

# 1 Information Travel and Epidemics

## 1.1 Introduction

Information spreading changes the behaviour of individuals and thus a population during a pandemic significantly. In this analysis, we want to study the effect of information spreading on containing an epidemic. Our work is based on Funk et al. 2009. The paper's conclusion is that high quality information can change human behaviour. To replicate their results, we extend a basic SIR-Model with an awareness component. The model we created is based on our research questions:

1. "Which conditions can lead to outbreaks being contained by the information spread?"
2. "What is the impact of self-isolation of knowingly infected individuals on the spread of a disease and under which conditions is it most impactful?"

Based on the study of Funk et al, we assume that information quality decreases after a certain time and number of individuals passing information on ([grapevine effect](#)). We extended the original model with further aspects like the isolation rate. We would like to investigate how the infection changes after an individual is infected and isolates under what circumstances the isolation would be particularly successful.

**This is how our program is structured:** The core files are main.ipynb, awareness.ipynb, simulation.ipynb and plot.ipynb. The file rk4.ipynb includes the numerical approach to solve partial differential equations. The folder data stores the results of the executed computations. They provide the results used in the model's diagrams. The figures-folder holds the created figures while the model-folder contains custom made figures describing the working of the model.

The final code can be downloaded from the Github Repository at the following link: <https://github.com/Mela/Information-Travel-and-Pandemics>

To requirements for the setup are listed in the requirements.txt file. If the program is running in the deepnote environment these will be installed automatically. If the program is running in another one you have to run the "pip install" command.

## 1.2 Reading order

Our suggested reading order for the different jupyter notebooks is as follows: 1. documentation.ipynb for the introduction, explanation of the model and the parameters 2. main.ipynb for run the main functions 3. rk4.ipynb for the algorithm to solve partial differential equations (PDEs) and a comparison between Runge-Kutta of 4th order (rk4) and the euler forward step method 4. optional: example\_heat.ipynb for a simple example using rk4 and a model for the heat equation 5. awareness.ipynb is the file where SIR and awareness is modeled 6. analysis.ipynb is the file to run and evaluate the all of the simulations.

## 1.3 Our team

- Lisa Beckmann, Georg-August-University of Göttingen, Physics MSc.
- Melanie Eckenfels, Karlsruhe Institute of Technology, European Culture and History of Ideas MA.
- Nathanael Gutknecht, Karlsruhe Institute of Technology, Physics MSc.

- Zuzanna Mielczarek, University of Lübeck, Media Informatics MA.

## 1.4 The model

Note: Depending on the situation, either the term “information” or “awareness” is used, however, in our model they are used interchangeably.

We use a basic SIR-model to model the spread of the epidemic (see Fig. 1). The rate for susceptibility to infection is modified by the decay constant  $\rho$  and the parameter  $i$ , which describes the awareness, with  $(1 - \rho^i)$ . It is to note that  $i = 0$  corresponds to first hand information and therefore the highest awareness.

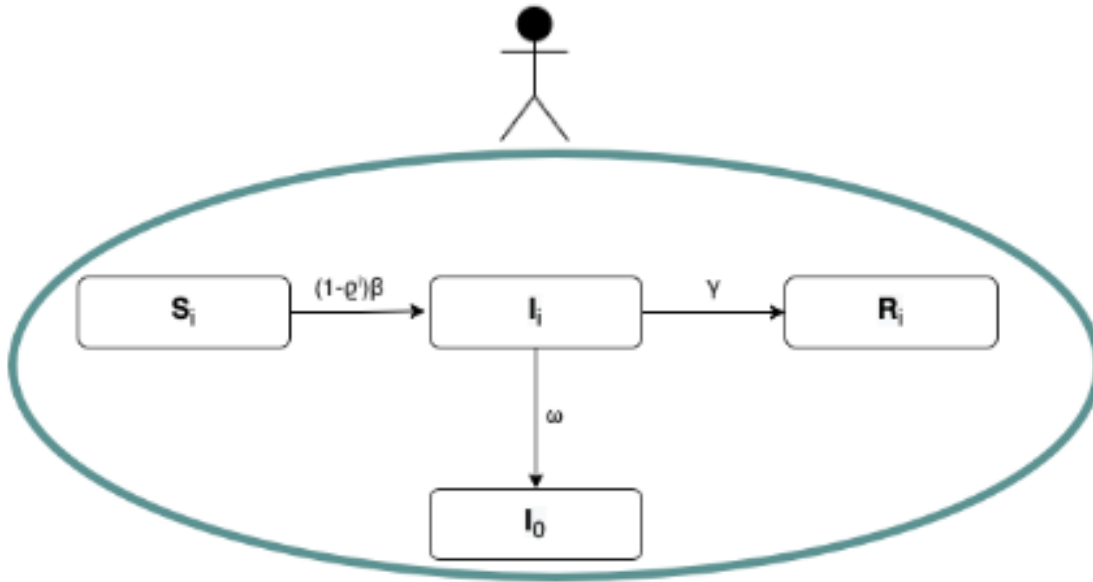


Fig. 1 Flowchart of the SIR model. The boxes describe individuals having one of three possible conditions: being susceptible, infected or recovered. The arrows show the transition rates

Information quality is characterized by the number of times it was passed on until arriving at individual  $X$  (see Fig. 2). Thus, the lower  $i$  is, the better the information quality and the less susceptible an individual is, Information is generated by infected individuals who can notice their own infection with a rate  $\omega$ . Information transmission can be achieved by interactions with the neighbors with a rate of  $\alpha$ , but the information quality wanes by  $i + 1$  for the receiving individual. Waning of information quality by time is modeled with a rate of  $\lambda$ .

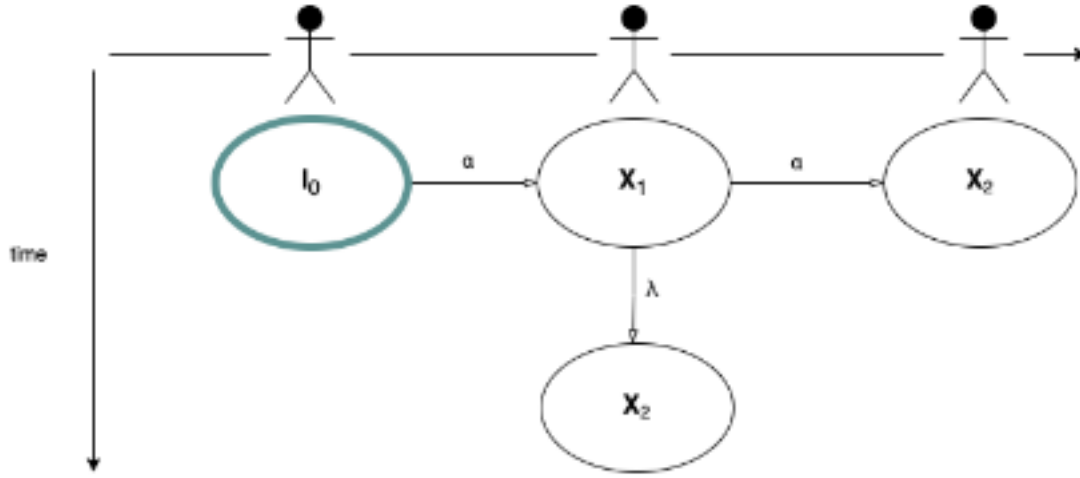


Fig. 2 Flowchart of the awareness generation. The boxes describe the individuals and the arrows describe the information transmission/fading with time. The index describes the grade of awareness.

An overview of the parameters we used:

Parameter	Description
$\beta$	normal infection rate
$\gamma$	recovery rate
$\rho$	decay constant
$\omega$	information generation rate
$\alpha$	information transmission rate
$\lambda$	information fading rate
$i$	awareness

To simplify the model and to match the individual-based analysis described by [Funk et al. 2009](#), we chose to model the population on a 2D-grid with each point having four neighbours.

Extension of the model In the original model as described above, only susceptible were acting on the information they were receiving. We extended the model with a self isolation rate of knowingly infected individuals of  $\kappa$ . This way, also the infected individuals were acting on information (although generated by themselves).

## 2 Results

Unless otherwise noted, these default parameters are assumed for all simulations:

### 2.0.1 Awareness with no New Infections

First, we had a look at how the awareness spreads and decays in our network when there are no new infections present, which could add more information/awareness to the system. For this, we



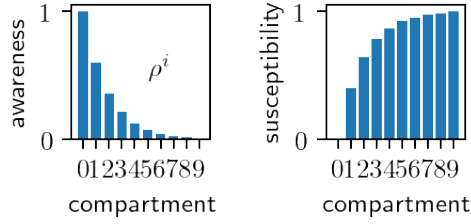


Fig. 4 Awareness and susceptibility for each awareness compartment.

## 2.0.2 Impact of Awareness on an Outbreak

**Outbreak without Awareness** As a baseline an outbreak was simulated without awareness in the population. The infections move outward in a circular shape until nearly everyone is in the recovered compartment in the end. The results can be seen in figure 5.

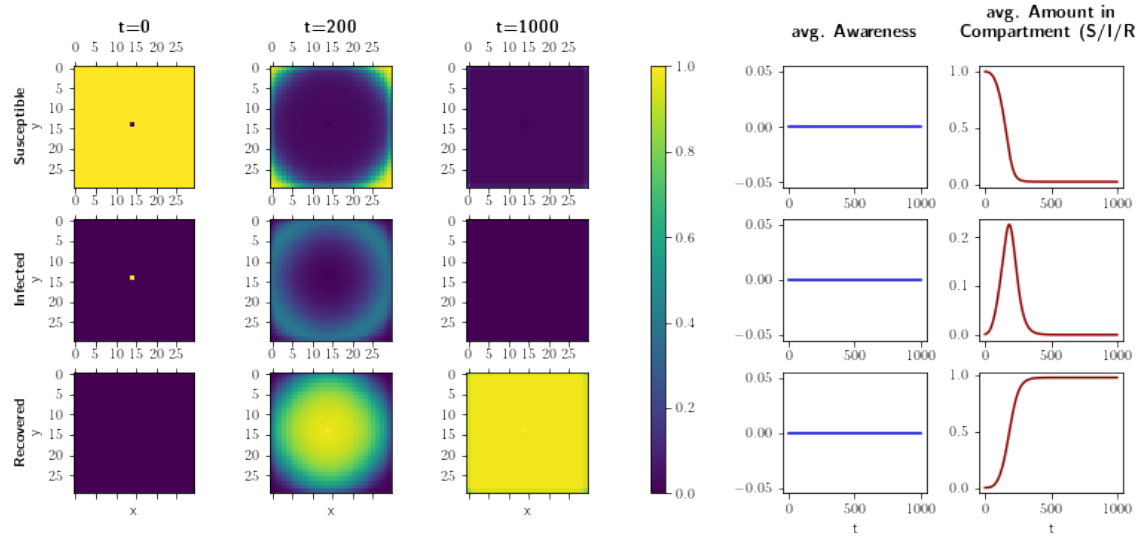


Fig. 5 Outbreak with no awareness. The outbreak is not contained and in the end the probability of having been infected at some point in time is close to one.

**Outbreak with Awareness** Next, the simulation was run with the same initial conditions, but this time with awareness being generated from infections, which spread throughout the network and reduce susceptibility. The results are displayed in figure 6.

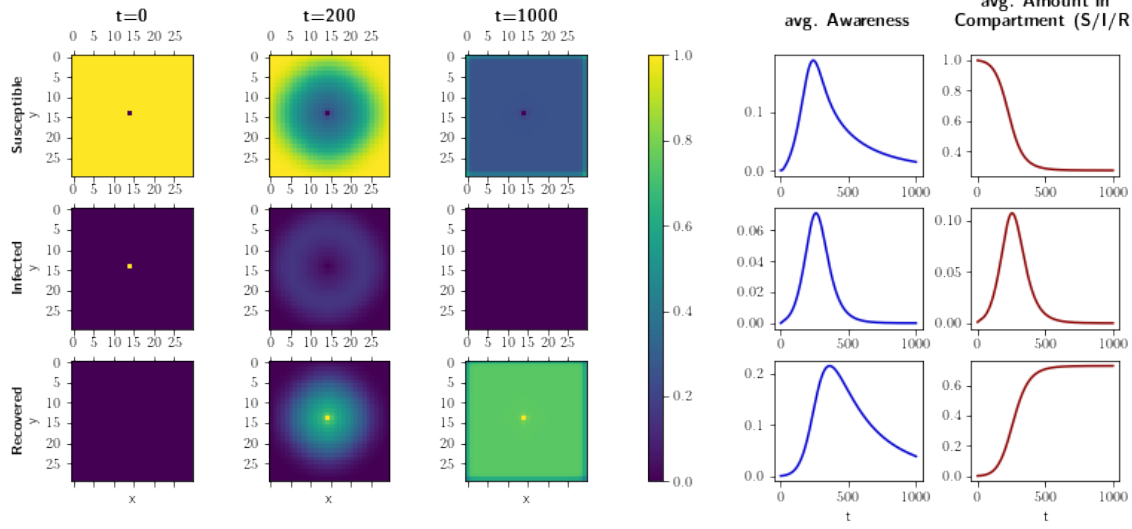


Fig. 6 Outbreak with awareness. The outbreak is reduced both in speed and in intensity.

**Comparison** When comparing the two scenarios one can notice the awareness having reduced the total number of infections as well as the maximum number of simultaneous infections as shown in figure 7.

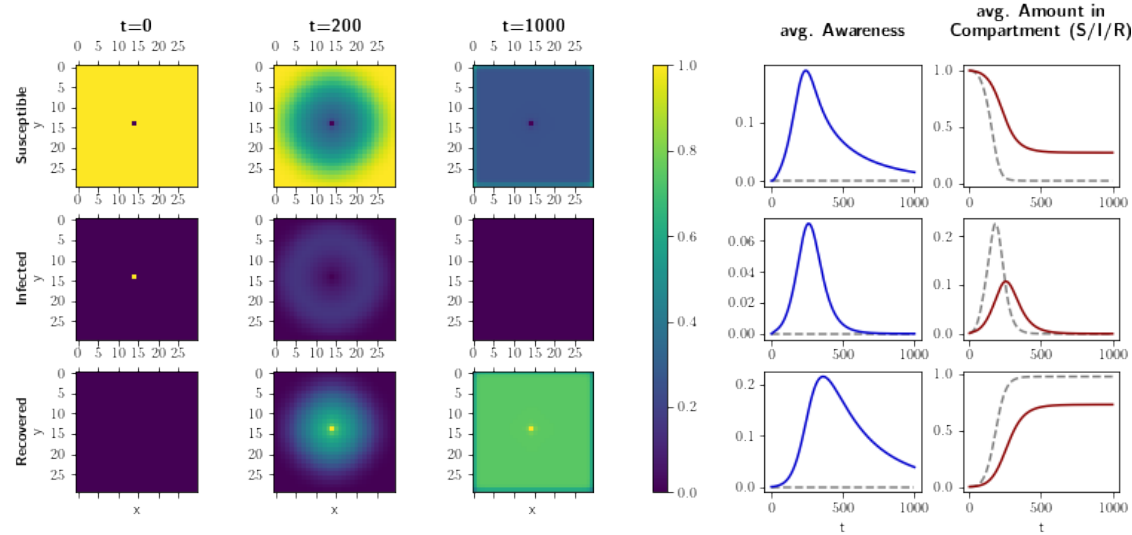


Fig. 7 Comparison of an outbreak with awareness (solid line) and without awareness (dashed line, gray). Here, the awareness reduces both the speed and the intensity.

### 2.0.3 Long Term Stability

In our model, that uses a superposition of SIR and information states for each individual, an outbreak won't completely die out just by chance. Instead, there will always be a small portion of infection remaining in the system. Thus, if the awareness level dropped to zero, a second outbreak might occur.

A stable condition without awareness is reached, when the rate of new infections is less or equal to the rate of recovery with no awareness being present. This happens when a certain number

of people have been removed from the susceptible compartment. For our model, where we used  $\beta = \gamma = 0.03$  this means that 3/4 of the population have to be removed from the susceptible pool as each individual has four neighbors that they can infect with a rate of  $\beta$ . This would correspond to a mean field SIR with  $\beta = 0.12$ ,  $\gamma = 0.03 \Rightarrow R_0 = \beta/\gamma = 4$ .

However, both in the mean field approach and in our grid based model the outbreak will lead to a much higher amount of infected people than the stability threshold requires. This is because when the threshold for stability is reached ( $R_{\text{eff}} = 1$ ) there are still many infections present in the population which still generate new cases, even when the effective reproduction number is below one. Therefore, a lot more than 75 % of people might become infected over time.

The awareness in our model acts against overshooting the stability threshold. For our default parameters the awareness reduces the number of recovered ones at  $t = 1000$  to about 73 % with some awareness still remaining in the susceptible population, which will slow down any further spread while being below the threshold. (The threshold might also be slightly reduced by edge effects.)

If the stability threshold is reached, then no further outbreaks will occur even when the awareness has faded away (figure 8).

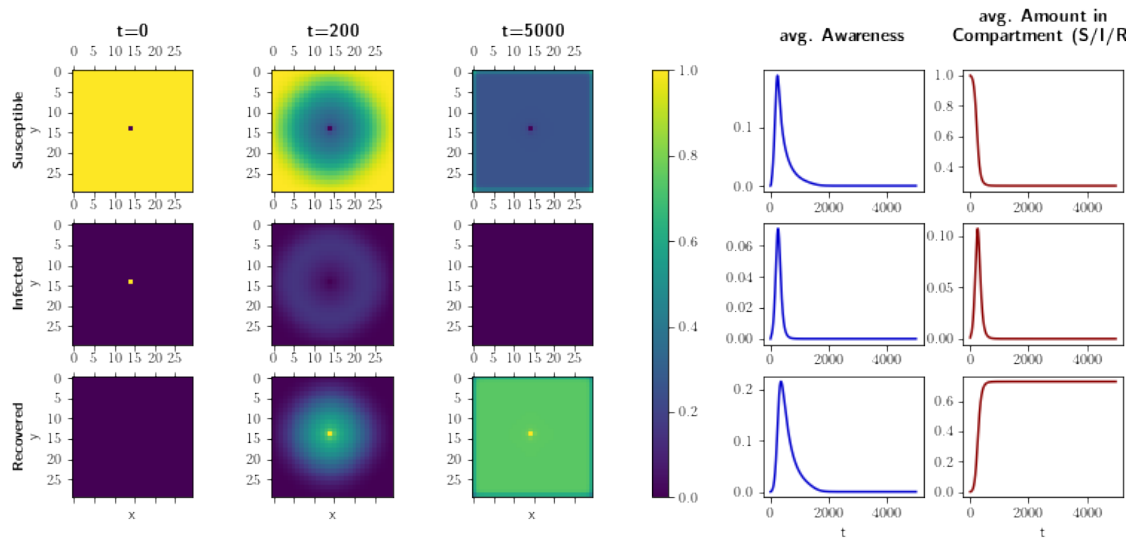


Fig. 8 Long term stability when a total of  $1 - 1/R_0$  have been infected.

When using more aggressive parameters for the awareness, the outbreak seems to be contained at an even lower value than the threshold. In that case, the awareness in the susceptible population will stay high so that the further spread of infections is kept at a very small rate (as depicted in Figure 9).

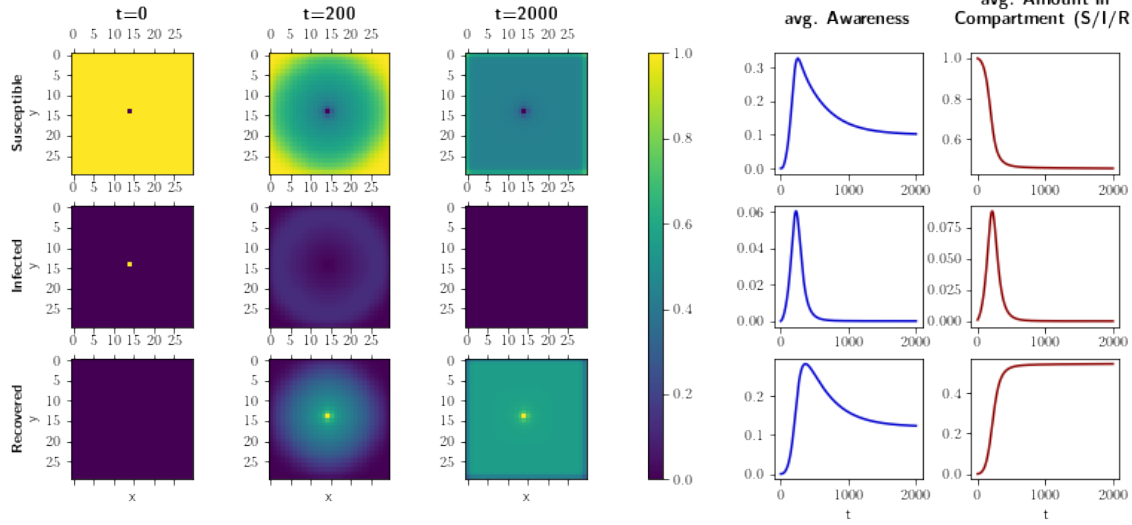


Fig. 9 Long term stability when a total of  $1 - 1/R_0$  have been infected.

However, if the awareness suddenly drops during that time, while the threshold has not been reached, a second wave is possible as figure 10 shows.

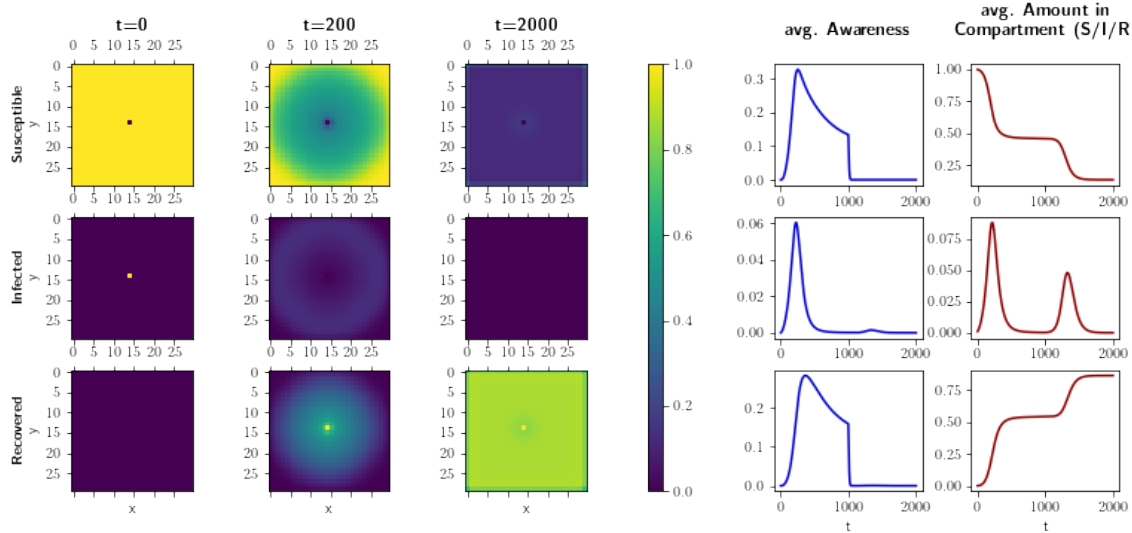


Fig. 10 Long term stability when a total of  $1 - 1/R_0$  have been infected.

## 2.0.4 Sensitivity Analysis

To get a better understanding on how much the different parameters influence the outcome of our system, we run a sensitivity analysis on our model. With regard to our default parameters as a point of reference, we take one parameter and adjusted it to see how much it will change the outcome. This is done for all parameters. As the model is always assumed to reach at least approximately 75 % recovered in the (very) long run and because the time of the end of the initial wave is influenced by the parameters, we choose the maximum number of simultaneously infected people to be our measure for the effectiveness of the awareness. The results from the sensitivity analysis are shown in figure 11.



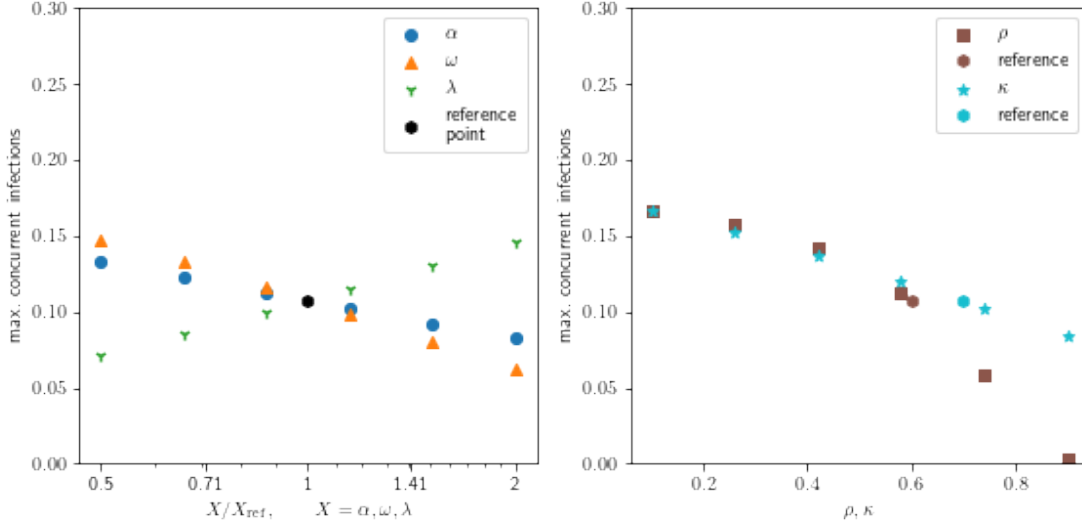


Fig. 11 Sensitivity analysis around the default values.

Before drawing any conclusions about real-world epidemics, one should remember that the sensitivity analysis depends on the chosen point of reference (compare figures 12 and 13). (The gradient varies depending on where you are in the parameter space.) The default values chosen for our model do not necessarily reflect a realistic view of the world. Rather they have been chosen in such a way that the different effects of the model become visible.

Furthermore, one should remember that the effectiveness of the different parameters heavily depends on the assumptions of the model, which might not represent reality in an accurate way (square lattice, etc.).

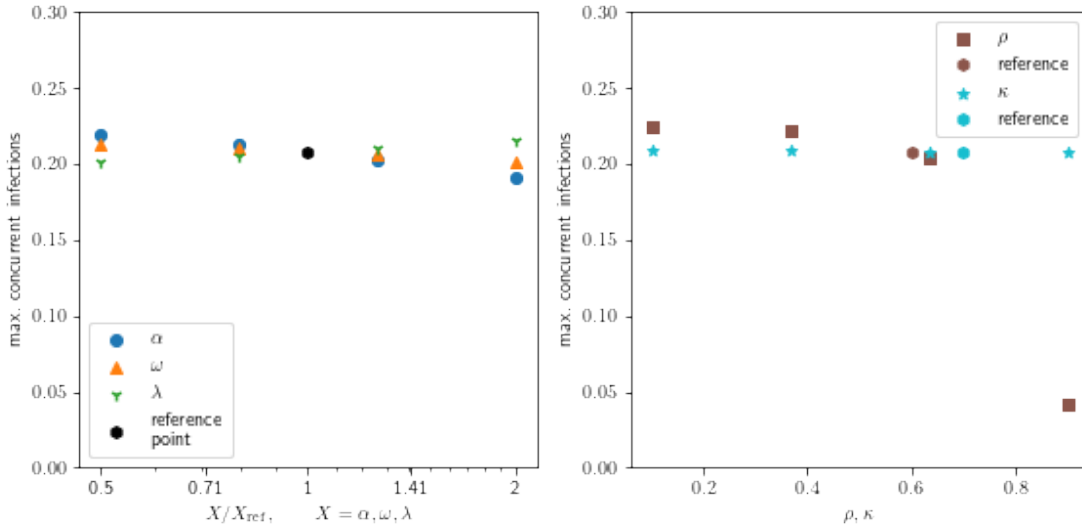


Fig. 12 Sensitivity analysis in a scenario where the spread is dominated by a fast transmission of information ( $\alpha = 0.1$ ) while the information is generated at a slow rate ( $\omega = 0.001$ ). Therefore self isolation only has little impact as the spreading of the virus is much faster than the virus is being detected.

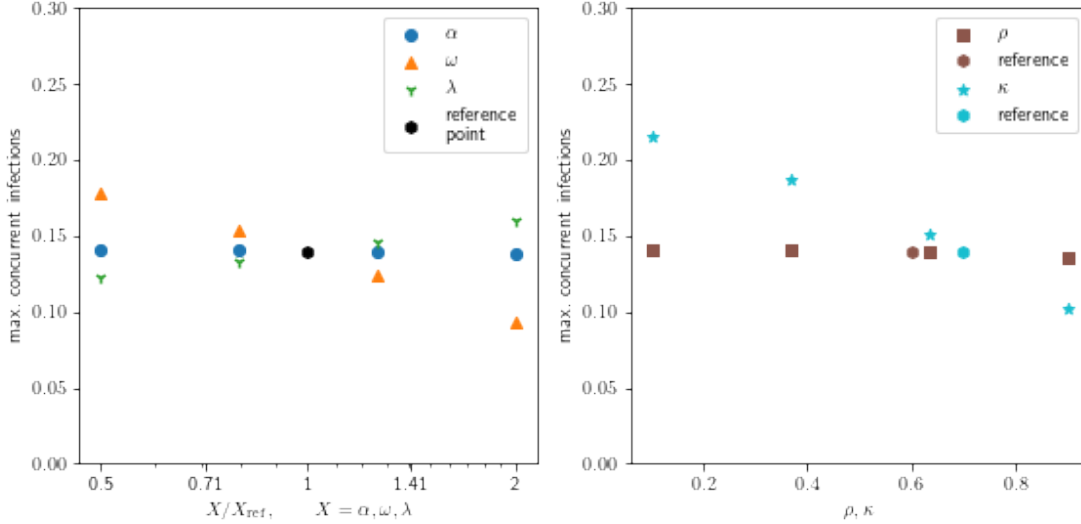


Fig. 13 Sensitivity analysis in a scenario where the transmission of information is very slow ( $\alpha = 0.001$ ) but the detection rate for generating new information is large ( $\omega = 0.1$ ), which leads to self isolation having a much greater impact.

## 2.1 Particular Features of Our Model

Note: This section has not been part of our regular group work but rather came about during debugging and trying to understand why our model behaves in that specific way. It has lead to some backup slides of our presentation in case of more detailed questions about our model. As the findings give some interesting insights into the workings of our model, they are included here. Furthermore, it shows how important it is to carefully consider which aspects to put into a model.

When deciding how to build our model we decided to use a square lattice to reflect a very simple, local contact structure. This distinguishes our model from a mean field approach, where everybody is connected to everybody. As another simplification, we assumed the contact structure for the spread of infections and awareness to be the same.

Instead of using a Monte Carlo simulation we decided to use the expectation values for the interactions instead and use a compartmentalized model, where each individual on the grid is in a superposition of different states.

We decided to have  $3 \cdot 10 = 30$  compartments for each individual (3 for SIR, 10 for awareness) to reflect the correlations between awareness and the SIR state. For example an individual A might have a high awareness probability because individual A is partly infected, but that awareness should not reduce individual A's susceptibility for the part of individual A that is still susceptible.

To keep things simple and to reduce computation time we do not respect correlations between the SIR states of neighbors. This, however, means that an individual A could infect their neighbor (individual B) and individual B in turn can increase individual A's probability of infection. In that way the local behavior of our model is similar to a mean field approach.

To better understand the differences between a mean field approach, a Monte Carlo simulation on a grid (which automatically respects correlations between neighbors) and our superposition model, a quick Monte Carlo simulation and a mean field simulation have been written for an

SIR model (without awareness, the models can be found in `model_inspection.ipynb`). To enable a better comparison, the Runge-Kutta method is dropped and the Euler forward method is used.

When running the simulation with our default parameters, there is a huge difference in the results from the Monte Carlo method compared to the mean field analysis and our own model. The outbreak from the Monte Carlo simulation happens considerably slower and has a real chance of dying out (see figure 14). Because the infections there either happen (with a set probability) or they don't happen (instead of using rates and averages), each individual effectively loses one neighbor to pass on the virus (the one that passed on the infection). Also, people in the same area often share their neighbors, thus further reducing the number of effective contacts.

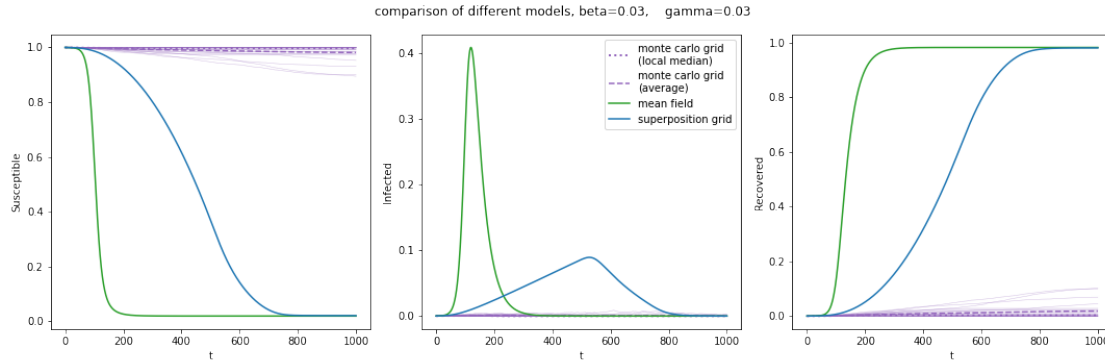


Fig. 14 Comparison between mean field approach (green), our model (blue) and a Monte Carlo simulation (purple) on a square lattice with size  $100 \times 100$  for an SIR model (no information/awareness).  $\beta = 0.03$ ,  $\gamma = 0.03$ . Our model leads to a linear increase in active cases because of the circular pattern moving outwards. The Monte Carlo simulation (25 passes) leads to a much slower spread as the correlations between neighbours are not respected in our model. Here, the Monte Carlo simulation has the possibility to freeze out at low numbers and thus significantly reduces the total number of infections.

When increasing the spreading rate  $\beta$  of the infections, the Monte Carlo method results in a lot more infections and is closer to the mean field analysis (figure 15).

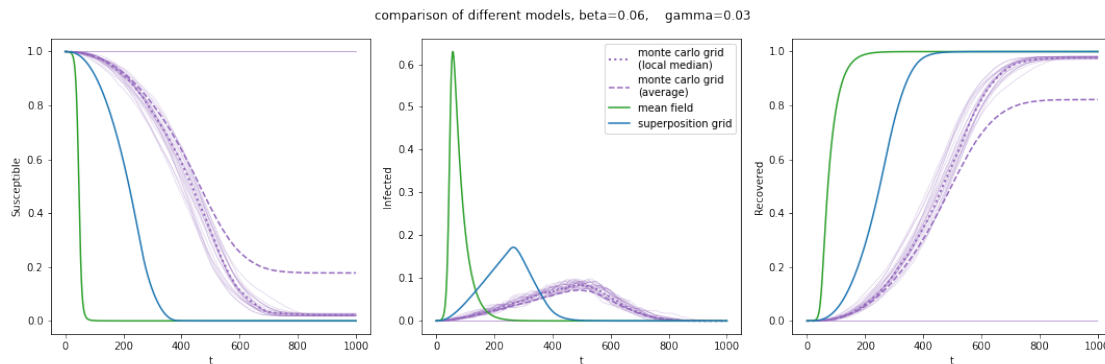


Fig. 15 Comparison between mean field approach (green), our model (blue) and a Monte Carlo simulation (purple) on a square lattice with size  $100 \times 100$  for an SIR model (no information/awareness).  $\beta = 0.06$ ,  $\gamma = 0.03$ . Our model leads to a linear increase in active cases because of the circular pattern moving outwards. The Monte Carlo simulation (25 passes) leads to a slower spread as the correlations between neighbours are not respected in our model.

To get around the back and forth infections and the shared neighbors, a 1D chain is simulated where the infections can only be passed on in one direction and the recovery rate is set to zero. Plotting the local spread at different times shows the speed of the infection wave traveling outward. When comparing our superposition approach to a Monte Carlo simulation, a clear difference can be seen (figure 16).

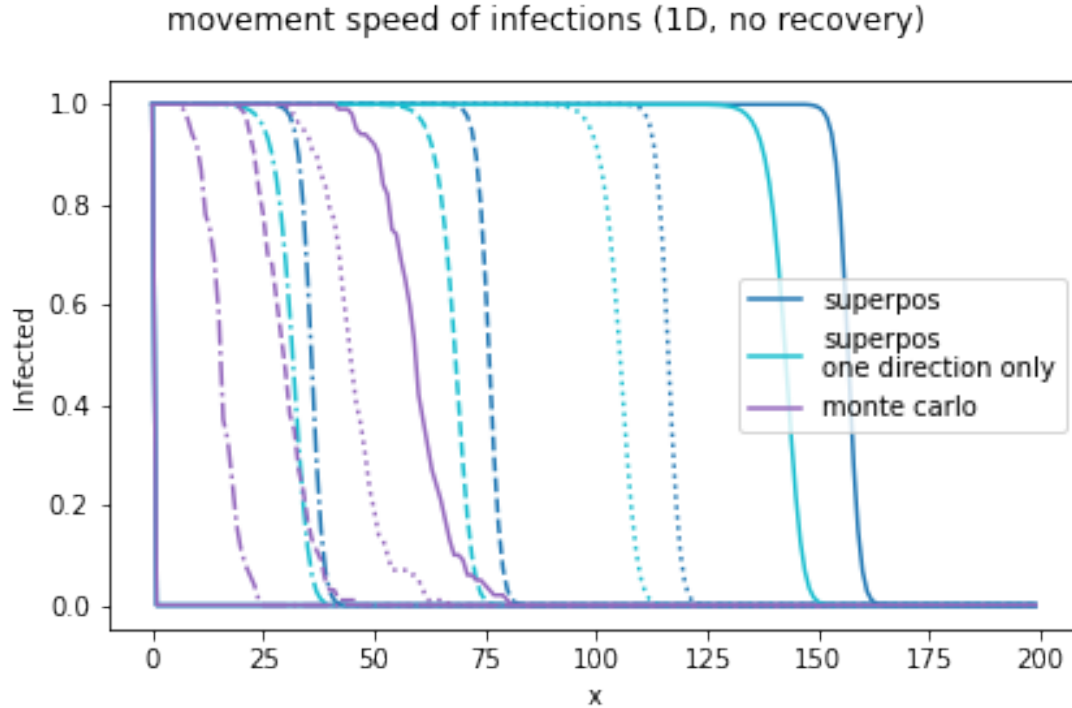


Fig. 16 Movement speed for a 1D chain of people. At  $t = 0$  only the individual at  $x = 0$  is infected. The figure shows the infected people at different time stamps ( $t = 250, 500, 750, 1000$ ) with linestyles: (mixed, dashed, dotted, solid). Blue: Our model. Cyan: A restricted model where infections can only travel from left to right. Purple: A Monte Carlo Simulation with 100 passes.  $\beta = 0.06$ ,  $\gamma = 0$ . It can be seen that our model leads to a much faster spread than the Monte Carlo Method.

If recovery is turned on, the infections along a 1D chain will quickly decay along the distance, as shown in figure 17. However, for a wave front in 2D, people can also infect each other laterally so the infection wave will not necessarily decay with distance.

movement speed of infections (1D, with recovery)

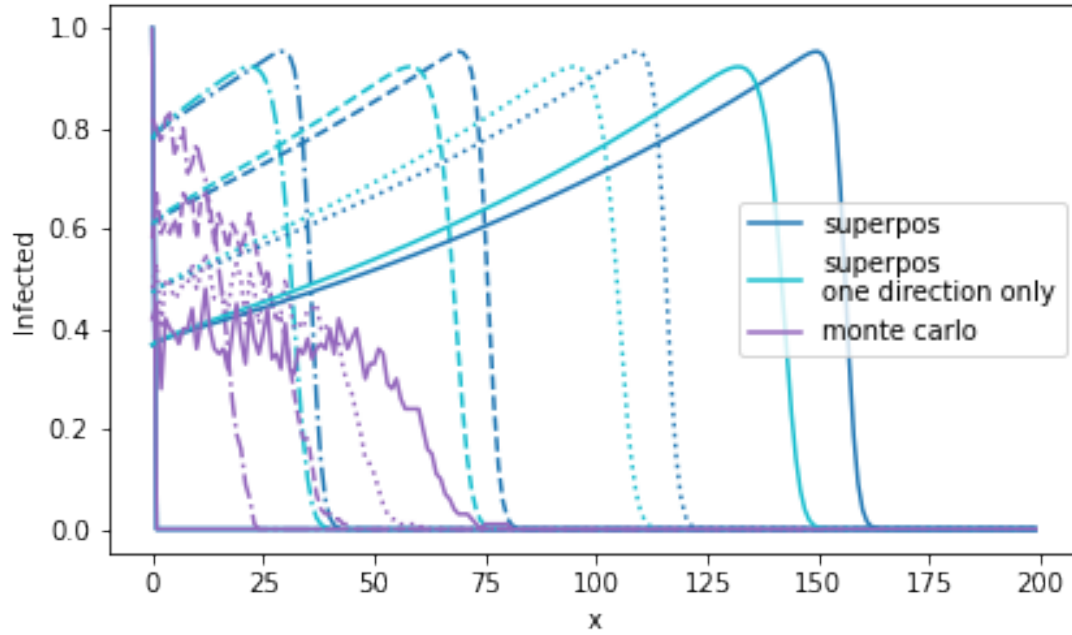


Fig. 17 Movement speed for a 1D chain of people. At  $t = 0$  only the individual at  $x = 0$  is infected. The figure shows the infected people at different time stamps ( $t = 250, 500, 750, 1000$ ) with linestyles: (mixed, dashed, dotted, solid). Blue: Our model. Cyan: A restricted model where infections can only travel from left to right. Purple: A Monte Carlo Simulation with 100 passes.  $\beta = 0.06$ ,  $\gamma = 0.001$ . It can be seen that our model leads to a much faster spread than the Monte Carlo Method. The recovery does not limit the speed of the spread but it can heavily reduce the reach of the spread as there is always a chance of the infection not being passed on.

The difference between the results from our superposition approach and the Monte Carlo simulation can be explained by the neglect of correlations between neighbors in our approach as indicated by figure 18.

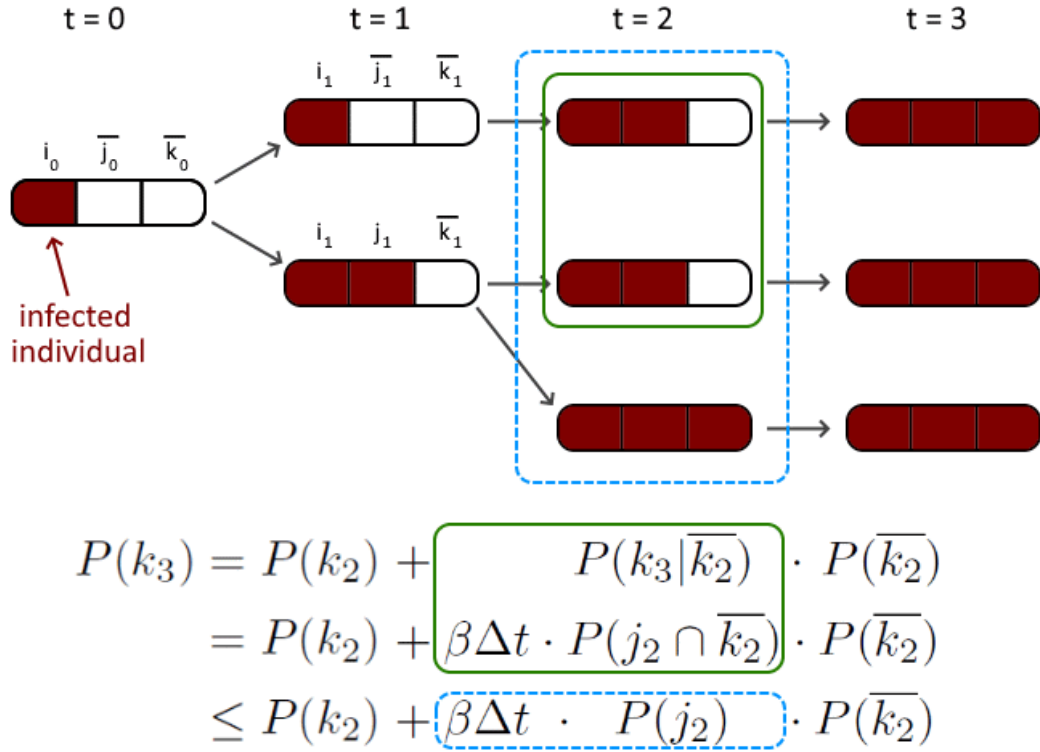


Fig. 18 Simplifications in the model (correlations between neighbors are not respected) lead to both a larger and faster epidemic outbreak compared to a Monte Carlo simulation. The example shown here is a chain of three people with the first one being infected. Blue (dashed) line: assumptions used in the model. Green (solid) line: actual expected behaviour, which corresponds to a Monte Carlo simulation.

### 3 Conclusion

#### 1. Which conditions can help to contain outbreaks through information spreading?

A high decay rate ( $\rho$ ) can “freeze” the outbreak if awareness is maintained for long.

#### 2. What is the impact of self-isolation of knowingly infected individuals on the spread of a disease and under which conditions is it most impactful?

Self-isolation will help limit the spread of the disease, yet it is harder to discern the ideal conditions.

The combination of parameters we chose for our model can lead to a containment of the disease. Self-isolation will lead to a decline in infections.

---

We think our model can be used to deepen the understanding how information influences infection patterns

Given more time to expand or enhance the model, it would have been interesting to add — for example — certain nodes enhancing higher quality information. This could have been influencers in

the community like sport coaches, teachers, social workers, cultural or spiritual leaders, Another approach would have been to include clusters with individuals that can not be reached by information using traditional methods, like people highly influenced by anti-vaccine disinformation ([anti-vaxxers](#)) or people lacking resources to deal with the anxiety that comes with the uncertainties of a pandemic, who tend to tune out information — called '[blunters](#)'.

The last worthwhile attempt we brainstormed was to expand the model to include the technique called "[pre-bunking](#)", meaning the dissemination of high-quality information prior to the first infection in a community.

### 3.1 Limitations

Due to time constraints we had to maintain a very basic model, lacking more complex approaches. The grid approach of four contacts per individual won't represent the communication, and infection, pattern of any community in a realistic fashion. We also had to limit the community size, to deal with limited computing power. The percentual approach would have run indefinitely, which is also a non-realistic outcome for a pandemic. Also, the neglectance of correlations between neighbours leads to deviations compared to a Monte Carlo Simulation. Randomization would have led to more accurate results.

Created in Deepnote