Estimate of the infection transmissibility of COVID-19 of Saudi Arabia outbreak: Estimation, and Simulating using Markov chain Monte Carlo

Salem Mubarak Alzahrani*
Faculty of Arts and Science in Almandaq, Al-Baha University, Al-Baha, Saudi Arabia* *Corresponding author; E-mail:* salemalzahrani@bu.edu.sa

Abstract:
The susceptible-infected-recovery-death model is used in this research project in order to replicate the illness epidemic that occurred in Saudi Arabia. The Markov chain Monte Carlo technique is used in the process of estimating and fitting the epidemic parameters and reproduction numbers from the transmission model to the real data. This is done with the help of the Markov chain Monte Carlo algorithm. A sensitivity analysis is performed with the use of the Sobol method in order to discover how sensitive the estimates are to changes in the values of the fixed parameters. The results of this analysis are then used to estimate the confidence intervals. The subsequent sensitivity analysis revealed that a number of parameters have an impact on the number of deaths, indicating that the authorities have a number of options available to them in order to lessen the harm caused by this virus. Additionally, the number of deaths was shown to be influenced by a number of other parameters.

Key words*: Heterogeneity, susceptible–infect1ed–recovered–dead, Markov Chain.*

1. INTRODUCTION

Since the COVID-19 epidemic broke out at the start of 2020, the whole globe has been in the midst of a serious and unprecedented international catastrophe (Zou et al., 2006, Diibendorfer, 2005, Elaiw, 2012.). One of the worst and potentially deadly respiratory disorders brought on by the virus is called acute respiratory distress syndrome (ARDS) (Elaiw, 2012). There has never before been an infectious illness with such a significant societal and economic consequence (Elaiw, 2014). It was a new coronavirus strain that caused the global outbreak of COVID-19 (Zou et al., 2006, Liang et al., 2018, Kephart et al., 1993, Marchette, 2004, Ren et al., 2012). Statistical and mathematical examination of the given data may shed light on the propagation of the virus's trend, allowing for the rapid implementation of a wide range of social actions. Analysis of epidemic data is also crucial for predicting future trends and understanding the underlying mechanisms involved in the development of the illness. This allows many groups to more precisely arrange their efforts to halt the spread.

These may be broken down into two broad categories: networked models (Jones, 1998, Elerian, et al., 2001, Eraker, 2001, Golightly, et al., 2006, Golightly, et al., 2008) and collective models (Mishra et al., 2011, Hosseini et al., 2018, Zhang et al., 2019, Bardhan et al., 2019, Batista et al., 2018, Deng and Chen, 2007, Deng et al., 2008). The susceptible-infected-recovered-dead (SIRD) model and its variants are instances of the former; together they are known as compartmental models (Ren et al., 2012, 14]. At

this juncture, it's worth noting that the kinetics of chemical processes in general, where a certain rate is connected with the transition from one state to another (Stramer et al., 2010), are very comparable to this occurrence. In epidemic modeling, the corresponding rates may be written in terms of the current counts of infected and cured people. Due to the complex interplay of various mechanisms driving infection transmission and recovery, a good mathematical model should be able to concurrently predict the temporal behavior of infected, recovered, and deceased people. The protective techniques need quarantine, confinement, social distancing lockdown measures, and so on, making a simple SIRD model unable to capture such complicated processes in general (Zhou et al., 2006),. The propagation of the current pandemic is modeled using a generalized version of the SIRD framework, which accounts for the percentages of the population that are exposed, quarantined, confined, actively infected, recovered, or dead at time instant t. We have shown that our model successfully accounts for the most recent data available for this country. Notably, the present approach is novel in that it accounts for all three reported data sets at once: a living, recouping, and departed populace. The model was originally implemented in Python before being incorporated into the Open Turns uncertainty calculation library. The data was retrieved from https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases, which is compiled by the Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE) from various sources, including the World Health Organization.

Talawar and Aundhakar (Talawar and Aundhakar, 2016) presented a model for infectious diseases, and we revisit it here. The eliminated component from the standard SIR model was decomposed into two categories, namely, the recovered and the dead. Additionally, the COVID19 model given here takes use of the infectious illness concept. For more discuccion about SIR models on epidemic diseases in KSA see Hassan, Taher S.; Elabbasy, et al.

In this study, we use a refined Susceptible-Infected-Recovered-Dead (SIRD) model to a time series of pandemic data for Saudi Arabia. In this investigation, the features and transmission mechanisms of COVID-19 were analyzed using a modified SIRD model that included the overall birth and death rates of the population. For this model, nine factors (including the beginning circumstances) had their probability laws estimated using data collected in Saudi Arabia from March 2, 2020 to January 4, 2022. The most crucial variables that regulate this pandemic's progression throughout time are estimated and compared. The SIRD model's parameters were determined using Markov chain Monte Carlo MCMC. The course of the outbreak was foreseen and accounted for by us. The use of a global sensitivity analysis offers a novel approach to resolving this issue. Indications of how far the epidemic's ravages were confined in Saudi Arabia were extracted using sensitivity analyses that made use of Taylor expansion and the creation of Sobol indices. In particular, we model the responses of public health officials and 1) estimate the virus's basic reproduction number, 2) forecast the epidemic's height, and 3) predict the disease's eventual extinction. Information for this research was given by the Saudi Arabian Ministry of Public Health. Results indicate that the highest incidence of COVID-19 in Saudi Arabia will occur

between June and August of 2020, with a reproduction rate of about 2.5 throughout the indicated time period. Moreover, with well-implemented public health measures, the epidemic curve might be lowered.

2. Method

2.1. Markov chain Monte Carlo

According to Gasparini's suggestion in Aundhakar (Gasparini, 1997), the Markov chain Monte Carlo (MCMC) method is as effective as any other method for estimating parameters. By using the MCMC method, we have a way to quickly pull samples from the posterior distribution. To avoid implicitly setting the essential, the samples may estimate the unknown parameters Aundhakar (Gilks, 1996). We anticipate a variety of possible outcomes $\pi(x)$. A nonperiodic and irreducible Markov chain with a simple distribution of may be constructed if $\pi(x)$ is too complicated to pattern directly $\pi(x)$. Assuming the Markov chain is sufficiently long, we may use it to infer key properties of by treating the simulated price as a pattern independent of the target distribution $\pi(x)$.

The posterior chance is determined using the prior chance and the probability function, as stated by Bayes' theorem. Let $\mathcal{M}$ represent the real cumulative broad variety of the unidentified infected individuals, with $\mathcal{M} = (\mathcal{M}_1.\mathcal{M}_2 \ldots, \mathcal{M}_n)$, and let Γ represent the actual cumulative wide variety of the unidentified infected computers, with $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$. It is important to remember that the inflammatory human population is consistent with the Poisson process, despite the fact that it is often believed that the population has made positive statistical errors and is incorrect. The probability of the cumulative wide variety of severely inflamed people is, thus, while the actual cumulative wide variety of inflamed people is m_t at the given time t, it is as follows:

$$p(\Gamma_t = z_t | \mathcal{M}_t = m_t) = \frac{m_t^{z_t}}{z_t!} e^{-m_t}.$$

Next, we assume that the parameter vectors are distributed uniformly throughout the set $\emptyset = (\beta, \gamma, \delta, \mu, \nu)$, where each parameter is treated as a separate vector. Let $L(\Gamma \lfloor \beta, \gamma, \delta, \mu, \nu)$ be the probability defining function, and choose a non-facts-earlier distribution, denoted by the constant $\emptyset^{\propto}$. As a result, we may express the likelihood of the posterior distribution as follows:

$$p(\beta, \gamma, \delta, \mu, \nu | \Gamma) = L(\Gamma \lfloor \beta, \gamma, \delta, \mu, \nu) p(\beta, \gamma, \delta, \mu, \nu | \Gamma)$$

$$\propto \prod_{j=1}^{n} p(z_j \lfloor \beta, \gamma, \delta, \mu, \nu)$$

$$\propto \prod_{j=1}^{n} \frac{m_t^{z_t}}{z_t!} e^{-m_t}.$$

First, given the prior distribution, and the initial time $t = 0$, derive the initial vector value $\emptyset^{(0)} = (\beta^{(0)}, \gamma^{(0)}, \delta^{(0)}, \mu^{(0)}, \nu^{(0)})$.

Second, draw the $\emptyset'$ from the concept distribution $q(\emptyset|\emptyset^{(t)})$ and draw a random number v from uniform $(0,1)$, calculate the popularity probability.

$$\alpha = min\left\{1, \frac{p(\emptyset\prime|z)}{p(\emptyset|z)}\right\}, \text{ if } \omega \leq \alpha,$$

Third, change t to $t+1$ and burn inside the potentially dangerous range of $\emptyset^{t+1}$.

Fourth. The loop closes when t becomes sufficiently large.

The Markov chain obtained using this method is very accurate in terms of convergence speed and stability, and the estimate outcomes of the version parameters are good as well.

2.2 Sobol Sensitivity Analysis

Sobol's method relies on a growing-dimensionality decomposition of the variance in the model's output into summands of variances in the input parameters (Saltelli et al., 1999, Sobol, 1993). Using a sensitivity analysis based on the Sobol function, one may ascertain the relative importance of each input parameter and the interactions between them in explaining the overall output variance. The purpose of a Sobol sensitivity analysis is to determine how much variation in model results depends on each of the input parameters, whether it be a single parameter or the interaction between many parameters. In a Sobol sensitivity analysis, the output variance is decomposed using the same principles as in a traditional analysis of variance in a factorial layout. Note that Sobol sensitivity analysis isn't necessarily aimed to identify the source of the entry variability. It just implies the kind and extent of the impact on version output. As a result, unlike a standard population PK/PD assessment, it cannot be utilized to identify the cause(s) of observed variation, such as the impact of demographic variables on total clearance. Selecting the appropriate version's output to utilize in the assessment is a crucial step in any sensitivity analysis, whether it local or worldwide. A positive time factor, such as Cmax for a concentration-time curve, or a metric that integrates changes withinside the model output of hobby over time, such as the duration of electrical depolarization, the included vicinity below a drug plasma concentration-time curve, or perhaps tumor size, are both acceptable choices. It is important to look at the question at hand while deciding the measure to employ. More often than not, the area under the concentration-time curve (AUC) is a more useful statistic of the version's performance. An included metric may expand over time for a variety of reasons, including but not limited to aging and disease (cf. ailment structures evaluation). The many uses of Sobol sensitivity analysis may be summarized as follows:

• Evaluation of the full spectrum of each input parameter variation and interactions between parameters;

• No assumptions made between version inputs and outputs

• The main downside is the high computational depth. Figure 2 depicts the key stages of the Sobol sensitivity analysis, which will be discussed in more depth in the following sections.

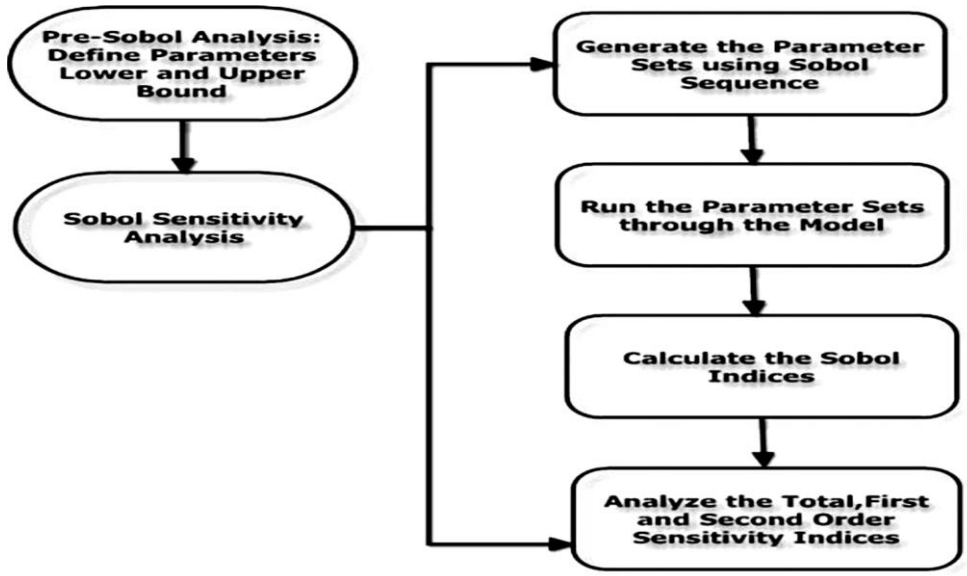

Figure 2: The method for carrying out a Sobol sensitivity analysis, including a drift chart.

Pre-Sobol and post-Sobol sensitivity analysis are the two most crucial phases. The Sobol sensitivity analysis consists of four stages: creating parameter sets, walking and simulating the version output with those settings, calculating and reading the total-, first-, and second-order and higher-order Sobol sensitivity indices, and last, reading the results. An initial parameter series is created through the Sobol series. I.M. Sobol of Russia initially suggested the idea of a quasirandomized, low discrepancy series now known as the Sobol series. Compared to random sequences, those with low discrepancies tend to pattern space more equally. Moreover, algorithms using such sequences may exhibit superior convergence (Sobol, 1967). It is ultimately possible to mimic version outputs using the created parameter units (Dalal et al., 2008). The following is an index to the popular Sobol series functions:

• The Sobol series, often known as the "quasi-random series," is a kind of low-discrepancy series.

• Distributed more consistently than the bogus random sequences.

• Faster convergence and better accuracy are the results of quasi-Monte Carlo integration.

• The large dimensionality of the integrals is a disadvantage.

Total-order, first-order, second-order, and higher-order sensitivity indices are calculated to accurately replicate the effect of the character enter, and the interaction between them (Sobol, 2001). This allows one to understand how the output variance can be attributed to character enter variables and the interplay among each of the enter variables. Let's say $x = (x_1, x_2, \ldots, x_s)$ represents the input parameters. After rescaling, it's safe to assume that all of the parameters in the finite c programming languages have values in the range [0,1]. Assuming that each parameter is independently determined and distributed equally on the interval [0,1] is helpful. An attribute of x, let's call it f, whose sensitivity to the entry arguments is being evaluated is the version's output $f(x)$. $f(x)$ is a random variable with a mean (f_0) and a variance (D) according to the probabilistic interpretation of the parameters:

$$f_0 = \int f(x)dx,$$

$$D = \int f(x)^2\, dx - f_0^2.$$

Each and every integral may be written as a pair of integrals with [0,1] as the limits in all dimensions. The Sobol method relies on breaking down D into its component parts, which might be the results of one parameter, the results of two parameters working together, and so on. In the first step, we decompose $f(x)$ into

$$f(x) = f_0 + \sum_{i=1}^{s} f_i(x_i) \; + \sum_{i=1}^{s} \; \sum_{i\neq j}^{s} f_{ij}(x_i, x_j) + \ldots\ldots + f_{1\ldots s}(1,2,\ldots.,\, x_s) \qquad (2.1)$$

Decomposition phrases are constructed as shown in [31]:

$$f_i(x_i) = \int f(x) \prod_{k\neq i} dx_k - f_0$$

$$f_{ij}(x_i, x_j) = \int f(x) \prod_{k\neq i,j} dx_k - f_0 - f_i(x_i) - f_j(x_j)$$

and so on.

As shown in Eq. (2.2), the assessment of variance illustration of $f(x)$ is dependent on the happiness of the scenario, as shown in (Sobol, 2001):

$$\int f_{i_1,i_2,\ldots\ldots i_s}(x_{i_1},\ldots.,x_{i_s})\, dx_k = 0, \quad \text{for } k = i_1, i_2, \ldots.. i_s. \qquad (2.2)$$

As a result of this characteristic, we get by squaring Eq. (2.1) and integrating:

$$D \; = \sum_{i=1}^{k} D_i \; + \sum_{i<j} D_{ij} + \sum_{i<j<m} D_{ijm} + \ldots + \sum_{i<2<\ldots,<k} D_{12\ldots k}, \qquad (2.3)$$

where

$$D_{i_1,i_2,\ldots\ldots i_s} = \int f^2_{i_1,i_2,\ldots\ldots i_s}(x_{i_1},\ldots.,x_{i_s})\, dx_{i_1}\ldots.dx_{i_s}$$

is the variance of $f_{i_1,i_2,\ldots..i_s}(x_{i_1},\ldots,x_{i_s})$, referred to as the partial variance similar to that subset of parameters. The Sobol sensitivity indices for that subset of parameters is then described as

$$S_{i_1,i_2,\ldots..i_s} = \frac{D_{i_1,i_2,\ldots..i_s}}{D}.$$

To determine the first-order contribution of the ith input parameter to the output variance, we may use $S_i = \frac{D_i}{D}$, while the second-order contribution of the interaction between the ith and jth parameters can be calculated using $S_{ij} = \frac{D_{ij}}{D}$. At last, general order sensitivity indices measure the overall effects of a single parameter at the model's output, and they may be thought of as the total of all the sensitivity indices: $S_{Ti} = S_i + S_{ij} + \ldots + S_{1..i\ldots s}$. From Eq. (2.3), we may get the individual sensitivity indices by dividing by D. Due to the fact that S_i connects the partial variance to the whole variance for each parameter, it can be shown using Eq. that the total of the sensitivity indices for all parameters must be at least one (2.4).

One may write this as:

$$\sum_{i=1}^{k} S_i + \sum_{i=1}^{k} S_{ij} + \ldots + \sum_{i=1}^{k} S_{ij\ldots k} = 1. \quad (2.4)$$

In conclusion, first-order sensitivity indices have the most bearing; they are used to measure the proportional contribution of a single parameter to the variance of the output. With the use of second-order sensitivity indices, we may quantify the relative importance of parameter interactions in determining the overall output variance. All first-order, second-order, and higher-order findings are recorded, as well as the overall assessment across all parameters, in total order sensitivity indices. Higher sensitivity index costs have a larger impact on the corresponding version parameters and related processes. Despite the fact that no exceptional cutoff value has been established, the rather arbitrary number of 0.05 is routinely used in this kind of assessment to differentiate between essential and inconsequential factors. Although this 0.05 criterion is usually used to more complex designs, it is crucial to remember that it may no longer be rigorous enough for designs that are reasonably easy and only have a few input parameters.

3. The Model Formulation

3.1. The SIRD model

In contrast to the standard SIR model, we are independently assessing the many different cases that ended in death and the recovered cases. Therefore, we may utilize variables like "Recovered" and "Deaths" instead of "Recovered + Deaths" in SIR's mathematical model $S \xrightarrow{\beta} I$

$\overset{\delta}{\rightarrow} R$, $I \overset{\delta}{\rightarrow} D$. Assumption: S = susceptible (= population - confirmed), I = infected (= confirmed - recovered - fatal), R = recovered, and D = deceased.

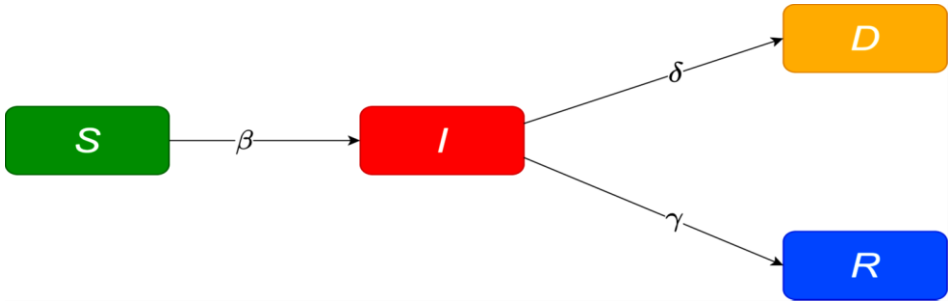

Figure 1 . The SIRD model shows the percentage of healthy (susceptible) people who develop inflammation after coming into touch with an inflamed person.

Illustration of the SIRD model. β represents the percentage of healthy (susceptible) folks that turn out to be inflamed after touch with an inflamed person. In different words, it's far the transmission price of the virus. γ denotes the healing price, and δ the mortality price related to the virus.

$$\frac{dS}{dt} = -\beta \frac{SI}{N},$$
$$\frac{dI}{dt} = \beta \frac{SI}{N} - (\gamma + \delta)I,$$
$$\frac{dR}{dt} = \gamma I,$$
$$\frac{dD}{dt} = \delta I,$$

where δ is the mortality rate [1/min], β the powerful touch rate in milliseconds, and the recovery rate [1/min]. $N = S + I + R + D$ represents the entire population, and t is the time since the start date. The virus's cost per duplicate is

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \delta}.$$

3.2 The SIRD model with births and deaths

In the SIRD version, everyone is assumed to remain unchanged during the pandemic. Due to the presumption that persons were healthy when they were born, we modified the first line of the SIRD version to read: $\mu(N - D) = \mu(S + I + R)$. For this reason, we decided to add

complexity to the SIRD model by including birth and death rates alongside initial capital investment. This figure shows how the aforementioned system behaves dynamically.

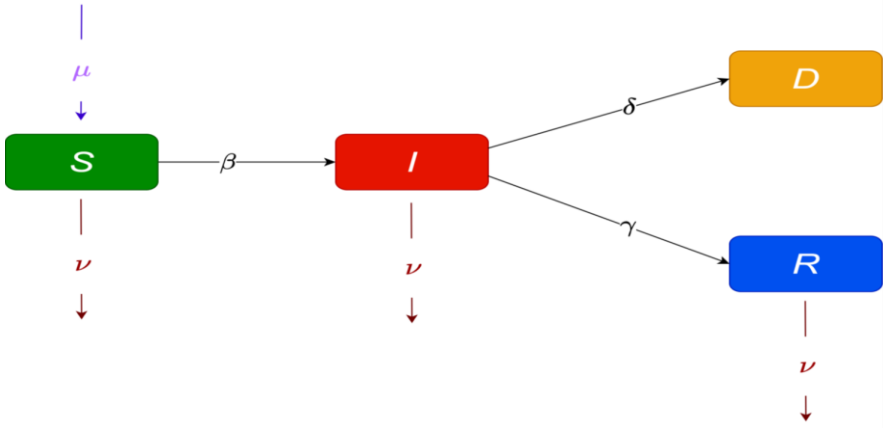

Figure 2 depicts the SIRD variant with the population's initial and final costs included.

Note that the definitions of and v are switched in this version, as opposed to the CNRS article (Bayette, and Monticelli, 2020). Summing the populations of the healthy (S), infected (I), and withdrawn (R) with the slanted (S) populations reveals that the newborns are obviously not dead. When S, I, and R populations are subtracted from those who died due to factors other than the virus, the following ODE system is generated:

$$\begin{cases} \frac{dS}{dt} = -\beta\frac{SI}{N} + \mu(S+I+R) - \nu S, \\ \frac{dI}{dt} = \beta\frac{SI}{N} - (\gamma+\delta)I - \nu I, \\ \frac{dR}{dt} = \gamma I - \nu R, \\ \frac{dD}{dt} = \delta I. \end{cases} \tag{3.1}$$

At $I = 0$, one gets $R = 0$. It is important to remember that in the SIRD model, the total population $N = S + I + R + D$ is not necessarily consistent, in contrast to the SIR and SIRD models. Sure enough, we have

$$\frac{dN}{dt} = (\mu - \nu)(S + I + R).$$

As a result, if $\mu > \nu$, the total population grows; if $\mu = \nu$, the population stays the same; and if $\mu < \nu$, the population shrinks. To get the virus's reproduction rate, $\mathcal{R}_0 = \frac{\beta}{\gamma+\delta+\nu}$, we solve the ODE for the number of infected (I) in the system (SIRD).

It is well knowledge that if $\mathcal{R}_0 > 1$, the virus will stop spreading, and if $\mathcal{R}_0 < 1$, the virus will start spreading again. Our current implementation is a Python program using a fourth-order Runge-Kutta algorithm. The subsequent repercussions model is provided by us.

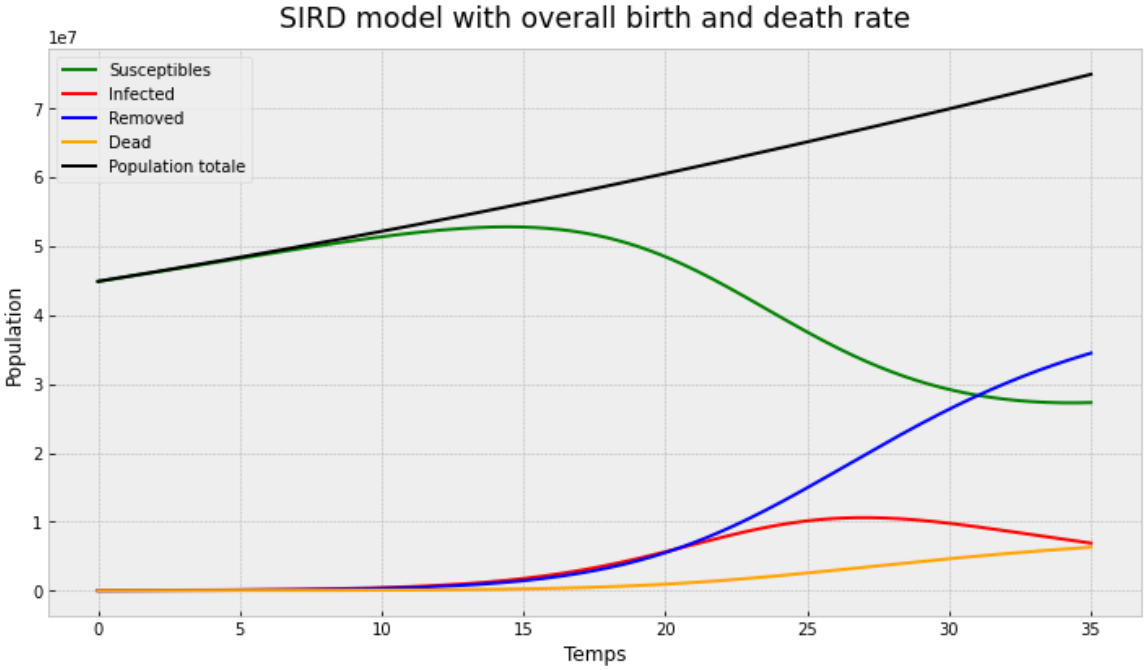

Figure 3 – The consequences of the implementation of the SIRD version with the coefficients $\beta = 0.615$, $\gamma = 0.193$, $\delta = 0.06$ fixed.

The initial conditions are $S_0 = 25216237 - 2e4, I_0 = 2e4$, $R_0 = 0$, $D_0 = 0$, and the simulation time is T = 35 days.

4. Results

4.1. Parameter estimation

The likelihood distributions shown here are estimated using the values of five version parameters: parameters $\beta, \gamma, \delta, \mu,$ and ν. This requires making distribution estimates for parameters $\beta, \gamma, \delta, \mu,$ and ν. New demographic data are needed to make this happen. The estimate of version parameters helps us predict the version with more precision. Parameter estimation poses significant challenges. Though many characteristics may be anticipated, the process is laborious and expensive. The SIRD Covid-19 version's unknown parameters are estimated using the Markov chain Monte Carlo (MCMC) technique in Equation (3.1).

4.1.1 Generation of the Input-Output Relationship

Starting in January 2020, at the beginning of the Coronavirus epidemic (see https://coronavirus.jhu.edu/data/new-cases) , the Johns Hopkins University CSSE 6 (CSSE, 2021a) started accumulating data. A associated GitHub repository with daily updates contains the relevant data (CSSE, 2021b). We obtained documents pertaining to Saudi Arabia between March 2, 2020 and January 4, 2022. The three time series that make up this document are shown in Figure 4; they are the ranges of cases that were shown, cases that were deceased, and cases that were retrieved.

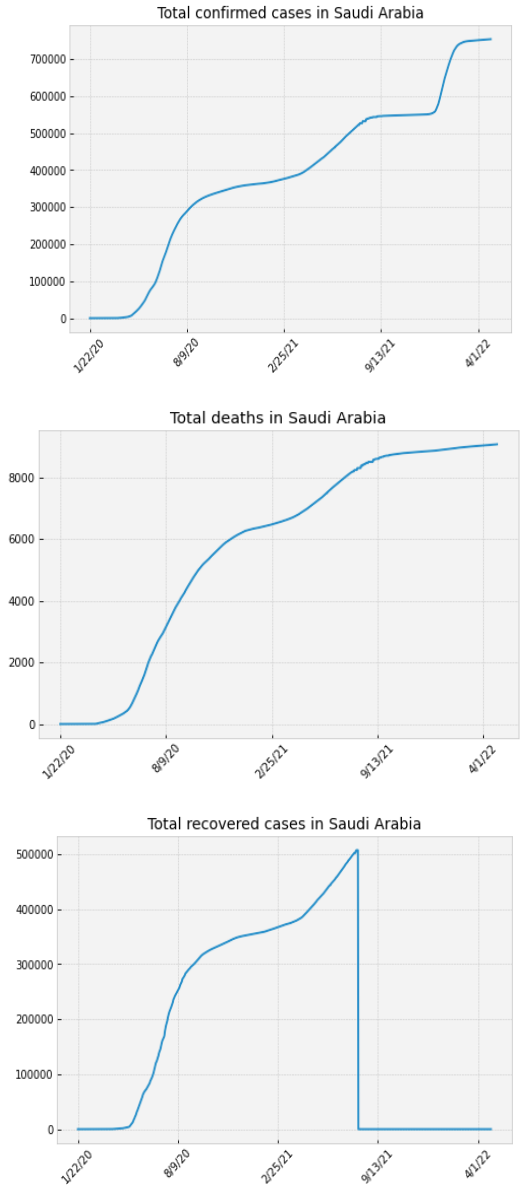

Figure 4 –The numbers of cumulative numbers.

Changes in COVID-19 in Saudi Arabia between January 22, 2020 and September 13, 2022 analyzed using data from the Johns Hopkins University database. Once the raw documents have been collected, we'll sort them into categories based on the characteristics of relevance to the SIRD model. No population census has been carried out in the area between the observation dates. On this day in 2022 (May 27), there are N=35,841,311 people living in Saudi Arabia. The great number of died instances, the large number of recovered instances, the large number of shown instances, and the large number of detected and recovered instances all have the following relationships with one another (discover 5). Do not forget that Susceptible is calculated by dividing the entire population by the actual number of occurrences.

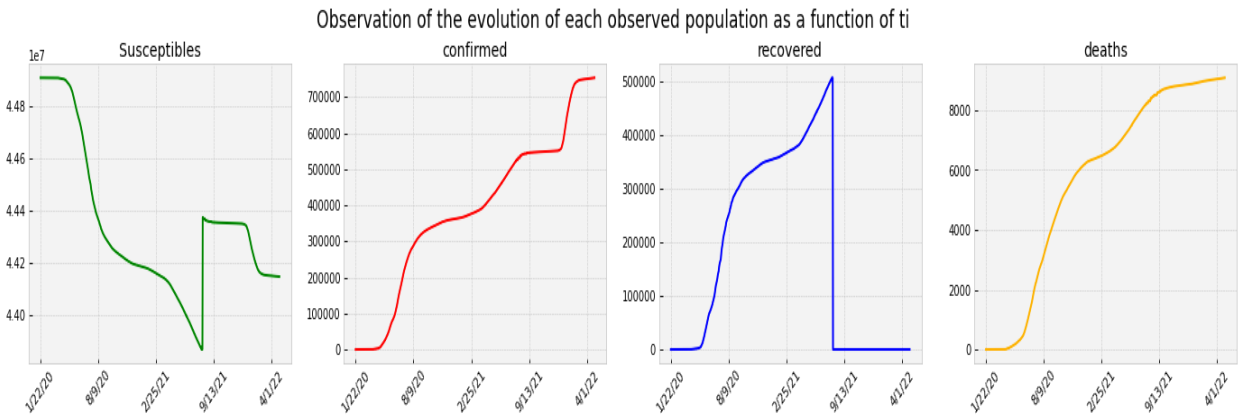

Figure 5. Pre-processed data in Saudi Arabia between January 22, 2020 and January 4, 2022.

After the data has been preprocessed and classified, the determined inputs and outputs may be made. But first, we need to zero down on a time period to perform our analysis. As was previously noted, the first COVID-19 instance was discovered in Saudi Arabia on March 2, 2020, and it was not seen again until January 4, 2022. Within this time frame, we undertake observations that are similar to those in a simulation over a period of 15 days. Out of them, we've been able to glean 85 that really work. Put another way, if the machine's country at time t is $X(t)$, then the value of the determined input is $X(t)$and the value of the determined output is $X(t+15)$. Consistent with the code below, whose results are shown in Figure 6. These data allow for an estimation of the parameters. Considering the brief period of the simulation (15 days) inside the observation window, it is vital to remember that the information has a linear shape, which we may take use of in the following.

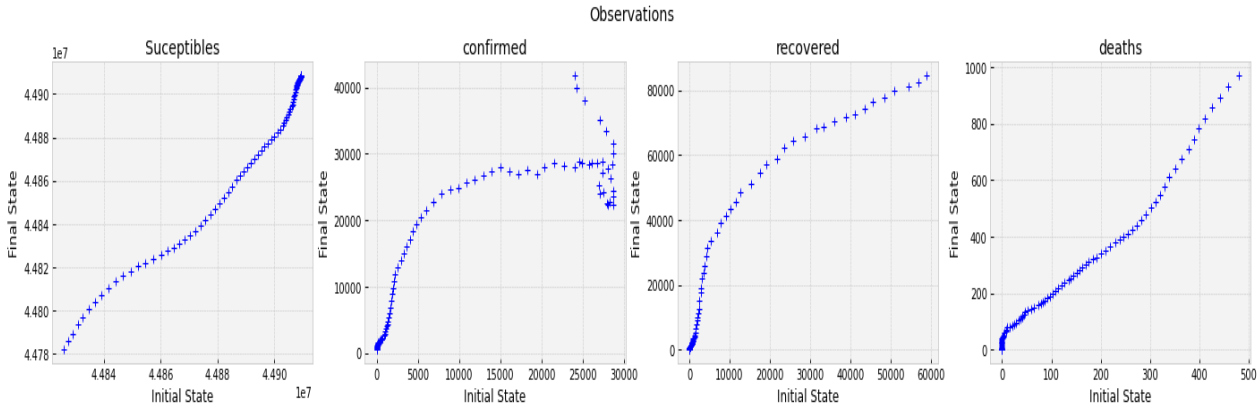

Figure 6. The observed processed data between March 2, 2020 and January 4, 2022.

4.1.2 Estimation of parameters by calibration

The data we've collected so far allows us to make estimates for the parameters in which we have the most confidence. To do this, we make use of Open Turns. As a first step in contributing our Python-based SIRD implementation to the Open Turns initiative, we create a Python Function type item. Then, we use the Linear Least Squares Calibration Method to make estimates for the parameters $\beta, \gamma, \delta, \mu, \nu$. Figure 6 shows that this estimator was chosen because of the linearity of the data it contains. By linearizing the model, this estimator helps to reduce the discrepancy between the estimated and observed values. Additionally, it generates a selectable state (named "prior") with values of 0.467, 0.4, 0.001, 0.0001, and 0.00019. After estimating these five parameters at random, the table below displays their chance distributions along with their corresponding normal distributions for means, standard deviations, and 95% highest and lowest CIs. We can estimate the parameters that have a high level of confidence in the records we have determined now that we have them.We use Open Turns for this. To begin, we develop a Python.

Parameters	Reference value	Mean	Standard deviation	Lower CI	Sup IC
β	0.181661	0.181661	0.00336377	0.155926	0.207396
γ	0.0554433	0.0554433	0.00116208	0.0465527	0.064334
δ	0.0148682	0.0148682	0.000431601	0.0115662	0.0181702
μ	0.00206816	0.00206816	0.00223205	-0.0150084	0.0191447
ν	0.00207206	0.00207206	0.00223195	-0.0150038	0.0191479

Table 1 – Results of parameter estimation of SIRD in KSA.

The "prior" price is the only one that the Open Turns linear least squares estimator has access to. The so-called "posterior" values, which were received after the estimate, are listed in the four columns that follow. It is exciting to note that the estimate is not always aberrant due to

the fact that the parameters μ and ν and are nearly identical and perfectly correlated after calibration, ensuring the experiment's general population's fidelity. Figure 7 shows how our parameter estimation system looks.

	β	γ	δ	μ	ν
β	1	0.948017	0.116023	0.988281	0.988251
γ	0.948017	1	-0.0328953	0.919966	0.919933
δ	0.116023	-0.0328953	1	0.0201733	0.0200674
μ	0.988281	0.919966	0.0201733	1	1
ν	0.988251	0.919933	0.0200674	1	1

Table 2 – Correlation matrix after estimating the different parameters of SIRD in KSA.

The following contains the five independent parameters that may be utilized with this evaluation matrix. Remember that a study quite comparable to ours was conducted by Nguemdjo et al. and published in August 2020. This research also used a SIR version to forecast the development of COVID-19 in Saudi Arabia from March 2, 2020 and January 4, 2022. The parameters β and γ, were determined using a maximum likelihood estimator and are listed below. Observe how the consequences shown above vary from ours in terms of the parameters β and γ and Notably, the virus spreads far more slowly amongst humans. That our selected time span, beginning on March 2, 2020, and ending on January 4, 2022, concludes just before the epidemic's peak, the lower in β provides a unique characterization of this differentiation. Between March 2, 2020 and January 4, 2022, the Saudi Arabian government has enough opportunity to beef up containment measures, lowering the chance of the virus spreading outside of the country.

4.2 Application of Sobol Sensitivity Analysis

In order to assess the robustness of our SIRD model, we shall conduct a sensitivity analysis. We decide to explore the model's impact on the initial parameters in addition to the parameters stated before. Finally, we may test how $S, I, R,$ and D vary depending on the values of the parameters $\beta, \gamma, \delta, \mu, \nu$. It is essential to initially collect distributions for all of these factors. Previous to this section, predictions were produced for the first five parameters. They will most likely act in accordance with standard protocol. The values of $S_0, I_0,$ $R_0,$ and D_0 in the beginning condition are shown in Figure 5 to adhere to the same patterns as $S, I, R, and\ D$. It is clear that Beta rules may be used as a close approximation for those regulations between March 2, 2020 and January 4, 2022. As an example, Table 3 shows the results of an open run when

this is implemented. Figure 9's QQ-plot convincingly demonstrates why we choose Beta distributions.

Parameters	**alpha**	**beta**	**a**	**b**
S_0	0.935242	0.348363	4.47806e+07	4.49106e+07
I_0	0.563381	1.02904	-405.235	42259.2
R_0	0.176244	0.586377	-830.588	85550.6
D_0	0.417359	1.34975	-9.52941	981.529

Table 3 – In light of the observations depicted in Figure 5, intuition-based estimation of the Beta laws and initial parameters.

Figures 10 and 11 show that the two most important elements affecting the distribution of vulnerable times for the Saudi Arabian population are the initial investment and the final investment v. The final inflamed range obtained from a simulation is primarily affected by the beginning inflamed range. This same range is critical for counting the final tally of survivors and casualties. Also, remember that the mortality rate δ of the virus is only one of several variables that affect the total number of deaths.

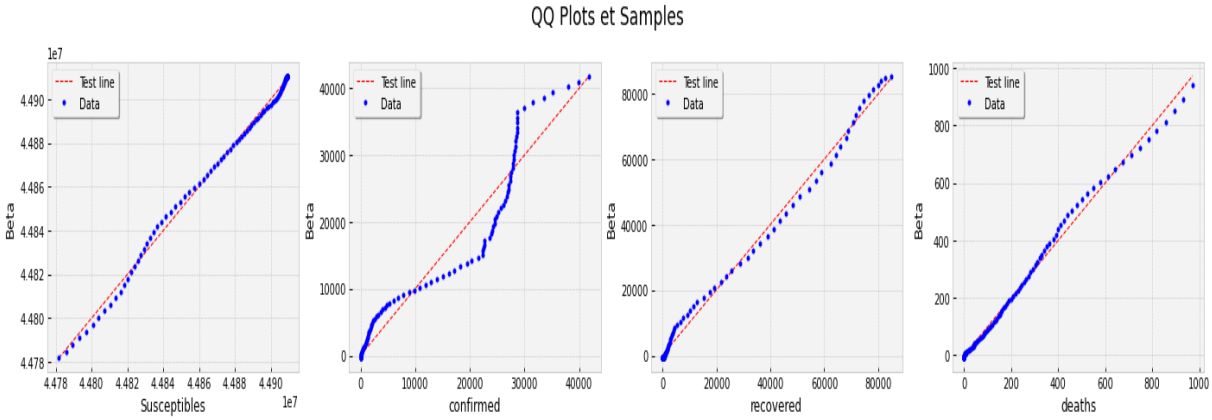

Figure 9 – Utilizing QQ-plots to confirm the selection of Beta laws for the initial conditions.

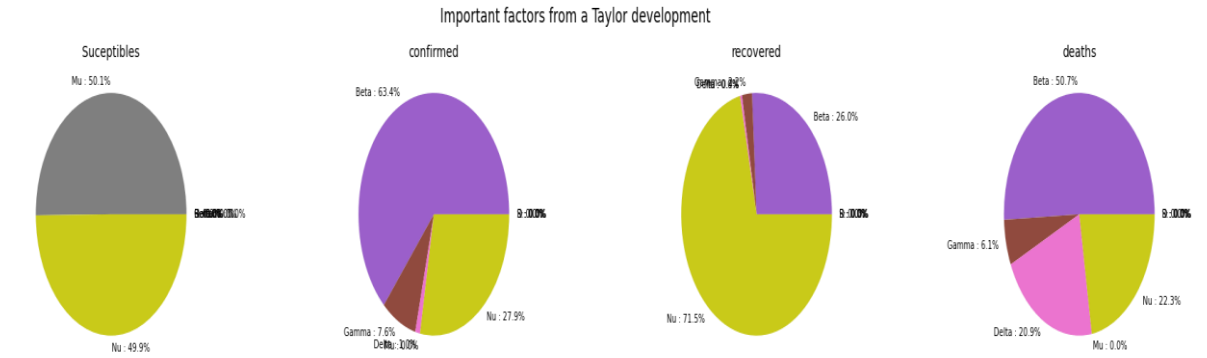

Figure 10 – Taylor expansion of the first order for sensitivity analysis.

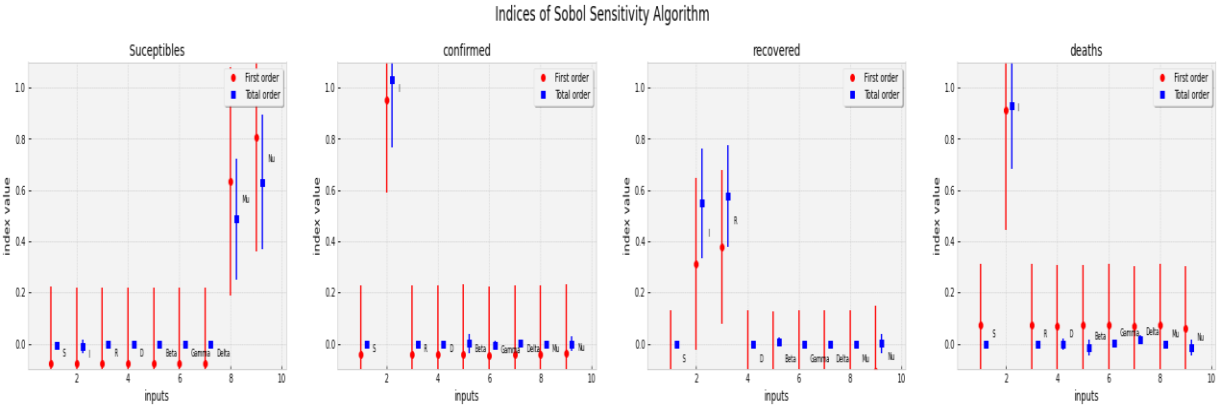

Figure 11 – The calculation of first-order and total Sobol indices for sensitivity analysis.

The results are independent of the beginning circumstances because of the linear character of this constrained growth. Table 2 shows that when two variables are evaluated and both have indices, the total index is usually different from the first-order index. This suggests that the variables interact. To test the robustness of our SIRD model, we shall conduct a sensitivity analysis. In addition to the parameters discussed before, we have opted to look at the effect the model has on its starting values. Last but not least, we'll have a look at how sensitive certain quantities are to changes in the initial values of the parameters $\beta, \gamma, \delta, \mu, \nu$. First, we need to collect data on the ranges within which these parameters tend to fall. The first five parameters $\beta, \gamma, \delta, \mu, \nu$ and, were estimated in the previous section. In this report's code, using the same study window as Nguemdjo et al. does work, but it decreases the amount of the observed data and introduces an accuracy loss when estimating parameters.

4.3 Epidemic Forecasting

4.3.1 Distribution of $\mathcal{R}_0$

Our projections of the virus's future evolution are based on its current spread and its reproduction rate, $\mathcal{R}_0 = \frac{\beta}{\gamma+\delta+\nu}$. The paper's second portion evaluated the parameters β, γ, δ, and ν by which the laws,, and were governed. Probability density of $\mathcal{R}_0$ is shown in the table below. For a more audacious forecast, we suggest simulating the behavior of each group for an additional 15 days after our research ends (between March 2, 2020 and January 4, 2022).

To rephrase: we know exactly how many people will call Saudi Arabia home on June 29, 2020. Here is the resulting picture. On the assumption that the distributions are uncorrelated and independent, the marginals of the posterior distribution are retrieved. There is a high probability that the virus will continue to spread, as it has been doing so in Saudi Arabia since January 4, 2022.

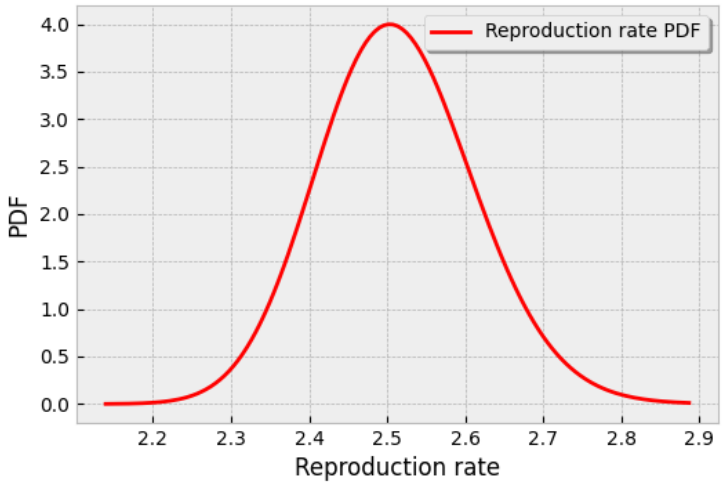

Figure 12. Propagation density function of a virus. With a mean of 1.04928 and a standard deviation of 0.0160945, this rule seems to be perfectly typical.

4.3.2 Determine the S, I, R, and D law parameters as a function of time.

When deciding which laws to follow, we rely on our gut feelings. These distributions will be used in the next sensitivity analysis. Both populations are shown in two separate figures due to the huge discrepancy in scales of measurement. This simulation comes extremely near to

describing the conditions in Saudi Arabia on the given day. By definition, the farther back in time you try to anticipate, the less accurate your results will be.

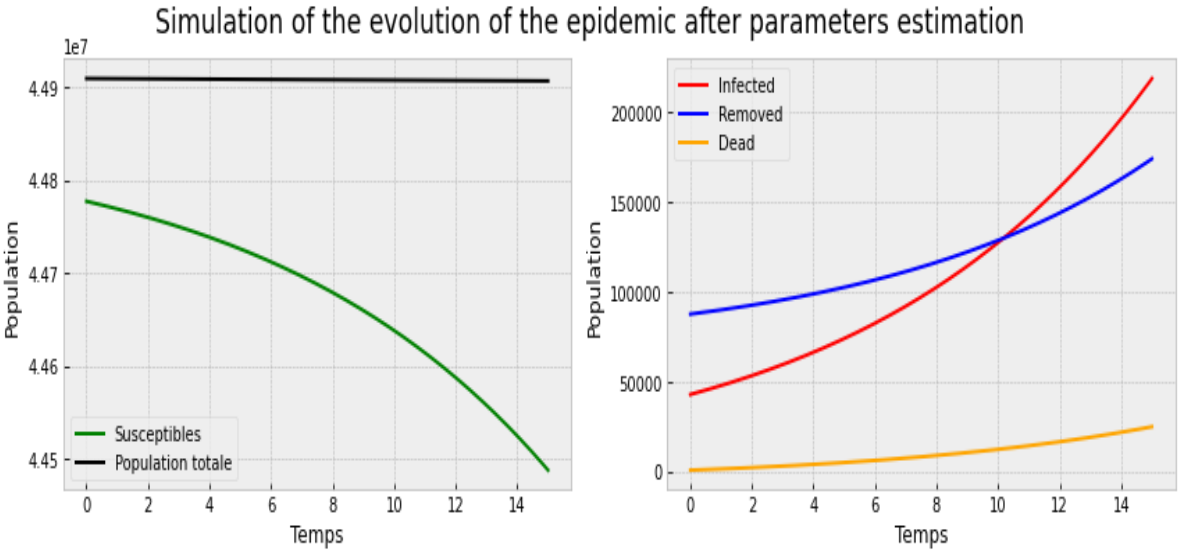

Figure 13 – Prediction of the population's condition on June 29, 2020, in Saudi Arabia.

Discussion

The SIRD COVID-19 model describes the properties and transmission mechanism of this virus. To enhance the study of the SIRD COVID-19 model, we use the MCMC technique to estimate the unknown parameters. The knowledge we have gained about COVID-19 will allow us to better manage and avoid it in the future. Inadequately chosen prior parameters may prevent or delay convergence in the MCMC method. Furthermore, if the previous parameter values were picked wrong, either the samples are oversampled around the boundary or the results do not converge. Because of the importance of the core problem, the above criteria must be chosen with care. A deterministic model, applicable to other viruses, describes the behavior of COVID-19. This approach is useful for dealing with a wide variety of viruses.

Conclusion

In this work, we presented a SIRD model to replicate the illness epidemic that occurred in KSA. We used the Markov chain Monte Carlo technique is used in the process of estimating and fitting the epidemic parameters and reproduction numbers from the transmission model to the real data. The SIRD system of the proposed SIRD system is carried out via the Sobol method. Also, we gave some useful conclusions at the end of this work.

The present COVID -19 pandemic in Saudi Arabia prompted a modification of the traditional SIRD model for the purpose of analyzing the temporal development of the transmission of infection.

- It is shown that the model accounts for the observed distribution of all three types of data for Saudi Arabia (active, recovered, and deceased).

- This model quantifies the importance of confinement and quarantine in preventing the spread of an infectious disease, and it takes into account real-world factors like asymptomatic people, the possibility of transmission through the surface or the air, etc.

- The data from Saudi Arabia demonstrates that there is an inverse association between delay and the rate of quarantine.

- This study demonstrates that the model is just as applicable to describing the epidemic in Saudi Arabia, despite the fact that the rate constants vary considerably. This could be because of variations in things like innate and acquired immunity, lifestyle choices, environment, food, societal preventative measures, etc. This fact requires more consideration in future research.

- in future work we will discuss the memory effect on the proposed SIRD model and Dynamic analysis for the proposed SIRD model should be stated in the future work

References

Bardhan, S., Montgomery, D., Filliben, J., et al., 2019. A general methodology for deriving network propagation models of computer worms, NIST Technical Note 2035.

Batista, F.K., Martn del Rey, N., Quintero-Bonilla, S., 2018. A SEIR model for computer virus spreading based on cellular automata. Advances in Difference Equations, p. 641–650.

Routh–Hurwitz Stability and Quasiperiodic Attractors in a Fractional-Order Model for Awareness Programs: Applications to COVID-19 Pandemic

Hassan, Taher S.; Elabbasy, et al.,2022 Odinaev, Ismoil. Discrete Dynamics in Nature and Society ; Article Dans Anglais | ProQuest Central | ID: covidwho-1807672

Bayette, C., Monticelli, M., 2020. Modélisation d'une épidémie, 1", available at https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-1.html.

Dalal, I. L., Stefan, D., Harwayne-Gidansky, J., 2008. Low discrepancy sequences for Monte Carlo simulations on reconfigurable platforms. International Conference on Application-Specific Systems, Architectures and Processors, Proceedings, Leuven, Belgium.

Deng, Y., Dai, G., Chen, S., 2007. A robust estimator for evaluating internet worm infection rate. 2007 International conference on computational intelligence and security, 12: 15–19, Harbin, China, IEEE, p. 682–686.

Deng, Y., Dai, G., Mu, D., 2008. Comparing robust estimator and least squares estimator in estimating worm infecting rate. J Northwest Polytech Univ, 26, p. 492–496.

Diibendorfer, T., Plattner, B., 2005. Host behaviour based early detection of worm outbreaks in Internet backbones, 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise (WETICE'05), p. 166-171.

Elaiw, A. M., 2012. Global properties of a class of virus infection models with multitarget cells. Nonlin Dynam, 69, p. 423–435.

Elaiw, A. A., 2014. Global properties of a cell mediated immunity in HIV infection model with two classes of target cells and distributed delays. Int J Biomath, 7, p. 119–143.

Elerian, O., Chib, S., Shephard, N., 2001. Likelihood inference for discretely observed non-linear diffusions. Econometrica, 69, p. 959–993.

Eraker, B., 2001. MCMC analysis of diffusion models with application to finance. J Am Acad Child Psychiatr, 19: 177–191.

Gasparini M., 1997. Markov chain Monte Carlo in practice. Technometrics, 39, p. 338–349.

Gilks, W. R., 1996. Introducing Markov chain Monte Carlo. Markov Chain Monte Carlo Pract, 91, p. 1–19.

Golightly, A., Wilkinson, D.J., 2006. Bayesian sequential inference for nonlinear multivariate diffusions. Stat Comput, 16, p. 323–338.

Golightly, A., Wilkinson, D.J., 2008. Bayesian inference for nonlinear multivariate diffusion models observed with error. Comput Stat Data Anal, 52, p. 1674–1693.

Hosseini, S., Abdollahi, Azgomi, M., 2018. The dynamics of an SEIRS-QV malware propagation model in heterogeneous networks. Phys A Stat Mech Appl, 512, p. 803–817.

Jones, C.S., 1998. A simple Bayesian method for the analysis of diffusion processes. SSRN Electron J, 111488.

Kephart J. O., White, S. R., 1993. Measuring and modeling computer virus prevalence. In: IEEE computer society symposium on research in security and privacy, Oakland, CA, USA, 5, p. 24–26.

Kirmani, E., Hood, C., S., 2010. Analysis of a scanning model of worm propagation. J Comput Virol, 6, p. 31–42.

Liang, X., Pei, Y., Lv, Y., 2018. Modeling the state dependent impulse control for computer virus propagation under media coverage. Phys A Stat Mech Appl. 491, p. 516–527.

Mishra, B. K., Pandey, S. K., 2011. Dynamic model of worms with vertical transmission in computer network. Appl Math Computat, 217: 8438–8446.

Ren, J., Yang, X., Zhu, Q., et al., 2012. A novel computer virus model and its dynamics. Nonlin Anal Real World Appl, 13, p. 376–384.

Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics. 41, p. 39–56.

[Sobol, I. M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comp. Exp. 1, p. 407–414.

Sobol, I. M., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. USSR Comp. Math. Math. 7, p. 86–112.

Sobol, I. M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math. Comput. Simul., 55, p. 271–280.

Stramer, O., Bognar, M., Schneider, P., 2010. Bayesian inference for discretely sampled Markov processes with closed-form likelihood expansions. J Financ Econ, 8, p. 450–480.

Talawar, A.S., Aundhakar, U. R., 2016. Parameter estimation of SIR epidemic model using MCMC methods. Global J Pure Appl Math, 2, p. 1299–1306.

Wierman, J. C., Marchette, D. J., 2004. Modeling computer virus prevalence with a susceptible-infected-susceptible model with reintroduction. Computat Stat Data Anal, 45, p. 3–23.

Zhang, Z., Kumari, S., Upadhyay, R. K., 2019. A delayed eepidemic SLBS model for computer virus. Adv Diff Equat, 1, p. 1–24.

Zhou, T., Dai, G., Ye, H., 2006. Research on Internet worm early warning system based on Kalman filter. J Northwest Univ Technol, 24, p. 19–22.

Zou, C. C., Towsley, D., Gong, W., 2006. On the performance of Internet worm scanning strategies. Perform Eval, 63, p. 700–723.

https://coronavirus.jhu.edu/data/new-cases.