RESEARCH

Inverse problem for parameters identification in a modified SIRD epidemic model using ensemble neural networks

Marian Petrica[1,2*] and Ionel Popescu[1,3]

*Correspondence:
marianpetrica11@gmail.com

[1] Faculty of Mathematics
and Computer Science,
University of Bucharest,
Bucharest, Romania
[2] Gheorghe Mihoc - Caius
Iacob Institute of Mathematical
Statistics and Applied
Mathematics of the Romanian
Academy, Bucharest, Romania
[3] Institute of Mathematics
of the Romanian Academy,
Bucharest, Romania

Abstract

In this paper, we propose a parameter identification methodology of the SIRD model, an extension of the classical SIR model, that considers the deceased as a separate category. In addition, our model includes one parameter which is the ratio between the real total number of infected and the number of infected that were documented in the official statistics. Due to many factors, like governmental decisions, several variants circulating, opening and closing of schools, the typical assumption that the parameters of the model stay constant for long periods of time is not realistic. Thus our objective is to create a method which works for short periods of time. In this scope, we approach the estimation relying on the previous 7 days of data and then use the identified parameters to make predictions. To perform the estimation of the parameters we propose the average of an ensemble of neural networks. Each neural network is constructed based on a database built by solving the SIRD for 7 days, with random parameters. In this way, the networks learn the parameters from the solution of the SIRD model. Lastly we use the ensemble to get estimates of the parameters from the real data of Covid19 in Romania and then we illustrate the predictions for different periods of time, from 10 up to 45 days, for the number of deaths. The main goal was to apply this approach on the analysis of COVID-19 evolution in Romania, but this was also exemplified on other countries like Hungary, Czech Republic and Poland with similar results. The results are backed by a theorem which guarantees that we can recover the parameters of the model from the reported data. We believe this methodology can be used as a general tool for dealing with short term predictions of infectious diseases or in other compartmental models.

Introduction

Infectious disease pandemics have had a major impact on the evolution of mankind and have played a critical role in the course of history. Over the ages, pandemics made countless victims, decimating entire nations and civilizations. As medicine and technology have made remarkable progress in the last century, the means of fighting pandemics have become significantly more efficient. A reality nowadays is that globalization, the

development of commerce, and the ease to travel all over the world facilitate the transmission mechanism of a new disease much more than it did in the past.

In 2019, Covid19, a virus from the coronavirus family appeared and spread around the world very quickly. This changed dramatically our world as we know it.

Covid pandemic

In 2019, COVID19, a virus from the coronavirus family appeared and spread around the world very quickly. The first cases of COVID19 were reported in China in December 2019 and, within less than 3 months, the outbreak became a global pandemic, spreading across almost all countries all over the world. In the period that followed, we witnessed many changes in political decisions including lockdowns, school closures and many other restrictions which were aimed to control the spreading of the virus.

The main ideas of this paper

Fighting COVID-19 was at first driven by quarantine and other restriction measures. These were imposed because the mechanisms of infection were not well understood and the hospitals were overwhelmed. Later on, various degrees of restrictions were imposed in order to control the spread and, at the same time, let the economy recover. Thus many political decisions affected for better or for worse the transmission of the virus.

Mathematical modeling is by now one of the scientific pillars on which we build our understanding of the world. It is a valuable tool that can be used in the assessment, prediction and control of infectious diseases, as it is the COVID-19 pandemic. There is a growing body of mathematical models used at the moment for the spread of virus which are related to our approach in this note. From the rich literature around this topic we sample the papers [1–8].

One popular choice used to mathematically model the epidemics is the *SIR* model, appeared in [9–14]. This is a compartmental model where an individual can be in one of the following 3 states, at any given time: susceptible (*S*), infected (*I*) or recovered (in some sources removed) (*R*). An extension of the *SIR* model is the *SIRD* model which considers the category of deaths as a separate compartment.

The epidemic *SIR*/*SIRD* models are non-linear and the identification of the parameters is challenging, compared to the linear models where analytical approaches exist (Godfrey and DiStefano, [15]). For the case of non-linear models there are methods for identifying the parameters, that are assumed to be constant, by using Taylor series (Gun et al, [16]; Pohjanpalo [17]) or differential algebra (Audoly et al, [18]; Eisenberg [19]). Differential algebra can also be an useful tool in the case of time dependent parameters (Hadeler, [20], Mummert, [21]). Marinov et al, [22], also proposed a numerical scheme for coefficient identification in SIR epidemic models using the Euler-Lagrange equations. In the case of the stochastic models, parameter estimation is a type of statistical inference, procedures as least square estimation (Banks et al. [23]) or maximum likelihood estimation (Julier, [24]) being the most used techniques.

Many studies have been done on the evolution of the pandemics around the world. For the case of Covid19 pandemic, the assessment and prediction of the spread as well as the identification of the parameters were rigorously and in detail studied in the papers [25–32].

The purpose of this work is to detail a predictive model, define a methodology of identifying its parameters and accurately assess the transmission dynamic of COVID-19, with potential applicability for other infectious models. At the same time we analyse the evolution of the pandemic in Romania and we validate the method, in a separate Appendix 2, for the case of Hungary, Czech Republic and Poland. The application of predictive models in the study of pandemics dynamics has been exhaustively addressed in the following studies [32–36].

The starting point of our study is the idea that, during the pandemic, it is hard to measure the number of infected, susceptible or recovered persons over time. Due to the large size population, it is very difficult to accurately know the number of infected or the number of recovered people. As Covid19 showed, the governmental institutions and the hospitals became rapidly overwhelmed, while many people became infected and treated themselves at home, without it being recorded in the official statistics.

The dynamic of the pandemic was driven by many factors, for instance lockdown, opening or closure of the restaurants, elections, opening or closing of schools, vacationing and other measures taken by authorities in order to mitigate the pandemic effects. Thus any reasonable parametric model is negatively impacted by the fact that the coefficients are not constant in time. This implies that the estimations can not be done on long term. In a previous paper [29] we considered a regime switch which was geared toward adapting the parameters to the political decisions of lockdown or relaxation.

In this paper the main idea is to use a dynamic model which does not take into account long term evolution or outside assumptions about the status of the pandemic. In this approach we consider the data on a short period of time, in our case we chose to work with seven days and estimate the parameters relying primarily on the number of deceased people.

We do this in several steps. The first one is the cleaning and smoothing of the data. We fixed some anomalies in the reporting and we take a moving average of two weeks time. The second step is to exploit the model and generate data with parameters chosen at random. Based on only seven days of data, we train neural networks to learn the parameters of the model. Furthermore, we used these neural networks to estimate the parameters based on the real data. The key here is a form of ensemble learning which is interesting in itself. A single neural network does not seem to predict the parameters very well, however the average has a much better prediction power. The third and last step is to predict the evolution for a number of future days and compare with the real data. For a range of ten days, we get very close results to the real data.

We should point out that in mathematical terms, our approach is a typical inverse problem. Generically, inverse problems can be ill posed. We show that our model is actually well posed, meaning that determination of the coefficients from the observation do identify the coefficient uniquely. This is the backbone of our analysis and estimation.

The paper is organized as follows. In The SIR and SIRD models section we briefly discuss the SIR model and we introduce the SIRD model. Here we state Theorem 1 whose proof is in Appendix 1. Also here we describe the guiding idea of our approach. In The data and the neural networks section we discuss the data and its cleaning and describe the construction of the neural networks. Then we proceed with the Discussion and conclusions section.

The SIR and SIRD models

One of the most important models that can describe infectious diseases is the SIR model. The first ones that developed SIR epidemic models were Bernoulli [37], Ross [10, 11], Kermack-McKendrick [38] and Kendall [39].

The SIR model is a mathematical model that can be used in epidemiology to analyze, at a given time for a specific population, the interactions and dependencies between the number of individuals who are susceptible to get an infectious disease, the number of people who are currently infected and those who have already been recovered or have died as a cause of the infection. This model can be used to describe diseases that can be contracted just one time, meaning that a susceptible individual gets a disease by contracting an infectious agent, which is afterward removed (death or recovery).

It is assumed that an individual can be in either one of the following three states: susceptible (S), infected (I) and removed/ recovered (R). This can be represented in the following mathematical schema:

$$\boxed{\text{Susceptible}} \xrightarrow{\beta} \boxed{\text{Infected}} \xrightarrow{\gamma} \boxed{\text{Recovered}}$$
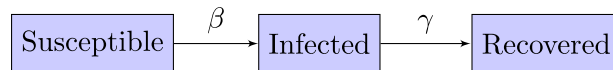
where:

- $\beta =$ infection rate
- $\gamma =$ recovery rate.

We consider N as the total population in the affected area. We assume N to be fixed, with no births or deaths by other causes, for a given period of n days. Therefore, N is the sum of the three categories previously defined: the number of susceptible people, the ones infected, and the ones removed:

$$N = \bar{S} + \bar{I} + \bar{R}.$$

Therefore, we analyze the following SIR model: at time t, we consider $\bar{S}(t)$ as the number of susceptible individuals, $\bar{I}(t)$ as the number of infected individuals, and $\bar{R}(t)$ as the number of removed/recovered individuals. The equations of the SIR model are the following:

$$\begin{cases} \frac{d\bar{S}(t)}{dt} = -\frac{\bar{\beta}\bar{S}(t)\bar{I}(t)}{N} \\ \frac{d\bar{I}(t)}{dt} = \frac{\bar{\beta}\bar{S}(t)\bar{I}(t)}{N} - \gamma\bar{I}(t) \\ \frac{d\bar{R}(t)}{dt} = \gamma\bar{I}(t) \end{cases} \tag{1}$$

where:

- $\frac{d\bar{S}(t)}{dt}$ is the rate of change of the number of individuals susceptible to the infection over time;
- $\frac{d\bar{I}(t)}{dt}$ is the rate of change of the number of individuals infected over time;
- $\frac{d\bar{R}(t)}{dt}$ is the rate of change of the number of individuals recovered over time.

Because there is no canonical choice of N, we will transform the system (1) by dividing it by N and considering $S(t) = \bar{S}(t)/N$, $I(t) = \bar{I}(t)/N$ and $\hat{R}(t) = \bar{R}(t)/N$. It is customary to choose $N = 10^6$ for convenience but this is just an arbitrary choice. For instance, analysis on smaller communities or cities involves less than 10^6, however, 10^6 is a common choice because countries number their populations in multiples of 10^6. With these notations, we translate (1) into

$$\begin{cases} \frac{dS(t)}{dt} = -\beta S(t)I(t) \\ \frac{dI(t)}{dt} = -\beta S(t)I(t) - \gamma I(t) \\ \frac{d\hat{R}(t)}{dt} = \gamma I(t) \end{cases} \tag{2}$$

where $\beta = \bar{\beta}/N$ and γ is the same as in (1). Observe that we actually have that $S(t) + I(t) + \hat{R}(t) = S_0 + I_0 + \hat{R}_0 = 1$ for all $t \geq 0$.

We use a slight change in the SIR model which accounts for the number of deceased people separately. The main idea here being that the number of deceased people might be more reliable than other data, as for instance the number of infected. In the plain vanilla SIR model the recovered and deceased are combined into the single category of recovered. The idea of the SIRD model is to use the provided data of deceased people in a significant way.

To this aim, we will work with four variables changing in time, namely $S(t)$, $I(t)$, $R(t)$ and $D(t)$ where $R(t)$ is the proportion of recovered and alive people while the $D(t)$ is the proportion of deceased people. We set the SIRD model as an interaction driven by the system of differential equations:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta S(t)I(t) \\ \frac{dI(t)}{dt} = \beta S(t)I(t) - (\gamma_1 + \gamma_2)I \\ \frac{dR(t)}{dt} = \gamma_1 I(t) \\ \frac{dD(t)}{dt} = \gamma_2 I(t). \end{cases} \tag{3}$$

Notice that in this setup the recovered population bifurcates into recovered ones, accounted by R and the dead ones accounted by D. We also point out that by taking sum of the two factors $\hat{R}(t) = R(t) + D(t)$ above we fall into the classical SIR model. The reason of accounting for $D(t)$ separately is that the data reports the number of deaths separately and the model above allows to fit the parameters using the data.

Even if the above system is satisfactory to a certain degree, we would like to point out that in practice, the number of infected as well as the number of recovered is not really known. The data we have at our disposal reports the number of infected and recovered which are *documented*. The real number of infected is not really observed. Therefore we will adjust the model by introducing another parameter α which measures the proportion of *observed* number of infected and recovered. Therefore, we denote

$$\tilde{I}(t) = \alpha I(t) \text{ and } \tilde{R}(t) = \alpha R(t).$$

In terms of these new quantities, the SIRD model becomes now

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta}{\alpha}S(t)\tilde{I}(t) \\ \frac{d\tilde{I}(t)}{dt} = \beta S\tilde{I}(t) - (\gamma_1 + \gamma_2)\tilde{I}(t) \\ \frac{d\tilde{R}(t)}{dt} = \gamma_1\tilde{I}(t) \\ \frac{d\tilde{D}(t)}{dt} = \frac{\gamma_2}{\alpha}\tilde{I}(t). \end{cases} \tag{4}$$

We are going to estimate the parameters $\alpha, \beta, \gamma_1, \gamma_2$ from data based on certain number of days. The advantage of getting an estimate on α is that we can in reality predict the real number of infected people and also the number of recovered people. In our adjusted model we have

$$S(t) + \frac{1}{\alpha}(\tilde{I}(t) + \tilde{R}(t)) + D(t) = 1 \tag{5}$$

for all times $t \geq 0$. To see this, we start by noticing that summing up all the equations in (3), we get that the derivative of $S(t) + I(t) + R(t) + D(t)$ is constant in time. Since $S_0 + I_0 + R_0 + D_0 = 1$, as they represent the proportion of the entire population, we get that the sum $S(t) + I(t) + R(t) + D(t) = 1$ and using the definition of α and $\tilde{I}, \tilde{R}$ we arrive at (5).

We present next the main mathematical result, with a proof in the Appendix 1. This guarantees that the problem is well posed and we can recover the main parameters of the model.

Theorem 1   *Given the observations of $\tilde{I}(0) = \tilde{I}_0, \tilde{R}(0) = \tilde{R}_0$ and $D(t)$ for $t = 0, 1, 2, 3, 4$, we can uniquely determine the parameters $\alpha, \beta, \gamma_1, \gamma_2$.*

This result is fundamental for our approach. It shows that given a number of daily observations, at least 5 days, we can uniquely determine the parameters of the model. In practice, given any day k of the pandemic, we will use the previous data on a number of days to determine the parameters.

The guiding idea: We can imagine the map from the daily data to the parameters as a function

$$\Phi(data_{gen}) = (\beta, \gamma_1, \gamma_2, \alpha) \text{ where } data_{gen} = (\tilde{I}_0, \tilde{R}_0, D(0), D(1), \dots, D(4))$$

generated by (4). The theorem ensures that this function is well defined on the set

$$Data_{gen} = \{(\tilde{I}_0, \tilde{R}_0, D(0), D(1), \dots, D(4)) \text{ solution of (4) for some } (\beta, \gamma_1, \gamma_2, \alpha)\}.$$

At this point, given an arbitrary data point $data = (\tilde{I}_0, \tilde{R}_0, D(0), D(1), \dots, D(4))$ it is difficult to check that this belongs to $Data_{gen}$. Thus our goal is to find the best approximation of the *data* with a point in $Data_{gen}$.

In general this is achieved using projection methods, for instance non-linear least square, as it is done in [5]. In our case we consider an extension problem, rather than the projection method, through the method of neural networks trained on simulated data, using the SIRD model. These are functions which construct approximations of Φ defined on the set $Data_{gen}$, that can naturally extend to the whole space, thus can be interpreted as (approximate) extrapolations of Φ to the whole space. As we will

describe below, one single neural network did not work very well in our numerical experimentation, while an ensemble of neural networks achieve a better performance.

Numerical simulations show that we get more robust results by considering a larger number of days for the deceased. We noticed that 7 days instead of 5 give more robust results. Also we exploit 2 days of data for infected and recovered to strengthen the robustness.

One of our main findings is that the prediction works very well as it is shown in the next pictures. For each day k we predict using our method the next 10 days and take the average for each category. These are plotted along the averages of the real data for 10 days starting from day k.

We can observe from Figs. 1, 2, and 3 that the prediction power is excellent for 10 forward days. By computing the MAE (Mean Absolute Error) for different timeframe predictions and reality data we notice that this observation is validated:

| Prediction	Case	MAE
10 days prediction	Deaths	66.77549
	Infected	1635.88403
	Recovered	752.07601
30 days prediction	Deaths	212.02360
	Infected	13731.60866
	Recovered	4733.52849
45 days prediction	Deaths	531.00102
	Infected	32418.20868
	Recovered	14693.97709

We detail and discuss more on these predictions in the next section.
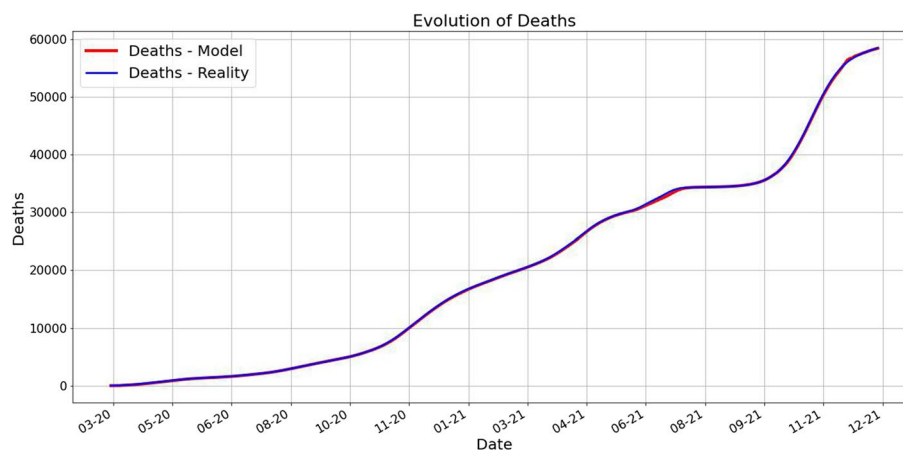
Fig. 1 Averages of deaths using 10 future days for predicted and real data. The red value at time k is computed as the average of future 10 days predictions, while the blue value at time k is the average of the 10 future days values
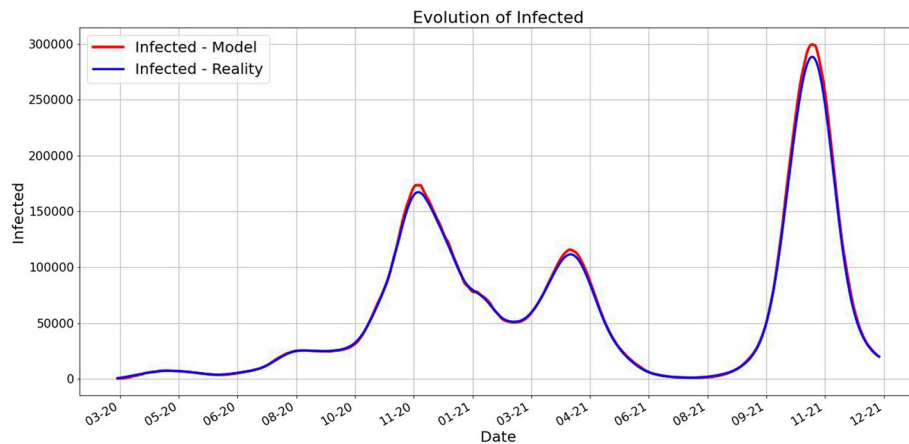
Fig. 2 Averages of infected using 10 future days for predicted and real data. The red value at time *k* is computed as the average of future 10 days predictions, while the blue value at time *k* is the average of the 10 future days values
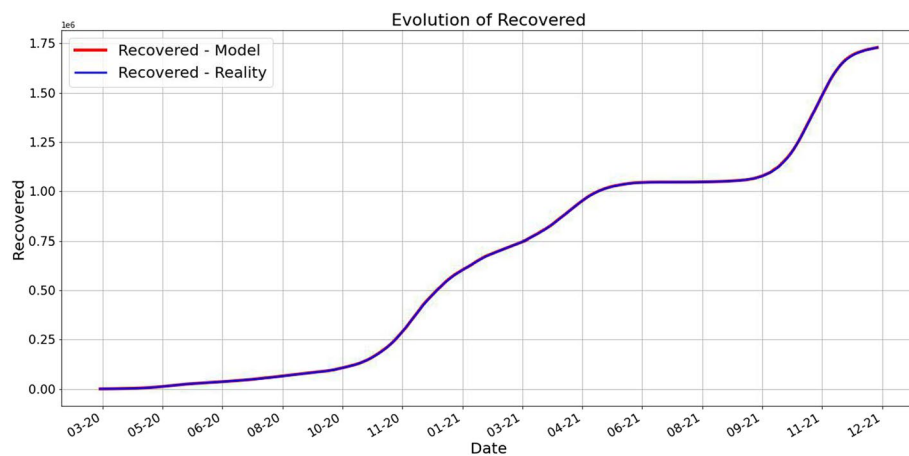
Fig. 3 Averages of recovered using 10 future days for predicted and real data. The red value at time *k* is computed as the average of future 10 days predictions, while the blue value at time *k* is the average of the 10 future days values

The data and the neural networks

In this section we present our strategy for the data cleaning and the estimation of the parameters $\beta, \gamma_1, \gamma_2, \alpha$ of the SIRD model.

We clean the raw data which has several anomalies and we use a regularization by averaging.

The basic strategy is the following. Based on the assumed SIRD model we generate data for 7 days and then train a neural network which learns the parameters from the data for this 7 days time interval.

Then, using the real data and the neural network we find the parameters in a dynamical way for any 7 consecutive days. Given these 7 days we can predict based on the model what is going to happen on the next few days.

In a real world the parameters do not stay constant, they change dynamically and we would like to catch part of this behavior.

Data

We took the data from https://datelazi.ro which keeps a record of all the data during the pandemic in Romania.

During the pandemic, the reported numbers and the methodology regarding the reporting changed several times causing delays or bad reporting. In October 2020, the definition of a recovered person changed thus causing a data anomaly in the reported number of recovered. Particularly, we can see a spike of 44000 new cases from one day to another, equivalent to the cumulative number of cases until then. To alleviate this anomaly, we distribute the extra number of cases proportionally to the previous days.

There are also periods of time in which the number of recovered people is actually 0 for almost three weeks.

We further analyse the data of the 102 weeks taken into assessment by day of the week and we compute the average of the reported number of infected.

Day of the week	Average number of infected (reported)
Monday	2461
Tuesday	4479
Wednesday	4525
Thursday	4453
Friday	4269
Saturday	4041
Sunday	2909

We notice a significant difference between the average numbers reported on weekdays and weekends and we perform an *One Way Anova Test* to validate this observation.

| Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	4.32E+08	6	72029925	2.286333	0.034118	2.111386
Within Groups	2.23E+10	707	31504561			
Total	2.27E+10	713				

We found a statistically significant difference between the averages of the reported cases according to the day of the week ($p < 0.05$). However, by law, the methodology of reporting from the health centers allows reporting cases within an interval of two weeks. In order to mitigate the above deficiencies, we replaced the data at time t with the average of the data during the previous two weeks preceding time t. We present in Fig. 4 the data before and after the cleaning.

The neural networks

To deal with the estimation of the parameters, we follow two main steps.

The first one is the generation of data.

We take the range of β in the interval $J_\beta = (0, 1)$, for γ_1 we consider the range in $J_{\gamma_1} = (0, 1)$, for γ_2 we consider the interval $J_{\gamma_2} = (0, 0.01)$ and for α we consider the interval $J_\alpha = (0.01, 1)$. Next we split each of these intervals into 7 sub-intervals of the same size which we will index accordingly as
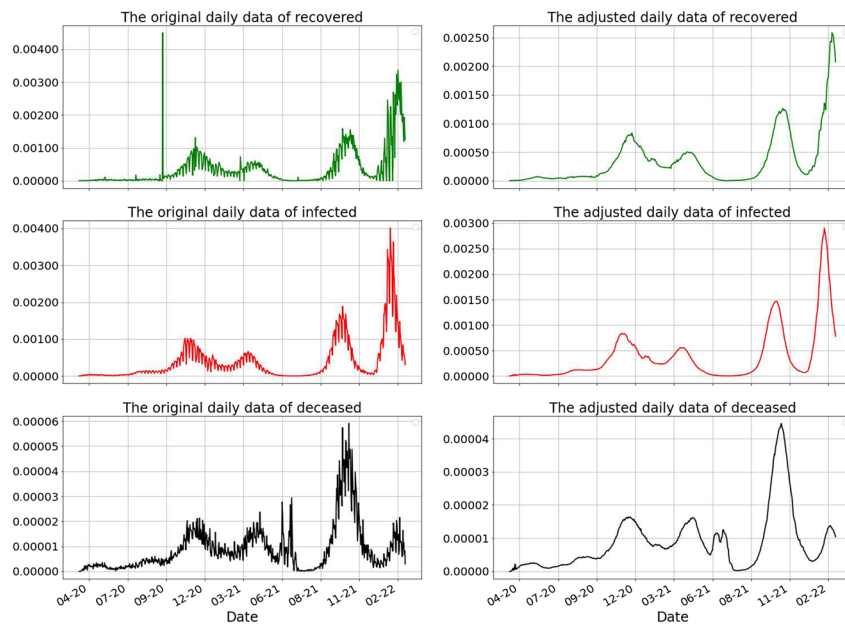
Fig. 4 Data adjustments according to the methodology of reporting the Covid19 numbers in Romania. The rows describe the data (from top to bottom) of recovered, infected and dead. The left column represents the raw data and the right column represents the adjusted data as we described above. Notice the scale and the spike in the first picture which is adjusted as we pointed out. The data we work with is scaled by 10, 000, 000

$$J_{\beta,i} = (i/7, (i+1)/7)$$

$$J_{\gamma_1,i} = (i/7, (i+1)/7))$$

$$J_{\gamma_2,i} = (0.01 * i/7, 0.01 * (i+1)/7)$$

$$J_{\alpha,i} = (0.01 + 0.99 * i/7, 0.01 + 0.99 * (i+1)/7)$$

for $i \in \{0, 1, \ldots, 6\}$. The splitting is motivated by the fact that we want to have good representation of the parameters and at the same time we want to avoid concentration of the parameters in one single region. We tried previously to use a simple uniform choice for each parameter in the whole interval, but we run into the problem of misrepresentation of small values of the parameters. It seems that this phenomena is due to some form of concentration of measure which is alleviated by using this splitting method. With this strategy we also avoid the overfitting problem of the neural networks. Therefore we obtain 7 sub-intervals for each of the 4 parameters which means that we get $7^4 = 2401$ combinations of sub-intervals.

Next, to generate the data we apply the following procedure:

1. Create the data set Δ to store the values obtained in the next steps
2. For $i_1 \in \{0, 1, \ldots, 6\}, i_2 \in \{0, 1, \ldots, 6\}, i_3 \in \{0, 1, \ldots, 6\}, i_4 \in \{0, 1, \ldots, 6\}$:

   (a) For $j \in \{1, \ldots, 10000\}$ pick at random

a. $\beta \in J_{\beta,i_1}$,
b. $\gamma_1 \in J_{\gamma_1,i_2}$,
c. $\gamma_2 \in J_{\gamma_2,i_3}$
d. $\alpha \in J_{\alpha,i_4}$,
e. I_0 in the interval (0, 0.2),
f. R_0 in (0, 0.6)
g. D_0 in (0, 0.007)

  (b) Solve the system (4) with all the parameters from step 1 and 2 for the time interval [0, 7] and add the row of $(I_0, I_1, \ldots, I_7, R_0, R_1, \ldots, R_7, D_0, D_1, \ldots, D_7)$ to Δ.

For each $i \in \{1, 2, \ldots, 10\}$, we create a training sample B_i from the dataset Δ of size 70% chosen at random without replacement. Using B_i sample we train a neural network, NN_i, $i \in \{1, 2, \ldots, 10\}$, having as input:

$$XTrain = (I_0, I_1, I_2, R_0, R_1, R_2, D_0, D_1, D_2, D_3, D_4, D_5, D_6, D_7)$$

and our parameters

$$YTrain = (\beta, \gamma_1, \gamma_2, \alpha)$$

as output.

According to our Theorem 1 we know that we can recover the parameters $\beta, \gamma_1, \gamma_2, \alpha$ only from $I_0, R_0, D_0, D_1, D_2, D_3, D_4$, from our dataset. However, we choose to use more data because the estimates are more robust. This choice can also be interpreted as a regularization which decreases the training/test loss.

The next step is the construction of the neural networks. We performed various tests, with different type of architectures, ones with multiple hidden layers and large number of neurons, as well as some architectures with 1-2 hidden layers and small number of neurons. Based on these tests we draw the conclusion that the large ones are expensive to train while the small ones do not perform very well. We decided to mitigate these disadvantages by choosing the below architecture, which is a medium one when it comes to its complexity. It helps us achieve great results at good performance with a reasonable consumption of resources. The training was performed on an Intel(R) Core(TM) i9-10885H CPU @ 2.40GHz x 16. The architecture of the neural network we use is of the following form:

- Layers:

    1. Dense 64, activation function 'ReLU', input dimentsion=14
    2. Dense 128, activation function 'ReLU'
    3. Dense 256, activation function 'ReLU'
    4. Dense 512, activation function 'ReLU'

- One output for each parameter, $(\beta, \gamma_1, \gamma_2, \alpha)$.
- Loss: Mean Absolute Error
- Optimizer: Adam (Kingma and Ba, [40])

The size of the training, respectively test data split for NN_i is 80%, respectively 20% from the sample B_i.

After training the neural networks, the predictions of our parameters is made by averaging the predictions from all the individual neural networks on the real data

$$(\hat{\beta}, \hat{\gamma_1}, \hat{\gamma_2}, \hat{\alpha}) = \frac{1}{10} \sum_{i=1}^{10} NN_i(data)$$

With this approach, we can achieve better performance of our model because we manage to decrease the variance, without increasing the bias. Usually, the prediction of a single neural network is sensitive to noise in the training set, while the average of many neural networks that are not correlated, is not sensitive. Bootstrap sampling is a good method of de-correlating neural networks, by training them with different training sets. If we train many neural networks on the same dataset, we will obtain strongly correlated neural networks.

In Fig. 5 we present the results of the estimated parameters $\beta, \gamma_1, \gamma_2, \alpha$.

Knowing the parameters for the model $(\beta_k, \gamma_{1,k}, \gamma_{2,k}, \alpha_k)$ at each time k and the values (I_k, R_k, D_k), from the real data, we can generate the predictions $P_{k,0}, \ldots, P_{k,10}$ using the system (4) for the time interval $[k, k+10]$. We take the average of $P_{k,0}, \ldots, P_{k,10}$ and call this P_k.

In Fig. 6 we plot for each day k the average of the real (smoothed) data for 10 days starting with k alongside with the average P_k computed above. As we already pointed out, the fit is very good.

The next images, in Fig. 7, show the prediction on the death for 30 and 45 days. In many cases the prediction is good, however there are regions in which the prediction ceases to be accurate. This highly depends on the timeframe we chose to make the predictions.
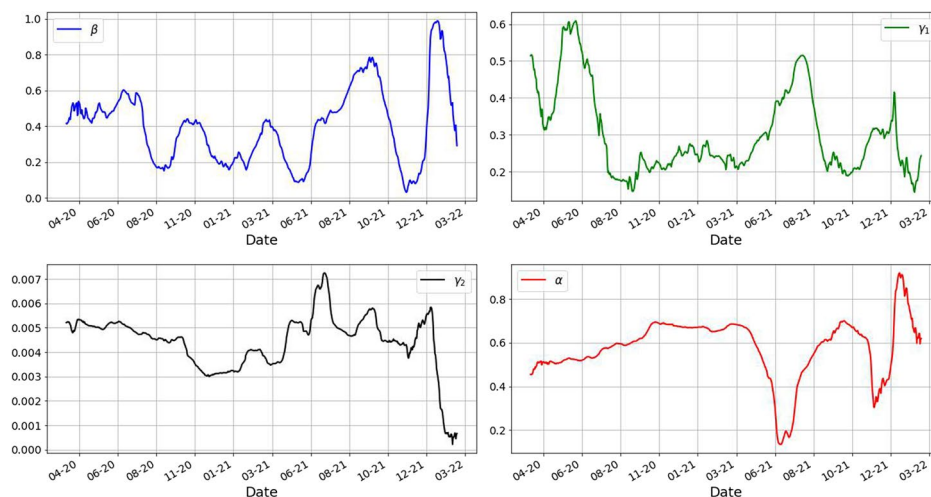

Fig. 5 Parameters of the spreading of Covid19 in Romania estimated using the above averaging of the neural networks. Notice the behavior of the parameters $\beta, \gamma_1, \gamma_2$ which tend to decrease over the period of almost two years. Interestingly we see the parameter α having large values during the Fall of 2020 and lower values during the summer of 2021. This suggests that the proportion of real infected people is between $[1/0.9, 1/0.1] = [1.11, 10]$ to the reported infected. This show that roughly only half of the infected get reported
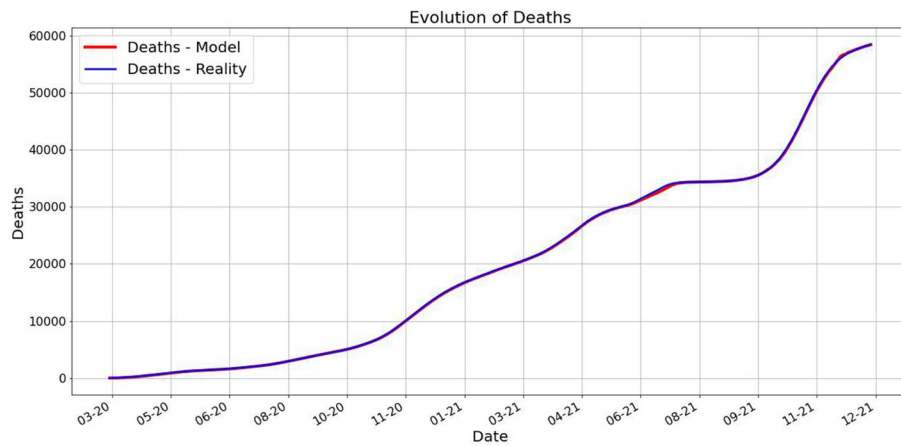
(a) Predicted and real deaths.
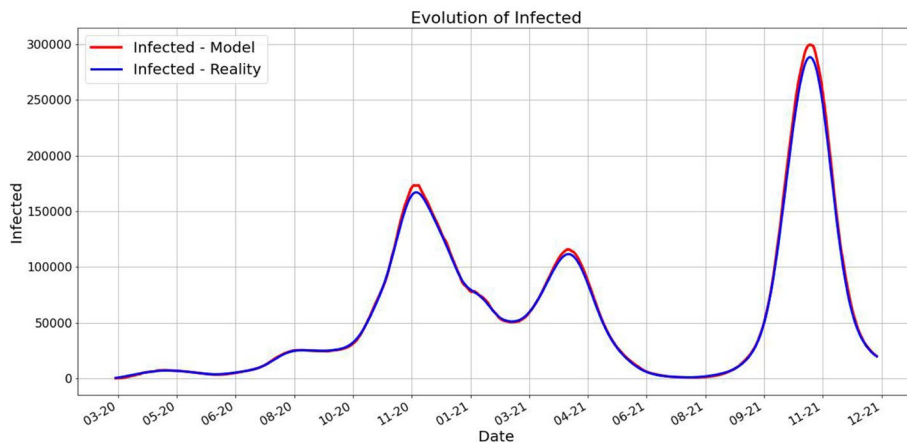
(b) Predicted and real infected.
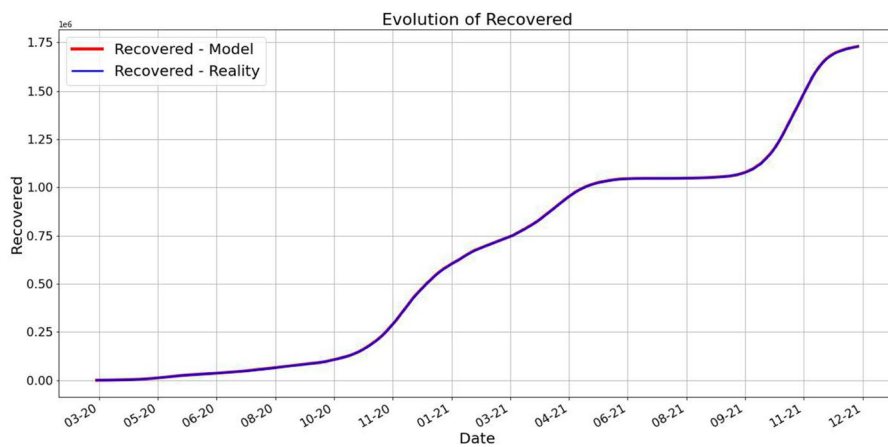
(c) Predicted and real recovered.

Fig. 6 The plots of the real data and the predicted averages for the next 10 days. As we described above, for each day k we compute the average of the real data for the next 10 days and the average of the predicted data for the next 10 days. Notice the important fact that each prediction is made in terms of the previous 7 days. The close match suggests a very good prediction power of our approach. A slight difference appears in the case of the infected number of people during the forth wave of the pandemic, namely the Fall of 2021
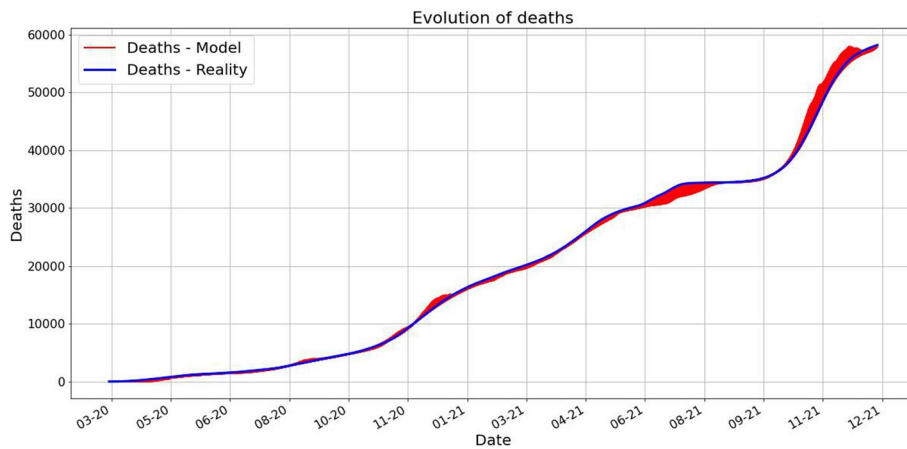
(a) For each day we predicted the deaths for 30 days.
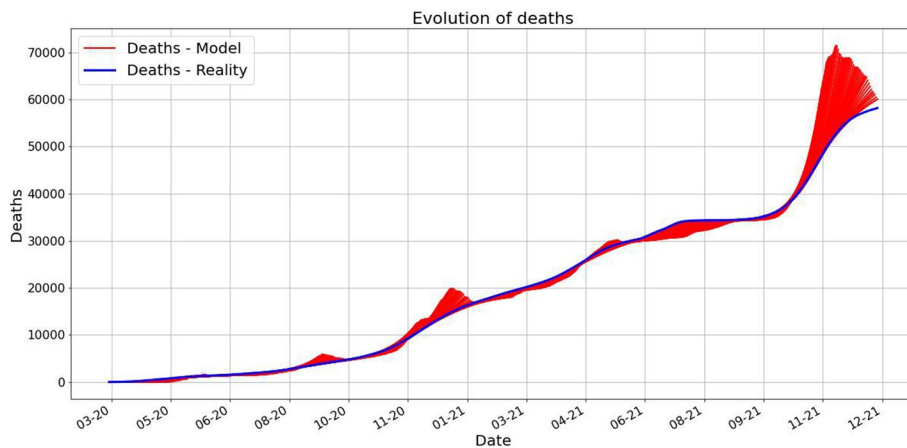
(b) For each day we predicted the deaths for 45 days.

Fig. 7 In both pictures, in blue is the real (reported) number of deaths. For each day *k*, we plotted, in red, the prediction of the deaths, starting with day *k*. The first picture shows the prediction is plotted for 30 days, while the second shows the predictions plotted for 45 days. Remark that the 30 days prediction is much better than the 45 days prediction

Next, in Figs. 8 and 9, we look in more details at the images above, to see the refined structure of the behavior. We do this for 10 days versus 30 days starting at different moments of time.

Discussion and conclusions

The dynamic of an infectious disease is highly impacted by numerous factors, including the measures imposed by governments or the attitude of population towards it, as the COVID-19 pandemic has shown. Therefore, it is very unlikely that the parameters of a model designed to asses the spread of a virus are constant over time. We have to be aware of the fact that, in the long term, the prediction is affected by all the restrictions/relaxations taken by most of the countries. We believe that this methodology has a high degree of generality to be used in many other cases of prediction,
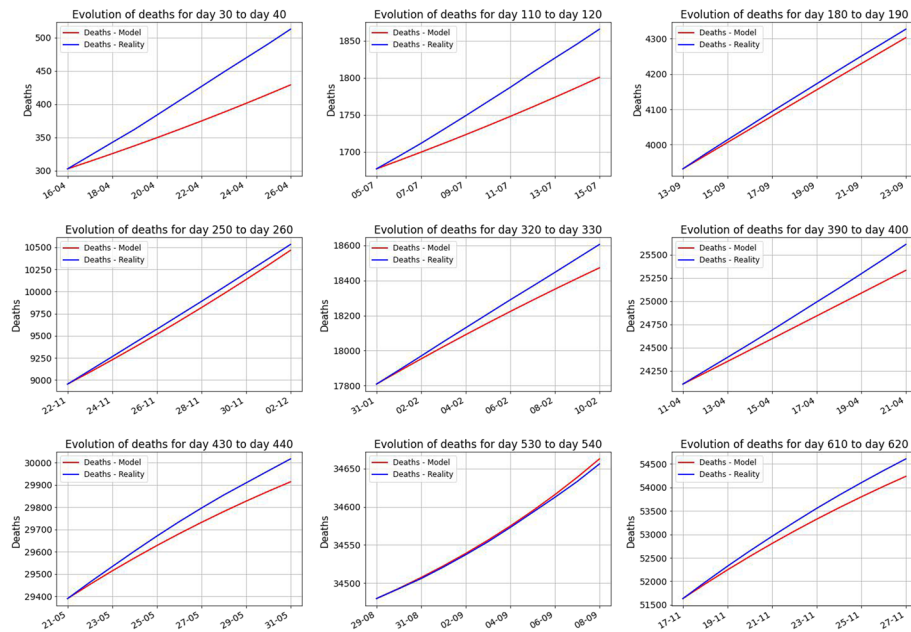
Fig. 8 The evolution of the predicted 10 days starting with the days 30, 110, 180, 250, 320, 390, 430, 530, 610
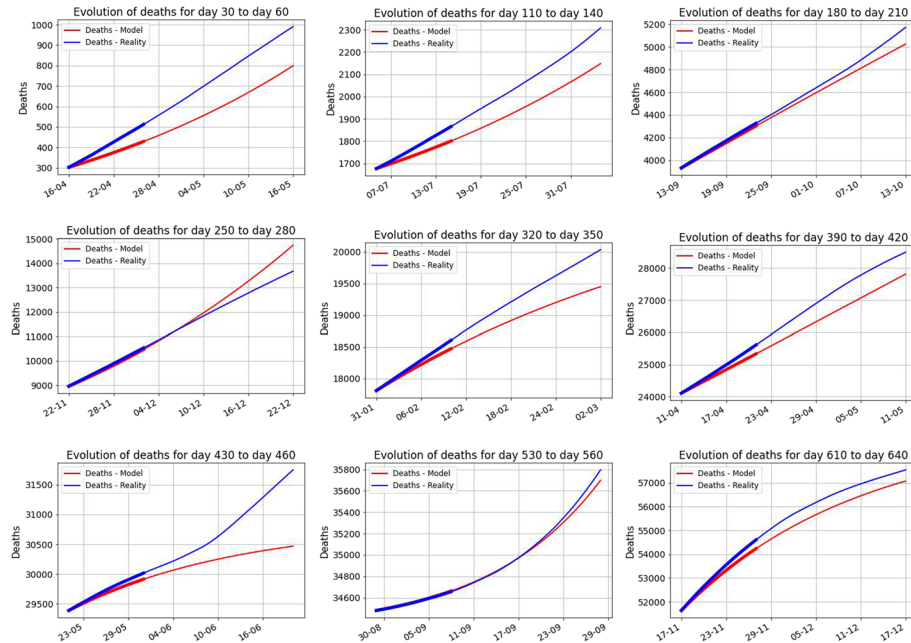
Fig. 9 The evolution of the predicted 30 days starting with the days 30, 110, 180, 250, 320, 390, 430, 530, 610. The predictions for the first 10 days were highlighted. This reinforces that the predictions are good for short periods of time and they loose the prediction power on longer periods of time. The main lesson we learn from the above figure is the fact that the parameters are not constant in time

particularly useful for those cases where the prediction depends on many other factors which change the behavior of the model.

We introduced on our model an extra parameter which accounts for the proportion of the infected and reported population versus the whole infected population. The point being that not every infected person is actually tested or reported. Thus the actual number of infected people should be higher.

We do not account for the vaccination campaign, though this does not affect our model since we are looking at the parameters on relatively short periods of time. The vaccination should in principle change the parameters, which is in fact exactly what we look for.

Regarding the limitations of the methodology presented in this paper, we can mention the followings. When it comes to the applicability of the model, one limitation could be that the number of recovered people is not included in the reporting by all of the countries, which would make the data incomplete. In the same time, one other limitation is caused by the changes that could appear in the reporting methodology of a specific country, such as the change of the definitions of infected/ recovered, that can have an impact in the results.

In order to test if this technique can be well generalized, we applied it to Covid19 data of 3 other countries: Hungary, Czech Republic and Poland. By replicating the approach, similar to Romania case, we are confident that the predictive model that we presented in this paper can be also applied to other countries in order to identify the parameters of the model and to accurately assess the transmission dynamic of the pandemic, other infectious diseases or other compartmental models.

Appendix 1

The goal of this section is to provide the proof of Theorem 1. Recall the system (4) given by

$$
\begin{cases}
\frac{dS(t)}{dt} = -\frac{\beta}{\alpha} S\tilde{I} \\
\frac{dI(t)}{dt} = \beta S\tilde{I} - (\gamma_1 + \gamma_2)\tilde{I} \\
\frac{d\tilde{R}(t)}{dt} = \gamma_1 \tilde{I} \\
\frac{dD(t)}{dt} = \frac{\gamma_2}{\alpha} \tilde{I}.
\end{cases} \tag{6}
$$

The statement of Theorem 1 is the following.

Theorem 2 *Given $\tilde{I}(0), \tilde{R}(0)$ and $D(0), D(1), D(2), D(3), D(4)$ we can uniquely determine the parameters $\alpha, \beta, \gamma_1, \gamma_2$ of (6).*

Proof

We will first reduce the analysis to a single equation, namely the one for D(t). To do this we will write each of the involved quantities as functions of D(t) as follows

$$S = u(D), \tilde{I} = v(D), \tilde{R} = w(D).$$

The easiest to deal with is $\tilde{R}$ because from the last two equations we get

$$\frac{d\tilde{R}(t)}{dt} = \frac{\gamma_1}{\alpha\gamma_2} \frac{dD(t)}{dt}$$

which leads to $\tilde{R}(t) = \frac{\gamma_1}{\alpha\gamma_2}(D(t) - D_0) + \tilde{R}_0$.

Now, we treat the function u which determines $S(t) = u(D(t))$. Dividing the first and the last from (6) we get

$$u'(D) = -\frac{\beta}{\gamma_2} u(D)$$

which can be integrated to give $S(t)$ in terms of $D(t)$ as

$$S(t) = S_0 \exp\left(-\frac{\beta}{\gamma_2}(D(t) - D_0)\right). \tag{7}$$

Furthermore, this allows us to solve for $\tilde{I}(t) = v(D(t))$. To see this, add the first two equations from (6) and then combine this with the last one to arrive at

$$\frac{dS}{dt}(t) + \frac{1}{\alpha}\frac{d\tilde{I}}{dt}(t) = -\frac{(\gamma_1 + \gamma_2)}{\alpha}\tilde{I}(t) = -\frac{\gamma_1 + \gamma_2}{\gamma_2}\frac{dD}{dt}(t)$$

from which we deduce that

$$S(t) + \frac{1}{\alpha}\tilde{I}(t) + \frac{\gamma_1 + \gamma_2}{\gamma_2}D(t) = S_0 + \frac{1}{\alpha}\tilde{I}_0 + \frac{\gamma_1 + \gamma_2}{\gamma_2}D_0.$$

Solving now for $\tilde{I}$ and using (7) we obtain

$$\tilde{I}(t) = \alpha S_0 + \tilde{I}_0 + \frac{\alpha(\gamma_1 + \gamma_2)}{\gamma_2}(D_0 - D(t)) - \alpha S_0 \exp\left(-\frac{\beta}{\gamma_2}(D(t) - D_0)\right)$$

which combined with the last equation of (6) shows that $D(t)$ satisfies the differential equation

$$\frac{dD}{dt}(t) = \frac{\gamma_2}{\alpha}\tilde{I}_0 - (\gamma_1 + \gamma_2)(D(t) - D_0) + \gamma_2 S_0\left[1 - \exp\left(-\frac{\beta}{\gamma_2}(D(t) - D_0)\right)\right]. \tag{8}$$

Before we move forward, we will treat a little bit a general problem. Assume we take a differential equation of the form

$$\frac{dX}{dt} = f(X) \text{ with } X(0) = X_0 \geq 0$$

where $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function with $f(X_0) > 0$. The solution X_t starts positive, and the derivative is positive, thus the solution is non-decreasing for a while. Moreover, the solution is defined for all $t \geq 0$ from general results for ordinary differential

equations. By continuity of the solution, we have that the solution stays in the region $f > 0$ and thus it is increasing for all times it stays inside the region $f > 0$. It can not hit in finite time a point where $f(X(t_c)) = 0$ since then, reverting the equation (looking at $X(t_c - s)$) and combining this with the uniqueness of the solution, we must have that $X_t = X_{t_c}$ which is a contradiction. Thus, the solution is increasing and we can integrate the equation as follows:

$$\phi(X(t)) = t \text{ where } \phi(x) = \int_{X_0}^{x} \frac{1}{f(s)} ds.$$

Notice here that the function ϕ is well defined on the interval of $f > 0$ which contains X_0. Therefore we have $\phi(X(t)) = t$ for all $t \geq 0$ and $X(t)$ is increasing from 0 to infinity.

Now assume that we have two differential equations

$$\frac{dX(t)}{dt} = f(X(t)) \text{ and } \frac{dY(t)}{dt} = g(Y(t)) \text{ with } X_0 = Y_0, f(X_0) > 0, g(X_0) > 0.$$

At this stage, the point is that if $X(t_i) = Y(t_i)$ for some sequence of points $0 = t_0 < t_1 < t_2 < t_3 < t_4$, then we obtain that

$$\phi(X(t_i)) - \psi(X(t_i)) = 0 \text{ where } \phi(x) = \int_{X_0}^{x} \frac{1}{f(s)} ds, \psi(x) = \int_{X_0}^{x} \frac{1}{g(s)} ds.$$

In particular, this implies that the function $\phi(x) - \psi(x)$ has at least five zeros. Since the function $\phi(x) - \psi(x)$ is C^1, this implies that the derivative has at least four zeros, in other words this means that $\frac{1}{f(x)} - \frac{1}{g(x)}$ has at least four zeros. Finally, this means that $f(x) = g(x)$ has at least four solutions.

Returning to our problem we take now some parameters $a, b, c, d, \tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}$ and consider

$$f(x) = a - bx + c(1 - e^{-dx}) \text{ while } g(x) = \tilde{a} - \tilde{b}x + \tilde{c}(1 - e^{-\tilde{d}x}).$$

In the case $f(x) - g(x) = 0$ has at least four solutions, we actually also get that $f'(x) - g'(x) = 0$ has at least three solutions, which then upon taking another derivative gives that $f''(x) - g''(x) = 0$ has at least 2 solutions. Now this means that

$$cd^2 e^{-dx} = \tilde{c}\tilde{d}^2 e^{-\tilde{d}x}$$

has at least two different solutions. The point is that if the above is satisfied for two different values of x, say x_1 and x_2, then

$$\frac{cd^2}{\tilde{c}\tilde{d}^2} = e^{(d-\tilde{d})x_1} = e^{(d-\tilde{d})x_2}$$

which then leads to the conclusion that we must have $d = \tilde{d}$ and $c = \tilde{c}$. Going now back the ladder, using the fact that $f'(x) = g'(x)$ for three distinct values of x we have

$$-b + cde^{-dx} = -\tilde{b} + cde^{-dx}$$

and thus $b = \tilde{b}$. Finally, having $f(x) = g(x)$ for five different values of x means that we also get that $a = \tilde{a}$, thus all the parameters must be equal.

Taking this back to our equation (8) and taking $X(t) = D(t) - D_0$, knowing the values $X(0), X(1), X(2), X(3), X(4)$, then we can uniquely determine the values of

$$\begin{cases} a = \frac{\gamma_2}{\alpha}\tilde{I}_0 \\ b = \gamma_1 + \gamma_2 \\ c = \gamma_2 S_0 \\ d = \frac{\beta}{\gamma_2}. \end{cases}$$

Knowing these values is not enough to determine all the values of $\gamma_1, \gamma_2, \beta, \alpha$ because S_0 we know that

$$S_0 = 1 - \frac{1}{\alpha}(\tilde{I}_0 + \tilde{R}_0) - D_0$$

which shows that we can solve now

$$\begin{cases} \alpha = \frac{c\tilde{I}_0 + a(\tilde{I}_0 + \tilde{R}_0)}{a(1 - D_0)} \\ \beta = \frac{d(c\tilde{I}_0 + a(\tilde{I}_0 + \tilde{R}_0))}{(1 - D_0)\tilde{I}_0} \\ \gamma_1 = \frac{(b(1 - D_0) - c)\tilde{I}_0 - a(\tilde{I}_0 + \tilde{R}_0)}{(1 - D_0)\tilde{I}_0} \\ \gamma_2 = \frac{c\tilde{I}_0 + a(\tilde{I}_0 + \tilde{R}_0)}{(1 - D_0)\tilde{I}_0}. \end{cases}$$

Consequently, knowing D_0, D_1, D_2, D_3, D_4 and $\tilde{I}_0, \tilde{R}_0$ we can determine the parameters $\beta, \gamma_1, \gamma_2, \alpha$ $\square$

Appendix 2

We apply the same methodology to 3 other countries and analyze the results. We use the COVID19 data of Hungary, Czech Republic and Poland. We have to mention that we don't have information about the procedure of reporting the number of cases for these countries, which can cause a bias in the results.

By replicating the technique we are confident that the predictive model that we presented in this paper can be also applied to other countries in order to identify the

parameters of the model and to accurately assess the transmission dynamic of the pandemic. The results for the 3 countries are detailed below.

Hungary
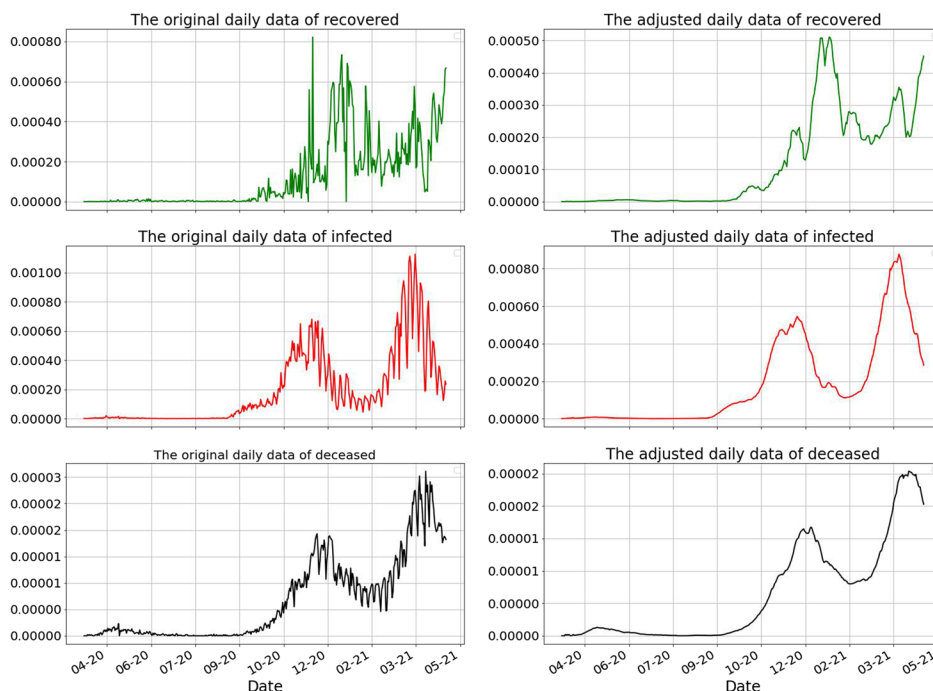
Fig. 10 Data adjustments according to the methodology presented into the article. The rows describe the data (from top to bottom) of recovered, infected and dead. The left column represents the raw data and the right column represents the adjusted data as we described above. Notice the scale and the spike in the first picture which is adjusted as we pointed out. The data was scaled by 10, 000, 000, exactly as in the case of Romania

In the next Figure we present the results of the estimated parameters $\beta, \gamma_1, \gamma_2, \alpha$, for Hungary:
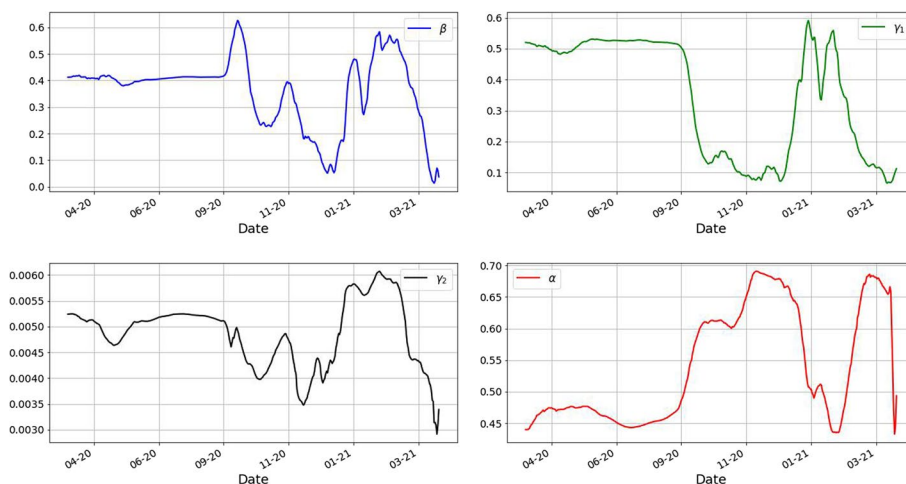
Fig. 11 Parameters of the spreading of Covid19 in Hungary estimated using the same methodology as we used

in the case of Romania

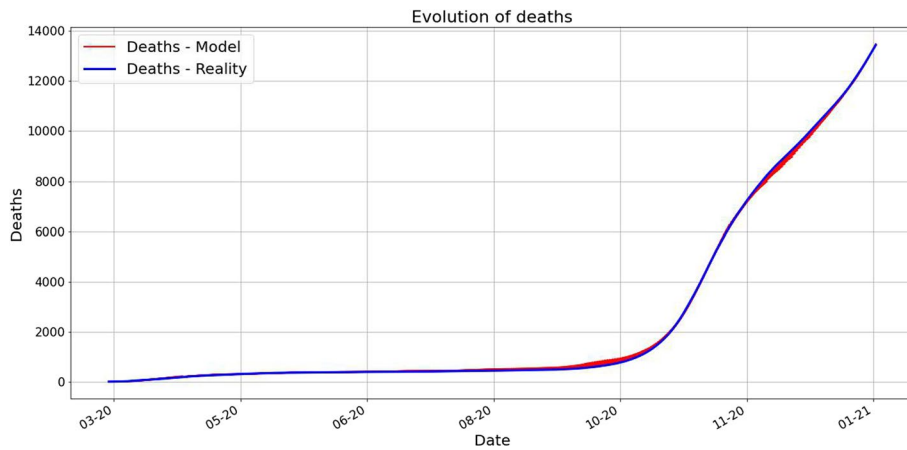Regarding the predictions:

1. Prediction of deaths, for 10 days:
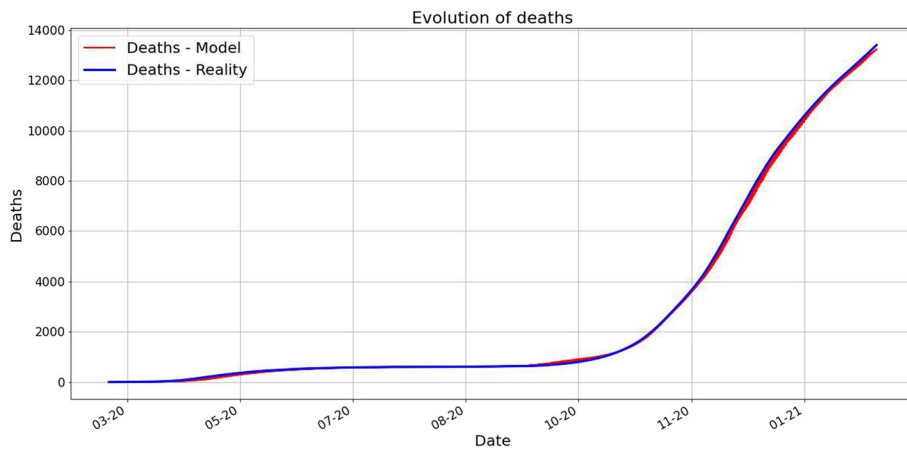
Fig. 12 In blue is the real (reported) number of deaths. For each day k, we plotted, in red, the prediction of the deaths, starting with day k, for 10 days

2. Prediction of deaths, for 30 days:
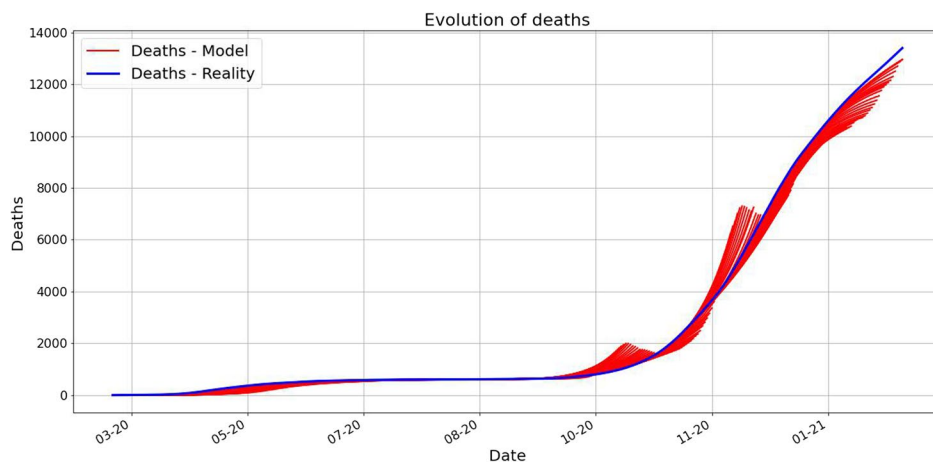
Fig. 13 In blue is the real (reported) number of deaths. For each day k, we plotted, in red, the prediction of the deaths, starting with day k, for 30 days
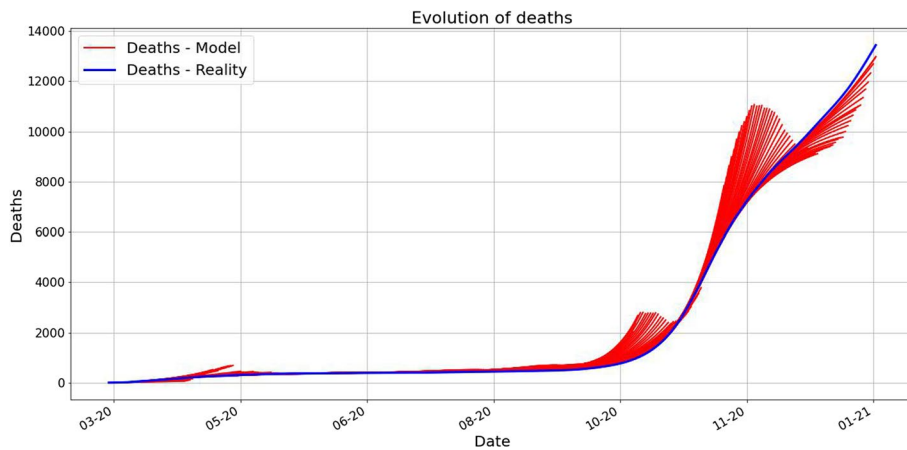
3. Prediction of deaths, for 45 days:
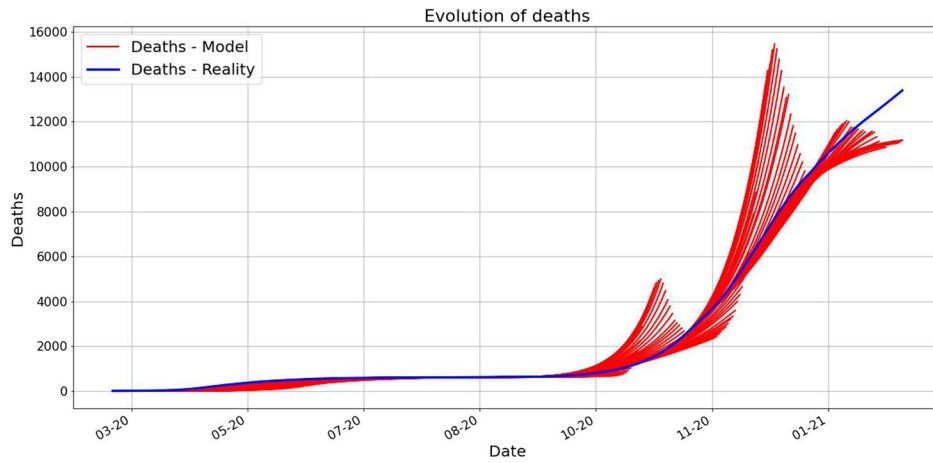
Fig. 14 In blue is the real (reported) number of deaths. For each day *k*, we plotted, in red, the prediction of the deaths, starting with day *k*, for 45 days

Regarding the MAE values, for Hungary:

Prediction	Case	MAE
10 days prediction	Deaths	40.36850
	Infected	699.77421
	Recovered	723.19638
30 days prediction	Deaths	153.05402
	Infected	6600.95112
	Recovered	3645.40171
45 days prediction	Deaths	309.09216
	Infected	15741.09146
	Recovered	6408.25157

Czech Republic

First of all we clean the data and we obtain:
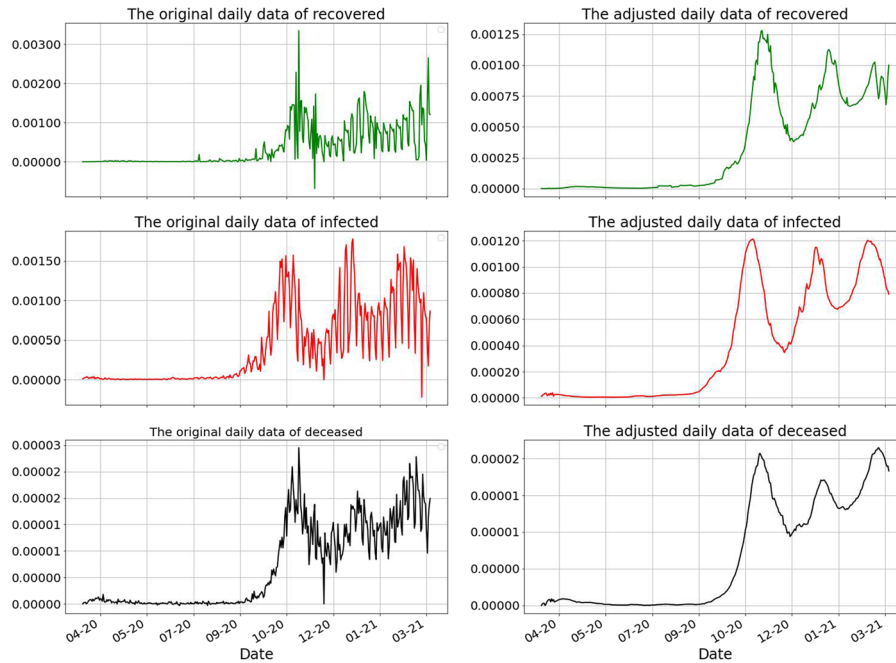
Fig. 15 Data adjustments according to the methodology presented into the article. The rows describe the data (from top to bottom) of recovered, infected and dead. The left column represents the raw data and the right column represents the adjusted data as we described above. Notice the scale and the spike in the first picture which is adjusted as we pointed out. The data was scaled by 10, 000, 000, exactly as in the case of Romania

In the next Figure we present the results of the estimated parameters $\beta, \gamma_1, \gamma_2, \alpha$, for Czech Republic:
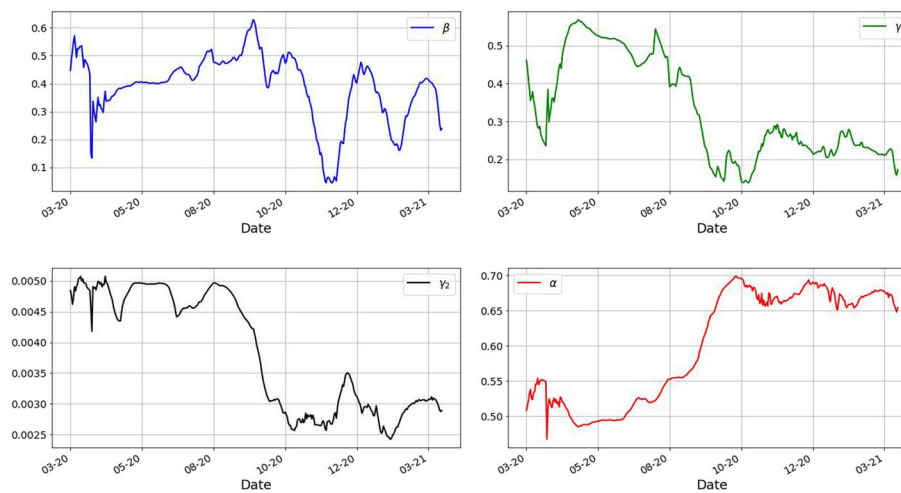
Fig. 16 Parameters of the spreading of Covid19 in Czech Republic estimated using the same methodology as we used in the case of Romania

Regarding the predictions:

1.   Prediction of deaths, for 10 days:

Fig. 17 In blue is the real (reported) number of deaths. For each day k, we plotted, in red, the prediction of the deaths, starting with day k, for 10 days
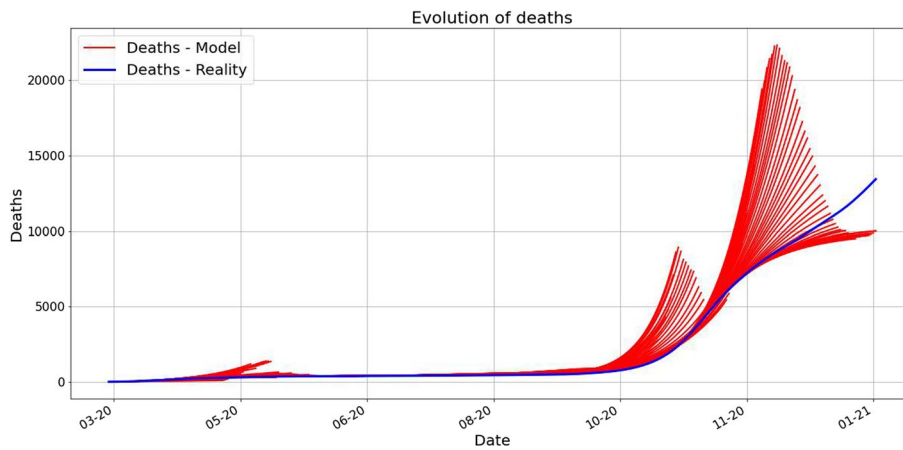
2.   Prediction of deaths, for 30 days:

Fig. 18 In blue is the real (reported) number of deaths. For each day k, we plotted, in red, the prediction of the deaths, starting with day k, for 30 days

3.   Prediction of deaths, for 45 days:

Fig. 19 In blue is the real (reported) number of deaths. For each day k, we plotted, in red, the prediction of the deaths, starting with day k, for 45 days

Regarding the MAE values, for Czech Republic:

Prediction	Case	MAE
10 days prediction	Deaths	30.50940
	Infected	2172.68655
	Recovered	921.83768
30 days prediction	Deaths	189.93063
	Infected	5308.06519
	Recovered	3645.40171
45 days prediction	Deaths	543.08844
	Infected	49255.44366
	Recovered	17283.88766

Poland

First of all we clean the data and we obtain:
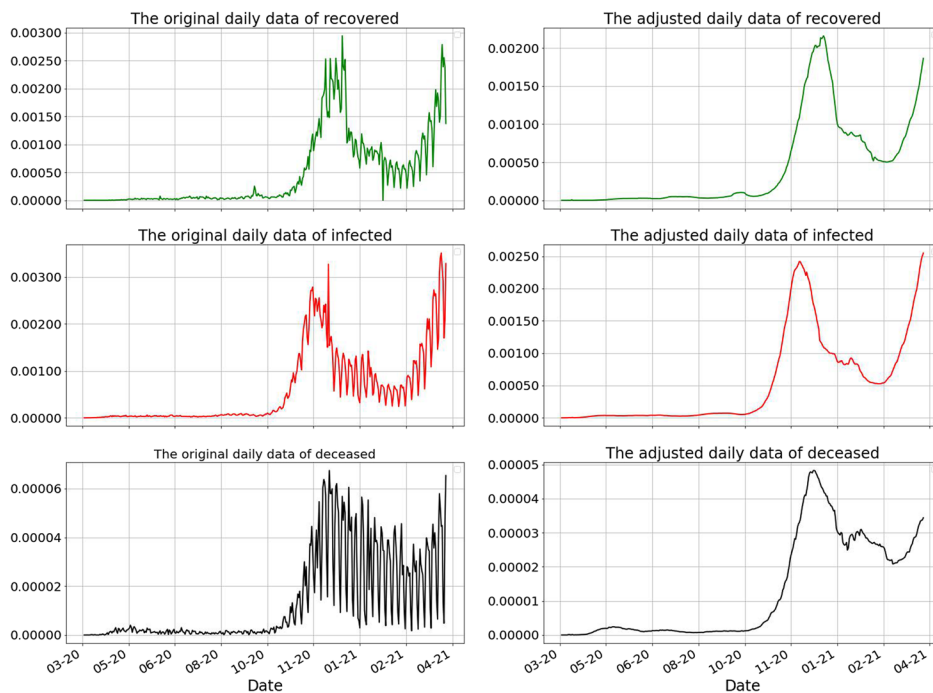
Fig. 20 Data adjustments according to the methodology presented into the article. The rows describe the data (from top to bottom) of recovered, infected and dead. The left column represents the raw data and the right column represents the adjusted data as we described above. Notice the scale and the spike in the first picture which is adjusted as we pointed out. The data was scaled by 10, 000, 000, exactly as in the case of Romania
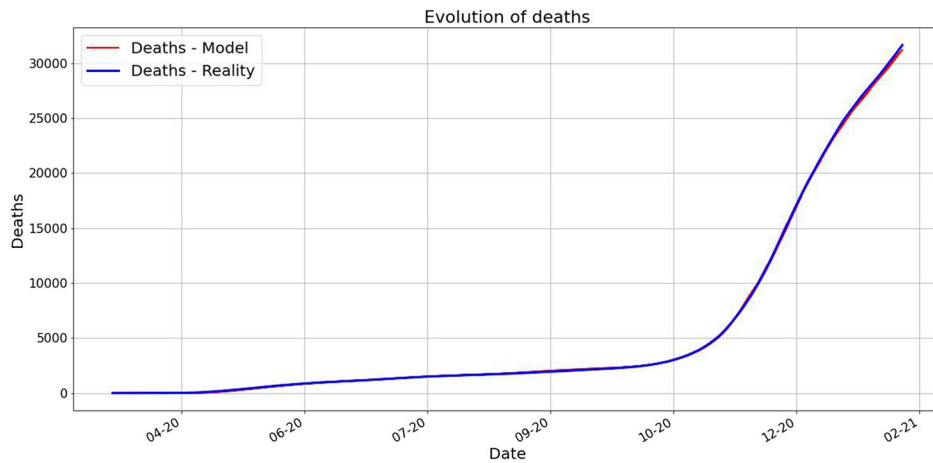
In the next Figure we present the results of the estimated parameters $\beta, \gamma_1, \gamma_2, \alpha$, for Poland:
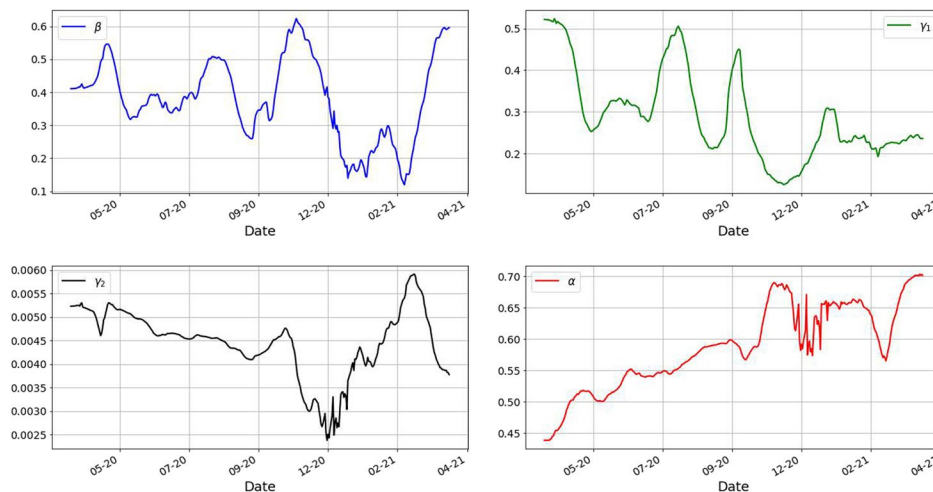
Fig. 21 Parameters of the spreading of Covid19 in Poland estimated using the same methodology as we used in the case of Romania
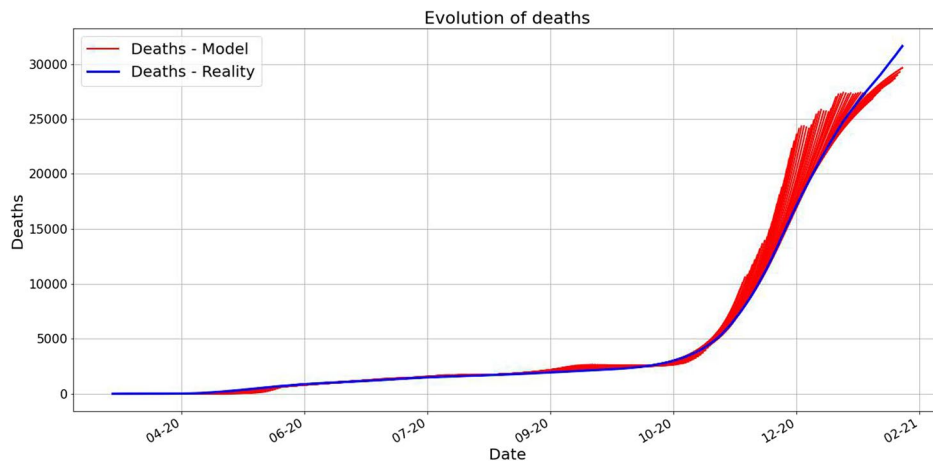
Regarding the predictions:
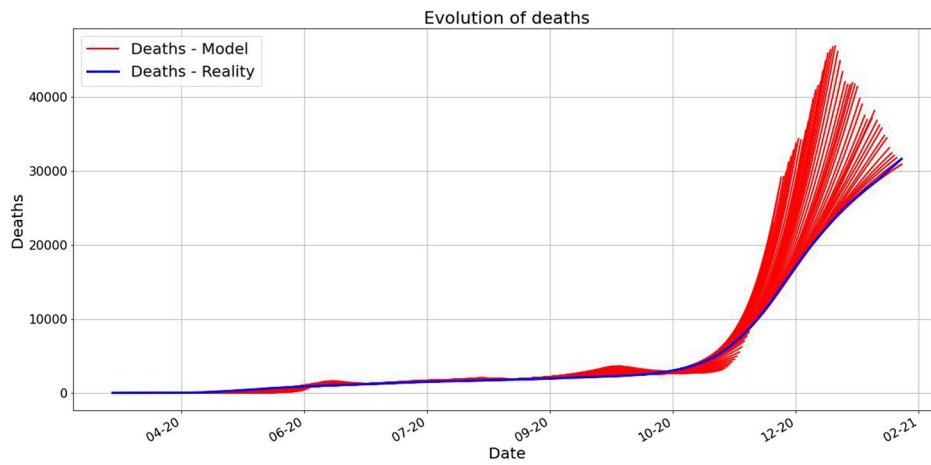
1.   Prediction of deaths, for 10 days:

Fig. 22 In blue is the real (reported) number of deaths. For each day *k*, we plotted, in red, the prediction of the deaths, starting with day *k*, for 10 days

2.   Prediction of deaths, for 30 days:

Fig. 23 In blue is the real (reported) number of deaths. For each day *k*, we plotted, in red, the prediction of the deaths, starting with day *k*, for 30 days

3.   Prediction of deaths, for 45 days:

Fig. 24 In blue is the real (reported) number of deaths. For each day *k*, we plotted, in red, the prediction of the deaths, starting with day *k*, for 45 days

Regarding the MAE values, for Poland:

Prediction	Case	MAE
10 days prediction	Deaths	42.08050
	Infected	3397.06115
	Recovered	1122.41742
30 days prediction	Deaths	270.48052
	Infected	30818.55274
	Recovered	5515.75961
45 days prediction	Deaths	842.45188
	Infected	72149.78502
	Recovered	16057.94535

Declarations

Ethics approval and consent to participate
We did not use any confidential data for the analysis in this paper and we do not have any ethical issues in this paper.

Consent for publication
We did not use any data which could possibly reveal any personal data of any patient.

Competing interests
The authors declare no competing interests.

References

1. Choisy M, Guégan J-F, Rohani P. Mathematical modeling of infectious diseases dynamics. Encycl Infect Dis Mod Methodologies. 2007;379–404
2. Wakefield J, Dong TQ, Minin VN. Spatio-temporal analysis of surveillance data. Handb Infect Dis Data Anal. Chapman and Hall/CRC. 2019;455–76.
3. Grassly NC, Fraser C. Mathematical models of infectious disease transmission. Nat Rev Microbiol. 2008;6(6):477–87.
4. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Eggo RM, Sun F, Jit M, Munday JD, et al. Early dynamics of transmission and control of covid-19: a mathematical modelling study. Lancet Infect Dis. 2020;(5):553–8
5. Sardar T, Nadim SkS, Chattopadhyay J. Assessment of 21 days lockdown effect in some states and overall india: a predictive mathematical study on covid-19 outbreak. arXiv preprint arXiv:2004.03487. 2020.
6. Schüttler J, Schlickeiser R, Schlickeiser F, Kröger M. Covid-19 predictions using a gauss model, based on data from april 2. Physics. 2020;2(2):197–212.
7. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand, 2020.
8. Tsay C, Lejarza F, Stadtherr MA, Baldea M. Modeling, state estimation, and optimal control for the us covid-19 outbreak. Sci Rep. 2020;10(1):1–12.
9. Ross R. The prevention of malaria, New York, E.P. Dutton & Company. 1910. https://archive.org/details/pr00eventi onofmalarossrich/mode/2up.
10. Ross R. An application of the theory of probabilities to the study of a priori pathometry.-part i. Proc R Soc Lond Ser A Containing Pap Math Phys Character. 1916;92(638):204–230.
11. Ross R, Hudson HP. An application of the theory of probabilities to the study of a priori pathometry.-part ii. Proc R Soc Lond Ser A Containing Pap Math Phys Character. 1917;93(650):212–225.
12. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics-i. Bull Math Biol. 1991;53(1–2):33–55.
13. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics-ii. the problem of endemicity. Bull Math Biol. 1991;53(1-2):57–87.
14. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. iii.-further studies of the problem of endemicity. Proc R Soc Lond Ser A Containing Pap Math Phys Character. 1933;141(843):94–122.
15. Godfrey KR, DiStefano JJ. Identifiability of Model Parameter. IFAC Proc. 1985;18(5):89–114.
16. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. Neuroimage. 1997;6(4):279–87.
17. Pohjanpalo H. System identifiability based on the power series expansion of the solution. Math Biosci. 1978;41(1–2):21–33.
18. Audoly S, Bellu G, D'Angiò L, Saccomani MP, Cobelli C. Global identifiability of nonlinear models of biological systems. IEEE Trans Biomed Eng. 2001;48(1):55–65.
19. Eisenberg MC, Robertson SL, Tien JH. Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease. J Theor Biol. 2013;324:84–102.
20. Hadeler KP. Parameter identification in epidemic models. Math Biosci. 2011;229(2):185–9.
21. Mummert A. Studying the recovery procedure for the time-dependent transmission rate(s) in epidemic models. J Math Biol. 2013;67(3):483–507.
22. Marinov TT, Marinova RS, Omojola J, Jackson M. Inverse problem for coefficient identification in SIR epidemic models. Comput Math Appl. 2014;67(12):2218–27.
23. Banks HT, Hu S, Thompson WC. Modeling and Inverse Problems in the Presence of Uncertainty. Chapman and Hall/CRC; 2014.
24. Julier SJ, Uhlmann JK. Unscented filtering and nonlinear estimation. Proc IEEE. 2004;92(3):401–22.
25. Soufiane B, Touaoula TM. Global analysis of an infection age model with a class of nonlinear incidence rates. J Math Anal Appl. 2016;434(2):1211–39.
26. Bentout S, Tridane A, Djilali S, Touaoula TM. Age-Structured Modeling of COVID-19 Epidemic in the USA, UAE and Algeria. Alex Eng J. 2021;60(1):401–11.
27. Bentout S, Chekroun A, Kuniya T. Parameter estimation and prediction for coronavirus disease outbreak 2019 (COVID-19) in Algeria. AIMS Public Health. 2020;7(2):306–18.
28. Djilali S, Bentout S, Kumar S, Touaoula TM. Approximating the asymptomatic infectious cases of the COVID-19 disease in Algeria and India using a mathematical model. Int J Model Simul Sci Comput. 2022;13(04):2250028.
29. Petrica M, Stochitoiu RD, Leordeanu M, Popescu I. A regime switching on covid19 analysis and prediction in romania. Sci Rep. 2022;12:15378.
30. Stochitoiu RD, Petrica M, Rebedea T, Popescu I, Leordeanu M. A self-supervised neural-analytic method to predict the evolution of COVID-19 in Romania. arXiv preprint arXiv:2006.12926. 2020.
31. Esquível ML, Krasii NP, Guerreiro GR, Patrício P. The multi-compartment si (rd) model with regime switching: An application to covid-19 pandemic. Symmetry. 2021;13(12):2427.
32. Lazebnik T. Computational applications of extended SIR models: A review focused on airborne pandemics. Ecol Model. 2023;483:110422.
33. Mohamadou Y, Halidou A, Kapen PT. A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. Appl Intell. 2020;50:3913–25.
34. Vega R, Flores L, Greiner R. SIMLR: Machine Learning inside the SIR Model for COVID-19 Forecasting. Forecasting. 2022;4:72–94.

35. Alexi A, Rosenfeld A, Lazebnik T. A Security Games Inspired Approach for Distributed Control Of Pandemic Spread. Adv Theory Simul. 2023;6:2200631.
36. Rahimi I, Chen F, Gandomi AH. A review on COVID-19 forecasting models. Neural Comput Appl. 2021;1–14.
37. Bernoulli D. Essai d'une nouvelle analyse de la mortalite causee par la petite verole. Mem Math Phys Acad Roy Sci Paris. 1766.
38. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc Royal Soc Lond Ser A Containing Pap Math Phys Character. 1927;115(772):700–21.
39. Kendall DG. Deterministic and stochastic epidemics in closed populations. Contributions to Biology and Problems of Health. University of California Press; 1956. pp. 149–66. https://www.degruyter.com/document/doi/10.1525/9780520350717-011/html.
40. Kingma D. Ba J. Adam: A method for stochastic optimization. Int Conf Learn Representations, 2014.

Publisher's Note