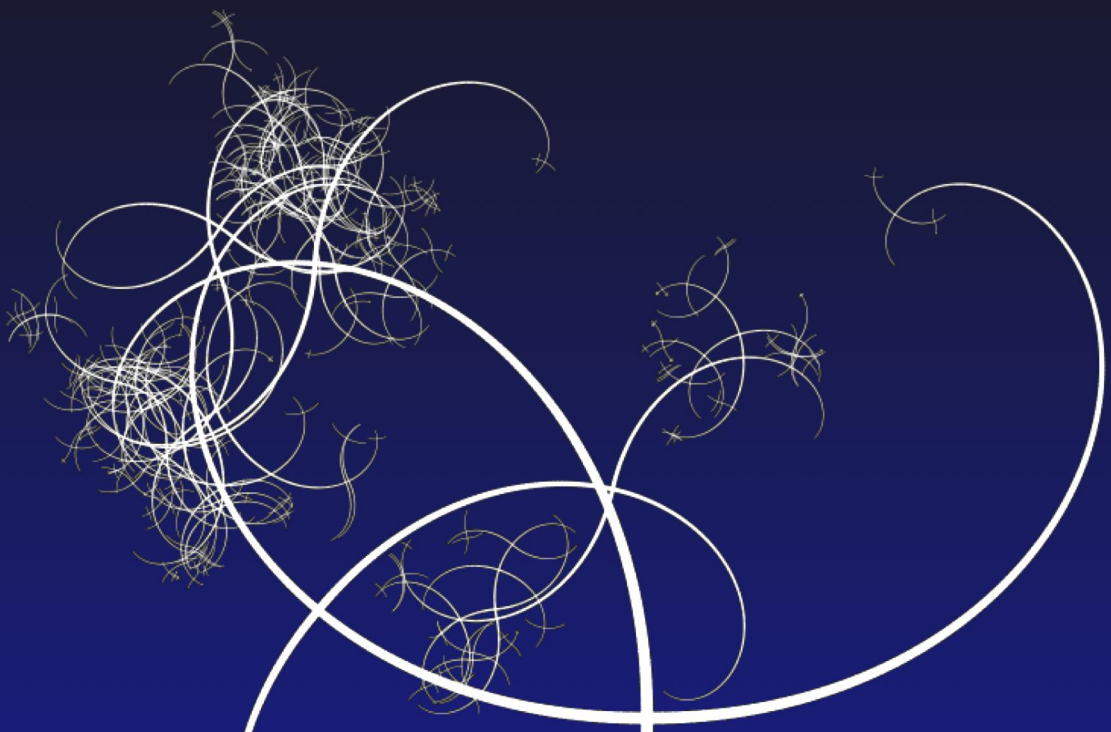
NISSUNA UMANA INVESTIGAZIONE SI PUO DIMANDARE VERA SCIENZIA
S'ESSA NON PASSA PER LE MATEMATICHE DIMOSTRAZIONI
LEONARDO DA VINCI



vol. 8 no. 3 2020

MATHEMATICS AND MECHANICS
of
Complex Systems

JEAN RÉMOND AND YVES RÉMOND

A NEW VIRUS-CENTRIC EPIDEMIC MODELING APPROACH 1. GENERAL THEORY AND MACHINE LEARNING SIMULATION OF 2020 SARS COV 2 (COVID-19) FOR BELGIUM, FRANCE, ITALY, AND SPAIN

msp

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 8, No. 3, 2020
https://doi.org/10.2140/memocs.2020.8.233

A NEW VIRUS-CENTRIC EPIDEMIC MODELING APPROACH 1. GENERAL THEORY AND MACHINE LEARNING SIMULATION OF 2020 SARS COV 2 (COVID-19) FOR BELGIUM, FRANCE, ITALY, AND SPAIN

JEAN RÉMOND AND YVES RÉMOND

We are trying to test the capacity of a simplified macroscopic virus-centric model to simulate the evolution of the SARS CoV 2 epidemic (COVID 19) at the level of a country or a geographical entity, provided that the evolution of the conditions of its development (behaviors, containment policies) are sufficiently homogeneous on the considered territory. For example, a uniformly deployed lockdown on the territory, or a sufficiently uniform overall crisis management. The virus-centric approach means that we favor to model the population dynamic of the virus rather than the evolution of the human cases. In other words, we model the interactions between the virus and its environment – for instance a specific human population with a specific behavior on a territory, instead of modeling the interactions between individuals. Therefore, our approach assumes that an epidemic can be analyzed as the combination of several elementary epidemics which represent a different part of the population with different behaviors through time. The modeling proposed here is based on the finite superposition of Verhulst equations commonly known as logistic functions and used in dynamics of population. Modelling the lockdown effect at the macroscopic level is therefore possible. Our model has parameters with a clear epidemiological interpretation, therefore the evolution of the epidemic can be discussed and compared among four countries: Belgium, France, Italy, and Spain. Parameter optimization is carried out by a classical machine learning process. We present the number of infected patients with SARS-CoV-2 and a comparison between data from the European Centre for Disease Prevention and Control and the modeling. In a general formulation, the model is applicable to any country with similar epidemic management characteristics. These results show that a simple two epidemics decomposition is sufficient to simulate with accuracy the effect of a lockdown on the evolution of the COVID-19 cases.

Communicated by Francesco dell'Isola.
MSC2020: 34C60, 68T20, 92-10, 92D25, 92D30.
Keywords: SARS-CoV-2, COVID-19, epidemic, modeling, simulation, machine learning, infected cases, logistic function.

1. Introduction

Significant progress has been made in epidemic modeling thanks to new capabilities of numerical simulation, improved mathematical modeling as well as artificial intelligence techniques. These modeling also benefit of continuous improvement in data quality. It is impossible to quote all the previous works done, and this is not the purpose of this article. However, the authors suggest consulting [Wiemken and Kelley April 2020] or [Colizza et al. December 2006], and among the most recent publications [Jia et al. 2020] or [Caccavo 2020] and the interesting state of the art concerning SIR and SEIR modelling published on the CNRS web site [Bayette and Monticelli 2020; Gonçalves 2020; Perra and Gonçalves 2015], as well as sites that allow interactive interfacing [Github]. The SIR model and their various developments are probably the cleverest approach at the micro and meso scale and permit to model a large range of macroscopic phenomena. Only a few specific studies used a single logistic function for modeling epidemics, especially for plant diseases and with interesting results, such as [Moral and Trapero 2009; Mesha and Hau 2008; Holb et al. 2005]. Some others used a double logistic curve for modeling HIV, including [Mahiane et al. 2017], or an r-hybrid model for the same virus [Eaton et al. 2019]. Another promising way consists in using PGD-like model reduction, as in [Chinesta and Cueto 2014], for analyzing epidemic kinetics by parametric optimization.

A large part of the existing approaches tries to model the epidemic at the individual scale, i.e. the microscopic scale, and consider the interactions between each individual. Then, it induces a theoretical epidemic evolution at the macroscopic level. A lot of contributions can be found in the literature using that method. This study takes the reverse way and try to find interesting conclusions depending on the microscopic scale, using a macroscopic modeling based on a generalization of the logistic function. It is a common approach developed in theoretical or applied mechanics or physics to use this type of homogenization methods to go from the macroscopic to the microscopic scale; see [Allaire 2001; Oleinik et al. 1992; Sanchez-Palencia 1980; Suquet 1987; Rémond et al. 2016]. Our simplified virus-centric macroscopic modeling is coupled with an automatic parameters optimization by machine learning and gives interesting results for the SARS-CoV-2 early 2020 pandemic. However, predicting the outcome of the epidemic across countries seems to be a lucky guess considering the variability of the containment policies through time and countries. Therefore, readers must be warned that our predictions, *mutatis mutandis*, cannot consider subsequent events such as a possible second epidemic, that could appear after the end of the lockdown, or other unexpected effects. However, the results obtained by this new and simplified approach seemed to us instructive enough to have explained it here.

2. General theory

2.1. ***Introduction to logistic function or sigmoid function.*** On a macroscopic scale, the elementary logistic function law, known as Verhulst's law [Verhulst 1838; Daley and Gani 1999], has been known for a long time (1838) as a law useful for classical modeling of epidemics at the macroscopic scale. It was first implemented in population dynamics. We can consider in a first approach that a population $y(t)$ - a read-valued function of time - of individuals evolves according to a very simple ordinary differential equation: $\frac{dy}{dt} = y(N(y) - M(y))$ where $N(y)$ is the birth rate and $M(y)$ represents the death rate. If $N(y)$ and $M(y)$ are linear functions, this equation can be written $\frac{dy}{dt} = ry\left(1 - \frac{y}{K}\right)$, a and K being strictly positive real numbers. K is conventionally called the carrying capacity in population dynamics theory, r is the growth rate which leads to an increase of population if $y(0) = y_0 < K$, to a decrease if $y_0 > K$ and is stable if $y_0 = K$.

The resolution of this simple ordinary differential equation allows to define the logistic function

$$y(t) = K\left[1 + \left(\frac{K}{y_0} - 1\right)e^{-rt}\right]^{-1}, \quad \lim_{t\to+\infty} y(t) = K. \tag{1}$$

This function is the solution defined on $[0, +\infty[$, of the system constituted by $y(0) = y_0$ and $y' = ry\left(1 - \frac{y}{K}\right)$.

2.2. ***General and macroscopic epidemic modeling.*** In this paper, we consider a virus centric epidemic modeling, which is the modeling of the evolution of virus through time in a specific environment. For this modeling we assume that the number cases of people can be assimilated to the number of virus. The human population of a given territory is the environment the virus has to survive into. The environment is more or less welcoming depending of the human behavior, the territory density and obviously the considered human sub-population. If we consider the whole human population, the number of virus is assimilated to the number of human COVID 19 cases. If we consider a more viable environment for the virus, for instance the sub-population of human susceptible enough to be hospitalized, the number of virus is assimilated of the number of hospitalizations. The same logical thinking could be applied for the number of intensive cares, death and recovery cases.

Therefore, we assume that the number of daily or cumulative cases of people is characterized by a function $E^k(t)$ of $\mathbb{R}$ in $\mathbb{R}$. $E^k(t)$ is defined over a given geographical territory with k the studied phenomena (infection, hospitalization, intensive care, death, recovery). Geographic territories should be chosen as territories in which we notice a similar epidemic management (for instance, territories with a similar lockdown intensity, as countries or other administrative entities, etc.).

This function $E^k(t)$ is itself the sum of continuous functions $E_i^k(t)$ of $\mathbb{R}$ in $\mathbb{R}$ with $i \in \{1, \ .., i, .., P\}$, characterizing the effect of the epidemic on P different populations belonging to the same geographic territory:

$$E^k(t) = \sum_{i=1}^{i=P} E_i^k(t)$$

Finally, each population may have a series of C_i different behaviors over time, and therefore

$$E_i^k(t) = \sum_{j=1}^{j=c_i} q_{ij}(t) E_{ij}^k(t),$$

$$E^k(t) = \sum_{i=1}^{i=P} \sum_{j=1}^{j=Ci} q_{ij}(t) E_{ij}^k(t) = q_{ij}(t) E_{ij}^k(t),$$

where the Einstein summation convention is in effect for i and j, and where $\sum_{j=1}^{j=C} q_{ij}(t) = 1$.

In the case of two different behaviors of a unique population i, the transition function $q(t)$ can be written as follow:

$$E_i^k(t) = q(t) E_{i1}^k(t) + (1 - q(t)) E_{i2}^k(t) \tag{2}$$

with $q(t)$ a monotonically increasing function defined from $[0, +\infty[$ to $]0, 1[$. Many functions may be suitable. We will define a specific one in the application paragraph.

3. Basic application to the 2020 Covid-19 epidemic

In the case of the Covid-19 epidemic which particularly occurred in Europe at the beginning of 2020, an interesting way to test the validity of such a model consists in taking particularly simplified hypothesis: one population per country ($i = 1$), two behaviors ($j = 2$), with a continuous passage from one to the other. Those two behaviors can be identified as the transition from a first behavior of the population before the containment measures to a second behavior considering the containment measures, especially the lockdown. Other subsequent behaviors could have been considered and identified, such as the end of containment measures, or a less rigorous behavior through time, lockdown, etc.

In our case, we define the elementary epidemic function $E^k(t)$ with an equation similar to (1):

$$E^k(t) = L_k[1 + e^{-r_k(t - t_{Sk})}]^{-1} \tag{3}$$

$E^k(t)$ thus, represents the cumulative number of cases of k (infected, hospitalized, in intensive care, deceased, cured) at time t. We will note L_k the final number of

cases of k in one epidemic, r_k the characteristic of the epidemic kinetics for the criterium k and t_S the time taken to reach the peak of the epidemic, described on a daily basis of new cases.

To model it, the following assumptions and development have been made:

H1: We assume that the global behavior of the population is given by the combination of two functions $E_1^k(t)$ and $E_2^k(t)$, as mention by equation (2), which characterize the difference of behavior, before and after the lockdown. As we will consider the unique criterium "number of infected cases" we will not use the exponent k anymore. Consequently, we have

$$E_1(t) = L_1[1 + e^{-r_1(t-t_{S1})}] \quad \text{and} \quad E_2(t) = L_2[1 + e^{-r_2(t-t_{S2})}].$$

The two functions $E_i(t)$ are distinguished by their respective coefficient r_i.

H2: Let us assume that the studied epidemic has a classical sigmoid / gaussian behavior. We have then the total number of cumulative case L and a "new infected case" peak t_s. We assume furthermore that the studied epidemy can be described as the combination of two epidemics with the same total number of cumulative cases L and the same peak t_s. Therefore, we keep the same value $L_i = L$ and we keep the same time $t_{Si} = t_S$ for the two functions $E_i(t)$.

H3: We choose the transition function $q(t)$ as $q(t) = \frac{1}{2}[\tanh(\alpha_c(t - t_i) + 1]$, defined on $[0, \ +\infty[$.

We obviously have $0 \leq q(t) \leq 1$. The coefficient α_c, positive number, represents the efficiency of the lockdown. t_i represents the duration between the start of the epidemic and the date of lockdown increased by the duration of appearance of its first effects. Due to the asymptotic variation of $q(t)$, $q(0) = \varepsilon \ll 1$ for the used values of t_i, which does not influence the results of the modeling.

With these assumptions, the evolution of the new infected cases of the epidemic is given by:

$$E(t) = q(t)E_1(t) + (1 - q(t))E_2(t)$$

The model is therefore defined by the six simple parameters given in Table 1.

4. Machine learning identification algorithm

To obtain the parameters of the model by supervised machine learning, a python code was developed using gradient descents conventionally used in machine learning. In our case, the Levenberg–Marquardt algorithm [Marquardt 1963; Levenberg 1944], was chosen to optimize the mean squared error function.

The learning datasets for each country have been compiled using official European information from the European Center for Disease Prevention and Control or

Parameters	Definition	Unit
L	Total number of cases	-
k_1	Characteristic of the epidemic without lockdown	Day^{-1}
k_2	Characteristic of the epidemic with lockdown	Day^{-1}
t_i	Duration between the start of the epidemic and the start of the lockdown + time of appearance of the first effects of the lockdown	Day
t_S	Duration from start of epidemic to date of inflection point*	Day
α_c	Lockdown efficiency coefficient	Day^{-1}

Table 1. List and definition of the model parameters. * The inflection point of the sigmoid corresponds to the maximum value of its derivative, which is often called the peak of the epidemic.

Parameters	Input value in the algorithm
L	$0.0031 \times$ Population of the country
k_1	0.3
k_2	0.1
t_i	Time from the start of the epidemic to the start of lockdown
t_S	Time between the start of the epidemic and the appearance of the inflection point + 15 days
α_c	0.1

Table 2. input values of the model parameters, i.e., before supervised learning.

ECDC [ECDC]. Considering the small amount of data, it was not possible to build a test data set.

To achieve the learning optimizations for each country, the following initial values have been set:

These values have been chosen through data analysis to be consistent with our model and the available experimental data.

For L, k_1, k_2, t_i, t_S, α_c, the input values influences the speed of the convergence of the algorithm and also ensure the bypassing of local minima:

L: the input value in the algorithm for the total number of people reached by Covid-19 was chosen to be 0.31% of the total population, following several data analysis.

k_1, k_2: their input values are linked to the analysis of the evolution of the epidemic as well as the analysis of experimental data.

t_i: the start dates of containment of the targeted countries are obviously available on government websites. They are as follows:

Country	Lockdown dates
Belgium	18 Mar 2020
France	17 Mar 2020
Italy	09 Mar 2020
Spain	14 Mar 2020

Ten days are added on those values, reflecting the delay needed to see the results of the lockdown. This duration, like the other parameters, is optimized by the algorithm.

t_S: the initial value for the inflection date t_S is chosen lockdown so t_S is greater than the lockdown date. 15 days were added on top of that after careful data analysis. α_c: the initial value was also set after data analysis and after the analysis of simulation curves.

Note that the total population of each country is also provided by the ECDC (https://www.ecdc.europa.eu/en); this information is reported to be from the World Bank Group (https://www.worldbank.org).

5. Results

5.1. ***The data.*** We took data from a single, reliable official source so that it could be compared across countries [ECDC]. It is obvious that regarding the detection of cases of infected persons, these data are highly dependent on the number of tests carried out and the identification protocols. In addition, it is regularly the case that the data is corrected later following updates. The figures used for Belgium, France, Italy, and Spain are shown in Figure 1. We can see that the raw data is unsurprisingly very noisy. We will not comment here on the roots of this fact. To increase the ability of the algorithm to quickly converge on a solution, we smoothed the raw data over several days (3 days, 5 days, 7 days). So, for a smoothing over three days, we have: $C_{i\,(\text{smoothed})} = \frac{1}{3}(C_{i-1} + C_i + C_{i+1})$. To minimize the noise effects and obtain the best possible convergence, we decided to use only the values smoothed over 7 days. The last data used dates from 21 April 2020. The smoothing over 7 days implies that the last valid dates for modeling correspond to 18 April 2020. We see on the following figures the raw data and the smoothed data for each country.

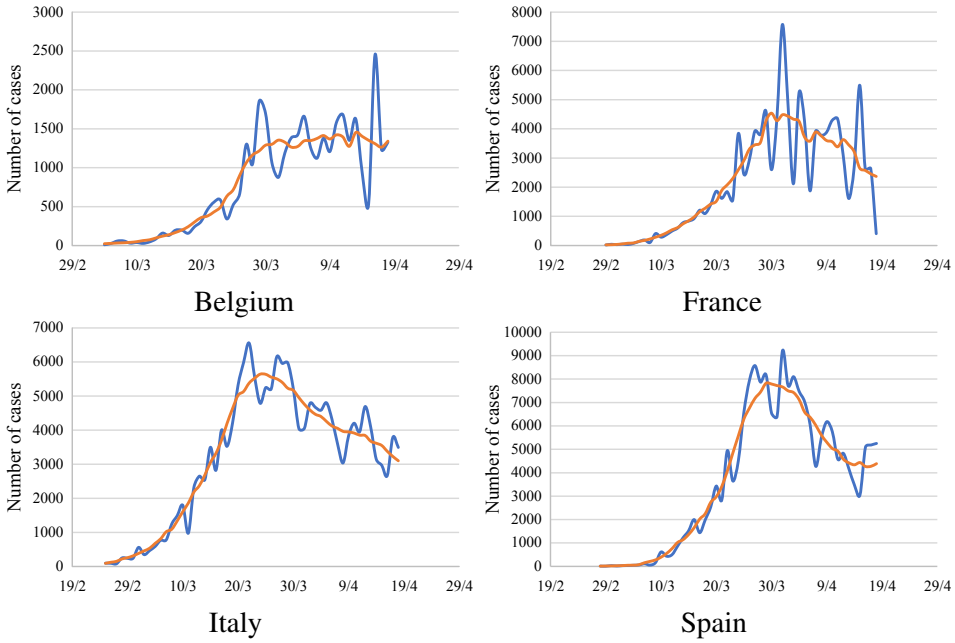


Figure 1. Number of detected cases of COVID-19 in 2020 in the four countries under consideration: raw data and data smoothed over 7 days. Values taken from [ECDC].

5.2. ***Comparison of the basic functions of the model by country.*** We see in Figure 2 comparisons of the graphs of the two epidemic functions $E_1(t)$ and $E_2(t)$ which frame the behavior of the epidemic between its standard evolution and its evolution with a lockdown from the start. A marked difference between these two graphs indicates a more significant effect of the lockdown. We could relate it to the efficiency of this lockdown in the considered territory.

The transition from the function $E_1(t)$ to the function $E_2(t)$ is done by the function $q(t)$ (Fig. 3). q(t) has been created to be a smoothed Heaviside step like function. The intensity of the slope is causally linked to the efficiency of the lockdown effect as the slope depends of the parameter α_c

5.3. ***Simulation results for past and current data.*** After the machine learning optimization of the parameters, the basic functions have been presented for each country. We now compare the smoothed data and the model in Figure 4.

5.4. ***Model predictions.*** So far, we presented the results up to the date of 18 April, i.e., up to the available data with which we trained the model. Obviously, one the main question is the quality of the predictions beyond this date. We present the results up to 18 May in Figure 5. Therefore, the reader can observe that the model

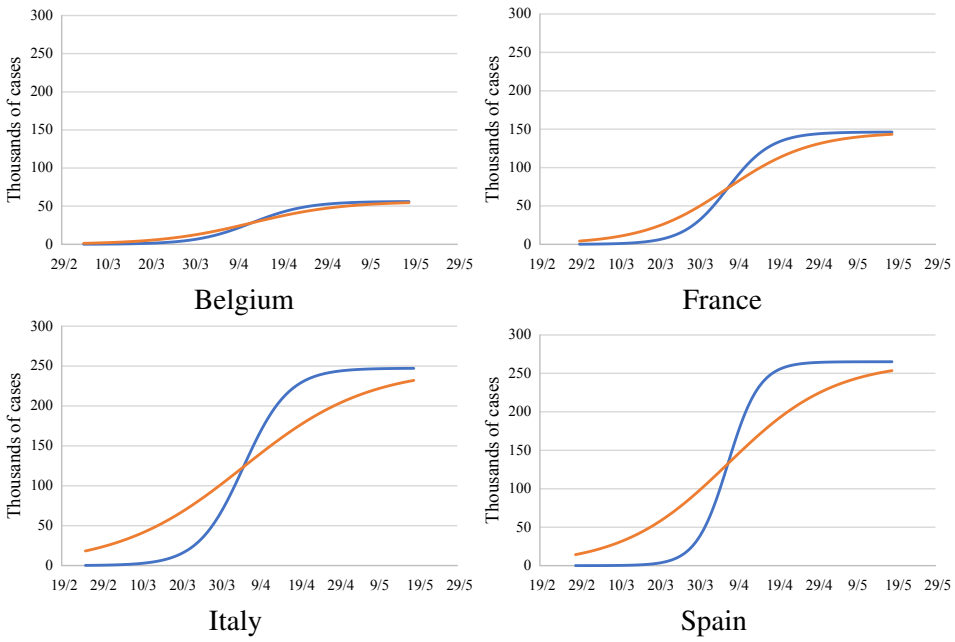


Figure 2. Graph of $E_1(t)$ (no lockdown, blue curve) and $E_2(t)$ (lockdown from the start, orange curve) for the four countries under consideration.

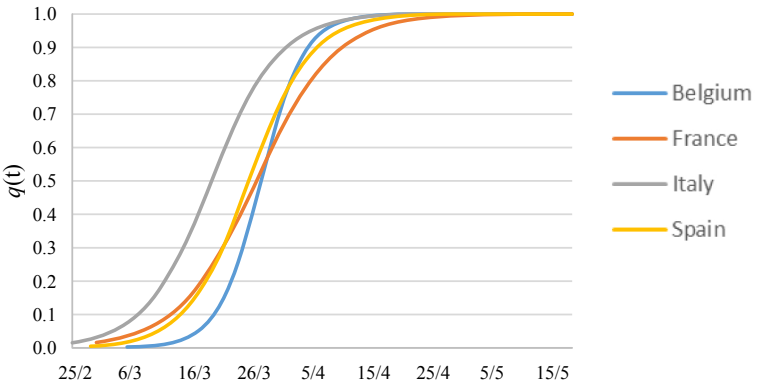


Figure 3. Comparison of the functions $q(t)$ for the four countries. The parameters were optimized by machine learning.

is based on data available up to 21 April (18 April once smoothed), we also draw the available data at the time of the publication, that is to say up to 3 May (30 April once smoothed). This allow the reader to assess the quality of the prediction.

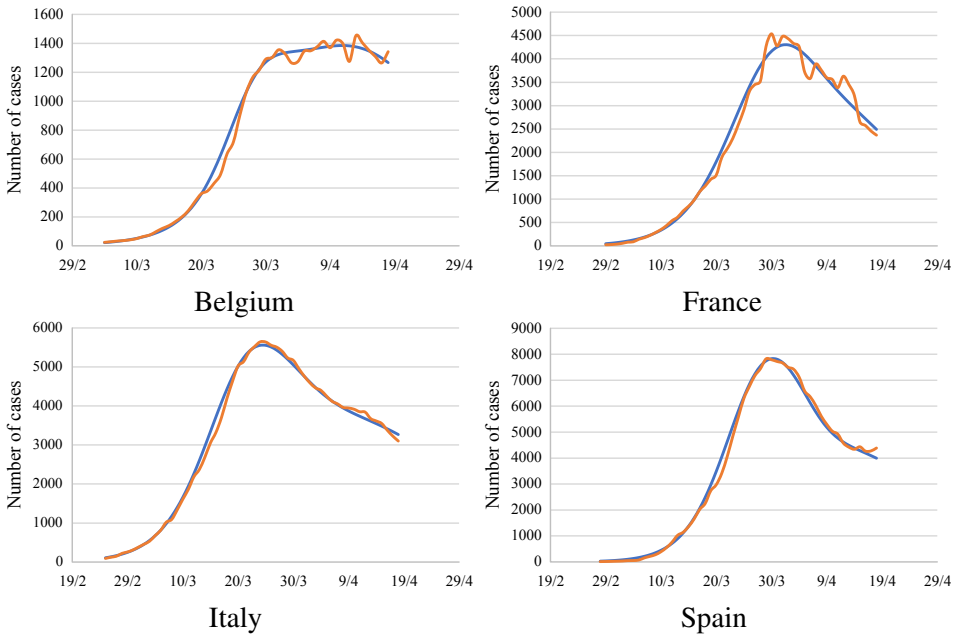


Figure 4. Comparison of our model's results with the smoothed data of the number of detected cases of COVID-19.

Final Value	t_i	α_c	L	k_2	t_S	k_1
Belgium	22.31	0.140	56271	0.098	38.23	0.158
France	26.53	0.076	146277	0.093	37.47	0.181
Italy	23.12	0.089	247302	0.063	39.95	0.174
Spain	26.27	0.096	265181	0.074	38.53	0.250
Error	t_i	α_c	L	k_2	t_S	k_1
Belgium	0.22	0.01	2089	0.00	0.77	0.00
France	1.26	0.01	11025	0.01	1.39	0.01
Italy	0.27	0.00	10356	0.00	1.22	0.01
Spain	0.78	0.00	17900	0.01	1.51	0.02

Table 3. Final values of the model parameters for the four countries, and analysis of the corresponding errors (standard deviation error) - t_i and t_S in days.

The reader will find below the corresponding value of the optimized parameters for each targeted country. Interestingly, those values are close for all countries, except for the total number of cases L, which obviously relies on country population.

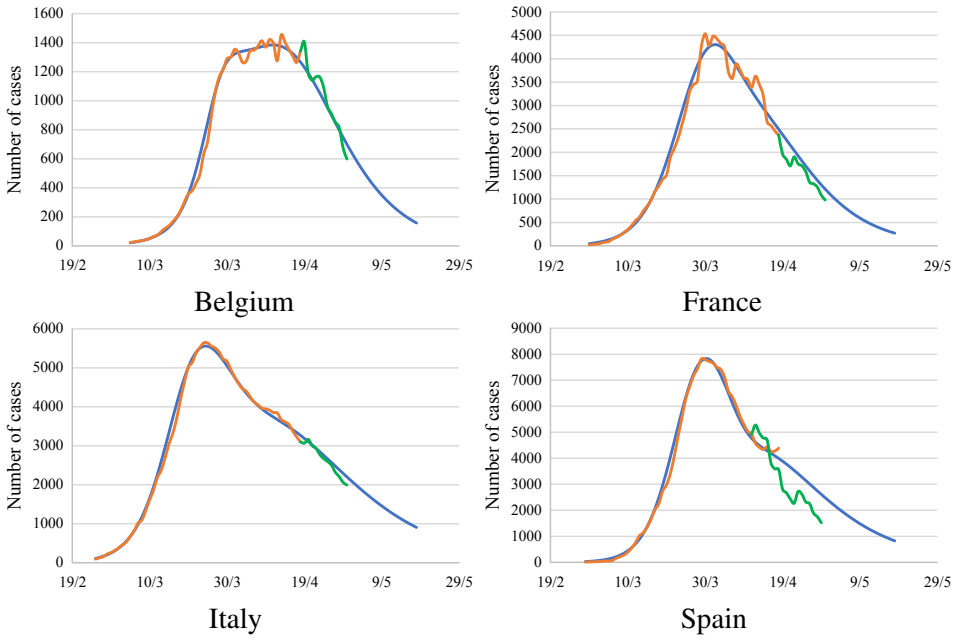


Figure 5. Comparison between the smoothed data of the number of detected cases of COVID-19 (orange: smoothed data available up to 18 April, green: up to 30 April) and our model (blue curve) based on the smoothed data of 18 April. (As discussed later, the data quality for Spain seems debatable since on the dataset up to 30 April, the raw data has been changed from that of 15 April with variation up to 160%; even stranger is that the number of cases on 19 April is negative (-1400 cases).

6. Discussion

We see that a simple macroscopic modeling of the Covid-19 epidemic in 2020, built with only two basic functions which permit to consider lockdown and its effects, allows to correctly model the evolution of the cases of people infected by the virus until the dates for which data are available. Simulation of other characteristics of the epidemic, such as the follow-up of hospitalizations or resuscitation, have also been carried out and will be published later [Rémond and Rémond]. They use the same model and the same optimization algorithm. For the forecasts, they are always to be taken with caution because on the one hand, unexpected events having biological, human causes or of management of the epidemic can occur which modify the form of it in a significant way, on the other hand the assumptions used may seem too summary to assign a level of probability to them. In

peculiarly, the unlockdown time can be considered as a new behavior for a given population and modeled with a new elementary function in addition to the two functions used. This unlockdown time is not considered in the simulations. The results are however interesting and show how a learning algorithm can allow a simple model to correspond well to the macroscopic effects of the epidemic.

It will also be noted that the inflection point of the t_S sigmoid occurs on average 38.65 days (from 37.47 to 39.95) after the start of the epidemic (Table 3) for the four countries. This means that as of this date, half of the people who will be affected by the epidemic have been infected. Analysis of the values of t_i (duration entered at the start of the epidemic and the date of lockdown + duration of appearance of the first effects of lockdown) shows that the duration of appearance of the effects of lockdown is very similar for the four countries, of the order of 10.30 days (between 9.53 for France and 11.27 for Spain). It should also be noted that the number of infected cases is half lower in France than in Italy and then in Spain.

The parameters α_c are associated to the velocity of changing of behavior after the lockdown. France, Italy, and Spain, with α_c between 0.076 and 0.096, have similar reactions for this evolution. By the way, to appreciate the step between the two behaviors, we must analyze the step between k_1 and k_2 for each country. The ratio k_1/k_2 represents the intensity of this change. In that case, Belgium appears to have the smallest change of behavior with a ratio of 1.62. France a ratio of 1.92 is better. Italy has a high ratio of 2.76 and Spain has strongly changed its behavior with a ratio of 3.36.

It is interesting to remember that the equation of the elementary basic functions $E_k(t)$ of the epidemic, given by the equation (2), or its initial form given in equation (1), are the solution of a elementary differential equation detailed in 2.1. This differential equation is solved for the boundary condition $y(0) = y_0$. The boundary condition y_0 characterizes the number of cases at time t_0. We show in the table 4 the values of y_0 given by the raw data and the smoothed data. They can be compared with the values of y_0 obtained by the modeling after the machine learning process. The differences of these values for raw and smoothed data are the intensity of the noise associated to these data. On the opposite, the differences between the values of y_0 for the data and the modeling are interesting. It shows that the number of cases at time t_0 were strongly underestimated by a factor 5 in Italy and in Belgium, and by a factor 10 for France and Spain. The more precise measurement of these boundary conditions at t_0 could have helped to better appreciate the intensity of this epidemic in a short time.

Note the extremely high value of y_0 for Belgium compared with the number of the inhabitants of this country, six times lower than the three other countries.

The coefficients R^2 of the model are given in the table below and by country, with smoothed and non-smoothed data. We see that the simulation is particularly

Country	y_0 **– Raw data** (Sum of new cases from $t = -3$ to $t = 0$)	y_0 **– Smoothed data** (Value of new cases at $t = 0$)	y_0 **– Given by the model**
Belgium	22	24	135
France	45	24	237
Italy	226	92	524
Spain	23	12	109

Table 4. Values of the number of cases y_0 at t_0 measured with the raw and smoothed data, comparing to the number of cases y_0 at t_0 given by the model.

Country	**Smoothed data**	**Raw data – 18th April**
Belgium	0.993	0.733
France	0.986	0.648
Italy	0.996	0.928
Spain	0.995	0.928

Table 5. R^2 coefficient of quality of simulations compared to raw data and smoothed data.

good for smoothed data. For raw data, the case of France is special given the positive data jumps that were recorded on certain days for administrative reasons. For Spain, despite of a good R^2 coefficient which show a good correlation between the data and the model, the official data given by this country after the identification of the model have changed strongly after 15 April. Then there is no peculiar significance on the analysis of this variation.

Therefore, this global modeling of the COVID-19 epidemic seems to be understandable as a sum of only two different elementary basis functions including the effects of the lockdown, and the development of such analysis will probably permit to analyze the specific behavior of population, in complement of the classical approaches by micro-macro analysis.

7. Conclusion

We have created a particularly simple virus-centric model of the Covid-19 epidemic, based on a decomposition in generic basic functions adaptable to all countries and to all the characteristic criteria of its development. Using a simple machine learning process, we show that only two basic elementary functions were sufficient to simulate the epidemic evolution for four European countries applying a lockdown, with accuracy. The results permit to quantify the difference of behavior

before and after the lockdown for these countries as well as the velocity of change and the intensity of change. We focused here on the model's ability to simulate the numbers of new cases reported in the Covid-19 epidemic over time. However, this model could be used for modeling hospitalizations, intensive care, and death. The prediction of a unlockdown effects should be also possible with this model, by adding a third elementary basic function describing the specific behavior of population after this event. The presented simulations are relevant and clearly show the effects of the various lockdowns carried out. The analysis of basic functions used in this decomposition could in any case allow us to have a macroscopic analysis of how the lockdowns were respected. Other characteristics and simulations of the 2020 SARS CoV 2 epidemic for other countries will be given in a sequel to this paper [Rémond and Rémond].

References

[Allaire 2001] G. Allaire, *Shape optimization by the homogenization method*, Springer, 2001.

[Bayette and Monticelli 2020] C. Bayette and M. Monticelli, "Modélisation d'une épidémie, 1", 2020, available at https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-1.html.

[Caccavo 2020] D. Caccavo, "Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model", 2020, available at https://tinyurl.com/caccavo-sird.

[Chinesta and Cueto 2014] F. Chinesta and E. Cueto, *PGD-based modeling of materials, structures and processes*, ESAform Series on Materials Forming, Springer, 2014.

[Colizza et al. December 2006] V. Colizza, A. Barrat, M. Barthelemy, and A. Vespignani, "The modeling of global epidemics: stochastic dynamics and predictability", *Bulletin of Mathematical Biology* **68** (December 2006), 1893–1921.

[Daley and Gani 1999] D. J. Daley and J. Gani, *Epidemic modelling*, Cambridge University Press, 1999.

[Eaton et al. 2019] J. W. Eaton et al., "The estimation and projection package age-sex model and r-hybrid model: new tools for estimating HIV incidence trends in sub-Saharian Africa", *AIDS* **33**:Suppl. 3 (2019), S235–S244.

[ECDC] ECDC, available at https://www.ecdc.europa.eu/en.

[Github] Github, available at https://github.com/uiuc-covid19-modeling/pydemic.

[Gonçalves 2020] B. Gonçalves, "Epidemic Modeling 102: all CoVID-19 models are wrong, but some are useful", 2020, available at https://tinyurl.com/goncalves-covid.

[Holb et al. 2005] I. J. Holb et al., "Analysis of summer epidemic progress of apple scab at different apple production systems in the Netherlands and Hungary", *J. of Phytopathology* **95**:9 (2005), 1001–1020.

[Jia et al. 2020] L. Jia, K. Li, Y. Jiang, X. Guo, and T. Zhao, "Prediction and analysis of coronavirus disease 2019", 2020, available at arxiv.org/abs/2003.05447v2.

[Levenberg 1944] K. Levenberg, "A method for the solution of certain problems in least squares", *Quart. Appl. Math* **2** (1944), 164–168.

[Mahiane et al. 2017] S. G. Mahiane et al., "Improvements in Spectrum's fit to program data tool", *AIDS* **31**:Suppl. 1 (2017), S23–S30.

[Marquardt 1963] D. W. Marquardt, "An algorithm for least-squares estimation of non linear parameters", *SIAM J. Appl. Math* **11** (1963), 431–441.

[Mesha and Hau 2008] Z. Mesha and B. Hau, "Effects of bean rust (*Uromyces appendiculatus*) epidemics on host dynamics of common bean", *Plant Pathology* **57** (2008), 674–686.

[Moral and Trapero 2009] J. Moral and A. Trapero, "Assessing the susceptibility of olive cultivars to anthracnose caused by *Colletottrichum acutatum*", *Plant Disease* **93**:10 (2009), 1028–1036.

[Oleinik et al. 1992] O. Oleinik, A. Shamaev, and G. Yosifian, *Mathematical problems in elasticity and homogenization*, Studies in Mathematics and its Application **26**, Elsevier, Amsterdam, 1992.

[Perra and Gonçalves 2015] N. Perra and B. Gonçalves, *Modeling and predicting human infectious diseases, social phenomena*, Springer, 2015.

[Rémond and Rémond] J. Rémond and Y. Rémond, "On a new virus-centric epidemic modeling, 2: Simulation of deaths from 2020 SARS CoV 2 in several countries", *Mathematics and Mechanics of Complex System*.

[Rémond et al. 2016] Y. Rémond, S. Ahzi, M. Baniassadi, and H. Garmestani, *Applied RVE reconstruction of heterogeneous materials*, Wiley-ISTE, 2016.

[Sanchez-Palencia 1980] E. Sanchez-Palencia, *Non-homogeneous media and vibration theory*, Lecture Notes in Physics **129**, Springer, 1980.

[Suquet 1987] P. M. Suquet, "Elements of homogenization for inelastic solid mechanics", pp. 193–278 in *Homogenization techniques for composite media*, Lecture Notes in Physics **105**, Springer, 1987.

[Verhulst 1838] P.-F. Verhulst, "Notice sur la loi que la population poursuit dans son accroissement", *Correspondance mathématique et physique* **10** (1838), 113–121.

[Wiemken and Kelley April 2020] T. L. Wiemken and R. R. Kelley, "Machine learning in epidemiology and health outcomes research", *Annual Review of Public Health* **41** (April 2020), 21–36.

Received 5 May 2020. Revised 12 May 2020. Accepted 16 Jun 2020.

JEAN RÉMOND: `j.remond@stanwell.fr`
Stanwell Consulting, Paris, France

YVES RÉMOND: `remond@unistra.fr`
ECPM - ICUBE Laboratory, University of Strasbourg / CNRS, Strasbourg, France