

How variance in parameters affect the reproduction number and why the SEIR model overshoots

Mathias L. Heltberg^{1,2,3}, Christian Michelsen^{1,3}, Kim Sneppen¹ and Troels Petersen^{1*}

¹ Niels Bohr Institute, University of Copenhagen,

Blegdamsvej 17, DK-2100,

² Laboratoire de Physique, Ecole Normale Supérieure,

Rue Lhomond, Paris 07505,

³ These authors contributed equally

(Dated: April 26, 2020)

The estimation of the spreading of COVID19 is a completely critical aspects of present science, politics and life. In this fundamental task, a standard procedure has become the implementation of the more or less detailed SIR models, that gives a very reasonable estimation of the predicted number of susceptible, infected and recovered individuals. However, surprisingly little is known about how variance in this fundamental model will affect the actual outcome of these numbers. In this paper, we will introduce an agent based network model, that will challenge the two fundamental assumptions of the SIR model - spatial homogeneity and rate homogeneity. We find that these significantly affect the number of diseased and thereby gives new estimates for the pressure on the health sector. We end by using this to construct an improved SIR model, taking these effects into account.

Keywords: Stochastic Networks, Epidemics, Numerical predictions.

Introduction

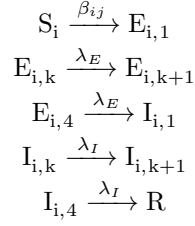
The dynamic of a single epidemic of a disease is often described in terms of the susceptible-infected-removed (SIR) model and its related variants. The main assumption of this model is that the population, in which a pathogenic agent is active, comprises three subgroups, described by the differential equations: $dS/dt = -\beta SI$, $dI/dt = \beta SI - \gamma I$ and $dR/dt = \gamma I$. In this model, individual members of the population susceptible to disease (S) become infected (I) and are subsequently removed from the pool spreading the disease due to either their death (D) or newly acquired immunity. The model examines only the temporal dynamics of the infection cycle and should thus be appropriate for the description of a well-localised epidemic outburst.

Model

We consider a network of N_0 nodes, each representing a household in a population. We will assume that each node, i , has a number of edges n_i , and that each link between node i and j has an assigned interaction strength $\beta_{ij} = \beta_{ji}$. For this network, the average degree is μ , and the network is constructed by generating $\mu \cdot N_0$ links.

*Electronic address: heltberg@nbi.ku.dk

From this we simulate an epidemic, starting from N_{init} initial points. In this agent based model, we assume that each node will make transition between infected states:



This is schematised in Figure 1B. The corresponding spread on a network, includes that an infected node, can only infect the subpopulation of nodes, which it interacts with. We will now compare this to the structure of a typical SEIR model. This model has been widely used to describe diseases and is at the time of this present paper being a core part of every countrys response to the COVID 19 outbreak. The SEIR model is deterministic and assumes a well mixed population. The corresponding model for a population of N_0 individuals take the form:

$$\dot{S} = -\beta_D SI \quad \text{with } \beta_D = \beta \frac{\mu}{2N_0} \quad (1)$$

$$\dot{E}_1 = \beta_D SI - \lambda_E E_1 \quad (2)$$

$$\dot{E}_i = \lambda_E E_{i-1} - \lambda_E E_i \quad (i = 2, 3, 4) \quad (3)$$

$$\dot{I}_1 = \lambda_E E_4 - \lambda_I I_1 \quad (4)$$

$$\dot{I}_i = \lambda_I I_{i-1} - \lambda_I I_i \quad (i = 2, 3, 4) \quad (5)$$

$$\dot{R} = \lambda_I I_4 \quad (6)$$

In the SEIR, some fundamental approximations has been made that are in some systems valid. These are that the system is well-mixed, meaning that all rates of infection between nodes are the same $\beta_{ij} = \beta_0$, and that every node on average has the same number of connections. Our first goal was to test the agreement between the agent based network model and the deterministic SEIR model. Especially we wanted to estimate the level of intrinsic noise that is not normally included, and the maximal number of infected. This specific number has been of crucial importance in the crisis of COVID-19 since it is the measure for whether the health care system of a nation will be over-exposed. Therefore in the

Effects of rates with variance and shifted node distribution

Our next goal was to include some effects that are present in the spreading of a disease, but which are not included in the SEIR model. First, we assumed that all rates are not equal, meaning that the strength of infection between two nodes have a significant spread. Physically this could be due to the wide spread in time that is spent with different connections and the normal distance to each individual. Furthermore there will be differences in viral load at the infected individual and the susceptibility at the receiving node. To account for all this we introduced a parameter, σ to measure the spread in variances. Therefore we now describe the infection rate as:

$$\beta_{ij} = \beta_0 + \sigma * \text{rand}_{[0,1]} \quad \sigma \in]0; 0.5[$$

This is visualized in Figure 2A and unless otherwise stated, we are using $\beta_0 = 1$.

To further understand the effects of variability in rates, we introduced a bias in the degree distribution between nodes. If all nodes are equally attractive, the resulting distribution is a poisson distribution, however to estimate the effects of people a lot of contacts, we introduce a bias in the probabilities to create contacts between two nodes. This is described by the parameter γ :

$$\rho_i = 0.5 + \gamma * \text{rand}_{[0,1]} \quad \gamma \in]0; 0.5[\quad (7)$$

Effects of distance and clusters

The final element we wanted to include into the model was the effects of space, such that there is a bias to interact with your nearest neighbours. To do this, we now constructed a network on a 2-dimensional surface, by using a random walker to measure the position of individuals:

$$X_i = X_{i-1} + \sqrt{2D_0}\eta_x \quad (8)$$

$$Y_i = Y_{i-1} + \sqrt{2D_0}\eta_y \quad (9)$$

Here D_0 is a characteristic length between individuals, and η is a normally distributed random number. Now we introduced the bias when we create the network, that two nodes can interact with a pseudo probability dependent on their distance between them:

$$p(r_{ij}) = e^{-\alpha \frac{r_{ij}}{D_0}} \quad \alpha \in [0; \infty] \quad (10)$$

Here it is clear that if $\alpha = 0$ there is no spatial effect and we obtain a similar scenario to the studied systems, and as alpha increases the system tends to include local interactions with a higher frequency.

In reality, individuals are not just spread around with approximately equal distance, but rather clustered together in cities of varying size. To account for this effect as well, we introduce another part to the generation of edges in the network, by including a time parameter:

$$\text{if } t_{i-1} < \tau \begin{cases} t_i = t_{i-1} + \Delta t \\ X_i = X_{i-1} + \sqrt{2D_0\Delta t}\eta_x \\ Y_i = Y_{i-1} + \sqrt{2D_0\Delta t}\eta_y \end{cases} \quad (11)$$

$$\text{else } \begin{cases} t_i = t_{i-1} + \Delta t \\ X_i = X_{i-1} + \sqrt{2D_0\Delta t}\eta_x \\ Y_i = Y_{i-1} + \sqrt{2D_0\Delta t}\eta_y \\ \tau = \tau + \left(\frac{1}{\text{rand}_{[0,1]}}\right)^{1/\psi} - 1 \end{cases} \quad (12)$$

Thus as $\psi \mapsto 0$

Combination of effects and how to account for this in the modelling

I. SUPPLEMENTARY MATERIAL

To go from one state to another, we assume that this occurs with a rate that is dependent on the time as a logistic function:

$$r(t) = \frac{A}{1 + \exp(-b(t - \tau))} \quad (13)$$

The time distribution for such an event to happen can be computed:

$$F(t) = \prod_i^t (1 - r(t)dt)r(t)dt \quad (14)$$

Using Volterras equation we obtain

$$F(t) = r(t)e^{-\int_0^t r(t')dt'} \quad (15)$$

Solving the integral gives:

$$-\int_0^t \frac{A}{1 + e^{-b(t'-\tau)}} dt' = \frac{A}{b} \int_{e^{b\tau}}^{e^{b(t-\tau)}} (1+u)^{-1} du = \frac{A}{b} \ln \left[\frac{1 + e^{-b(t-\tau)}}{1 + e^{b\tau}} \right] \quad (16)$$

Putting it all together gives:

$$\frac{A}{1 + \exp(-b(t - \tau))} e^{-\frac{A}{b} \ln \left[\frac{1 + e^{-b(t-\tau)}}{1 + e^{b\tau}} \right]} \quad (17)$$

$$= \mathcal{C} (1 + e^{\psi(t)})^{-\frac{A-b}{b}} e^{\psi(t)} \quad (18)$$

$$\text{where } \psi(t) = b(t - \tau) \quad \text{and} \quad \mathcal{C} = \frac{A}{(1 + e^{-b\tau})^{A/b}} \quad (19)$$

For this distribution we can calculate the mean value:

$$\int_0^x F(t)dt = \frac{1}{2} \quad (20)$$

$$\frac{\mathcal{C}}{b} \int_{1+e^{b\tau}}^{1+e^{-b(x-\tau)}} v^{-\frac{A+b}{b}} dv = \frac{1}{2} \quad \text{where } v = 1 + e^{-b(t-\tau)} \quad (21)$$

$$\mu = -\frac{\ln \left[\frac{A}{2\mathcal{C}} - (1 + e^{b\tau})^{\frac{b}{A+b}} - 1 \right]}{b} + \tau \quad (22)$$

We imagine three different submodels, all describing the distribution in number of connections by only one parameter μ :

$$P(n_i) = \begin{cases} e^{-\mu} \frac{\mu^{n_i}}{n_i!} & \text{Model 1} \\ \frac{1}{\mu} e^{-n_i/\mu} & \text{Model 2} \end{cases} \quad (23)$$

This network we simulate by placing all nodes on a surface, with a circular boundary. To further characterize this

network, we introduce a spatial dependence parameter α , so the relative probability for two nodes to interact is given by the function

$$\mathcal{P}(r_{ij}) = e^{-\alpha r_{ij}} \quad (24)$$

Now we include this network into a model with structure where a node can be susceptible, infected and recovered, as is previously described in the SIR model. To include the time-delay that is usually incorporated by the exposed

Effects of Variability in rates

Normally we consider the We consider N particles, that all can have a specific decay rate r_i , but that this follows a probability distribution $P(r)$. We will adopt the convention that;

$$\int_0^\infty r P(r) dr = r_\mu \quad (25)$$

This means that after a time τ the number of remaining particles will be:

$$N(\tau) = N_0 \int_{i=0}^\infty e^{-r\tau} P(r) dr \quad (26)$$

If we assume the well mixed system, we impose the condition that $P(r) = \delta(r - r_\mu)$. Therefore in the well mixed system, the the number of particles after τ will be:

$$N(\tau) = N_0 e^{-r_\mu \tau} \quad (27)$$

If we now consider $P(r) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r-r_\mu)^2}{2\sigma^2}}$ we obtain:

$$N(\tau) = N_0 \int_{i=0}^\infty e^{-r\tau} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r-r_\mu)^2}{2\sigma^2}} dr = N_0 e^{-\frac{1}{2}\sigma^2\tau^2 - r_\mu\tau} \quad (28)$$

Therefore the relative decrease in the observed decay rate, by having just a width in the distribution will be:

$$\frac{N_\sigma}{N_\delta} = e^{\frac{1}{2}\sigma^2\tau^2} \quad (29)$$