Can mathematical modeling help to understand COVID-19 data?

Pierre Magal

University of Bordeaux, France

pierre.magal@u-bordeaux.fr

Summer Bootcamp of Infectious Disease Modelling 2023

Kyoto University, August, 2023.

Abstract

The COVID-19 outbreak, which started in late December 2019 and rapidly spread around the world, has been accompanied by an unprecedented release of data on reported cases. Our objective is to offer a fresh look at these data by coupling a phenomenological description to the epidemiological dynamics. We use a phenomenological model to describe and regularize the reported cases data. This phenomenological model is combined with an epidemic model with a time-dependent transmission rate. The time-dependent rate of transmission involves changes in social interactions between people as well as changes in host-pathogen interactions. Our method is applied to cumulative data on reported cases for eight different geographic areas. In the eight geographic areas considered, successive epidemic waves are matched with a phenomenological model and are connected to each other. We find a single epidemic model that coincides with the best fit to the data of the phenomenological model.

Scope & aims

- How to recover information from the **cumulative number of reported cases data**?
- We want to use simple models with a **limited number of parameters**.
- Here the parameters include part of the **initial conditions**.
- We want to **reconstruct** and **forecast** the epidemic.

Outline

Example of unreported cases

A published study[1] traced COVID-19 infections resulting from a business meeting in Germany attended by a person who was infected but had no symptoms at the time. Four people were eventually infected from this single contact.

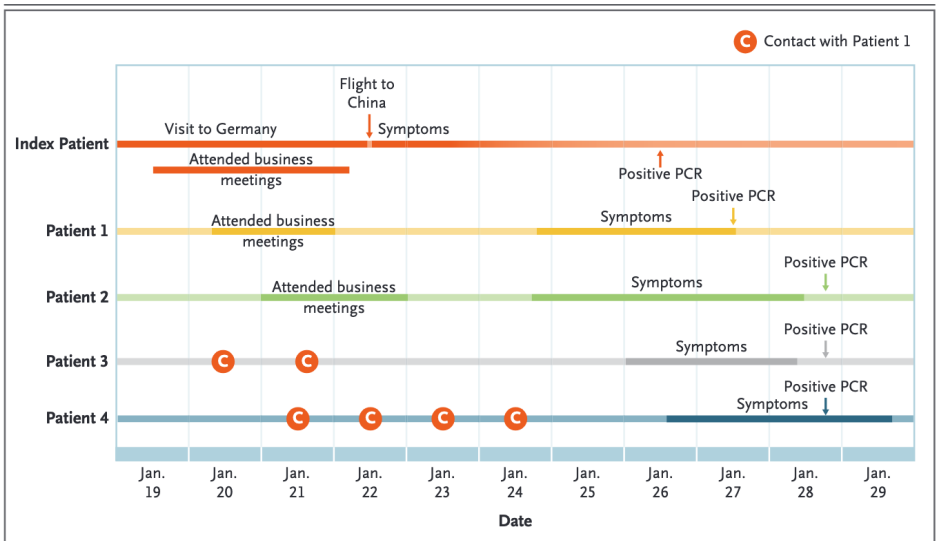


Figure 1. Timeline of Exposure to Index Patient with Asymptomatic 2019-CoV Infection in Germany.

[1] **Rothe, et al.** (2020), Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*, **382(10)**, 970-971.

Example of unreported cases

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

On the French *aircraft carrier Charles de Gaulle*, clinical and biological data for all 1739 crew members were collected on arrival at the Toulon harbor and during quarantine: 1121 crew members (64%) were tested positive for COVID-19 using RT-PCR, and among these, 24% were asymptomatic[3].

[2] **H. Nishiura, N. M. Linton, & A. R. Akhmetzhanov** (2020), Serial interval of novel coronavirus (COVID-19) infections, *Int. J. Infect. Dis.*, **93**, 284-286.

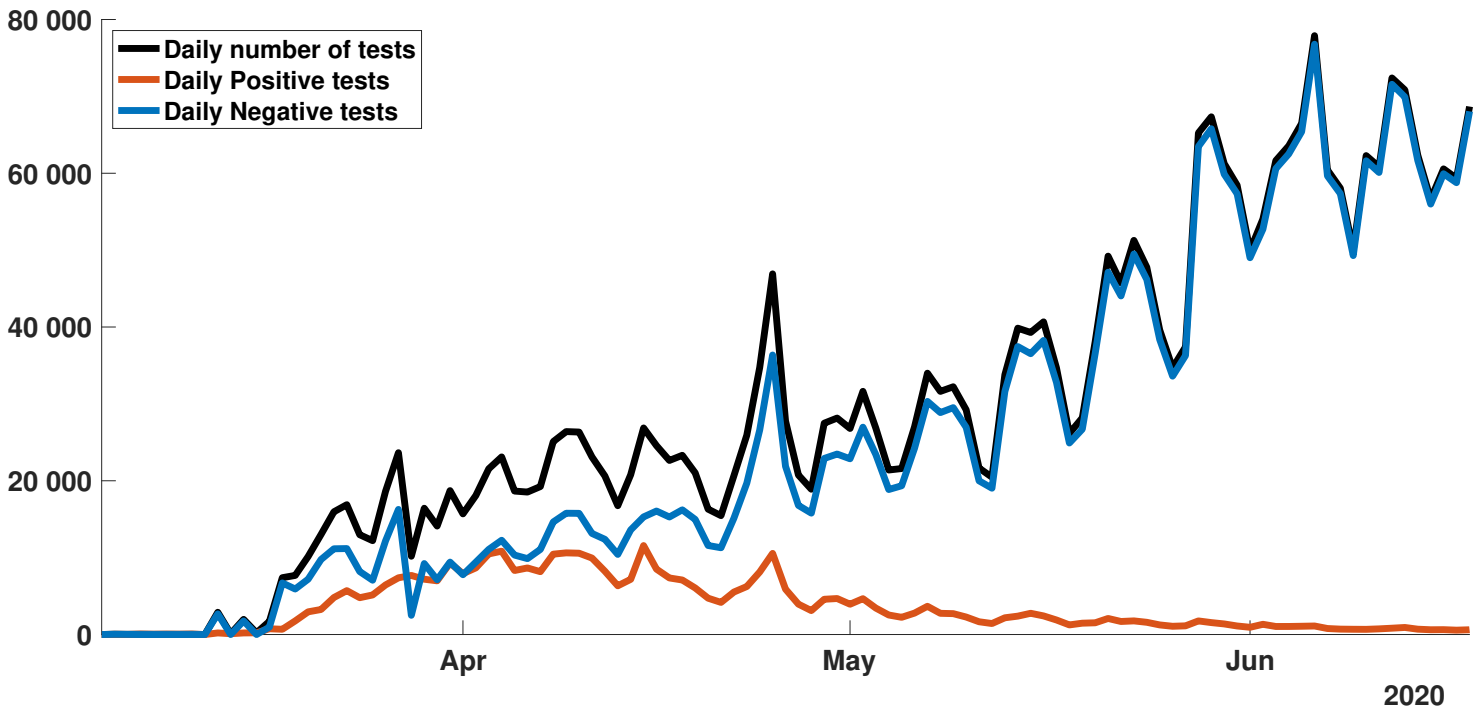
[3] **O. Bylicki, N. Paleiron, and F. Janvier** (2021), An Outbreak of Covid-19 on an Aircraft Carrier. *New Engl. J. Med.*, **384(10)**, 976–977.

What are the unreported cases?

- **Mild symptoms** induce unreported cases because people will only get tested in case of severe symptoms.
- Unreported cases are partly due to a low daily number of tests.

Testing data for New York state

The dynamic of the daily number of tests is connected to the dynamic of the daily number of reported cases in a complex way[4].



[4] **Q. Griette and P. Magal** (2021) Clarifying predictions for COVID-19 from testing data: the example of New York State, *Infectious Disease Modelling*, **6**, 273-283.

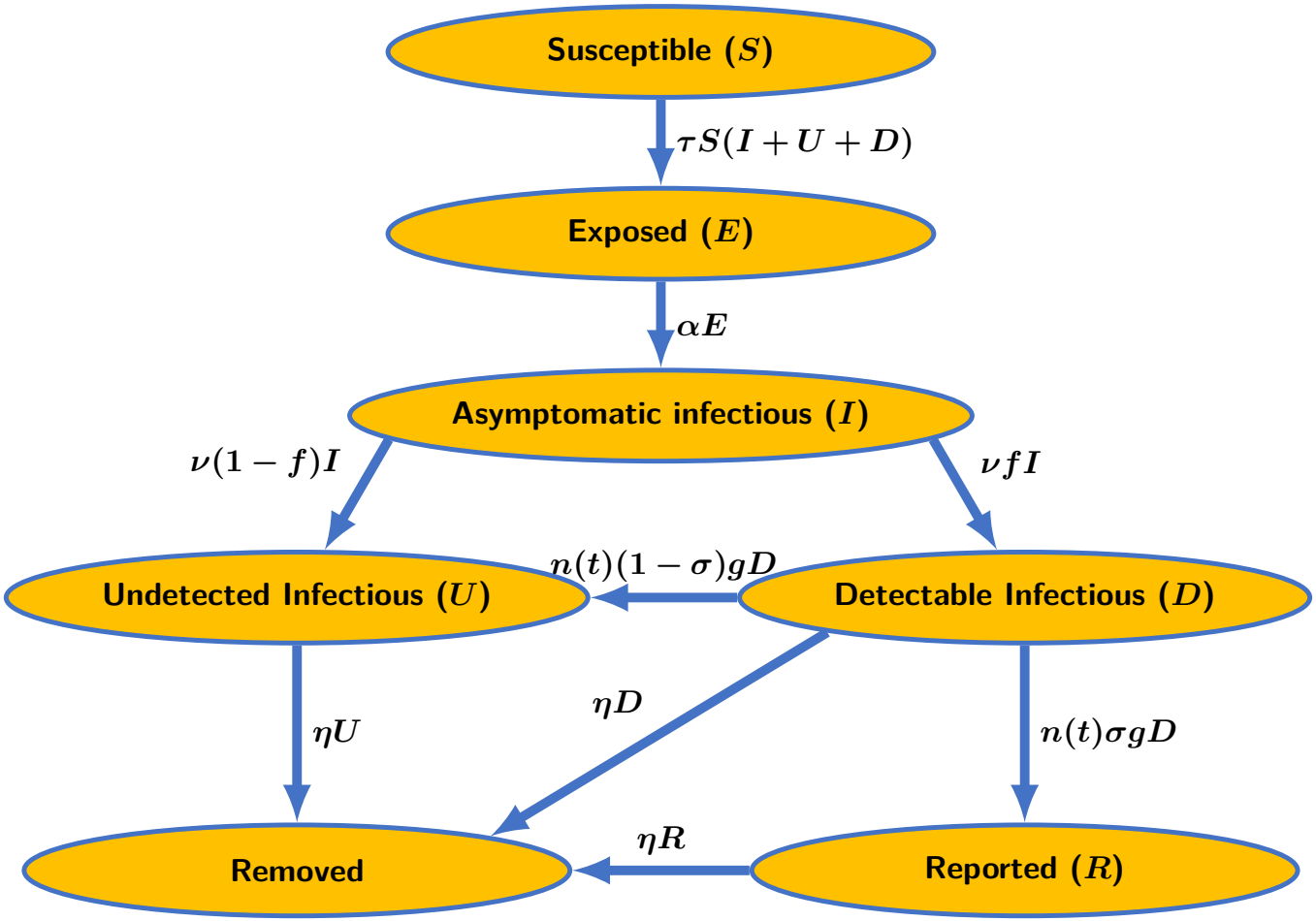
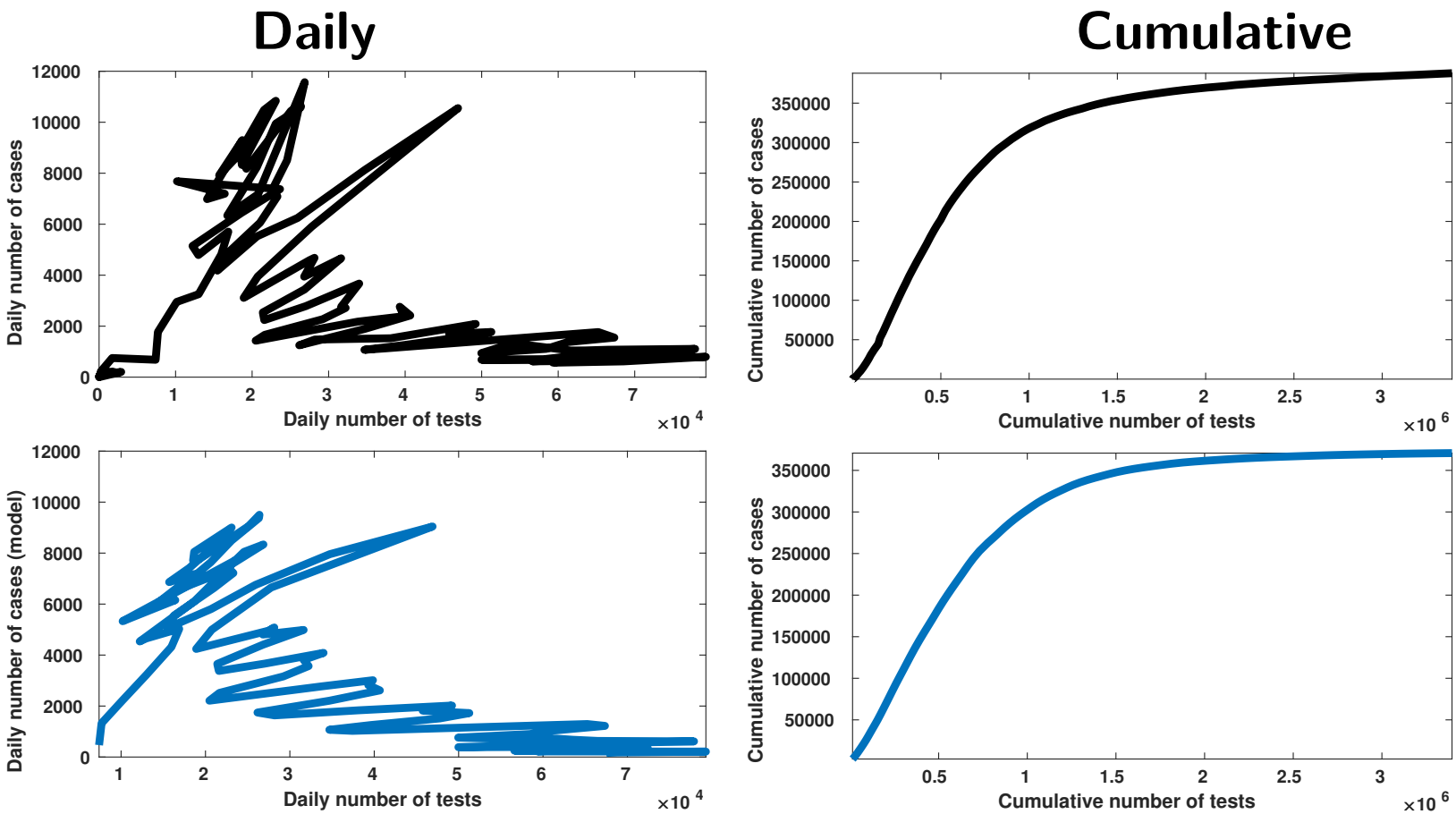


Figure: *Flow chart of the epidemic model with tests.*

Testing data for New York state[5]



The **black curves** are produced by using **the data only**. The **blue curves** are produced by using **the model** with **the testing data**.

[5] **Q. Griette and P. Magal** (2021) Clarifying predictions for COVID-19 from testing data: the example of New York State, *Infectious Disease Modelling*, **6**, 273-283.

PCR tests can lead to significant bias for false positives and true positives

- Mathematical model[6];
- Mathematical model combined with experimental studies [7];
- Experimental studies [8,9].

---

[6] **J. Peccoud, & C. Jacob** (1996) Theoretical uncertainty of measurements using quantitative polymerase chain reaction. *Biophysical journal*, **71(1)**, 101-108.

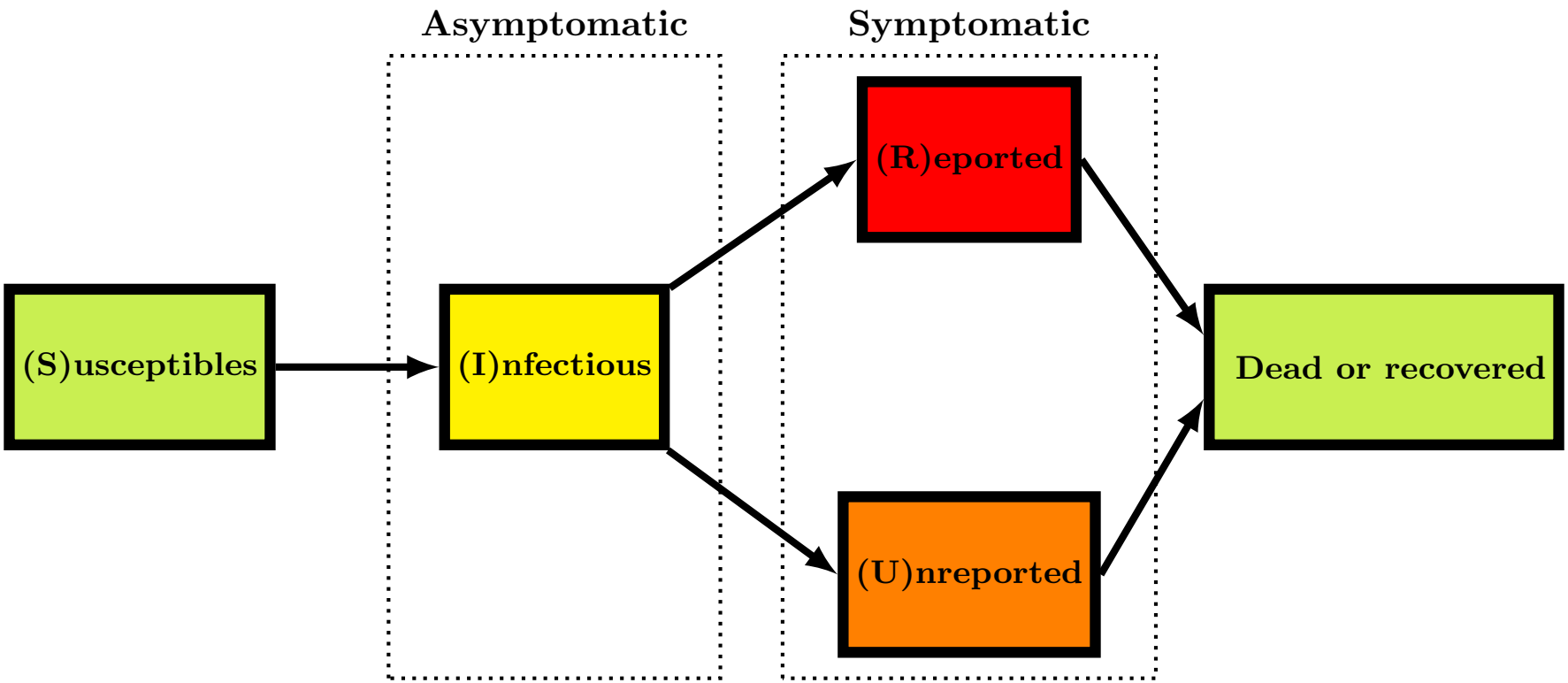
[7] **J. Hellewell, T. W. Russell, R. Beale, G. Kelly, C. Houlihan, E. Nastouli, & A.J. Kucharski** (2021) Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections, *BMC medicine*, **19(1)**, 1-10.

[8] **L.M. Kucirka, S.A. Lauer, O. Laeyendecker, D. Boon, & J. Lessler,** (2020) Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. *Annals of internal medicine*, **173(4)**, 262-267.

[9] **P. P. Salvatore et al.** (2023) Transmission potential of vaccinated and unvaccinated persons infected with the SARS-CoV-2 Delta variant in a federal prison, *Vaccine*, **41** 1808-1818.

Outline

Epidemic model with Unreported Cases[10,11]



[10] **Z. Liu, P. Magal, O. Seydi, and G. Webb** (2020), Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9(3), 50**.

[11] **J. Arino, F. Brauer, P. van den Driessche, J. Watmough and J. Wu** (2006), Simple models for containment of a pandemic, *Journal of the Royal Society Interface*, **3(8)**, 453-457.

Epidemic model

Transmissions between infectious and susceptible individuals are described by

$$\begin{cases} S'(t) = -\tau(t)\, S(t)\, I(t), \\ I'(t) = \tau(t)\, S(t)\, I(t) - \nu\, I(t), \end{cases} \tag{1}$$

where

- $\tau(t)$ is the rate of transmission.
- $1/\nu$ is the average duration of the asymptomatic infectious period.
- $\tau(t)\, S(t)\, I(t)$ is the flux of S-individuals becoming infected at time t.
- $\nu\, I(t)$ is the flux of I-individuals leaving the I-compartment.

Initial distribution of the model

The system (1) is complemented with the initial distribution of the model

$$S(t_0) = S_0 \geq 0,\ I(t_0) = I_0 \geq 0. \tag{2}$$

The parameter

$$t_0$$

is also unknown.

That is the time t_0 from which the epidemic model (1) becomes applicable.

Connecting the data and the model[12,13]

To connect the data and the model (1) we use the following equation

$$\mathrm{CR}'(t) = f\,\nu\,I(t), \text{ for } t \geq t_0, \tag{3}$$

where f is the **fraction of reported individuals**.

We assume that

- f is the fraction of patients with **severe symptoms**.
- $1 - f$ is the fraction of of patients with **mild symptoms**.

[12] **Z. Liu, P. Magal, O. Seydi, and G. Webb** (2020), Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9(3), 50**.

[13] **P. Magal, and G. Webb** (2018) The parameter identification problem for SIR epidemic models: Identifying Unreported Cases, *Journal of Mathematical Biology* **77(6-7)**, 1629–1648.

Given Parameters

- Number of susceptible individuals when the epidemic starts

$$S_0 = 67 \text{ millions for France.}$$

- Time from which the epidemic model starts to be valid, also called initial time of the model t_0.
- The average duration of the infectiousness $\frac{1}{\nu} = 3$ days.
- The fraction of reported individuals $f = 0.9$.

Computed parameters

- I_0 the number of asymptomatic infectious patients at the start of the epidemic.

- $\tau(t)$ the rate of transmission.

What factors govern the transmission rate $\tau(t)$?

As explained in Magal and Ruan[14] by using stochastic individual based models

$$\tau(t) = \frac{\text{The probability of transmission}}{\text{The average duration of a contact}}.$$

Contact patterns are impacted by **social distancing measures**.

The average number of contacts per unit of time depends on the **density of population**[15,16].

[14] **P. Magal and S. Ruan** (2014), Susceptible-Infectious-Recovered Models Revisited: From the Individual Level to the Population Level, *Mathematical Biosciences* **250**, 26-40.

[15] **J. Rocklöv, & H. Sjödin.** (2020), High population densities catalyse the spread of COVID-19. *J Travel Med*, **27(3)** , taaa038.

[16] **H. Seligmann, N. Vuillerme & J. Demongeot** (2020), Summer COVID-19 third wave: faster high altitude spread suggests high UV adaptation, *medRxiv*.

What factors govern the transmission rate $\tau(t)$?

- The probability of transmission depends of the virulence of the pathogen which can depend on the **temperature, the humidity, and the Ultraviolet**[17,18].
- The probability of transmission depends of the susceptibility of the individuals
  - **Blood group**[19] : Blood group O is associated with a lower susceptibility to SARS-CoV2;
  - **Genetic lineage**[20] A gene cluster inherited from Neanderthal has been identified as a risk factor for severe symptoms.

[17] **J. Demongeot, Y. Flet-Berliac, & H. Seligmann** (2020), Temperature Decreases Spread Parameters of the New Covid-19 Case Dynamics, *Biology*, **9**, 94.

[18] **J. Wang, et al** (2020), High temperature and high humidity reduce the transmission of COVID-19. *Available at SSRN 3551767*.

[19] **P. Guillon, et al.** (2008), Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by anti-histo-blood group antibodies, *Glycobiology* **18.12**, 1085-1093.

[20] **H. Zeberg and S. Pääbo,** (2020), The major genetic risk factor for severe COVID-19 is inherited from Neanderthals, *Nature*.

Outline

Modeling the exponential phase

At the early stage of the epidemic, we can assume that $S(t)$ is constant, and equal to S_0. We can also assume that $\tau(t)$ remains constant equal to $\tau_0 = \tau(t_0)$. Therefore, by replacing these parameters into the I-equation of system (1) we obtain

$$I'(t) = (\tau_0 S_0 - \nu) I(t).$$

Therefore

$$I(t) = I_0 \exp\left(\chi_2 \left(t - t_0\right)\right),$$

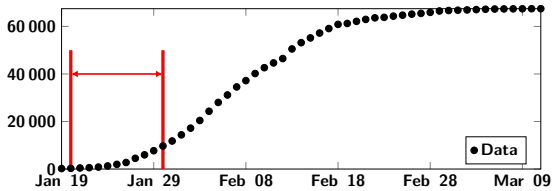
where

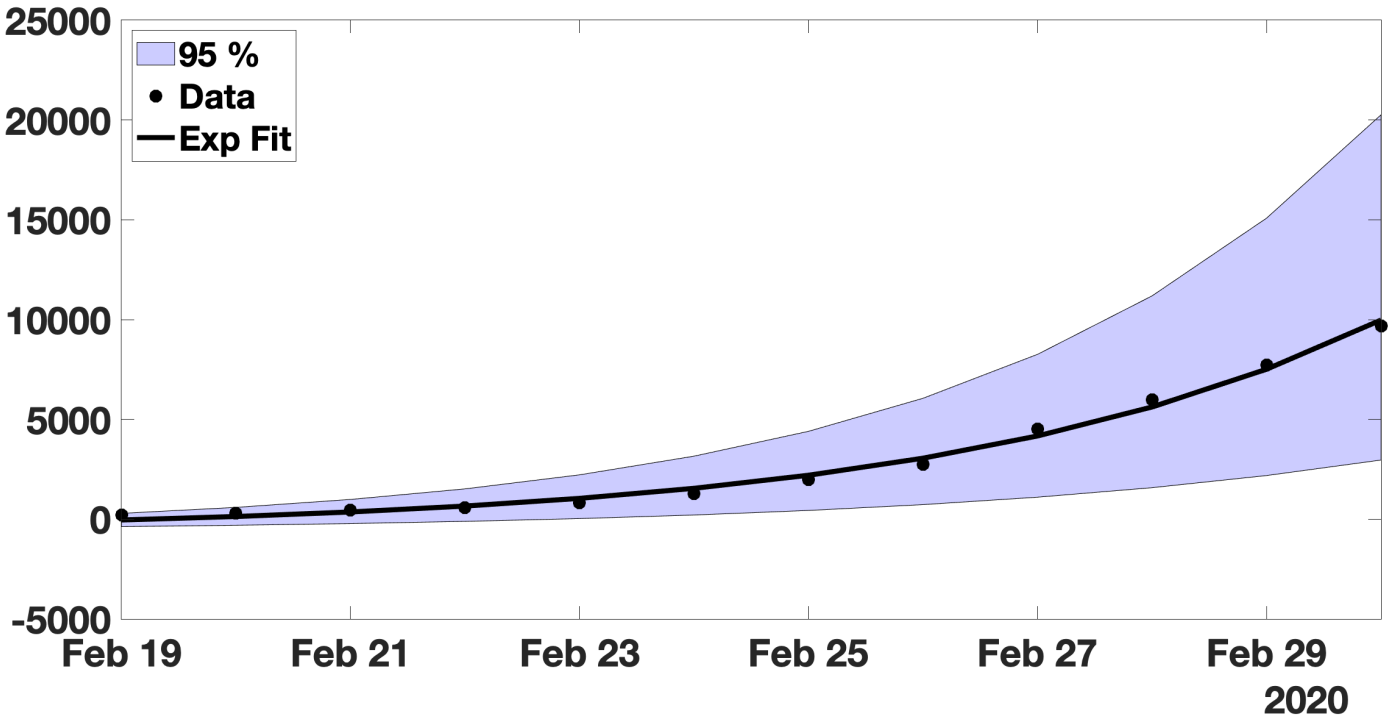
$$\chi_2 = \tau_0 S_0 - \nu.$$

By using (3), we obtain

$$\mathrm{CR}(t) = \chi_1 e^{\chi_2 t} - \chi_3. \tag{4}$$

Application to COVID-19 in mainland China [21,22]





[21] **Z. Liu, P. Magal, O. Seydi, and G. Webb** (2020), Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9(3), 50**.

[22] **J. Demongeot, Q. Griette and P. Magal** (2020), SI epidemic model applied to COVID-19 data in mainland China, *Royal Society Open Science* 7:201878.

Initial number of infected and transmission rate

Remember that (3) and (4) are respectively

$$\mathrm{CR}'(t) = f\,\nu\,I(t), \text{ for } t \geq t_0, \quad \text{and} \quad \mathrm{CR}(t) = \chi_1 e^{\chi_2 t} - \chi_3.$$

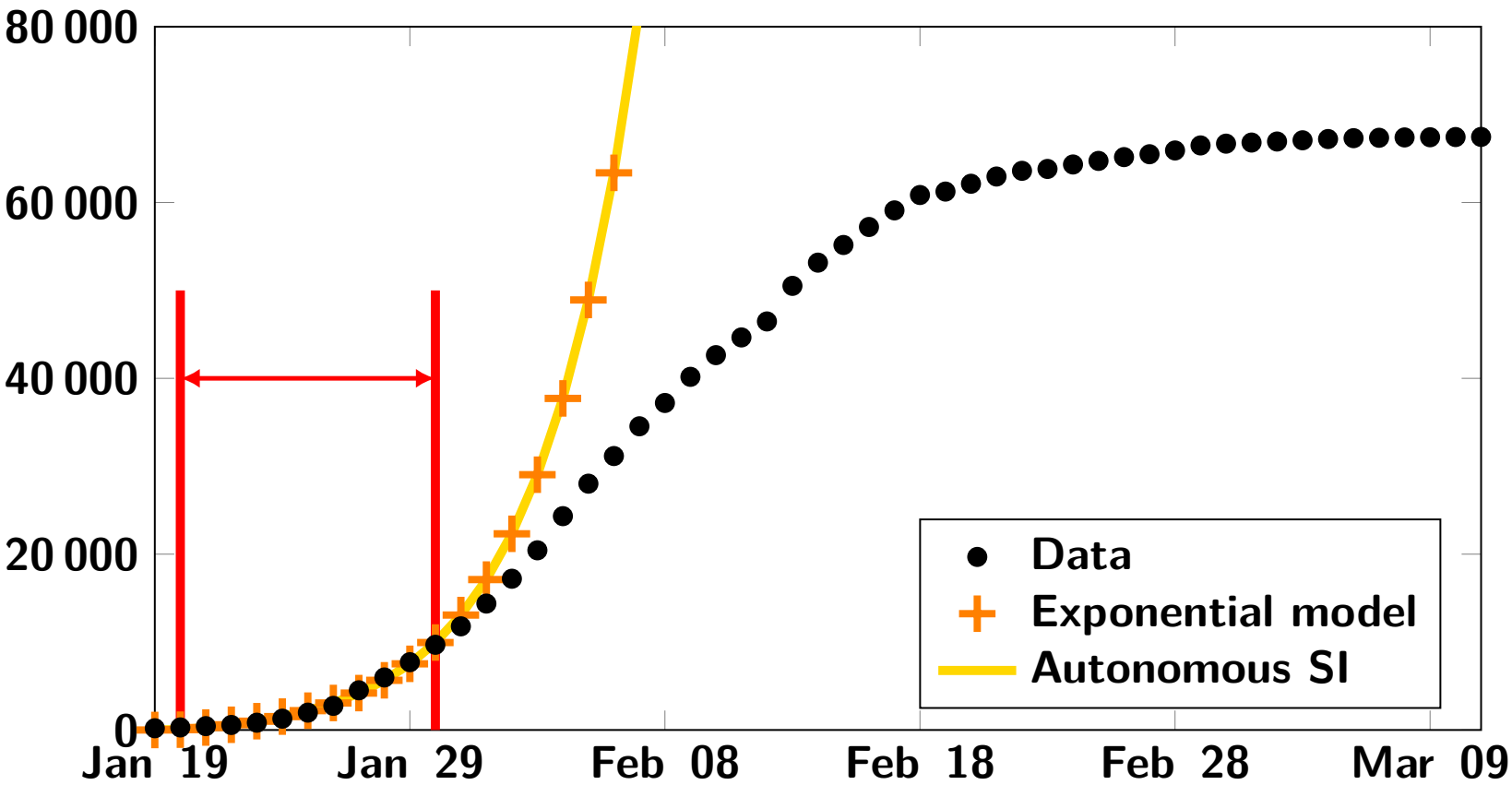
By using (3) and (4) we obtain

$$I_0 = \frac{\mathrm{CR}'(t_0)}{\nu\,f} = \frac{\chi_1\,\chi_2 e^{\chi_2\,t_0}}{\nu\,f},$$

and by using (4)

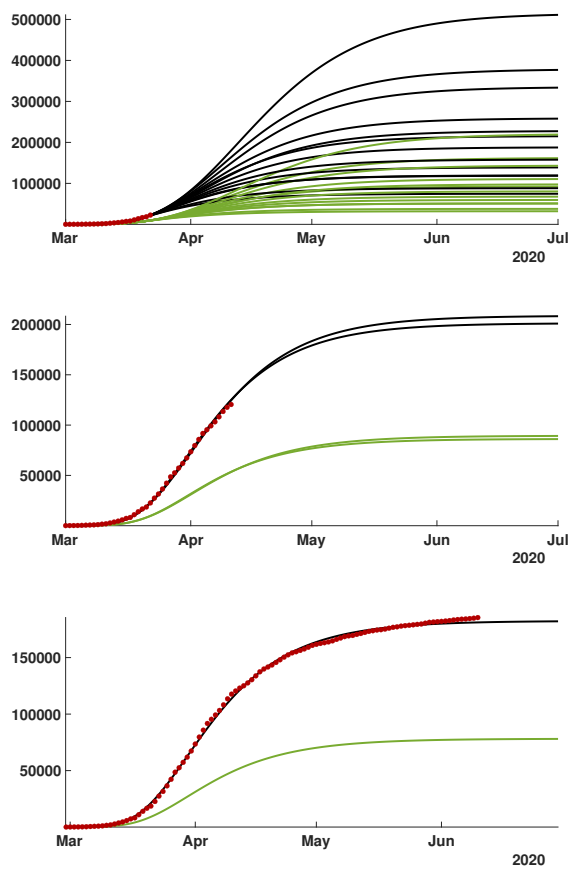
$$\tau_0 = \frac{\chi_2 + \nu}{S_0}.$$

Why do we need a time-dependent transmission rate?

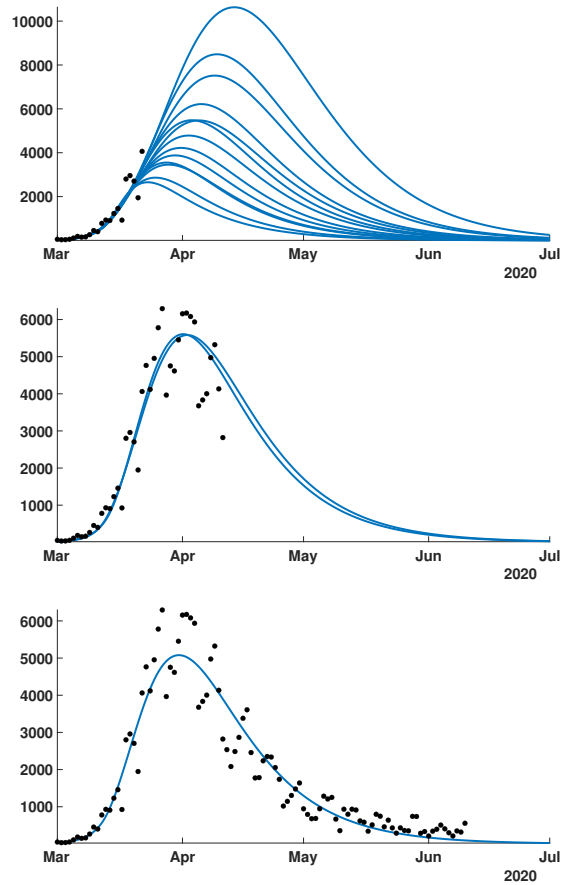


Uncertainty due to the period chosen to fit the data[23]

Cumulative number cases



Daily number of cases



[23] **Z. Liu, P. Magal, G. Webb** (2021) Predicting the number of reported and unreported cases for the COVID-19 epidemics in China, South Korea, Italy, France, Germany and United Kingdom, *Journal of Theoretical Biology* **509, 21**.

Outline

Earlier results with a transmission rate reconstructed from the data

This problem has already been considered in several articles. In the early 70s, London and Yorke[24,25] discussed the time dependent rate of transmission in the context of measles, chickenpox and mumps.

Motivated by applications to the data for COVID-19 the group of Bakhta, Boiveau, Maday, & Mula[26] also obtained some new results about reconstructing the rate of transmission.

---

[24] **W. P. London, and J. A. Yorke** (1973), Recurrent outbreaks of measles, chickenpox and mumps: I. Seasonal variation in contact rates. *Am J Epidemiol*, **98(6)**, 453-468.

[25] **J. A. Yorke, and W. P. London** (1973), Recurrent outbreaks of measles, chickenpox and mumps: II. Systematic differences in contact rates and stochastic effects. *Am J Epidemiol*, **98(6)**, 469-482.

[26] **A. Bakhta, T. Boiveau, Y. Maday, & O. Mula** (2021), Epidemiological Forecasting with Model Reduction of Compartmental Models. Application to the COVID-19 Pandemic. *Biology*, **10(1)**, 22.

Results for mainland China

Results

Our results show that

- It is not possible to estimate the exact value of the mean duration of the infectious period using the model.
- Multiple values of $1/\nu$ give exactly the fit to the data.
- We can obtain a significant upper limit for the average duration of the infectious period using the model.
- In the case of Sars-CoV2 in mainland China, this upper limit is 3.3 days.

Results

Our results show that

- Our results show that it is impossible to estimate the exact value of the fraction of individuals reported using the model.
- Multiple values of f give the same fit the data.

Phenomenological Model for the Epidemic Phase[27,28]

Second phenomenological model

Consider the model of **Bernoulli-Verhulst**

$$\mathrm{CR}_{\mathrm{Data}}(t) = N_{\mathrm{base}} + \frac{e^{\chi(t-t_0)} N_0}{\left[1 + \frac{N_0^\theta}{N_\infty^\theta}\left(e^{\chi\theta(t-t_0)} - 1\right)\right]^{1/\theta}}, \text{ for } t \in [t_0, t_1]. \quad (5)$$

[27] **D. Bernoulli**, Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. *Mémoire Académie Royale des Sciences*, Paris (1760).

[28] **P.F. Verhulst**, Notice sur la loi que la population poursuit dans son accroissement. *Correspondance Mathématique et Physique*, vol.X (1838), 113-121.

Bernoulli-Verhulst equation

Namely, by setting

$$N(t) = \mathrm{CR}(t) - N_{\mathrm{base}},$$

we obtain

$$N'(t) = \chi\, N(t) \left[1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right],$$

completed with the initial value

$$N(t_0) = N_0.$$

In the model, $N_{\mathrm{base}} + N_0$ corresponds to the value $\mathrm{CR}(t_0)$ of the cumulative number of cases at time $t = t_0$. The parameter $N_\infty + N_{\mathrm{base}}$ is the maximal value of the cumulative reported cases after the time $t = t_0$.

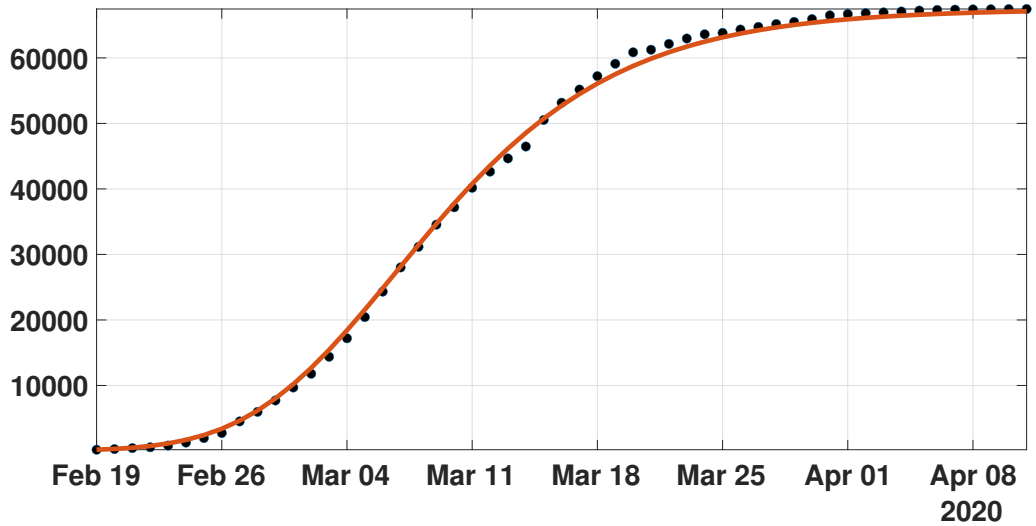


Figure: *In this figure, we plot the best fit of the Bernoulli-Verhulst model to the cumulative number of reported cases of COVID-19 in China. We obtain* $\chi_2 = 0.66$ *and* $\theta = 0.22$*. The black dots correspond to data for the cumulative number of reported cases and the blue curve corresponds to the model.*

Computed parameters[29]

Results

The number of **infectious asymptomatic** individuals is

$$I_0 = \frac{\mathrm{CR}'_{\mathrm{Data}}(t_0)}{\nu\, f} \tag{6}$$

The **rate of transmission** $\tau(t)$ at time t is

$$\tau(t) = \frac{\nu f \left(\dfrac{\mathrm{CR}''_{\mathrm{Data}}(t)}{\mathrm{CR}'_{\mathrm{Data}}(t)} + \nu \right)}{\nu f \left(I_0 + S_0\right) - \mathrm{CR}'_{\mathrm{Data}}(t) - \nu \left(\mathrm{CR}_{\mathrm{Data}}(t) - \mathrm{CR}_{\mathrm{Data}}(t_0)\right)}. \tag{7}$$

[29] **J. Demongeot, Q. Griette and P. Magal** SI epidemic model applied to COVID-19 data in mainland China, *Royal Society Open Science* 7:201878 (2020).

Estimated initial number of infected

By combining (6) and the Bernoulli-Verhulst equation (5) for $t \to \mathrm{CR}(t)$, we deduce the initial number of infected

$$I_0 = \frac{\mathrm{CR}'_{\mathrm{Data}}(t_0)}{\nu\, f} = \frac{\chi_2\, \mathrm{CR}_{\mathrm{Data}}(t_0)\left(1 - \left(\dfrac{\mathrm{CR}_{\mathrm{Data}}(t_0)}{\mathrm{CR}_{\mathrm{Data}}(\infty)}\right)^{\theta}\right)}{\nu\, f}.$$

Remark 4.1

We fix $f = 0.5$, from the COVID-19 data in mainland China and formula (6) (with $\mathrm{CR}_{\mathrm{Data}}(t_0) = 198$), we obtain

$$I_0 = 1909 \text{ for } \nu = 0.1,$$

and

$$I_0 = 954 \text{ for } \nu = 0.2.$$

Compatibility condition with the data

Since $\mathrm{CR}(t) < \mathrm{CR}_{\mathrm{Data}}(\infty)$, by considering the sign of the numerator and the denominator of (8), we obtain the following proposition.

Proposition 4.2

The rate of transmission $\tau(t)$ given by (8) is non negative for all $t \geq t_0$ if

$$\nu \geq \chi_2 \, \theta, \tag{8}$$

and

$$f\,(I_0 + S_0) + \nu \mathrm{CR}_{\mathrm{Data}}(t_0) > \mathrm{CR}_{\mathrm{Data}}(\infty)\,(\chi_2 + \nu)\,. \tag{9}$$

Compatibility of the model SI with the COVID-19 data for mainland China

The model SI is compatible with the data only when $\tau(t)$ stays positive for all $t \geq t_0$. From our estimation of the Chinese's COVID-19 data we obtain $\chi_2\,\theta = 0.14$. Therefore from (8) we deduce that model is compatible with the data only when

$$1/\nu \leq 1/0.14 = 3.3 \text{ days.}$$

This means that the average duration of infectious period $1/\nu$ must be shorter than 3.3 days.

Similarly the condition (9) implies and since we have $\mathrm{CR}_{\mathrm{Data}}(t_0) = 198$ and $\mathrm{CR}_{\mathrm{Data}}(\infty) = 67102$, we obtain

$$f \geq \frac{67102 \times 0.66 + (67102 - 198) \times 0.14}{1.4 \times 10^9} \geq 3.83 \times 10^{-5}.$$

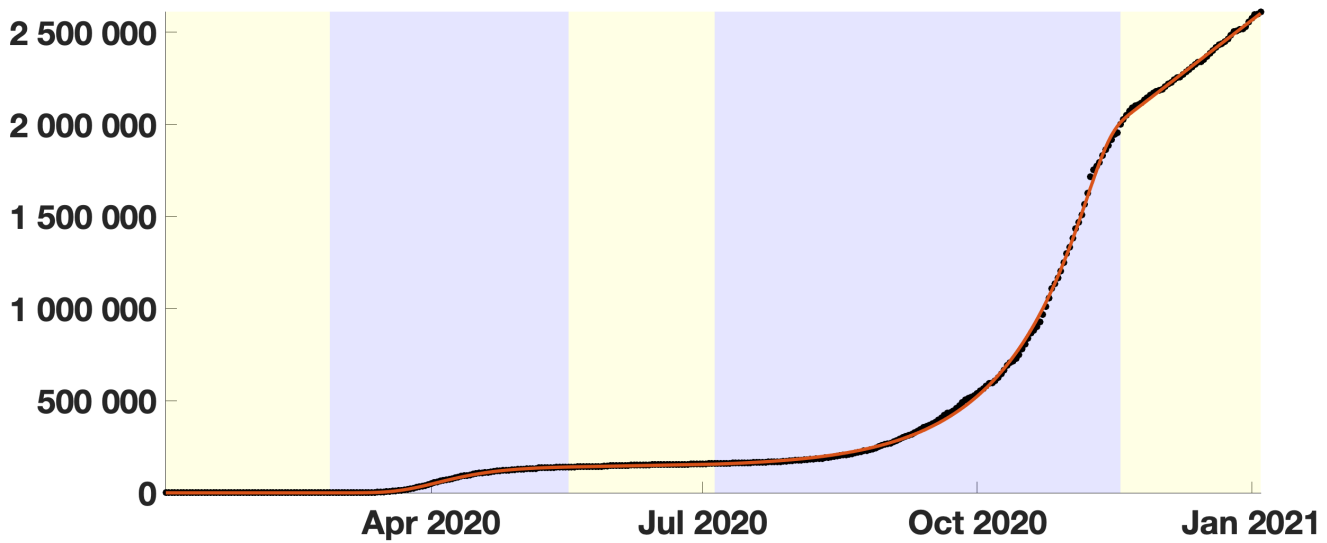
So according to this estimation the fraction of unreported $0 < f \leq 1$ can be almost as small as we want.

Outline

Epidemic and Endemic phases in France

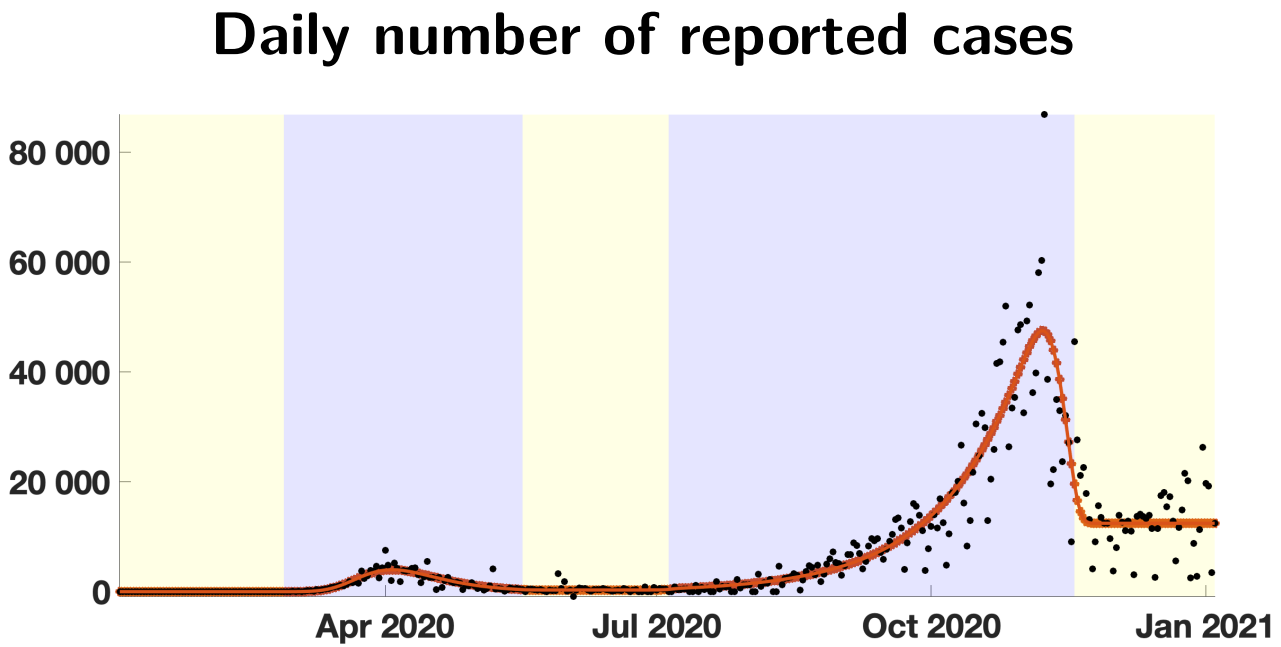
We fit a Bernoulli-Verhulst model during each epidemic phase. Then we extend the model by lines outside the epidemic phases. We regularize the junction points by a convolution with a Gaussian function with standard deviation of 7 days.

Cumulative number of reported cases



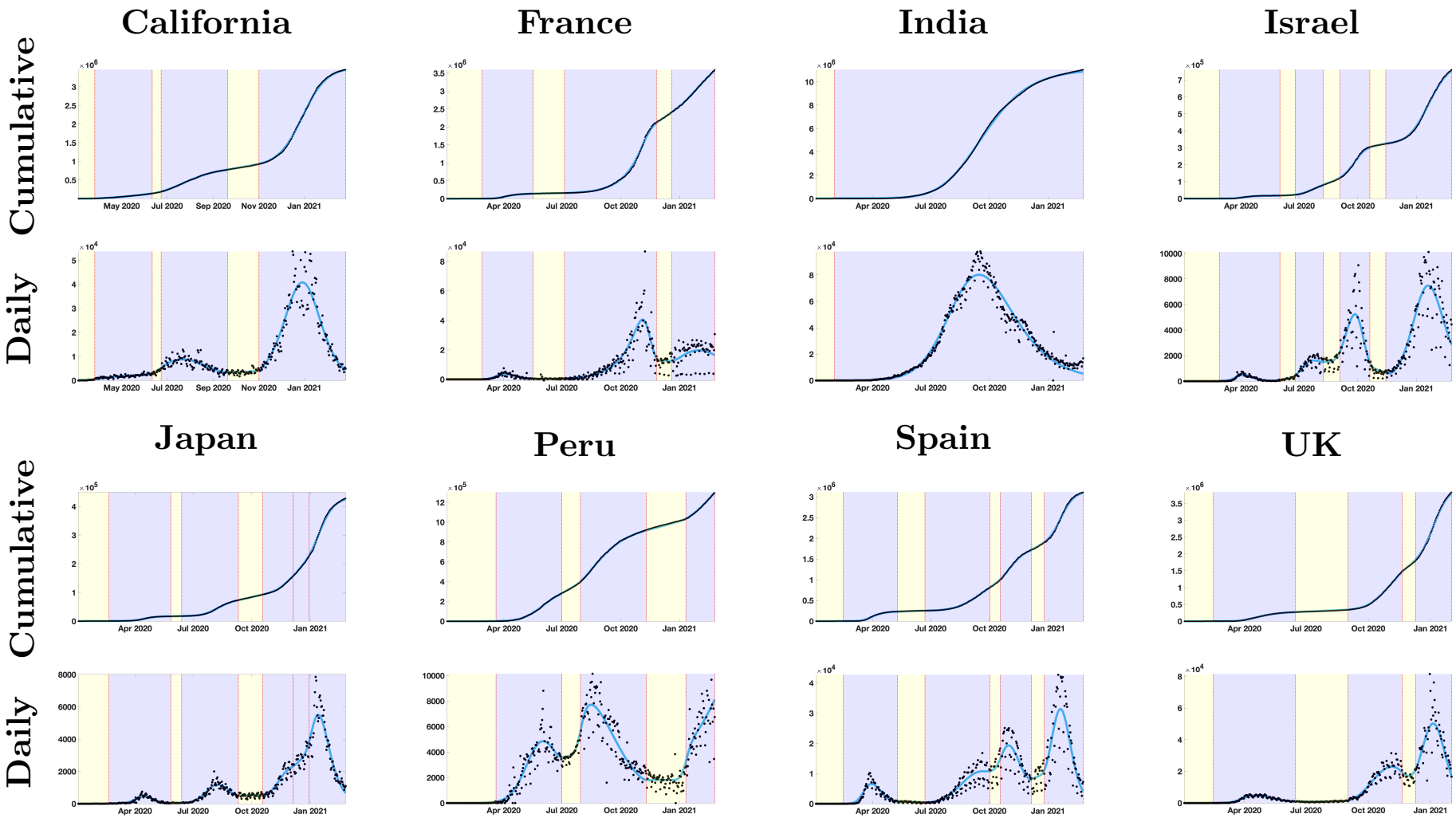
The red curve corresponds to the phenomenological model and the black dots correspond to the data of the number of cumulative cases. **We use** $16 = 2 \times 5 + 3 \times 2$ **parameters for more than 365 points.**

Epidemic and Endemic phases in France



The red curve corresponds to the first derivative of the phenomenological model and the black dots correspond to the data of the daily number of cases.

Phenomenological Model[30]



[30] **Q. Griette, J. Demongeot and P. Magal** (2021), What can we learn from COVID-19 data by using epidemic models with unidentified infectious cases? *Mathematical Biosciences and Engineering*, **19(1)**: 537–594.

Instantaneous reproduction number

We use our method to compute the transmission rate, and we consider the **instantaneous reproduction number**

$$\mathbf{R_e(t)} = \tau(\mathbf{t})\mathbf{S(t)}/\nu,$$

and the **quasi-instantaneous reproduction number**

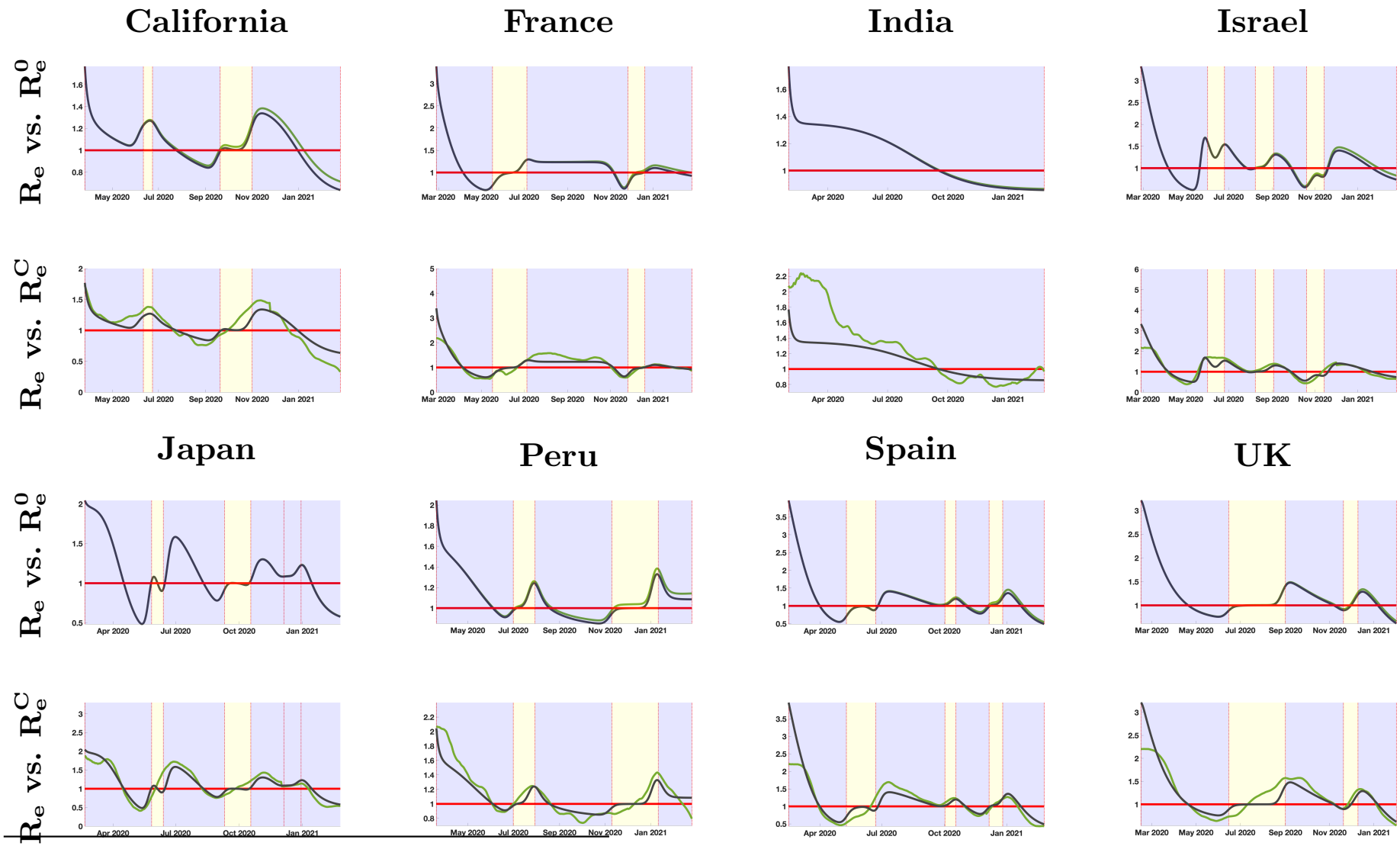
$$\mathbf{R_e^0(t)} = \tau(\mathbf{t})\mathbf{S_0}/\nu,$$

We compare the above indicators with $\mathbf{R_e^C(t)}$ the classical notion of **instantaneous reproduction number**[31,32].

[31] **T. Obadia, R. Haneef, & P. Y. Boëlle** (2012), The R_0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC medical informatics and decision making*, **12(1)**, 1-9.

[32] **A. Cori, N. M. Ferguson, C. Fraser, & S. Cauchemez** (2013), A new framework and software to estimate time-varying reproduction numbers during epidemics. *American journal of epidemiology*, **178(9)**, 1505-1512.

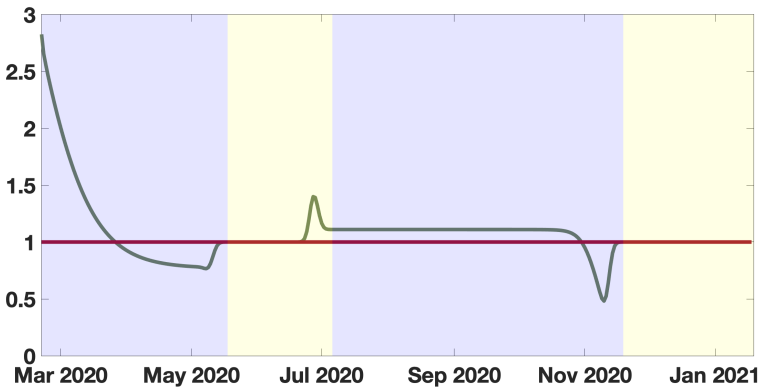
Instantaneous reproduction numbers[33]



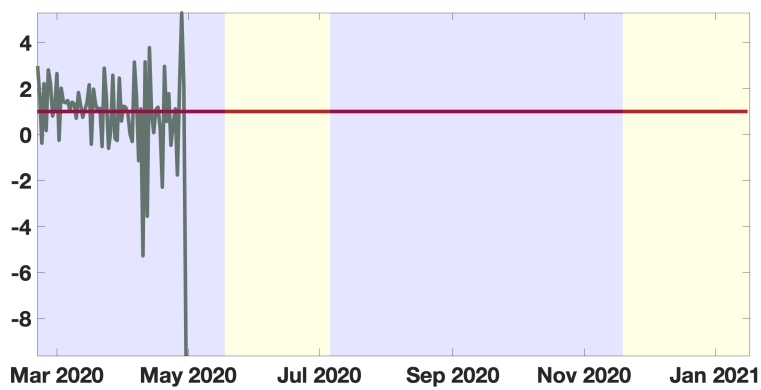
[33] **Q. Griette, J. Demongeot and P. Magal** (2021), What can we learn from COVID-19 data by using epidemic models with unidentified infectious cases? *Mathematical Biosciences and Engineering*, **19(1)**: 537–594.

Why do we need a phenomenological model to regularize the data?

With phenomenological model



Without phenomenological model



Conclusions

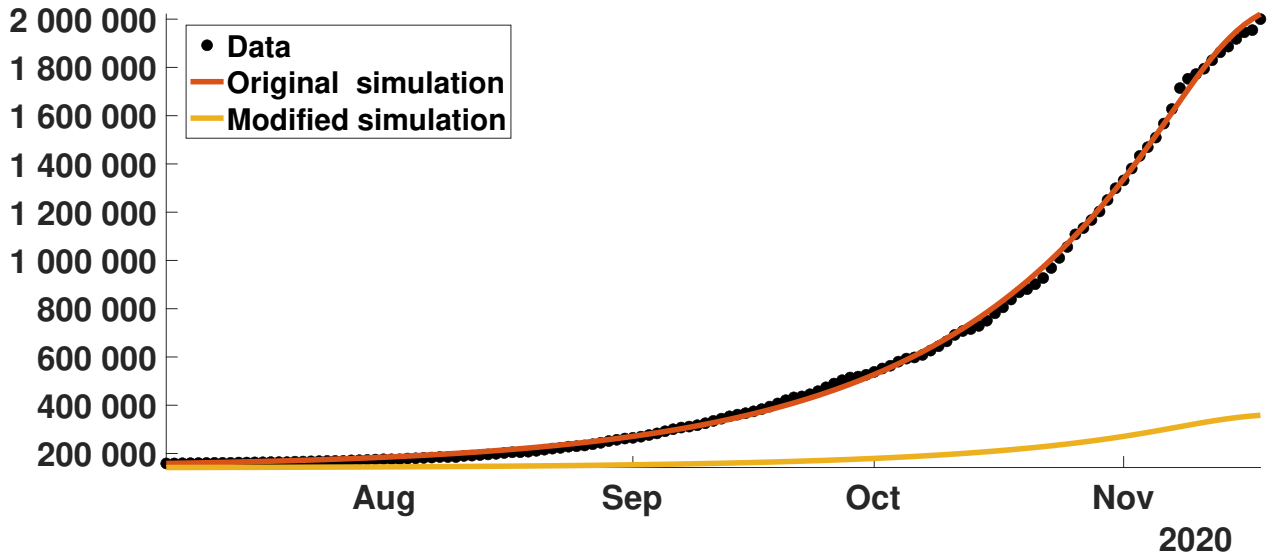
The population of susceptible patients is almost unchanged after the epidemic passed. Therefore, the system behaves almost like the non-autonomous system

$$I'(t) = \tau(t)S_0 I(t) - \nu I(t), \forall t \geq t_0, \text{ and } I(t_0) = I_0,$$

This means that $I(t)$ depends linearly on I_0.

Conclusions

The average daily number of cases during the endemic phases matters a lot.[34]



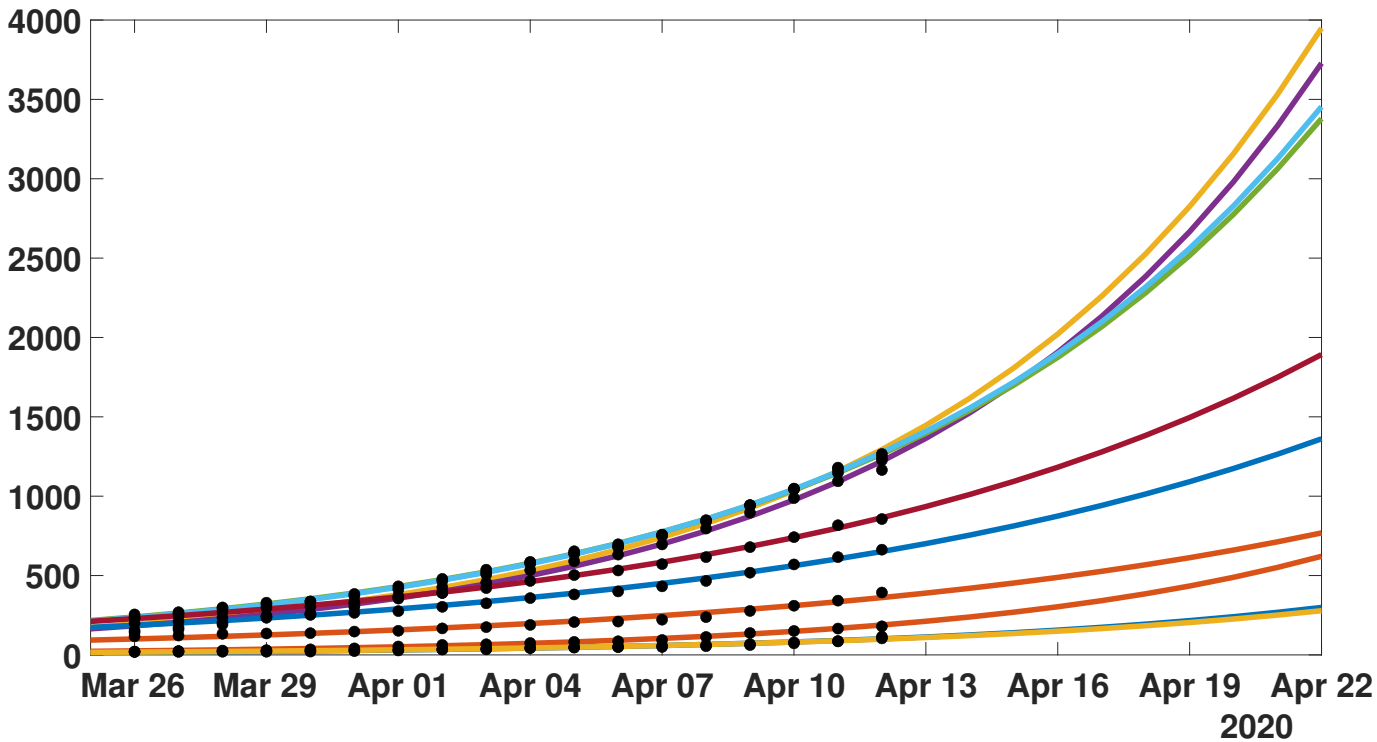
We start the simulation at time $t_0 =$ July 05 with the initial value $I_0 = \dfrac{\mathrm{CR}'(t_2)}{\nu f}$ for **red curve** and with $I_0 = \dfrac{1}{10}\dfrac{\mathrm{CR}'(t_2)}{\nu f}$ for **yellow curve**.

[34] **Q. Griette, J. Demongeot and P. Magal** (2021) A robust phenomenological approach to investigate COVID-19 data for France, *Mathematics in Applied Sciences and Engineering*, **2(3)**, 149-218.

Outline

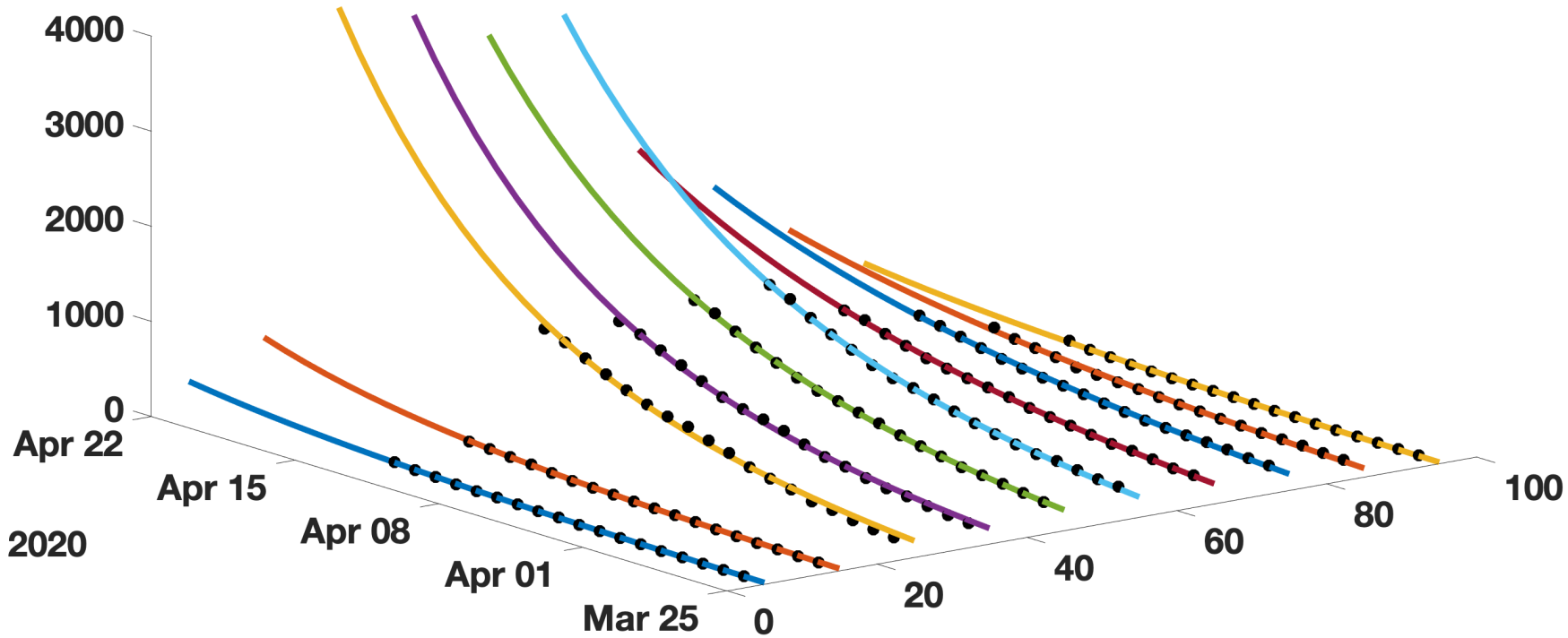
Cumulative reported cases with age structure in Japan

In the figure below we use an exponential fit for age group data for Japan[35].



The exponential growth depend on the age groups.

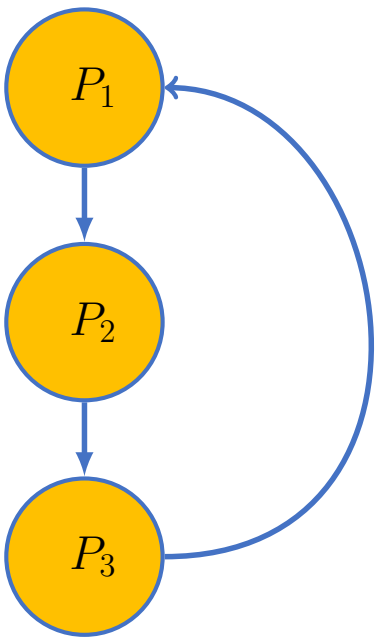
[35] **Q. Griette, P. Magal and O. Seydi** (2020), Unreported cases for Age Dependent COVID-19 Outbreak in Japan, *Biology* **9, 132**.



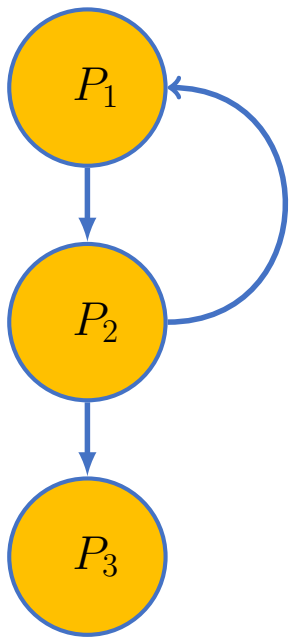
The growth rates being different, there is a weak coupling between the age groups.

Graphs of cooperative systems[36]

Non-primitive irreducible matrix



Non-irreducible matrix



[36] **A. Ducrot , Q. Griette , Z. Liu , P. Magal** (2022), *Differential Equations and Population Dynamics I: Introductory Approaches*, Springer.

Epidemic model with unreported cases[37]

The consider for example the case of a single age group, we obtain the following model which was first considered for COVID-19

$$\begin{cases} I'(t) = \tau S_0 \left(I(t) + U(t)\right) - \nu I(t), \\ U'(t) = \nu \left(1 - f\right) I(t) - \eta U(t), \end{cases}$$

for $t \geq t_0$, and with initial distribution

$$I(t_0) = I_0 \text{ and } U(t_0) = U_0. \tag{10}$$

[37] **Z. Liu, P. Magal, O. Seydi, and G. Webb** (2020), Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9(3)**, 50.

We can reformulate this system using a matrix formulation

$$\begin{pmatrix} I'(t) \\ U'(t) \end{pmatrix} = A \begin{pmatrix} I(t) \\ U(t) \end{pmatrix}, \forall t \in [t_0, t_1],$$

where

$$A = \begin{pmatrix} \tau S_0 - \nu & \tau S_0 \\ \nu (1-f) & -\eta \end{pmatrix}.$$

Then the matrix A is **irreducible** if and only if

$$\nu (1-f) > 0 \text{ and } \tau S_0 > 0.$$

Remember that the model to connect the data and the epidemic model is the following

$$\mathrm{CR}'(t) = f\,\nu\,I(t), \text{ for } t \geq t_0. \tag{11}$$

Consider the **exponential phase of the epidemic**. That is,

$$\mathrm{CR}'(t) = \chi_1\chi_2 e^{\chi_2 t}, \forall t \in [0, \tau],$$

for some $\tau > 0$. Combining the two previous equations, we obtain

$$f\,\nu\,I(t) = \chi_1 e^{\chi_2 t}$$

- Remember that χ_1 and χ_2 are computed by using the data.
- More preciserly, these parameters are obtained by fitting $t \to \chi_1 e^{\chi_2 t} - \chi_3$ to the **cumulative number of cases data** during a period of time $[0, \tau]$.

We can rewrite $f\,\nu\,I(t) = \chi_1 e^{\chi_2 t}$ by using a inner product

$$\left\langle y_0, \begin{pmatrix} I(t) \\ U(t) \end{pmatrix} \right\rangle = \chi_1 e^{\chi_2 t}, \text{ with } y_0 = \begin{pmatrix} \nu\, f \\ 0 \end{pmatrix},$$

where $\langle ., . \rangle$ is the **Euclidean inner product** is define in dimension 2

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2.$$

One may generalized the **Euclidean inner product** to the dimension n as follows

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Theorem 6.1

Let $\chi_1 > 0$, $\chi_2 > 0$, and $\tau > 0$. Let A be a n by n real matrix. Assume that the off-diagonal elements of A are non-negative, and A is irreducible. Assume that there exist two vectors $y_0 > 0$, and $x_0 > 0$ such that

(Linear model) $$\dot{x}(t) = Ax(t), \text{ and } x(0) = x_0,$$

satisfies

(Connection with the data) $$\langle y_0, x(t) \rangle = \chi_1 e^{\chi_2 t}, \forall t \in [0, \tau].$$

*Then χ_2 must be the **dominant eigenvalue** A (i.e., with the largest real part). Moreover, we can choose a vector $x_0 \gg 0$ (i.e., with all its components strictly positive), satisfying*

$$Ax_0 = \chi_2 x_0.$$

Multiplying x_0 by a suitable positive constant we obtain $\langle y_0, x_0 \rangle = \chi_1$. Under this extra condition, we will have

$$\langle y_0, x(t) \rangle = \chi_1 e^{\chi_2 t}, \forall t \in [0, \tau].$$

Returning back to the example of an **epidemic model with unreported cases**, we must find $I(0) > 0$ and $U(0) > 0$ such that

$$\begin{pmatrix} \tau S_0 - \nu & \tau S_0 \\ \nu(1-f) & -\eta \end{pmatrix} \begin{pmatrix} I(0) \\ U(0) \end{pmatrix} = \chi_2 \begin{pmatrix} I(0) \\ U(0) \end{pmatrix}.$$

After a few computations (see the supplementary in Liu et al. [38]), we obtain

$$\tau = \frac{\chi_2 + \nu}{S_0} \frac{\eta + \chi_2}{\nu(1-f) + \eta + \chi_2},$$

and

$$U_0 = \frac{\nu(1-f)}{\eta + \chi_2} I_0 = \frac{(1-f)\nu}{\eta + \chi_2} I_0.$$

[38] **Z. Liu, P. Magal, O. Seydi, and G. Webb** (2020), Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9(3)**, 50.

Let $\chi_1 > 0$, $\chi_2 > 0$, $\phi_1 > 0$, $\phi_2 > 0$, and $\tau > 0$. Assume that $x_0 > 0$, $y_0 > 0$ and $z_0 > 0$ satisfy

$$\dot{x}(t) = Ax(t), \text{ and } x(0) = x_0,$$

and

$$\langle y_0, x(t) \rangle = \chi_1 e^{\chi_2 t}, \forall t \in [0, \tau],$$

$$\langle z_0, x(t) \rangle = \phi_1 e^{\phi_2 t}, \forall t \in [0, \tau].$$

If $\chi_2 \neq \phi_2$ the matrix A must be **reducible**. That is, up to a re-indexation of the components of $x(t)$, the matrix A reads as

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{ij} are block matrices. The matrix A presents a **weak coupling** between the last block's components and the first block's components[39].

[39] **P. Magal, O. Seydi, and G. Webb** (2018), Final size of a multi-group SIR epidemic model: Irreducible and non-irreducible modes of transmission, *Mathematical Biosciences* **301**, 59-67.

Thank you for listening