Understanding unreported cases in the COVID-19 epidemic outbreak and the importance of major public health interventions

Quentin Griette and Pierre Magal

University of Bordeaux, France

Modelling the propagation of COVID-19, EHESS Mai 20 2020.

Co-authors

- **Zhihua Liu**, *Beijing Normal University, China*
  
- **Ousmane Seydi**, *Ecole Polethnique de Thies, Senegal*
  
- **Glenn Webb**, *Vanderbilt University, USA*
  

Abstract

We develop a mathematical model to provide epidemic predictions for the COVID-19 epidemic in China. We use reported case data from the Chinese Center for Disease Control and Prevention and the Wuhan Municipal Health Commission to parameterize the model. From the parameterized model we identify the number of unreported cases. We then use the model to project the epidemic forward with varying level of public health interventions. The model predictions emphasize the importance of major public health interventions in controlling COVID-19 epidemics.

Outline

1 Introduction

2 Results

3 Numerical Simulations

4 Age dependency in COVID-19 for Japan

What are unreported cases?

Unreported cases are missed because authorities are not doing enough testing on people showing symptoms, or 'preclinical cases' in which people are incubating the virus but not yet showing symptoms.

Research published[1] traced COVID-19 infections which resulted from a business meeting in Germany attended by someone infected but **who showed no symptoms at the time**. Four people were ultimately infected from that single contact.

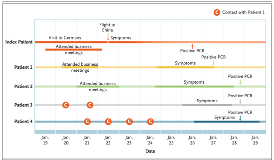


Figure 1. Timeline of Exposure to Index Patient with Asymptomatic 2019-CoV Infection in Germany.

[1] Rothe, et al., Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine* (2020).

Why unreported cases are important?

A team in Japan[2] reports that 13 evacuees from the *Diamond Princess* were infected, of whom 4, or 31%, **never developed symptoms**.

A team in China [3] suggests that by 18 February, there were 37,400 people with the virus in Wuhan whom authorities didn't know about.

[2]Nishiura et al. Serial interval of novel coronavirus (COVID-19) infections, *Int. J. Infect. Dis.* (2020).

[3]Wang et al. Evolving Epidemiology and Impact of Non-pharmaceutical Interventions on the Outbreak of Coronavirus Disease 2019 in Wuhan, China, *medRxiv* (2020)

Early models designed for the COVID-19

- Wu et al. [4] used a susceptible-exposed-infectious-recovered metapopulation model to simulate the epidemics across all major cities in China.
- Tang et al. [5] proposed an SEIR compartmental model based on the clinical progression based on the clinical progression of the disease, epidemiological status of the individuals, and the intervention measures which did not consider unreported cases.

[4] Wu, Joseph T., Kathy Leung, and Gabriel M. Leung, Nowcasting and forecasting the potential domestic and international spread of the COVID-19 outbreak originating in Wuhan, China: a modelling study, *The Lancet*, (2020).

[5] Biao Tang, Xia Wang, Qian Li, Nicola Luigi Bragazzi, Sanyi Tang, Yanni Xiao, Jianhong Wu, Estimation of the transmission risk of COVID-19 and its implication for public health interventions, *Journal of Clinical Medicine*, (2020).

Early results on identification the number of unreported cases

Identifying the number of unreported cases was considered recently in

- Magal and Webb[6]
- Ducrot, Magal, Nguyen and Webb [7]

In these works we consider an SIR model and we consider the Hong-Kong seasonal influenza epidemic in New York City in 1968-1969.

[6] P. Magal and G. Webb, The parameter identification problem for SIR epidemic models: Identifying Unreported Cases, *J. Math. Biol.* (2018).

[7] A. Ducrot, P. Magal, T. Nguyen, G. Webb. Identifying the Number of Unreported Cases in SIR Epidemic Models. *Mathematical Medicine and Biology*, (2019)

Outline

1 Introduction

2 Results

3 Numerical Simulations

4 Age dependency in COVID-19 for Japan

The model

Our model consists of the following system of ordinary differential equations

$$\begin{cases} S'(t) = -\tau S(t)[I(t) + U(t)], \\ I'(t) = \tau S(t)[I(t) + U(t)] - \nu I(t), \\ R'(t) = \nu_1 I(t) - \eta R(t), \\ U'(t) = \nu_2 I(t) - \eta U(t). \end{cases} \tag{2.1}$$

Here $t \geq t_0$ is time in days, t_0 is the beginning date of the epidemic, $S(t)$ is the number of individuals susceptible to infection at time t, $I(t)$ is the number of asymptomatic infectious individuals at time t, $R(t)$ is the number of reported symptomatic infectious individuals (i.e. symptomatic infectious with **sever symptoms**) at time t, and $U(t)$ is the number of unreported symptomatic infectious individuals (i.e. symptomatic infectious with **mild symptoms**) at time t. This system is supplemented by initial data

$$S(t_0) = S_0 > 0,\ I(t_0) = I_0 > 0,\ R(t_0) \geq 0 \text{ and } U(t_0) = U_0 \geq 0. \tag{2.2}$$

Compartments and flow chart of the model.

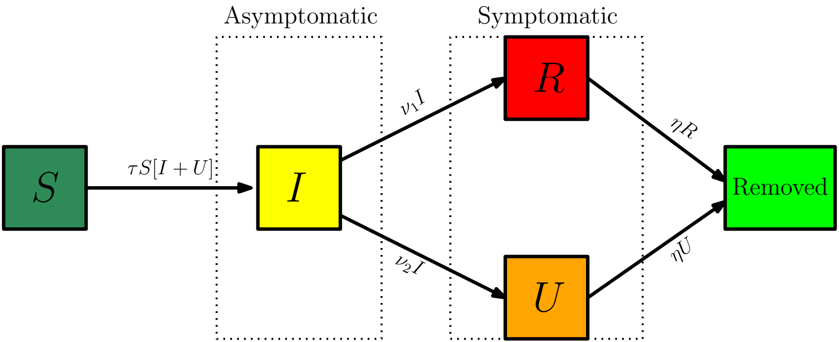


Figure: *Compartments and flow chart of the model.*

Why the exposed class can be neglected?

Exposed individuals are infected but not yet capable to transmit the pathogen.

A team in China[8] detected **high viral loads** in 17 people with COVID-19 soon after they became ill. Moreover, another infected individual never developed symptoms but shed a similar amount of virus to those who did.

In Liu et al. [9] we compare the model (2.1) with exposure and the best fit is obtained for an average exposed period of 6-12 hours.

[8] Zou, L., SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *New England Journal of Medicine*, (2020).

[9] Z. Liu, P. Magal, O. Seydi, and G. Webb, A COVID-19 epidemic model with latency period, *Infectious Disease Modelling (to appear)*

Parameters of the model

Symbol	Interpretation	Method
t_0	Time at which the epidemic started	fitted
S_0	Number of susceptible at time t_0	fixed
I_0	Number of asymptomatic infectious at time t_0	fitted
U_0	Number of unreported symptomatic infectious at time t_0	fitted
R_0	Number of reported symptomatic infectious at time t_0	fixed
τ	Transmission rate	fitted
$1/\nu$	Average time during which asymptomatic infectious are asymptomatic	fixed
f	Fraction of asymptomatic infectious that become reported symptomatic infectious	fixed
$\nu_1 = f\,\nu$	Rate at which asymptomatic infectious become reported symptomatic	fixed
$\nu_2 = (1-f)\,\nu$	Rate at which asymptomatic infectious become unreported symptomatic	fixed
$1/\eta$	Average time symptomatic infectious have symptoms	fixed

Table: *Parameters of the model.*

Estimation of the parameters for the model (2.1)

We fit the data by using a **phenomenological model** for the cumulative number of reported $CR(t)$

$$CR(t) = \chi_1 \exp(\chi_2 t) - \chi_3. \tag{2.3}$$

By using our model the cumulative number of reported is given by

$$CR(t) = \nu_1 \int_{t_0}^{t} I(s) ds. \tag{2.4}$$

By fixing $S(t) = S_0$ in the I-equation of system (2.1), we obtain

$$t_0 = \frac{1}{\chi_2}\left[\ln(\chi_3) - \ln(\chi_1)\right]$$

$$I_0 = \frac{\chi_1 \chi_2 \exp\left(\chi_2 t_0\right)}{f\,\nu} = \frac{\chi_3 \chi_2}{f\,\nu}, \tag{2.5}$$

$$\tau = \frac{\chi_2 + \nu}{S_0} \frac{\eta + \chi_2}{\nu_2 + \eta + \chi_2}, \tag{2.6}$$

and

$$U_0 = \frac{(1-f)\nu}{\eta + \chi_2} I_0 \text{ and } R_0 = \frac{f\nu}{\eta + \chi_2} I_0. \tag{2.7}$$

Outline

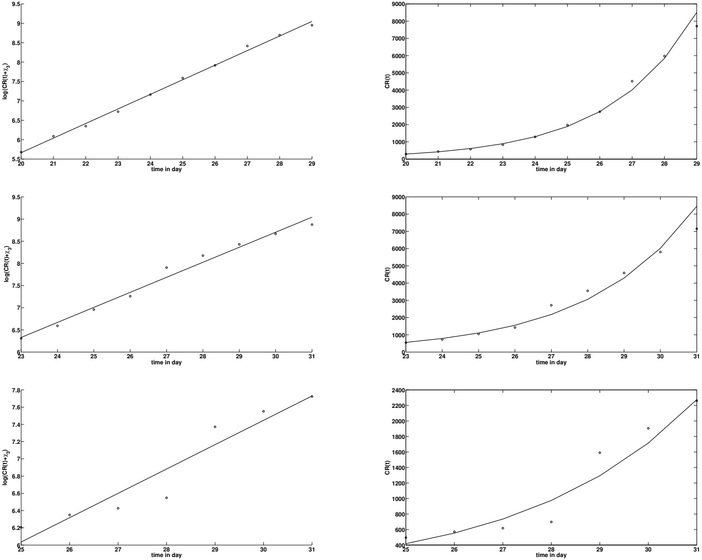
1. Introduction
2. Results
3. Numerical Simulations
4. Age dependency in COVID-19 for Japan

Numerical Simulations

We can find multiple values of η, ν and f which provide a good fit for the data. For application of our model, η, ν and f must vary in a reasonable range. For the corona virus COVID-19 epidemic in Wuhan at its current stage, the values of η, ν and f are not known. From preliminary information, we use the values

$$f = 0.8,\ \eta = 1/7,\ \nu = 1/7.$$

Fit of the exponential model (2.4) to the data for China (top) Hubei province (middle) and Wuhan City (bottom)



Time dependent transmission rate $\tau(t)$

The formula for $\tau(t)$ during the exponential decreasing phase was derived by a fitting procedure. The formula for $\tau(t)$ is

$$\begin{cases} \tau(t) = \tau_0, \, 0 \leq \, t \, \leq \, N, \\ \tau(t) = \tau_0 \, \exp\left(-\mu \left(t - N\right)\right), \, N < t. \end{cases} \tag{3.1}$$

The date N is the **first day of the confinement** and the value of μ is the **intensity of the confinement**. The parameters N and μ are chosen so that the cumulative reported cases in the numerical simulation of the epidemic aligns with the cumulative reported case data during a period of time after January 19. We choose $N = 25$ (January 25) for our simulations.

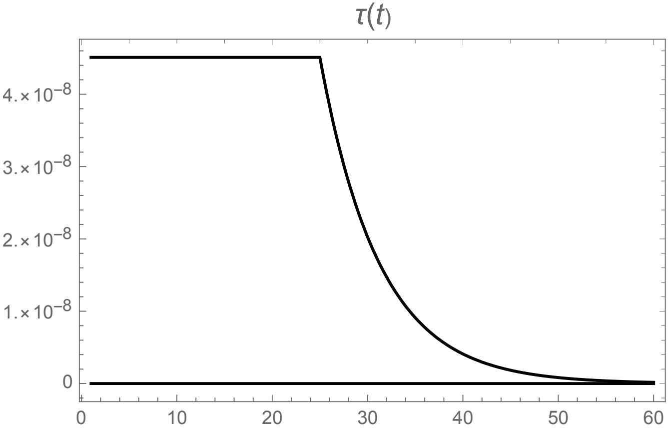
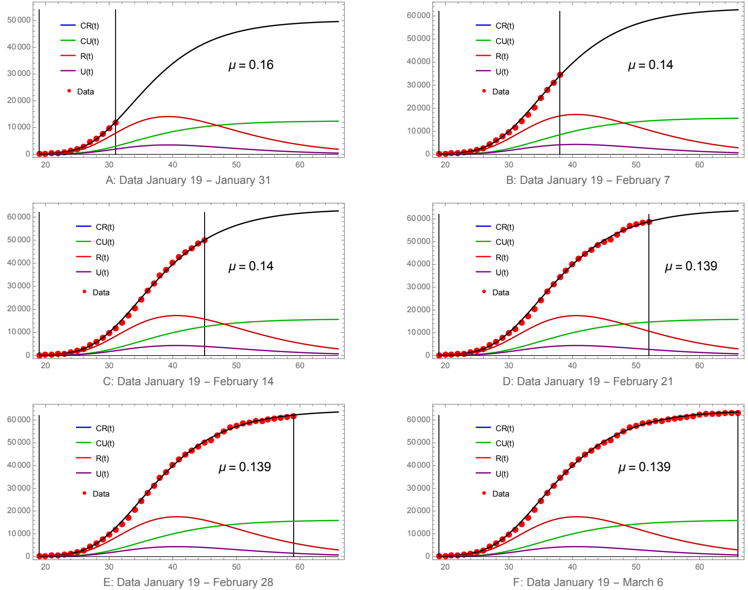


Figure: *Graph of* $\tau(t)$ *with* $N = 25$ *(January* 25*) and* $\mu = 0.16$*. The transmission rate is effectively* 0.0 *after day* 53 *(February* 22*).*

Predicting the epidemic in China with $f = 0.8$

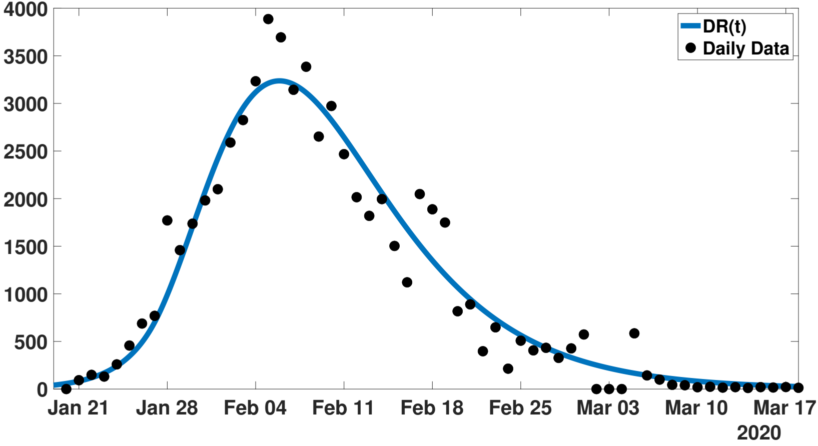


Daily number of cases

The daily number of reported cases from the model can be obtained by computing the solution of the following equation:

$$DR'(t) = \nu\, f\, I(t) - DR(t), \text{ for } t \geq t_0 \text{ and } DR(t_0) = DR_0. \tag{3.2}$$

Predicting the weekly data in China



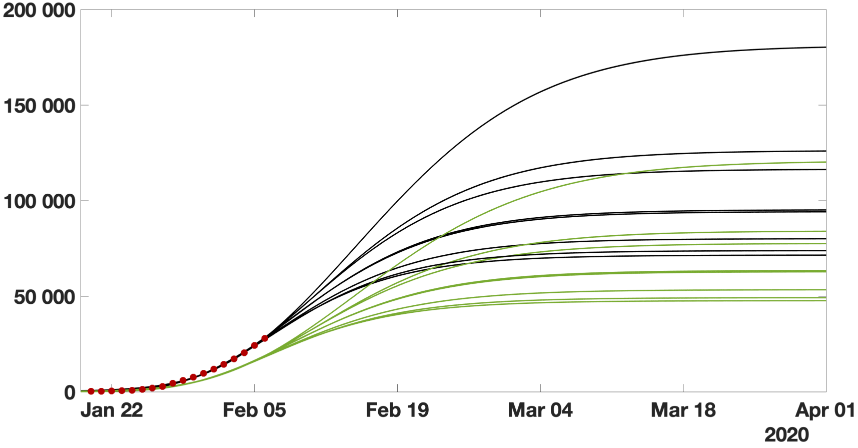
Multiple good fit simulations

We vary the time interval $[d_1, d_2]$ during which we use the data to obtain χ_1 and χ_2 by using an exponential fit. In the simulations below we vary the first day d_1, the last day d_2, N (date at which public intervention measures became effective) such that all possible sets of parameters (d_1, d_2, N) will be considered. For each (d_1, d_2, N) we evaluate μ to obtain the best fit of the model to the data. We use the mean absolute deviation as the distance to data to evaluate the best fit to the data. We obtain a large number of best fit depending on (d_1, d_2, N, f) and we plot the smallest mean absolute deviation $\mathrm{MAD}_{\min}$. Then we plot all the best fit with mean absolute deviation between $\mathrm{MAD}_{\min}$ and $\mathrm{MAD}_{\min} + 5$.

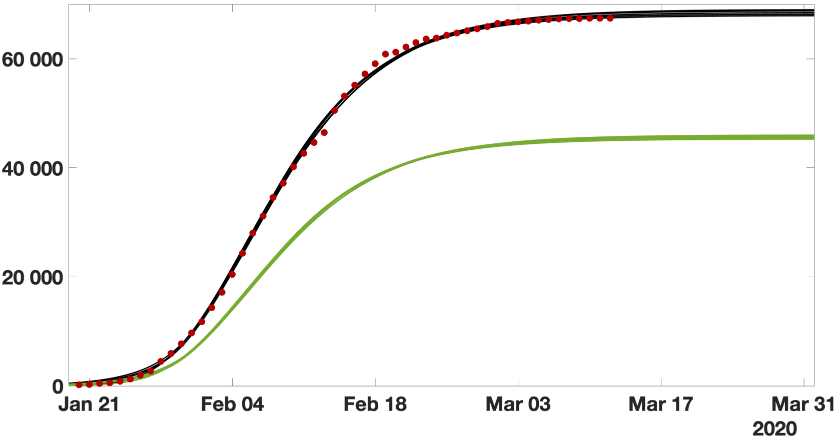
Remark 3.1

The number 5 chosen in $\mathrm{MAD}_{\min} + 5$ is questionable. We use this value for all the simulations since it gives sufficiently many runs that are fitting very well the data and which gives later a sufficiently large deviation.

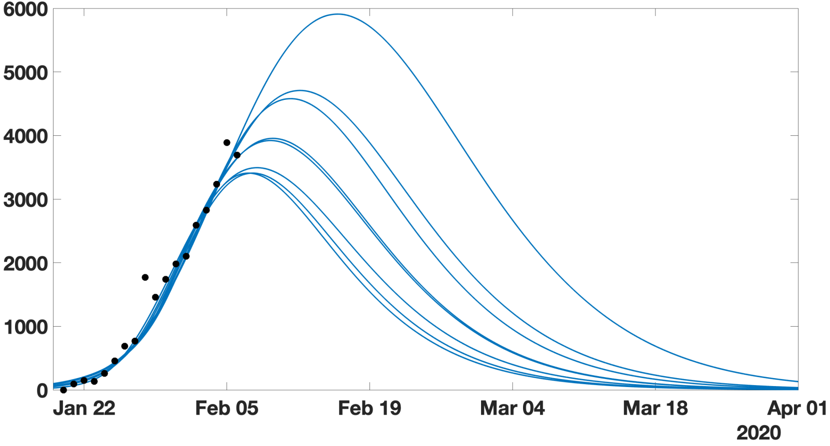
Cumulative data for China until February 6 with $f = 0.6$



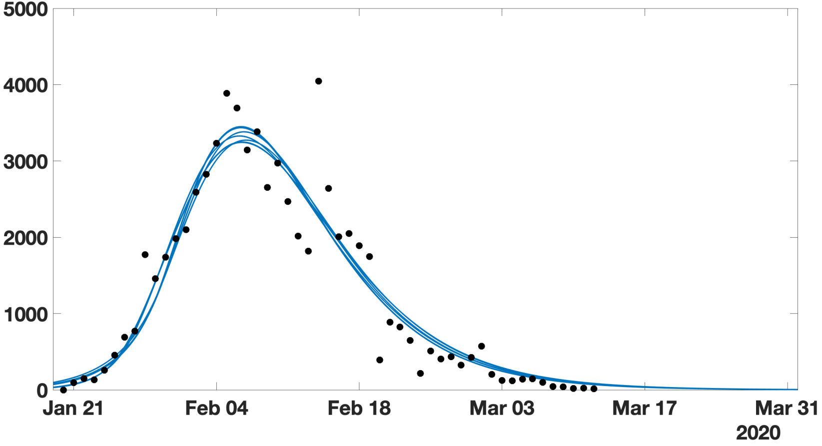
Cumulative data for China until March 12 with $f = 0.6$



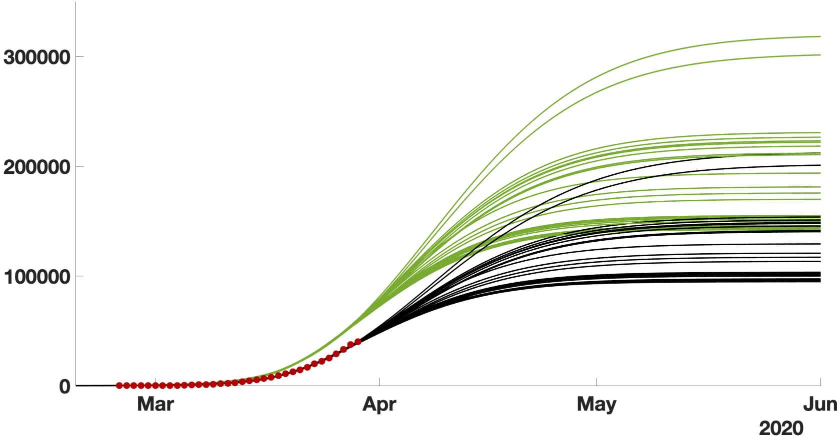
Daily data for China until February 6 with $f = 0.6$



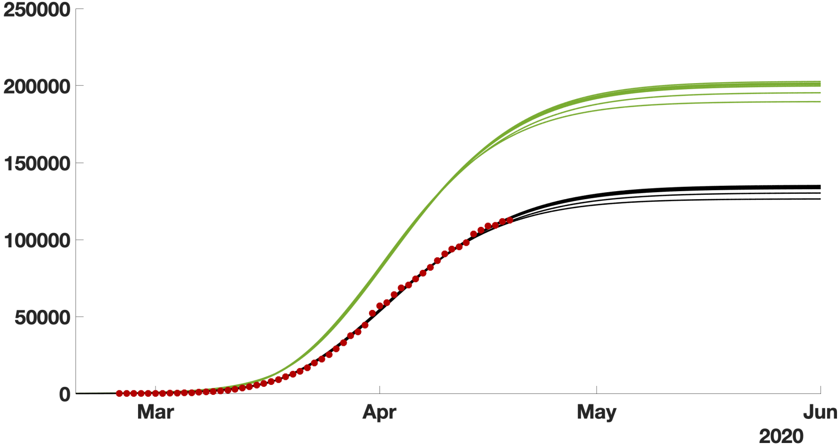
Daily data for China until March 12 with $f = 0.6$



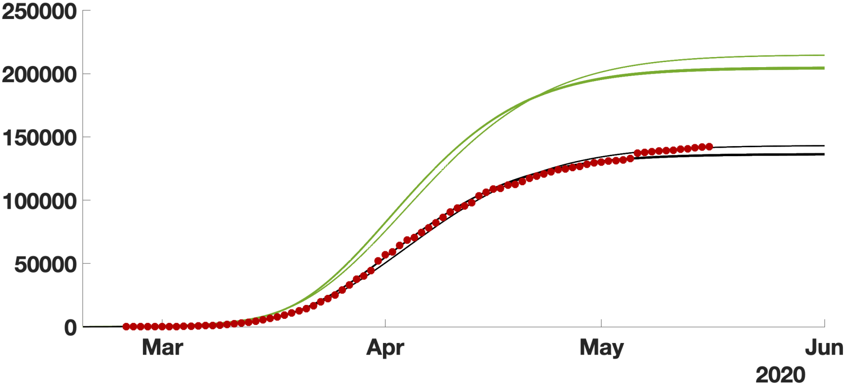
Cumulative data for France until Mars 30 with $f = 0.4$



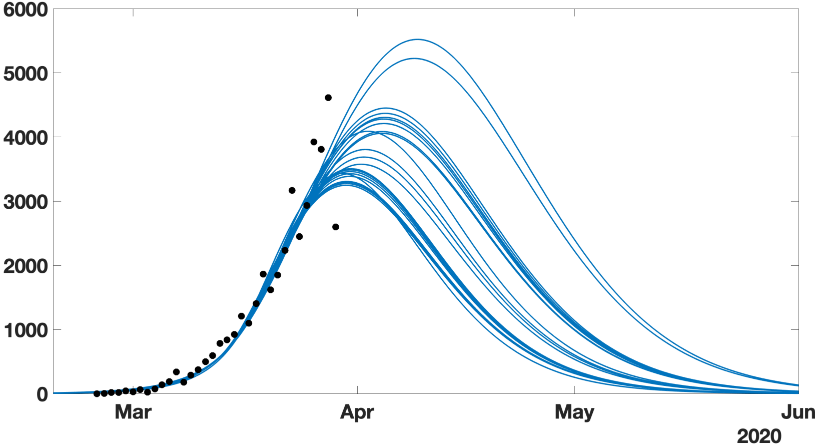
Cumulative data for France until April 20 with $f = 0.4$



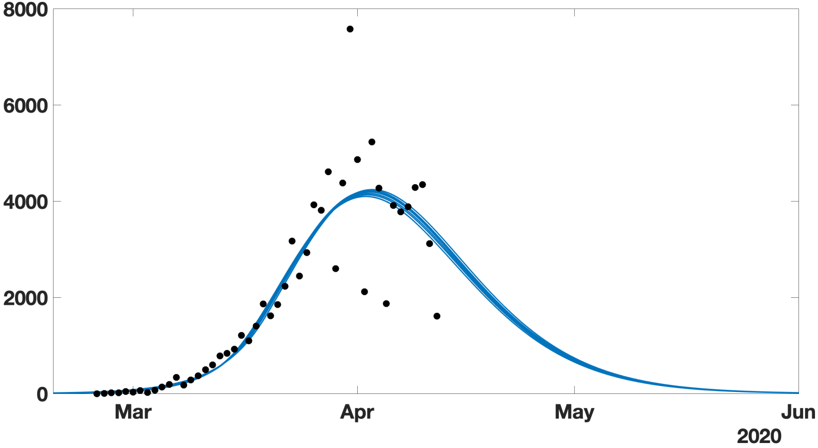
Cumulative data for France until Mai 17 with $f = 0.4$



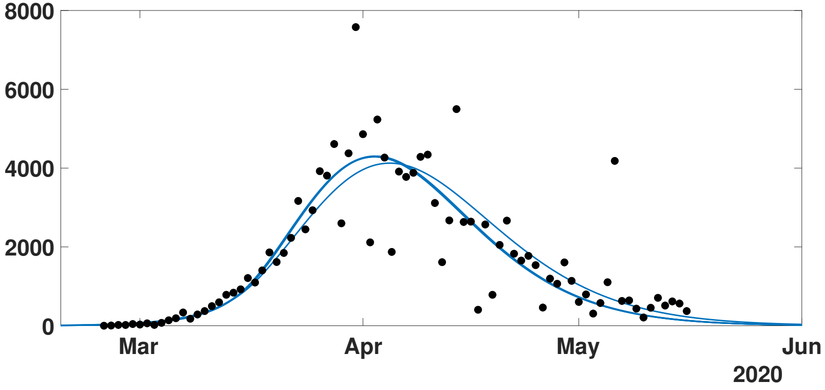
Daily data for France until Mars 30 with $f = 0.4$



Daily data for France until April 20 with $f = 0.4$



Daily data for France until Mai 17 with $f = 0.4$



Outline

1 Introduction

2 Results

3 Numerical Simulations

4 Age dependency in COVID-19 for Japan

Early models with age designed for the COVID-19

- Prem, Liu, Russell, *et al.*, [10] They use an SIR model with age classes. They use a matrix of contacts which is obtained from real data. No comparison of their model with time dependent age structured data is presented.
- There are more results about age and COVID-19 Ayoub et al. [11] and Chikina and Pegden [12]

[10] K. Prem, Y. Liu, T. W Russell, *et al.*, The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study, *The Lancet Public Health* **5(5)** (2020).

[11] H. H. Ayoub, *et al.*, Age could be driving variable SARS-CoV-2 epidemic trajectories worldwide, *medRxiv* (2020).

[12] M. Chikina and W. Pegden, Modeling strict age-targeted mitigation strategies for COVID-19, *arXiv* (2020).

Age dependence on the number of reported case of COVID-19 in Japan [13]

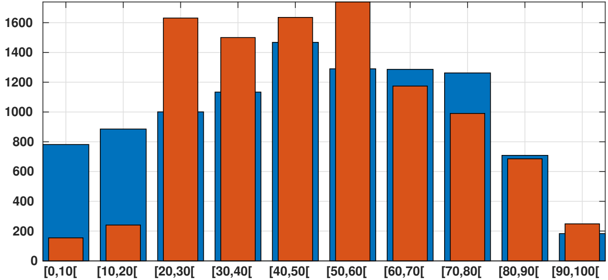


Figure: *In this figure we plot in blue the age distribution of the Japanese population for* $10\,000$ *people and we plot in orange the age distribution of the number of reported cases of SARS-CoV-2 for* 13660 *patients on April 29. We observe that* 77% *of the confirmed patients belong to the 20–60 years age class.*

[13] `https://covid19japan.com/`

Multiple exponential growth of cumulative reported number of case per age classe

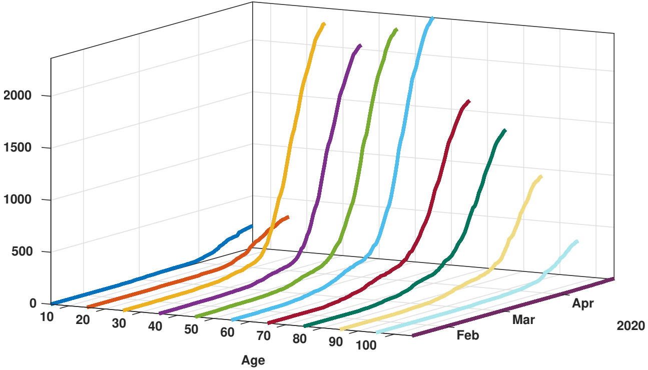


Figure: *Time evolution of the cumulative number of reported cases of SARS-CoV-2 per age class.*

Multiple exponential growth of cumulative reported number of case per age classe

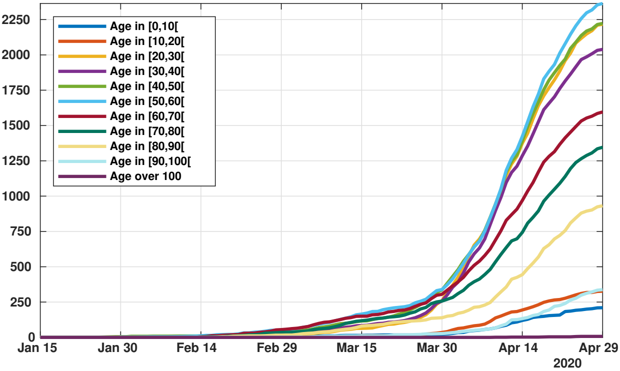


Figure: *Time evolution of the cumulative number of reported cases of SARS-CoV-2 per age class.*

Age dependence on the number of death due to COVID-19 in Japan [14]

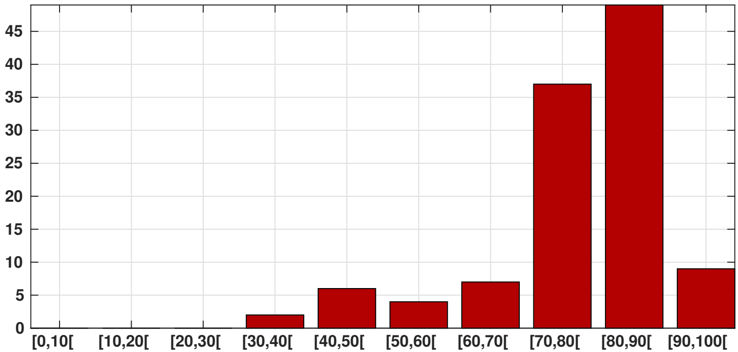


Figure: *Cumulated number of SARS-CoV-2-induced deaths per age class. We observe that* 83% *of death occur in between* 70 *and* 100 *years old.*

[14] `https://covid19japan.com/`

Model with age structure

We consider $N_1, \dots, N_{10}$ the number of individuals respectively for the age classes $[0, 10[, \dots,$
The model for the number of susceptible individuals $S_1(t), \dots, S_{10}(t)$, respectively for the age classes $[0, 10[, \dots, [90, 100[$, is the following

$$\begin{cases} S_1'(t) = -\tau_1 S_1(t)\left[\phi_{1,1}\dfrac{(I_1(t)+U_1(t))}{N_1} + \ldots + \phi_{1,10}\dfrac{(I_{10}(t)+U_{10}(t))}{N_{10}}\right], \\ \vdots \\ S_{10}'(t) = -\tau_{10} S_{10}(t)\left[\phi_{10,1}\dfrac{(I_1(t)+U_1(t))}{N_1} + \ldots + \phi_{10,10}\dfrac{(I_{10}(t)+U_{10}(t))}{N_{10}}\right]. \end{cases} \tag{4.1}$$

The model for the number of asymptomatic infectious individuals $I_1(t), \dots, I_{10}(t)$, respectively for the age classes $[0, 10[, \dots, [90, 100[$, is the following

$$\begin{cases} I_1'(t) = \tau_1 S_1(t)\left[\phi_{1,1}\dfrac{(I_1(t)+U_1(t))}{N_1} + \ldots + \phi_{1,10}\dfrac{(I_{10}(t)+U_{10}(t))}{N_{10}}\right] - \nu I_1(t), \\ \vdots \\ I_{10}'(t) = \tau_{10} S_{10}(t)\left[\phi_{10,1}\dfrac{(I_1(t)+U_1(t))}{N_1} + \ldots + \phi_{10,10}\dfrac{(I_{10}(t)+U_{10}(t))}{N_{10}}\right] - \nu I_{10}(t). \end{cases} \tag{4.2}$$

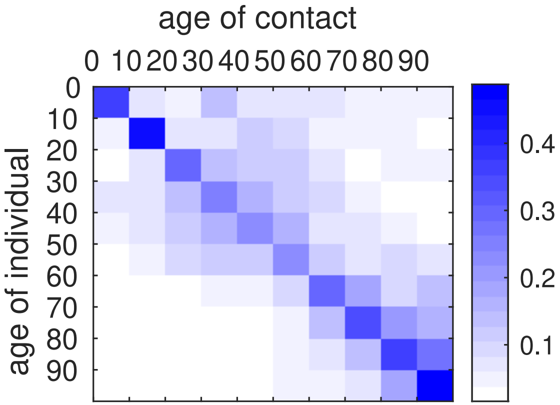
The model for the number of reported symptomatic infectious individuals $R_1(t), \dots, R_{10}(t)$, respectively for the age classes $[0, 10[, \dots, [90, 100[$, is

$$\begin{cases} R_1'(t) = \nu_1^1 \, I_1(t) - \eta R_1(t), \\ \vdots \\ R_{10}'(t) = \nu_1^{10} \, I_{10}(t) - \eta R_{10}(t). \end{cases} \tag{4.3}$$

Finally the model for the number of unreported symptomatic infectious individuals $U_1(t), \dots$, $U_{10}(t)$, respectively in the age classes $[0, 10[, \dots, [90, 100[$, is the following

$$\begin{cases} U_1'(t) = \nu_2^1 \, I_1(t) - \eta U_1(t), \\ \vdots \\ U_{10}'(t) = \nu_2^{10} \, I_{10}(t) - \eta U_{10}(t). \end{cases} \tag{4.4}$$

Thanks to Prem, Cook and Jit [15] we obtain the matrix of conditional probability $\phi_{i,j}$ of contact between age classes which is the following



[15] K. Prem, A.R. Cook, M. Jit, Projecting social contact matrices in 152 countries using contact surveys and demographic data, *PLoS Computational Biology* **13(9)** (2017)

Dependency of other parameters

In order to describe the confinement for the age structured model (4.1)-(4.4) we will use for each age class $i = 1, \ldots, 10$ a different transmission rate having the following form

$$\begin{cases} \tau_i(t) = \tau_i,\, 0 \leq \, t \, \leq \, D_i, \\ \tau_i(t) = \tau_i \, \exp\left(-\mu_i \left(t - D_i\right)\right),\, D_i < t. \end{cases} \tag{4.5}$$

The date D_i is the first day of public intervention for the age class i and μ_i is the intensity of the public intervention for each age class.
The parameter f_i (probability to become reported) is also assumed to be dependent on the age class.

Best fit to the data from Japan

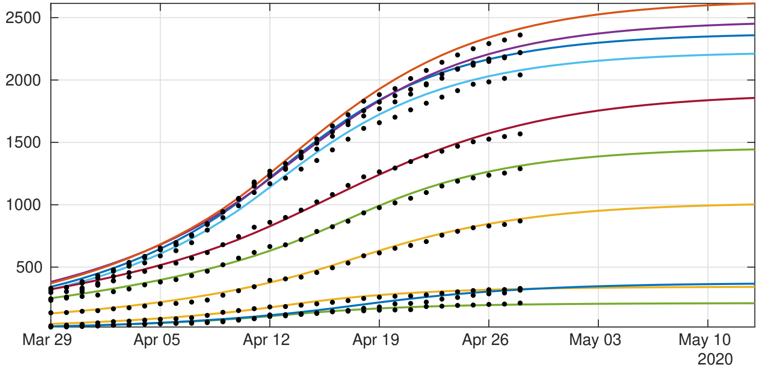


Figure: *In this figure we compare the 10 age classes coming to the data (black dots) and the 10 age classes coming for the model (color curves)*

Best fit to the data from Japan

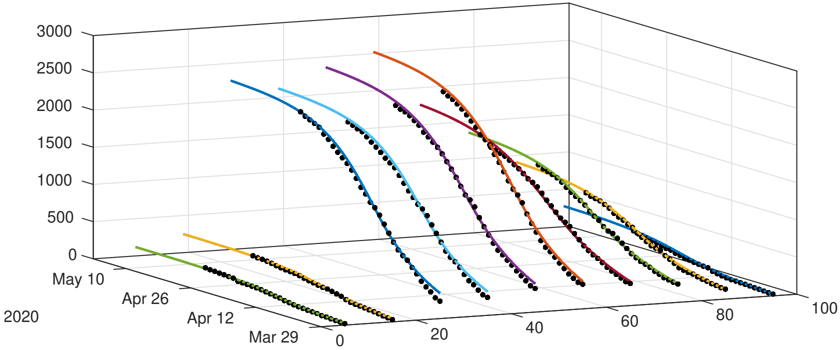
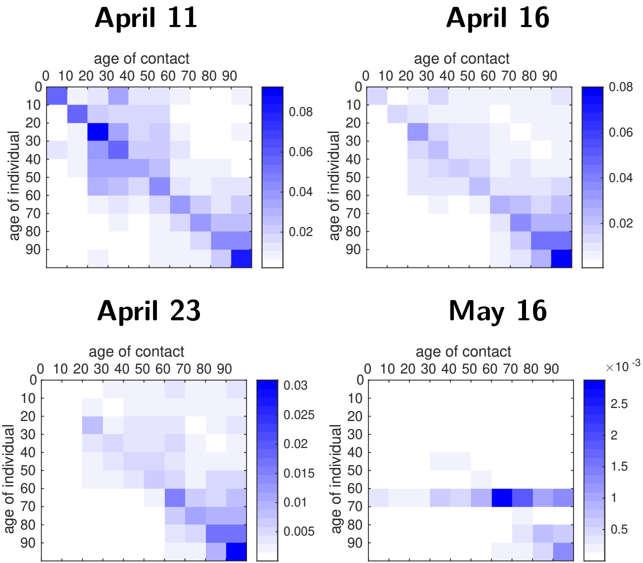


Figure: *In this figure we compare the 10 age classes coming to the data (black dots) and the 10 age classes coming for the model (color curves)*

Transmission matrices



References

- **Z. Liu, P. Magal, O. Seydi, and G. Webb**, Understanding unreported cases in the COVID-19 epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *MPDI Biology* 2020, 9(3), 50.
- **Z. Liu, P. Magal, O. Seydi, and G. Webb**, Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data, *Mathematical Biosciences and Engineering* **17(4)** 2020, 3040-3051.
- **Z. Liu, P. Magal, O. Seydi, and G. Webb**, A COVID-19 epidemic model with latency period, *Infectious Disease Modelling (to appear)*
- **Z. Liu, P. Magal, O. Seydi, and G. Webb**, A model to predict COVID-19 epidemics with applications to South Korea, Italy, and Spain, *SIAM News (to appear)*
- **Z. Liu, P. Magal and G. Webb**, Predicting the number of reported and unreported cases for the COVID-19 epidemic in China, South Korea, Italy, France, Germany and United Kingdom, *medRxiv*
- **Q. Griette, Z. Liu and P. Magal**, Estimating the last day for COVID-19 outbreak in mainland China, *medRxiv*.
- **Q. Griette, Z. Liu and P. Magal**, Unreported cases for Age Dependent COVID-19 Outbreak in Japan, *medRxiv*