A self-supervised neural-analytic method to assess the evolution of COVID-19 in Romania

Radu D. Stochitoiu

Faculty of Automatic Control and Computers
University Politehnica of Bucharest
radu.stochitoiu@gmail.com

Marian Petrica

Gheorghe Mihoc - Caius Iacob Institute of
Mathematical Statistics and Applied Mathematics
of the Romanian Academy
Faculty of Mathematics and Computer Science
University of Bucharest
marianpetrica11@gmail.com

Traian Rebedea

Faculty of Automatic Control and Computers
University Politehnica of Bucharest
traian.rebedea@cs.pub.ro

Ionel Popescu

Faculty of Mathematics and Computer Science
University of Bucharest
Institute of Mathematics of the
Romanian Academy
ionel.popescu@fmi.unibuc.ro
ionel.popescu@imar.ro

Marius Leordeanu

Institute of Mathematics of the
Romanian Academy
University Politehnica of Bucharest
leordeanu@gmail.com

Abstract

Analyzing and understanding the transmission and evolution of an infectious disease is crucial to be able to design the best social and medical policies, foresee their results and deal with all subsequent social and economic effects. We address this important problem from a computational and machine learning perspective. More specifically, we want to statistically estimate all relevant parameters of the spread of COVID-19, such as the reproduction number, the death rate or the duration of the infectiousness period, based on Romanian patients, as well as to be able to predict future outcomes. This endeavor is important, since it is well known that these factors vary across the globe and might be dependent on many causes, including social, medical, age and genetic factors. At the core of our computational approach lies the paper, state-of-the-art work **?** that proposes an improved version of SEIR, which is one of the classic and established models for infectious diseases. We want to infer all the parameters of the model that govern the evolution of the pandemic based on the most reliable measurement, which is the number of deaths. The true number of infected people is impossible to precisely know.

Once the model parameters are estimated, we are able to predict all the other relevant measures, such as the number of exposed and infectious people and many other factors, as shown in this paper. To achieve this, we propose a self-supervised approach to train a deep convolutional network to estimate the correct set of Modified-SEIR model parameters, given the observed number of daily fatalities. Lastly, starting from these initial parameters, we refine the solution with a stochastic coordinate descent approach. We compare our deep learning optimization scheme with the classic grid search approach and show great improvement in both computational time and prediction accuracy. We find an optimistic result in the case of fatality rate for Romania which may be around 0.3% and we also demonstrate that our model is able to correctly predict the number of daily fatalities for up to three weeks in the future, while staying around the intervals defined by (**?**) that was also used in the United States of America and the predictions from IHME (**?**).

1 Introduction

Early understanding of the dynamics of an infectious disease is fundamental to being able to act in time and take the best safety measures for the population. Existing powerful mathematical models based on differential equations are able to assess reasonably well the evolution of the different curves (e.g. number of infected and hospitalized people, fatalities), given the correct set of model parameters. However, the inverse learning problem of finding the best parameters given the observed data is not an easy task, especially when the problem is not convex and several distinctive sets of parameters constitute relatively good local optima. The existence of multiple solutions is an important aspect, as we want to be able to predict future outcomes and learn fundamental parameters, such as the reproduction number and the fatality rate.

We take a dual neural-analytic approach, which effectively combines the power of the analytical solutions to model and predict data with a relatively small set of meaningful parameters with the power of deep neural networks to learn the inverse problem, that of estimating the correct parameters from the observed data. We propose an effective self-supervised training and prediction scheme, in which the two pathways, one classic, analytical (using differential equations) and the other based on machine learning (using deep convolutional networks) can feed each other, in tandem. During the self-supervised training phase, we start from random parameters of an improved, state-of-the-art modified SEIR model (**?**, **?**) (which we will refer to as Modified-SEIR), to predict fatality curves for a given period. Then we train the neural network to predict the known generator set of parameters, given the generated curves, using the analytical model. At test time, the network is used to rapidly estimate the correct set of parameters from the real, observed, curve of daily fatalities. Then, the set of parameters is further optimized by stochastic coordinate descent to minimize the L2-norm between the predicted curve (by the analytical Modified-SEIR model **?**) and the real curve of daily fatalities. Note that we estimate the correct set of model parameters solely from the number of fatalities, since, as mentioned above, that number is the only one that could be measured correctly. The number of true infected people is impossible to know, given mainly the limitations in testing and the relatively large portion of the population that is asymptomatic.

Among the first measures taken by the authorities in Romania was to impose a very strict social distancing plan. This restriction has a major impact on the basic reproduction number of the SEIR model, generally reducing it by a percentage between 40% and 80% (**?**, **?**). As social distancing norms were alleviated on 15 May 2020, we considered two simulated scenarios, so we could analyze and compare the impact of the heavy vs. moderate isolation restrictions. We now know, based on **?**,

that there are Gaussian simulations that could help us make a more educated guess for a date when to lift the lockdown completely.

The two main contributions that we make in this paper are:

1. For prediction, we implement a proposed mathematical model based on SEIR (**??**) (referred to as Modified-SEIR) to estimate the evolution of infectious diseases. We optimize the model to fit the data provided by Romanian health authorities for the COVID-19 pandemic, using a novel deep learning approach, trained in a self-supervised fashion.

2. We introduce a novel self-supervised deep learning approach for fast optimization and learning of the parameters of the anlytical Modified-SEIR model. More specifically, the convolutional network is trained on many outputs of the Modified-SEIR model, generated by random sets of parameters, to predict precisely the same generator parameters set, in each case. After being trained on hundreds of thousands of such synthetic cases, the neural network becomes able to take us directly in the neighborhood of the best fitting parameters, when presented with the real data curve. Then, a refinement coordinate descent procedure is applied at the end, to obtain the final solution. Our experiments clearly show that the proposed deep learning approach to optimization greatly improves speed and accuracy over the baseline optimization approach (grid search with coordinate gradient).

2   Related work

Shortly after the first cases of COVID-19 appeared in the world, an important research movement began, which aims at finding bounds for the characteristics of the infectiousness of this new coronavirus. In **?** authors show that the basic reproduction number is significantly influenced by travel restrictions, ranging from 2.35 to 1.05. In their procedures, they used an estimate of the incubation period equal to 5.2 days, but which can be as low as 2 days, according to a study conducted in Wuhan (**?**). This study also found that the value of 2.2 is a good approximation for the basic reproduction number, which is also similar to the findings of our research reported here.

In **?** authors use a SEIR (Susceptible-Exposed-Infectious-Recovered) model to estimate the basic reproduction number at 2.68 on January 25 in Wuhan. At the same time, it reveals a worldwide incubation time that ranges from 4.6 to 6.9. In this article we see the positive correlation between the basic reproduction number and the probability of creating an exponential epidemic starting with a single infected person.

The official report of the World Health Organization (**?**) introduces several observed parameters, including a basic reproduction number in the range of $2 - 2.5$, an incubation period with an average in the range of $5 - 6$, a minimum hospitalization rate around 6.1%, represented by critically ill patients, and a maximum of $\simeq 20\%$, to which are added those 13.8% in severe condition, a period from incubation to death in a wide range of $2 - 8$ weeks and a recovery time for mild cases topped by 12 days and for severe cases up to 14 days of hospitalization.

In several studies dealing with the estimation of the incubation period of viruses such as 2019-nCoV, SARS or MERS (**?**, **?**, **?**) we notice values in the range of $4.4 - 6.9$.

Another important factor studied is how strict the rules of social distancing must be in order to stop the number of cases from increasing. In **?** it is estimated that a reduction in the basic reproduction number by $58 - 76\%$ is needed to stop the increase in the number of infected people, considering the base reproduction number equal to 3.11.

The core of our prediction model is the Modified-SEIR (**?**), with a freely available toolbox, the Epidemic Calculator (**?**), which we use as a stepping stone in our own implementation and validation of our model and experiments.

To solve a very difficult prediction problem, such as the one relating to COVID-19, **?** proposes AI-powered ways to manage limited healthcare resources, develop personalized patient management and treatment plans, inform policies, enable effective collaboration, and better understand and account for uncertainty.

A limited number of published medical studies **??** present and discuss the epidemiology, clinical preparedness, and medical challenges of the COVID-19 pandemic in Romania. As shown next, our model accurately fits the real data, predicts future events (not seen during training) and estimates important core characteristics of the virus specific to Romania, such as the reproduction number, the length of infectiousness, time to recovery, and fatality rate. Note that the related work presented above provides specific ranges, established in the medical literature, within which we optimize the parameters of the Modified-SEIR.

Numerous studies have addressed the global evolution of the pandemics. In the context of the COVID-19 pandemic, the characterization and forecasting of transmission dynamics, as well as the rigorous estimation of model parameters, have been comprehensively examined in the papers **???????**.

3   A Modified-SEIR prediction model

In order to assess and understand key factors of COVID-19 we approach the recent mathematical model (**?**) based on the classic SEIR (**S**usceptible, **E**xposed, **I**nfectious, **R**emoved; **?**), which is a widely accepted standard for modeling the evolution of infectious diseases. The Modified-SEIR model follows the usual steps in which an infectious disease evolves. The evolution along with key elements and measures are fully described by a set of differential equations (Table **??**) that we present in this section. We also offer a visual representation of the model in Figure **??**.
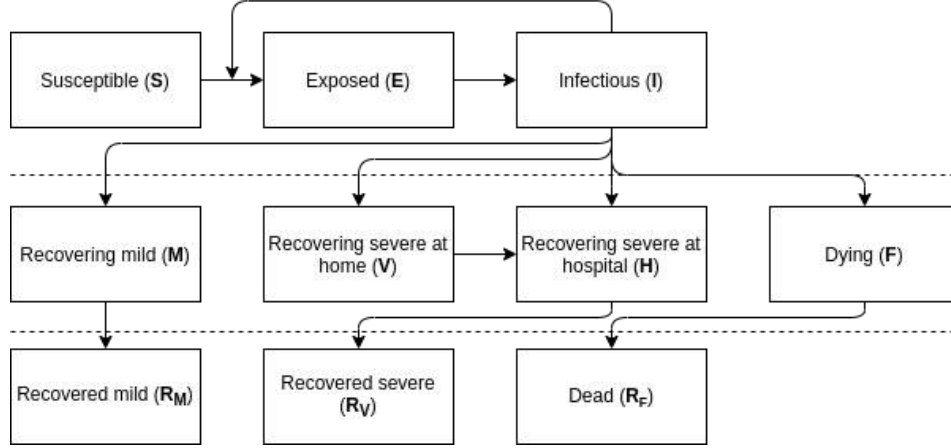
Figure 1: Modified SEIR model: the "Removed" case is split into: 1) Recovered from mild symptoms; 2) Recovered from severe symptoms; 3) Deceased. The diagram follows the differential equations of the Modified-SEIR model presented in Table **??**.

We further divide the *Removed* compartment into three categories: recovered from mild symptoms (R_M), recovered from severe symptoms (R_V) and deceased (R_F). There is also an extra layer of differential equations in the middle which helps us better shape and understand the dynamics of the disease. Following the Modified-SEIR model (**?**) and the Epidemic Calculator (**?**), we assume that all fatalities come from hospitals, and that all severe cases are admitted to hospitals immediately after the infectious period ends.

Table 1: Modified SEIR: System of differential equations, which describe the evolution of key characteristics and measures of pandemics over time. In our work we consider the number of fatalities (F, R_F) as the only real variable that could be measured correctly and we optimize the model according to it.

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta I S \qquad\qquad \frac{\mathrm{d}E}{\mathrm{d}t} = \beta I S - \sigma E$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \gamma I \qquad\qquad \frac{\mathrm{d}M}{\mathrm{d}t} = P_M \gamma I - \frac{1}{T_M} M$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = P_V \gamma I - \frac{1}{T_H} V \qquad\qquad \frac{\mathrm{d}H}{\mathrm{d}t} = \frac{1}{T_H} V - \frac{1}{T_V} H$$

$$\frac{\mathrm{d}F}{\mathrm{d}t} = P_F \gamma I - \frac{1}{T_F} F \qquad\qquad \frac{\mathrm{d}R_M}{\mathrm{d}t} = \frac{1}{T_M} M$$

$$\frac{\mathrm{d}R_V}{\mathrm{d}t} = \frac{1}{T_V} H \qquad\qquad \frac{\mathrm{d}R_F}{\mathrm{d}t} = \frac{1}{T_F} F$$

$$\beta = \begin{cases} \frac{R_0}{T_{inf}}, & \text{before } T \\ \frac{(1-P_T)R_0}{T_{inf}}, & \text{otherwise} \end{cases} \qquad \sigma = \frac{1}{T_{inc}}$$

$$\gamma = \frac{1}{T_{inf}} \qquad\qquad P_M = 1 - P_V - P_F$$

The advantage of having an explicit analytic model versus a pure deep learning approach is that each parameter has a clear meaning that is easy to interpret. In this case, the meaning of each parameter used is described and summarized in Table **??**. The ranges in which we look for optimal parameters are fixed according to recently published medical research and their confidence intervals, as presented

in Section **??**. Note that we use constant values for some parameters, such as the total size of the population or the time from severe symptoms to hospitalization (according to **?**).

We choose a continuous-time ODE framework considering the following reasons:

- ODEs admit well-known tools for steady-state and stability analysis (eigenvalues, manifolds) that underpin our sensitivity study and control formulation.

- Embedding the continuous dynamics directly in the loss function allows back-propagation through time via adjoint or neural-ODE methods, which is more difficult in a discrete model.

- Although data arrive at daily intervals, the underlying processes (transmission, progression, recovery) occur continuously. A continuous model naturally averages over day-to-day variability and mitigates day-to-day reporting noise.

Table 2: Modified SEIR parameters. "Deduced" means that the values are deduced from the system of differential equations found at Table **??**.

Name	Description	Initial value	Range
S	Susceptible population	$N - I_0$	Deduced
E	Exposed population	0	Deduced
I	Infectious population	I_0	Deduced
M	Recovering at home with mild symptoms	0	Deduced
V	Recovering at home with severe symptoms	0	Deduced
H	Recovering in hospital with severe symptoms	0	Deduced
F	Dying	0	Deduced
R_M	Recovered from mild symptoms	0	Deduced
R_V	Recovered from severe symptoms	0	Deduced
R_F	Dead (Fatal)	0	Deduced
P_M	Mild symptoms rate		Deduced
P_V	Severe symptoms rate		$[0.04 - 0.2]$
P_F	Case fatality rate		$[0.1\% - 3\%]$
T_{inc}	Length of incubation period (days)		$[2 - 14]$
T_{inf}	Length of infectiousness period (days)		$[3 - 14]$
T_M	Recovery time for mild cases (days)		$[4 - 12]$
T_V	Recovery time for severe cases (days)		$[7 - 35]$
T_H	Time from severe symptoms onset to hospitalization (days)		$[5]$
T_F	Time from end of infectiousness to death (days)		$[14 - 35]$
R_0	Basic reproduction number		$[1.5 - 2.5]$
T	Intervention time to reduce R_0 (days)		$[20 - 22]$
P_T	Percentage to decrease transmission by after intervention		$[40\% - 80\%]$
β	Transmission rate		Deduced
σ	Rate of getting infectious from being exposed		Deduced
γ	Recovery rate		Deduced
I_0	Number of initial infections		$[1500 - 2500]$
N	Total size of population		$[20175912]$

In order to solve the time-based differential equations and produce the different evolution curves we use a fourth order Runge-Kutta integrator (**?**).

4 Learning the model parameters

As the number of new fatalities per day is easily known and is not influenced by the number of actual tests run in any location, we consider it as ground truth in our experiments. Even though this number mixes patients with comorbidities (whose health is also influenced by other conditions) with those without comorbidities, the actual measured number of infected people who are dying is ultimately the only estimation that can be considered certain in the current research. We use the data uploaded daily by Johns Hopkins CSSE.

Our purpose is to find a set of parameters that best approximates a curve modeled by the modified SEIR, the real curve of the daily number of fatalities. We denote D_{real} as the vector (curve) of the reported daily number of deaths in a specific time interval and D_θ as the vector (curve) generated by the model with parameters θ, for the same time interval. Thus, the cost function that we minimize to find the best fitting parameters is the square root of the sum of squared errors between the real and the predicted curve, as shown in Equation (**??**). We denote the optimal set of parameters by θ^*, as defined in (**??**). The valid search (optimization) ranges for each parameter are presented in Table **??**, as mentioned previously, and constitute relatively large search regions, as unions over the ranges published in recent medical literature.

We test and compare two main ways to find the optimal parameters. One is a baseline, which starts with a classic grid search procedure followed by a stochastic coordinate descent refinement. The second optimization method, which is our main technical novelty, is to make an initial guess of the parameters using the self-supervised trained convolutional network followed by the same refinement step using stochastic coordinate descent. Each optimization module (grid search, neural network, and coordinate descent) is described in the next sections.

$$J(\theta) = \sqrt{\sum_i \left(D_{real}(i) - D_\theta(i)\right)^2} \tag{1}$$

$$\theta^* = \arg\min_\theta J(\theta) \tag{2}$$

Throughout calibration we use the first difference of the cumulative fatal compartment, $D(i) = \Delta R_F(i) = R_F(i) - R_F(i-1)$. The cost function in eq. **??** then measures the fit between this model-generated daily fatality curve and the reported data.

All compartmental variables in our ODE system (e.g. $S(t), E(t), I(t), R_F(t)$) are defined as *fractions* of the total population N. To compare the model to observed daily-death counts $D_{\text{real}}(i)$, we simply multiply the model's fatality compartment by N:

$$\widetilde{R}_F(t) = N \cdot R_F(t) \implies D_\theta(i) = \Delta\widetilde{R}_F(i) = N\big[R_F(i) - R_F(i-1)\big].$$

4.1 Optimization by grid search

There are 11 parameters that we optimize over: $I_0, R_0, T_{inc}, T_{inf}, P_F, T_F, T_M, T_V, P_V, P_T, T$. Therefore, a full and very fine grid search is computationally infeasible. However, we can divide each range into (2-4) smaller ranges and look for a decent approximation to start with. The grid search module is followed by the coordinate descent refinement procedure (Section **??**).

4.2 Optimization by a neural network

Because of the long computational time required by grid search optimization, we propose a deep learning approach, using a convolutional neural network trained in a self-supervised manner, as

discussed previously, that is able to bring the solution to the neighborhood of the optimum very fast. The neural network optimization module is also followed by the same final coordinate descent procedure. Interestingly enough, it turns out that the results, when using the neural net optimization, are vastly superior in both speed and accuracy to the grid search approach.

4.3   Self-supervised learning of neural net optimization

Below we present the exact steps taken for the self-supervised scheme in which the neural network learns to guess the right set of parameters, given a curve of daily deaths for a given period of time.

1. Create a dataset
    - Take 100000 random samples from a uniform distribution of the 11 parameters in the ranges indicated in Table **??**;
    - Generate a curve of daily fatalities using the Modified-SEIR model (Table **??**) for each set of model parameters picked at the previous step. Although the estimated total number of deaths is cumulative (the R_F variable), we take the number of daily deaths as daily increments in the total number of fatalities.
    - For each curve, we randomly select a fixed number L of consecutive days of daily deaths (L is defined by the number of days in certain time ranges in our experiments, such that for [March 22, May 3], $L = 43$). This vector of L consecutive numbers, representing the fatalities for the corresponding L days, modeled by Modified-SEIR, along with the corresponding set of parameters, will constitute the 100K training pairs used in training the neural net optimizer presented in the following. Note that the start of the L-day sequence is chosen randomly (so it could be at the beginning or towards the end of the pandemic). We try to mirror the real case, when we really do not know which day should be considered the first of the pandemic.

2. Multi-head neural network optimizer, trained in a self-supervised way, to predict the Modified-SEIR model parameters, given the L-element curve of daily fatalities (generated by precisely the same set of parameters that should be predicted by the network).
    - Hidden layers
        - (a) Conv1D (512, 5, 'relu')
        - (b) MaxPooling1D
        - (c) Conv1D (128, 5, 'relu')
        - (d) MaxPooling1D
        - (e) Conv1D (32, 5, 'relu')
        - (f) MaxPooling1D
        - (g) Flatten
        - (h) Dense (512)
        - (i) Dense (256)
        - (j) Dense (128)
    - One output for each parameter
    - Loss: Mean Squared Error
    - Optimizer: Adam (**?**)

We compare the grid search approach with neural network predictions. The advantage of the latter approach is that it offers almost instantaneous predictions. In Figure **??** we show the percentage of tasks (problems that are randomly generated by the model using random initialization) where the samples (sets of parameters) found by grid search produce superior curves (closer to the truth) than

the ones predicted by the neural network. The percentage is a function of time since for an infinite amount of time we expect grid search, with a sufficiently fine grid, to beat the neural network. The plot shows how vastly superior in terms of speed the neural network is. Almost 7 hours of running for grid search, on an Intel© Core™ i9-9980HK CPU @ 2.40GHz x 8, are not enough to surpass the neural net predictions.
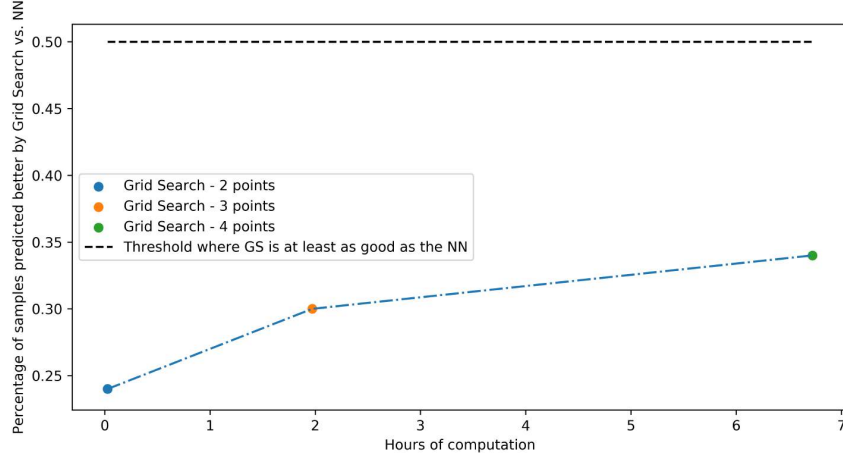

Figure 2: Comparison between the grid search and the neural network solutions. We present how often grid search produces better solutions than the neural net. We clearly see that 7 hours of computation are far from sufficient for grid search to beat the neural net. Note that in the plot, the final refinement procedure is not used by either of the two approaches.

4.4 Final coordinate descent refinement

We can expect that neither of the two approaches (grid search and neural net optimization) will directly produce an optimal solution, even though we do expect them to output a set of parameters that are close to a local optimum. In order to refine our results, we further apply an iterative *stochastic coordinate descent* approach similar to **?**. Starting from the best predictions of a given first-stage module (neural net or grid search), we take random subsets of two parameters at a time, divide their search ranges into 20-40 parts around the current best solution, and replace the next values of the chosen parameters with the ones that minimize the cost function. We iterate the procedure until we reach a convergence of 10^{-3} absolute error.

We make use of the stochastic property iteratively because we assume that some parameters influence the cost more than others. Thus, by choosing random subsets of parameters to optimize over, we avoid the risk of spending valuable time optimizing over subsets of parameters that do not bring much value.

While our model is formulated as a deterministic ODE system, one could extend it to an SDE. Such stochastic formulations enable explicit representation of reporting noise, but at the cost of significantly greater computational and inferential complexity. In particular, likelihood-based calibration of SDEs often requires particle-filtering or Monte Carlo methods, and analytical steady-state and sensitivity-analysis results become less tractable. We view the deterministic approach as a pragmatic compromise: it allows efficient gradient-based training and clear stability insights, while still capturing the dominant epidemic trends observed in aggregate fatality data.

5 Experimental analysis

5.1 On data dependency

We consider three particular data sets on which we optimize and search for the best parameters of our model.

- Daily fatalities from March 22 to May 3 (2020);
- Daily fatalities from March 22 to May 14 (2020);
- Daily fatalities from March 22 to May 21 (2020).

We know that on May 15 2020 the Romanian authorities changed the policy from state of emergency to state of alert. Because the time from incubation to death is greater than one week, we assume the reported daily fatalities from May 15 2020 to May 21 2020 are not influenced by the changed policy. This enables us to search for optimal parameters for the interval March 22 2020 to May 21 2020, as it can be modelled by Modified-SEIR.

For every data set, we present in Table **??** the optimal model parameters found with neural net optimization followed by the refinement step, as defined in Section **??**. As stated previously, they minimize the L2 distance between the real and the predicted curves of fatalities from March 22 to May 3, May 14, and May 21, respectively (all in the year 2020, of course). For each set of optimal parameters, we compute the error of prediction for the following dates: May 3, May 15, May 21, June 3, June 8, June 9, June 10, June 11 (all in 2020).

As the prediction errors are inversely proportional to the data set size, we prove that our approach is data-driven and provides better solutions for bigger data sets. In other words, every new day of observations is important in finding the real parameters that shape the evolution of the COVID-19 infectious disease.

Table 3: Best parameters found by our neural net optimization followed by the final coordinate descent refinement.

Name	Description	May 3	May 14	May 21
I_0	Initial infectious population	2450	1600	1600
R_0	Basic reproduction number	2.63	2.63	2
T_{inc}	Length of incubation period (days)	2	2	2
T_{inf}	Length of infectiousness period (days)	7.4	7.4	3
P_F	Case fatality rate	0.39%	0.68%	0.39%
T_F	Time from end of infectiousness to death (days)	14	16.1	28
T_M	Recovery time for mild cases (days)	4	13.28	4
T_V	Recovery time for severe cases (days)	7	7	7
P_V	Severe symptoms rate	5%	15.6%	10%
P_T	Decrease in transmission after intervention	60%	62%	56%
T	Intervention time to reduce R_0 (days)	21	21	21
Err. May 3	Prediction absolute error on May 3, 2020	3.04%	3.92%	3.16%
Err. May 15	Prediction absolute error on May 15, 2020	4.21%	3.55%	0.19%
Err. May 21	Prediction absolute error on May 21, 2020	9.95%	8.22%	1.99%
Err. Jun 3	Prediction absolute error on June 3, 2020	24.92%	20.14%	6.87%
Err. Jun 8	Prediction absolute error on June 8, 2020	31.22%	24.94%	8.44%
Err. Jun 9	Prediction absolute error on June 9, 2020	31.83%	25.26%	8.12%
Err. Jun 10	Prediction absolute error on June 10, 2020	33.31%	26.40%	8.60%
Err. Jun 11	Prediction absolute error on June 11, 2020	34.48%	27.25%	8.77%

5.2 On finding optimal parameters

It is worth mentioning that in general the reversed problem of finding the correct model parameters from the observation of partial data (in our case we observe only the values of one output variable, the number of daily fatalities) is ill posed, since the problem is not necessarily convex and many different sets of parameters can produce similar output. However, this is a common case in machine learning, in which several valid solutions exist. The AI system usually learns to predict the most probable output (in this case, the set of parameters) given the observed data, based on its training experience. One good example is the case of vision, in which many different 3D worlds with different semantic interpretations could produce the same 2D image. Nevertheless, the visual system learns to pick the most likely interpretation (given its prior experience) out of infinitely many. In our specific case, we expect that the synthetic generation of many curves from different sets of parameters will help the neural network implicitly learn the priors in the data model, such that the network will learn to output, from the many different solutions, one that is most likely to have produced the given curve. We can easily imagine that there are certain distinct neighborhoods of parameters that generate similar curves and that a set of parameters coming from a larger neighborhood is more likely than one coming from a smaller one. In other words, sets of model parameters for which the output curve is more stable (low curve gradient w.r.t parameters) are probably more likely than sets of parameters for which the curve changes rapidly in their immediate neighborhood. Such subtle priors in the space of parameters should be implicitly learned during the self-supervised training if sufficiently many pairs of model-generated curves (input) - parameters (output) are presented to the network.

It is clear by now that searching for the best parameters is a complex task, especially if the cost function is not convex, thus admitting multiple local minima. Besides the intuitive discussion, we also experimentally analyze this issue by comparing different 4D plots where the 3D space is defined

by three different parameters and the fourth dimension, the final cost function, is defined by color. In Figure **??** one can see that there are regions where there is a linear dependency between two parameters preserving the minimum cost and that there are multiple intervals with local minima for sets of three parameters. The latter finding is especially interesting as it tells us that our prediction might not be the correct one even if it has the lowest cost. Or, in other words, different distinctive sets of parameters may generate the same final cost. Thus, learning from larger sets of reported data, over longer periods of time, may shift the balance towards a different set of optimal parameters which can immediately modify the foreseen dynamics of the coronavirus infectiousness and fatality. This situation, which is commonly encountered in many AI tasks, reveals the inner ambiguity and difficulty of the problem tackled. What is, however, important and relevant here, is that we are able to learn sets of parameters which are plausible, often matching many independent findings in the literature, which predict with surprising accuracy the evolution of COVID-19 in Romania.
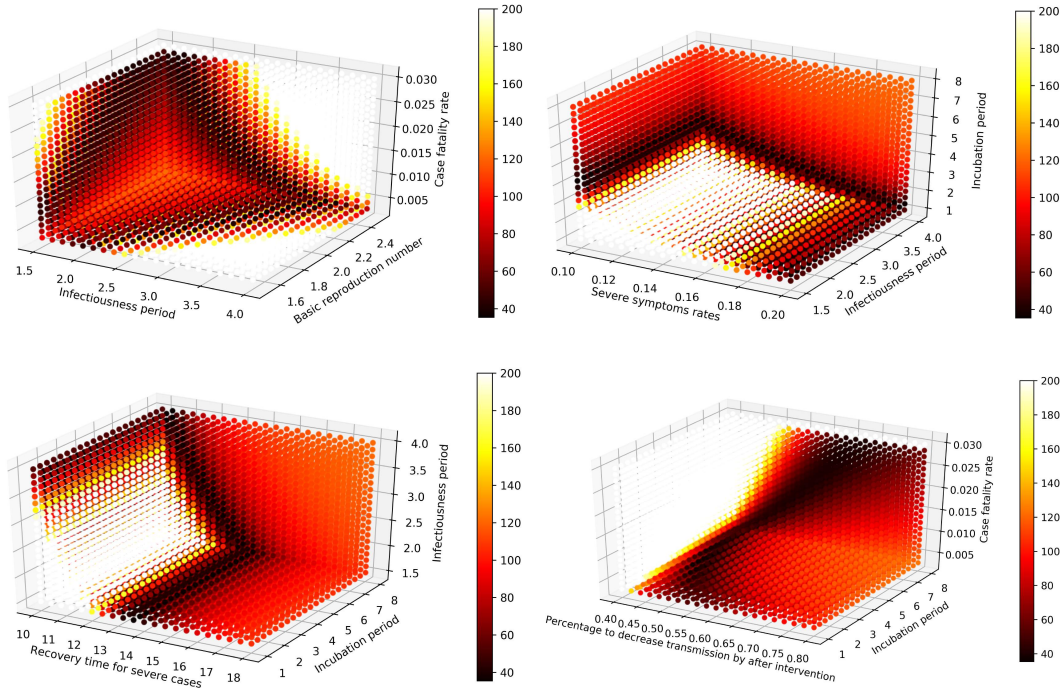

Figure 3: Plots of the cost function, represented by color, computed for joint distributions of three parameters at a time. The cost function goes from black (optimal/smallest) to white (least optimal/greatest). For better visualization we upper bound the cost at 200.

Starting from the knowledge that a particular parameter set might not necessarily be the correct one even if it has the lowest cost on the limited training data, we compare the predictions of two different sets of parameters that have similar L^2 costs (close to the lowest cost found for the interval from March 22, 2020, to May 3, 2020). In Figure **??** we present the evolution curves corresponding to these two sets of parameters. Both of them fit the given data equally well.
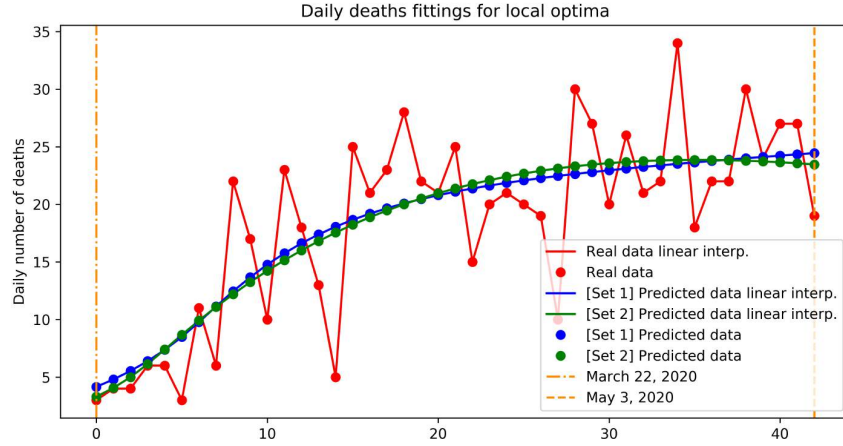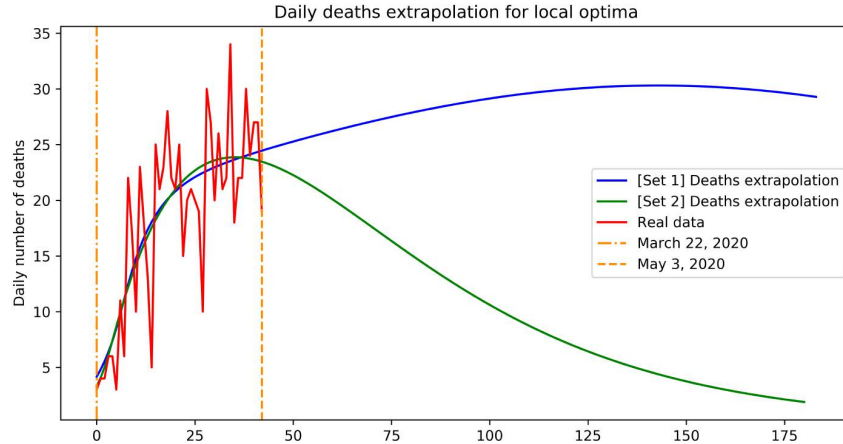
Figure 4: Daily deaths fitted curves for two sets of parameters with similar cost. The red dots represent the officially reported daily number of deaths in Romania. The blue and green lines with dots show the fitting of the Modified-SEIR models through the real data.

Surprisingly, when we extrapolate the curves into the future (by running the models according to the corresponding parameters into the future), we notice substantially different evolution curves. In Figure **??** we notice how one approximation goes down quickly, while the other continues to increase for another three months.

Figure 5: Daily deaths fitted curves extrapolation for two sets of parameters with similar costs (on seen data). The red dots represent the officially reported daily number of deaths in Romania. The blue and green lines show the extrapolation of the Modified-SEIR models through the real data. Note how different the future predictions are, between the two sets of parameters.

This subtle change in the way the two fitted curves diverge at the end of the data set has drastic outcomes regarding, among other characteristics, the number of active infections and the total number of fatalities. Figure **??** shows us that while the fitted curves for the observed interval (March 22, 2020, to May 3, 2020) are similar, one set of parameters predicts a total of 2290 deaths, while the other predicts a total of 10225. We compare their predictions in Table **??**.
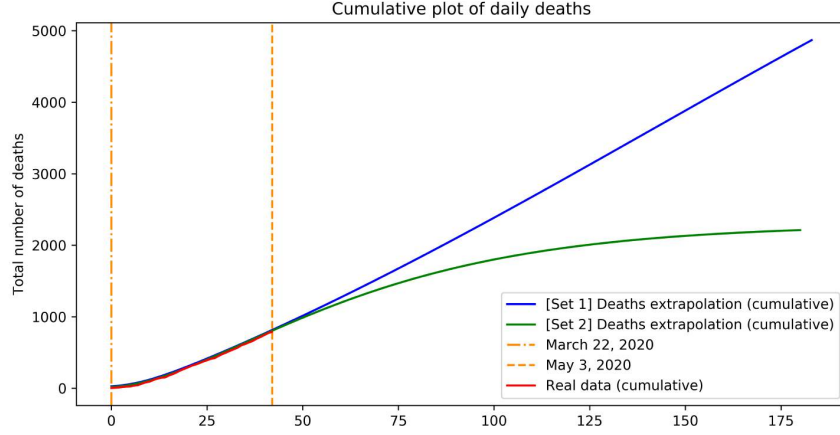
13

Figure 6: Cumulative extrapolation of the total number of fatalities for two sets of parameters with similar cost. The red dots represent the officially reported total number of deaths in Romania. The blue and green lines show the cumulative extrapolation of the Modified-SEIR models through the real data.

Table 4: Two sets of parameters for the Modified-SEIR model with similar costs but very different evolutions.

Name	Description	Set 1	Set 2
I_0	Initial infectious population	2450	1548
R_0	Basic reproduction number	2.63	2.32
T_{inc}	Length of incubation period (days)	2	2
T_{inf}	Length of infectiousness period (days)	7.4	3.48
P_F	Case fatality rate	0.39%	0.15%
T_F	Time from end of infectiousness to death (days)	14	25.9
T_M	Recovery time for mild cases (days)	4	10
T_V	Recovery time for severe cases (days)	7	10
P_V	Severe symptoms rate	5%	10%
P_T	Decrease in transmission after intervention	60%	59.2%
T	Intervention time to reduce R_0 (days)	21	21
$J(\theta)$	L2 cost of the fitting	33.285	33.597
Err. May 3	Prediction absolute error on May 3, 2020	3.04%	1.65%
Err. May 15	Prediction absolute error on May 15, 2020	4.21%	0.28%
Err. May 21	Prediction absolute error on May 21, 2020	9.95%	5.99%
Err. Jun 3	Prediction absolute error on June 3, 2020	24.92%	10.80%
Err. Jun 8	Prediction absolute error on June 8, 2020	31.22%	13.22%
Err. Jun 9	Prediction absolute error on June 9, 2020	31.83%	13.07%
Err. Jun 10	Prediction absolute error on June 10, 2020	33.31%	13.75%
Err. Jun 11	Prediction absolute error on June 11, 2020	34.48%	14.02%
$\sum$ **Fatalities**	**Prediction of total number of fatalities**	**10225**	**2065**

5.3 Key observations about the model parameters found

As discussed in the previous section, the approach based on a neural network optimization followed by a stochastic alternative coordinate descent seems to overfit the data set. In order to help our model

learn parameters that generalize better in the future, we add a relative future cost based on the real numbers to the cost presented in Section **??**, so its new formula is introduced in Equation **??**.

$$J(\theta) = \sqrt{\sum_i (D_{real}(i) - D_\theta(i))^2} + \lambda_{May\,21} + \lambda_{June\,3} + \lambda_{June\,11} \tag{3}$$

$$\lambda_{date} = 100 \cdot |\frac{R_{date} - P_{date}}{R_{date}}| \tag{4}$$

$$R_{date} \equiv Reported\ number\ of\ fatalities\ on\ \textbf{date} \tag{5}$$

$$P_{date} \equiv Predicted\ number\ of\ fatalities\ on\ \textbf{date} \tag{6}$$

We present in Table **??** the best model parameters found with neural net optimization followed by the refinement step. As stated previously, they minimize the L2 distance between the real and predicted curves of fatalities, from March 22, 2020, to May 21, 2020, while trying not to overfit the data set by looking ahead until the last available reported date. Several interesting observations are worth making: the estimated basic reproduction number was found to be 2.21, which is very close to the value estimated in the literature. Since it is above 1, it defines an exponential growth in the number of infected people unless reduced by the social distancing measures. The imposed measures of containment indeed reduced the reproduction number by 60% (from 2.2 to 0.884), so the curve starts decreasing towards zero. The continuous decrease definitely helps our medical staff to manage patients better and safer, which may also explain why the fatality rate found is so small, of only 0.245%. Note that this value is significantly lower than other numbers reported in the literature so far, which is very good news but it is highly dependent on the number of tested people and on the social dynamics of the observed population.

Table 5: Best parameters found by our neural net optimization + refinement, in the case of heavy social distancing assumption through the real data.

Name	Description	Value
I_0	Initial infectious population	1725
R_0	Basic reproduction number	2.21
T_{inc}	Length of incubation period (days)	2
T_{inf}	Length of infectiousness period (days)	3.47
P_F	Case fatality rate	0.245%
T_F	Time from end of infectiousness to death (days)	20.3
T_M	Recovery time for mild cases (days)	10
T_V	Recovery time for severe cases (days)	10
P_V	Severe symptoms rate	10%
P_T	Percentage to decrease transmission by after intervention	60%
T	Intervention time to reduce R_0 (days)	21

Based on this set of parameters, we analyze two cases that influence the daily deaths curve:

1. Heavy social distancing, meaning that the enforced norms will not be diminished until the end of the pandemic period;

2. Moderate social distancing, meaning that on May 15, 2020, the social interaction increased by 10%.

5.4 Heavy social distancing

Here, we assume that people will adopt a careful behavior and the heavy social distancing rules will apply long after they were proposed (R_0 becomes 40% of its initial value). In Figure **??** we plot our best fitting for the daily deaths approximation using the parameters from Table **??**. You can notice a change in the convexity of the prediction (blue) curve when the heavy social distancing norms have been adopted (on intervention day).
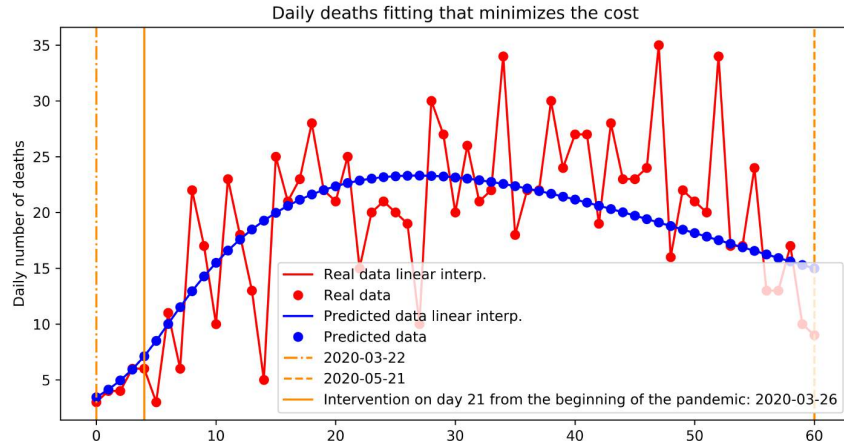

Figure 7: Daily deaths fitted curves. The red dots represent the officially reported daily number of deaths in Romania. The blue line with dots shows our best fit of the Modified-SEIR model through the real data.

Using the RK4 integrator, as presented in Section **??**, we extrapolate around 200 days from the first day of reported data. In this way, we predict when there are going to be less than 5, 3 or 1 deaths per day, as you can see in Figure **??**. A key insight is that the maximum number of daily deaths, the peak of our extrapolation, has already passed on April 18, 2020 (which is right before the Orthodox Easter on April 19-21), meaning that the curve is going to stay under 24 deaths per day.
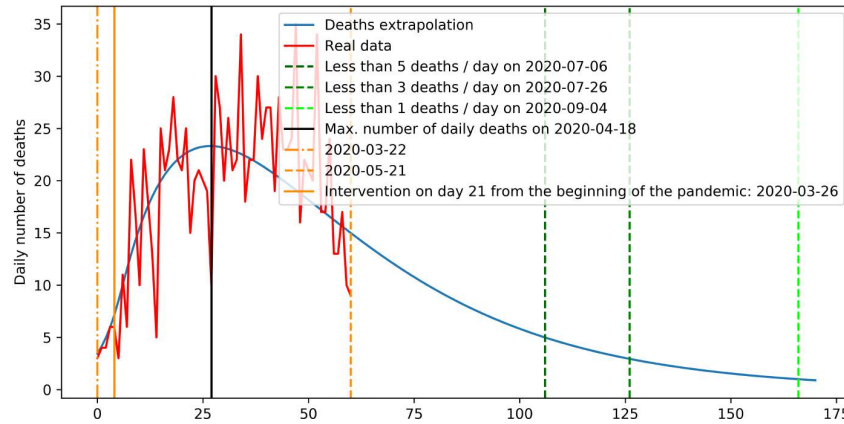

Figure 8: Extrapolation of daily number of fatalities. The red plot represents the officially reported daily number of deaths in Romania. Our extrapolation of daily deaths is the blue line and the green lines represent the days when we estimate less that 5, 3 or 1 daily fatalities.

16

A subject that is of interest is the predicted total number of fatalities caused by the virus. Using the daily numbers extrapolation, we create a cumulative curve which tells us the total number of fatalities by a certain date. Thus, we predict a total of around 1730 using data from March 22, 2020, to May 21, 2020, as presented in Figure **??**.
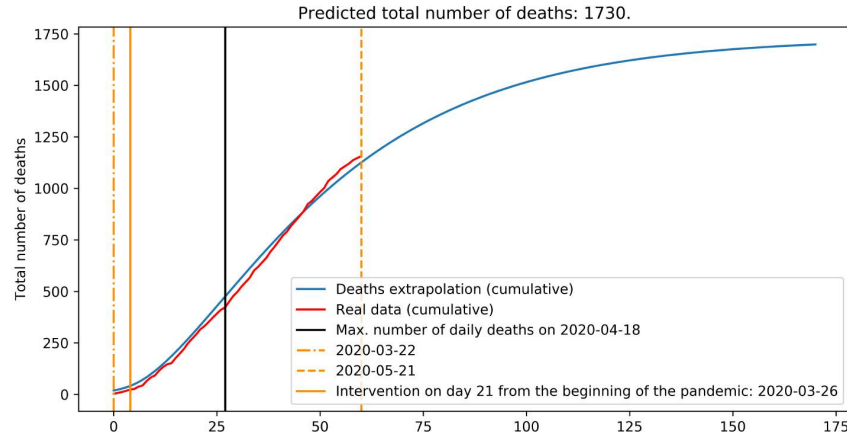

Figure 9: Extrapolation of cumulative fatalities. The red plot represents the officially reported cumulative number of deaths in Romania. Our extrapolation of the cumulative number is the blue line.

There is the possibility of optimizing the parameters for fitting the cumulative number of fatalities instead of their daily number. The reason we opted for the latter is three-fold. First, the amount of information that each specific day brings to the cumulative function becomes smaller and smaller each day, converging to 0 at infinity. The cumulative number thus increases with each day and the meaningful information for a given day becomes much smaller than the total cumulated number. It is thus expected that learning could suffer from numerical issues. Second, the number of daily fatalities contains a certain amount of noise that can help us generalize better for our predictions. And third, the curve of daily numbers shows better and clearer when the pandemic peaks and when it is expected to diminish to a non-threatening state.

We know that the real number of actively infectious people is hard to obtain, as the number of tested people at a time is just a fraction of the whole population. However, once the right parameters are learned, the model can estimate a total number of infectious people. The results are shown in Figure **??**.
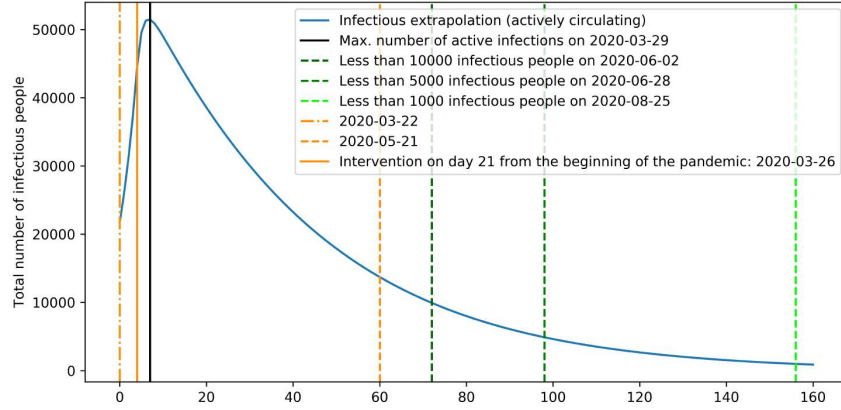
Figure 10: Total number of infectious people extrapolation. We predict less than 10000 infections on June 2, 2020, less than 5000 infections on June 8, 2020, and less than 1000 infections on August 25, 2020.

5.5 Moderate social distancing

While we do not know for sure how the social dynamics influence the evolution of the coronavirus, we conduct an experiment assuming that on May 15, 2020, when the social distancing norms became less constraining, the already reduced basic reproduction number (by the heavy social distancing norms) increased by 10%, while the rest of the parameters remained the same.

In Figure **??** we show in our prediction that the period until there will be less than 5 daily deaths is prolonged until the end of August, 2020. We do not see a second peak because the basic reproduction number does not get to be over 1 again.
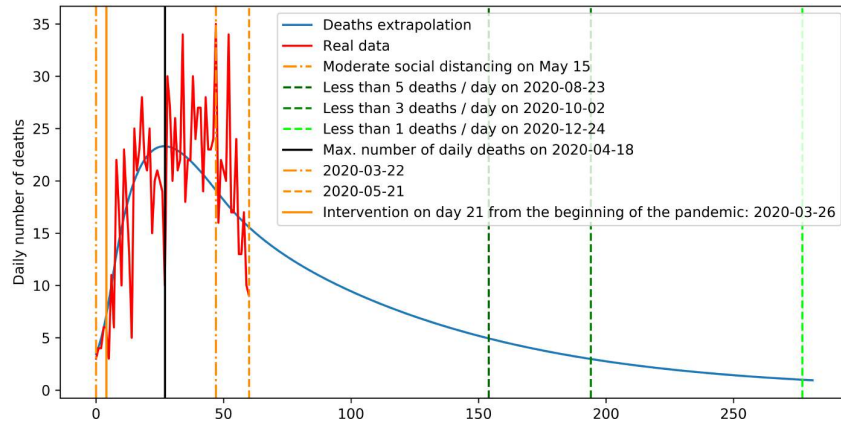

Figure 11: Daily fatalities extrapolation with increased mobility (moderate social distancing) from May 15, 2020. The red plot represents the officially reported daily number of deaths in Romania. Our extrapolation is the blue line and the green lines represent the days when we estimate less that 5, 3 or 1 daily fatalities.

The total number of deaths jumps to 2361, shown in Figure **??**. Compared to **?** we seem a bit pessimistic as the number of deaths predicted by us by Aug 1, 2020, is 1861 while their machine learning approach predicts 1776, as of June 12, 2020. Thus, we are still inside their confidence bounds: [1586 - 2187]. Please note that their approach is also data dependent and can offer significantly different predictions by using more data.
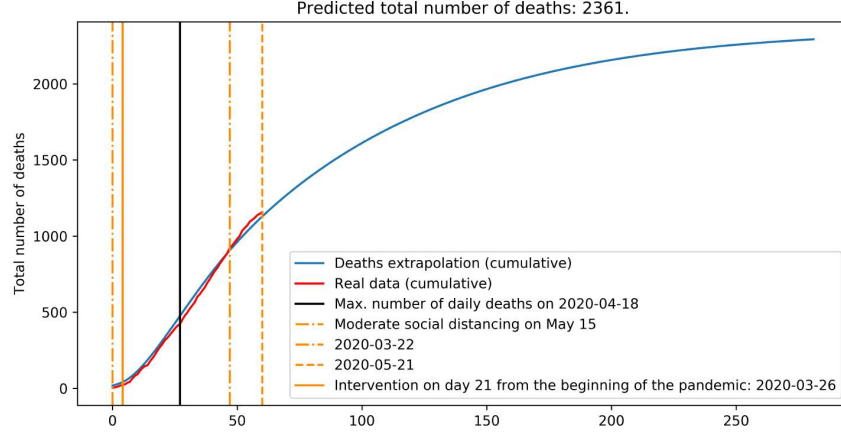
Figure 12: Cumulative deaths extrapolation with increased mobility from May 15, 2020. The red plot represents the officially reported cumulative number of deaths in Romania. Our extrapolation of the cumulative number of deaths is the blue line.

In Figure **??** one could see the effect of a slight change in the prevention norms, as the total number of actively infected people takes way more time to go down.
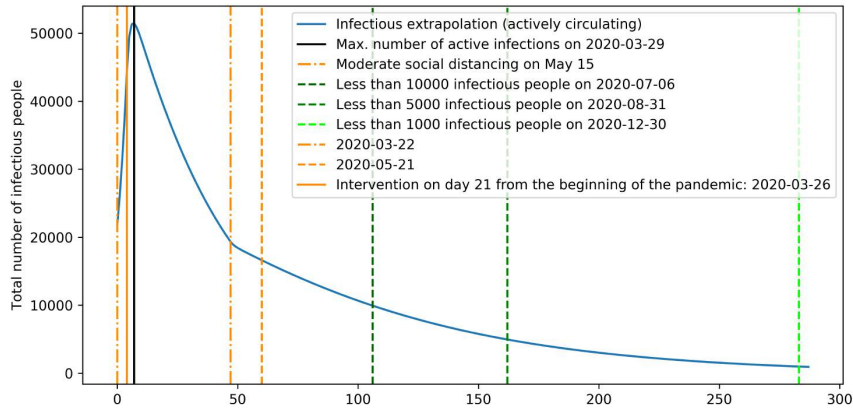

Figure 13: Total number of infectious people extrapolation with less social distancing from May 15, 2020.

5.6 Validation for our approach

Ultimately, such modeling is important not only to fit the available observed data and estimate various model parameters, such as the fatality rate, but to predict future events. The ability to predict future outcomes is its real value, in order to prepare the best courses of action in advance. Acting on time, especially in the face of a pandemic, is vital. Thus, in order to test the validity of our model (as it is usually done, in fact, in machine learning) we compare the future predictions made on May 21, 2020 (based on observed data until that date), with the information available until June 11, 2020. Our model is surprisingly accurate with the heavy social distancing assumptions. The predicted total number of fatalities is close to the reported values: 1296 true reported deaths on June 3, 2020, versus 1294 (predicted by the heavy social distancing model) and 1312 (predicted by the moderate social distancing model) and 1369 true reported fatalities on Jun 11, 2020, versus 1375 (predicted by the heavy social distancing model) and 1411 (predicted by the moderate social distancing model).

Regarding the number of infected people, we already know that the number of reported active cases is lower than the real one, so, as mentioned previously, we can use our model to guess the actual number of active infections. Testing our predictions against future data is shown in Table **??**.

Table 6: Predictions verification. HSD means heavy social distancing and MSD means moderate social distancing.

Criterion	Reported	HSD prediction	MSD prediction
Total number of deaths on May 3, 2020	790	807 (2.15%)	807 (2.15%)
Total number of deaths on May 15, 2020	1070	1031 (3.64%)	1031 (-3.64%)
Total number of deaths on May 21, 2020	1156	1125 (2.68%)	1128 (2.42%)
Total number of deaths on Jun 3, 2020	1296	1294 (0.15%)	1312 (1.23%)
Total number of deaths on Jun 11, 2020	1369	1375 (0.44%)	1411 (3.07%)
Total number of deaths on Jun 19, 2020	1484	1442 (2.83%)	1501 (1.15%)
Active infections on May 3, 2020	7504	22066	22066
Active infections on May 15, 2020	5997	16093	17673
Active infections on May 23, 2020	5494	13000	16278
Active infections on Jun 3, 2020	4573	9666	14478
Active infections on Jun 11, 2020	4530	7778	13259
Active infections on Jun 19, 2020	5361	6253	12117

5.7 Predictions for the future

Now that we have two models, one following heavy social distancing norms and the other following moderate social distancing norms, that have prediction errors < 4% so far, we can use them as bounds for our future predictions. We summarize our findings in Table **??**.

Table 7: Predictions. HSD means heavy social distancing and MSD means moderate social distancing.

Criterion	HSD prediction	MSD prediction
Total number of deaths on July 1, 2020	1521	1620
Total number of deaths on August 1, 2020	1640	1861
Total number of deaths on September 1, 2020	1692	2027
Total number of deaths on October 1, 2020	1713	2137
Total number of deaths on November 1, 2020	1723	2213
Total number of deaths on December 1, 2020	1727	2262
Total number of deaths	1730	2361
Less than 5 deaths per day	Jul 6, 2020	Aug 23, 2020
Less than 3 deaths per day	Jul 26, 2020	Oct 2, 2020
Less than 1 deaths per day	Sep 4, 2020	Dec 24, 2020
Active infections on July 1, 2020	4499	10548
Active infections on August 1, 2020	1912	7252
Active infections on September 1, 2020	808	4900
Active infections on October 1, 2020	351	3315
Active infections on November 1, 2020	148	2196
Active infections on December 1, 2020	64	1467
Less than 10000 active infections	Jun 2, 2020	Jul 6, 2020
Less than 5000 active infections	Jun 28, 2020	Aug 31, 2020
Less than 1000 active infections	Aug 25, 2020	Dec 30, 2020

6 Conclusions

In this paper we propose a computational model to predict the evolution of COVID-19 in Romania and estimate key factors of the pandemics such as the fatality rate, incubation period, infectiousness period and reproduction number, based on the state of the art Modified-SEIR model **?**. Our technical novelty consists in the way we optimize the parameters of the model, through a self-supervised deep learning approach, in which a convolutional neural network learns from synthetic data, produced by the analytical Modified-SEIR model for random sets of parameters, to predict the correct parameter set - which is known, since it is the one used to generate the synthetic data. Our results show beyond any doubt that our novel self-supervised approach is effective and learning a set of parameters which are not only able to fit the observed data but also to accurately predict in the future, for the three weeks period tested (which is a relatively large period in the case of a rapidly evolving pandemics).

At the conclusion of our study, we highlight some important findings comprising the total number of fatalities by following the heavy social distancing norms (1730), the total number of deaths following a small decrease in the prevention norms on May 15, 2020, (2361) and the fact that we already passed the peak of the daily number of deaths on April 18, 2020 (one day before the Orthodox Easter). Our predictions are right inside and around the bounds predicted by IHME (1614 deaths by August 1, as of June 12, 2020; **?**) and the ML-based approach presented in Section **??** (1776 deaths by August 1, as of June 12, 2020; **?**). This and the fact that our set of found parameters are close to the ones presented in the latest literature (for example, an optimal basic reproduction number of 2.21) only empowers the idea that our novel approach can be useful in a fast-paced pandemic, maybe not only for the case of Romania. A notable finding is that the case fatality rate in all the local optima sets seems to be significantly less than 1%, mostly around 0.245% and 0.3% - and this is a optimistic surprise, when compared to the estimates in the literature.

Considering that our model trained on data collected until May 21, 2020, accurately describes the future evolution (future unseen data) of the number of fatalities until June 11, 2020, we conclude that both the model and its inner parameters found provide answers that are very close to the true ones. The results strongly indicate that we should seriously consider data-driven computational approaches, in combination with machine learning, in the analysis and decision making process, with respect to fundamental aspects of our lives (such as it is the case of COVID-19 pandemics), for the future and greater good of society.

7 Declaration

7.1 Ethics approval

We did not use any confidential data for the analysis in this paper, and we do not have any ethical issues in this paper.

7.2 Consent for publication

We did not use any data that could possibly reveal any personal data of any patient.

7.3 Availability of data and material

The code used in this paper is available online in a Colab repository: `https://colab.research.google.com/drive/1940gRu6cZOhken1ki-PZxeX_VOQTPZ39?usp=sharing`

7.4 Competing interests

The authors declare that they have no competing interests.

7.5 Funding

There are no funding sources for this paper.

7.6 Authors' contributions

All authors contributed equally to this paper.

References

Jantien A Backer, Don Klinkenberg, and Jacco Wallinga. Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, china, 20-28 january 2020. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 25(5), February 2020. ISSN 1025-496X. doi: 10.2807/1560-7917.ES.2020.25.5.2000062. URL `https://europepmc.org/articles/PMC7014672`.

Rajiv Chowdhury, Kevin Heng, Md Shajedur Rahman Shawon, Gabriel Goh, Daisy Okonofua, Carolina Ochoa-Rosales, Valentina Gonzalez-Jaramillo, Abbas Bhuiya, Daniel Reidpath, Shamini Prathapan, et al. Dynamic interventions to control covid-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries. *European journal of epidemiology*, pages 1–11, 2020.

Stanca Ciupe and Necibe Tuncer. Identifiability of parameters in mathematical models of sars-cov-2 infections in humans. *Scientific Reports*, 12, 08 2022. doi: 10.1038/s41598-022-18683-x.

Iulian Gherghel and Mihai Bulai. Is romania ready to face the novel coronavirus (covid-19) outbreak? the role of incoming travelers and that of romanian diaspora. *Travel Medicine and Infectious Disease*, 2020.

Gabriel Goh. Epidemic calculator, 2020. URL `https://gabgoh.github.io/COVID/index.html`.

Youyang Gu. Covid-19 projections using machine learning, 2020. URL `https://https://covid19-projections.com/`.

Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000. doi: 10.1137/S0036144500371907. URL `https://doi.org/10.1137/S0036144500371907`.

Murray IHME COVID-19 health service utilization forecasting team, Christopher JL. Forecasting the impact of the first wave of the covid-19 pandemic on hospital demand and deaths for the usa and european economic area countries. *medRxiv*, 2020. doi: 10.1101/2020.04.21.20074732. URL `https://www.medrxiv.org/content/early/2020/04/26/2020.04.21.20074732`.

Joel Kelso, George Milne, and Heath Kelly. Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza. *BMC public health*, 9:117, 05 2009. doi: 10.1186/1471-2458-9-117.

Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations*, 12 2014.

Olga Krivorotko, Sergey Kabanikhin, M. Sosnovskaya, and D. Andornaya. Sensitivity and identifiability analysis of covid-19 pandemic models. *Vavilov Journal of Genetics and Breeding*, 25:82–91, 03 2021. doi: 10.18699/VJ21.010.

Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, and Rosalind M Eggo. Early dynamics of transmission and control of covid-19: a mathematical modelling study. *medRxiv*, 2020. doi: 10.1101/2020.01.31.20019901. URL `https://www.medrxiv.org/content/early/2020/02/18/2020.01.31.20019901`.

Eric Lau, C Hsiung, Benjamin Cowling, Chang-Hsun Chen, Lai-Ming Ho, Thomas Tsang, Chiu-Wen Chang, Christl Donnelly, and Gabriel Leung. A comparative epidemiologic analysis of sars in hong kong, beijing and taiwan. *BMC infectious diseases*, 10:50, 03 2010. doi: 10.1186/1471-2334-10-50.

Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy S.M. Leung, Eric H.Y. Lau, Jessica Y. Wong, Xuesen Xing, Nijuan Xiang, Yang Wu, Chao Li, Qi Chen, Dan Li, Tian Liu, Jing Zhao, Man Liu, Wenxiao Tu, Chuding Chen, Lianmei Jin, Rui Yang, Qi Wang, Suhua Zhou, Rui Wang, Hui Liu, Yinbo Luo, Yuan Liu, Ge Shao, Huan Li, Zhongfa Tao, Yang Yang, Zhiqiang Deng, Boxi Liu, Zhitao Ma, Yanping Zhang, Guoqing Shi, Tommy T.Y. Lam, Joseph T. Wu, George F. Gao, Benjamin J. Cowling, Bo Yang, Gabriel M. Leung, and Zijian Feng. Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*, 382(13):1199–1207, 2020. doi: 10.1056/NEJMoa2001316. URL `https://doi.org/10.1056/NEJMoa2001316`. PMID: 31995857.

Jie Long, A. Q. M. Khaliq, and K. M. Furati. Identification and prediction of time-varying parameters of covid-19 model: a data-driven deep learning approach. *International Journal of Computer Mathematics*, 98(8):1617–1632, 2021. doi: 10.1080/00207160.2021.1929942.

Tchavdar T. Marinov and Rossitza S. Marinova. Dynamics of covid-19 using inverse problem for coefficient identification in sir epidemic models. *Chaos, Solitons & Fractals: X*, 5:100041, 2020. ISSN 2590-0544. doi: https://doi.org/10.1016/j.csfx.2020.100041. URL `https://www.sciencedirect.com/science/article/pii/S2590054420300221`.

Gemma Massonis, Julio R. Banga, and Alejandro F. Villaverde. Structural identifiability and observability of compartmental models of the covid-19 pandemic. *Annual Reviews in Control*, 51:441–459, 2021. ISSN 1367-5788. doi: https://doi.org/10.1016/j.arcontrol.2020.12.001. URL `https://www.sciencedirect.com/science/article/pii/S1367578820300778`.

Ahmed Alaa Mihaela van der Schaar. How artificial intelligence and machine learning can help healthcare systems respond to covid-19. 3 2020.

Marian Petrica, Radu D Stochitoiu, Marius Leordeanu, and Ionel Popescu. A regime switch analysis on covid-19 in romania. *Scientific Reports*, 12(1):15378, 2022.

Corneliu Petru Popescu, Alexandru Marin, Violeta Melinte, George Sebastian Gherlan, Filofteia Cojanu Banicioiu, Adelina Dogaru, Sebastian Smadu, Ana Maria Veja, Elena Nedu, Delia Stanciu, et al. Covid-19 in a tertiary hospital from romania: Epidemiology, preparedness and clinical challenges. *Travel Medicine and Infectious Disease*, 2020.

Zhaozhi Qian, Ahmed Alaa, and Mihaela Schaar. When to lift the lockdown? global covid-19 scenario planning and policy effects using compartmental gaussian processes. 05 2020.

Jonathan M Read, Jessica RE Bridgen, Derek AT Cummings, Antonia Ho, and Chris P Jewell. Novel coronavirus 2019-ncov: early estimation of epidemiological parameters and epidemic predictions. *medRxiv*, 2020. doi: 10.1101/2020.01.23.20018549. URL `https://www.medrxiv.org/content/early/2020/01/28/2020.01.23.20018549`.

Arun Sikder, Md. Biplob Hossain, and Md Hamidul Islam. Compartmental modelling in epidemic diseases: a comparison between sir model with constant and time-dependent parameters. *Inverse Problems*, 39, 02 2023. doi: 10.1088/1361-6420/acb4e7.

Chen Zheng Tan Delin. On a general formula of fourth order runge-kutta. *Journal of Mathematical Science & Mathematics Education*, 2012. URL `http://w.msme.us/2012-2-1.pdf`.

Victor Virlogeux, Vicky Fang, Minah Park, Jianhong Wu, and Benjamin Cowling. Comparison of incubation period distribution of human infections with mers-cov in south korea and saudi arabia. *Scientific Reports*, 6, 10 2016. doi: 10.1038/srep35839.

World Health Organization WHO. Report of the who-china joint mission on coronavirus disease 2019 (covid-19), 2020. URL `https://www.who.int/publications-detail/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19)`.

Stephen J. Wright. Coordinate descent algorithms. *Mathematical Programming*, 2015. doi: 10.1007/s10107-015-0892-3. URL `https://arxiv.org/abs/1502.04759`.

Joseph T Wu, Kathy Leung, and Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395(10225):689 – 697, 2020. ISSN 0140-6736. doi: https://doi.org/10.1016/S0140-6736(20)30260-9. URL `http://www.sciencedirect.com/science/article/pii/S0140673620302609`.

8   Appendix

The equations of the proposed model are:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta IS \qquad\qquad \frac{\mathrm{d}E}{\mathrm{d}t} = \beta IS - \sigma E$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - \gamma I \qquad\qquad \frac{\mathrm{d}M}{\mathrm{d}t} = P_M \gamma I - \frac{1}{T_M} M$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = P_V \gamma I - \frac{1}{T_H} V \qquad\qquad \frac{\mathrm{d}H}{\mathrm{d}t} = \frac{1}{T_H} V - \frac{1}{T_V} H$$

$$\frac{\mathrm{d}F}{\mathrm{d}t} = P_F \gamma I - \frac{1}{T_F} F \qquad\qquad \frac{\mathrm{d}R_M}{\mathrm{d}t} = \frac{1}{T_M} M$$

$$\frac{\mathrm{d}R_V}{\mathrm{d}t} = \frac{1}{T_V} H \qquad\qquad \frac{\mathrm{d}R_F}{\mathrm{d}t} = \frac{1}{T_F} F$$

$$\beta = \begin{cases} \frac{R_0}{T_{inf}}, & \text{before } T \\ \frac{(1-P_T)R_0}{T_{inf}}, & \text{otherwise} \end{cases} \qquad \sigma = \frac{1}{T_{inc}}$$

$$\gamma = \frac{1}{T_{inf}} \qquad\qquad P_M = 1 - P_V - P_F$$

We point out that the parameters $(\beta, \sigma, \gamma, P_M, P_V, P_F, T_M, T_H, T_V, T_F)$ are positive constants.

8.1   Identification of the steady states

A steady state occurs when all time derivatives are zero.

- From $\frac{dR_M}{dt} = \frac{1}{T_M} M = 0$, since $T_M > 0$, we have $M = 0$.
- From $\frac{dR_V}{dt} = \frac{1}{T_V} H = 0$, since $T_V > 0$, we have $H = 0$.
- From $\frac{dR_F}{dt} = \frac{1}{T_F} F = 0$, since $T_F > 0$, we have $F = 0$.
- From $\frac{dM}{dt} = P_M \gamma I - \frac{1}{T_M} M = 0$, with $M = 0$, we get $P_M \gamma I = 0$. Since $P_M, \gamma > 0$, we have $I = 0$.
- From $\frac{dI}{dt} = \sigma E - \gamma I = 0$, with $I = 0$, $\sigma E = 0$. Since $\sigma > 0$, we have $E = 0$.
- From $\frac{dS}{dt} = -\beta IS = 0$, with $I = 0$, S is unconstrained.
- From $\frac{dE}{dt} = \beta IS - \sigma E = 0$, with $I = 0$, $E = 0$, this is satisfied.
- From $\frac{dV}{dt} = P_V \gamma I - \frac{1}{T_H} V = 0$, with $I = 0$, $\frac{1}{T_H} V = 0$. Since $T_H > 0$, we have $V = 0$.
- From $\frac{dH}{dt} = \frac{1}{T_H} V - \frac{1}{T_V} H = 0$, with $V = 0$, $H = 0$, this is satisfied.
- From $\frac{dF}{dt} = P_F \gamma I - \frac{1}{T_F} F = 0$, with $I = 0$, $F = 0$, this is satisfied.

Now, consider the equations for R_M, R_V, and R_F:

- $\frac{dR_M}{dt} = \frac{1}{T_M} M = 0$ since $M = 0$, so R_M is constant (denote it by R_M^*).
- $\frac{dR_V}{dt} = \frac{1}{T_V} H = 0$ since $H = 0$, so R_V is constant (denote it by R_V^*).
- $\frac{dR_F}{dt} = \frac{1}{T_F} F = 0$ since $F = 0$, so R_F is constant (denote it by R_F^*).

Finally, equation **??** ($-\beta IS = 0$) with $I = 0$ imposes no constraint on S, suggesting S can be any non-negative value (since S likely represents a population, $S \geq 0$), say S^*.

Thus, the steady states are:

$$(S, E, I, M, V, H, F, R_M, R_V, R_F) = (S^*, 0, 0, 0, 0, 0, 0, R_M^*, R_V^*, R_F^*)$$

where $S^* \geq 0$ and R_M^*, R_V^*, R_F^* are arbitrary constants.

We should also notice that the sum of all variables

$$S + E + I + M + V + H + F + R_M + R_V + R_F = N \tag{1}$$

remains constant in time and equals the total population at time 0. Thus we also obtain that

$$S^* + R_M^* + R_V^* + R_F^* = N.$$

The three recovered compartments incorporate the mass that has accumulated up to the moment the infection dies out.

8.2 Jacobian at a disease-free equilibrium

To find the eigenvalues, we compute the Jacobian matrix of the system at the steady state. Define the state vector as $(S, E, I, M, V, H, F, R_M, R_V, R_F)$, and the system as $\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x})$, where:

$$\mathbf{f} = \left(-\beta IS, \beta IS - \sigma E, \sigma E - \gamma I, P_M \gamma I - \frac{1}{T_M} M, P_V \gamma I - \frac{1}{T_H} V, \frac{1}{T_H} V - \frac{1}{T_V} H, P_F \gamma I - \frac{1}{T_F} F, \frac{1}{T_M} M, \frac{1}{T_V} H, \frac{1}{T_F} F \right)$$

The Jacobian $J = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ is a 10×10 matrix. We evaluate it at the steady state $(S^*, 0, 0, 0, 0, 0, 0, R_M^*, R_V^*, R_F^*)$.

With the above ordering of variables, the Jacobian $J(S_0) = \partial f / \partial x$ evaluated at any disease-free equilibrium is block–lower-triangular:

$$J = \begin{pmatrix}
0 & 0 & -\beta S^* & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\sigma & \beta S^* & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sigma & -\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & P_M\gamma & -\frac{1}{T_M} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & P_V\gamma & 0 & -\frac{1}{T_H} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{T_H} & -\frac{1}{T_V} & 0 & 0 & 0 & 0 \\
0 & 0 & P_F\gamma & 0 & 0 & 0 & -\frac{1}{T_F} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{T_M} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{T_V} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{T_F} & 0 & 0 & 0
\end{pmatrix}.$$

Since R_M, R_V, and R_F do not appear in the equations for other variables, the Jacobian is block triangular:

$$J = \begin{pmatrix} A & 0 \\ B & D \end{pmatrix}$$

- A: 7×7 submatrix for (S, E, I, M, V, H, F),
- D: 3×3 diagonal submatrix for (R_M, R_V, R_F) with entries $0, 0, 0$,
- B: coupling terms,
- 0: zero matrix.

The eigenvalues of J are the eigenvalues of A union those of D. Since $D = \text{diag}(0, 0, 0)$, it has three zero eigenvalues. For A:

$$A = \begin{pmatrix} 0 & 0 & -\beta S^* & 0 & 0 & 0 & 0 \\ 0 & -\sigma & \beta S^* & 0 & 0 & 0 & 0 \\ 0 & \sigma & -\gamma & 0 & 0 & 0 & 0 \\ 0 & 0 & P_M\gamma & -\frac{1}{T_M} & 0 & 0 & 0 \\ 0 & 0 & P_V\gamma & 0 & -\frac{1}{T_H} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{T_H} & -\frac{1}{clsT_V} & 0 \\ 0 & 0 & P_F\gamma & 0 & 0 & 0 & -\frac{1}{T_F} \end{pmatrix}$$

A is block triangular:

- First block: $A_1 = \begin{pmatrix} 0 & 0 & -\beta S^* \\ 0 & -\sigma & \beta S^* \\ 0 & \sigma & -\gamma \end{pmatrix}$,

- Second block: $-\frac{1}{T_M}$,

- Third block: $\begin{pmatrix} -\frac{1}{T_H} & 0 \\ \frac{1}{T_H} & -\frac{1}{T_V} \end{pmatrix}$,

- Fourth block: $-\frac{1}{T_F}$.

The eigenvalues of A are the union of the eigenvalues of these blocks.

- The first block, A_1:

$$A_1 - \lambda I = \begin{pmatrix} -\lambda & 0 & -\beta S^* \\ 0 & -\sigma - \lambda & \beta S^* \\ 0 & \sigma & -\gamma - \lambda \end{pmatrix}$$

The characteristic polynomial is:

$$(-\lambda)[(-\sigma - \lambda)(-\gamma - \lambda) - (\beta S^*)(\sigma)] = (-\lambda)[\lambda^2 + (\sigma + \gamma)\lambda + \sigma(\gamma - \beta S^*)] = 0$$

The roots of the characteristic polynomial are

$$\lambda_1 = 0$$

$$\lambda_{2,3} = \frac{-(\sigma + \gamma) \pm \sqrt{(\sigma + \gamma)^2 - 4\sigma(\gamma - \beta S^*)}}{2}$$

- The second block, for $-\frac{1}{T_M}$, gives:

$$\lambda = -\frac{1}{T_M}$$

- The third block, $\begin{pmatrix} -\frac{1}{T_H} & 0 \\ \frac{1}{T_H} & -\frac{1}{T_V} \end{pmatrix}$, has the characteristic polynomial:

$$\det\begin{pmatrix} -\frac{1}{T_H} - \lambda & 0 \\ \frac{1}{T_H} & -\frac{1}{T_V} - \lambda \end{pmatrix} = (-\frac{1}{T_H} - \lambda)(-\frac{1}{T_V} - \lambda) = 0$$

which gives the solutions

$$\lambda = -\frac{1}{T_H}$$

$$\lambda = -\frac{1}{T_V}$$

27

- The fourth block, For $-\frac{1}{T_F}$, gives

$$\lambda = -\frac{1}{T_F}$$

Hence, the eigenvalues of J are:

$$
\begin{array}{|l|}
\hline
\lambda = 0 \quad \text{(multiplicity 4: } S,\ R_M,\ R_V,\ R_F\text{)}, \\[2mm]
\lambda_{2,3} = -\dfrac{\gamma + \sigma}{2} \pm \dfrac{1}{2}\sqrt{(\gamma - \sigma)^2 + 4\beta\sigma S^*}, \\[2mm]
\lambda = -\dfrac{1}{T_M},\ -\dfrac{1}{T_F},\ -\dfrac{1}{T_H},\ -\dfrac{1}{T_V}. \\
\hline
\end{array}
$$

8.3 Interpretation

The system admits many equilibrium points; in fact, the set of critical points forms a 3-dimensional manifold, parameterized by the four "susceptible/recovered" variables.

$$(S^*,\ R_M^*,\ R_V^*,\ R_F^*)$$

subject to the overall population constraint

$$S^* + R_M^* + R_V^* + R_F^* + E^* + I^* + H^* = N.$$

Linearizing around any such point, the Jacobian has:

1. **Four zero eigenvalues** corresponding to (S, R_M, R_V, R_F). These modes are neutral (rate 0), reflecting the fact that once on the manifold of *susceptible* and *recovered*, there is no restoring force to push it away from it.

2. **Four strictly negative eigenvalues** for the fast compartments (M, V, H, F), implying these populations relax exponentially quickly onto the slow manifold.

3. **Two eigenvalues** $\lambda_{2,3}$ governing (E, I), with

$$\lambda_2 + \lambda_3 = -(\sigma + \gamma) \quad \text{and} \quad \lambda_2\,\lambda_3 = \sigma\,(\gamma - \beta\,S^*).$$

8.4 Stability of the model

This section is dedicated to proving that the main system of the paper is indeed globally stable. Here is the formal result.

Theorem 1. *The system* **??** *is globally stable for the parameters*

$$\beta, \sigma, \gamma, P_M, P_V, P_F, T_M, T_H, T_V, T_F > 0$$

and the initial conditions

$$S(0) > 0, I(0), E(0), M(0), V(0), H(0), F(0), R_M(0), R_V(0), R_F(0) \geq 0,$$

with at least $I(0) > 0$ *or* $E(0) > 0$.

Proof. Let us have a closer look at the subsystem corresponding to (S, E, I), being the system responsible for the stability of the model:

$$\begin{cases} \dfrac{dS}{dt} = -\beta IS \\[2mm] \dfrac{dE}{dt} = \beta IS - \sigma E \\[2mm] \dfrac{dI}{dt} = \sigma E - \gamma I \\[2mm] \text{with } E(0) \geq 0, I(0) \geq 0. \end{cases} \qquad (2)$$

The subsystem is locally Lipschitz on the positive region $\{(S, E, I) \in \mathbb{R}^3 : S \geq 0,\ E \geq 0,\ I \geq 0\}$, so solutions can be continued as long as they stay nonnegative and do not explode.

Knowing that $E \geq 0$, $I \geq 0$, $S \geq 0$ and $S(t) = S_0 e^{-\beta \int_0^t I(s)ds}$, we obtain $S(t) > 0$ for $I(s)$ well defined. From the initial conditions, we have $E(0) \geq 0$ and $I(0) \geq 0$ with $E(0) > 0$ or $I(0) > 0$.

Now, take a time t_0 such that $E(t_0) \geq 0$ or $I(t_0) \geq 0$. Then

$$e^{\sigma t} E(t) = e^{\sigma t_0} E(t_0) + \int_{t_0}^t \beta S(u) I(u) du$$

which implies that for $\epsilon > 0$ and $t \leq t_0 + \epsilon$ the inequality $E(t) > 0$ holds. Furthermore.

$$e^{\gamma t} I(t) = e^{\gamma t_0} I(t_0) + \sigma \int_{t_0}^t e^{\gamma u} E(u) du > 0, \text{ for } t \in [t_0, t_0 + \epsilon].$$

As a remark, as long as the solution exists and $I(0) \geq 0$ and $E(0) \geq 0$ with at least one of them being positive, for the whole time $t \geq 0$ we will have $E(t) > 0$ and $I(t) > 0$.

Using the summation of the equations of the subsystem **??**, we obtain:

$$\frac{d}{dt}(S(t) + E(t) + I(t)) = -\gamma I(t) \leq 0$$

So, $S(t) + E(t) + I(t) \leq S(0) + E(0) + I(0)$. In particular, this also implies that all quantities $S(t), E(t), I(t)$ are bounded for all $t \geq 0$, and this implies that the solution is defined for all $t \geq 0$.

Next, $S(t) = S(0) e^{-\beta \int_0^t I(u)du}$ and we can notice that $S(t)$ is a decreasing function; hence,

$$S(t) \xrightarrow{t \to \infty} S^*.$$

We know that $S(t) + E(t) + I(t)$ is also decreasing, thus $S(t) + E(t) + I(t)$ is convergent when $t \to \infty$ and automatically we can deduce that $E(t) + I(t)$ is convergent when $t \to \infty$.

Let us observe now that:

$$e^{\gamma t} I(t) = e^{\gamma t_0} I(t_0) + \int_{t_0}^t \sigma e^{\gamma u} E(u) du \geq e^{\gamma t_0} I(t_0)$$

Thus $e^{\gamma t_0} I(t_0) \geq \epsilon$ and for $t \leq t_0 + \frac{\ln(2)}{\gamma}$ we have $I(t) \geq \frac{\epsilon}{2}$. So, there exists $\epsilon > 0$ and a sequence $t_n \xrightarrow[n \to \infty]{} \infty$ with $t_{n+1} - t_n \to \infty$, such as $I(t_n) \geq \epsilon$. By taking $t - t_n \leq \frac{\ln(2)}{\gamma}$ we have that $I(t) \geq \frac{\epsilon}{2}$.

Therefore,

$$S(t) + E(t) + I(t) = S(t_n) + E(t_n) + I(t_n) - \gamma \int_{t_n}^t I(u) du$$

$$\leq S(t_n) + E(t_n) + I(t_n) - \gamma(t - t_n)\frac{\epsilon}{2}$$

For $t = \bar{t}_n = t_n + \frac{\ln(2)}{\gamma}$, we have:

$$S(\bar{t}_n) + E(\bar{t}_n) + I(\bar{t}_n) \leq S(t_n) + E(t_n) + I(t_n) - \frac{\epsilon \cdot \ln(2)}{2}.$$

So, because $t_{n+1} - t_n \to \infty$ we have

$$0 \leq S(t_{n+1}) + E(t_{n+1}) + I(t_{n+1}) \leq S(t_n) + E(t_n) + I(t_n) - \frac{\epsilon \cdot \ln(2)}{2}$$

that results in $S(t_n) + E(t_n) + I(t_n) \to -\infty$, which is a contradiction. Hence, $I(t) \xrightarrow{t \to \infty} 0$

Using the second equation of the system **??**, we have

$$e^{\sigma t} E(t) = e^{\sigma t_0} E(t_o) + \beta \int_{t_0}^{t} e^{\sigma u} I(u) S(u) du.$$

For $I(t) \xrightarrow{t \to \infty} 0$ and knowing $0 \leq S(u) \leq S(0) + E(0) + I(0)$ we obtain

$$0 \leq e^{\sigma t} E(t) \leq e^{\sigma t_\epsilon} E(t_\epsilon) + \beta \epsilon (S(0) + E(0) + I(0))(e^{\sigma t} - e^{\sigma t_\epsilon}), \text{ for } t \geq t_\epsilon \text{ with } |I(t)| \leq \epsilon.$$

This concludes that $E(t) \to 0$.

Next, from the first equation of the system **??**, we obtain

$$S(t) = S_0 e^{-\beta \int_0^t I(u) du}.$$

At the same time $S(t) + E(t) + I(t) = S(0) + E(0) + I(0) - \gamma \int_o^t I(u) du$, so

$$S(t) = S(0) e^{\frac{\beta}{\gamma}(S(t) + E(t) + I(t) - S(0) - E(0) - I(0))}$$

from which we deduce that

$$\frac{\beta}{\gamma} S(t) e^{-\frac{\beta}{\gamma} S(t)} = \frac{\beta}{\gamma} S(0) e^{-\frac{\beta}{\gamma}(E(t) + I(t) - S(0) - E(0) - I(0))}$$

obtaining, in the limit,

$$S^* = S_0 e^{\frac{\beta}{\gamma}(S^* - S(0) - E(0) - I(0))}.$$

Combining the last two equations of the system **??** gives:

$$\frac{d}{dt}(E(t) + I(t)) = (\beta S(t) - \gamma)I$$

We know that $\beta S(t) - \gamma \to \beta S^* - \gamma$. If $\beta S^* - \gamma \geq 0$ then $\beta S(t) - \gamma \geq 0$ because $S(t) \searrow S^*$, when $t \to \infty$. Hence, $\frac{d}{dt}(E(t) + I(t)) \geq 0$ and automatically $E(t) + I(t) \geq E(0) + I(0) > 0$ if $E(0) + I(0) > 0$, from which we conclude that $E(t) + I(t)$ does not converge to 0. The conclusion of this is that

$$\beta S^* < \gamma$$

and that, in regards to the stability of the main system **??** implies that the system is stable.

Thus, so far we proved that $S(t)$ converges, $E(t)$ and $I(t)$ converge to 0. Next, we will prove that the other quantities of the initial system also converge.

Let us consider a general equation:

$$\frac{dW}{dt} = aI - bW$$

with $a > 0, b > 0$ and $W(0) \geq 0$, then $W(t)$ converges to 0. Indeed, notice that

$$e^{bt}W(t) - a \int_0^t e^{bu}I(u)du = W(0), \text{ for all } t > 0.$$

Then, in the first place $W(t) \geq 0$ and moreover,

$$W(t) = e^{-bt}W(0) + ae^{-bt} \int_0^t e^{bu}I(u)du. \tag{3}$$

If $I(t) \xrightarrow{t \to \infty} 0$ we have:

$$\left| e^{-bt} \int_0^t e^{bu}I(u)du \right| \leq \left| e^{-bt} \int_0^{t_\epsilon} e^{bu}I(u)du \right| + e^{-bt} \int_0^t \epsilon \cdot e^{bu}du$$

$$\leq e^{-bt} \int_0^{t_\epsilon} e^{bu}|I(u)|du + \frac{\epsilon(1 - e^{-bt})}{b}$$

for $|I(t)| \leq \epsilon, \forall t \geq t_\epsilon$. Therefore, we can conclude that $W(t) \xrightarrow{t \to \infty} 0$.

This argument shows that all quantities M, V, H, F converge to 0.

Finally, if $\frac{dU}{dt} = W$ with $W(t) \geq 0$ and such that U is bounded, then U is convergent. From equation **??** we know that $W(t) \geq 0$, thus U is an increasing function, which combined with the boundedness of U guarantees that U converges.

In conclusion, the disease-free equilibrium is globally asymptotically stable,

$$\lim_{t \to \infty} I(t) = \lim_{t \to \infty} E(t) = 0, \quad \lim_{t \to \infty} S(t) = S^* < \frac{\gamma}{\beta},$$

and all other compartments driven by I likewise vanish or converge. $\qquad \square$