Pandemic Simulation and Forecasting of exit strategies: Convergence of Machine Learning and Epidemiological Models

Salah Ghamizi[1], Renaud Rwemalika[1], Maxime Cordy[1], Mike Papadakis[1], and Yves Le Traon[1]

[1]University of Luxembourg, SnT

May 2020

Abstract

The COVID-19 pandemic has created a public health emergency unprecedented in this century. The lack of accurate knowledge regarding the outcomes of the virus has made it challenging for policymakers to decide on appropriate countermeasures to mitigate its impact on society, in particular the public health and the very healthcare system.

While the mitigation strategies (including the lockdown) are getting lifted, understanding the current impacts of the outbreak remains challenging. This impedes any analysis and scheduling of measures required for the different countries to recover from the pandemic without risking a new outbreak.

Therefore, we propose a novel approach to build realistic data-driven pandemic simulation and forecasting models to support policymakers. Our models allow the investigation of mitigation/recovery measures and their impact. Thereby, they enable appropriate planning of those measures, with the aim to optimize their societal benefits.

Our approach relies on a combination of machine learning and classical epidemiological models, circumventing the respective limitations of these techniques to allow a policy-making based on established knowledge, yet driven by factual data, and tailored to each country's specific context.

1 Introduction

On December 31st, 2019 a pneumonia of unknown cause was detected in Wuhan, China. It was later reported to be the first confirmed instance of the Coronavirus disease 2019 (COVID-19). It is caused by a betacoronavirus in the same subgenus as the severe acute respiratory syndrome (SARS) virus designated sever acute respiratory syndrome coronavirus 2 (SARS-CoV-2). A month after the first case was reported, the outbreak was declared a public health emergency of international concern by the World Health Organization (WHO). On March 11th, WHO assessed that the outbreak of COVID-19 can be characterized as a pandemic [1]. As of May 8th, 2020, date of the writing of this report, there are about four millions confirmed cases in 215 countries accounting for a death toll over 250 000 death world wide [2].

In face of the rapid spread of the COVID-19 and the non-existence of a vaccine, decision-makers rely on social distancing regulation and other non-pharmaceutical intervention. Indeed, a report published in March 2020 [3], WHO is describing the coordinated global research road map for the development of a virus, shows estimates first clinical trials around November 2020. Many countries worldwide have been applying social distancing and lock down measures in order to control the spread of COVID-19, and avoid a peak of cases leading to the over saturation of healthcare facilities. However, they face uncertainty, not only because of the epidemiological characteristic of SARS-CoV-2, but also with program uncertainties related to feasibility and effectiveness of mitigation strategies and their worldwide socio-economic impacts. Despite these uncertainties decisions must be made, hence the need of tools to help the decision process.

In this context, mathematical modelling offers public health planners frameworks to make predictions about the spread of emerging diseases and assess the impact of possible mitigation strategies. This is particularly important when dealing with infectious diseases, such as COVID-19 where mass interventions *e.g* screening, social distancing and vaccination can lead to effects at a population level, including herd immunity, changes in the infection rate or even changes in the pathogen ecology as consequence of selective pressure.

Researchers based their models on two types of approaches: static cohort models and transmission dynamic models [4]. Static models, typically relying on decision trees and Markov processes, assume a force

of infection that is independent of the proportion of the population that is infected and therefore is of little use in response of highly infectious diseases like COVID-19. Transmission dynamic model on the other hand see a force of infection varying depending on the proportion of the population which is infected. Compared with static cohort models, transmission dynamic models are usually more complex to parametrize requiring epidemiological information on the infectious disease and demographic and economic information from the population affected.

Different approaches exist to implement dynamic approaches. Agent Based Model (ABM) are simulations composed of agent that can interact with each other and their environment. Because each agent can makes its own rules, this type of approach can capture aggregate phenomena derived from the behavior of single agents. These models offer a great explainability of the root causes leading to the propagation of a disease, but are costly to develop and are typically region specific. Indeed, the behaviour and the interaction of each type of agent needs to be fully defined in order for the model to be useful. These rules are typically really case specific and are not applicable from region to region.

Maybe the most common approach to model the spread of infectious disease is the suceptible-infected-removed (SIR) model and its elaborations *i.e* SIRD, MSIR, SIER, SEIS, MSEIR and MSEIRS. They are equation based where each equation defines the rate to go from one state to the other. The SEIR model separates the population into four groups, susceptible, exposed, infected and removed and simulates the time evolution of each of these subpopulation and the transition rate are defined by the time scale to which an individual can transmit the disease, the time to recovery and the number of newly infected people due to one infected individual.

These methods are faced scalability issues, in the case of ABM, and are really dependent on the representativeness of the input parameters and data, for both SEIR and ABMs. In this technical report, we propose an approach to alleviate these constraints in transmission dynamic model by relying on available large dataset and machine learning to estimate the parameters needed for SEIR models.

More recently, to circumvent the limitation of these approaches, recently, researchers started to take advantage of the advances made in machine learning in order to create models based on large datasets available [5,6]. In this work we call these approaches ML based approaches. Our work falls in ML approaches family. The model is relying on a SEIR model to predicted the evolution of the subpopulation (susceptible, exposed, infectious and removed) over time. While SEIR is a very powerful model, it presents two limitations (1) it requires hyper-parameters that are hard to observe such as the infection rate of an individual, and (2) unlike agent based models, evaluating scenarios can be difficult given the input parameters of the model. To tackle these limitations, our approach combines the SEIR with a machine learning predictor, based on a deep learning model, to estimate the effective reproductive number (R_t) over time. This number represent the number of people that can be contaminated by an infectious individual over a period of time. The machine learning predictor relies on demography and mobility features to predict an effective reproductive number. For each time increment, R_t is updated by our approach. This behavior allows to elaborate strategies which can be fed to our approach and observe the impact in term of number of cases over time.

2 Methodology

Our approach presented in Figure 1 combines both an epidemiological model (SEIR) and a Deep learning predictor. While the first one can be fitted to the past data using optimization algorithms, the second is trained to learn how to predict the the SEIR hyper-parameters (mainly the reproduction rate R_t) based on the population behaviour. The fitted SEIR model provides the initial training and validation dataset of the ML model, while the second is used to forecast the future parameters to be used by the SEIR model. These parameters are used to forecast the future cases, hospitalized and deaths for different scenarios.

2.1 Epidemiological Model

We rely on the SEI-HCRD compartmental model - Susceptible (S) $\rightarrow$ Exposed (E) $\rightarrow$ Infectious (I) $\rightarrow$ Removed (Hospitalized (H), Critical (C), Recovered (R), Dead (D)). We describe the model with the following system of ordinary differential equations:

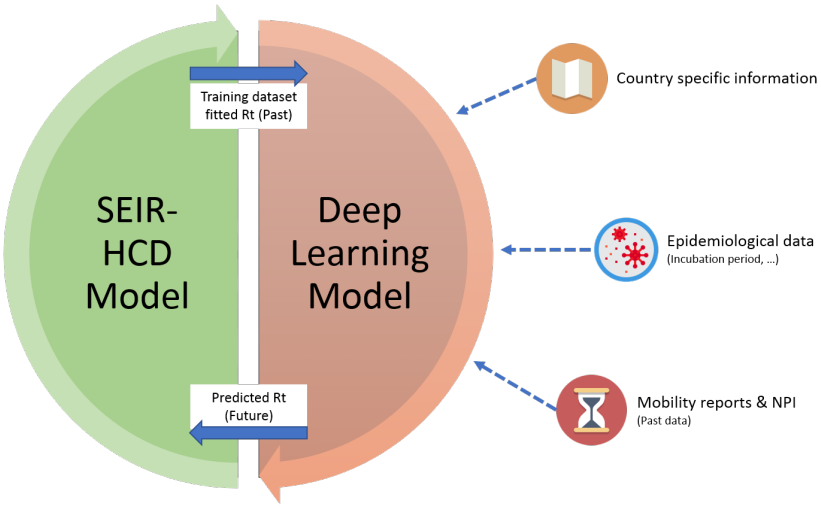


Figure 1: Our proposed approach combines both a Deep Learning predictor and an epidemiological model to predict the country-specific reproduction rate based on population mobility behaviours.

$$\frac{dS}{dt} = -\frac{R_t}{t_{inf}} \cdot IS \tag{1}$$

$$\frac{dE}{dt} = \frac{R_t}{t_{inf}} \cdot IS - \frac{1}{t_{inc}} \cdot E \tag{2}$$

$$\frac{dI}{dt} = \frac{1}{t_{inc}} \cdot E - \frac{1}{t_{inf}} \cdot I \tag{3}$$

$$\frac{dH}{dt} = \frac{1-m}{t_{inf}} \cdot I + \frac{1-f}{t_{crit}} \cdot C - \frac{1}{t_{hosp}} \cdot H \tag{4}$$

$$\frac{dC}{dt} = \frac{c}{t_{hosp}} \cdot H - \frac{1}{t_{crit}} \cdot C \tag{5}$$

$$\frac{dR}{dt} = \frac{m}{t_{inf}} \cdot I + \frac{1-c}{t_{hosp}} \cdot H \tag{6}$$

$$\frac{dD}{dt} = \frac{f}{t_{crit}} \cdot C \tag{7}$$

where $S + E + I + H + C + R + D$ is equal to the total population. R_t is the reproduction number over time. The parameters t_{suffix} denote the transition time estimated to transit from one population to the other. t_{inc} is the average incubation period, t_{inf} is the average infectious period, t_{hosp} is the average time a patient is in hospital before either recovering or becoming critical and t_{crit} is the average time a patient is in a critical state before either recovering or dying. Finally, the last set of parameters in this system of equation, m, c and f, regulate the proportion of contamination that are either asymptomatic or mild(m), the proportion of severe cases that turn critical (c) and finally, the proportion of critical cases that are fatal (f).

In this work all the transition time parameters are considered constant. We fix their values based on the number reported by Liu *et al.* in the Journal of Travel medicine [7].

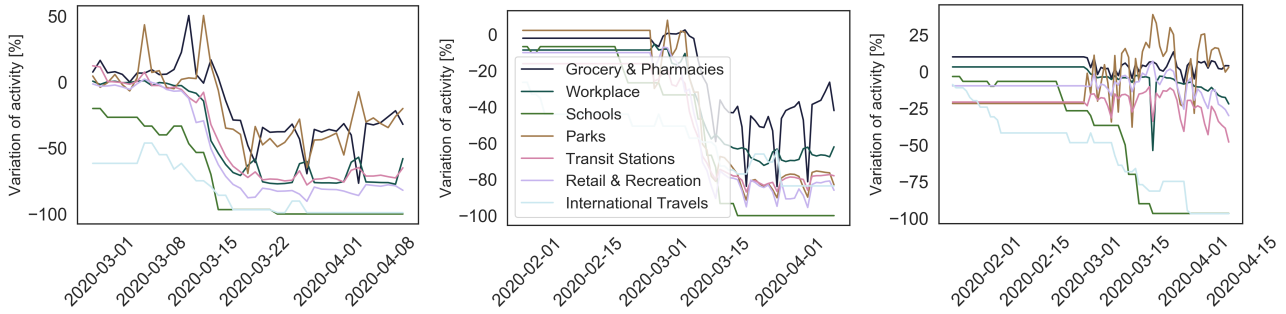
Finally, for parameters m, c and f, we use real-time cases and death data from the ECDC (European Centre of Disease Control) and WHO to fit the model. It is worth noting that both the case and death data are highly unrepresentative of the actual number of infections due to under reporting as well as country specific testing strategy. Similarly, post-mortem testing and reporting is not systematic and it is also underestimated. Therefore, we take into account both information in our model. Table 1 depicts the values for the parameters that were used to build the model.

2.2 Machine Learning Model

Our objective is to predict the reproduction rate Rt based solely on time-independent inputs. Given a 33 features vector, our model is expected to predict the associated Rt. The dataset is large enough to be handled by a multi-layer neural network.

Parameter	Value
t_{inc}	5.6 days
t_{inf}	2.9 days
t_{hosp}	4 days
t_{crit}	14 days
m	80%
c	10%
f	30%

Table 1: Value for the parameters used in the SEIHCRD model.



(a) Luxembourg (b) Italy (c) Japan

Figure 2: Evolution of the mobility indicators for Luxembourg, Italy and Japan. A value of 0 means that the activity is at the same level as before the confinement, a value of -100% is a total stop of the activity and a positive values shows an increase of the activity compared to the referrence value.

2.2.1 Features for the machine learning model

We start from Google Mobility Reports, that track for 105 countries the movement trends over time, across different categories of places. For each feature, the value indicates the relative daily difference to the Baseline: the median value, for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020.

According to their reports[1] the available features are:

- Grocery & pharmacies:

  Mobility trends for places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies.

- Parks

  Mobility trends for places like local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens.

- Transit stations

  Mobility trends for places like public transport hubs such as subway, bus, and train stations.

- Retail & recreation

  Mobility trends for places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters.

- Residential

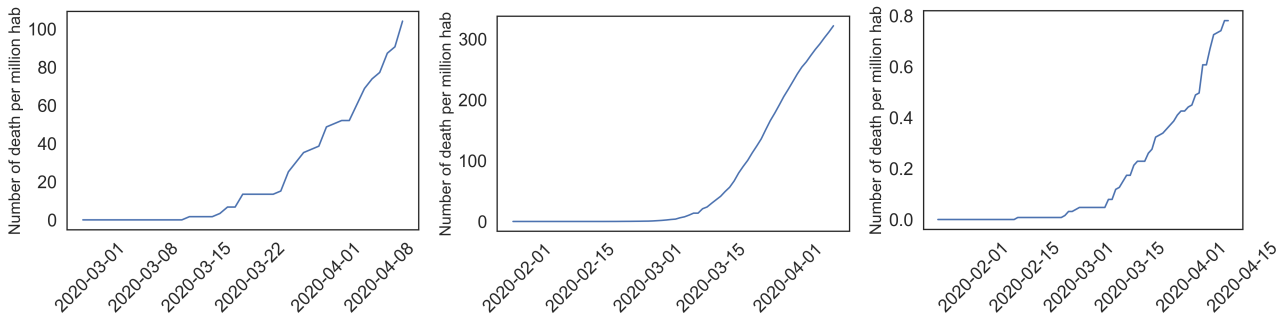
  Mobility trends for places of residence.

- Workplaces

  Mobility trends for places of work.

We augmented this dataset with two additional features extract from the Oxford dataset XXX, related to the school activity and international activity. All the results for thress countries, *i.e*, Luxembourg, Italy and Japan,

[1]https://www.google.com/covid19/mobility/



(a) Luxembourg (b) Italy (c) Japan

Figure 3: Evolution of the number of cumulative deaths for Luxembourg, Italy and Japan. For each point, this number represents the total number of deaths attributed to COVID-19 for each country.

Feature category	Feature	Smoothing (days)	Values
Mobility	Grocery/pharmacy	5,10,15,30 days	[-100,0]
	Parks	5,10,15,30 days	[-100,0]
	Residential	5,10,15,30	[-100,0]
	Retail/recreation	5,10,15,30	[-100,0]
	Transit stations	5,10,15,30	[-100,0]
	Workplace	5,10,15,30	[-100,0]
NPI	International travel control	7	[-100,0]
	School opening	7	[-100,0]
Demographics	GDP (Gross domestic product)	-	$\mathbb{N}$
	Population	-	$\mathbb{N}$
	Density	-	$\mathbb{R}$
	Area	-	$\mathbb{N}$
	P14 (rate of population under 15yrs)	-	[0,1]
	P65 (rate of population over 64yrs)	-	[0,1]
	Day (Day of week)	-	0,1,2,3,4,5,6

Table 2: Features of our Deep Learning model

are presented in Figure 2. The figure shows that not all countries adopted the same mitigation strategies to fight the spread of COVID-19. Where Italy and Luxembourg see their activities drastically reduced, Japan does not show a significant reduction of its activity except for the case of Schools and International travels.

Pushing our analysis further, we investigate whether mobility measures are sufficient to explain the impact of COVID-19. Looking at Figure 3, we see that Japan sees a similar trend in the number of cases, but that the numbers of death per millions of inhabitants are much lowers, while as Figure 2c it did not apply strong mitigation strategies in terms of mobility. This result can be explain by other demographic and cultural factors such as the density of the population, the hygiene habits of the population (wearing a mask in crowded environment), the quality of healthcare facilities and so on. To account for these variations from country to country, we included in our feature set data related to the demography of each country and the area of the countries in order to try to capture local culture trends.

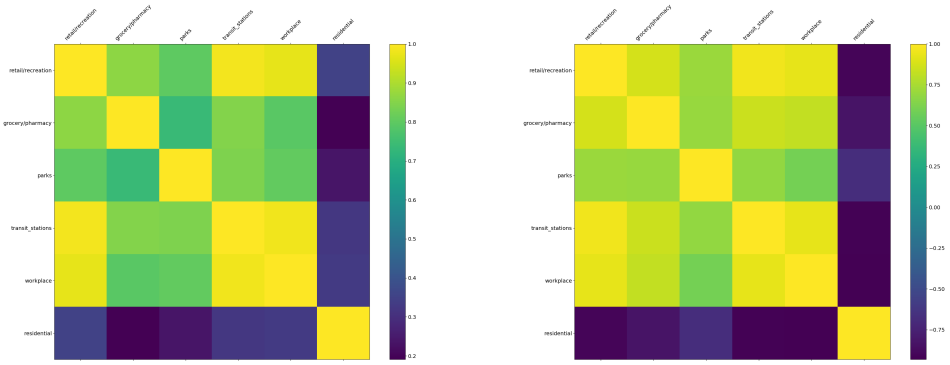
All these features are cleaned, and for the missing days, interpolated between the closest days with available information. The features are then augmented using downsampling and smoothing (over 5, 10, 15 and 30 days).

We complete our dataset with demographic features like population, density ... and with the index of the current day to take into account the weekly fluctuatioons of the data

Our feature engineering led to a 5,000 inputs of 34 features each, recorded for 105 countries from February 17th to April 14th. Table 2 summarize the features

2.2.2 Training & validation

The model can be seen as a supervised predictor taking as input mobility and demography features to predict a R_t value. Thus, in order to be trained, we need to produce a ground truth, R_t for each country for each day in the training period. To do so, we model the reproduction rate with a Hill decaying function to take



(a) Luxembourg (b) All Countries

Figure 4: Mobility feature correlation.

into account the mitigation measures that have been enacted in during the training period. Hill decay uses 2 parameters k & L, where L is the rate of decay, i.e. time to half decay and k is a shape parameter. The function takes the following form:

$$R_t = \frac{1}{(1 + (\frac{t}{L})^k)} \cdot R_0$$

Here we selected the Hill decay function because it showed good results in the literature. However, other decay such as Exponential decay and Weibull decay, linear decay and smooth-compact decay are also available. We provide in annex (Figure 11 a comparaison of the behavior of different decay functions.

In order to choose the best set of hyperparameter to model the past behaviour of COVID-19, we use an L-BFGs optimization equation to find for each country the set of parameters (k,L, m, c, f) that allow the SEIR model to fit the best their number of declared cases and deaths (to minimize the mean of Mean Square Error (MSE) over cases and deaths). Using this initial approach we have for each country a time serie of the past values of R_t.

Once the training set with the expected values is established, we train the model on all the countries except the target country and compare the predictions with the observed number of cases and deaths. Indeed, Figure 4a shows that when working on a country basis, all the mobility features can be highly correlated (over 0.75). For instance in Luxembourg where all the activity reduction and closure were enacted in most sectors at the same time. Being collinear, it will be hard to distinguish their individual impact unless we are capable of learning the feature interactions across many countries, where some have a slightly different behaviour. Figure 4b shows a much more nuanced picture. The features are still correlated when studied across all the countries of our dataset, however the strength of correlation is much lower (mid to high), showing the importance of keeping all the mobility features within the model. However, instead of training the model country by country, it is trained using all the countries at once.

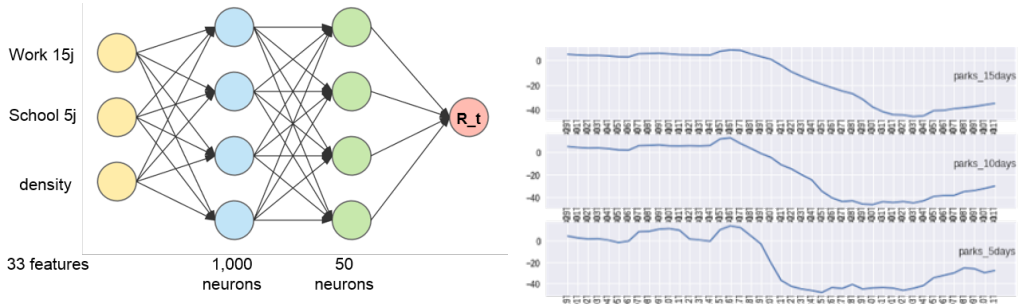
Finally, to optimize the performances of the model, we use a grid search and cross validation to select the architecture and hyper parameters that lead to the lowest MSE. The preferred architecture contains 2 hiddens layers, with respectively 1,000 and 50 neurons. We use Adam optimization and early stopping in the training process.

2.3 Intervention scenarios

We investigated a range of non-pharmaceutical interventions that can be associated with the feature of the model, which we assume would impact the Reproduction Rate:

School opening: This measure impacts both the feature “Public Transit” and the “School Activity” feature Business-focused activity opening: This measure impacts both the features “Public Transit” and “Workplace” Leisure activity opening: This measure impacts the features “Public Transit”, “Retail and recreation”, “Workplace” and “Parks”

For each intervention, we evaluate the impact of going back to Baseline values, i.e. the values before the non-pharmaceutical interventions were implemented.



(a) 2 layers neural network

(b) The "Park" feature smoothed over 3 ranges

Figure 5: Our Deep Learning model and a sampled feature

3 Motivations & Research questions

Our end goal is to provide decision-makers with a tool to easily generate exit strategies and scenarios and evaluate their impact. To achieve this, we rely on a synergetic approach that that uses a deep neural network as proxy to evaluate the hyper parameters of a SEIR Model based on mobility and NPI data.

In view of this, our first research question is:

> **RQ1:** Can a SEIHCRD model fitted on the past number of cases & deaths provides sufficiently accurate model of the covid19 disease over our 15 countries?

Having developped a sufficiently good model of the past behaviours, we turn our focus to the possibility to predict the reproduction rate that has been previously fitted with objective data. We rely as feature on demographic information of the countries, on mobility data and on the history of Non-pharmaceutical measures that have been implemented. Our second research question is thus as follow:

> **RQ2:** Can we predict the Reproduction rate based on mobility and NPI (Non-pharmaceutical interventions) data?

Ultimately our model is intended to be used to evaluate and select NPI scenarios and schedules that will allow the policy-makers to evaluate the impact of their scenario. This means that our approach should be able to model existing and popular exit strategies and also be available and easily customizable by the end users.

We evaluate this by predicting the impact of various scenarios and discussing some approaches to recommend actionnable scenarios. Thus, we ask the following question:

> **RQ3:** Can we build reliable non pharmaceutical intervention and exit scenarios?

4 Results

4.1 Short term fitting of SEIR-HCD models

We fitted the hyper-parameters of our SEIR-HCD model to the observed values of cases and deaths and evaluated the RMSE of the regression. The best fitted values for each country are presented in table 3. We evaluted the performance of the fit using the Root Mean Square Error (RMSE).

The countries that have beeen best fitted are Italy, Greece and Spain (with respectively 0.03, 0.04 and 0.05 error), which indicates that the decrease of Reproduction Rate of these countries can be best fitted with a Hill Decay function. It is worth nothing that the values found by our optimization algorithm for the hyper-parameters t_{hosp}, t_{crit}, m, c and f are far from the reference values (4 days, 14 days, 0.8, 0.1 and 0.3 respectively). While in Luxembourg 94% of the infections are mild (did not require hospitalization), it drops to 50

Figure 6 shows a vizualization of the cases and deaths predicted for Luxembourg. In appendix, we provide a similar visualization of all of the countries.

Country	R_0	t_{hosp}	t_{crit}	m	c	f	k	L	RMSE
Belgium	3.02	9.39	14.43	0.50	1.00	1.00	5.0	58.07	0.22
France	3.71	4.06	13.99	0.50	0.98	0.57	1.56	50.36	0.17
Germany	3.39	10.00	15.14	0.56	0.66	0.27	5.0	56.0	0.15
Greece	5.44	0.5	2.09	0.50	0.11	0.98	2.84	18.77	0.04
Italy	7.93	0.5	2.58	0.85	1.00	1.00	1.41	15.03	0.03
Latvia	3.94	10.00	20.00	0.50	1.00	0.05	5.00	23.41	0.16
Luxembourg	6.70	0.5	6.39	0.94	0.43	1.00	3.36	16.42	0.06
Netherlands	9.83	0.51	2.00	0.87	1.00	1.00	2.05	12.96	0.09
Spain	3.90	0.58	2.00	0.50	0.24	1.00	4.94	48.26	0.05
Switzerland	7.70	0.56	7.67	0.94	0.99	0.96	3.74	18.90	0.06
Brazil	6.15	0.51	8.73	0.51	0.28	1.00	1.84	23.45	0.35
Cameroon	4.43	1.56	8.71	0.52	0.14	0.92	2.26	21.80	0.26
Canada	3.07	10.00	19.18	0.65	0.99	0.58	1.00	95.16	0.43
Japan	3.07	4.28	14.11	0.53	1.00	0.07	1.00	50.42	0.17
United Kingdom	3.73	7.2	14.25	0.50	1.00	0.99	1.32	52.29	0.29

Table 3: Hyper-parameters and performance of our best fitted SEIR-HCD regression

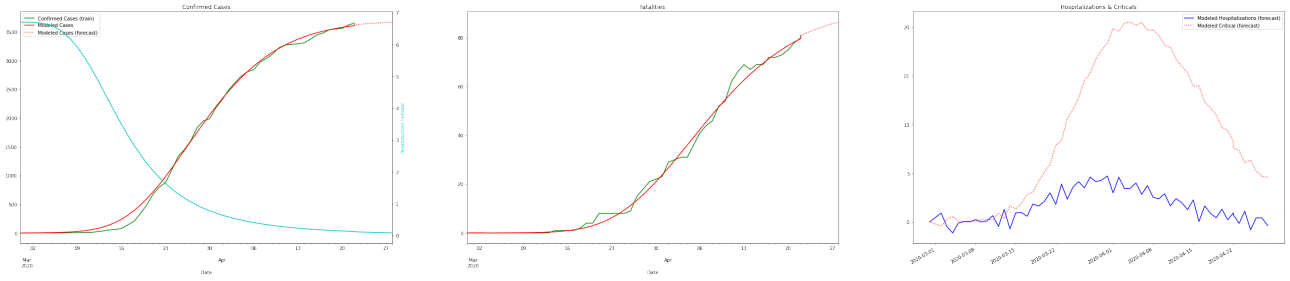


Figure 6: Predicted cases, hospitalizations, critical and deaths of our fitted model for Luxembourg.

4.2 Short-term predictions of using the DN-SEIR model

Performance metrics of the DN-SEIR model After the identification of the best hyper-parameters with cross validation, we train again on all countries and evaluate on our 10 reference countries. Both the training and test were evaluated on the same time span (February 22nd - April 11th). On the test set, the performance metrics are:

- R^2 score: 0.98
- Mean absolute error : 0.19
- Mean squared error: 0.01
- Median absolute error: 0.15
- RMSE: 0.43

Comparison of a fitted SEIR model and our DN-SEIR model We show in figure 7 the fitted Reproduction rate of the 10 European countries. We evaluate the confidence interval by evaluating the mean and standard deviation of a Bayesian Ridge Regressor on each element of the test set (using the same training set as the Neural Network). We use a grid search over 3 values for each of its 2 hyper-parameters α_{init} (0.1, 1, 1.9) and λ_{init} (1, 0.1, 0.01).

Belgium, Germany and Spain have all an initial plateau before the Reproduction Rate decreases. However, while Germany's reproduction rate drops under 1 by April 11th (end of data), Spain and Belgium still have a reproduction rate that may cause lead to an second outbreak.

We compare in table 4 the predicted cases of the time-dependant SEIR (Fitted on past cases/deaths) and the DN-SEIR approach. The DN-SEIR approach has a lower error to the ground truth values of cases for 10 over 15 countries in comparison with a time-dependant approach. Besies, the ground truth falls within the confidence interval of the DN-SEIR model for 10 of the 15 countries evaluated. It is worth nothing that Canada, France and Japan are the ones display a relative high rate of error (19%, 24% and 33% respectively), while 6/15 countries lie under 5% error and 12 over 15 countries do no exceed 15% error.

The prediction of deaths is trickier. Our approach assume that the critical rates and the mortality rates of the infected remains stable, while it can only be decreasing with the increase of the testing capacity and better handling of the infected.

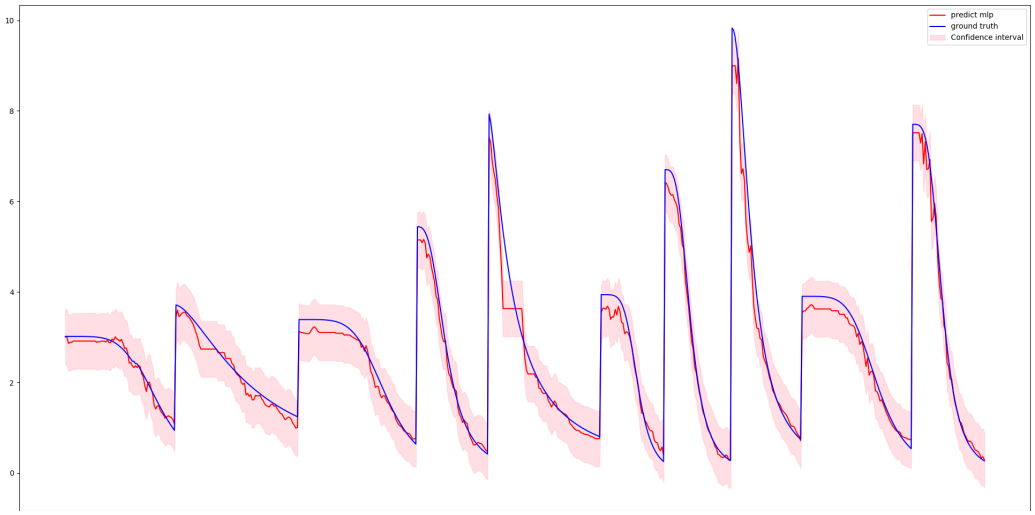


Figure 7: Predicted Reproduction Rate Rt using the mobility and demographic features for 10 countries. From left to right: Belgium, France, Germany, Greece, Italy, Latvia, Luxembourg, Netherlands, Spain and Switzerland.

Country	SEIR cases	DN-SEIR cases	True cases	DN-SEIR ϵ_r
Belgium	47,219	51,532 [45,130-60,464]	47,859	0.07
France	213,103	191,188 [166,043-224,554]	128,442	0.33
Germany	161,383	181,438 [163,333-207,151]	157,641	0.13
Greece	2,429	2,478 [2,343-2,666]	2,576	0.03
Italy	201,802	202,369 [187,205-223,194]	203,591	¡0.01
Latvia	767	769 [748-821]	858	0.12
Luxembourg	3,668	3,724 [3,578-3,936]	3,769	0.01
Netherlands	37,371	36,216 [32,980-40,767]	38,802	0.07
Spain	209,646	230,794 [211,299-257,782]	240,743	0.04
Switzerland	28,412	28,244 [27,684-29,601]	29,407	0.04
Brazil	60,714	68,271 [55,268-86,164]	78,162	0.14
Cameroon	1,432	1,645 [1,375-2,003]	1,832	0.11
Canada	77,614	63,520 [51,977-78,913]	51,597	0.19
Japan	12,250	11,353 [10,010-13,288]	14,088	0.24
United Kingdom	201,701	160,442 [134,251-188,088]	165,221	0.03

Table 4: Total cases on April 29th as predicted by the fitted SEIR model, by the DN-SEIR model and the actual cases. Both models trained until April 11th. The last column is the absolute relative error of the DN-SEIR model. In bold, the countries where the DN-SEIR approach beats the SEIR time-fitted approach

Averted deaths In order to evaluate the impact of the non-pharmaceutical interventions, we compare how the disease would have spread in each country if the Reproduction Rate stayed the same as of February 15th. We compare the numbers of deaths predicted by the SEIR model at this reproduction rate, the deaths predicted with our DN-SEIR deaths and the actual deaths. We focused on 4 european countries and we provide in appendix values of the remaining countries.

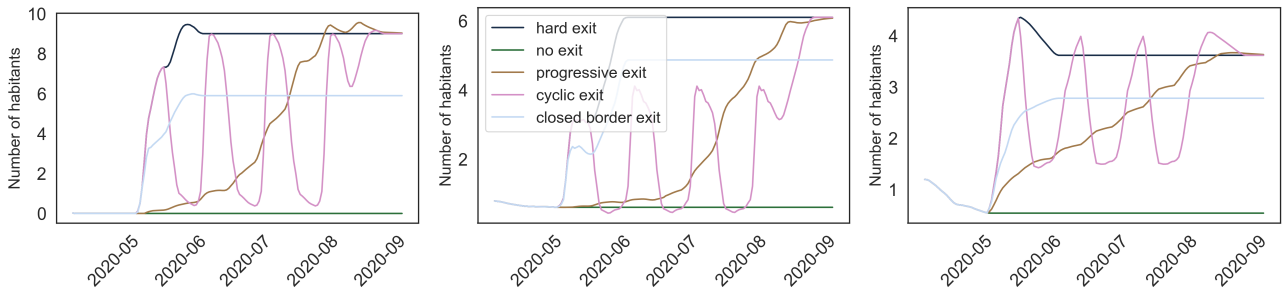
The DN-SEIR model predicts relatively well the death cases of the countries, for instance with less than 6% error on France deaths. It also shows that 4,889 deaths have been averted in France thanks to the NPI that have been undertaken. In Italy, more than 200,000 deaths have been averted thanks to the mitigation measures.

4.3 Mid-term predictions with exit strategies

In this section we investigate four prediction scenarios for an exit from the lockdown that is current taking place all over the world. The goal of the exit strategies is to allow a return to normal activities while mini-

Country	SEIR deaths ($R_t of 15-02-2020$)	DN-SEIR deaths (with NPIs)	True deaths	DN-SEIR ϵ_r
France	30,382	25,493 [25,314-25,695]	24,087	0.06
Germany	48,724	7,293 [7,242-7,353]	6,115	0.16
Italy	230,989	22,476 [22,348-22,623]	27,682	0.19
Spain	81,347	19,449 [19,338-19,577]	24,543	0.21

Table 5: Total deaths on April 29th as predicted by the DN-SEIR model and the actual deaths. We compare the numbers with the predicted one by a SEIR model if the Reproduction rate stayed the same as of 15th.



(a) Luxembourg (b) Italy (c) Japan

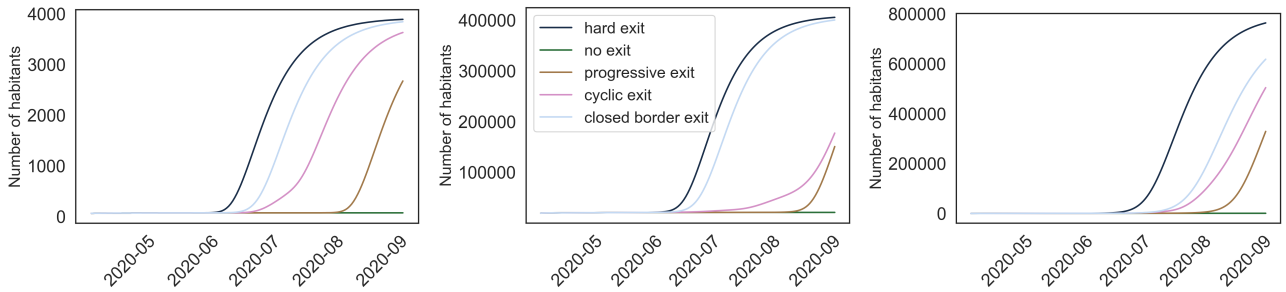
Figure 8: Evolution of the R_t values for the four exit strategies modelled, *i.e* a hard exit, a progressive exit, a cyclic exit and no exit in Luxembourg, Italy and Japan.

mizing the impact on the number of death and avoiding peaks in hospitalization that would saturate current healthcare facilities. For all the scenario, the starting point is the mobility activity recorded by Google Mobility Reports presented in section 2.2.1. The scenarios that we are investigating are the following:

- **Hard exit**: In this scenario, all mobility activities are resumed to normal on May 11, 2020.
- **Progressibe exit**: In this scenario, mobility activities are gradually restored, with an increase of 15% of the activity every 2 weeks until normal activity is resumed (the activity record for the period before lockdown).
- **Cyclic exit**: Every two weeks, activity is resumed to normal then brought back to lockdown situation. The process is repeated for four cycles, thus ending on August 3, 2020.
- **No exit**: The current situation is maintaind for the entirety of the period under evaluation.

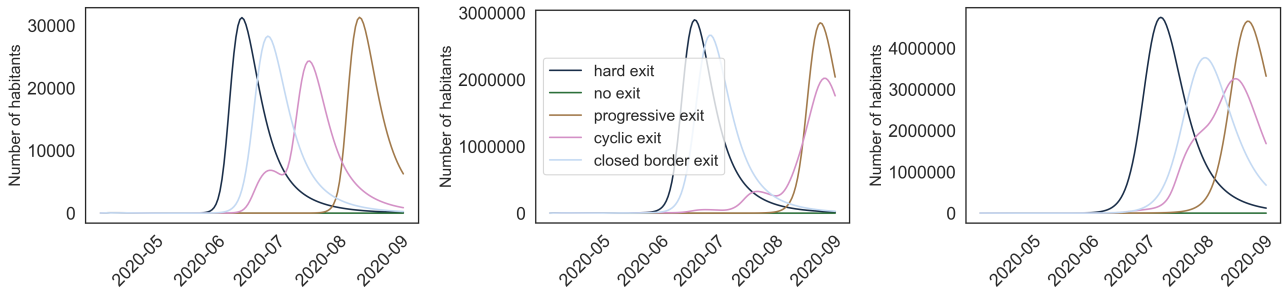
Figure 8 shows the evolution of R_t values for Luxembourg, Italy and Japan. As expected we see no evolution from the initial value in the no exit case, a cyclic fluctuation in the case of a cyclic exit, an soft increase of the R_t values in the when applying a progressive exit and finally, a brutal jump in the case of a hard exit.

Applying these scenarios on the three countries Luxembourg, Italy and Japan we obtain the results depicted by Figure 9 for the prediction of death due to COVID-19 and Figure 10 for the number of people in hospitals treated for COVID-19.



(a) Luxembourg (b) Italy (c) Japan

Figure 9: Evolution of the cumulative number of death for the four exit strategies modelled, *i.e* a hard exit, a progressive exit, a cyclic exit and no exit in Luxembourg, Italy and Japan.



(a) Luxembourg (b) Italy (c) Japan

Figure 10: Evolution of the number of people hospitalized for the four exit strategies modelled, *i.e* a hard exit, a progressive exit, a cyclic exit and no exit in Luxembourg, Italy and Japan.

References

[1] W. H. Organization, "Rolling updates on coronavirus disease (covid-19)," 2020.

[2] W. H. Organization, "Coronavirus disease (covid-19) pandemic," 2020.

[3] World Health Organisation, *a Coordinated Global Research Roadmap: 2019 Novel Coronavirus*. No. March, 2020.

[4] M. Jit and M. Brisson, "Modelling the Epidemiology of Infectious Diseases for Decision Analysis," *PharmacoEconomics*, vol. 29, pp. 371–386, may 2011.

[5] M. A. C. Vollmer, S. Mishra, H. J. T. Unwin, A. Gandy, T. A. Mellan, H. Zhu, H. Coupland, I. Hawryluk, M. Hutchinson, O. Ratmann, P. Walker, C. Whittaker, L. Cattarino, C. Ciavarella, L. Cilloni, M. Baguelin, S. Bhatia, A. Boonyasiri, N. Brazeau, G. Charles, V. Cooper, Z. Cucunuba, G. Cuomo-dannenburg, A. Dighe, B. Djaafara, J. Eaton, L. V. Elsland, R. Fitzjohn, K. Fraser, K. Gaythorpe, W. Green, S. Hayes, N. Imai, E. Knock, D. Laydon, J. Lees, T. Mangal, A. Mousa, G. Nedjati-gilani, P. Nouvellet, D. Olivera, K. V. Parag, M. Pickles, H. A. Thompson, R. Verity, H. Wang, Y. Wang, O. J. Watson, L. Whittles, X. Xi, and A. Ghani, "Report 20 : Using mobility to estimate the transmission intensity of COVID-19 in Italy : A subnational analysis with future scenarios," Tech. Rep. May, Imperial College COVID-19 Response Team, 2020.

[6] N. Soures, D. Chambers, Z. Carmichael, A. Daram, D. P. Shah, K. Clark, L. Potter, and D. Kudithipudi, "SIRNet: Understanding Social Distancing Measures with Hybrid Neural Network Model for COVID-19 Infectious Spread," tech. rep., University of Texas, apr 2020.

[7] Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv, "The reproductive number of COVID-19 is higher compared to SARS coronavirus," *Journal of Travel Medicine*, vol. 27, pp. 1–4, mar 2020.

5 Appendix

Country	SEIR deaths (R_t as of 15-02-2020)	DN-SEIR deaths (with NPIs)	True deaths
Belgium	2,180	7,049 [6,996-7,112]	7,501
France	30,382	25,493 [25,314-25,695]	24,087
Germany	48,724	7,293 [7,242-7,353]	6,115
Greece	107	107 [106-108]	139
Italy	230,989	22,476 [22,348-22,623]	27,682
Latvia	32	32 [32-33]	15
Luxembourg	78	78 [78-79]	89
Netherlands	3,364	3,348 [3,317-3,384]	4,711
Spain	81,347	19,449 [19,338-19,577]	24,543
Switzerland	1,391	1,390 [1,383-1,399]	1,716
Brazil	2.358	2,359 [2,312-2,414]	5,466
Cameroon	30	30 [29-31]	61
Canada	1,106	2,047 [2,011-2,088]	2,996
Japan	363	505 [498-512]	415
United Kingdom	22,703	19,882 [19,659-20,086]	26,771

Table 6: Total deaths on April 29th as predicted by the DN-SEIR model and the actual deahs. We compare the numbers with the predicted one by a SEIR model using the Reproduction rate of each country on February 15th.

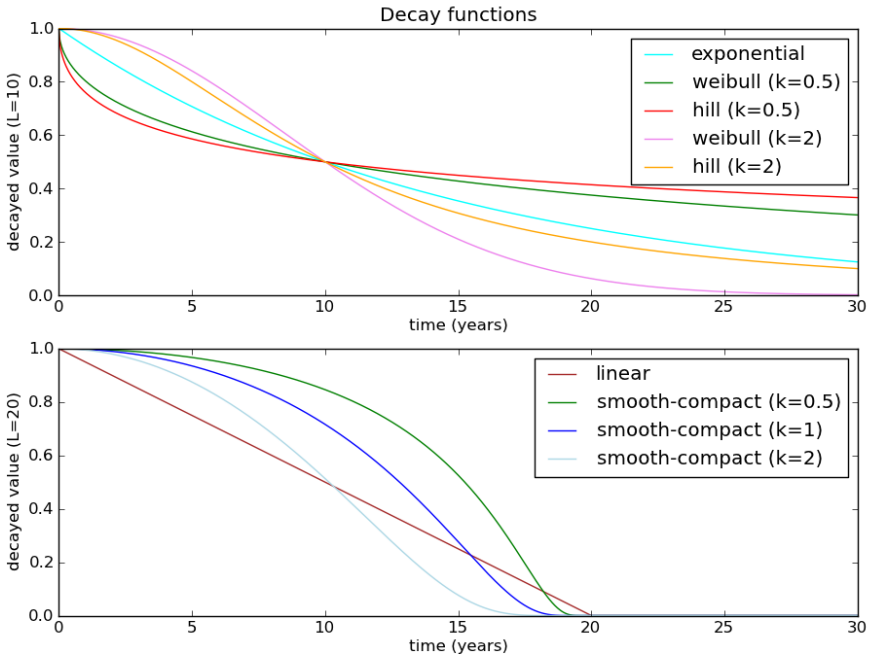


Figure 11: Various decay functions with L=10 and L=20

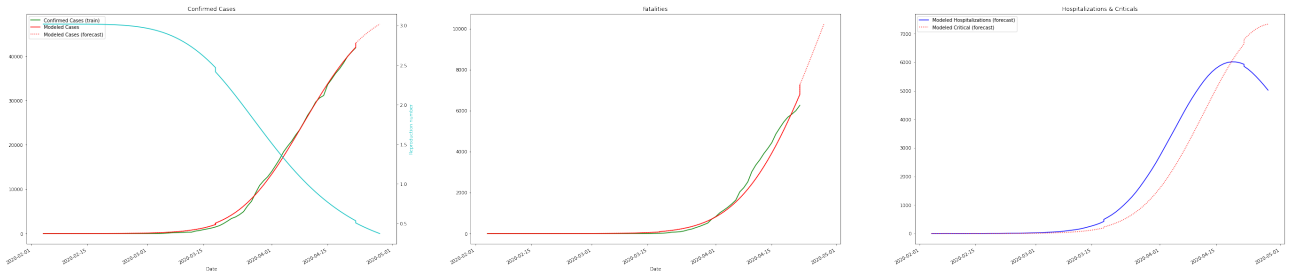


Figure 12: Predicted cases, hospitalizations, critical and deaths of our fitted model for Belgium.

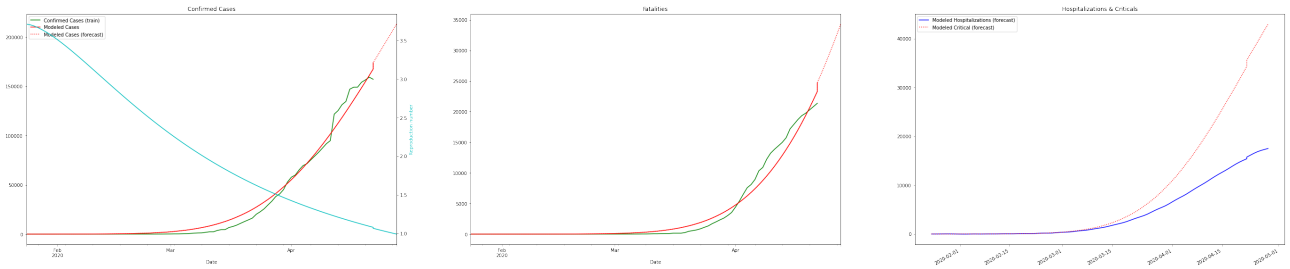


Figure 13: Predicted cases, hospitalizations, critical and deaths of our fitted model for France.

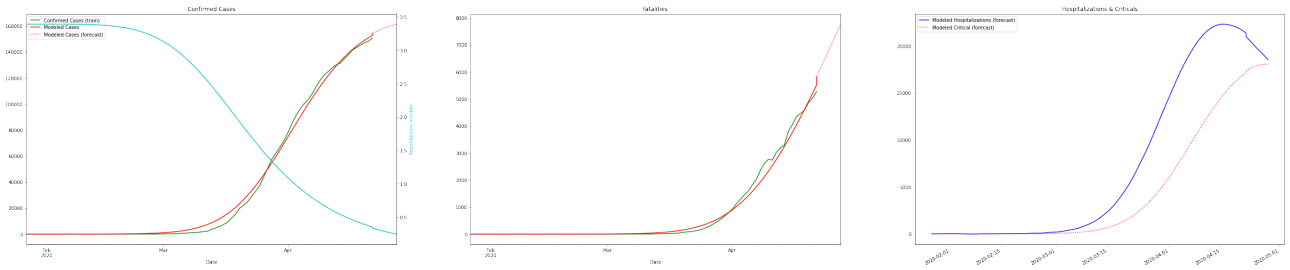


Figure 14: Predicted cases, hospitalizations, critical and deaths of our fitted model for Germany.

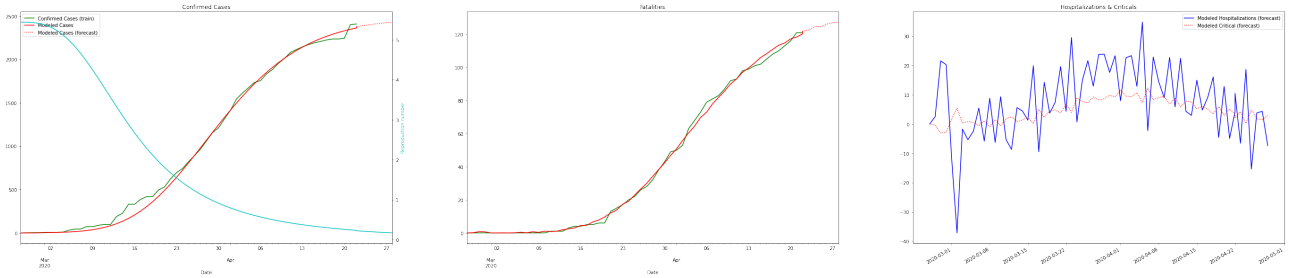


Figure 15: Predicted cases, hospitalizations, critical and deaths of our fitted model for Greece.

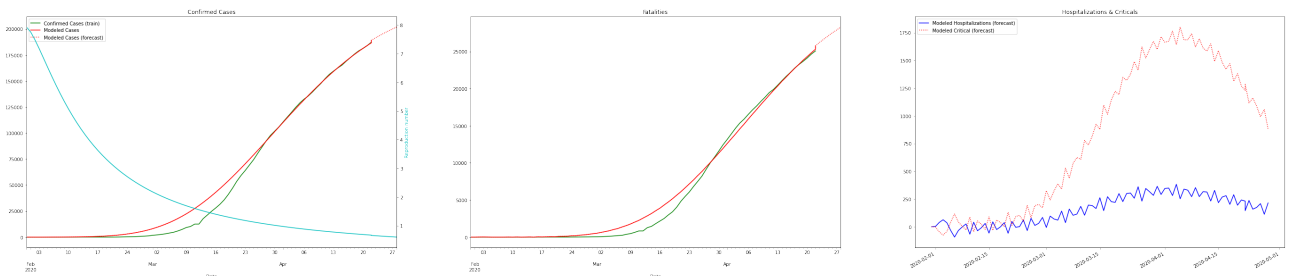


Figure 16: Predicted cases, hospitalizations, critical and deaths of our fitted model for Italy.

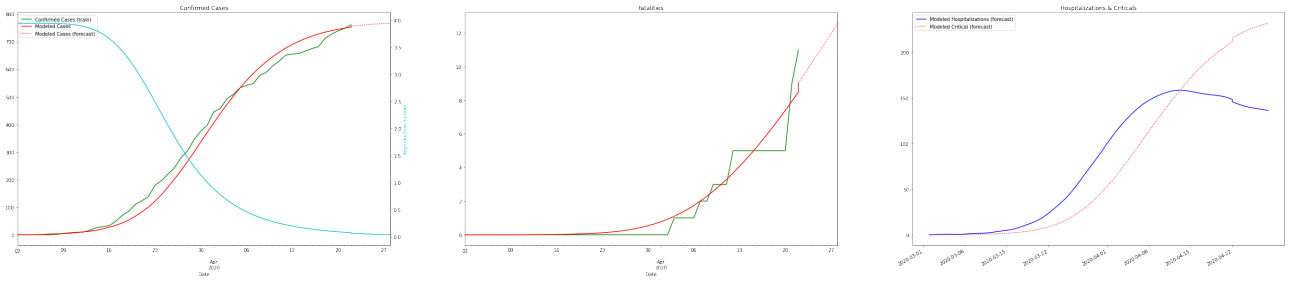


Figure 17: Predicted cases, hospitalizations, critical and deaths of our fitted model for Latvia.

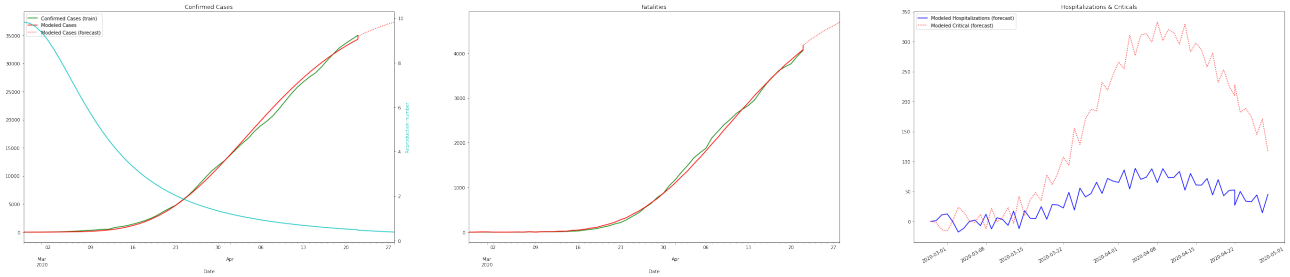


Figure 18: Predicted cases, hospitalizations, critical and deaths of our fitted model for Netherlands.

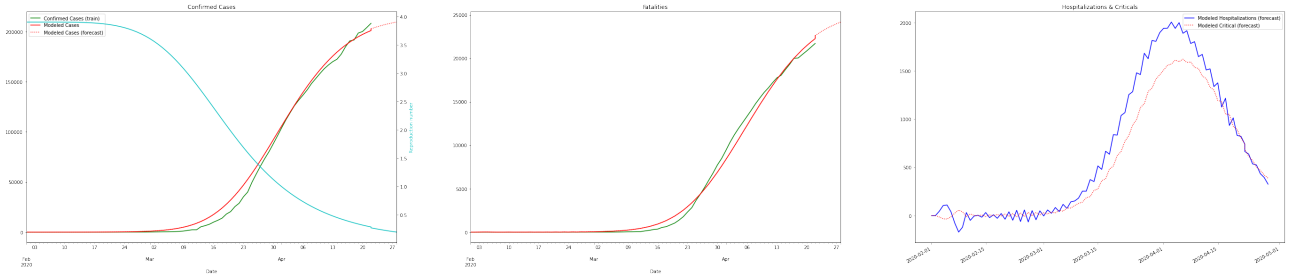


Figure 19: Predicted cases, hospitalizations, critical and deaths of our fitted model for Spain.

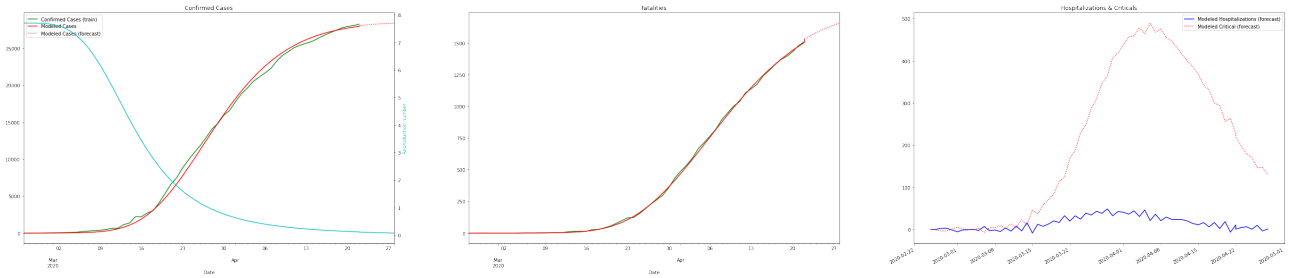


Figure 20: Predicted cases, hospitalizations, critical and deaths of our fitted model for Switzerland.

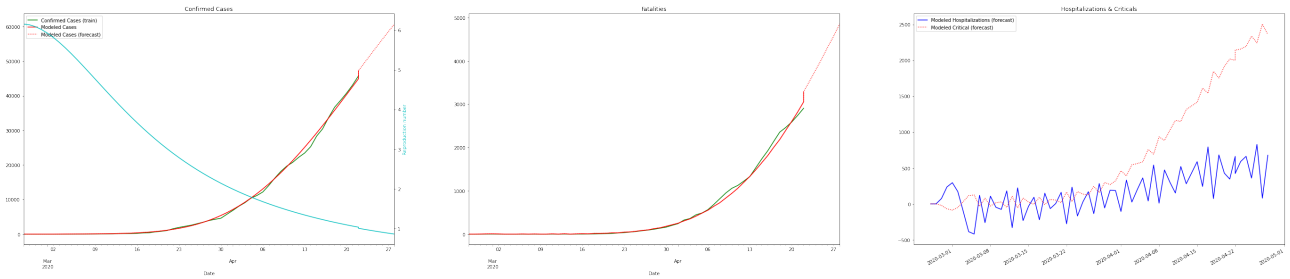


Figure 21: Predicted cases, hospitalizations, critical and deaths of our fitted model for Brazil.

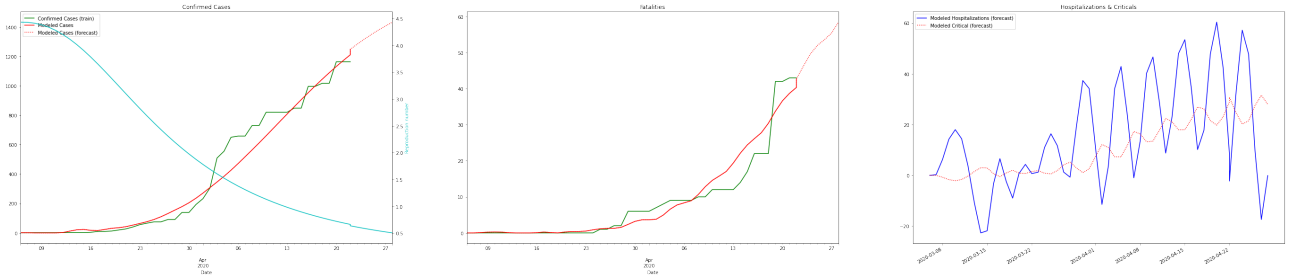


Figure 22: Predicted cases, hospitalizations, critical and deaths of our fitted model for Cameroon.

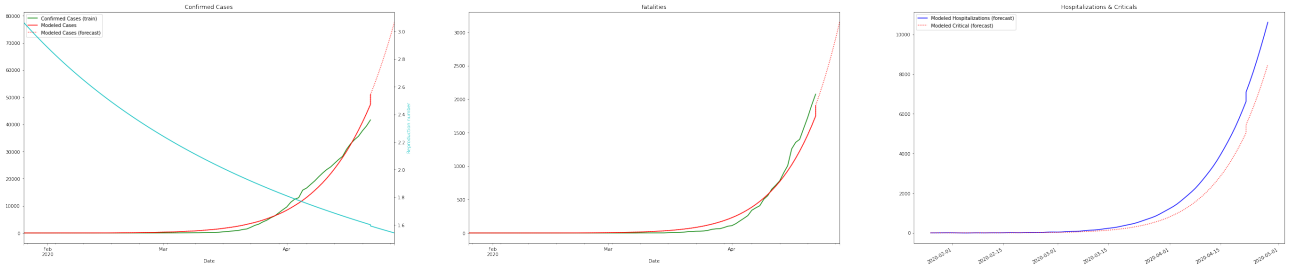


Figure 23: Predicted cases, hospitalizations, critical and deaths of our fitted model for Canada.

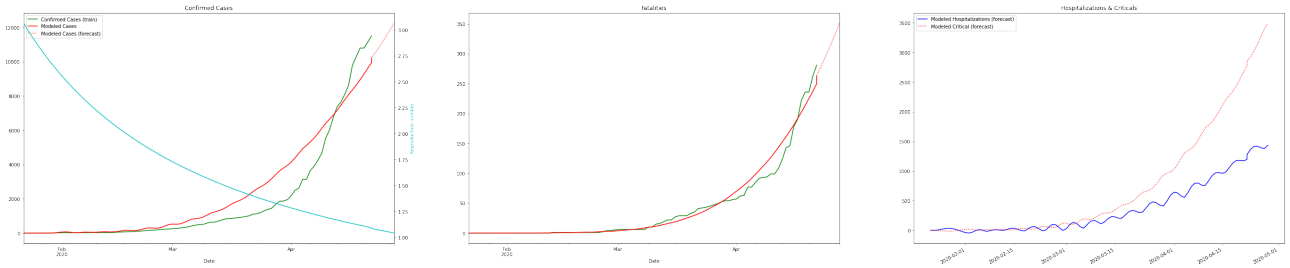


Figure 24: Predicted cases, hospitalizations, critical and deaths of our fitted model for Japan.

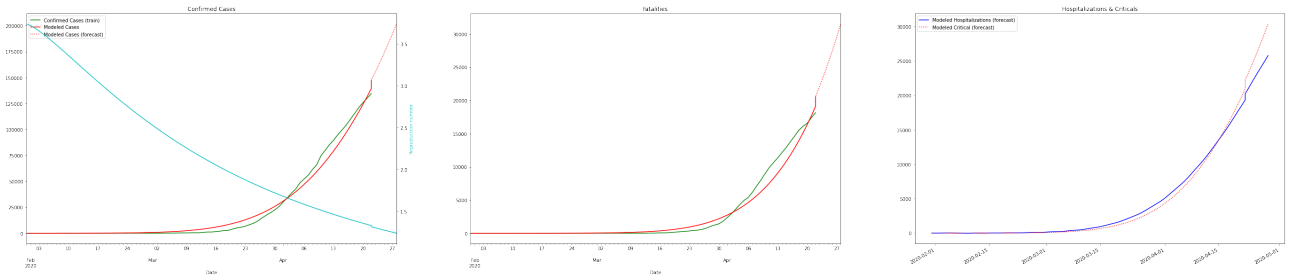


Figure 25: Predicted cases, hospitalizations, critical and deaths of our fitted model for United Kingdom.