

Physico-mathematical Modeling for Decision-Making against COVID-19 in Cuba

Modelación físico-matemática para la toma de decisiones frente a la
COVID-19 en Cuba

Dr. C. Héctor Eduardo Sánchez Vargas^{1*} <http://orcid.org/0000-0003-0640-6151>

Dr. Luis Beltrán Ramos Sánchez¹ <http://orcid.org/0000-0002-6403-1396>

Dr. Pablo Ángel Galindo Llanes¹ <http://orcid.org/0000-0003-1738-8761>

MSc. Amyrsa Salgado Rodríguez² <http://orcid.org/0000-0002-6923-8661>

¹Department of Chemical Engineering, Faculty of Applied Sciences *Ignacio Agramonte Loynaz* University of Camagüey, Cuba

²Center for Business and Territorial Studies, Faculty of Economic Sciences, *Ignacio Agramonte Loynaz* University of Camagüey, Cuba

*Corresponding author: hector.sanchez@reduc.edu.cu

ABSTRACT

Objective: To apply physico-mathematical modeling to the dynamics of COVID-19 that enables decision-making associated with mitigation and eradication of the pandemics in Cuba.

Methods: Modeling was applied in order to characterize the predicted peak timing, and the reproductive performance of the pandemic, through MATLAB tools and functions. Peak timing was determined with the application of the SIR model, after adjustments. It was fit using the *GlobalSearch* optimization strategy. Function *ode23tb* was used in the solution, including a Runge-Kutta algorithm combined with another trapezoidal rule algorithm. To determine reproductive performance, an exponential model was fit using *Curve Fitting*.

Main results: The parameters of the SIR model were identified, and the peak forecast was completed rapidly and accurately, two weeks before it occurred. The susceptible, accumulated infected and recovered patients were predicted. The calculated basic reproduction number (R_0) helped conclude that to eradicate the pandemic by vaccination, the immunized population should be over 72 %. The effective reproduction number (R_{ef}) helped evaluate the efficacy of mitigation measures. Remarks are made concerning the proper conduct to follow for eradication.

Conclusions: The SIR model proved its capacity to predict the peak timing of the pandemic. The R_0 of SARS-CoV-2 corroborated the high transmissibility of the virus. The mitigation measures have been effective, and should be kept until the pandemic is eradicated, even with $R_{ef} < 1$, until 72 % of the population is immunized.

Key words: COVID-19; SARS-CoV-2; decision-making; mathematical modeling; reproduction number.

RESUMEN

Objetivo: Aplicar la modelación físico-matemática a la dinámica de la COVID-19 para la toma de decisiones asociadas a la mitigación y erradicación de la pandemia en Cuba.

Métodos: La modelación se aplicó para caracterizar el pronóstico del pico y el comportamiento reproductivo de la pandemia, mediante herramientas y funciones de MATLAB. El pico se determinó aplicando el modelo SIR, luego de adecuaciones. Este se ajustó con la estrategia de optimización *GlobalSearch*. Para su solución se empleó la función *ode23tb* que usa un algoritmo combinado de Runge-Kutta con otro de regla trapezoidal. Para el comportamiento reproductivo se realizó el ajuste de un modelo exponencial empleando la herramienta *Curve Fitting*.

Principales resultados: Se identificaron los parámetros del modelo SIR y se logró el pronóstico del pico con dos semanas de anticipación y una precisión satisfactoria. Se pronosticaron los susceptibles, infectados acumulados y recuperados. El número de reproducción básico (R_0) calculado permitió determinar que, para erradicar la pandemia por vacunación, la población inmunizada debe ser superior al 72 %. El número de reproducción efectivo (R_{ef}) permitió evaluar la eficacia de las medidas de mitigación. Se

reflexionó sobre la conducta a seguir para su erradicación.

Conclusiones: El modelo SIR demostró capacidad para predecir el pico de la pandemia. **El R_0 del SARS-CoV-2 permitió corroborar su elevada transmisibilidad.** Las medidas de mitigación han sido efectivas y deben mantenerse hasta erradicar la pandemia, **incluso para $R_{ef} < 1$** , mientras no se inmunice el 72 % de la población, para lograr la erradicación irreversible.

Palabras clave: COVID-19; SARS-CoV-2; toma de decisiones; modelación matemática; número de reproducción.

Received: 05/06/2020

Accepted: 27/07/2020

INTRODUCTION

Epidemics are community health events that take place when the number of affected individuals is higher than expected for a period of time. When epidemics spreads over vast geographical areas of continents or even the whole globe, it is called a pandemic.

COVID-19 originated in the city of Wuhan, capital of the Chinese province of Hubei, in December 2019, and was cataloged as a pandemic by the World Health Organization (WHO), on March 11th, 2020 (WHO,2020b) It has become the largest pandemic event after the Spanish Flu. Causing devastating effects to human health globally —183 countries with cases of COVID-19, 2 982 688 confirmed cases, and 210 193 deaths until April 28— (MINSAD, 2020), the social and economic damages produced are immeasurable.

The rapid propagation of COVID-19 throughout the world, and its high morbidity and lethality have led to the collapse of health systems in an important number of countries. Hence, governments and local authorities of the nations affected have been confronted with the challenge of the pandemic, since the effectiveness or absence of it in the policies implemented is expressed in the number of contagions, deceased victims, and recovered patients. Additionally, the pertinence or not of these policies reflects in the

economy.

The access to accurate and timely information that helps predict the possible behavior of the disease in any instances of a nation may contribute to both decision-making that ensures the effectiveness of such policies and the efficacy of these actions. Hence, the utilization of physico-mathematical models can help predict the behavior of COVID-19. These models might help in short-term and long-term planning to control the dynamic of the disease, and evaluate public health interventions (Brauer, Castillo, and Feng, 2019; Roda, Varughese, Han & Li, 2020).

Some key elements should be considered for decision-making in these scenarios. One of them is related to the so-called peak of the pandemic; that is, the moment in which the largest number of infected people will occur. Being able to predict the highest number of infected people, and the date in which it will occur, are very important to ensure the availability of material and human resources, and achieve more effective actions against the disease in the most critical moment. Another key element that must be considered is the reproductive behavior of the pandemic, characterized by the effective reproduction number (R_t). The determination of both elements will surely contribute to effective measures in terms of resource management during its evolution and the control of the disease for eradication.

To predict the behavior of the pandemic of COVID-19 in different parts of the world, several predictive models have been used. The predictions have included, among others, the determination of the date and magnitude of the peak of the pandemic, as well as the estimation of R_t . The models used have shown great variability (Roda *et al.*, 2020), so they may have been influenced by behavioral differences inherent to environmental, genotypic, and cultural differences, shaped also by the measures taken by local authorities in the different developmental stages of the pandemic.

Cuba has also been affected by the pandemic, and has bet on the utilization of science and technological innovation, with the use of scientific results and physico-mathematical models, under the coordination of the National Temporary Group to Fight COVID-19. One of the models is known as SIR (susceptible, infected, recovered patients), published by Kermack and McKendrick (1927), which has been extensively used in different variants, and acknowledged for its simplicity, and predictive capacity. It

requires fit of the real data input in relation to the pandemic, and permits simulating the behavior of the number of infected, susceptible, and recovered patients through time.

Although the R_t from the pandemic can be determined with the utilization of mathematic modeling methods, including SIR, other statistical and mathematical methods can be used for determination.

This paper shows the results from the application of SIR fit to the Cuban case, estimations and comparative analysis of the reproductive factor in various countries, and discussion of these results in relation to decision-making. Accordingly, this paper explains the way the physical model and mathematical deduction were made. It fits the model that implies the determination of characteristic parameters, and it simulates the dynamic of the pandemic in Cuba, with which the peak is identified. Besides, the data concerning the development of the pandemic in Cuba and several countries, as well as the basic reproduction number (R_0), and the R_t of the pandemic were determined. Based on the results from SIR application and R_t determination, the efficacy of pandemic mitigation actions in Cuba were evaluated. Then, the recommended conduct to follow to eradicate the disease in two scenarios one without a vaccine, and another in which an effective vaccine could be used, were discussed.

DEVELOPMENT

Epidemic modeling and other decision-making processes

Authors such as Arrendondo (2013), De Kohan (2015), Miranda (2015) and Stoner, Freeman, and Gilbert (1996) have contributed to decision-making processes. In addressing a problem, the need of information for analysis, the design of alternatives for solution, and the selection of a more effective alternative, are among the most frequently evaluated elements by these authors.

Martín, Loredó, and Álvarez (2016) also said that when the problem addressed has no similar preceding situations, decisions are made with increased uncertainty and difficulty. Robbins and Coutler (2010) added that the complexity of situations in which decisions are made demands five essential elements: information, knowledge, experience, analysis, and judgment.

Sharing pertinent and timely information has a large-scale influence on the effect of analyses to make decisions; therefore, it is not always sufficient. When similar situations are unknown, and the problem is characterized by a high complexity, new tools must be found to predict the behavior of the main variables. Mathematical models emerged as tools that help make decisions.

A mathematical model is an idealized representation expressed in terms of symbols, mathematical expressions, and related equation systems that described the essence of the problem, thus enabling n quantifiable decisions which are interrelated, and are known as decision variables (Hillier and Lieberman, 1997). Linares, Ramos, Sánchez, Sarabria, and Vitoriano (2021) also determined a simplified way in which these models can express a complex reality. They also added that they can balance the need of gathering all the details of reality with proper feasible techniques for solutions. The analysis and detection of relations among data, the establishment of assumptions and approximations in the representation of problems, and the development or utilization of specific algorithms to provide solutions, are elements that characterize these powerful tools, according to these authors.

Deduction is a mathematical model that requires abstraction from reality; that is, a representation of the system studied, whose rules and symbols can be known. This representation is called physical model, and its conception is necessary to explain phenomena that take place in association with certain processes.

In the particular case of contagious diseases, mathematical modeling has been a powerful tool to understand and predict the dynamic of transmission. The most commonly models used to predict epidemics are phenomenological or mechanistic (Chowell, 2017; Ma, 2020). They are particularly appropriate in situations where it is hard to formulate a mechanistic approach, due to the existence of multiple transmission ways, interactions of spatial influences, and other aspects related to uncertainty (Brauer *et al.*, 2019).

Mechanistic models include physical laws or mechanisms involved in the dynamic of the problem studied. Among them, the widely used compartment models are based on systems of ordinary differential equations that focus on the dynamic evolution of a population with different epidemic states (Chowell, 2017). Compartment models are

powerful mathematic tools used to characterize complicated epidemic behaviors. The pioneers of these models were Kermack and McKendrick (1927), with the creation of the SIR model. Since, the SIR model has been used as the base of many other behavioral models (SIRS, SEIR, SEIRS, etc.), which introduce new compartments in the classic SIR model (Wang *et al.*, 2018).

Despite its simplicity, the SIR model is generally capable of describing the behavior of epidemics, with prediction curves that fit well to the real data (Galindo, Rodríguez, and Cervantes, 2014). It is appropriate when the individuals recovered acquire immunity for life (Prakash, Setia, and Alapatt, 2017), and when the disease can be passed on to an individual relatively quick (Khan and Zaman, 2018), as is the case of viral infections.

The SIR model has been conceived with three compartments: individuals susceptible of infection (S), infected individuals (I), and recovered individuals (R). The number of susceptible individuals will depend on the immunological capacity of the population to fight the disease. In the case on novel diseases, such as COVID-19, the entire population of a country is considered susceptible. The increased number of infected cases (I) will depend on the coexistence of infected and susceptible individuals, as well as their numbers, and transmission capacity of the infectious agent, expressed in R_0 . The recovered patients will reach this condition after overcoming the disease, and will acquire some level of immunity that could be permanent or temporary, depending on the characteristics of the disease.

The mathematical model included in the SIR model is obtained from the application of mass balances between all the population groups in the physical model. The general equation of a mass balance applied to a given population group i (S, I, R), is represented in the following equation.

$$\left[\begin{array}{c} \text{Accum. speed} \\ \text{of population} \\ i \end{array} \right] = \left[\begin{array}{c} \text{Input speed} \\ \text{of } i \\ \text{through boundary} \end{array} \right] - \left[\begin{array}{c} \text{Output speed} \\ \text{of } i \\ \text{through boundary} \end{array} \right] + \left[\begin{array}{c} \text{Increase} \\ \text{speed} \\ \text{of} \\ i \end{array} \right] - \left[\begin{array}{c} \text{Decrease} \\ \text{speed} \\ \text{of} \\ i \end{array} \right] \quad (1)$$

If the mass balance is made for every population group represented by the physical model, the system is integrated by three ordinary differential equations. The solution to this equation system requires the utilization of numerical methods, which permits scholars to determine how the number of individuals in the population group i changes

through time. The number of individuals in the susceptible population will be reduced throughout the epidemic, and the number of recovered patients will increase. Regarding the infected people, there will be an increase, and then, when the number of susceptible patients is significantly reduced, there will be a decrease; that will be the moment when the so called peak will be surpassed.

Application of SIR in Cuba

Physical model

To contrast the number of deceased from recovered patients, and to describe the behavior of the two groups separately, a fourth compartment was added to the SIR model, the deceased (D), which has been acknowledged by some authors (Brauer *et al.*, 2019; Getz, Salter, Muellerklein, Yoon, and Tallam 2018). Fig. 1 shows the physical model with the mentioned fit.

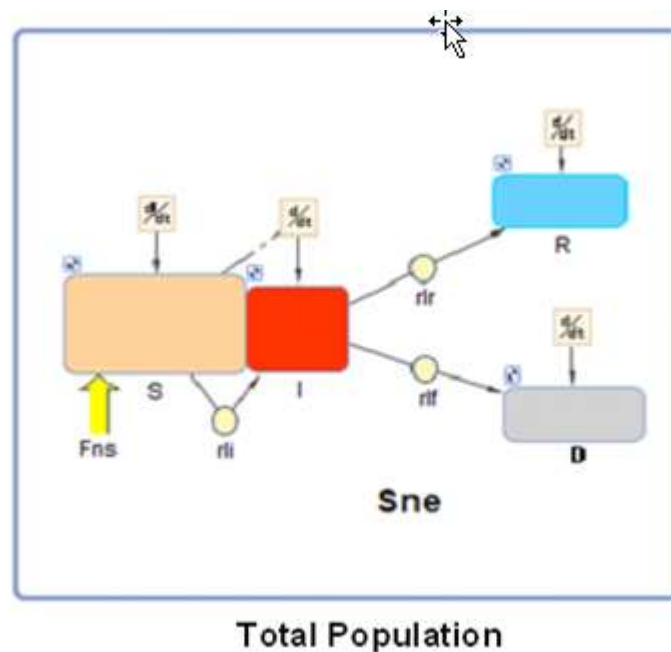


Fig.1. Structure of the modified SIR model

The chart, which includes all the compartments, represents the population of the entire country. In the particular case of Cuba, the total population considered $N = 11\,193\,470$ (ONEI, 2019). The four compartments have distinctive colors. The ones susceptible to

be infected or exposed to the pandemic (S) are in the orange compartment; the infected (I) are in red; the recovered, who acquire immunity, (R) in blue; and the deceased (D) in gray. In this model, the exposed susceptible population was separated from the non-exposed susceptible population (Sne), which surrounds the compartments.

The sum of individuals from each group makes the population of the entire country (N):

$$\mathbf{N = Sne + S + I + R + D} \quad (2)$$

The transit from one compartment to another was represented in three speeds: infection speed (r_{ii}), that is the transit of exposed susceptible individuals (S) into the infection (I) zone; recovery speed of infected patients (r_{ir}); mortality speed (r_{ir}). Another modification done by the authors of this paper is associated to the introduction of the Fns parameter, symbolized by a yellow arrow (Fig. 1), which represents the transit or net mean flow of people between the exposed susceptible (S) compartments into the non-exposed susceptible (Sne) people.

Depending on the Cuban conditions at the onset of the pandemic, and in order to establish mathematical relations in the model, the following considerations were set:

- a. During the pandemic, just one infection can cause an infectious process.
- b. The outcome of the disease is death or complete immunity.
- c. All the population has the same likelihood of being exposed and infected. There is a homogeneous mix between exposed infected and susceptible.
- d. The infection rate is proportional to the number of sick people.
- e. The speed of new contagions is proportional to the number of infected (I) and exposed susceptible (S) in contact.
- f. The total population is constant, and the system is closed; that is, it does not consider births or migrations.
- g. The number of infected people only grows through r_{ii} , not by incoming flow.

The consideration (b) rationale was based on the fact that a recovered person acquires immunity that protects from further infections; that is, this person must not be infected

again. Moreover, consideration (c) is applicable to this case, since the entire Cuban population is susceptible to contagion, based on the fact that this is a novel disease, and consequently, individuals do not have an immunological response. The rate of contagions referred to (d) is the proportionality factor that determines the speed of infection. This infection depends on the existence of infected and susceptible people. The mathematical value of the proportionality factor will depend on the effectiveness of the interaction between the sick and susceptible that can generate new infected people. Considerations (f) and (g) stem from the fact that the Cuban government limited the flow of people coming and leaving the country strictly, and because of the rapid advance of the pandemic, the total population change by births and deaths was not included.

Mathematical model

The mathematical model derived from the physical model. Accordingly, a mathematical model of the equation (1) was developed. All the previous considerations were taken into account.

The equation (1) responds to the mathematical expression of the mass conservation law, and was applied to all the population groups represented in the compartments of the physical model. The term on the left represents the accumulation speed of a population in a given compartment, so i is mathematically defined as $\left(\frac{di}{dt}\right)$ for a population group. The input and output boundary speeds refer to the flow of people that enter and exit the compartment without changing their status (S, I, R, D). The decrease and increase speeds refer to speeds of transformation (status change).

The closed system means the absence of flows between compartments and the exterior (outside the total population). Following the mass compartment balances, the flows between the non-exposed susceptible (Sne), and compartments I, R, and D are also zero, so in these three cases

$$\left[\begin{array}{c} \text{Input speed} \\ \text{of } i \\ \text{through boundary} \end{array} \right] = \left[\begin{array}{c} \text{Output speed} \\ \text{of } i \\ \text{through boundary} \end{array} \right] = 0 \quad (3)$$

In the S compartment, the mean net flow (Fns) is defined as:

$$Fns = \left[\begin{array}{c} \text{Input speed} \\ \text{of } i \\ \text{through boundary} \end{array} \right] - \left[\begin{array}{c} \text{Output speed} \\ \text{of } i \\ \text{through boundary} \end{array} \right] \quad (4)$$

Fns may take positive or negative values. The physical meaning of a negative value indicates, favorably, the existence of more people exiting the exposed susceptible compartment. A positive value indicates the opposite.

Considering (f), the susceptible speed, and the recovery and deceased speeds were considered zero.

Upon mass balance performed to every population group, a mathematical model made of these differential equations was obtained.

Exposed susceptible people (S)

$$\frac{dS}{dt} = Fns - r_{ds} \quad (5)$$

Infected people group (I)

$$\frac{dI}{dt} = r_{Ii} - r_{di} \quad (6)$$

Recovered people group (R)

$$\frac{dR}{dt} = r_{Ir} \quad (7)$$

Deceased people group (D)

$$\frac{dF}{dt} = r_{If} \quad (8)$$

The infection occurs from susceptible individuals; hence, the increase speed of infected people equals the decrease speed of susceptible people.

$$r_{Ii} = r_{ds} \quad (9)$$

Since the status changes only due to recovery or death,

$$r_{di} = r_{Ir} + r_{If} \quad (10)$$

Based on the conception of the model, the interaction between infected and susceptible generates new infected through a mechanism similar to a second order elementary chemical reaction dependent on mass action. This law establishes that the speed of a chemical reaction is proportional to the concentrations of reactants (Leib and Pereira, 2008). That way, the generation of infected, mathematically represented by the speed of infection r_{ii} , depends on a proportionality factor β as described in the consideration (d), and represented in the equation (11):

$$r_{ii} = \beta SI \quad (11)$$

Similarly, the recovery and death speeds are mathematically represented with the following equations:

$$r_{Ir} = \gamma I \quad (12)$$

$$r_{If} = \mu I \quad (13)$$

Combining the differential equations (from 5 to 8), with the algebraic equations (from 9 to 13), the mathematical model described in the behavior of the variables of the physical model corresponding to the population groups, is obtained.

$$\frac{dS}{dt} = -\beta SI \quad (14)$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I \quad (15)$$

$$\frac{dR}{dt} = \gamma I \quad (16)$$

$$\frac{dF}{dt} = \mu I \quad (17)$$

If terms recovery and decrease are removed from equation (15), a differential equation can be defined to determine the accumulated infected (I_{ac}):

$$\frac{dI_{ac}}{dt} = \beta SI \quad (18)$$

Model fit

The identification of parameters from the model was done by minimizing the error between predictions and true data (Bartholomew, 2008). The objective function used was the average percent variation coefficient from the residues (equation 19).

$$\text{Error} = \frac{\sum_{i=1}^m \sqrt{\frac{\sum_{l=1}^n (y_{Ci} - y_{Ei})^2}{n}}}{m} \times 100 \quad (19)$$

Where: m- the number of response variables fit, n- number of experimental points, y_{Ci} - values of the response variable calculated with the model in the experimental point “i”, y_{Ei} -true values of the response variable in the experimental point “i”.

The response variables defined were accumulated infected (Iac), infected (I), deceased (D), and recovered (R). The y_{Ci} were determined with the solution to the system of ordinary differential equations, which are the mathematical model of the system, using function *ode23tb* of MATLAB, with a Runge-Kutta algorithm combined with trapezoidal rule algorithm. The y_{Ei} were taken from the daily WHO reports (OMS, 2020a) shown in table 1.

For fitting, the MATLAB GlobalSearch overall strategy was used with a *fmincon* function. This function used an interior-point predetermined algorithm, based on the gradient method.

The traditional parameters to be identified in the SIR model (β , γ , μ) were established as limits of the 0 and 1 intervals. Besides, other two parameters to be fit were defined: the initial value of exposed susceptible (S_0) and el Fns. The interval limits of S_0 , $0 < S_0 < N$, being N the entire population of Cuba, and for $-\frac{N}{10} < \text{Fns} < \frac{N}{10}$, considering that Fns is a daily net flow of people, and a fraction of the total population.

Table1. Daily summary of public health reports used in modeling

Date	Time (days)	New cases	Accumulated infected (Iac)	Infected (I)	Deceased (D)	Recovered (R)
3-11-2020	0	3	3	3	0	0
3-12-2020	1	1	4	4	0	0
3-13-2020	2	0	4	4	0	0
3-14-2020	3	0	4	4	0	0
3-15-2020	4	0	4	4	0	0
3-16-2020	5	1	5	5	0	0
3-17-2020	6	2	7	6	1	0
3-18-2020	7	4	11	10	1	0

Date	Time (days)	New cases	Accumulated infected (Iac)	Infected (I)	Deceased (D)	Recovered (R)
3-19-2020	8	5	16	15	1	0
3-20-2020	9	5	21	20	1	0
3-21-2020	10	14	35	33	1	1
3-22-2020	11	5	40	38	1	1
3-23-2020	12	8	48	46	1	1
3-24-2020	13	9	57	55	1	1
3-25-2020	14	10	67	63	2	2
3-26-2020	15	13	80	73	2	5
3-27-2020	16	39	119	111	3	5
3-28-2020	17	20	139	131	3	5
3-29-2020	18	31	170	161	4	5
3-30-2020	19	16	186	175	6	5
3-31-2020	20	26	212	193	6	13
4-1-2020	21	21	233	213	6	14
4-2-2020	22	36	269	247	6	16
4-3-2020	23	19	288	266	6	16
4-4-2020	24	32	320	296	8	16
4-5-2020	25	30	350	322	9	19
4-6-2020	26	46	396	356	11	29
4-7-2020	27	61	457	416	12	29
4-8-2020	28	58	515	470	15	30
4-9-2020	29	49	564	496	15	53
4-10-2020	30	56	620	525	16	79
4-11-2020	31	49	669	557	18	94
4-12-2020	32	57	726	582	21	123
4-13-2020	33	40	766	611	21	134
4-14-2020	34	48	814	637	24	153
4-15-2020	35	48	862	662	27	173
4-16-2020	36	61	923	698	31	194
4-17-2020	37	63	986	725	32	229
4-18-2020	38	49	1035	744	34	257
4-19-2020	39	52	1087	764	36	287
4-20-2020	40	50	1137	788	38	311
4-21-2020	41	52	1189	806	40	343
4-22-2020	42	46	1235	825	43	367
4-23-2020	43	50	1285	818	49	418
4-24-2020	44	52	1337	847	51	439
4-25-2020	45	32	1369	812	54	503

Date	Time (days)	New cases	Accumulated infected (Iac)	Infected (I)	Deceased (D)	Recovered (R)
4-26-2020	46	20	1389	806	56	527
4-27-2020	47	48	1437	802	58	577
4-28-2020	48	30	1467	790	58	619

Source: WHO (2020a)

It is known that the progress of the pandemic remarkably depends on the policies and measures taken by government, and individual and collective behavior of citizens. Accordingly, the values of the parameters in the model change, being the mathematical expression that marks the phenomenological behavior. Therefore, the above grounds compel scholars to perform periodic model fitting. The quality of fitting was acceptable, marked by a steady fit error below 4%. Fig. 2 shows the fit of the four response variables, and evidences the quality of the fit done on April 8th, less than a month after the onset of the pandemic in Cuba, with 2.406% error. The black straight line represents the value of the variable calculated by the model. The blue circles represent the real values of the variables.

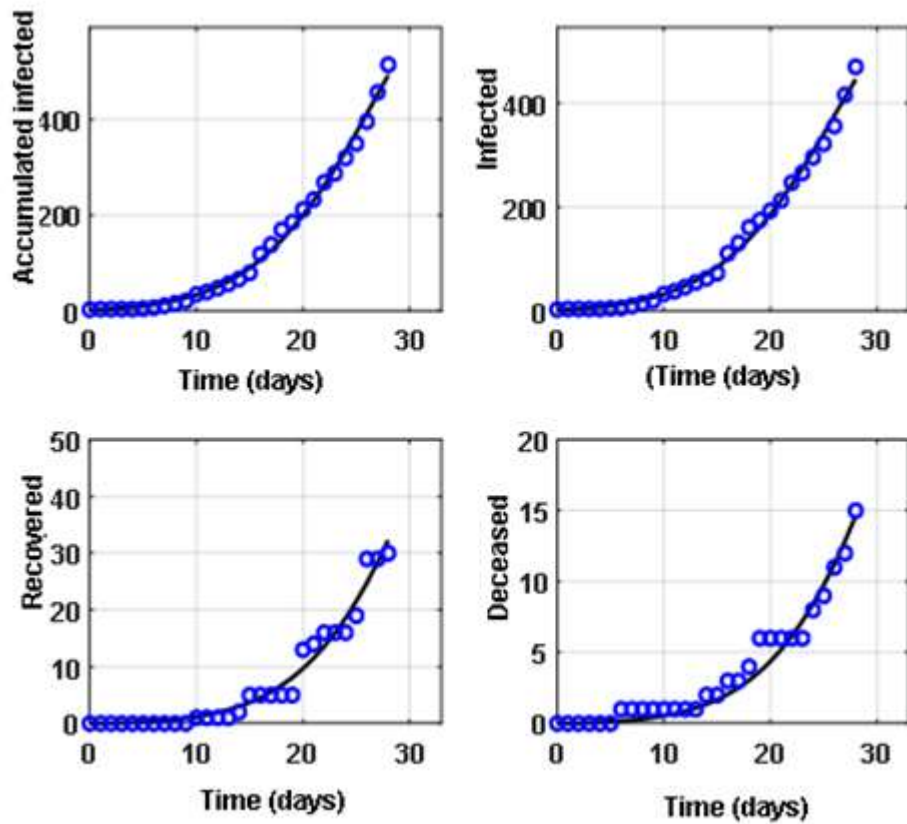


Fig.2. SIR fit to pandemic data in Cuba

Table 2 shows the values of parameters β , γ , μ , S_0 and F_{ns} , resulting from model fitting on April 8th.

Table2. Estimated parameters of the kinetic model

Parameter	Value estimated	Units
β	$5.78 \cdot 10^{-05}$	(susceptible individuals) ⁻¹ (day) ⁻¹
γ	0.00912364	day ⁻¹
μ	0.0041302	day ⁻¹
S_0	4 852	susceptible individuals
F_{ns}	-99.993	(susceptible individuals) (day) ⁻¹

Identification of the peak of the pandemic in Cuba

The peak of the pandemic in Cuba was identified by simulation after model fitting. The days following April 8th, and until April 24th, when the peak was produced, the model was capable of predicting its occurrence date between April 21st and 24th, see Fig. 3), and the number of infected (between 757 and 982), with a mean value of 830 people infected. The simulations were run by solving the mathematical model of the system, using function MATLAB's *ode23tb*.

The parameters obtained following the daily fitting between April 9th and 24th are shown in table 3. Fig. 3 shows a family of curves that represent the predictions of people infected by the pandemic. The legend indicates the adjustment and simulation days in every curve.

Table 3 Results of model parameter fitting

Date	S_0	β	γ	δ	Fns
4-9-2020	17 559	$1.61 \cdot 10^{-05}$	0.01017937	0.0040761	-385.9
4-10-2020	17 558	$1.60 \cdot 10^{-05}$	0.01193508	0.0039905	-371.9
4-11-2020	9 208	$2.98 \cdot 10^{-05}$	0.01401356	0.0040041	-177.08
4-12-2020	9 038	$2.98 \cdot 10^{-05}$	0.01688339	0.0041330	-158.5
4-13-2020	10 571	$2.70 \cdot 10^{-05}$	0.01672686	0.0037764	-217.7
4-14-2020	5 424	$5.04 \cdot 10^{-05}$	0.01887329	0.0038903	-93.9
4-15-2020	5497	$5.00 \cdot 10^{-05}$	0.02000690	0.0038691	-95.2
4-16-2020	6 030	$4.58 \cdot 10^{-05}$	0.02082728	0.0038904	-106.1
4-17-2020	12 851	$2.19 \cdot 10^{-05}$	0.02209713	0.0039048	-246.4
4-18-2020	9 083	$3.06 \cdot 10^{-05}$	0.02360913	0.0039516	-164.3
4-19-2020	9 085	$3.05 \cdot 10^{-05}$	0.02456923	0.0038901	-161.3
4-20-2020	10 947	$2.56 \cdot 10^{-05}$	0.02514310	0.0037910	-200.5
4-21-2020	12 852	$2.17 \cdot 10^{-05}$	0.02605555	0.0037928	-234.9
4-22-2020	10 564	$2.64 \cdot 10^{-05}$	0.02716303	0.0037743	-187.3
4-23-2020	10 558	$2.63 \cdot 10^{-05}$	0.02814454	0.0037989	-184.8
4-24-2020	10 555	$2.62 \cdot 10^{-05}$	0.02904362	0.0038246	-181.6

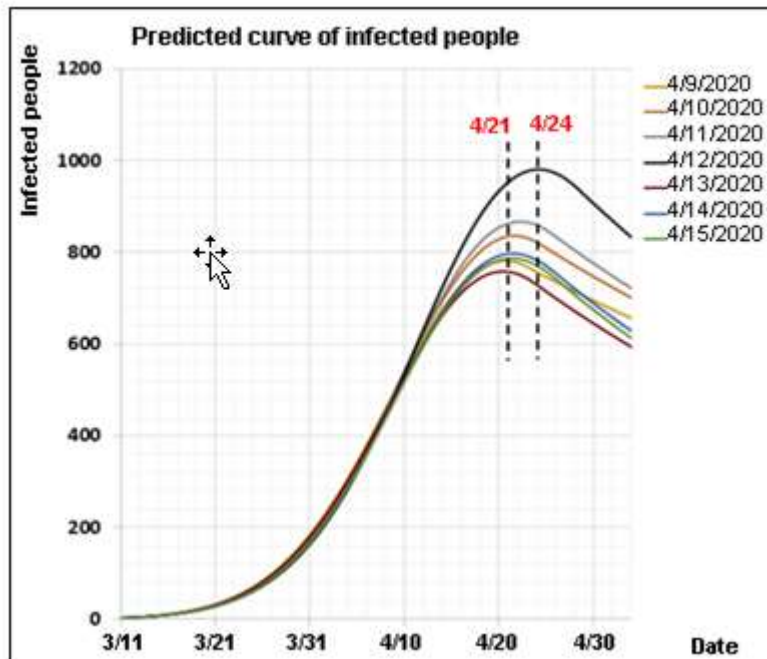


Fig.3. Family of curves that show the prediction of the peak of the pandemic

The model also was used to simulate the behavior of variables defined as response variables (I_{ac} , I , R , D). Fig. 4 shows an example of these predictions, from April 11th, 13 days in anticipation of the occurrence of the peak of the pandemic. The straight lines represent the prediction of the model, whereas the circles show the true values.

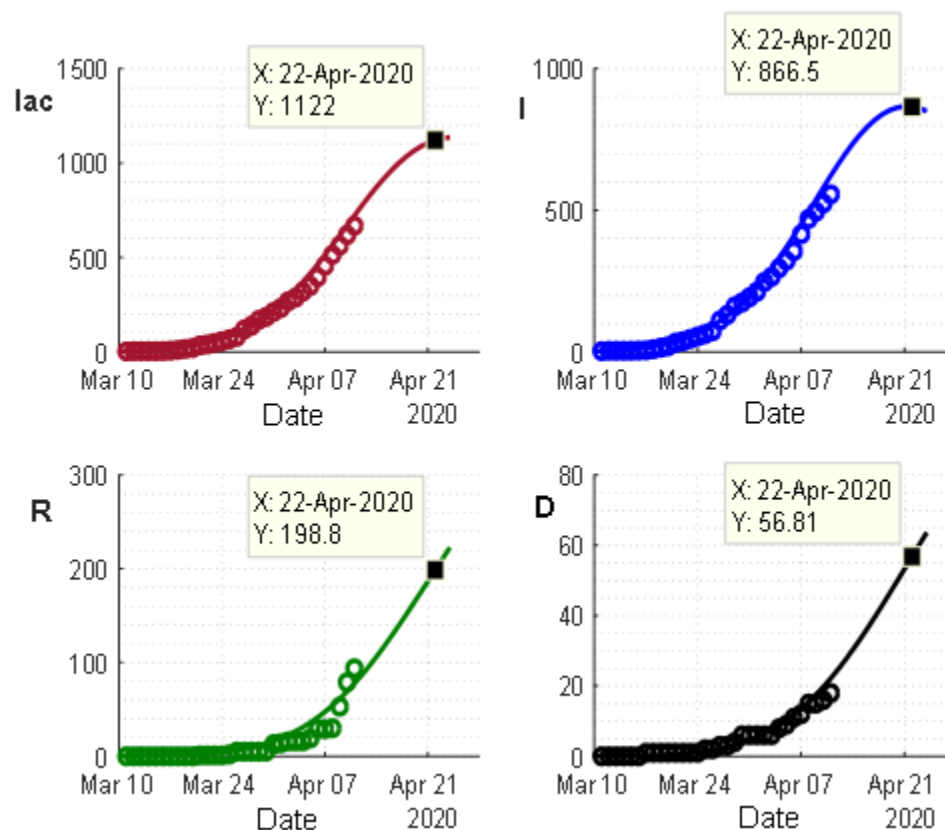


Fig.4. Prediction curves of response variables of the model

The graph labels show the values of variables predicted by the model on April 22nd, predicted as a peak since April 11th. Table 1 shows the true data, April 24th appears in red, when the peak of the pandemic was reached.

Determination of the reproduction factor applied to the pandemic in Cuba

The reproduction number of a pandemic R_t , also named factor or effective reproduction number, is defined as the mean number of secondary infections caused by a case in a given moment (Cauchemez, Hoze, Cousien, Nikolay & Ten Bosch, 2019). An epidemic can affect a substantial part of the population, only if $R_t > 1$ (Cauchemez *et al.*, 2019; Leung, Wu, Liu & Leung, 2020).

Different from R_t , the basic reproduction number R_0 , is the number of secondary infections produced by a primary case in a completely susceptible population, without intervention (Venkatramanan, Lewis, Chen, Higdon, Vullikanti & Marathe, 2018; Wang, Wang, Ye & Liu 2020). The term R_0 is specific for an infection (Kwok *et al.*, 2019). A

pathogenic agent with $R_0 > 1$ is expected to spread throughout an entire population; however, if $R_0 < 1$, the eradication of the pandemic could be expected (Galindo *et al.*, 2014). Hence, R_0 has been used in the prediction, and to compare the expected performance of strategies for epidemic control, such as vaccination (Lavine, Poss, and Grenfell, 2008).

In turn, the number of control reproduction R_c , is analogous to R_0 , but conditioned by the control measures in a given place, after an epidemic has been identified (Brauer *et al.*, 2019). On the onset, an epidemic that takes place in a completely susceptible population without intervention has equal R_t and R_0 values. If, on the contrary, some control measures are implemented since the beginning, the value of R_t equals R_c . After that moment, R_t decreases as the susceptible population is reduced (Chowell, 2017), according to the equation:

$$R_t = R_0 \frac{S}{N} \quad (20)$$

Where:

S- Number of exposed susceptible people

N- Totally exposed population to the epidemic

The equation (20) can be used alternately to determine R_t , from R_0 or R_c . It should be noted that if R_0 is used for determination, the R_t value calculated will be independent from the control measures; that is, when $R_t < 1$ is obtained, the epidemic will be eradicated, even in the absence of isolation control measures and movement restrictions. On the contrary, if R_t is determined using R_c , its value will depend on the control measures; then the epidemic could only be eradicated by keeping these measures, and maintaining $R_t < 1$.

The equation (2) conceives the values of R_0 o R_c , as constants; however, external values like interventions, climate, or social factors, might change the value of R_t , without having significant changes in the susceptible population. In that sense, several statistical-mathematic methods have been developed to estimate this trend (Cauchemez *et al.*, 2019; Cori, Ferguson, Fraser & Cauchemez, 2013; Wallinga and Teunis, 2004). To contrast them, the R_t conditioned to different factors is given the denotation R_{ef} .

Eradication of the pandemic via vaccination

As previously mentioned, the calculation of R_0 is used in vaccination strategies. Its value may be determined using various methods: 1) direct count (Cauchemez *et al.*, 2019); 2) statistical methods (Cori *et al.*, 2013; Wallinga and Teunis, 2004); 3) mathematical models used as exponential growth models or compartment models like SIR and SEIR (Liu, Gayle, Wilder, & Rocklöv, 2020).

The R_0 value for SARS-CoV-2 has been reported by several authors. In relation to the first 425 cases confirmed in Hubei, authors Li, Guan, & Wu. (2020) estimated a value of 2.2. According to Liu *et al.* (2020), the two studies based on stochastic models, estimated a R_0 between 2.2 and 2.68. Another important group of researchers led by the University of New York (Tang *et al.*, 2020) proposed a deterministic model with more parameters (SEIR), and determined a R_0 value as high as 6.47. However, Yang and Wang (2020) estimated the R_0 value through mathematical fit, resulting in $R_0 = 4.25$.

Despite the dispersion of results, the studies done through statistical methods and the exponential growth methods have demonstrated to provide reasonably comparable values between 2.2 and 3.58 (Liu *et al.*, 2020).

To determine the effective reproduction factor R_t throughout the pandemic in Cuba, it was important to first determine R_0 , the basic reproduction number, which is specific for SARS-CoV-2. However, the R_0 value could vary among different regions and countries, mainly influenced by differences in population density, and cultural aspects.

The adjustment method in the exponential growth applied to the evolution of the disease was used to determine R_0 in the province of Hubei, China, where the pandemic began. Data from WHO (OMS, 2020a) were used, and the data from the first two weeks were included for fitting, thus covering only the exponential growth phase, and avoiding using data influenced by measures to control the pandemic. The fitting model was identified as,

$$I = I_0 e^{(\beta - \gamma)t} \quad (21)$$

β was set as the mean speed of transmission, and $\frac{1}{\gamma}$ was the mean period of infection, resulting in:

$$R_0 = \frac{\beta}{\gamma} \quad (22)$$

Fitting of the exponential model was done using MATLAB's *Curve Fitting*. The *Trust-Region* algorithm was used for fitting. The determination coefficient of R^2 was 0.98. The β and γ values were 0.3388 and 0.0937, respectively. After applying the equation (22), the R_0 value determined was 3.62, very near the ones reported by Liu *et al.* (2020) for this method.

Fig. 5 shows proper fitting of the exponential growth model to the true data. Also, from day 14 on, the data moved away from exponential behavior, thus stressing the need of careful fitting to determine R_0 , making sure that the data were taken at an early stage. The change in data behavior from this moment on evidenced the beginning of the effect of the measures to mitigate the pandemic.

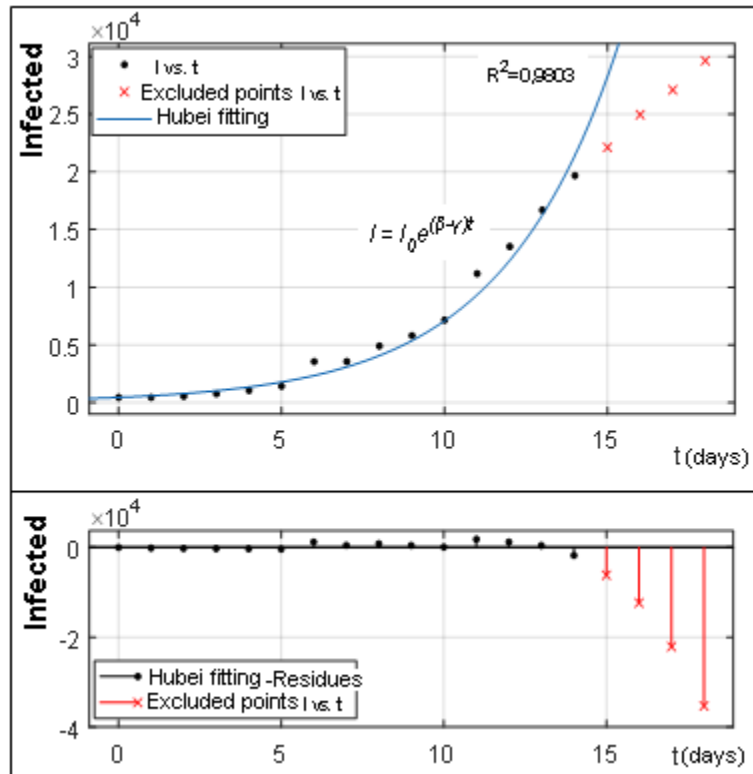


Fig.5. Adjustment curves of the exponential model to determine β and γ

From the equation (20) on, R_t can be determined in the Cuban case. Considering that almost the entire population of Cuba (N) is susceptible, except for the few cases

recovered from the disease, $S \approx N$, so:

$$R_t \approx R_0 \approx 3.62 \quad (23)$$

Considering that the pandemic is eradicated when $R_t < 1$, and based on equation 20, the eradication is said to take place when:

$$\frac{S}{N} < \frac{1}{R_0} \quad (24)$$

Therefore, it was determined that the fraction of susceptible people in Cuba $\left(\frac{S}{N}\right)$ must be less than 0.28, so that the pandemic can be eradicated; that is, less than 28% of the Cuban population.

Eradication of the pandemic in the absence of a specific pandemic vaccine

Using the same method to determine R_0 , the R_c values were determined in a group of the countries affected first, plus Cuba. The results are shown in Table 4.

Table4. R_c values determined in several affected countries, including Cuba

Country	R_c
Spain	2.06
Italy	3.27
The USA	2.39
Russia	1.65
Cuba	1.20

The R_c values of these countries —which apart from Cuba, have been highly affected by COVID-19— are significantly lower than the one estimated in the Chinese province of Hubei = 3.62. It can be explained by the fact that these values are determined by some measures to control the pandemic that were taken at the beginning, because it was already known, in contrast to the case of Hubei. The low value of R_c in Cuba demonstrated the effectiveness of mitigation measures from the first moment of the pandemic.

Later, the extinction of the pandemic without vaccination requires the reduction of the

R_{ef} value. In that case:

If $R_{ef} > 1$, the pandemic spreads.

If $R_{ef} < 1$, the pandemic is eradicated.

Upon the application of the statistical-mathematic method of Cori *et al.* (2013), the behavior of R_{ef} was obtained for the Cuban case, in the March 19-June 2, 2020 period (Fig. 6).

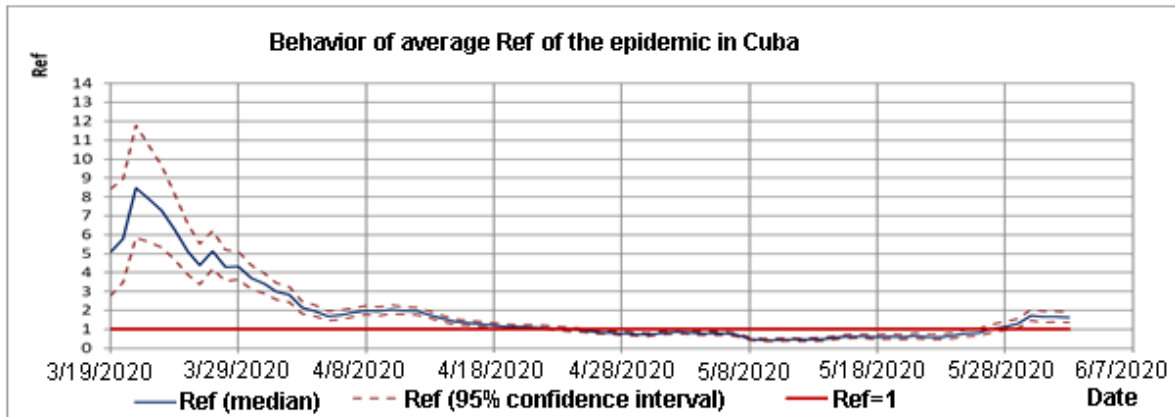


Fig.6. R_{ef} behavior curve in Cuba

Fig. 6 shows that after a period of high R_e , it was lowered to 1, in the April 23-May 27 period. Unfortunately, R_e has surpassed the threshold of 1 in recent days, which demonstrates the need to keep a strict observation of the control measures initially set. It must be considered that though $R_{ef} < 1$ is obtained, these results will not be consistent. The pandemic will be eradicated if the containment measures that produced the earlier values are maintained. Lifting them might lead to an epidemic outbreak if a positive case appears again in the country.

Reflections about the results of modeling for decision-making

In this paper, mathematical modeling has been used in two key issues that contribute to decision-making: the conception, fitting, and the utilization of the SIR model for prediction, as well as determination and utilization of R_0 , R_c , and R_{ef} .

The utilization of SIR permitted the prediction of the peak of the pandemic two weeks in advance. This prediction can be very useful to make decisions in human and material

resource planning, so the negative effects of the pandemic can be mitigated. The family of prediction curves obtained indicates the most likely interval of dates for the occurrence of the peak, and the amount of infected people at that moment. The accuracy observed was adequate, when comparing the prediction with the true data of the peak, in red, in Table 1.

In further research, the application of the SIR model could be done by territories, so resource planning could not only be made temporally, but spatially, ensuring planning based on the needs of each province.

Meanwhile, the calculated R_0 helped determine that to eradicate the epidemic through vaccination, the Cuban susceptible population must be under 28%, which represents 3.18 million people. It means that, at least, the rest of the population of the country (72%) should be immunized (approximately 8.16 million), a representation of the entire population homogeneously distributed as a whole. As long as this is not achieved, the mitigation measures indicated by the government should not be called off.

The determination of R_c allowed for corroboration of the effectiveness of the control measures in Cuba, since the very onset of the pandemic. Keeping a low R_{ef} value in the country will entirely depend on the observation of the control measures. Fig. 7 shows the effect of R_t or R_{ef} values on the propagation of the disease. Three R_t values were used to facilitate scheme development: four, which is close to the initial value observed in Hubei; 1.25, close to the Cuban R_c ; 0.75, which is the desired value, under 1.

Fig. 7a) shows that with $R_t = 4$, the number of cases infected increases 16-fold in only two steps of the transmission chain. In Fig. 7b) with $R_t = 1.25$ (near the Cuban R_c), the number of infected people is not doubled (1.95 cases), following three steps in the transmission chain, with an average of 2.44 cases after four steps, which is way below the 256 cases after four steps, if $R_t = 4$. This explains the enormous difference in terms of propagation of the disease in Cuba, and a number of important countries in the world. In the desired case shown in Fig. 7c), with $R_t = 0.75$, the number of infected people is reduced from 4 to 1.27 on average, and the pandemic is eradicated after five steps of the transmission chain, with no new cases (0.95 cases). Column 1.b in the figure, is simply the sum of cases on column 1.a), to form whole numbers.

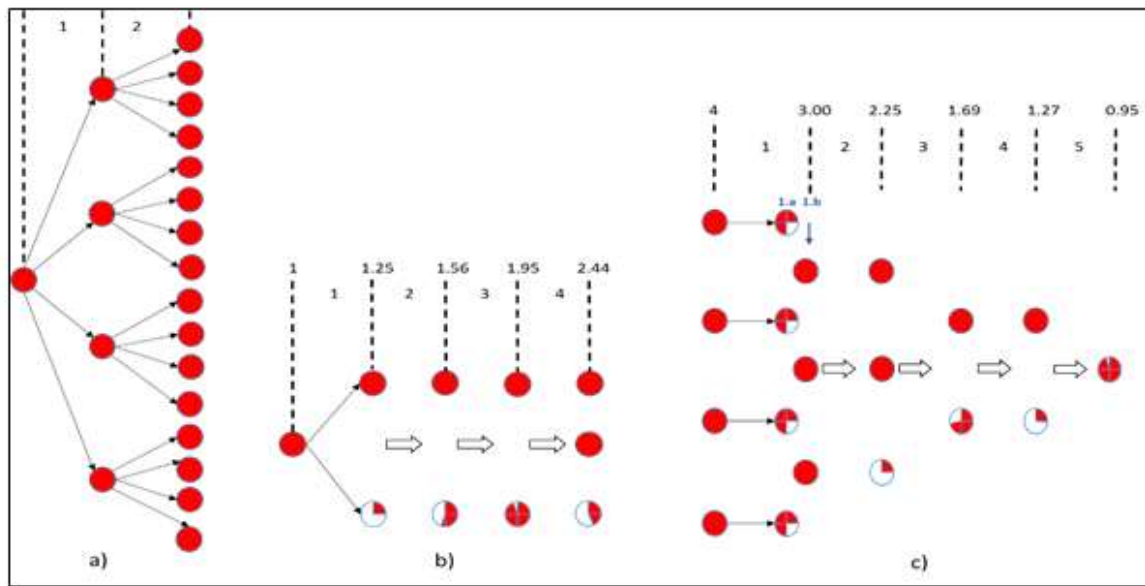


Fig.7. Reproduction dynamic of the pandemic according to the effective reproduction factor (R_t or R_{ef}) a) $R_t = 4$, b) $R_t = 1.25$, c) $R_t = 0.75$

Based on the analysis of Fig. 7c), the number of steps in the transmission chain expected to take place before the infected people disappear, is reduced, thus indicating that the total eradication of the pandemic ($R_{ef} < 1$) would be effective in quite a long period of time. This, along with the waiting time before the appearance of a specific vaccine, may be a warning sign of the need of long-term planning in the fight against the COVID-19 epidemic in Cuba, and any other parts of the world.

CONCLUSIONS

The mathematical models are powerful decision-making tools in face of complex situations; with demonstrated potential and feasibility of application, especially in understanding and predicting the dynamic of epidemic transmission.

The SIR model demonstrated the capacity of predicting the time and magnitude of the peak of the pandemic, as well as the dynamic of infected, recovered, and deceased, two weeks in advance, and a reasonable accuracy to make effective decisions in terms of resource availability and mitigation measures.

The determination of the basic number of reproduction of the pandemic (R_0) of COVID-19 allowed for corroboration of the high level of SARS-CoV-2 transmission, comparing it with infectious agents. Its application in the Cuban case indicated that more than 72% of the Cuban population should be vaccinated to eradicate the pandemic via vaccination.

The comparison between the dynamics of the pandemic and the values of the number of control reproduction in Cuba, and the values of several of the first infected countries, showed that since the beginning, the mitigation measures were efficacious to fight COVID-19 in the Cuban context.

The differences between the number of basic reproduction (R_0) and the control and effective reproduction numbers, demonstrate that to eradicate the pandemic without a specific vaccine against COVID-19, the control measures should be in place, even with R_{ef} values below 1. Meanwhile, the eradication of the disease under these conditions will not be sustained if after the elimination of the measures a positive case is detected in the country, leading to an epidemic outbreak.

SARS-CoV-2 transmissibility, the international situation, the absence of a vaccine, and the number of steps in the transmission chain required to eradicate the pandemic, demonstrated in this research, implies the need to plan long-term control measures, both to eradicate the pandemic in Cuba and to prevent an outbreak.

REFERENCES

- Arrendondo, F. G. (2013). Un modelo de análisis racional para la toma de decisiones gerenciales, desde la perspectiva elsteriana. Cuadernos de Administración, 26(46), 135-158.
- Bartholomew, M. (2008). *Nonlinear Optimization with Engineering Applications* (vol. 19). New York, USA: Springer.
- Brauer, F., Castillo, C., & Feng, Z. (2019). Mathematical Models in Epidemiology. En, *Texts in Applied Mathematics* (p. 619). USA: Springer Science. DOI:10.1007/978-1-4939-9828-9
- Cauchemez, S., Hoze, N., Cousien, A., Nikolay, B., & Ten Bosch, Q. (2019). How Modelling Can Enhance the Analysis of Imperfect Epidemic Data. *Trends in*

- Parasitology*, 35(5), 369-379. DOI:10.1016/j.pt.2019.01.009
- Chowell, G. (2017). Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling*, 2(1), 379-398. DOI:10.1016/j.idm.2017.08.001
- Cori, A., Ferguson, N. M., Fraser, C., & Cauchemez, S. (2013). A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of epidemiology*, 178(9), 1505-1512.
- De Kohan, N. C. (2015). Los sesgos cognitivos en la toma de decisiones. *International Journal of Psychological Research*, 1(1), 68-73.
- Galindo, S., Rodríguez, M. A. y Cervantes, J. L. (2014). Las matemáticas de las epidemias: caso México 2009 y otros. *CIENCIA ergo -sum*, 20(3), 238-246.
- Getz, W. M., Salter, R., Muellerklein, O., Yoon, H. S., & Tallam, K. (2018). Modeling epidemics: A Primer and Numerous Model Builder Implementation. *Epidemics*, 25(2018), 9–19. DOI:10.1016/j.epidem.2018.06.001
- Hillier, F. S. y Lieberman, G. J. (1997). *Introducción a la investigación de operaciones*. Mexico: McGraw-Hill Interamericana Editores.
- Kermack, W. O., & McKendrick, A. G. (1927). Contributions to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London*, 115(772), 700-721.
- Khan, A. & Zaman, G. (2018). Global analysis of an age-structured SEIR endemic model. *Chaos, Solitons and Fractals*, 108(2018), 154-165. DOI:10.1016/j.chaos.2018.01.037
- Kwok, K. O., Tang, A., Wei, V. W. I., Park, W. H., Yeoh, E. K., & Riley, S. (2019). Epidemic Models of Contact Tracing: Systematic Review of Transmission Studies of Severe Acute Respiratory Syndrome and Middle East Respiratory Syndrome. *Computational and Structural Biotechnology*, 17(2019), 186-194. DOI:10.1016/j.csbj.2019.01.003
- Lavine, J. S., Poss, O., & Grenfell, B. T. (2008). Directly transmitted viral diseases: modeling the dynamics of transmission. *Trends in Microbiology*, 16(4), 165-172.
- Leib, T. M., & Pereira, C. J. (2008). Reaction Kinetics. En J. O. Maloney (Ed.), *Perry Chemical Engineers Handbook* (8th ed., sección 7, pp. 7-39). USA: McGraw-Hill.
- Leung, K., Wu, J. T., Liu, D., & Leung, G. M. (2020). First-wave COVID-19

- transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *Lancet*, 395(2020), 1382-1393. DOI:10.1016/S0140-6736(20)30746-7
- Li, Q., Guan, X., & Wu, P. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *New England Journal of Medicine*, 382, 1199-1207. DOI:10.1056/NEJMoa2001316
- Linares, P., Ramos, A., Sánchez, P., Sarabria, A. y Vitoriano, V. (2001). *Modelos matemáticos de optimización*. Madrid, España: Universidad Pontificia Comillas.
- Liu, Y., Gayle, A. A., Wilder, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27(2), 1-4. DOI:10.1093/jtm/taaa021
- Ma, J. (2020). Estimating epidemic exponential growth rate and basic reproduction number. *Infectious Disease Modelling*, 5(2020), 129-141. DOI:10.1016/j.idm.2019.12.009
- Martín, C. E., Loredó, N. A. y Álvarez, N. M. (2016). Diagnóstico de la toma de decisiones no estructuradas por directivos camagüeyanos. *Retos de la Dirección*, 10(2), 1-17
- MINSAP. (2020). *COVID-19 Parte de cierre del día 28 de abril a las 12 de la noche*. Retrieved from <https://salud.msp.gob.cu/parte-de-cierre-del-dia-28-de-abril-a-las-12-de-la-noche/>
- Miranda, E. (2015). Perfil de competencias del profesional que cumple funciones de monitoreo, evaluación y gestión de evidencias de programas y proyectos de desarrollo: un aporte para las decisiones basadas en evidencias. *Anales de la Facultad de Medicina*, 76 (spe), 67-76.
- OMS. (2020a). *Coronavirus disease (COVID-19), Situation Report – 134*. Retrieved from <https://www.who.int>
- OMS. (2020b). *COVID-19: Cronología de la actuación de la OMS*. Retrieved from <https://www.who.int>
- ONEI. (2019). *Anuario demográfico de CUBA*. La Habana, Cuba: Centro de Estudios de Población y Desarrollo.
- Prakash, B., Setia, A., & Alapatt, D. (2017). *Numerical solution of nonlinear fractional*

- SEIR epidemic model by using Haar wavelets*. India: Department of Mathematics, BITS Pilani, K K Birla Goa Campus, India.
- Robbins, S. P. y Coutler, M. (2010). *Administración*. México: Pearson Education.
- Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling*, 5(2020), 271-281. DOI:10.1016/j.idm.2020.03.001
- Stoner, J. A., Freeman, R. E., & Gilbert, D. R. (1996). *Administración* (6ta ed.). México: Prentice-Hall Hispanoamericana S.A.
- Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., & Xiao, Y. (2020). Estimation of the transmission risk of 2019-nCoV and its implication for public health interventions. *Journal of Clinical Medicine*, 9(2020), s.p.
- Venkatramanan, S., Lewis, B., Chen, J., Higdon, D., Vullikanti, A., & Marathe, M. (2018). Using data driven agent based models for forecasting emerging infectious diseases. *Epidemics*, 22(2018), 43–49. DOI:10.1016/j.epidem.2017.02.010
- Wallinga, J., & Teunis, P. (2004). Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. *American Journal of Epidemiology*, 160(6), 509-516.
- Wang, L., Wang, Y., Ye, Y., & Liu, Q. (2020). Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents*. 55(6), s.p. DOI:10.1016/j.ijantimicag.2020.105948
- Wang, X., Peng, H., Shi, B., Jiang, D., Zhang, S., & Chen, B. (2018). Optimal Vaccination Strategy of a Constrained Time-Varying SEIR Epidemic Model. *Communications in Nonlinear Science and Numerical Simulation*. (2018), 1-29. DOI:10.1016/j.cnsns.2018.07.003
- Yang, C., & Wang, J. (2020). A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Mathematical Biosciences and Engineering*, 17(3), 2708-2724. DOI:10.3934/mbe.2020148

Conflicts of interest and conflict of ethics statement

The authors declare that this manuscript is original, and it has not been submitted to another journal. The authors are responsible for the contents of this article, adding that it contains no plagiarism, conflicts of interest or conflicts of ethics.

Author contribution statement

1. Héctor Eduardo Sánchez Vargas. Conceptualization, redaction-proofreading, and editing. Formal analysis research, methodology, software Redaction of the original draft.
2. Luis Beltrán Ramos Sánchez. Conceptualization, formal analysis, resources. Redaction-proofreading, and editing.
3. Pablo Ángel Galindo Llánez. Data curation. Project management. Resources Redaction-proofreading, and editing.
4. Amyrsa Salgado Rodríguez. Conceptualization and redaction of the original draft.