Department: Head
Editor: Name, xxxx@email

Discrete-Time Modeling of COVID-19 Propagation in Argentina with Explicit Delays

Mariana Bergonzi
CIFASIS-CONICET, FCEIA-UNR
Rosario, Argentina
bergonzi@cifasis-conicet.gov.ar

Ezequiel Pecker-Marcosig
ICC-CONICET, FI-UBA
Buenos Aires, Argentina
emarcosig@dc.uba.ar

Ernesto Kofman
CIFASIS-CONICET, FCEIA-UNR
Rosario, Argentina
kofman@cifasis-conicet.gov.ar

Rodrigo Castro
ICC-CONICET, FCEyN-UBA
Buenos Aires, Argentina
rcastro@dc.uba.ar

***Abstract*—We present a new deterministic discrete time compartmental model of COVID-19 that explicitly takes into account relevant delays related to the stages of the disease, its diagnosis and report system, allowing to represent the presence of imported cases. In addition to developing the model equations we describe an automatic parameter fitting mechanism using official data on the spread of the virus in Argentina. The result consistently reflects the behaviour of the disease with respect to characteristic times: latency, infectious period, report of cases (confirmed and dead) and allows for detecting automatically changes in the reproductive number and in the mortality factor. We also analyse the model's prediction capability and present simulation results assuming different future scenarios. We include a usage of the model in a closed loop control scheme, where the explicit presence of delays plays a key role in projecting more realistic dynamics than that of classic continuous time models.**

THE OUTBREAK of the coronavirus disease 2019 (COVID-19) as of late December in the city of Wuhan, mainland China, took eventually three months to be classified as a pandemic [1] by the World Health Organization. About the same date, on March 3rd., the first case of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was officially reported in Argentina by the Ministry of Health [2]. The Argentine national administration was very early in taking quick measures to curb the spread of the virus. Throughout the first three weeks of March several public activities were progressively banned, ending up in a nationwide mandatory lockdown (so called Mandatory Preventive Social Isolation (or ASPO for its initials in Spanish) as of March 19th. The compliance with the new regulations was exemplary during approximately three weeks, followed by a mixture of relaxation of controls combined with a loss of commitment by some social sectors. This resulted in a very particular slow but sustained exponential growth of cases. As of the time of writing, more than 650,000 positive cases have been confirmed, more than 500,000 have recovered and around 15,000 died.

During the early weeks of the epidemic in the country, several health-related systems were stressed at unprecedented levels, including the information systems meant to monitor and track the status of the epidemic nationwide. After 5 months of operation the result converged into a publicly available subset of information updated twice a day. The dataset bears several peculiarities in the informed dates, which are a consequence of several factors such as people reporting symptoms very late, varying delays in the testing system, slow processing and data entry of new cases and test results, etc.

This situation impacts directly on data-driven approaches that try to model, simulate and forecast the evolution of the disease relying on the official datasets as primary sources of information.

The primary application of simulation models during an on-going, previously unknown, and highly infectious disease is to assist health policy makers with tools to assess the impact of potential interventions within pressing time horizons. Without the possibility of waiting for enhanced data sources, and with delayed decisions potentially costing lives, it becomes highly relevant to apply modeling techniques that are robust to available data inaccuracies.

In this work, we present an approach that relies on discrete time difference equations with explicit delays, tailored to best assimilate the daily data updates reported for Argentina. The use of explicit delays has the advantage of improving the model fitting using short data series, which occurs not also at the beginning of the outbreak but also after non pharmaceutical interventions (like strict lockdown). Additionally, it permits using more *instantaneous* type of data (e.g. daily reported cases) in a more reliable way. The use of this type of *noisy* data avoids the need to wait for the stabilization of the time series of cases (typically resorting to the date of symptom onset) which imposes a lag in the order of 1 to 2 weeks in the analysis. Another advantage of the approach is that it yields an explicit and straightforward representation of the model parameters. The combination of these facts with the discrete time formulation results in a model that can be easily understood by non expert policy makers.

We will compare this approach with a classic continuous time differential equation-based model without delays, and discuss the relevance of both approaches.

Models of infectious diseases and previous work on models for COVID-19 in Argentina

Modeling the spread of contagious diseases is a broad field [3], [4], [5], with approaches including both deterministic or stochastic dynamics, discrete time or continuous time, compartmental or agent-based, with or without explicit georeferencing, to name a few. A widely adopted classification for epidemiological models is to state the possible stages through which an agent can evolve in the population model [4]. This way, the agents are conceptualized as belonging to mutually exclusive compartments, originating the category of *compartmental models*. The most common compartments are the Susceptible, Infected and Recovered (the SIR model) while finer grained classifications include stages such as Exposed (infected but not yet contagious), Asymptomatic, Quarantined, Hospitalized and Dead (obtaining

for instance a SEIRDH type of model).

There is already a number of published efforts to model the COVID-19 evolution for the Argentinian case. In [6] the authors propose an agent-based, spatially explicit, and age-structured SIRD model exploring intervention scenarios reflected by demographic, medical, social and institutional parameters. In [7] a connection is established between short-time time series statistical forecasts and structural parameters of a continuous time SIRD model, with special attention on the occurrence of epidemic waves and focusing on the City of Buenos Aires. In [8] the authors developed a continuous time, stochastic SEIARQ model including age structure, a distinction according to the severity infected cases and focusing on the Metropolitan Area of Buenos Aires. Different interventions are studied considering varied percentages of the population going quarantined. In [9] the spread of the disease is modeled with a SEIHRD structure that is used as the target of a predictive closed-loop control strategy based on a proportional controller. The control goal is to avoid the collapse of the health system while reducing economic impact. In [10] the authors chaperon the reader in a guided walk-through of modeling experiences, sharing their lessons learned from studying the evolution of COVID-19 in Argentina. They span varied types of SEIR models (homogeneous and inhomogeneous, stochastic and deterministic, spatially lumped and distributed, continuous and discrete time) and also include a valuable study to estimate mobility within and between districts from cell phone data.

An adequate choice of a combination of model type and represented disease stages depends largely on the questions to be answered by the model, and consequently by the spatio-temporal scale at which answers are expected to hold valid.

In the most abstract type of model structure, the compartments in a compartmental model represent the whole population for a given spatial scale (neighbor, city, country). This approach involves the least amount of parameters, but ignores potentially relevant dynamics at smaller scales. Such populations can in turn be subdivided into ”classes” (e.g. age cohorts, essential vs. non-essential personnel, etc.) obtaining so-called ”metapopulation” models. This subdivision into n classes comes usually at the expense of multiplying the number of required parameters by a factor between n and n^2 (depending on the interaction structure among classes). The aforementioned structures can in turn be made more spatially explicit, for instance considering networked populations (e.g. interconnected cities) or continuous regions where adjacent patches influence each other according to their neighboring context (e.g. cellular automata models). This model structure can provide more fine grained answers regarding the spread of the spatial density of the disease, but requires a thorough understanding of population mobility (which is a challenging data processing task even when reliable, non-biased data is available) and are in general much more difficult to validate. The degree of detail can be taken a step further by considering agent-based models where individuals in a population are modeled with autonomous behavior. Here, the macroscopic indicators of the disease in the population emerge from the interactions among said agents, each of them undergoing the typical stages of the disease. Agents can be either fixed or mobile. In the first case, the underlying (usually random) connectivity graph determines the possibilities of contact between each pair of persons, while in the second case agents move in space according to given rules and the opportunities for interaction (e.g. contagious encounters) emerge from the aggregated pattern of mobility. Finally, stochastic dynamics can be factored in within all types of models, to cope with the inherent uncertainties found in most model parameters (e.g. incubation period, recovery delay, effective transmission rate, etc.). Models with stochastic behavior provide more information on the statistical distribution of relevant variables (such as daily new infections, recoveries, deaths) besides their averaged values. In many cases, and particularly in the presence of unprecedented system dynamics as in the case of COVID-19, reliable statistical descriptions of model parameters are not available. In this context, a reasonable approach is to conduct scenario studies, i.e. testing for the robustness of the simulated results by considering plausible sets of perturbed values for model parameters. We refer the reader to [11] where the authors provide a very illustrative overview and discussion of

model types in the context of a vector-host virus transmission (including the possibility of adopting simultaneously different model structures for the vector and the host).

A Discrete time model with explicit delays

Compartmental models are most typically formulated in continuous time, in the form of sets of ordinary differential equations, with fewer efforts devoted to discrete time variants. For the particular case of the COVID-19, examples of works that fit continuous models to data can be found in [12], [13], [14] while approaches relying on discrete time models can be found in [15], [16].

A salient feature of the COVID-19 virus is its long incubation time that delays both the start of the infection period and then the time of detection. For this reason, it seems appropriate to model explicitly the presence of such *delays*. Moreover, in the case of an unprecedented pandemic outbreak, during early spreading stages in a given country imported cases are very frequent and strongly influential. This was particularly true in Argentina, where more than a month after the confirmation of the first case half of the cumulative detected cases were still imported ones.

The incorporation of explicit delays and imported cases complicates the approach based on differential equations, as it implies using delayed differential equations, whose numerical resolution can be challenging [17].

Therefore, we propose a discrete time model with a one-day time step that, unlike traditional discrete time models [5], takes explicitly into account time delays (e.g. incubation, detection, recovery and death intervals since the instant of exposure to the virus). In addition, we incorporate imported cases as a forcing input signal.

We later fit the parameters of this model to publicly available data provided by the Ministry of Health of Argentina. As it will be seen, the parameters that minimize the mean square error between data and the simulated trajectories are consistent with known information about COVID-19 in our country (incubation time, infectious period, recovery times, death, detection and reporting).

The origin of the structure of compartmental models goes back to the work of Kermack and McKendrick in 1927 [18] and consists of classifying the population into three different groups (or *compartments*):

- $S(t)$: Population *susceptible* to the virus in time t. (Individuals who are not infected, nor have immunity to the virus).
- $I(t)$: Population *infected* in time t. (Individuals who are infected and can spread the virus to those who are susceptible).
- $R(t)$: Population *removed* in time t. (Individuals who have already been removed from the dynamics, either because they have recovered by acquiring immunity to the virus, or have died).

Each individual in the population belongs to only one of these three groups and can evolve from one to the other. One of the hypotheses is that the recovered individuals acquire lifelong immunity to the virus, so the allowed evolutions are: a) an individual of S can become infected by entering in contact with one of the I group (becoming part of this group), b) an individual of the I group ends up passing to the R group when they recover or die. The total number of individuals in the population is assumed to be constant with $S(t) + I(t) + R(t) = N$.

There are other mathematical models based on the SIR model, to which other compartments or additional characteristics are added: In the SEIR model, the population category E (*exposed*) is added to model individuals that are incubating the virus but do not yet have the capacity to infect. Similarly, the SEIRD model adds the category D corresponding to the *dead* population due to the epidemic.

Continuous time models In large populations the variables can be taken as continuous and the relationship between the three groups can be represented with a system of ordinary differential equations:

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta \cdot S \cdot I \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta \cdot S \cdot I - \gamma \cdot I \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma \cdot I\end{aligned}$$

Where β and γ are parameters that represent the infection and recovery rates respectively.

From the relation of these parameters a fundamental indicator arises: R_0, the basic reproduction number. This factor indicates, on average, how many people get infected by a previously infected person.

Discrete time models with delay Discrete time models [19], [5] support the incorporation of explicit delays to model behaviors that occur after a known period of time [5]. An example could be the 'latency time' of a disease, i.e. the time since the individual's first exposure to the virus until he or she starts being infective.

With respect to the incorporation of delays, there is an antecedent in [20] where a SEIR structure is proposed but a small distinction is made in the compartments. Here, the compartment I represents all infected (both detected and undetected) who are already contagious, but do not yet have symptoms. In addition, R represents those infected who have been detected (the R is for 'reported') and U represents those infected but not yet detected.

In this work, two different options are proposed to represent latency periods. On the one hand, using rates to represent the flows between compartments, and on the other by means of explicit delays obtaining *Delay Difference Equations*.

Model description

In this section, we first introduce the proposed model SEIRD and then we compare it with a classic continuous time SEIR model.

Model Equations

Following the idea of classic SEIRD models, we use the following state variables:

- $S(t)$: Susceptible population
- $E(t)$: Exposed population
- $I(t)$: Infectious population
- $R(t)$: Removed population
- $D(t)$: Dead population

Note that the Dead population (D) is part of the Removed population (R). In order to explicitly consider the effect of delays, we also use as state variables the number of *Daily Exposed* people in the last T days: $N^E(t), N^E(t-1), \ldots, N^E(t-T)$, where T is the maximum delay in the model.

In addition, in order to account for imported cases, we consider a signal $U(t)$ that represents the number of *imported cases* that were exposed at time t while they were abroad and were later detected in the population. This signal is assumed to be known. The model parameters, that will be later fitted to the available data, are the following:

- $R_0(t)$: Basic Reproductive Number. Average number of individuals exposed to the virus by a single infectious person.
- τ_I: Latent Time: The time elapsed since an individual is exposed until becoming infectious.
- τ_R: Removal Time. The time elapsed since an individual is exposed until he or she no longer infects others. Notice that $\tau_R - \tau_I = \tau_{\text{Inf}}$ is the duration of the infectious period. This time depends not only on the time it takes to reduce the virus load, but also on the policies applied to isolate individuals after detecting or suspecting the possibility of infection.
- d_{f}: Diagnostic Factor. The fraction of positive cases that are eventually diagnosed. Notice that this value depends on testing and diagnosis policies.
- $c_{\text{fr}}(t)$: Case Fatality Rate: The fraction of confirmed cases that eventually die. We assume that c_{fr} may change over time.
- τ_D: Diagnosis Time: Time elapsed since exposure until diagnosis. This delay depends not only on biological features (such as symptoms onset) but also on testing and reporting policies.
- τ_M:Time of Death: Time elapsed since exposure until death.

We shall also compute the average number of individuals exposed by a single infected person per day as $\beta(t) = R_0(t)/(\tau_R - \tau_I)$.

With these state variables and parameters, the model has the following dynamics:

$$
\begin{aligned}
N^E(t+1) =& \beta(t) \cdot I(t) \cdot S(t) + U(t+1)/d_{\text{f}} \\
S(t+1) =& S(t) - \beta(t) \cdot I(t) \cdot S(t) \\
E(t+1) =& E(t) + N^E(t+1) \\
& - N^E(t+1-\tau_I) \\
I(t+1) =& I(t) + N^E(t+1-\tau_I) \\
& - N^E(t+1-\tau_R) \\
R(t+1) =& R(t) + N^E(t+1-\tau_R)
\end{aligned}
$$

In order to fit parameters and perform projections, the model computes the following signals corresponding to reported data:

$$D(t+1) = D(t) + (c_{\text{fr}} \cdot d_{\text{f}}) \cdot N^E(t+1-\tau_M)$$
$$I_D(t+1) = I_D(t) + d_{\text{f}} \cdot N^E(t+1-\tau_D)$$

representing the number of accumulated deaths and diagnosed cases, respectively.

The model relies on the following assumptions:

- The daily number of exposed people is proportional to the number of infectious people and to the fraction of susceptible population (without taking into account the number exposed from abroad).
- The fraction of diagnosed cases d_{f} is the same for both local and imported cases, and it does not change over time.

Notice that when the susceptible population remains nearly constant, the model is almost linear. Moreover, assuming also null initial conditions such that the dynamics is first started by the inflow of imported cases, it turns out that the evolution of the observed signals $I_D(t)$ and $D(t)$ is independent on the detection factor d_{f}. This factor appears dividing the input $U(t)$ and multiplying the outputs $I_D(t)$ and $D(t)$. This feature implies that the dynamics can be inferred without actually knowing the detection factor d_{f}.

Parameter fitting

In order to fit the parameters we implemented the model described before in Octave. We used as input the data of imported cases in Argentina. The initial time was set to $t = 0$ on 20/2/2020, since the first reported cases (and the first death) date from the beginning of March (thus bringing the first exposure event back to around February 20th.)

We take as the output of the model the number of deaths $D(t)$ and the number of total cases detected $I_D(t)$. These numbers were compared with the corresponding reported data (of deaths and case reports) from 20/2/2020 to 25/4/2020 using the Case Report Date (CRD). All data were obtained from the daily reports of the Ministry of Health at https://www.argentina.gob.ar/coronavirus/informe-diario.

For each set of integer values of the delays τ_I, τ_D, τ_M, τ_R, the values of $R_0(t)$ (reproductive number) and $c_{\text{fr}}(t)$ (case fatality rate, i.e. fraction of deaths per detected and reported case) were adjusted by least squares, and then it was sought which set of delay values minimized the resulting cost. For the least-squares adjustment the built-in `leasqr` Octave function was used, attempting to minimize the difference between the accumulated reported cases and deaths in the dataset and those computed by the model. In order to minimize the norm of the relative error, the `leasqr` function was invoked using a weighting matrix that is inversely proportional to the square root of each data point.

Denoting with $y^{\text{data}}(t) \triangleq [D^{\text{data}}(t), I_D^{\text{data}}(t)]^T$ and $y^{\text{sim},p}(t) \triangleq [D^{\text{sim},p}(t), I_D^{\text{sim},p}(t)]^T$ to the output data and the simulated outputs for the set of parameters p, respectively, the `leasqr` function tries to find the set of parameters p^* that minimizes the following error:

$$e = \sum_{t=0}^{t_f} \frac{(y^{\text{data}}(t) - y^{\text{sim},p}(t))^2}{|y^{\text{data}}(t)|}$$

The set of parameters p contains an array of values of $R_0(t)$ and $c_{\text{fr}}(t)$ at different times. In this case, we allowed both trajectories to change twice, so p was formed by 3 values of each parameter. The instants of change for those parameters were the ones that allowed to minimize the error after the optimization procedure. As a result, we obtained the following expressions for the reproduction and mortality rates:

$$R_0(t) = \begin{cases} 2.267 & t < 17/3 \\ 0.886 & 17/3 \leq t < 3/4 \\ 1.104 & t \geq 3/4 \end{cases}$$

$$c_{\text{fr}}(t) = \begin{cases} 0.054 & t < 17/3 \\ 0.061 & 17/3 \leq t < 3/4 \\ 0.072 & t \geq 3/4 \end{cases}$$

In all cases the least squares algorithm was executed with initial parameter guess $R_0(t) = 1$ and $c_{\text{fr}}(t) = 0.05$, which are values within the order of magnitude of what is known about the virus. Convergence to this minimum was also

verified for varied initial guess values. These parameters were obtained using the delays $\tau_I = 5$, $\tau_D = 11$, $\tau_R = 12$, and $\tau_M = 17$, which minimize the quadratic cost, and are also consistent with what is known about the virus: 5 days until the patient becomes contagious, 11 days until detection and report, 12 days for the effective contagious period (considering that upon detection the person is isolated) and 17 days until recovery or death [20], [10], [21]. As it was already mentioned, the time until detection, report and isolation are strongly dependant on the local health system in charge of processing swab tests. Likewise, it is highlighted that the time t_1 that is obtained for minimizing errors is almost coincident with the enforcement of the first strict lockdown (t_1 corresponds to March 17th., while the lockdown was enforced on March 19th.). The obtained time t_2 corresponds to April 3rd., when the population began to relax its commitment to strict confinements.

The relative mean error between measured and simulated data is $\|x_{\text{data}} - x_{\text{sim}}\|/\|x_{\text{data}}\| = 0.0613$ (6.13%). Figures 1-2 compare the simulation results with the data corresponding to the number of deaths and the number of detected cases.

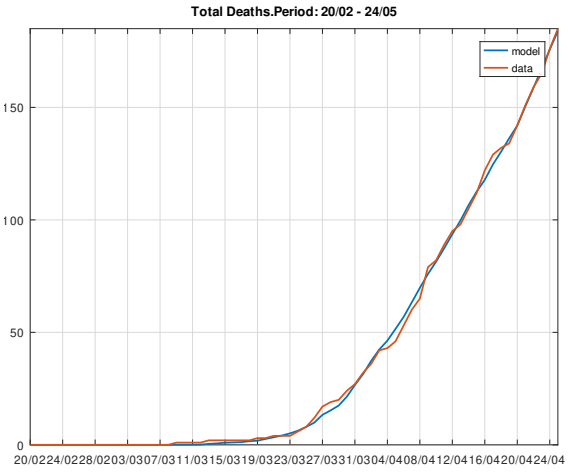


Figure 1: Total number of deaths (data and model)

Experiments and Results

In this section, we show first that the developed model can be used to analyze the past evolution of the contagion figures and to project future scenarios. Then, we demonstrate its usefulness in evaluating non pharmacological interventions and finally we cross-check the model by comparing it against a classic continuous time SEIR model.

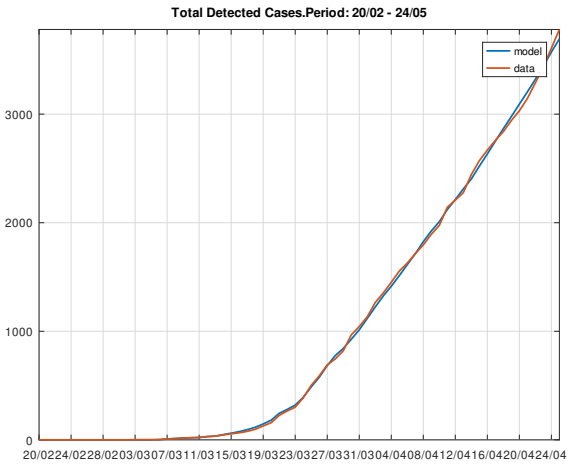


Figure 2: Total number of detected cases (data and model)

Analysis and Projections

We show next the use of the model developed to analyze the evolution of factors of contagion and mortality in the past, on the one hand, and also to project the evolution of the curves in the short and medium term.

We will fit the parameters $R_0(t)$ and $c_{\text{fr}}(t)$ over a longer period, using data from 20/2 to 04/6. We will consider that these parameters can change every 14 days. The main reason for choosing this period between changes is that the Argentine government updated the policies once every two weeks. A more precise adjustment can be obtained by allowing weekly or even daily changes, but optimizing on a large set of parameters would eventually lead to over-fitting. Keeping the delay parameters of the previous experiment ($\tau_I = 5$, $\tau_D = 11$, $\tau_R = 12$ and $\tau_M = 17$) we get then the following paths for $R_0(t)$ and $c_{\text{fr}}(t)$:

$$R_0(t) = \begin{cases} 2.396 & t < 16/3 \\ 0.898 & 16/3 \leq t < 29/3 \\ 1.051 & 30/3 \leq t < 12/4 \\ 1.078 & 13/4 \leq t < 26/4 \\ 2.042 & 27/4 \leq t < 10/5 \\ 1.523 & t \geq 11/5 \end{cases} \quad (1)$$

$$c_{\mathrm{fr}}(t) = \begin{cases} 0.054 & t < 16/3 \\ 0.061 & 16/3 \leq t < 29/3 \\ 0.064 & 30/3 \leq t < 12/4 \\ 0.068 & 13/4 \leq t < 26/4 \\ 0.031 & 27/4 \leq t < 10/5 \\ 0.018 & t \geq 11/5 \end{cases} \quad (2)$$

The relative mean error between measured and simulated data results in $\|x_{\mathrm{data}} - x_{\mathrm{sim}}\|/\|x_{\mathrm{data}}\| = 0.015$ (1.5%).

Figures 3-6 compare the simulation results with data corresponding to the number of deaths, number of detected cases accumulated and per day. In this case, we extended the final simulation time to project beyond the data used to adjust the model. We initially consider that since 12/5 the parameters $R_0(t)$ and $c_{\mathrm{fr}}(t)$ remain constant and equal to the last adjusted value. We also include two scenarios considering a $\pm 10\%$ variation of these parameters from 25/5 onwards (i.e. a period not observed from data due to the delays). These scenarios allow for considering small changes in the social contact and/or in the death rate, as well as some inaccuracies in the parameter adjustment. Due to the unstable nature of the model when $R_0 > 1$, those small changes in the parameter values imply a large difference in the trajectories as time evolves (see Figure 6 where those small changes imply that the number of daily cases is modified by a factor of 3 after six weeks).

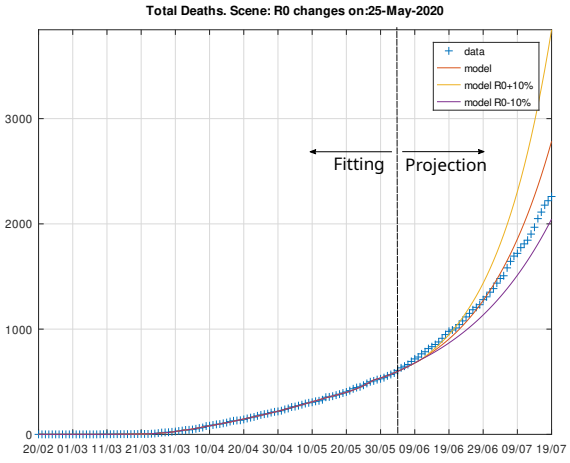


Figure 3: Total number of deaths (data and model)

It can be seen that the fitted model allows us first to draw conclusions about the curve from

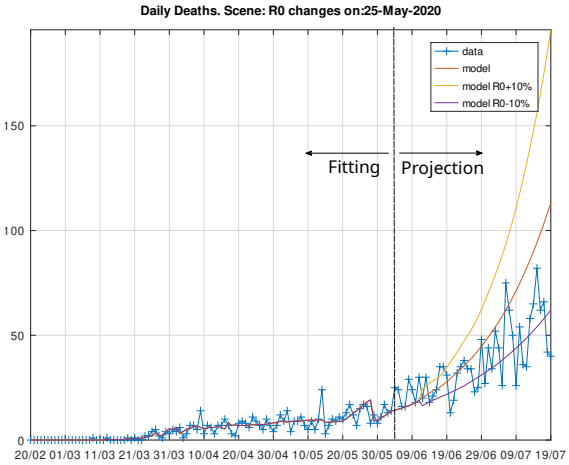


Figure 4: Daily number of deaths (data and model)

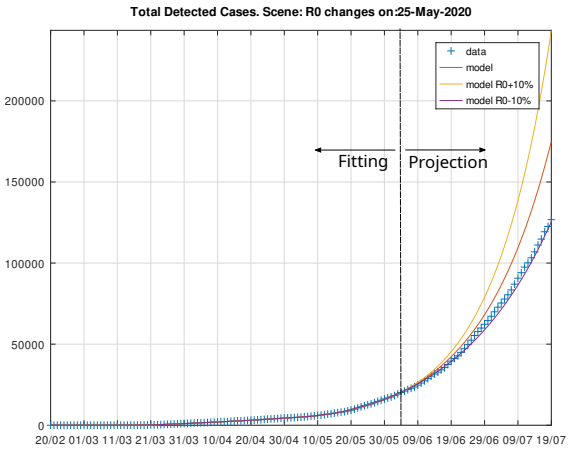


Figure 5: Total number of detected cases (data and model)

past data. Equations (1)-(2) show the evolution of the factor of contagion and the effects of a quarantine and its subsequent relaxation, as well as the decrease in mortality per case that reflects the greater spread of the virus among the younger population.

On the other hand, the projections from June 4 (final date of the data used for the adjustment) show first the possibility of using the model to estimate the future the spread of the virus under different scenarios. On the other hand, it allows to analyze the potential effect of the increase or decrease of the social distancing reflected in the factor R_0.

The coincidence between the projections made at the beginning of June with the subsequent data up to 6 weeks later reflects not only the

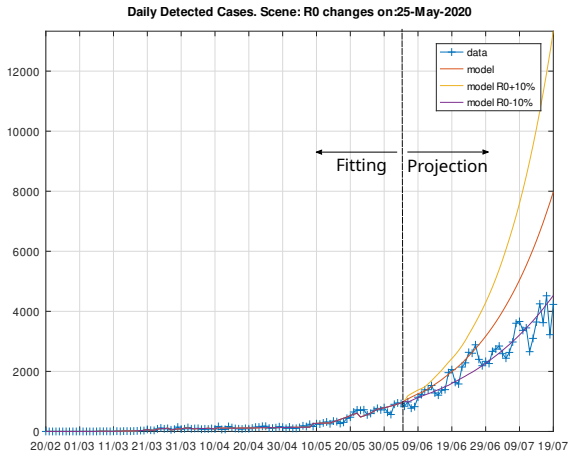


Figure 6: Daily number of detected cases (data and model)

correctness of the model and the accuracy in its estimated parameters, but also the lack of change in the social behavior. We have used the model in different large cities of Argentina (Buenos Aires and Rosario, for instance) where the initial projections were accomplished for more than 10 weeks.

Applying continuous and discrete models for non-pharmaceutical interventions.

In this section we will address a comparison between a classic Continuous-time SEIR Model (CM) and the Discrete-time SEIRD Model (DM) presented in this work.

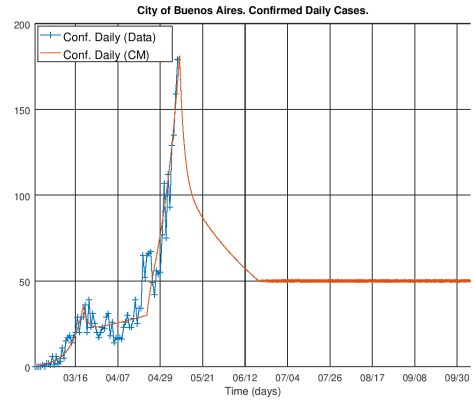
Different approaches were proposed to contain the spread of COVID-19 in the absence of a vaccine, resorting to public-health measures known as non-pharmaceutical interventions [22]. We present here a *Controlled Intermittent Lockdowns* (CIL) strategy combining periods of suppression $(R0 < 1)$ and mitigation $(R0 > 1)$. The control policy is parameterized by a threshold (`maxCases`) which defines when to switch between strict lockdowns and relaxed phases. Next, we will compare the results of applying this control technique using the discrete (DM) and continuous (CM) models. We shall explore the effects of setting $\texttt{maxCases} = 50$ cases.

Figure 7 illustrates the result of applying a CIL strategy in Buenos Aires City. The CM is fitted to detected cases from February 23rd to August 15th. We considered cases aggregated by their Symptoms Onset Date (SOD), as the usage of cases grouped according to Case Report Date (CRD) may lead to inconsistent estimates. Using SOD requires cutting out the time series over the last ten days, as this is the typical time it takes to collect all the notifications of cases that share the same SOD in the past. In the CM we considered a population proportional to the number of symptomatic cases using the detection factor $df = 10\%$

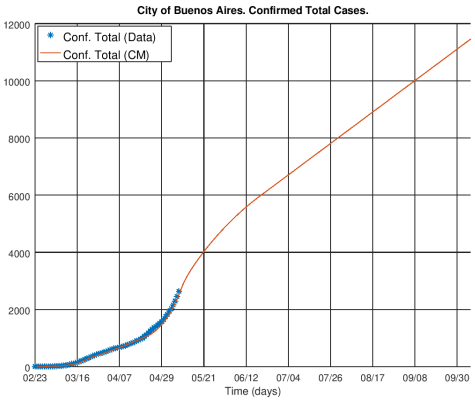
Figure 7 illustrates the result of applying the CIL control to Buenos Aires City. Next, in Figure 9 we repeat the experiment but using the DM instead. Comparing both figures, we can notice that for the DM the trajectory of daily cases moves away from the threshold swhitching between phases with a period of about one month. Conversely, the corresponding trajectory for the CM remains close to its threshold switching unrealistically fast between phases. This difference in closed-loop behaviour can be explained by noting that in the DM there is a delay between the instant when the daily cases cross the threshold and the moment this crossing is detected. For a fair comparison, Figure 8 illustrates the result of applying the CIL control to Buenos Aires City but incorporating a 11-days delay between threshold crossing and control action. It's worth noticing that the results in Figure 8 and 9 have a high degree of similarity.

The advantage of the DM here is that it does not require any modification to reflect the actual closed loop behavior, while the CM requires to add an additional delay that transforms the model into a set of *Delay Differential Equations*. In addition, the DM exhibits an interesting phenomena after each lockdown: some days after the sudden change in R_0 the number of cases has a *rebound* (notice it around May 15 in Figure 9). This phenomena can be explained by taking into account that the number of infectious people continues growing for some days after the lockdown (the reduction in R_0 only reduces the new exposed individuals), and that growth is reflected back in the number of new exposed, which depends on the number of infectious people.

It is worth mentioning that the DM parameters of this experiment were fitted without considering imported cases, as they were not available in the dataset for the city of Buenos Aires. Instead, we considered as imported all the cases detected

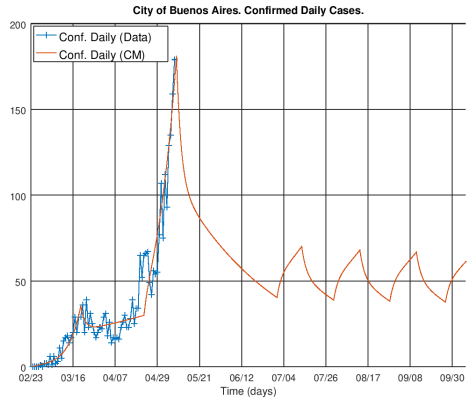


(a) Daily Cases

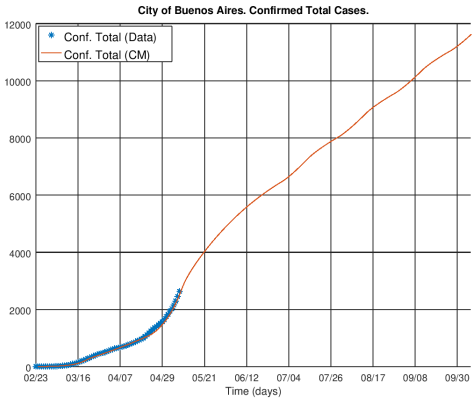


(b) Total Cases

Figure 7: Controlled Intermittent Lockdowns (CIL) for Buenos Aires City using CM.



(a) Daily Cases



(b) Total Cases

Figure 8: Controlled Intermittent Lockdowns (CIL) for Buenos Aires City using CM with an 11-day measuring delay.

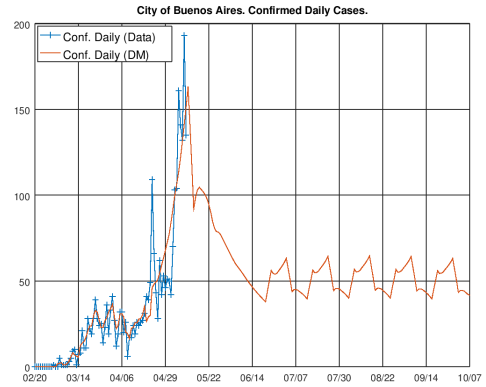
prior to certain date (April 20th.) so they acted as initial conditions for the state variables $N_E(t)$, $E(t)$, and $I(t)$ starting on that date. The parameter adjustment procedure detected that R_0 was null before April 9 (11 days before the initial date), and obtained valid values for R_0 after that. The actual fraction of imported cases detected in Argentina after April 20th. was negligible, so we considered that $U(t)$ was null from then on without introducing further relevant errors.

The same procedure was used in different experiments with datasets of different cities and regions of Argentina using only data from detected daily cases without distinguishing between imported and local cases. The idea can be straightforwardly applied to any city, region or country provided that there is a daily record of detected cases and deaths.

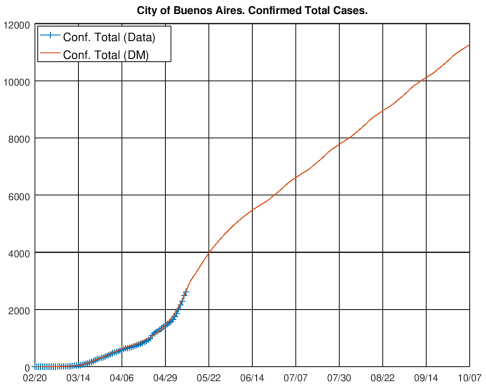
Discrete and Continuous Model Comparison

In order to cross-check the model with a classic continuous time SEIR model, we adjusted the parameters in order to fit both models using the diagnosis dates for the DM and the symptoms onset dates for the CM. In both cases, we used the official data series at national level.

Simulation results and official data series are compared in Figure 10, with both approaches achieving a high degree of similarity. In both cases, we measured the relative mean square error between the real and the simulated trajectories for the total number of cases, obtaining values of 1.04% and 1.08% for the CM and DM, respectively. It is worth mentioning that the diagnosis date data series has a much higher dispersion than the symptom onset data series (it can be easily appreciated in Figure 10a), so the fact that

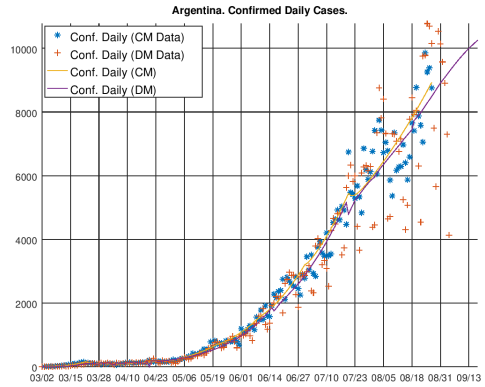


(a) Daily Cases

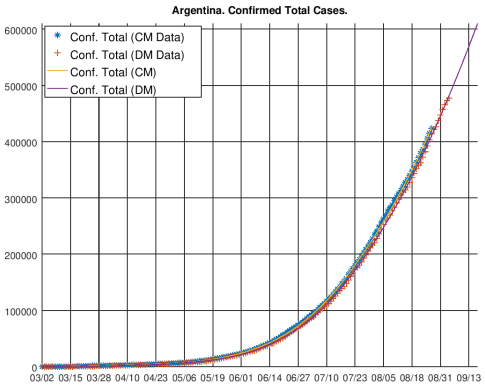


(b) Total Cases

Figure 9: Controlled Intermittent Lockdowns (CIL) for Buenos Aires City using the DM.



(a) Daily Cases



(b) Total Cases

Figure 10: Model fitting for Argentina using DM and CM.

both approaches have a very similar error exhibits in fact a further advantage of the discrete time model.

Conclusions

A simple discrete model of COVID-19 propagation based on first principles was presented, which does not require complex hypotheses, developments or interpretations. Its parameters are transparent and explicitly represent the essential characteristics of the underlying phenomena. In addition, the model distinguishes imported cases, which eliminates the overestimation of the infection rate at the early stages of the pandemic outbreak.

Another salient aspect of the methodology is that there are no assumed values for the parameters; they are instead obtained by fitting procedures with respect to the actual measurements. That is, using only the data reported on the number of cases and deaths, the model allows for inferring biological, social, and even administrative dynamics (by means of the parameters) that influence the evolution and spread of the virus.

Having explicit and directly interpretable parameters, it is possible to study scenarios through their modification that are otherwise cumbersome to represent in classic models of continuous time. For instance, the effect of a strategy of tracing and isolating close contacts of confirmed cases (during the N days prior to detection) can be captured simply by reducing the parameter τ_R.

Likewise, in terms of control strategies (non-pharmacological interventions), having a model with explicit delays between action and measurement is essential to be able to study closed-loop performance.

The current work can be extended in several

directions. Firstly, multi population models based on age and/or geographical features can be easily represented with this methodology, using different delay parameters for each population. We are also exploring the possibility of adding available statistical information. For instance, instead of considering a single delay between detection and death, we might use a weighted sum of delayed signals.

A simple variant of this model can be also used for studying vaccination policies where the explicit delays can be used to represent the delayed effectiveness of a vaccine. Similarly, if reinfection is likely after a certain period of time, a SEIRDS variant of this model can be straightforwardly proposed.

References

1. W. H. Organization, "Novel Coronavirus (2019-nCoV): situation report, 11," Technical documents.
2. National Ministry of Health, "Daily Report (March 5th)," 2020, https://www.argentina.gob.ar/sites/default/files/5-03-2020-nuevo-coronavirus-covid-19-reporte-diario_1.pdf, accessed 15th September. [In Spanish].
3. H. W. Hethcote, "Mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.
4. E. Vynnycky and R. White, *An introduction to infectious disease modelling*. OUP oxford, 2010.
5. F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*. Springer, 2012, vol. 2.
6. C. J. Romero, A. Tisnés, and S. Linares, "Agent-based COVID-19 simulation model. Application to the argentine case," Tech. Rep., 2020. [Online]. Available: https://ri.conicet.gov.ar/handle/11336/107719
7. H. Ahumada, S. Espina, and F. Navajas, "COVID-19 with Uncertain Phases: Estimation Issues with An Illustration for Argentina," *SSRN Electronic Journal*, jun 2020. [Online]. Available: https://papers.ssrn.com/abstract=3633500
8. R. A. Borracci and N. D. Giglio, "Forecasting the effect of social distancing on covid-19 autumn-winter outbreak in the metropolitan area of buenos aires," *Medicina*, vol. 80, pp. 7–15, 2020.
9. F. Pazos and F. E. Felicioni, "A control approach to the Covid-19 disease using a SEIHRD dynamical model," *medRxiv*, p. 2020.05.27.20115295, may 2020. [Online]. Available: https://doi.org/10.1101/2020.05.27.20115295
10. E. Tagliazucchi, P. Balenzuela, M. Travizano, G. B. Mindlin, and P. D. Mininni, "Lessons from being challenged by COVID-19," *Chaos, Solitons and Fractals*, vol. 137, p. 109923, aug 2020.
11. A. Wiratsudakul, P. Suparit, and C. Modchang, "Dynamics of Zika virus outbreaks: An overview of mathematical modeling approaches," *PeerJ*, vol. 2018, no. 3, 2018. [Online]. Available: 10.7717/peerj.4526
12. J. T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P. M. de Salazar, B. J. Cowling, M. Lipsitch, and G. M. Leung, "Estimating clinical severity of covid-19 from the transmission dynamics in wuhan, china," *Nature Medicine*, pp. 1–5, 2020.
13. W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li, "Why is it difficult to accurately predict the covid-19 epidemic?" *Infectious Disease Modelling*, 2020.
14. Y. Fang, Y. Nie, and M. Penny, "Transmission dynamics of the covid-19 outbreak and effectiveness of government interventions: A data-driven analysis," *Journal of medical virology*, 2020.
15. C. Anastassopoulou, L. Russo, A. Tsakris, and C. Siettos, "Data-based analysis, modelling and forecasting of the covid-19 outbreak," *PloS one*, vol. 15, no. 3, p. e0230405, 2020.
16. J. Dehning, J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, and V. Priesemann, "Inferring change points in the spread of covid-19 reveals the effectiveness of interventions," *Science*, 2020.
17. R. Castro, E. Kofman, and F. E. Cellier, "Quantization-based integration methods for delay-differential equations," *Simulation Modelling Practice and Theory*, vol. 19, no. 1, pp. 314–336, 2011.
18. W. O. Kermack and A. G. 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, pp. 700–721, 1927.
19. L. J. Allen, "Some discrete-time si, sir, and sis epidemic models," *Mathematical biosciences*, vol. 124, no. 1, pp. 83–105, 1994.
20. Z. Liu, P. Magal, O. Seydi, and G. Webb, "A covid-19 epidemic model with latency period," *Infectious Disease Modelling*, 2020.
21. S. Flaxman, , S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, and S. Bhatt, "Estimating the effects of non-pharmaceutical interventions on COVID-19 in europe," *Nature*, vol. 584, no. 7820,

pp. 257–261, Jun. 2020. [Online]. Available: https://doi.org/10.1038/s41586-020-2405-7

22. N. G. Davies, A. J. Kucharski, R. M. Eggo, A. Gimma, W. J. Edmunds, T. Jombart, K. O'Reilly, A. Endo, J. Hellewell, E. S. Nightingale, B. J. Quilty, C. I. Jarvis, T. W. Russell, P. Klepac, N. I. Bosse, S. Funk, S. Abbott, G. F. Medley, H. Gibbs, C. A. Pearson, S. Flasche, M. Jit, S. Clifford, K. Prem, C. Diamond, J. Emery, A. K. Deol, S. R. Procter, K. van Zandvoort, Y. F. Sun, J. D. Munday, A. Rosello, M. Auzenbergs, G. Knight, R. M. Houben, and Y. Liu, "Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study," *The Lancet Public Health*, vol. 5, no. 7, pp. e375–e385, 2020.