POLITECNICO DI TORINO
DISMA Dipartimento di Scienze Matematiche G. L. Lagrange
ECCELLENZA 2018 • 2022

The SIR Model for Spread of Disease and its analysis for COVID-19

Candidate:
Shannon Mc Mahon

Supervisor:
Prof. L. Preziosi

Bachelor's degree course in Mathematics for Engineering
September 21, 2020

Part I

The SIR Model for spread of disease

The SIR Model

The SIR compartmental model divides the population into three categories:

1. Susceptibles **(S)**: individuals that can contract the disease, and are therefore susceptible to it;
2. Infectious **(I)**: individuals that have contracted the disease, and that can hence infect susceptibles that will themselves become infected;
3. Removed **(R)**: individuals that have contracted the disease and have either recovered or died from it.

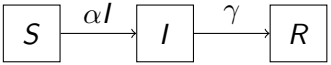
Differential Equations of the model

The dynamics of the model can be represented by:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\alpha S(t)I(t) \quad , \quad S(0) = S_0 \\ \frac{dI(t)}{dt} &= \alpha S(t)I(t) - \gamma I(t) \quad , \quad I(0) = I_0 \\ \frac{dR(t)}{dt} &= \gamma I(t) \quad , \quad R(0) = 0 \end{aligned} \tag{1}$$

where:

- γ rate of *removal*.
- β: the average number of people contacted by an infected individual, per unit of time;
- $\alpha = \frac{\beta}{N}$ rate of *infection*.



What can be deduced?

By dividing the second equation of (1) by the first, one finds

$$\frac{dI}{dS} = -\left(1 - \frac{\gamma}{\alpha S}\right) \tag{2}$$

It follows that

$$I_{\max} = -\frac{\gamma}{\alpha} + \frac{\gamma}{\alpha}\ln\left(\frac{\gamma}{\alpha}\right) + I_0 + S_0 - \frac{\gamma}{\alpha}\ln\left(S_0\right) \tag{3}$$

Moreover, it is also possible to find an expression for S by dividing the first equation by the third and integrating:

$$S = S_0 e^{\frac{\alpha}{\gamma}(R - R(0))} \tag{4}$$

The Basic Reproduction Number R_0

The *Basic reproduction number* R_0 represents the average number of people that an infected individual will infect (during his infectivity period) assuming that all other people in the population are susceptible.
We can express it as:

$$R_0 = \frac{\beta}{\gamma} = N\frac{\alpha}{\gamma} \tag{5}$$

- if $R_0 > 1$ then the number of infected cases will grow exponentially,
- if $R_0 < 1$ then the disease dies out.

Model Visualisation

The figure below shows the behaviour of S, I and R as time goes by, and focuses specifically on what happens as the value of R_0 increases.

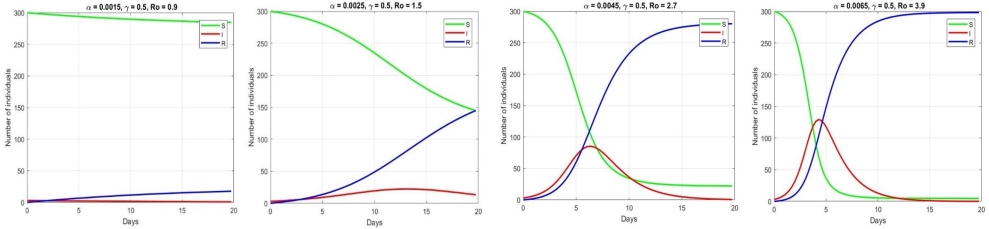


Figure: SIR Model for various values of R_0.

As seen in (5) R_0 depends on the infection and removal rates, and in an epidemic we want to lower the value of this variable.

"Flattening the Curve"

If the number of infected individuals becomes too high, the healthcare system is overwhelmed. We need to flatten the curve of $I(t)$ and lower the value of R_0.

As shown in the figure in the previous slide:

- **Recovery rate** γ: assuming the disease is novel ad that no vaccine is available, the recovery rate is a parameter that **cannot vary**.
- **Infection rate** α: it is possible to **decrease** the infection rate eg. by introducing social distancing and/or lockdown measures, and with it decrease R_0.

Part II

Using Real Data in a SIR Model

Example I: Influenza Epidemic in an English Boarding School

The data in the table below is from an influenza epidemic in an English boarding school in 1978 with N=763 students and average recovery period of 2 days (as stated in the British Medical Journal).

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
I	1	3	8	28	75	221	291	255	235	190	125	70	28	12	5

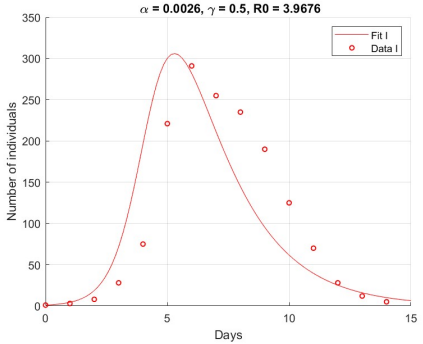
Table: For each day the number of active infected cases is reported.

Day 1: 1 infected $\Longrightarrow$ $S(0)$=762, $I(0)$=1 and $R(0)$=0.
Day 2: 3 infected $\Longrightarrow$ $\alpha = \frac{2}{762} = 0.0026$.

Removal rate: $\gamma = \frac{1}{2} = 0.5$.

Plotting the resulting SIR Model

The figure below shows the obtained fit of for the infected data.



Although the fitting is appreciable we need a well defined method for determining the values of α and γ that give the best possible fit.

Solving Optimisation Problems with the *fmincon* Algorithm

To solve our optimisation problem we make use of MATLAB's *fmincon* algorythm, which is gradient based. It has the following structure,

$$x = fmincon(fun, x0, [], [], [], [], lb, ub, [])$$

where:

- **fun** is the function to be minimized;
- **x0** is the initial guess vector that contains the starting values of α and γ for the algorithm;
- **x** is the solution vector that minimizes *fun*;
- **lb** and **up** are lower and upper bound vectors such that $lb \leq x \leq ub$.

Choosing the *fun* function to minimize

In order to obtain *fun* we calculated the 2-norm of the residual, squared, for each category of the SIR model and sum up the quantities:

$$fun = norm(S - S_{\text{exp}}, 2)^2 + norm(I - I_{\text{exp}}, 2)^2 + norm(R - R_{\text{exp}}, 2)^2 \qquad (6)$$

where

- the subscript *"exp"* stands for experimental, to indicate that we are referring to the actual input data.
- S, I and R contain the prediction obtained by means of the SIR Model.

The value that the function *fun* acquires at the optimal solution x will be referred to as **Error**.

Illustrative Example of the use of the Optimisation Algorithm

Consider the following data of an outbreak in a population of N=500.

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
I	3	6	10	20	38	38	80	100	97	80	61	44	31	26	21	16
R	0	0	5	12	27	52	84	137	190	254	300	340	365	397	402	420

Table: Data of an epidemic.

- **Step 1**: we make an initial guess for x by defining x0=[α γ]=[0.0015 0.35] which lead to Error=2.1511e+07;
- **Step 2**: we run the algorithm previously defined and find that x=[0.0025 0.59] which lead to Error decreasing to 1.3785e+07.

Plot with x0 *vs* Plot with Optimal solution x

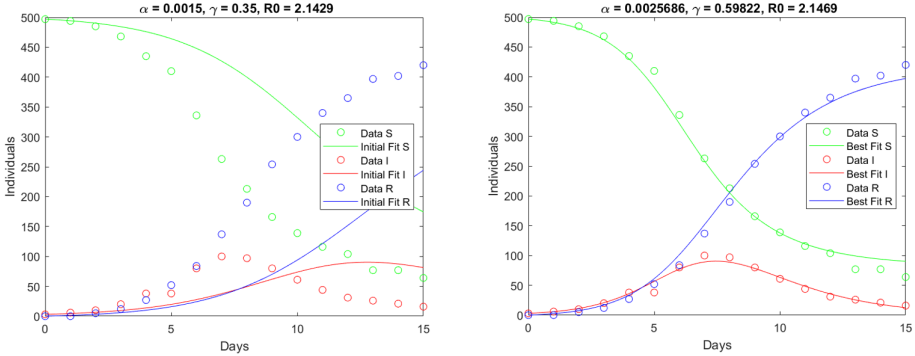


Figure: Data fitting with the initial estimate x0 (left) and optimal value x (right).

SIR applied to COVID-19

Next we applied the algorithm to a recent infectious disease: COVID-19.

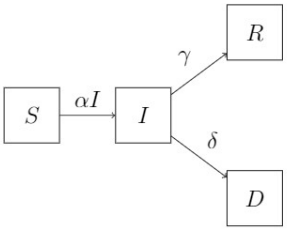
- Coronavirus disease 2019 is a disease caused by the novel SARS-CoV-2 virus. The first reported case of this disease was on December 31st 2019, in Wuhan, China.
- The outbreak was declared a public health emergency by WHO (World Health Organization) on January 30th 2020, and subsequently on March 11th a pandemic.
- Using the available online data we will model the COVID-19 outbreak in Italy in the time-frame between 20th February and 31st of July.
  - Note that on March 9th lockdown was imposed.

Defining the COVID-19 Model

We applied both the SIR and the SIRD variation of the model. The following slides focus on the SIRD model and results applied to the COVID-19 outbreak. SIR results have been included in the Appendix.

In SIRD we have that:

- R is the *Recovered* compartment;
- γ is the *recovery rate*;
- D is the *Dead* compartment
- δ is the *mortality rate*.



COVID-19 Parameters and lessons learnt

In the first set of results from this study (SIR) the parameters were kept constant and yielded a poor fit.

For this reason we chose to consider the key aspects of the COVID-19 outbreak in Italy and make the parameters for SIRD **time-dependant**.

However since there is no cure or vaccine available, γ will remain a constant.

Defining the Parameters for COVID-19

- Once draconian measures are taken, the transmission rate of the disease decreases, so we define:

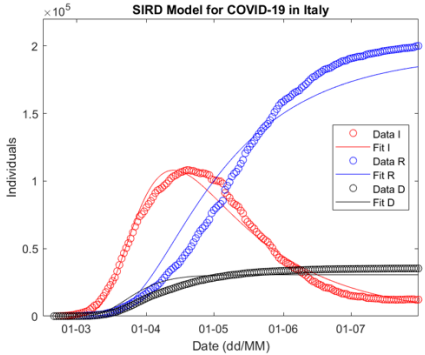
$$\alpha(t) = \begin{cases} \alpha_0 & t \leq t_{lock} \\ \alpha_0 e^{\frac{-t+t_{lock}}{\tau\alpha}} & t > t_{lock} \end{cases} \tag{7}$$

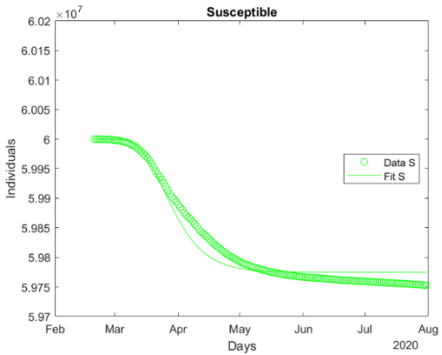
- At the beginning of an epidemic the more severe cases are detected, so the mortality rate is higher than in subsequent moments when less severe cases are added, hence:

$$\delta(t) = \begin{cases} \delta_0 & t \leq t_{lock} \\ \delta_0 e^{\frac{-t+t_{lock}}{\tau\delta}} & t > t_{lock} \end{cases} \tag{8}$$

Applying the algorithm to COVID-19 Italian Data

By applying the optimisation method on the Italian data-set the algorithm returns the optimal values for α_0, τ_α, γ, δ_0 and τ_δ.

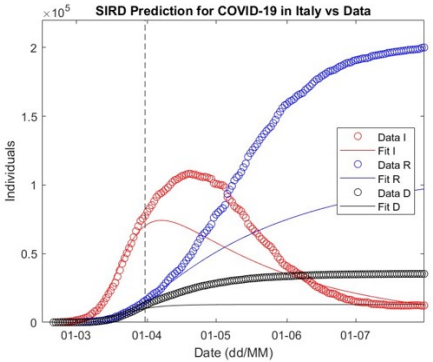


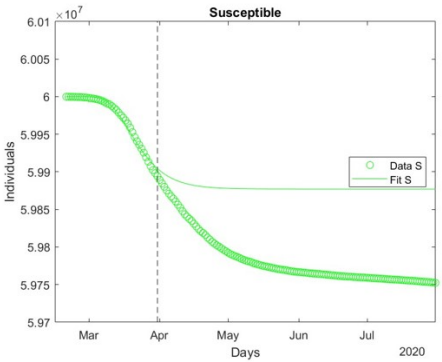


α_0	τ_α	δ_0	τ_δ	γ	$max(R_0)$	$min(R_0)$	*Error*
8.45e-9	13.13	0.10	10.68	0.024	3.97	0.97	4.08e+14

Predicting COVID-19

Let us imagine that it is March 30th, 2020. What happens if we use only the data up this point in time?





α_0	τ_α	δ_0	τ_δ	γ	$max(R_0)$	$min(R_0)$	*Error*
7.74e-9	10	0.04	10	0.01	7.34	0.98	6.83e+14

Tukey's Bisquare Function

- It is possible to improve the prediction by using Tukey's Bisquare function ρ, that for large errors tapers off. If u_i is the i-th residual we have:

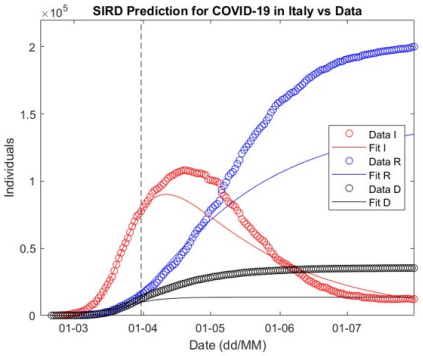
$$\rho(u_i) = \begin{cases} \frac{u_i^2}{2} - \frac{u_i^4}{2c^2} + \frac{u_i^6}{6c^4} & |u_i| \leq c \\ \frac{c^2}{6} & |u_i| > c \end{cases} \tag{9}$$

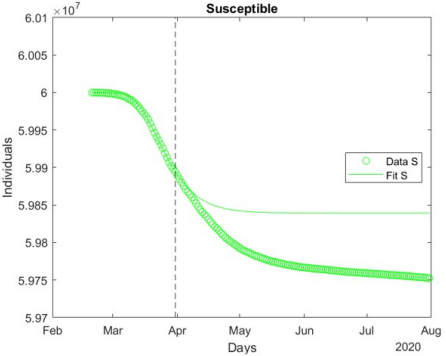
- Our new *fun* will be:

$$fun = \sum_i \rho(S(i) - S_{\exp}(i)) + \sum_i \rho(I(i) - I_{\exp}(i)) + \sum_i \rho(R(i) - R_{\exp}(i)) + \sum_i \rho(D(i) - D_{\exp}(i)) \tag{10}$$

Improved Prediction

The use of Tukey's function yields the following result:





α_0	τ_α	δ_0	τ_δ	γ	$max(R_0)$	$min(R_0)$	*Error*
7.72e-9	11.8	0.05	10	0.02	6.26	0.98	5.21e+14

Overview of Results

- The SIRD model can describe the evolution of the COVID-19 epidemic in Italy accurately when provided with the full data-set.
- SIRD model performance in terms of predictive capability was not robust
- Employing time-dependant parameters is key to obtaining a good fit

Limitations of the SIRD model

Our model is based on the assumptions that:

- Once an infected individual has recovered from the disease he/she is permanently immune to it and this is not true for COVID-19.
- The time between exposure to an infection and the appearance of symptoms is not taken into account.
- The recovery period is highly variable in COVID-19, yet it is a constant parameter in the model.
- No distinction is made between symptomatic and asymptomatic cases of infected individuals which leads to a bias in the data and model results.

To address some of the above limitations other compartmental models have been introduced eg.the SEIIRD model.

Limitations of numerical integration approach

- Focus on the average behaviour of a person, not allowing any kind of variability
- They are deterministic, so a specific initial state always leads to the same output.
- A continuous representation is used, which means that the starting point of the epidemic could be a fractional person, which is clearly not possible.

These limitations are what justifies **moving to a stochastic approach** where I) the number of infected individuals is no longer described by a continuous variable but by a positive integer, II) the transitions between compartments occur randomly.

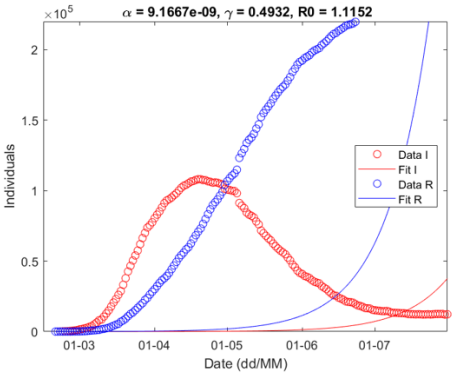
Potential next steps in modelling COVID-19

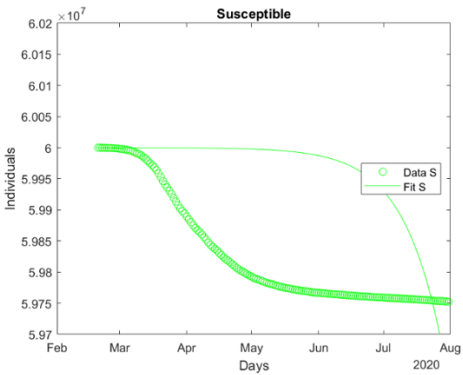
Based on the study and considerations we would recommend the following next steps:

- Explore the use of stochastic modelling approach to the same Italian data set to provide more robust predictive results to evaluate policies to avoid a second lockdown
- Apply the preferred modelling approach to other countries data to evaluate the pandemic at a global level

Appendix

SIR model results for COVID-19 with constant parameters





References I

- Anonymous, *Influenza in a boarding school*, British Medical Journal **1** (1978), 578.
- Milan Batista, *Estimation of the final size of the coronavirus epidemic by the sir model*, 02 2020.
- Sally Blower and D Bernoulli, *An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. 1766.*, Reviews in medical virology **14** (2004), no. 5, 275–288.
- Xavier Bardina, Marco Ferrante, and Carles Rovira, *A stochastic epidemic model of covid-19 disease*, arXiv preprint arXiv:2005.02859 (2020).
- Filippo Buoncompagni, *Modelli epidemici sir*, 05 2019.
- Diego Caccavo, *Chinese and italian covid-19 outbreaks can be correctly described by a modified sird model*, 03 2020.

References II

- Francesco Casella, *Can the covid-19 epidemic be managed on the basis of daily data?*, 03 2020.
- Kaustuv Chatterjee, Kaushik Chatterjee, Arun Kumar, and Subramanian Shankar, *Healthcare impact of covid-19 epidemic in india: A stochastic mathematical model*, Medical Journal Armed Forces India (2020).
- Yongli Cai, Yun Kang, and Weiming Wang, *A stochastic sirs epidemic model with nonlinear incidence rate*, Applied Mathematics and Computation **305** (2017), 221–240.
- Yi-Cheng Chen, Ping-En Lu, and Cheng-Shang Chang, *A time-dependent sir model for covid-19*, arXiv preprint arXiv:2003.00122 (2020).
- Coursera, *Science matters: Let's talk about covid-19*, `https://www.coursera.org/learn/covid-19#about`, 2020, [Online; accessed 3 September 2020].

References III

- Ben S Cooper, Richard J Pitman, W John Edmunds, and Nigel J Gay, *Delaying the international spread of pandemic influenza*, PLoS Med **3** (2006), no. 6, e212.
- D. J. Daley and J. Gani, *Epidemic modeling: An introduction*, NY: Cambridge University Press, 2005.
- Marco Ferrante, Elisabetta Ferraris, and Carles Rovira, *On a stochastic epidemic seihr model and its diffusion approximation*, Test **25** (2016), no. 3, 482–502.
- Geoffrey P Garnett and Roy M Anderson, *Sexually transmitted diseases and sexual behavior: insights from mathematical models*, Journal of Infectious Diseases **174** (1996), no. Supplement_2, S150–S161.
- Bruno Gonçalves, *Epidemic modeling 101: Or why your covid-19 exponential fits are wrong*, Medium (2020), [Online; accessed 3 September 2020].

References IV

- _____, *Epidemic modeling 102: All covid-19 models are wrong, but some are useful*, Medium (2020), [Online; accessed 5 September 2020].
- _____, *Epidemic modeling 103: Adding confidence intervals and stochastic effects to your covid-19 models*, Medium (2020), [Online; accessed 5 September 2020].
- BT Grenfell, *Chance and chaos in measles dynamics*, Journal of the Royal Statistical Society: Series B (Methodological) **54** (1992), no. 2, 383–398.
- W Hamer et al., *Epidemiology old and new.*, Epidemiology Old and New. (1928).
- Nicolas Hoertel, Martin Blachier, Carlos Blanco, Mark Olfson, Marc Massetti, Marina Sánchez Rico, Frédéric Limosin, and Henri Leleu, *A stochastic agent-based model of the sars-cov-2 epidemic in france*, Nature Medicine (2020), 1–5.

References V

- Herbert W Hethcote, *The mathematics of infectious diseases*, SIAM review **42** (2000), no. 4, 599–653.
- Shaobo He, Yuexi Peng, and Kehui Sun, *Seir modeling of the covid-19 and its dynamics*, Nonlinear Dynamics (2020), 1–14.
- Sha He, Sanyi Tang, and Libin Rong, *A discrete stochastic model of the covid-19 outbreak: Forecast and control*, Math. Biosci. Eng **17** (2020), 2792–2804.
- Matt J Keeling and Ken TD Eames, *Networks and epidemic models*, Journal of the Royal Society Interface **2** (2005), no. 4, 295–307.
- William O Kermack and Anderson G McKendrick, *Contributions to the mathematical theory of epidemics—i*, Bulletin of mathematical biology **53** (1991), no. 1-2, 33–55.

References VI

- Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, Fiona Sun, Mark Jit, James D Munday, et al., *Early dynamics of transmission and control of covid-19: a mathematical modelling study*, The lancet infectious diseases (2020).
- Kenji Karako, Peipei Song, Yu Chen, and Wei Tang, *Analysis of covid-19 infection spread in japan based on stochastic transition model*, Bioscience trends (2020).
- Hillary Leung, *Can you be re-infected after recovering from coronavirus? here's what we know about covid-19 immunity*, Time (2020), [Online; accessed 10 September 2020].
- Charles J Mode and Candace K Sleeman, *Stochastic processes in epidemiology: Hiv/aids, other infectious diseases, and computers*, World Scientific, 2000.

References VII

- Merriam-Webster, *Merriam-webster.com dictionary*, https://www.merriam-webster.com/dictionary/epidemiology, August 2020, [Online; accessed 21 August 2020].
- Hiroshi Nishiura, Tetsuro Kobayashi, Ayako Suzuki, Sung-Mok Jung, Katsuma Hayashi, Ryo Kinoshita, Yichi Yang, Baoyin Yuan, Andrei Akhmetzhanov, Natalie Linton, and Takeshi Miyama, *Estimation of the asymptomatic ratio of novel coronavirus infections (covid-19)*, International Journal of Infectious Diseases **94** (2020).
- Madan K Oli, Meenakshi Venkataraman, Paul A Klein, Lori D Wendland, and Mary B Brown, *Population dynamics of infectious diseases: a discrete time model*, Ecological Modelling **198** (2006), no. 1-2, 183–194.

References VIII

- Bastian Prasse, Massimo A Achterberg, Long Ma, and Piet Van Mieghem, *Network-based prediction of the 2019-ncov epidemic outbreak in the chinese province hubei*, arXiv preprint arXiv:2002.04482 (2020).
- Andrea Pugliese, *Cenni su teoria ed utilizzo di modelli matematici per le epidemie*.
- Pejman Rohani, David JD Earn, and Bryan T Grenfell, *Impact of immunisation on pertussis transmission in england and wales*, The Lancet **355** (2000), no. 9200, 285–286.
- Ronald Ross, *The prevention of malaria*, J. Murray, 1910.
- Yuliana Susanti, Hasih Pratiwi, Sri H., and Twenty Liana, *M estimation, s estimation, and mm estimation in robust regression*, International Journal of Pure and Applied Mathematics **91** (2014).

References IX

- Henry C Tuckwell and Ruth J Williams, *Some properties of a simple stochastic epidemic model of sir type*, Mathematical biosciences **208** (2007), no. 1, 76–97.
- Pauline van den Driessche and James Watmough, *A simple sis epidemic model with a backward bifurcation*, Journal of Mathematical Biology **40** (2000), no. 6, 525–540.
- Cécile Viboud, Kaiyuan Sun, Robert Gaffey, Marco Ajelli, Laura Fumanelli, Stefano Merler, Qian Zhang, Gerardo Chowell, Lone Simonsen, Alessandro Vespignani, et al., *The rapidd ebola forecasting challenge: Synthesis and lessons learnt*, Epidemics **22** (2018), 13–21.
- Worldometer, *Italy coronavirus*, [Online; accessed 28 August 2020].
- Wikipedia, The Free Encyclopedia, *Mathematical modelling of infectious disease - history*, [Online; accessed 21 August 2020].

Thank you for the attention

POLITECNICO
DI TORINO