Spread of Disease

Kaixin Hu and Suzan Blommestijn

October 24, 2015

Abstract

For centuries people have tried to prevent diseases from spreading. One tool that can be used to prevent epidemics is simulation. This report explains how Cellular Automaton and the SIR Model can help simulate the spread of disease. Five different simulations are discussed: One Deterministic Model, three different Stochastic Models and an extra set of models with a moving Typhoid Mary. Whether or not an epidemic occurs in these models turned out to depend on the probability of catching a disease and on the initial distribution of susceptible, infected and recovered people. Though the models discussed might be too simple to represent the real world, they give an idea to how these models can be used to analyse the spread of disease.

1 Introduction

Everyone gets sick every now and then. Whether it is just a cold, a touch of the flu, or something more serious. Diseases are all around us, but only on rare occasions does the spreading of a disease result in an epidemic. Though the definition of an epidemic is not agreed on [2], in this report the following description will be used: when over a time period of a hundred days the population of infected people does not reach zero.

There are many ways to model the spread of disease. Sadilek, Kautz and Silenzio (2012), for example, studied the spread of the flu by analysing the social media network Twitter. [4] Another method is to use mathematical equations. [3] In this report, however, a Cellular Automaton (CA) environment will be used to investigate when an epidemic does and does not occur.

When modelling the spread of a disease there are a lot of factors to take into consideration. Culture, sexuality and connectedness of people, for instance, could be taken into account. Because the main goal of this paper is to model the spread of disease in general, those factors will not be included in this project. The paper will discuss different CA models. First a short explanation of CA and the variables used in this report will be given. Second a deterministic model is discussed. Where only the initial number of susceptible, infected and immune people is dependent on probabilities. Next three different stochastic models will be discussed. Where the difference between the models lies in the way the

susceptible people can be infected by their neighbors. Finally a set of models with one moving person will be described.

2 Methods

In this section first a description of the terminology and the variables used in the simulations will be given. An explanation of how CA can be used to help model the spread of disease will be provided. Next both the deterministic and stochastic simulations will be discussed. Finally a simulation where one moving cell is infecting other cells will be described.

2.1 CA and SIR model

For this project the python language was used to set up cellular automaton simulations. Cellular Automata (CA) can be described as optimizations of systems where both time and space are discrete. [6] In a CA simulation cells are set on a grid. These cells can have an infinite possibility of states. Cells adjacent to the cell of interest are called the cell's neighbors. A set of rules can be applied to determine the state of the cell at a next time step. Each time step (t) every cell is checked and updated to the new state according to the set of rules. [5] Important in modeling with CA are the neighbors of each cell. For this project the von Neumann neighbourhood was used. This means that only the four cells that touch sides with a cell are considered the neighbors of this cell. [1]

This project uses the SIR model to simulate the spread of disease. In this model S stands for the susceptible, I for the infected, and R for the recovered or immune people. [5] In our CA environment the susceptible people were represented by green cells, the infected people by shades of blue cells and the immune people by shades of red cells. The lighter the shade of blue or red, the further along in infection or immunity the cell was. For example, a cell in day two of infection had a lighter shade of blue than a cell in day one of infection. Two probabilities were used to set the initial state of the grid: probSusceptible and probInfected. ProbSusceptible is the probability that in the initialisation state (at t=0) a person is susceptible (value 0). ProbInfected determines which of those not susceptible at the initialisation state are infected. The day or level of infection (value 1 or 2) will be uniformly distributed over those infected cells. Finally, those cells that are not susceptible or infected will be uniformly distributed over five days or levels of immunity (value 3 - 7). The size of the grid for all the simulations was set to 50 x 50, to represent 2500 people as cells.

After the initialization state it differs per model how the states changes per time step. In the stochastic models two more probabilities are introduced: probCatch and probBeSusceptible. ProbCatch is the probability that a susceptible individual will get infected. This probability can be set or be solely based on the neighbourhood of the cells. ProbBeSusceptible is the probability that an individual who has been immune for at least five days becomes susceptible again.

This probability will only be used in Stochastic Model 1. The time period that was chosen for the models was a hundred days, with time steps of a day.

2.2 SIR Deterministic Model

The first model that was simulated was a deterministic SIR model. Infection lasted two days, immunity lasted five days. It was deterministic in the sense that each person, represented by a cell on a grid, followed a pattern of susceptibility, infection and recovery. Only in the initialization state there were probabilities to determine who was susceptible, infected or recovered.

The deterministic process of a cell through time was as follows. If a cell has the value '0' and one of its direct neighbors had a value of '1' or '2' at time is t+1 the susceptible cell would become infected (1) itself. After day one of infection it would go to day two (2), and after that go through the five days of immunity (3 to 7). After day five of immunity it would become susceptible again (0).

The chosen probabilities for both the initial susceptibility and initial infection rate were set to the values 0.1, 0.5 and 0.9. Because one probability was held stable while the other varied, this resulted in nine separate simulations. The simulation time was set to 100. This number was chosen to make statistically sound conclusions, while not letting the time it took to run the simulation be too long.

Figure 1 shows one possible image of the initial state for the deterministic model. The probSusceptible is set at 0.5 and the probInfected at 0.9. With these probabilities you can see that there are a lot of susceptible and infected, and not so many recovered cells in the initial state.

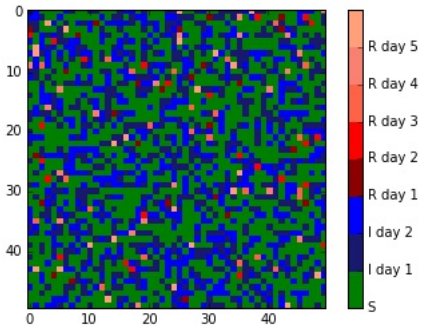

Figure 1: Example of Initial Grid for the Deterministic Model

2.3 SIR Stochastic Model

2.3.1 SIR Stochastic Model 1: infection depends on given probability

The second model is a stochastic model. It is stochastic in the sense that the infection of a susceptible individual and susceptibility of a recovered individual depend on probabilities.

In Stochastic Model 1, the rule for setting the initial state is the same as in the Deterministic model. But in the Deterministic model, individuals exposed to sick neighbors will surely get sick, and individuals who have been immune for five days will automatically become susceptible again. In order to model the reality more accurately, in this model two new variables, probCatch and probBeSusceptible, are set to make the model stochastic. ProbCatch is the probability that a susceptible individual who has a sick neighbor will get sick, and probBeSusceptible is the probability that an individual who has been immune for fiv days will become susceptible.

Similarly as in the Deterministic Model, in this part the values of 0.1, 0.5 and 0.9 are set respectively to the initial susceptibility and initial infection rate. In addition, the values for probCatch and probBeSusceptible were set to 0.1, 0.5 and 0.9. This resulted in 81 separate simulations for this model. As in the Deterministic Model, the simulation was run for a 100 times.

2.3.2 SIR Stochastic Model 2: infection depends on number of infected neighbors

The third model was a stochastic model where the infection of a susceptible person was dependant on the number of infected neighbors. If four neighbors were infected, it was certain that the susceptible person would get infected. If, however, only one out of four neighbors was infected, there was only a 0.25 probability that the susceptible person would get infected. The initialization state was the same as in the deterministic simulation.

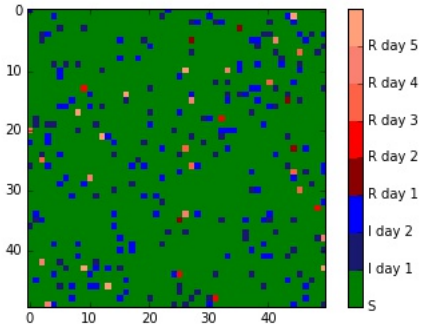

Figure 2: Example of initial grid for Stochastic Model 2

Figure 2 shows an example of the initial state, where both probSusceptible and probInfected are set to 0.9. This results in a high number of susceptible cells, a lower number of infected cells, and only a couple of recovered cells.

2.3.3 SIR Stochastic Model 3: infection depends on day of infection of neighbors

The forth model was a stochastic model where the infection of a susceptible person was dependant on both the number of infected neighbors, and the level of infection. If, for instance, a susceptible person had two neighbors and one was in day two of infection and the other in day one of infection, the susceptible person would have a chance of $(1 + 0.5)/2$ to get infected. However, if both of the neighbors were in day one of infection he would have a chance of $(1 + 1)/2 = 1$ to get infected.

2.3.4 Models with Typhoid Mary

For the final model, Typhoid Mary is included. Typhoid Mary is a carrier who never gets sick. Mary walks at random through the grid, and at each step she changes places with the individual in whose cell she steps. The values of probSusceptible and probInfected were set to 0.1, 0.5 and 0.9. This again resulted in nine separate simulations. Each simulation was run for a 100 times.

2.3.4.1 SIR Deterministic Model with Typhoid Mary

In this model, Typhoid Mary is included, but the other rules are the same as in the Deterministic Model. That is, probSusceptible and probInfected determine the initial state, and after that the model is deterministic. The only difference is that Typhoid Mary changes position with one of her neighbors randomly, and her cell value stays '8' at every time step.

2.3.4.2 SIR Stochastic Model 3 with Typhoid Mary

In this model, Typhoid Mary is included, but the other rules are the same as in Stochastic model 3. That is, probSusceptible and probInfected determine the initial state, and probCatch of a susceptible cell depends on the level of infection of the neighbors of the cell. Typhoid Mary counts as a level '1' infected cell in calculating the probCatch. For example, if a susceptible cell has three neighbors and one of those neighbors is Typhoid Mary, but no other neighbors are infected, probCatch = $(1 + 0 + 0)/3$. Typhoid Mary changes position with one of her neighbors randomly, and her cell value stays '8' at every time step.

2.3.4.3 SIR Deterministic Model with Typhoid Mary, initially no one is sick

In this model, initially no one is sick, so initially every cell is set with the value of '0'. Only one cell is randomly chosen and set with the value of '8', representing

Typhoid Mary. For the following time steps the updating rule of each cell is the same as in the Deterministic Model with Typhoid Mary.

2.3.4.4 SIR Stochastic Model 3 with Typhoid Mary, initially no one is sick

In this model, initially no one is sick, so initially every cell is set with the value of '0'. Only one cell is randomly chosen and set with the value '8', representing Typhoid Mary. For the following time steps, the updating rule of each cell is the same as in SIR Stochastic Model 3 with Typhoid Mary.

3 Results

For every model the graphs of the means of the hundred simulations were plotted to show the population size for S, I and R over a hundred day time period. Error bars were added to represent the 95 percent confidence interval, which was calculated by taking twice the standard deviation of the hundred simulations per time step of a hundred days. If the error bars are wide, this means that each time the simulation ran, the population size was different. Or in other words, that there is a lot of variance between simulations. A small error bar means that the results in each simulation are consistent, or there is little variance.

3.1 Deterministic SIR Model

The simulation was run for the nine different sets of combinations of probabilities for probSusceptible and probInfected. Figure 3 shows the graphs of these simulations. When probSusceptible is low or intermediate and probInfected is low or intermediate an epidemic occurs: The population of infected cells stays above zero for at least a hundred days and the error bars are small. When one of the probabilities is high, however, the error bars become so wide that it is hard to draw conclusions from the graphs. The results varied to much between those simulations. Another interesting observation is that in the first set of graphs, where the error bars are low, the graph seems to repeat itself after a twenty day time period. 4 shows the population for S, I and R over a twenty day time period, for probSusceptible is 0.5 and probInfected is 0.5.

Another interesting finding in this model is the following. As can be seen in figure 5, the spread of disease follows a strange pattern. It does not look natural and the stochastic models did not show this behaviour.

3.2 SIR Stochastic Model 1

With the combination of three values set to probSusceptible, probInfected, probCatch, and probBeSusceptible respectively, there are 81 situations in this model, as showed in figure 6. The 81-situation-results in Stochastic Model 1 can be roughly classified into three groups, as shown in figure 7. The first kind of

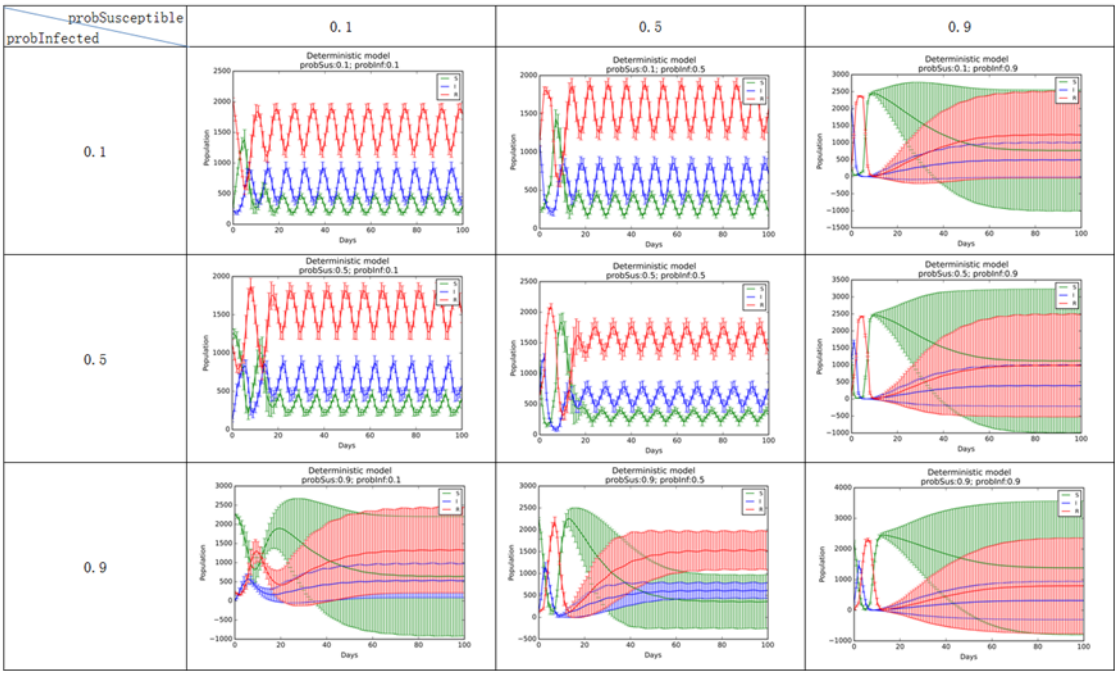

Figure 3: Results of Deterministic Model

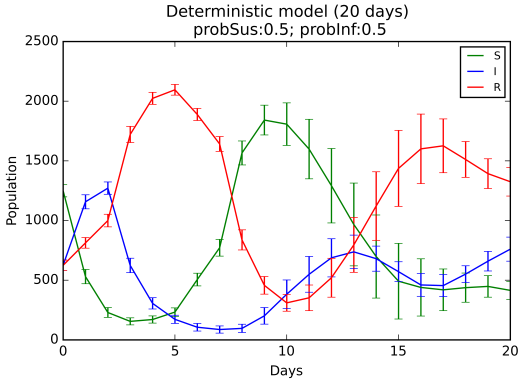

Figure 4: Deterministic model: 20 day period. With probSus = 0.5 and probInf = 0.5.

outcome means that there is no spread of disease, such as in the situation with the combination of :probSus=0.1; probInf=0.1; probCatch=0.1; probBeSusceptible=0.1. The second kind of outcomes in the graph means that there are large deviations between the 100 simulations. Statistically speaking, it is not

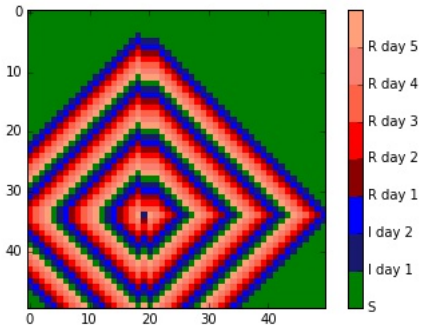

Figure 5: Example of the pattern for the Deterministic Model

clear whether an epidemic occurs or not. For example, as in the situation of: probSus=0.1; probInf=0.1; probCatch=0.5; probBeSusceptible=0.1. The third kind of outcomes means there is an epidemic, such as in the situation of: probSus=0.1; probInf=0.1; probCatch=0.5; probBeSusceptible=0.5.

As shown in the figure 6, the situation-grids colored with green, there is no spread of disease. When colored with blue, there is large deviations between each iteration, the result of the breakout of the disease is hard to predict. When colored with red, it is said that there is an epidemic. The distribution of three kinds of results in the table shows that the smaller probCatch and probBeSusceptible, and the larger the probSusceptible, the more likely there will be an epidemic.

3.3 Stochastic SIR Model 2

For the second Stochastic Model there were nine combinations of probSusceptible and probInfected, which resulted in nine different sets of simulations. The resulting graphs are shown in 8. As can be seen in the graphs in none of the probability combinations does an epidemic occur: the number of infected cells does not stay above zero. It seems that, based on this model, when the probability of getting infected depends solely on the number of neighbors, the disease will soon die out.

3.4 Stochastic SIR Model 3

The results of the third Stochastic Model are shown in Figure 9. Again nine graphs are shown for all the possible probability combinations. And like in Stochastic Model 2 there is no spread of disease: an epidemic does not occur in any of these simulation sets. This makes sense, since the probability of catching the infection does not only rely on the number of neighbors, but also on the level of infection. Which can result in an even lower probCatch then in Stochastic

probInfected \ probSusceptible		0.1			0.5			0.9		
	probBeSusceptib \ probCatch	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
0.1	0.1	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease
	0.5	large deviations	disease spreading	disease spreading	large deviations	disease spreading	disease spreading	some deviations	large deviations	large deviations
	0.9	large deviations	disease spreading	disease spreading	large deviations	disease spreading	disease spreading	no spread of disease	large deviations	large deviations
0.5	0.1	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease
	0.5	disease spreading	disease spreading	disease spreading	some deviations	disease spreading	disease spreading	large deviations	disease spreading	disease spreading
	0.9	disease spreading	disease spreading	disease spreading	large deviations	disease spreading	disease spreading	some deviations	large deviations	large deviations
0.9	0.1	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease	no spread of disease
	0.5	disease spreading	disease spreading	disease spreading	some deviations	disease spreading	disease spreading	large deviations	disease spreading	disease spreading
	0.9	large deviations	large deviations	large deviations	large deviations	large deviations	large deviations	some deviations	large deviations	large deviations

Figure 6: Results of Stochastic Model 1

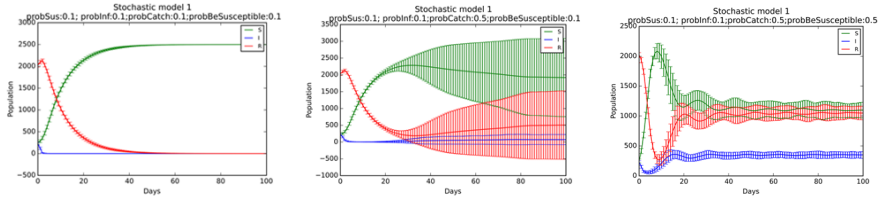

Figure 7: Classification of results in Stochastic Model 1

Model 2. The results show that there is simply to low a chance for a susceptible cell to get infected.

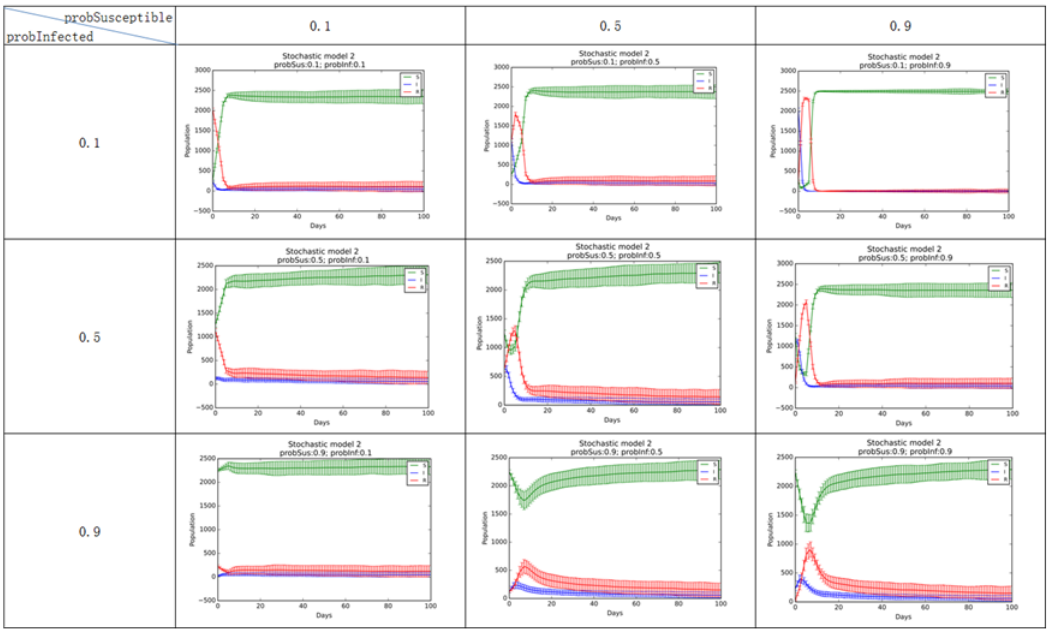

Figure 8: Results of Stochastic Model 2

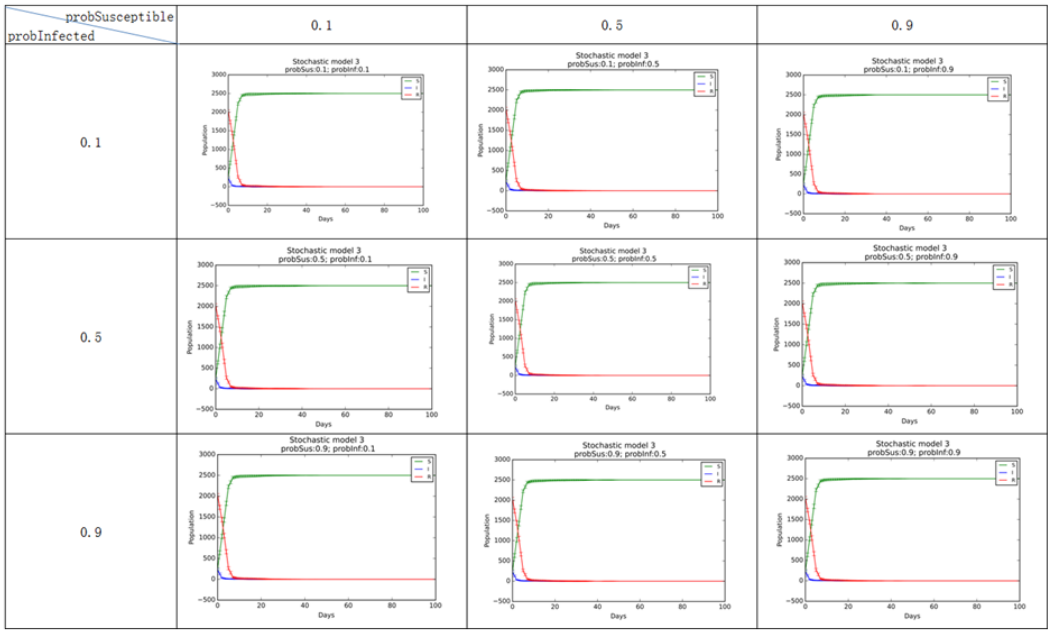

Figure 9: Results of Stochastic Model 3

3.5 Models with Typhoid Mary

3.5.1 SIR Deterministic Model with Typhoid Mary

With the combination of three values set to probSusceptible and probInfected, there are nine situations in this model. As shown in figure 3 and figure 10, the outcomes of Deterministic Model with Typhoid Mary are similar to the related model without Mary, except for one difference. After Mary is included in, the 5 situations where there is large deviations between 100 iterations, show smaller deviations. This means that with Typhoid Mary walking randomly in the crowd in the Deterministic Model, the tendency for the spreading of disease increases.

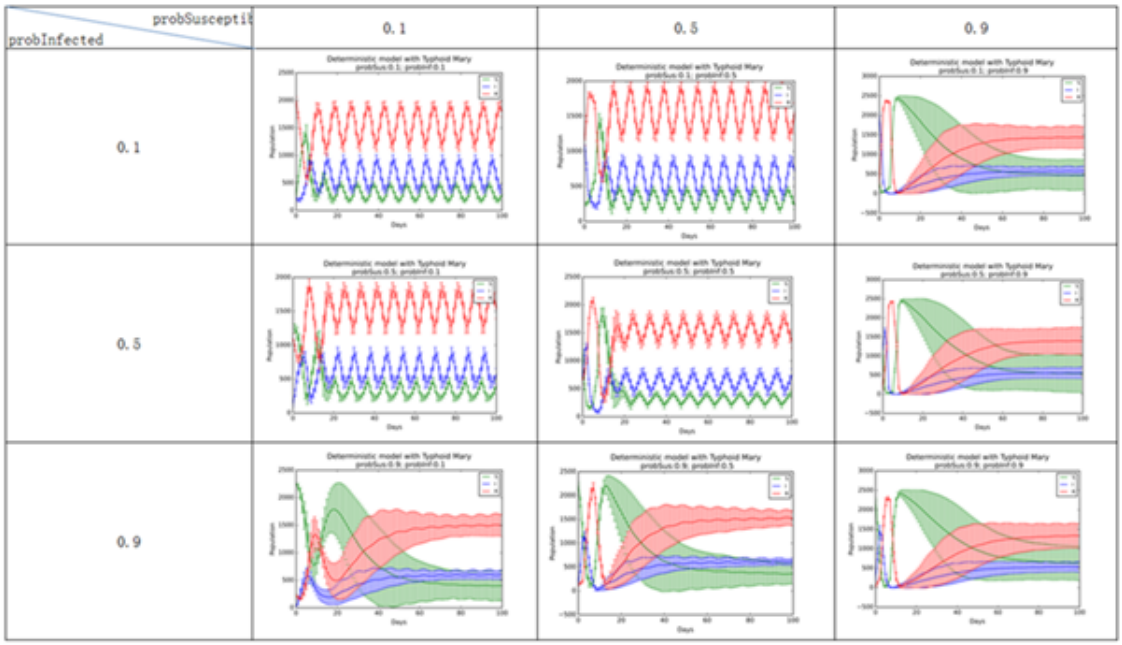

Figure 10: Results of Deterministic Model with Typhoid Mary

3.5.2 SIR Stochastic Model 3 with Typhoid Mary

With the combination of three values set to probSusceptible and probInfected, there are also nine situations in this model.

As shown in figure 9 and figure 11, the outcomes of Stochastic Model 3 with Typhoid Mary are similar to the corresponding model without Mary, except for one difference. Although there is still no spread of disease, the standard deviations are a little bigger around the 20th day. But the error bars narrow to close to zero afterwards. It means that the existence of Typhoid Mary in the crowd, to some extend, makes the situation more complex. But in this model, Mary just influences several cells around her, so the disease is confined to a small area. The area is always moving with the move of Mary. An epidemic does not occur due to the relatively small intrinsically determined value of probCatch, compared with the Deterministic Model with Typhoid Mary.

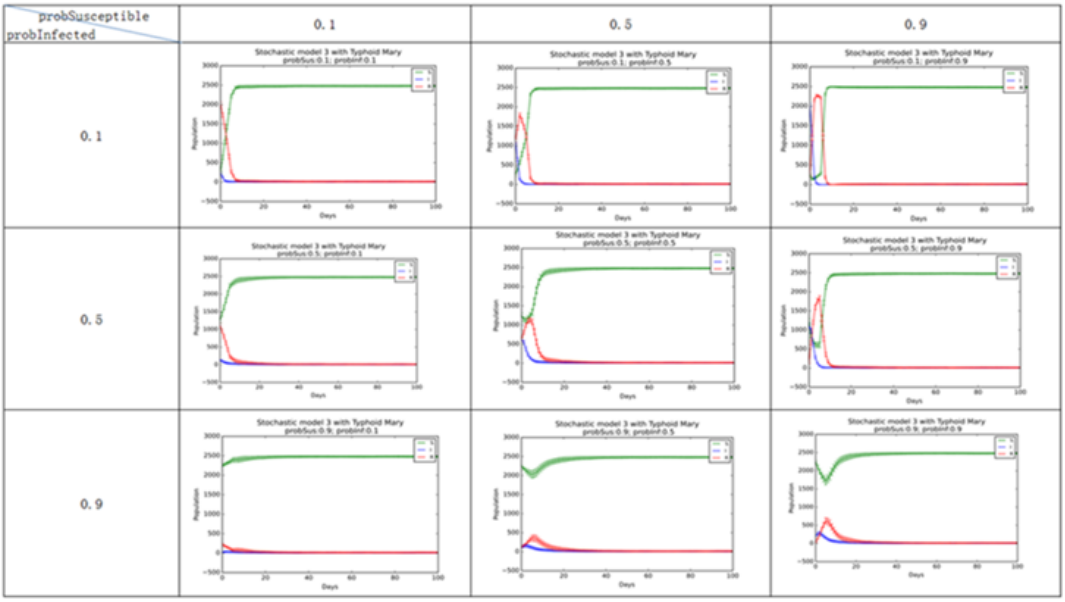

Figure 11: Results of Stochastic 3 Model with Typhoid Mary

3.5.3 SIR Deterministic Model with Typhoid Mary, initially no one is sick

In the Deterministic Model with Typhoid Mary, initially no one is sick, there is still an epidemic, as shown in Figure 12. So even if the grid initially only exists of susceptible cells, Typhoid Mary makes sure an epidemic occurs. This is because in the deterministic models the probability of catching the infection is set to 1.

3.5.4 SIR Stochastic Model 3 with Typhoid Mary, initially no one is sick

In Stochastic Model 3 with Typhoid Mary, where initially no one is sick, an epidemic does not occur (Figure 13. The disease is confined to a small area around Typhoid Mary. This result in the 20-to-100-day time step is similar to the Stochastic Model 3 with Typhoid Mary, initially some individuals are infected.

3.5.5 Comparison of the Four Different Models

All the results discussed above look at each model separately. Here a short comparison between the four different models is given. Figure 14 shows the different graphs where the probabilities probSusceptible and probInfected are

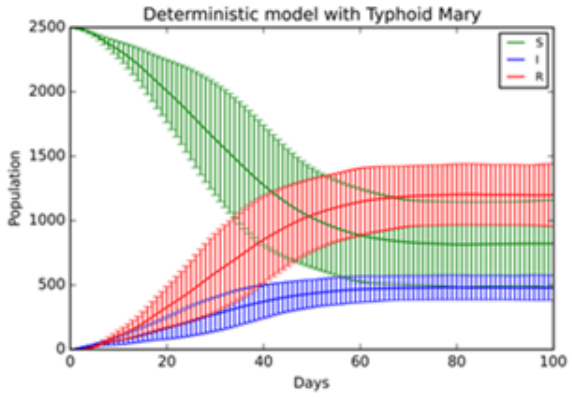

Figure 12: Results of Deterministic Model with Typhoid Mary, no one is sick initially

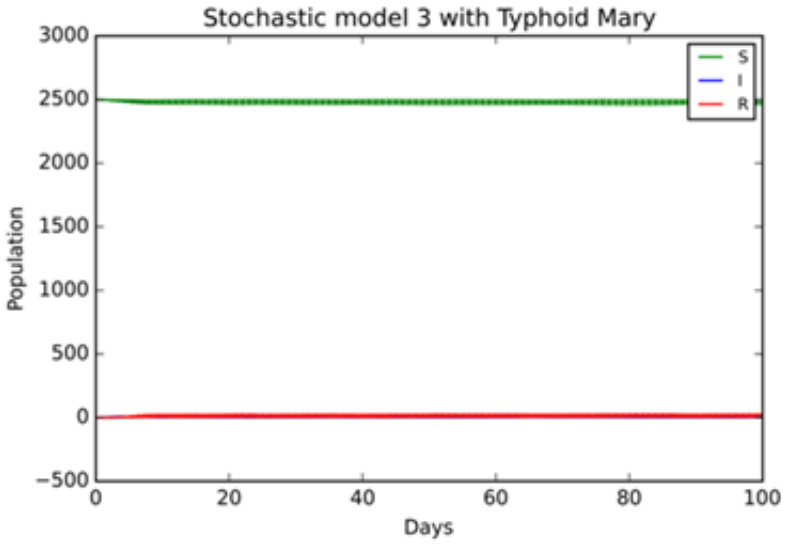

Figure 13: Results of Stochastic Model 3 with Typhoid Mary, no one is sick initially

set to 0.1. The deterministic model clearly shows an epidemic occurs, while the first stochastic model only does when the probCatch is set high (0.9). What can be seen in the graphs is that when probCatch declines, at some point an epidemic cannot occur. In addition, the inclusion of Typhoid Mary in the Deterministic Model or Stochastic Model 3, does not influence the result of the spread of

disease.

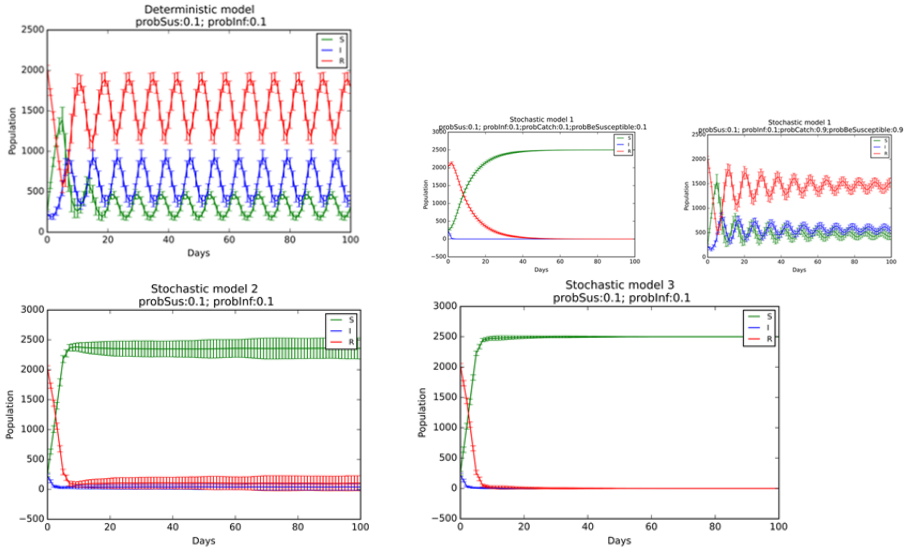

Figure 14: Results of different Models when initial state is set equal

4 Conclusion

The main goal of this project was to investigate when an epidemic does and doesn't occur. From the results it can be concluded that the probability of catching a disease and the initial distribution of susceptible, infected and recovered people determine whether or not an epidemic occurs. When the probability of catching a disease is high and the population has a nice spread of susceptible, infected and recovered people, an epidemic is likely to happen according to the findings of this report.

This project only shows a small part of how the SIR model can be used to study the spread of disease. In the next two parts the limitations and suggestions for further research will be given.

4.1 Limitations and Further Research

The values set for the two pairs of probabilities for the simulations are limited. This article only selected three values for the probSusceptible and probInfectious to create the initial state. Similarly, in Stochastic Model 1, the values 0.1, 0.5, and 0.9 are assigned to probCatch and proBeSusceptible to analyse the situation. In order to analyse under which situations an epidemic occurs or not, one should set more values to the two pairs of probabilities. For example, with an interval

of 0.01 for all the probability values, it can be determined what the critical point of the spread of disease is.

Also, instead of running the simulation a hundred times, it would be more statistically sound to run it a thousand times. Due to the amount of time it took to run each simulation it was not possible to set the simulation time this high for this project. Either parallel programming, or a stronger computer might be able to solve this issue in further research.

In addition, there is a gap between these models and reality. In the real world people are not surrounded by a maximum of four neighbors. To build a model that can simulate the reality more accurately, we may try to adopt other methods for comparison, such as graph theory to set social networks. Or use ODE's to model the spread of disease. Furthermore, it would be even better to compare those models with real data.

Finally, in the models discussed in this report only the Von Neumann neighbourhood is used. For the deterministic model this resulted in unnatural behaviour (see figure 5. This might be explained by the fact that it was a deterministic model, which itself is an unnatural representation of reality. However, it might be interesting to add different neighbourhood methods (for instance Moore's neighbourhood) to analyze if this results in more natural behavior.

References

[1] CHJ Davies. The effect of neighbourhood on the kinetics of a cellular automaton recrystallisation model. *Scripta metallurgica et materialia*, 33(7):1139–1143, 1995.

[2] MS Green, T Swartz, E Mayshar, B Lev, A Leventhal, PE Slater, and J Shemer. When is an epidemic an epidemic? *The Israel Medical Association journal: IMAJ*, 4(1):3, 2002.

[3] Leonid A Rvachev and Ira M Longini. A mathematical model for the global spread of influenza. *Mathematical biosciences*, 75(1):3–22, 1985.

[4] Adam Sadilek, Henry A Kautz, and Vincent Silenzio. Modeling spread of disease from social interactions. In *ICWSM*, 2012.

[5] Angela B Shiflet and George W Shiflet. *Introduction to computational science: modeling and simulation for the sciences*. Princeton University Press, 2014.

[6] Stephen Wolfram. *Cellular automata and complexity: collected papers*, volume 1. Addison-Wesley Reading, 1994.