Infectious Diseases Modelling with Ordinary Differential Equations

1st Gabriel Ordoñez Vergel
Facultad de ingenierias fisicomecanicas
Universidad Industrial de Santander
Bucaramanga, Colombia.
Gabrielandre2001@hotmail.com

2nd Tatiana Flórez Acosta
Facultad de ingenierias fisicomecanicas
Universidad Industrial de Santander
Bucaramanga, Colombia.
Tatianaflorez179@gmail.com

***Abstract*—This work studies the behavior of an infectious disease in a specific population using compartmental models and their solution through numerical methods in computational tools. In Colombia, this models are already used to study the behavior of COVID-19, however, the mortality of the disease in the country is being overlooked. The objective of this work is to improve the models used by the government to model the virus behavior. For this we have implemented a SIRD model in a population. Experimental data corresponding to the population of Bucaramanga was obtained. Subsequently, the compartmental model corresponding to Bucaramanga was solved using Runge Kutta 4 and adjusted using non-linear least squares and an overlapped window process. Finally, different experiments were carried out to test scenarios of strict quarantine, vaccination, and a situation without preventive measures. Our model satisfactorily resembles reality, and it is possible to use it to study the behavior of many situations of the pandemic that we are facing.**

***Index Terms*—Runge-Kutta 4, COVID-19, SIRD, non-linear least squares, ODE**

I. Introduction

Since December 2019, mankind has faced an unforeseen public health issue that has grown rampant, produced by the virus Sars-covid-2019. Nowadays, the number of deaths attributed to these pandemic amounts to 2.56 million worldwide and the number of infected people increases by thousands daily, an expected behavior from a highly contagious disease.

In these scenarios, it is essential to study the behavior of the disease, in order to be able to take measures to help protect the population effectively. One approach used to achieve this objective in an optimal way is the implementation of compartmental models.

Compartment models in epidemiology simplify the mathematical modeling of infectious diseases [7]. In these models, the population is assigned to different compartments. Individuals who belong to a certain compartment share specific characteristics. Such individuals progress through compartments when their initial characteristics change. Once established, the relationship between these compartments is represented by means of differential equations, which make it possible to study the disease by analyzing the compartmental changes over time.

In April 2020, the Colombian National Institute of Health adjusted the SIR compartmental model [3] to the Colombian population [5]. In this model, the population is categorized in three compartments. Compartment **S** for susceptible population, compartment **I** for infected population, and compartment **R** for the recovered population.

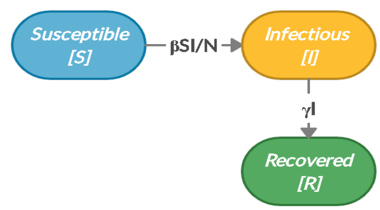

Fig. 1. SIR Model for Ordinary Differential Equations

The relation between said compartments is described by the following differential equations:

$$\frac{dS}{dt} = \frac{-\beta SI}{N} \tag{1}$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \tag{2}$$

$$\frac{dR}{dt} = \gamma I \tag{3}$$

Where β is the infection rate and γ is the recovery rate.This model does not take into account patient mortality; thus, it is necessary to introduce it in a new model, the SIRD model.

The SIRD model [2], a variant of the SIR model, in which the population is distributed in four compartments, instead of three as in the previous SIR model. The main difference between these models is the new group of individuals, compartment D for the deceased population. Which means now the Infected can pass to compartment **D** or compartment **R**.

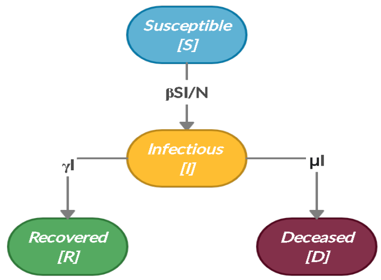

Fig. 2. SIRD Model for Ordinary Differential Equations

The model is expressed from differential equations of the following structure:

$$\frac{dS}{dt} = \frac{\beta IS}{N} \tag{4}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I - \mu I \tag{5}$$

$$\frac{dR}{dt} = \gamma I \tag{6}$$

$$\frac{dD}{dt} = \mu I \tag{7}$$

Where β is the recovery rate, γ is the infection rate and μ is the mortality rate

The SIRD model allows the optimal study of a population under an infectious disease in which the number of deaths represents an extensive amount, as is the case of COVID-19, thus creating a compartment for said population from which its behavior may be studied.

The adjusted model will be used to find a solution to scenarios of high interest to public health entities. These models are difficult to solve analytically due to the multiple variables that must be taken into account, therefore, numerical integration techniques will be used to simplify the work in order to create a feasible numerical tool.

This document consists of the following sections: an introduction to the situation caused by COVID-19 and its impact in Colombia is presented. The motivation for the development of the project is discussed below. Subsequently, the proposal and the algorithms used for the creation of a computationally efficient SIRD model are described and, lastly, the results and conclusions obtained.

II. Motivation

The main motivation of this project is to improve the techniques with which Colombian state entities apply compartmental models to the Colombian population, through the implementation of the SIRD model, which better adjusts to the virus that we face, a deadly COVID- 19.

As of March 3, 2021, in Colombia, the country has a cumulative 59,972 deaths from COVID-19, and 36,490 active cases, according to the data recovered from the Ministry of Health and Social Protection [6]. These numbers are not negligible; therefore, it is crucial to implement the SIRD model instead of the SIR model used by the Colombian government.

The SIRD model adjusted to the population of Bucaramanga, makes it possible to carry out an analysis of its behavior concerning the disease during the last 10 months, contemplating different scenarios that allow a clear understanding of the current panorama and act accordingly when choosing strategies of prevention that favor the citizens of Bucaramanga.

III. Approach

Our proposal is to solve the SIRD compartmental model using the fourth-order Runge-Kutta (RK4) [1] and identify the parameters β (recovery rate), γ (infection rate), and μ (mortality rate), which are correctly adjusted to the population from the city of Bucaramanga.

To obtain the experimental data, the information provided by the Colombian National Health Institute will be used, in this way a theoretical SIRD model will be created. With this model, the required parameters can be found by fitting curves using the nonlinear least squares numerical method [10].

A. *Runge Kutta 4*

Developed at the beginning of the 20th century by the German mathematicians Carl David Tolmé Runge and Martin Wilhelm Kutta, the Runge-Kutta methods are a family of implicit and explicit iterative methods to find the numerical solution of differential equations, mainly used to solve initial value problems.

The best-known method among the Runge-Kutta methods is the so-called fourth-order Runge-Kutta or classical Runge-Kutta method and is widely known for being a high precision method. Based on an initial value problem, defined as:

$$\frac{dy}{dt} = f(t, y), \qquad y(t_0) = y_0 \tag{8}$$

Where y is a function of time t which we want to approximate.

It is known that $\frac{dy}{dt}$ is the variation at which y changes, and is a function of both the time t, and y itself. In the initial time t_0 the value corresponding to y es y_0. The function f and the initial conditions are given. Finally, a step size $h > 0$ is chosen.

Therefore, the fourth order Runge-Kutta method is defined as

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4) \tag{9}$$

$$t_{n+1} = t_n + h \tag{10}$$

For $n = 0, 1, 2, 3, ...$ where:

$$
\begin{aligned}
k_1 &= f(t_n, y_n) \\
k_2 &= f(t_n + \frac{h}{2}, y_n + h\frac{k_1}{2}) \\
k_3 &= f(t_n + \frac{h}{2}, y_n + h\frac{k_2}{2}) \\
k_4 &= f(t_n + h, y_n + hk_3)
\end{aligned} \tag{11}
$$

Here y_{n+1} is the next value, determined by the present value y_n plus the product of the interval size h, multiplied by the approximate slope. The slope is a weighted average of slopes, where k_1 is the slope at the beginning of the interval; k_2 is the slope at the midpoint, using k_1; k_3 is the slope at the midpoint, but now using k_2 and finally k_4 is the slope at the end of the interval, using k_3. When the four slopes are averaged, the slopes at the midpoint are given greater weight, as it follows.

The Runge-Kutta 4 method is a method of order $O(h^4)$, which remains stable in lower resolution scenarios, compared to other methods for solving systems of differential equations, such as Euler's method, making it the designated tool for the development of this project.

B. National Institute of Health Experimental Data

For the current proposal, all the experimental data corresponding to the 4 populations contemplated in the SIRD model, are those corresponding to the inhabitants of the municipality of Bucaramanga.

The susceptible population corresponds to the urban population projections for 2018-2020 recovered from the National Administrative Department of Statistics (DANE) [4] for the given territorial entity.

The infected population corresponds to daily cases according to the onset of symptoms according to the Colombian National Health Institute (INS) for the population of the municipality of Bucaramanga. This project considered the data in the date range from May 15, 2020 to March 10, 2021.

Lastly, the data corresponding to the compartments of the recovered population and the deceased population correspond to synthetic profiles created based on the parameters implemented in the SIR model of COVID-19 transmission and lethality by the Colombian National Institute of Health (INS).

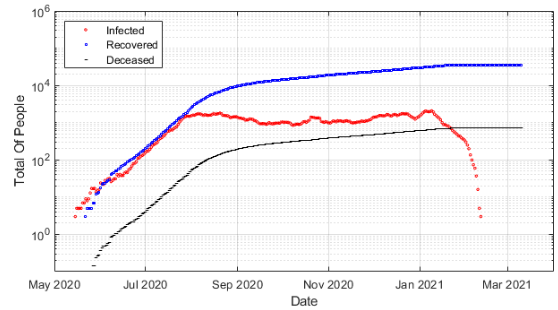

Fig. 3. Profiles of infected population, recovered population and deceased population corresponding to the citizens of the municipality of Bucaramanga in the date range from May 15, 2020 to March 1, 2021

C. Non linear least squares

The nonlinear least squares method is the variant of the least squares approach used in some forms of nonlinear regression, to fit a data set with a nonlinear model with unknown parameters. This is based on the approximating the model by a linear method and later refining the parameters through successive iterations.

To fit our model solved by RK4 it was necessary to use MATLAB's lsqcurvefit function [9] *Lsqcurvefit* is a nonlinear least squares curve fitting problem solver in MATLAB. This solver finds the x parameters that solve the problem

$$\min_x \|F(x,xdata)-ydata\|_2^2 = \min_x \sum_i (F(x,xdata_i)-ydata_i)^2$$

Given the input data $xdata$, and the experimental data $ydata$, where $xdata$ and $ydata$ are matrices or vectors, and $F(x, xdata)$ is a function with matrix or vector values of the same size as $ydata$.

D. Overlapping windows

Due to the extensive nature of the time period used in our model, different dynamics are present throughout it. There are moments in time with a high rate of infection and moments with a rather stable rate of infection, therefore, different rates of infection, mortality and recovery are observed.

This causes the parameters obtained by lsqcurvefit not to fit the model in coherence to reality. A window adjustment is necessary to better describe changes in population dynamics.

A distribution by windows of forty days with an overlap of ten days was carried out. Ten windows were generated in total, nine forty-day windows and a thirty-day window.

In this way, it was possible to obtain a model solved by RK4, and with the parameters obtained in the distribution by windows that correctly adjusted the reality of the population of Bucaramanga.

E. Experiments

- Experiment 1: How would the model behave in the face of a strict quarantine in the window that contains the first Duty free day? (June 19, 2020) [8]
- Experiment 2: How would the model behave if faced with a strict quarantine in the window that contains the month of December and the beginning of January?
- Experiment 3: How would the model behave in the face of a single-dose vaccination process from day 150? 500 vaccines per day.
- Experiment 4: How would the model behave in a situation in which prevention measures are not taken in the city of Bucaramanga?

IV. Results

The results obtained through our proposal were as follows. By means of the overlapping windows and Matlab *lsqcurvefit*, the parameters of Fig 4 were obtained

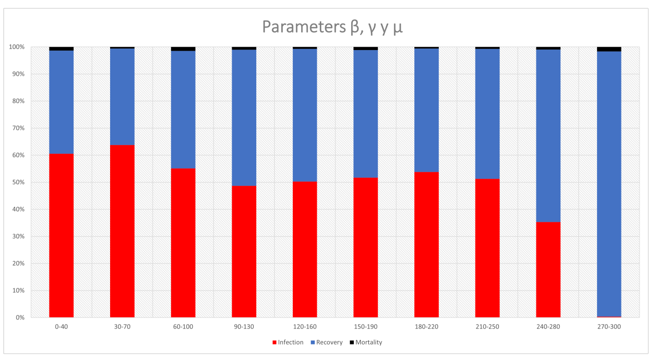

Fig. 4. Parameters obtained for each window of the model solved by RK4

The figure 5 presents the model resolved by RK4 with distribution by overlapping windows of 40 days and parameters obtained by non-linear least squares.

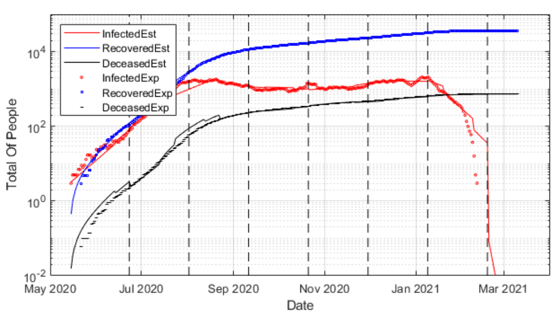

Fig. 5. Experimental data vs Estimated data from 15/05/2020 to 1/03/2021

The model fits satisfactorily to reality. There are slight discrepancies to be observed, such as that in the model recovered and deceased are generated earlier than in the experimental data. The profile corresponding to the recovered population shows an indistinguishable difference from reality as time progresses. On the other hand, the line of the deceased population presents small peaks at the beginning and the approximation improves as it advances. Finally, the line corresponding to the infected population shows slight differences with real data in all sections, this is due to the fact that it is the variable with the greatest noise.

A. Experiment 1

The figure 6 displays the problematic caused by the conglomeration of people in the main shopping centers of the country due to the first of a series of duty free days

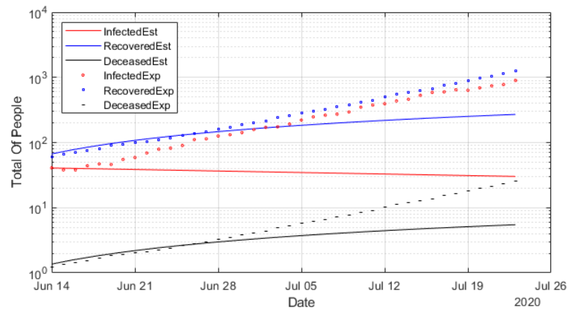

Fig. 6. Duty free day experiment. The model shows a diminution in the expected infected population (solid red line) vs the events (dashed red line)

The model predicts that the infected population would have decreased significantly, by approximately 87%. The model also indicates that mortality would have been 54% lower compared to reality.

B. Experiment 2

The figure 7 displays the problematic caused by the lack of preventive measures for the Christmas and new year's eve celebrations. Plenty of people gathered in their homes unconsciously becoming a medium for the spread of the virus.

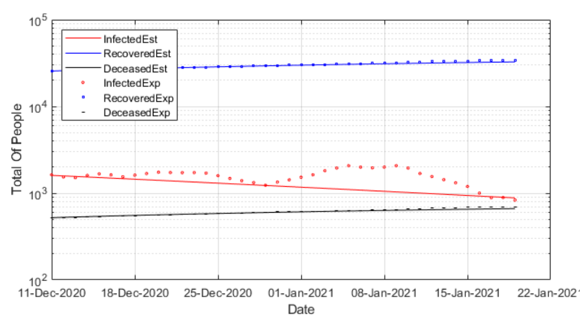

Fig. 7. Holidays experiment. The model shows a diminution in the expected infected population (solid red line) vs the events (dashed red line). Specially at the start of the new year.

The model predicts that those infected would have decreased by approximately 22% and there would have been 1,31% fewer deaths.

C. Experiment 3

The figure 8 displays what would have happened if the government implemented a vaccination plan of 500 doses per day from October 15, 2020. For this experiment it was necessary to make a modification to the equation 4 and 6:

$$\frac{dS}{dt} = \frac{\beta IS}{N} - 500 \tag{12}$$
$$\frac{dR}{dt} = \gamma I + 500 \tag{13}$$

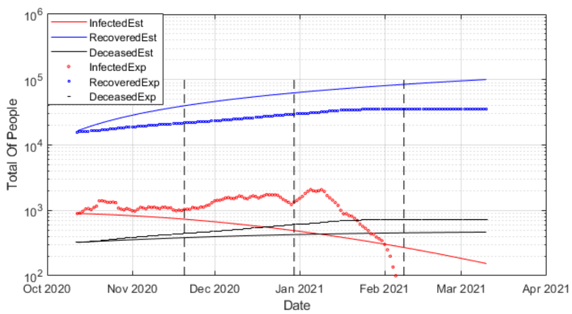

Fig. 8. Vaccines experiment. The model shows an increase in the expected recovery population (solid blue line) vs the events (dashed blue line). It also shows a diminution in the expected infected and deceased population vs the events.

The model predicts that the recovered population would increase by almost half an order of magnitude as can be seen in the graph. Furthermore, there would have been at least 200 fewer deaths a day, a fact to take into account

D. Experiment 4

The figure 9 displays an unthinkable situation: the government does not declare preventive measures, therefore, a quarantine was never carried out.

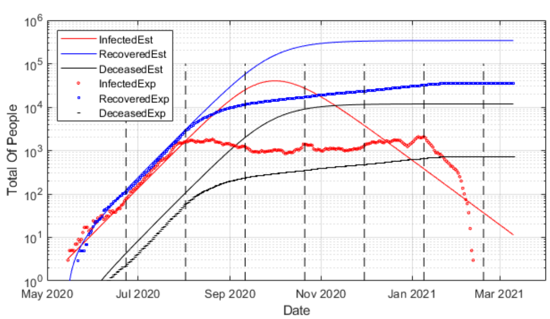

Fig. 9. No quarantine experiment. The model shows an enormous increase in the expected deceased population (solid black line) vs the events (dashed black line) .

The model predicts that the toll of deaths would have increased by a full order of magnitude, which means that instead of having 1,000 deaths per day, it would have reached 10,000 deaths per day.

V. Conclusions

Compartmental models in engineering allow us to represent and study the world. This project proved that, although perfect precision will never be achieved, it is possible to model reality with remarkable precision, obtaining great results. Numerical methods are fast and efficient tools that simplify complex mathematical processes, such as solving the systems of ordinary differential equations that compartmental models use. By implementing both the compartmental models and the numerical integration techniques, the project showed it is possible to simulate an infinity of different scenarios that allow an in-depth study of real world dynamics. Finally, the main advantage the authors of this project want to point out, is that the construction of this model that resembles reality with a low computational cost was possible thanks to the use of numerical methods.

References

[1] Uri M Ascher and Linda R Petzold. "Computer Methods for Ordinary Di erential Equations and Di erential-Algebraic Equations". In: (1997).

[2] Norman T.J. Bailey. *The mathematical theory of infections diseases and its applications*. 2nd. London, 1975, p. 413. ISBN: 0-85264-231-8.

[3] Ross Beckley et al. "GroupProjectSIR.pdf". In: (2013), pp. 1–9.

[4] DANE. $Proyecciones_{d}e_{P}oblacion_{2}018 - 2020$.

[5] Molero Garc et al. "Modelo de transmision de coronavirus COVID-19". In: *Reseña científica* 53.11 (2020), p. 11. ISSN: 1098-6596. arXiv: arXiv:1011.1669v3. URL: https://apps.who.int/iris/bitstream/handle/10665/333390/WHO-2019-nCoV-Sci_Brief-Transmission_modes-2020.3-spa.pdf?%0Ahttp://dx.doi.org/10.1016/j.jsames.2011.03.003%0Ahttps://doi.org/10.1016/j.gr.2017.08.001%0Ahttp://dx.doi.org/10.1016/j.precamres.2014.12..

[6] Instituto Nacional de Salud. *COVID-19 en Colombia*. URL: https://www.ins.gov.co/Noticias/Paginas/coronavirus-filtro.aspx.

[7] A.G Kermack, W.O ; McKendrick. "A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London . Series A , Containing Papers of a Mathematical and physical Character, Vol. 115, No. 772." In: *Society* 115.772 (1927), pp. 700–721.

[8] Luisa Fernanda Ruiz Villamizar. *Lo bueno y lo malo del primer Día sin IVA*. 2020. URL: https://www.vanguardia.com/economia/local/lo-bueno-y-lo-malo-del-primer-dia-sin-iva-NB2517807.

[9] MatLab. *Lsqcurvefit*. URL: https://www.mathworks.com/help/optim/ug/lsqcurvefit.html.

[10] Nonlinear Programming. "Nonlinear programming 4". In: *Mathematics and Computers in Simulation* 25.1 (1983), pp. 86–87. ISSN: 03784754. DOI: 10.1016/0378-4754(83)90042-3.