years, who have been previously vaccinated or are considered to have natural immunity, given that a measles outbreak have occurred five years earlier, and also for children under 1 year, who are supposed to have maternal immunity [3]. Consequently, population group aged between 1 and 19 years was defined as the susceptible population, comprising 220 individuals. Summarizing, at the initial moment, total population was divided as $S(0) = 220$, $I(0) = 1$ and $R(0) = 233$. The measles epidemic in this location lasted 60 days and reached a total of 50 individuals, with no deaths occurring.

Despite measles is a highly contagious disease, conditions for contagion and spread depend on several factors, such as population density, age group, vaccination policies, and others [14]. Due to this setting, there exist a range of values for the basic reproduction number $\mathscr{R}_0$, which is not a

universal constant. Finding this value for a particular epidemic case is tricky and very important for modeling. With no previous information on $\mathscr{R}_0$ or the contagion and recovering rates, these parameters can be computed through several analytical or numerical techniques, including least square principle [7], Next Generation Matrix, graph theory and others (see reference [4] and references therein). In this study, it is used the basic reproduction number $\mathscr{R}_0 = 6.92$ as it was defined in the reported case study, reference [3].

The average duration of infection was 9.5 days, which is the mean time an individual takes to recover. This value is used to estimate the recovery rate, so $\gamma = 2/19$. Consequently, the infection rate was set to $\beta = \mathscr{R}_0\gamma \approx 0.7284$.

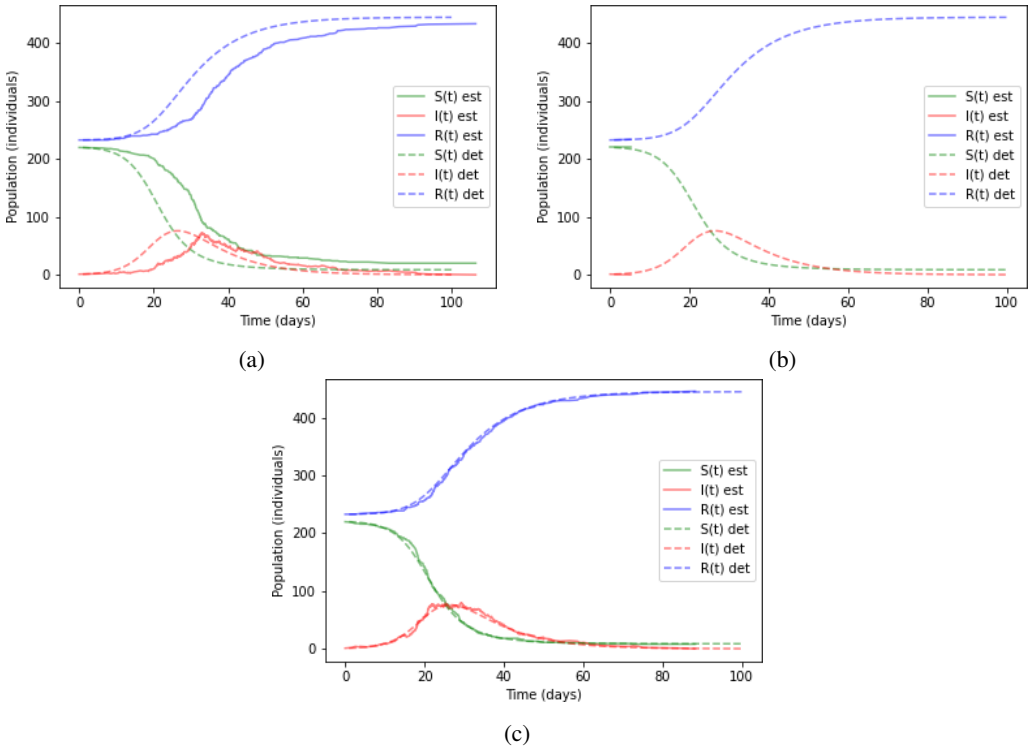


Figure 2: Diversity of possible scenarios obtained from CTMC SIR model (solid line) versus the unique evolution of deterministic SIR model (dashed line) given the same initial condition in both models for all the simulations. While in Figure 2a, results from both models differ, the outcomes in Figure 2c are very similar. At the same time, in each case, epidemic occurs at different moments in time. In Figure 2b, CTMC SIR evolves for a non-epidemic condition.

Numerical simulations were performed in Python. Some of the possible scenarios obtained from numerical simulation for the evolution of measles are shown in Figure 2. In each sub-figure, a single simulation of the disease dynamics is depicted, comparing the outcome of the stochastic

CTMC SIR model (solid line) with the result of the deterministic SIR model (dashed line). All these simulations were performed for the same initial condition.

Figure 2 illustrates the fact that, while the deterministic model offers a unique dynamics for the evolution of the epidemic outbreak given a specific initial condition, for the stochastic modeling a diversity of possible situations may be obtained. Besides the fact of having different results from both models in Figure 2a, as well as highly similar outcomes in Figure 2c, stochastic modeling brings the scenario of having the epidemic situation with different periods of time. This information could be useful for evaluating control strategies. On the other hand, a dynamics which does not evolve to an epidemic phase, also generated by the stochastic model, was illustrated in Figure 2b. For this reason, for stochastic modeling, it is necessary to perform a high number of simulations and analise them statistically. It is worth noting that, changing the initial condition, the deterministic model will offer a different solution. Analysing a unique solution is not the purpose of the epidemiological modeling.

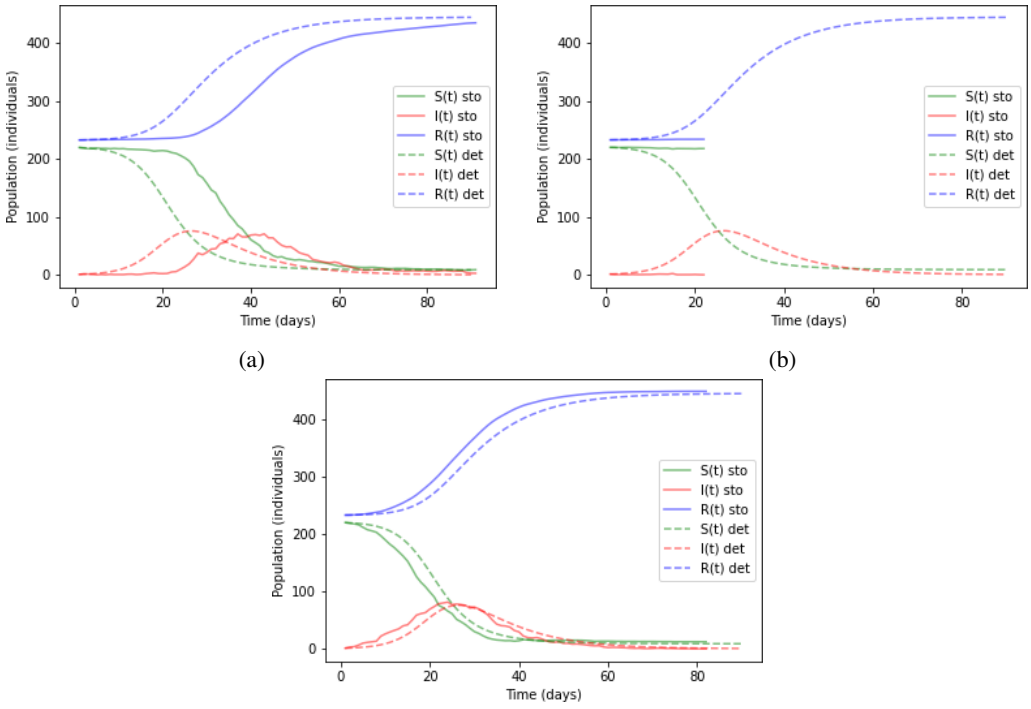
Numerical simulations for SDE SIR model defined by equations (2.7) were also performed. It was used Euler-Maruyama scheme for integrating the SDE equations. As expected from a stochastic model, several different scenarios were obtained for its dynamics, as depicted in Figure 3. Conditions like non-epidemic evolution (Figure 3b), visible different (Figure 3a) or highly similar (Figure 3c) to deterministic SIR evolutions are once again illustrated. Once again, as it was the case in Figure 2, epidemic may occur in different time periods for the SDE modeling. It can be see in Figures 3a and 3c and could help to evaluate possible control strategies.

Although through stochastic differential equations models, it is also possible to compute many specific properties for epidemic evolution, this work focused on computing these properties by using CTMC SIR model.

From equation (2.4), it is possible to determine the probability of occurrence of the disease outbreak in the population. With the previously defined value of $\mathscr{R}_0$, we obtain a result of 0.85 for this property, that is, the probability of the disease spreading in the population is of 85%.

In stochastic modeling, all possible scenarios have a certain probability of occurrence, including that of non epidemic condition depicted in Figure 2b. Thus, to obtain valid results using the stochastic model, a large set of simulations must be performed. In order to carry out a valid statistical analysis of the results, we performed 1.000 simulations of the CTMC SIR model.

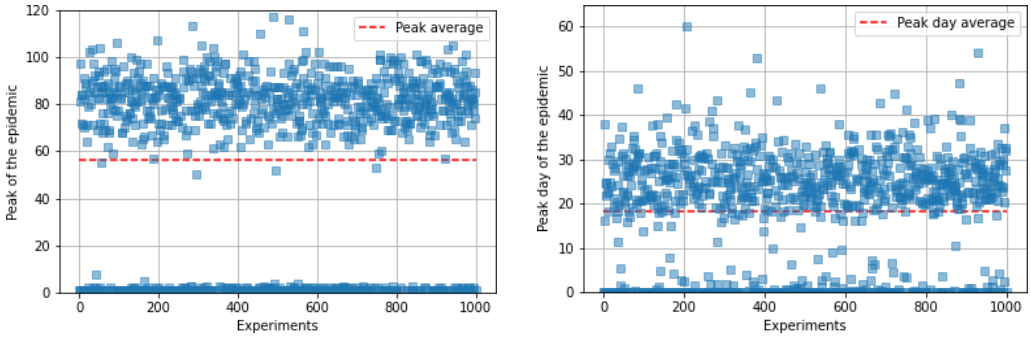
As several scenarios may result from the CTMC SIR model, averages of some quantities of interest were computed from the results of the 1.000 simulations carried out. It was found that the peak of the epidemic would happen, on average, with the infection of 56 individuals around the twentieth day, a reasonable number for the description of the reported measles outbreak, considering that 50 individuals were infected in total during the timeframe. In contrast, the deterministic SIR model provided an estimate of 75 infected individuals on the peak day, appearing as the 25th day. These results are shown in Figures 4a and 4.



(a) (b)

(c)

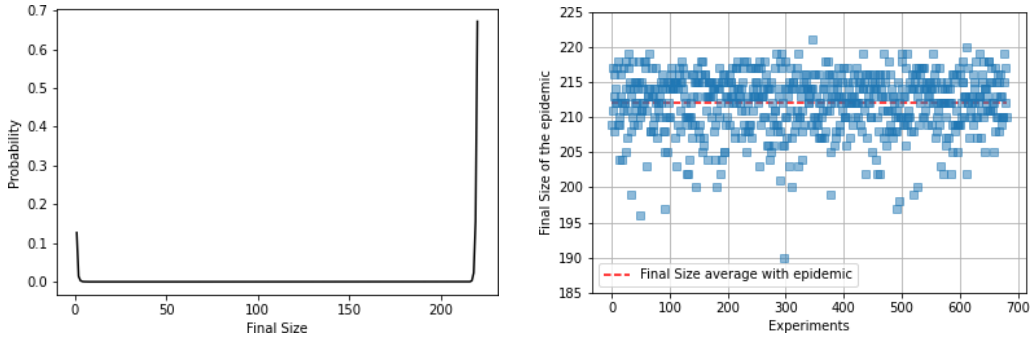
Figure 3: Diversity of possible scenarios obtained from SDE SIR model (solid line) versus the unique evolution of deterministic SIR model (dashed line) given the same initial condition in both models for all the simulations. While in Figure 3a, results from both models differ, the outcomes in Figure 3c are similar. At the same time, in each case, epidemic occurs at different moments in time. In Figure 3b, SDE SIR evolves for a non-epidemic condition.



(a) (b)

Figure 4: Average values computed from 1.000 CTMC SIR simulations carried out. 4a) Mean (red dashed line) of epidemic peaks for measles outbreak. 4b) Average day (red dashed line) for the ocurrence of the epidemic peak.

Another important advantage of stochastic modeling is the possibility of generating a probability distribution for the final size of the epidemic. This distribution, in the case under study, is shown in Figure 5a. Analyzing this probability distribution, an initial peak can be observed, which represents the probability of not having an epidemic outbreak in the population. The probability of having an outbreak is distributed for all possible final size values of the epidemic, with the highest probability associated with the final size of approximately 221 contagions. This last value is related to the original susceptible population and follows from the SIR model dynamics. These results are illustrated in Figures 5a and 5b.



(a) (b)

Figure 5: 5a) Probability distribution of final size of the epidemic. 5b) Average (red dashed line) for the final size.

Finally, stochastic modeling offers the possibility of estimating the expected duration of the epidemic. From a theoretical approach, the value for the expected duration, computed from equation (2.5) with $(s,i) = (220,1)$ as the initial state, is approximately 57.56 days.

It is also possible to compute the average duration of the epidemic based on the results of the performed simulations. Computed from the simulations, the obtained result was 59.84 days, in accordance with the theoretical result. This value is depicted as a red dashed line in Figure 6. Both theoretical and experimental results obtained for the expected duration of the epidemic are very close to the real value of the duration of the measles outbreak, that was of 60 days, as reported in [3].

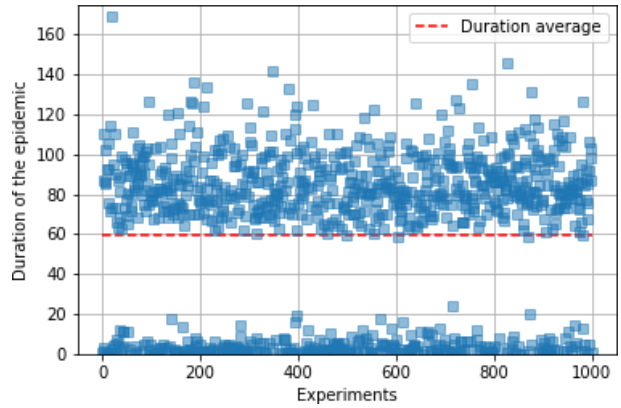


Figure 6: Mean value for the epidemic duration (red dashed line), 59.84 days.

4 CONCLUSIONS

Considering the incidence of measles in Brazil in recent years and the broad use of mathematical modeling for improving social and public policies, developing and adjusting stochastic models is of major importance. In this work, we have studied and implemented stochastic versions of the deterministic compartmental SIR model, using Continuous Time Markov Chain and Stochastic Differential Equations. We modeled a measles outbreak that took place in Ilha Grande, Rio de Janeiro, in 1976. Computer programming was done in Python, allowing code reusability.

By comparing the evolution of all the considered models, deterministic SIR, CTMC SIR and SDE SIR models, it was possible to validate their dynamics. The SIR CTMC model was used to extract specific information related to the possible evolution of an epidemic, which is not accessible through deterministic models. These properties are, basically, the probability of occurrence of the disease outbreak, the probability distribution for the final size and the expected duration of the epidemic. We carried out 1.000 simulations of the CTMC SIR model and averages values for those quantities of interest were computed. Through the stochastic modeling, we obtained results in excelent agreement with those values reported from the measles outbreak in Ilha Grande. In this way, advantages from stochastic modeling in relation to deterministic modeling became evident.

As future work, SDE SIR model will be used for computing averages values for the studied properties and the computational cost will be compared to that of the CTMC SIR model used in this work. Afterwards, it will be useful to extend the simulation to other epidemic outbreaks, adding effects of some measures (pharmacological or not) used to control the spread of the disease or even the population characteristics. Finally, studying seasonality of measles outbreaks to predict the strength of possible outbreaks will also be a future topic to be explored, from a stochastic modeling perspective.

Acknowledgments

The authors would like to thank Flavio C. Coelho and the anonymous reviewers, for their helpful comments and suggestions. Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001 are acknowledged for partial financial support.

REFERENCES

[1] L.J.S. Allen. An introduction to stochastic epidemic models. In "Mathematical Epidemiology". Springer (2008), p. 81–130.

[2] L.J.S. Allen. "An Introduction to Stochastic Processes with Applications to Biology". Chapman and Hall/CRC, Texas, USA (2010), 496 p.

[3] N.A. Araújo Filho & J.R. Coura. Surto epidêmico de sarampo na Ilha Grande, Rio de Janeiro, Brasil. *Revista da Sociedade Brasileira de Medicina Tropical*, **13**(1) (1980), 147–155.

[4] S.R. Arsal, D. Aldila & B.D. Handari. Short review of mathematical model of measles. In "AIP Conference Proceedings", volume 2264. AIP Publishing LLC (2020), p. 020003.

[5] F. Brauer. Compartmental Models in Epidemiology. In "Mathematical Epidemiology". Springer (2008), p. 19–79.

[6] M.L. de Almeida. "Modelagem epidemiológica com um modelo SIR estocástico utilizando Cadeia de Markov de tempo contínuo". Dissertação de Mestrado, Universidade do Estado do Rio de Janeiro, Rio de Janeiro (2020). Rede Sirius Biblioteca CTC-A, CDU 576.851.

[7] E. Dishmema & L. Hanelli. A SIR Model for Measles Disease Case for Albania. *International Journal of Innovative Research in Engineering and Management*, **6** (2019).

[8] A. Düx, S. Lequime, L.V. Patrono, B. Vrancken, S. Boral, J.F. Gogarten, A. Hilbig, D. Horst, K. Merkel, B. Prepoint, S. Santibanez, J. Schlotterbeck, M.A. Suchard, M. Ulrich, N. Widulin, A. Mankertz, F.H. Leendertz, K. Harper, T. Schnalke, P. Lemey & S. Calvignac-Spencer. Measles virus and rinderpest virus divergence dated to the sixth century BCE. *Science*, **368**(6497) (2020), 1367–1370.

[9] D. Ferraz & L. Monteiro. The impact of imported cases on the persistence of contagious diseases. *Ecological Complexity*, **40** (2019), 100788. doi:https://doi.org/10.1016/j.ecocom.2019.100788. URL `https://www.sciencedirect.com/science/article/pii/S1476945X19300972`.

[10] F.G. Foster. A note on Bailey's and Whittle's treatment of a general stochastic epidemic. *Biometrika*, **33** (1955), 123–125.

[11] W.O. Kermack & A.G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *The Royal Society*, (1927), 700–721. Vol. 115, No. 772.

[12] N.A. Kudryashov, M.A. Chmykhov & M. Vigdorowitsch. Analytical features of the SIR model and their applications to COVID-19. *Applied Mathematical Modelling*, **90** (2021), 466–473.

[13] E.A.S. Medeiros. Entendendo o ressurgimento e o controle do sarampo no Brasil. *Acta Paulista de Enfermagem [online]*, **33**(e-EDT20200001) (2020).

[14] E. Szusz, L. Garrison & C. Bauch. A review of data needed to parameterize a dynamic model of measles in developing countries. *BMC Research Notes*, **3**(75) (2010).

[15] G.T. Tilahun, S. Demie & A. Eyob. Stochastic model of measles transmission dynamics with double dose vaccination. *Infectious Disease Modelling*, **5** (2020), 478–494. doi:https://doi.org/10.1016/j.idm.2020.06.003. URL `https://www.sciencedirect.com/science/article/pii/S246804272030021X`.

[16] R. Wang, W. Jing, M. Liu & J. Liu. Trends of the Global, Regional, and National Incidence of Measles, Vaccine Coverage, and Risk Factors in 204 Countries From 1990 to 2019. *Frontiers in medicine*, **8** (2021).

[17] H.H. Weiss. The SIR model and the foundations of public health. *Materials matematics*, (2013), 0001–17.

[18] P. Whittle. The Outcome of a Stochastic Epidemic – A Note on Bailey's Paper. *Biometrika*, **42**(1/2) (1955), 116–122.

[19] World Health Organization. Measles (2019). URL `https://shre.ink/EOH`. Acesso em: 21-06-2022.

[20] L. Zhou, Y. Wang, Y. Xiao & M.Y. Li. Global dynamics of a discrete age-structured SIR epidemic model with applications to measles vaccination strategies. *Mathematical Biosciences*, **308** (2019), 27–37. doi:https://doi.org/10.1016/j.mbs.2018.12.003. URL `https://www.sciencedirect.com/science/article/pii/S0025556418302013`.