

Mathematical analysis of epidemic systems- comparison of different models

Torbjørn Paschen Seland

Dec 11, 2014

Contents

0.1	Introduction	3
1	ODE models	5
1.1	Simple Epidemic Models	5
1.2	Threshold phenomenon	6
1.3	English Boarding School 1978	10
1.4	Zombification	13
1.4.1	Parameters used in the model	14
1.5	Discussion	26
2	PDE models	27
2.1	Simple system for spatial spread	27
2.1.1	Travelling wave 1D	28
2.1.2	Verifying the solution	30
2.1.3	Constant solution	31
2.1.4	Manufactured solution	32
2.1.5	Convergence rate	34
2.1.6	The spatial error	34
2.1.7	The temporal error	35
2.1.8	Travelling wave in 2D	35
2.1.9	A gaussian wave	36
2.2	English Boarding School	40
2.2.1	Introducing a Gaussian distribution of infected	40
2.3	Zombiefication	44
2.3.1	The position of the <i>Infected</i> group	47
2.3.2	Free areas for the <i>Susceptible</i> group	50
2.3.3	Ten minutes at Frederikkepllassen	52
2.3.4	Random distribution	54
2.4	Appendix	58
2.4.1	Sympy to find manufactured solution	58
2.4.2	Discretization	58
2.4.3	2D Gaussian function from $x=0$	59
2.4.4	2D Gaussian function from $x=0, y=0$	61
2.4.5	2D Gaussian function from $x=0, y=0$ with higher initial value	61

2.4.6	English Boarding School	61
2.4.7	Zombieification	63
3	Random walk	67
3.1	Monte Carlo methods	67
3.1.1	Random variable	67
3.1.2	probability distribution functions (PDF)	68
3.1.3	moments of a PDF	69
3.1.4	The pertinent variance σ^2	70
3.2	Random walks	70
3.2.1	Diffusion equation and random walks	70
3.2.2	Random walker	72
3.3	English Boarding School	73
3.3.1	A Lower average value	75
3.3.2	Threshold value	76
3.4	Zombieification	76
3.4.1	Random walk	77
3.4.2	Moving smart	79
3.4.3	Three phases in Walking Dead	80
3.5	Discussion and conclusion	81

0.1 Introduction

Throughout history large epidemic diseases have spread around the world, often over large geographical areas. These diseases have done great harm on the population and millions of people have died. The Black Death and Cholera are epidemics that have moved over large distances into Europe. When the Black Death came to Europe in 1347, it killed about a third of the population, which at that time was about 85 million people [?, p. 315]. These diseases often gave physical symptoms, which have given important knowledge through history to prevent new outbursts and to cure already infected humans. They have various outbreaks, but are often related to connections between humans and animals. Malaria is an example of a disease that transmits from mosquitoes to humans. There have also been various explanations of the spread and cause of epidemics. AIDS(autoimmune deficiency syndrome) has as an example been ascribed by many as a punishment sent by God.[?, p. 316].

The first major epidemic in the U.S.A was the Yellow Fever, discovered in 1793 in Philadelphia. 5000 of a population of 50 000 died. About 20 000 fled the city and the situation was quite chaotic[?, p. 316]. This had a major impact on the subsequent life and politics of the country. The power of a disease can do larger damage, with respect to death, than a war.

After World War II, public health strategy has focused on elimination and control of organisms which cause disease. In 1978 United Nations sat a goal of eradicating all diseases by year 2000. A large job has been done and smallpox is an example of a disease that was last seen in Somalia in 1977[Ref:cdc.gov]. AIDS was later recognized and has been difficult to stop the spread.

Another important aspect in the current spread of diseases, is the displacement of human populations. About a million people cross international borders daily. The growth of human population, especially in underdeveloped countries, is a factor that affects the spread. These conditions played a key role in the spread of HIV(human immunodeficiency) in the 1980's. World Health Organization has estimated that around 32.6 million are infected with the HIV virus today[web:<http://www.who.int/features/factfiles/hiv/en/>].

Knowledge through history is important for the control of different epidemics, but also important in detecting new diseases. The plague of Athens had been studied in great detail by Thucydides in 430-438 BC. Similar had been done with the 'sweating sickness' in the late 15th and first of the 16th centuries in England. The symptoms of 'sweating sickness' was detected in 1993 in the Southwest U.S.A. Here the disease was called hanta virus. There is likely that this is the same disease, but that the 'sweating sickness' has been dormant for couple of hundred years.[?, p. 317].

There have been done several mathematical studies on different diseases. HIV/AIDS is a field which has been studied by several scientists through the years. *Mathematical Modelling of the Transmission Dynamics of HIV Infection and AIDS: a Review* was published in 1988 by Valerie Isham [?]. This paper focuses on modelling transmission of infection in the context of AIDS epidemic. In 1999 Alan S. Perelson and Patrick W. Nelson published *Mathematical Analysis of HIV-1 Dynamics in Vivo* [?]. They studied the dynamics of HIV-1 pathogenesis to AIDS, where they looked at rapid dynamical processes that occur in short time scales, as hours, while AIDS occurs on a time scale of about 10 years. This affect the way that AIDS patients are treated with drugs. *Predicting the HIV/AIDS epidemic and measuring the effect of mobility in mainland China* by Xiao et al. [?] is another study done on HIV/AIDS where they look at the geographic variation in the severity of the epidemic in China. The models in this study are built on the basic ODE model shown in the next section. These three papers show some of the variation in both time and research when it comes to the study of diseases. The next section will describe a simple epidemic model and analyse the data this model can give.

Chapter 1

ODE models

This chapter will be split into two different parts. The part includes two sections, which will be based on the chapter *Dynamic of Infectious Diseases* from Mathematical Biology by J.D Murray [?]. *Epidemic models* will give a historic perspective on different epidemic diseases and their effect on the human population. Furthermore a basic ODE system will be shown and studied to see how this model can give information about the disease. The section will check if a disease is severe for the human population, and based on this called an epidemic disease.

The last part will be based on a scenario where the population faces a zombification, one of the most critical and devastating epidemic diseases that can occur. Here, the TV series *Walking Dead* will be used as reference, and the series will be tested against a model based on the SIR model explained in section *Simple Epidemic models*. There have been a couple of papers on this model earlier, and this part will be based on these models and try to adjust the system to the TV series.

1.1 Simple Epidemic Models

Most of the models shown here will have a constant population. The zombie model shown later will have a slight increase considering newborn, but this will be close to negligible. This may differ from the real world, where the population in different areas will vary with population flow. Reasons for doing this are, first of all to simplify the model and second to be able to model a closed system. How the population interacts is another assumption that has to be done. Here this is set to be similar for the whole area that is modeled. To simplify the population can be divided into three different groups.

- *Susceptible (S)*, who are humans that are healthy and at risk of becoming infected.

- *Infected* (I), who are humans who have the disease or are carriers of the disease. This group can infect the *Susceptible*.
- *Removed* (R), who are dead or recovering humans, often people that already have had the disease.

The natural order for a human is,

$$S \rightarrow I \rightarrow R. \quad (1.1)$$

This model is called *SIR* model, but the number of groups can be changed. *SI* only consists of the two first groups and a *SEIR* model has added an extra group *Exposed*, E , where the disease is latent. This can be used to model the incubation time.

The transmission of the infection and incubation period are elementary factors in the spread of a disease. These are reflected in the terms of the equations. Since this is a SIR model, the incubation time is negligible. The amount of people in each group can be seen as a function of time, expressed as $S(t)$, $I(t)$ and $R(t)$. The growth of I caused by *Susceptible*, can be viewed as a rate proportional to the number of *Infected* and *Susceptible* multiplied by a constant, rSI , where $r > 0$. This constant controls the efficiency of the transmission from S to I . This will appear as a reduction in the function $S(t)$. The rate of removal from *Infected* to *Removed* can be viewed as the number of *Infected* times a constant, aI , where $a > 0$ controls the time spent in the *Infected* group. The dynamic model will be,

$$\begin{aligned} \frac{dS}{dt} &= -rSI \\ \frac{dI}{dt} &= rSI - aI \\ \frac{dR}{dt} &= aI \end{aligned} \quad (1.2)$$

This model is called the Kermack-McKendrick(1927) model [?, p. 320]. It is considered that the groups are uniformly mixed and that there is equal probability of contact for all individuals. These assumptions will not be correct for all diseases, especially sexually transmitted diseases. The total number of the population will stay constant, since this is a closed system. This can be seen on the total change.

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \quad (1.3)$$

Therefore the total size of the population, N , will be constant.

$$S(t) + I(t) + R(t) = N \quad (1.4)$$

1.2 Threshold phenomenon

The threshold value is essential when studying an epidemic model. To cause an epidemic situation, the model needs to fulfill $I(t) > I_0$ for some $t > 0$, where I_0

describes the initial condition of the *Infected* group. The initial conditions can be given as,

$$S(0) = S_0 > 0, \quad I(0) = I_0 > 0, \quad R(0) = 0. \quad (1.5)$$

These initial conditions given in Eq.(1.5) combined with r and a controls the epidemic situation. These will affect the spread of the infection. From Eq.(1.2) the function for the *Infected* group at initial time is,

$$\left[\frac{dI}{dt} \right]_{t=0} = I_0(rS_0 - a) \quad (1.6)$$

The expression inside the brackets controls the change in I . The function will increase if $S_0 > \frac{a}{r}$, this will therefore be the threshold value for the function. The threshold value will be described by the variable ρ ,

$$\rho = \frac{a}{r} \quad (1.7)$$

This can be shown with some phase trajectories of the *Susceptible* and the *Infected* in Fig.(1.1).

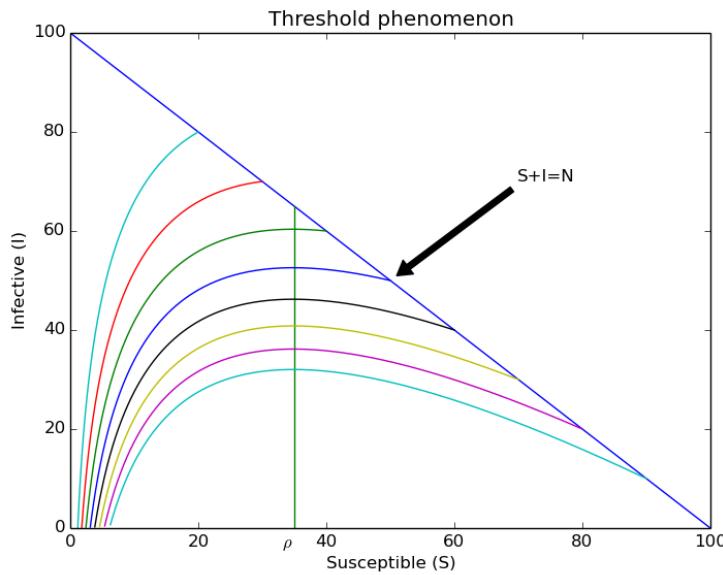


Figure 1.1: Simulations of the SIR model (1.2) with start positions along the blue line. I increases until S is equal to the threshold value ρ , which is set to 35. Then I is reduced to 0. In the simulations where $S_0 < \rho$, no epidemic situation is achieved.

The simulation shows that I is based on the relation between S and ρ . This can be described with a reproduction rate,

$$R_0 = \frac{rS_0}{a} \quad (1.8)$$

It will cause an epidemic reaction if $R_0 > 1$. This parameter is crucial in the understanding of the work with the disease. To prevent a dispersion, the value of R_0 has to be under 1. An effective way to get control is by global vaccination programs. Smallpox is an example on a disease that nearly has been eradicated around the world. This is due to a reduction of *Susceptible*. However there is always a small chance of side effects when using vaccination, and therefore some people choose to skip it. This is quite critical for the fight of total eradication. Not only is it a big risk for the specific person, but it also increases the number of *Susceptible*. An epidemic situation can quickly grow again if the reproduction rate reaches the threshold.

Some analytical studies can be done on the simple model(1.2).

$$\frac{dI}{dS} = -\frac{(rS - a)I}{rSI} = -1 + \frac{\rho}{S}, \quad \rho = \frac{a}{r}, (I \neq 0). \quad (1.9)$$

The singularities will all lie on the $I=0$ axis. This equation can be integrated and will then give phase plane trajectories in the (I,S) plane. This can be seen in Fig.(1.1).

$$I + S - \rho \ln S = \text{constant} = I_0 + S_0 - \rho \ln S_0 \quad (1.10)$$

All initial values satisfy $I_0 + S_0 = N$ since $R(0) = 0$. This will change when $t > 0$. If a disease appears it would be important to know the severity of the disease and the chance of developing to an epidemic disease. Therefore it is crucial to know the maximum value I_{\max} which occurs when $S = \rho$. At this point, $\frac{dI}{dt} = 0$. This can be found by using (1.10)

$$\begin{aligned} I + S - \rho \ln S &= I_0 + S_0 - \rho \ln S_0 \\ I_{\max} + \rho - \rho \ln \rho &= I_0 + S_0 - \rho \ln S_0 \\ I_{\max} &= -\rho + \rho \ln \rho + I_0 + S_0 - \rho \ln S_0 \\ I_{\max} &= N - \rho + \rho \ln \frac{\rho}{S_0} \end{aligned} \quad (1.11)$$

The different trajectories in Fig.(1.1) shows the difference between $S_0 > \rho$ and $S_0 < \rho$. An increasing of the *Infected* group will occur in the cases where S_0 is higher. While a decreasing will happen when S_0 is lower. An example can be shown. The ρ in the simulation in Fig.(1.1) is set to 35, while $N = 100$ for all trajectories. A calculation can be done on the lowest trajectory which has the

initial conditions $S_0 = 90$ and $I_0 = 10$

$$\begin{aligned} I_{\max} &= N - \rho + \rho \ln \frac{\rho}{S_0} \\ I_{\max} &= 100 - 35 + 35 \ln \frac{35}{90} \\ I_{\max} &= 31.94 \end{aligned}$$

This situation causes an epidemic situation since I_{\max} is much higher than the initial condition I_0 . The Fig.(1.1) shows that the trajectory of this function starts decreasing after this point. In the two upper trajectories where $S_0 < \rho$, the *Infected* group starts decreasing from the initial condition. The *Infected* group will decrease towards zero as $t \rightarrow \infty$.

The *Susceptible* will always have a decreasing solution since $\frac{dS}{dt} < 0$ when $S \neq 0$ and $I \neq 0$. From the ODE system(1.2) some integration can be done,

$$\begin{aligned} \frac{dS}{dR} &= -\frac{S}{\rho} \\ S &= S_0 e^{-R/\rho} \end{aligned} \tag{1.12}$$

Then the following term is true,

$$S = S_0 e^{-R/\rho} \geq S_0 e^{-N/\rho} > 0 \tag{1.13}$$

As $t \rightarrow \infty$, the total number of *Susceptible* will be in the range $0 < S(\infty) \leq N$. This range can be reduced even more by knowing that I will increase as long as $S > \rho$. The number of *Susceptible* will be in the range $0 < S(\infty) \leq \rho$. Since I will be zero when the time goes towards infinity, the *Removed* group can be described $R(\infty) = N - S(\infty)$. Now this can be insert in eq(1.12), which gives,

$$S(\infty) = S_0 \exp \left(-\frac{N - S(\infty)}{\rho} \right) \tag{1.14}$$

$S(\infty)$ can be found as the positive root in the transcendental equation. This can be used to find the total number of people who catch the disease.

$$I_{total} = I_0 + S_0 - S(\infty) \tag{1.15}$$

This analysis is based on the implication that the disease dies out because the *Infected* group goes towards zero, and not because of the lack of *Susceptible*. It will affect the number of *Susceptible* that can be infected. The reproduction rate (R_0), will vary with respect to different parameters as population density, incubation time and the length of the period of sickness. The two equations shown in Eq.(1.11) and Eq.(1.15) gives an understanding of the maximum and total number in the *Infected* group, but the methods demand the exact numbers of ρ, I_0, S_0 and S , which is quite hard to get in a real situation. The challenging thing is often to know how many infected there are at each time. The number

of *Removed* is often the easiest group to control.. This group is assisted with medical help. So to model a realistic situation, the number of *Removed* as a function of time dR/dt is a realistic model. Here the equations from (1.2) ,(1.4) and (1.12) can be used.

$$\begin{aligned}\frac{dR}{dt} &= aI \\ &= a(N - R - S) \\ &= a(N - R - S_0 e^{-R/\rho}), R(0) = 0,\end{aligned}\tag{1.16}$$

This solution demands several parameters as a , r , S_0 and N to solve this numerically. It is normal to adjust the parameters after the epidemic situation to get the best result as possible. However, if the epidemic is not to large, R/ρ will be quite small, at least under 1. Then another model from Kermack and McKendrick(1927) can be used. J.D Murray does a deeper study of this in his book [?, p. 324]. This chapter will look at an example from a boarding school in England and do some studies on the change in ρ .

1.3 English Boarding School 1978

The British medical journal published a report from a boarding school in England in 1978. One of the boys had brought with him a disease back to the school. Since this was a boarding school, they were totally isolated from others and had a closed system to model [?, p. 325]. The simulation can be seen in Fig.(1.2)

The parameter value ρ has a major impact on the result. The epidemic disease could turn out quite differently than in the situation in Fig.(1.2), caused by variations in a and r . Fig.(1.3) consists of some examples where ρ varies from 50 to 400.

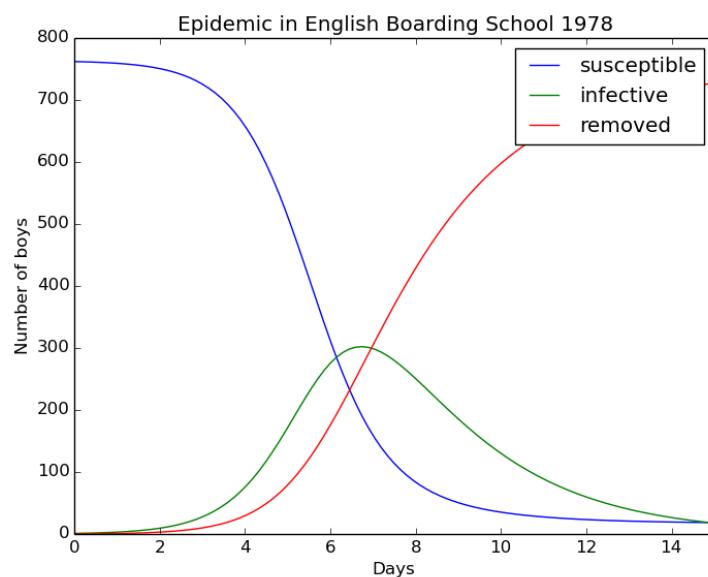


Figure 1.2: An English boarding school is modeled for 15 days with the following parameters: $N = 763$, $S_0 = 762$, $I_0 = 1$, $R_0 = 0$, $\rho = 202$ and $r = 2.18 \times 10^{-3}$. An increase in the *Infected* group can be seen since $S_0 > \rho$.

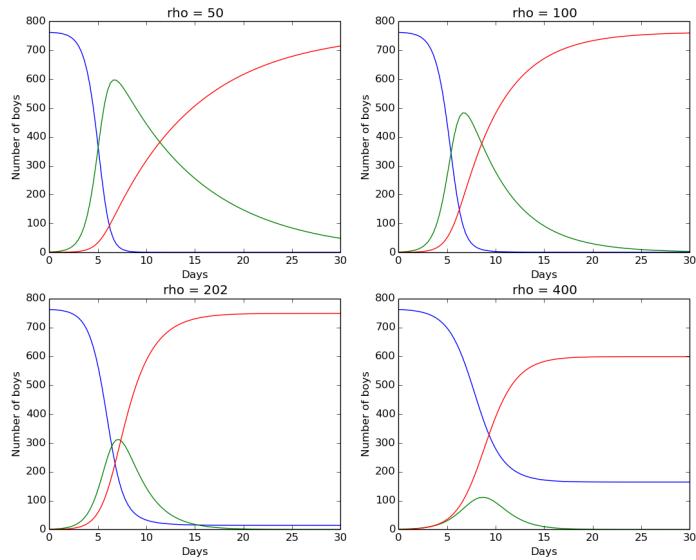


Figure 1.3: In the first plot where $\rho = 50$, the *Infected* group will increase until the number of *Susceptible* falls down to 50. This will result in a majority of infected students. In the last plot where $\rho = 400$, the total number of *Susceptible* stays around 170 students and will go towards a steady number as $I(\infty) = 0$.

1.4 Zombification

One of the worst epidemics that can affect the human population is a zombie attack. This will have a huge impact on the way humans live today. Several movies and series has illustrated this type of situation, but it is time that the scientists also take this threat seriously. There have been written a couple of papers about this. Munz et. al[?] used the SEIR model to simulate a possible upcoming zombiefication, where the latent group(E), is replaced with an *Infected* group(I) and the *Infected* group(I) is replaced with a *Zombie* group(Z). Here it is important to know that the *Infected* group in the *SIZR* is not the same as in the *SEIR* model. The following model was used,

$$\begin{aligned}\frac{dS}{dt} &= \Sigma - \beta SZ - \delta S \\ \frac{dI}{dt} &= \beta SZ - \varrho I - \delta I \\ \frac{dZ}{dt} &= \varrho I + \zeta R - \alpha SZ \\ \frac{dR}{dt} &= \delta S + \delta I + \alpha SZ - \zeta R\end{aligned}$$

This is a bit more complicated than the standard *SEIR* model. A presentation of the parameters;

- Σ describes the birthrate for new *Susceptible*. $\frac{dS}{dt}$ is now able to be positive. This is now not a closed system.
- βSZ describes the numbers of *Susceptible* that become infected , based on interactions between zombies and humans. Similar to the case for rSI in the *SIR* model.
- δ describes the number of natural deaths in the group. This is used in the *Susceptible* and the *Infected* group
- ϱI describes the probability for an infected human to wake up as a zombie.
- ζR controls the number of *Removed* that arises as *Zombie*.
- αSZ describes the number of zombies killed by humans in the zombie attacks.

This model was challenged by Langtangen, Mardal and Røtnes [?] now referred to as LMR, where they developed another model. They had three objections to the model from Munz et al. [?]. LMR argue that dead zombies cannot become functioning zombies again. Therefore ζ will be zero, if magic is not introduced. They let the parameters in the model change with time, according to different

phases. LMR argue that the behavior will change with time during a zombie attack. The parameters in the model is based on the movie *The Night of The Living Dead*. This is done to reproduce its scenarios and then predict how a zombie outbreak would appear. There is also added a function $\omega(t)$, which creates a massive attack from the humans. This is controlled by a time variable and give the *Susceptible* a chance to fight back. The system can be seen in eq(2.34):

$$\begin{aligned}\frac{dS}{dt} &= \Sigma - \beta SZ - \delta_S S \\ \frac{dI}{dt} &= \beta SZ - \varrho I - \delta_I I \\ \frac{dZ}{dt} &= \varrho I - (\alpha + \omega(t)) SZ + \zeta R \\ \frac{dR}{dt} &= \delta_S S + \delta_I I - \zeta R + (\alpha + \omega(t)) SZ\end{aligned}\tag{1.17}$$

The main change is the $\omega(t)$. This is a Gaussian curve and can be seen in eq(1.18).

$$\omega(t) = a \sum_{i=0}^m \exp \left(\frac{1}{2} \left(\frac{t - T_i}{\sigma} \right)^2 \right)\tag{1.18}$$

$\omega(t)$ controls the attacks from the *Susceptible*, which will be fired at predefined time steps. These are controlled by the three parameters.

- a here works as a similar parameter as α , but will only be activated when the *Susceptible* group is organized and ready to attack.
- T contains a list of numbers, which controls the time when the attacks will occur.
- σ controls the length of the attack.

This function will be modeled later when it is used in section 1.4.1

1.4.1 Parameters used in the model

The parameter values are essential factors when modelling a zombie attack. Data from the movie *The Night of The Living Dead* was used as basis for the parameters in the ODE system from LMR [?]. This thesis is based on a thorough study of the TV series Walking Dead [?]. The data will be based on the first five episodes and are constructed after having watched the episodes carefully. The three phases in a zombie attack will be based on the form used in the paper from LMR, but with an extended version in the *Counter attack phase*.

The initial phase. The disease is not yet known in this phase and humans try to save the sick ones by taking them to hospitals or getting some kind of treatment. Because of this ignorance related to the disease, the number in the *Infected* group is high. This phase is often quite short and humans soon start to realise that the risk of getting infected by saving others is really high. *Walking Dead* never shows anything from this phase, but the viewer sees the results when the main character sheriff Rick Grimes wakes up at the local hospital. What he sees is the major damage caused in the initial phase, while the society has moved to the hysterical phase.

To determine the values for each group in each phase, the length of Ricks coma is essential. There are several factors that give an indication of the time aspect. When Rick wakes up at the hospital, he has grown a smooth beard of about 1 cm. This would correspond with 1 month in average for a male of European origin. He also has some flowers that have dried out. These also give the impression that some weeks have gone by. The hospital is running on its emergency generator. This would probably not last for many days with a fully operational hospital, but the hospital is as well as shut down when Rick wakes up and can give the emergency generator a longer lifetime. Dr. Edwin Jenner gives the viewer some information in episode 5 where he tells the videolog that it was 63 days since the epidemic started spreading. By studying the first five episodes in detail, one gets an impression that the time aspect has not been in focus. Therefore the different phases are constructed from the information that has been given. Rick Grimes has probably been in a coma for a month and what he meets the first days will be the basis for the number in each group. The total amount of objects in the model will be based on the number of humans, dead and zombies seen in the first five episodes.

- The number of humans has been estimated to **62**. 20 living in the camp with Rick, 40 humans in the old nursing home and the father and son in episode 1.
- The number of dead is estimated to **200**. This is based on the amount of dead outside the hospital where Rick woke up.
- The number of zombies are assumed to be **360**. These are based on the 30 outside the house of Morgan Jones and his son Duane, 300 zombies in the city Atlanta and 30 zombies attacking the camp.

The total number will be **622** and the time aspect around a month, which means that these numbers are for the hysterical phase. Over the three first days when Rick is awake, 1 human and 20 zombies are killed in battles. This can be used to find the final number in the initial phase by calculating backwards. By going nine similar periods backwards, the number of killed zombies is 190. The same can be done for humans, which then results in 9 killed humans in this period. The final number for the initial phase can then be set to 71 humans, 540 zombies/infected and 20 dead. This is the same number as for the initial values for the hysterical

phase, since the phases are connected.

Another issue to discuss is the incubation time. Here there are two examples that can be used. The first transformation from human to zombie happens for the character Amy, who was bit in the arm by a zombie. The transformation happens in about 12 hours. The other example is character Jim who has a slower transformation. This lasts for about two days before the rest of the group leave him alongside the road on their way to CDC(Center for Disease Control). An estimate of the incubation time can be set to 24 hours based on these two transformations.

Now the ODE system (2.34)can be used to model the initial phase. The expected results are $S_0 = 621$ and $Z_0 = 1$ while the two other groups are set to zero. The value of β can be found with the expression $\beta\Delta tSZ$ from the first ODE equation. After three days about 90 percent of the humans are killed.

$$\begin{aligned}\beta\Delta tSZ &= 0.9S \\ 3\beta &= 0.9 \\ \beta &= 0.3\end{aligned}\tag{1.19}$$

The probability of a human being infected will be set to $\beta = 0.3$. The natural death rate is set to $\delta_S = 2.2 * 10^{-5}$ based on numbers from CDC [?]. It is quite hard to find similar realistic data for infected humans, so $\delta_I = \delta_S$. The number of births is set to $\Sigma = 3.45 * 10^{-5}$. This is based on data from CDC from 2012 [?]. Since this is data for the initial phase, zombies are seen as infeteced humans that can be saved. Therefore $\alpha = 0$. And the two last parameters are also zero, $a = \zeta = 0$. The initial phase is modeled in Fig.(1.4):

The result from Fig.(1.4) shows that the human population is eradicated in about a half day. This is not the case, and some adjustments need to be done. There are three parameters that are interesting to study. The first one is β , which describes how many humans that get infected in a human-zombie collision. The second one is ϱ . This parameter controls the incubation time. The last parameter that can affect the number in each group is α . This describes the number of zombies killed in a human-zombie collision. These variables are plotted separately and combined in Fig.(1.5). The idea here is to produce results that fulfill the final number for the groups *Susceptible* and *Removed* , which is 71 and 20. The blue dot in each plot describes this value. A rough estimate has been done for each parameter before using it. This is why they all lie in different regions than the parameter value in Fig.(1.4)

By choosing $\beta = 0.01155$, $\varrho = 1.37$ and $\alpha = 0.00044$, the following plot can be seen in Fig.(1.6):

It is possible to argue for the changes done in Fig.(1.6). Increasing ϱ to 1.37 reduces the incubation time. Now the average time will be about 17.5 hours, which is realistic. The probability β is sensitive and has a major effect only by small variations. This is due to the term that it is a part of $\Delta tSZ\beta$. A couple of examples demonstrate this. One hour can be estimated by setting

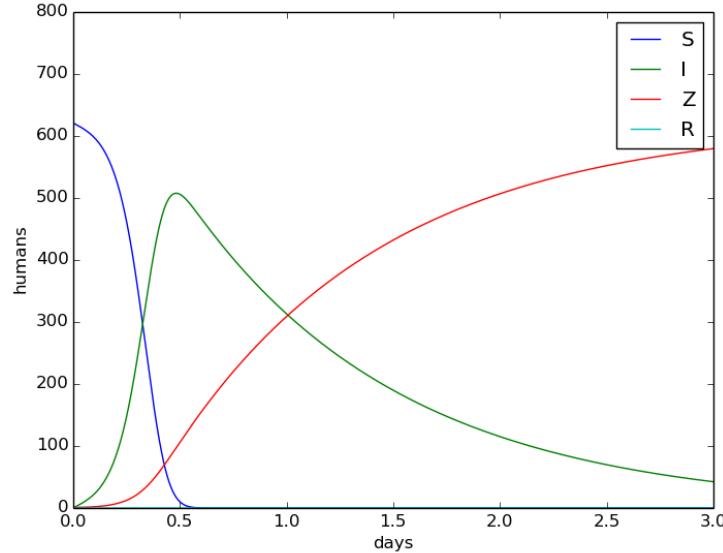


Figure 1.4: Initial phase for *Walking Dead*. $\beta = 0.3$, $\varrho = 1$ and $\alpha = 0$ leads to eradication.

$\Delta t = 1/24$. When using the initial values for the groups S and Z and $\beta = 0.01155$ from Fig.(1.6), the number in the *Infected* group in the first hour will be $(1/24) * 721 * 1 * 0.01155 = 0.34$. About one-third of a human in the first hour seems as a slow and not very aggressive disease. However when the number of *Zombies* slowly increases, the number of infected will be affected. By looking at the hour when the values are equal between humans and zombies, about 200 in each group, the number of infected will be 19.25 per hour. This result in about 10 percent of the humans. By changing β to the value from Fig.(1.4), the number of infected will be 500 per hour and it is quite easy to see that this will lead to eradication in a short amount of time. The last parameter α controls the number of zombies that dies in collisions between zombies and humans. While humans still think that the infected can be saved, it is still a chance that the result from a collision can end with a zombie kill. These results can therefore be seen as realistic values.

The hysterical phase. Now the humans start to avoid the infected and some try to fight them. The humans often gather in groups and try to find safe spots away from the zombies. Important supplies as weapons and food are their main priorities. Barricades are built and the guarding is strict. When Rick Grimes wakes up, the hospital is abandoned and the halls are filled up with dead people. Quite fast he understand that he needs to reach safety. After a couple of days

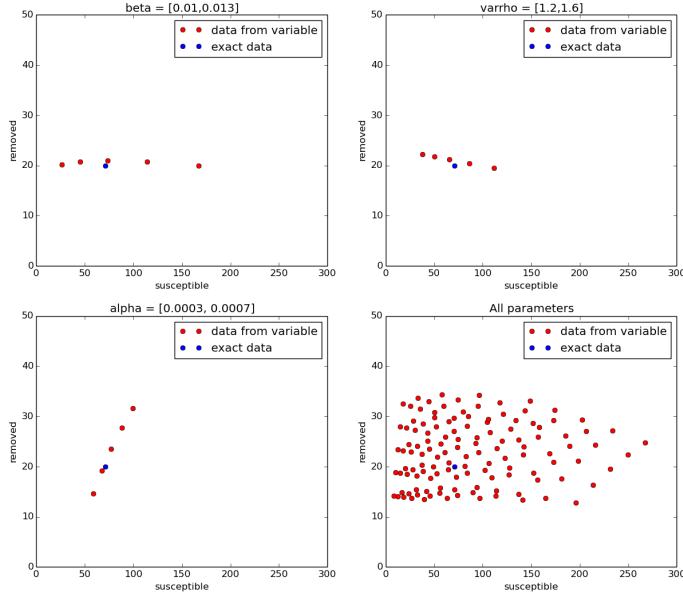


Figure 1.5: These plots give an important knowledge in the effect of varying the parameters. β and ϱ mainly affect the number of *Susceptible* while α affect them both.

he ends up in a camp outside Atlanta city. A couple of elementary changes has happened with the interaction between humans and infected/zombies. In the initial phase, the humans tried to help the infected humans. This resulted in a high percent of infected. Now they understand this risk and keep distance to those who are infected. This will give β a lower value. The morality for a zombie kill has dramatically changed. While this was seen as no option in the initial phase, this is now okay. The humans have started to treat zombies and infected as enemies instead of sick allies. This results in a higher death rate among the zombies, which is described by α .

The hysterical model can be constructed based on the data found in the initial phase.

hysterical phase	initial values	final values
S	71.3	62
I	230.8	-
Z	298.9	360
R	21	200

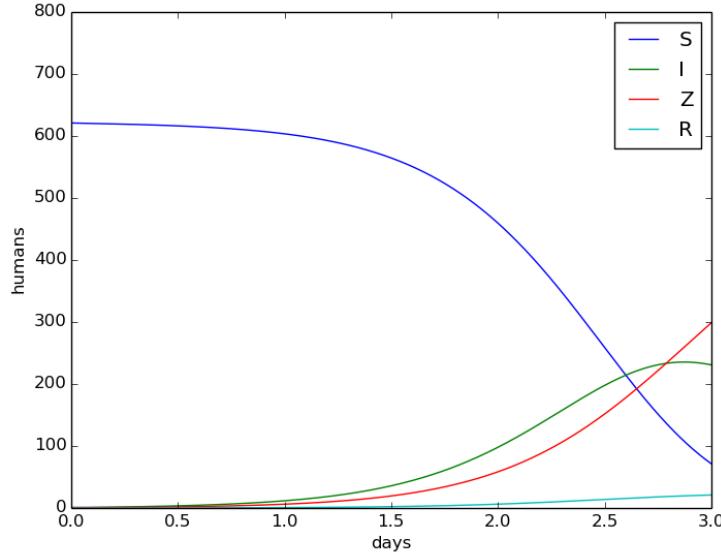


Figure 1.6: The final values are $S_n = 71.3$, $I_n = 230.8$, $Z_n = 298.9$ and $R_n = 21$, which is quite close to the result from the movie.

Here, the infected and zombies are counted as one group for the final values, since it is difficult to separate these groups in the series. The time aspect will be modeled for 30 days, which results in a ten times longer simulation. Since the final results are known here, a similar adjustment of the parameters can be done. The range of the parameter values have been found by some test simulations. In Fig.(1.7) the parameters of β, ϱ and α have been simulated with different values.

Fig.(1.7) gives insight in how variations in parameters affect the final result. By decreasing β , it will essentially increase the number of *Susceptible* that will survive, but it will also increase the number of deaths. This may at a first glance seem quite strange. Should not the number of deaths decrease when the number of surviving humans increase? This can be explained with the idea that was shown for β in the initial phase. Since βSZ gets smaller when β gets smaller, the combination of SZ will stay higher for a longer time. This will again affect αSZ , which regulates the number of killed zombies. The larger this combination is, the more zombies will die.

By increasing α , the number of *Removed* also increase. However similar to the increasing of the *Removed* it also has a slight increase on the number of *Susceptible*. Here the argument for β above can be reversed. Since a higher α leads to a higher death rate among the zombies, the combination SZ will be smaller, which makes βSZ smaller.

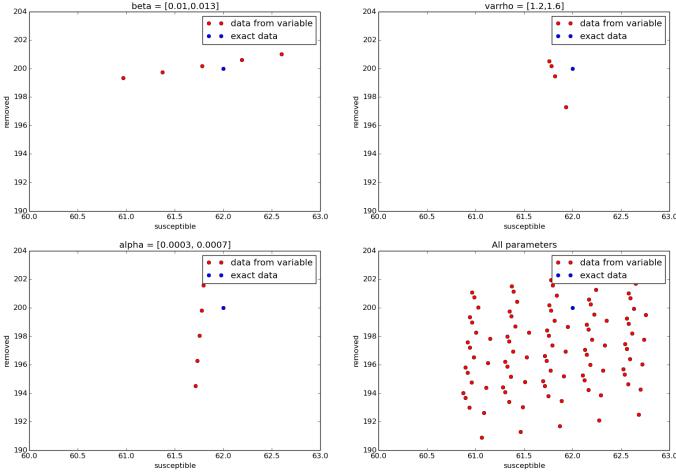


Figure 1.7: Variation in three different parameters in the hysterical phase. The last subplot consist of the combination of all. $\beta = [1 * 10^{-5}, 1.2 * 10^{-5}]$, $\varrho = [1, 2]$ and $\alpha = [2 * 10^{-4}, 2.2 * 10^{-4}]$

The last parameter, ϱ , has nearly no effect. The red dots show a small decreasing of *Removed* and increasing of *Susceptible* when ϱ is increased. This parameter varies most, but has least impact. This can be explained with the long time aspect and the low number of infected compared to the number of zombies. Since the number of zombies is much higher, the transformation length from infected to zombie is almost negligible.

Fig.(1.8) fulfills the result that was predicted based on the series. These final numbers correspond with the number in each group when Rick woke up at the hospital. The plot shows that the number of zombies increases quickly and reaches its maximum value after a couple of days in this phase, similar to the number of infected that dramatically decreased. Here the humans have been able to stabilize. Since the clashes between humans and zombies are dramatically decreased, nearly no humans get infected. And in the cases where humans have to face zombies, the killing rate has increased. The increase of *Removed* is close to proportional to the decrease of *Zombies*, which means that it is mostly zombies that die.

The counter attack. This counter attack is more complicated to model, since this phase appears simultaneously as the hysterical phase in *Walking Dead*. The group of humans are trying to avoid the zombies, but when the zombies get too close, the humans need to fight back. These situations are caused by a high density of zombies in some areas, which force the zombies to spread. In

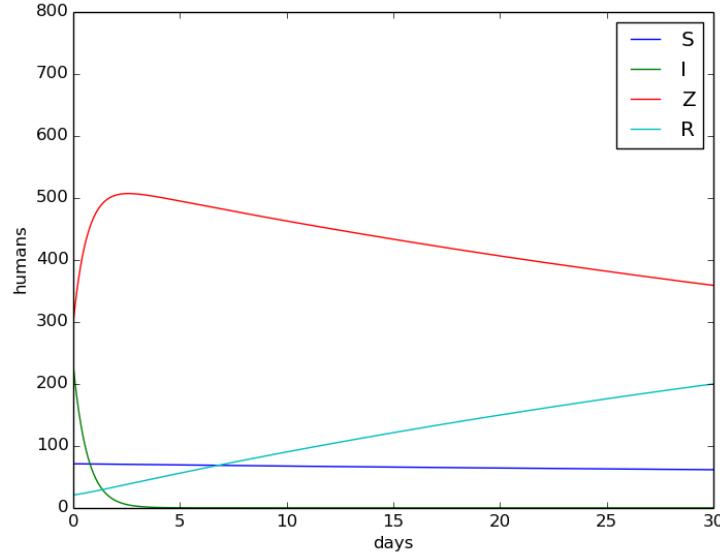


Figure 1.8: Hysterical phase with parameter values $\beta = 0.000011$, $\varrho = 1.5$ and $\alpha = 0.000208$.

Walking dead the counter attack appears when a group of 30 zombies reach the camp. This triggers a fight where all the zombies are killed and 4 of the humans get bitten. This shows that a counter attack from the humans causes a lot of damage. The time aspect is set to 6 hours.

Now the function $\omega(t)$ will be used. This can be seen in Fig.(1.9):

To get some start values, $SZ\omega(t) = 30$ can be used. Where $\omega(t)$ is the area under the function. By inserting the final values from hysterical phase for S and Z , the area shall be $\omega(t) = 1.34 * 10^{-3}$. This result can be reproduced by using $a = 0.00103$ and $\sigma = 0.005$ in $\omega(t)$. The counter attack is set to appear during the last part of the day $[0.75,1]$. The value of T is then set to $T = [0.875]$.

This simulation in Fig.(1.10) results in some deaths, but the total number should be higher. Another problem is that no humans died during this battle. The ODE model (2.34) is based on *The Night of the Living Dead*, where the amount of humans who are killed is close to zero. This is not the case in *Walking dead*. Therefore the risk of dying is higher for human during a counter attack. This is solved by adding $\mu\omega(t)SZ$, where μ is the risk for a human getting infected compared to a zombie kill during this attack. The model (2.34) can then be expanded to system(2.27),

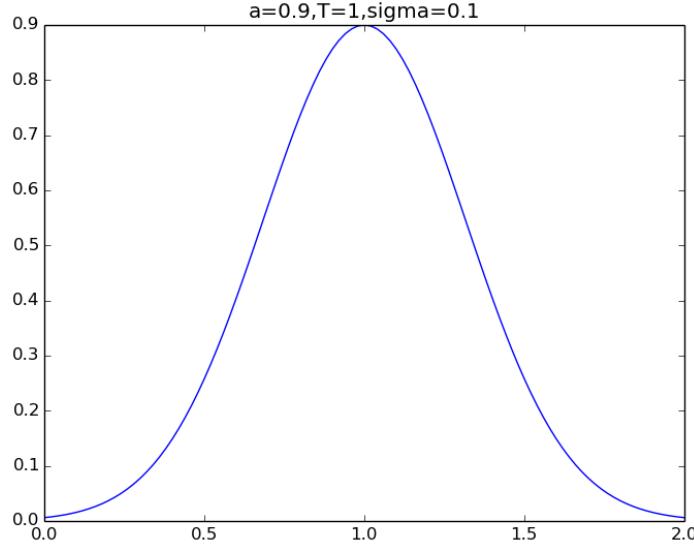


Figure 1.9: $\omega(t)$ is a Gaussian function where a controls the maximum value, T controls the time for maximum value and σ controls the length of the attack.

$$\begin{aligned}\frac{dS}{dt} &= \Sigma - (\beta + \mu\omega(t))SZ - \delta_S S \\ \frac{dI}{dt} &= (\beta + \mu\omega(t))SZ - \varrho I - \delta_I I \\ \frac{dZ}{dt} &= \varrho I - (\alpha + \omega(t))SZ + \zeta R \\ \frac{dR}{dt} &= \delta_S S + \delta_I I - \zeta R + (\alpha + \omega(t))SZ\end{aligned}$$

Fig.(1.11) is modeled with the initial values given when Rich woke up, explained in the initial phase. The result after this day is that the humans are reduced to 58 humans. The number in the *Infected* group is increased to 2.47, which can be explained with the two characters in the series, Amy and Jim. The number of *Removed* is increased to 231, and is a combination of killed zombies and humans who are attacked. By modelling this for another day, the *Removed* group will increase with a couple of new deaths.

It would be interesting to check what would happen if this counter attack was repeated over time. Who would survive? An attack every other day will give the following result shown in Fig.(1.12)

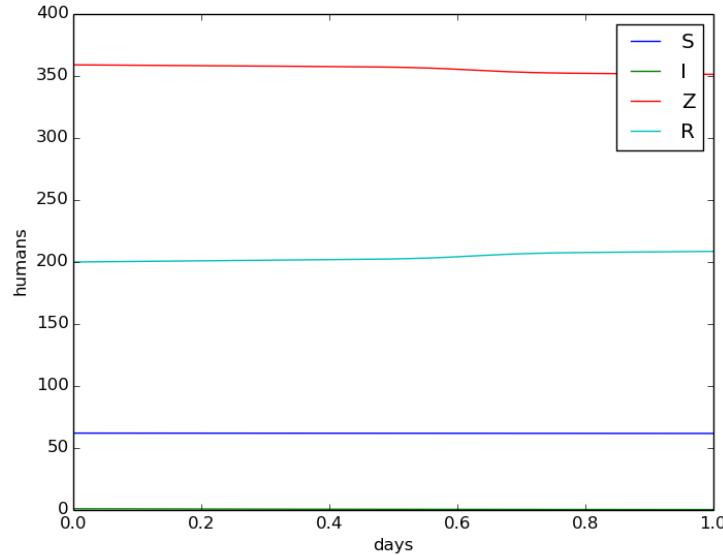


Figure 1.10: The counter attack. 8-9 zombies are killed and all humans survive

After 200 days there would be about 15 humans and 12 zombies left. Then the humans would be able to survive since they are more efficient in battles. There are of course several things to discuss. How will the number of battles effect the humans and zombies. Will they be tired or more efficient? What about weapons? What would happen if the group of zombies was much larger? Due to different variations in behavior of humans and zombies, one can never predict all situations accurately.

The three phases in Walking Dead. By adding these three phases together, the final result after the attack should be possible to match. The simulation here will be done with the parameters used in the earlier sections. This can lead to a small error since the result of the final number in each phase is given with decimals and the initial values are based on assumptions and round off numbers. The different parameter values are listed in Tab.(2.3).

parameter	Initial phase	hysterical phase	counter attack
β	0.01155	0.000011	0.00011
ϱ	1.37	1.5	1.5
α	0.00044	0.000208	0.000208
a	0	0	0.0073
σ	0	0	0.005
μ	0	0	0.14

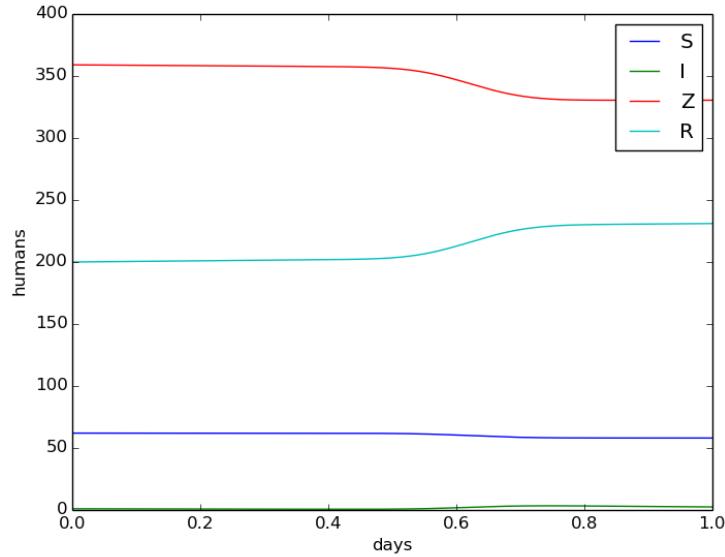


Figure 1.11: Counter attack with the new model. Parameter values are set to $a = 0.0073$ and $\mu = 0.14$.

The simulation is run for 34 days. The three first days are in the initial phase, the resisting days are in the hysterical phase. The counter attack is released on day 33 and lasts for about 6 hours. The plot is shown in Fig.(1.13)

Fig.(1.13) clearly shows that the change in parameter values affect the different phases. The different values are shown in the Tab.(1.4.1), where the values are given at the initial time. The last column consist of the final values after 34 days.

	Initial phase	hysterical phase	counter attack	final values
S_0	621	71	62	58
I_0	0	231	0	1
Z_0	1	299	359	332
R_0	0	21	202	231

Considering the uncertainty of the parameters, this simulation gives a result close to the expected result.

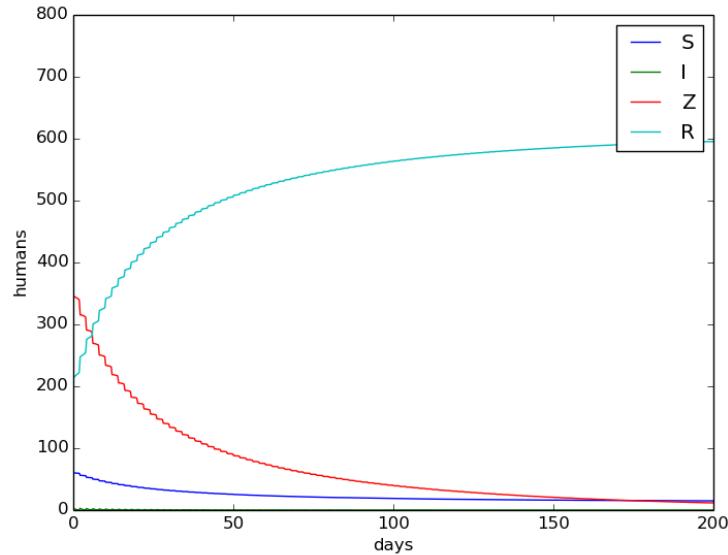


Figure 1.12: 100 counter attacks during 200 days will result in a higher population of humans.

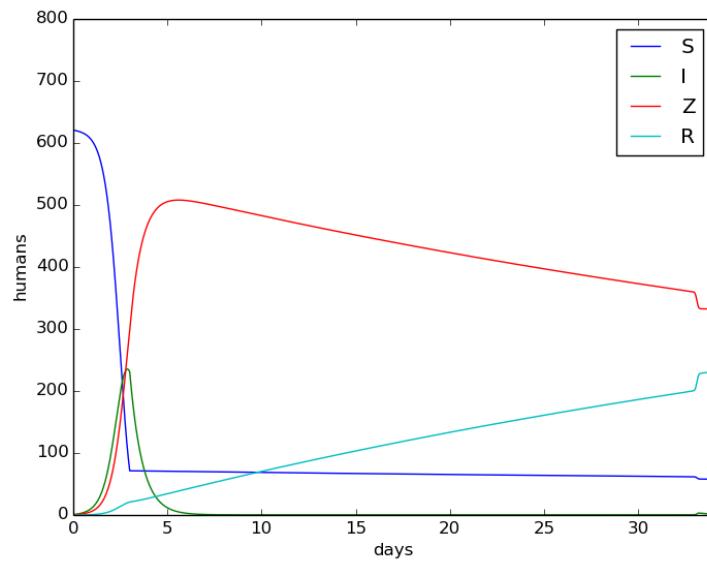


Figure 1.13: Walking Dead simulated after 5 episodes. Based on the three different phases.

1.5 Discussion

This model will be used as the basic model, which the two other models presented later will be compared to. This model is straight forward to calculate and is easy to use. The model is good if the goal is to find the total amount of each group measured over an area. It demands that the parameters are given. But the model gives no information about the spatial spread of the disease. This will therefore be useless in describing how a disease can spread abroad countries and borderlines, or the spatial differences in the chance of being infected. The two next chapters will introduce more complicated models which will be compared to this ODE-system.

Chapter 2

PDE models

This chapter will introduce a spatial model for epidemic diseases. The ODE system from the previously chapter can be expanded with a term for geographic spread of the disease. The first section *Simple system for spatial spread* will be based on the simple SIR model presented in previously chapter, and the chapter *Geographic spread and Control of epidemics* by Murray [?]. The parameters from the *English Boarding School* in the previously chapter will be used for the model and the results will be compared. The position of the infected student will be studied to see if it affects the number in the *Infected* group. The last section, *Zombiefication*, will study and expand the system from Langtangen, Mardal and Røtnes [?]. The results and parameter values used to calculate *Walking Dead* will be compared with the previously ODE system and variations of this PDE.

2.1 Simple system for spatial spread

A spatial variable, \mathbf{x} , will be introduced to the model. This results in both temporal and spatial variations. The difference from a standard ODE system will be the diffusion part added to each equation. The system can be seen in Eq.(2.1).

$$\begin{aligned}\frac{\partial S}{\partial t} &= -rIS + D\nabla^2 S \\ \frac{\partial I}{\partial t} &= rIS - aI + D\nabla^2 I \\ \frac{\partial R}{\partial t} &= aI + D\nabla^2 R\end{aligned}\tag{2.1}$$

Here S describes the *Susceptible* group, I describes the *Infected* group and R describes the *Removed* group. The following conditions are set for the boundary and initial values:

$$\begin{aligned}u_x(0, t) &= u_x(X, t) = 0, \quad u = S, I, R \\ u(x, 0) &= f_u(x), \quad u = S, I, R\end{aligned}\tag{2.2}$$

This results in Neumann conditions at the boundary. The following implementation can be used at the boundary

$$\begin{aligned} \frac{u_{-1}^n - u_1^n}{2\Delta x} &= 0 \\ u_{-1}^n &= u_1^n \end{aligned} \quad (2.3)$$

This is solved by adding an extra point on each side, called ghost points. The values in these ghost points are updated every time step with values from u_1^n and u_{X-1}^n . All three groups, S, I, R in Eq.(2.1) have the same diffusion coefficient, D . This gives the three groups the same diffusion speed. This can vary between systems. Later in the chapter, in section *Zombiefication*, different diffusion terms are given to the groups. The two probabilities rIS and aI will work in the same way as in the ODE system. Since this model takes the position into account, a group of *Infected* that move into a uniform population with *Susceptible* can be model. The group of *Susceptible* has the density S_0 . A simulation can show the geotemporal spread of the disease. The problem can first be considered as one-dimensional. The system can be nondimensionalised by writing

$$\begin{aligned} I^* &= \frac{I}{S_0}, & I^* &= \frac{I}{S_0}, & R^* &= \frac{R}{S_0}, \\ x^* &= \left(\frac{rS_0}{D}\right)^{1/2} x, & t^* &= rS_0 t, & \lambda &= \frac{a}{rS_0}, \end{aligned} \quad (2.4)$$

S_0 is used as a representative population. Now Eq.(2.1) can be expressed as in Eq.(2.5). The asterisks have been dropped to make it easier to read.

$$\begin{aligned} \frac{\partial S}{\partial t} &= -IS + \frac{\partial^2 S}{\partial x^2}, \\ \frac{\partial I}{\partial t} &= IS - \lambda I + \frac{\partial^2 I}{\partial x^2}, \\ \frac{\partial R}{\partial t} &= \lambda I + \frac{\partial^2 R}{\partial x^2}, \end{aligned} \quad (2.5)$$

The three parameters r , a and D have been replaced by λ . The *reproduction rate* that was presented for the ODE model can be seen as $1/\lambda$. This has equivalent meanings. $1/\lambda$ can be seen as the number of secondary infections produced by one primary infected. It can also be used to measure two different time scales. The first one, $1/(rS_0)$, measures the contagious time of the disease. The second one looks at the life expectancy for an infected. This can be described as $1/a$ [?].

2.1.1 Travelling wave 1D

In this case the travelling wave describes how a group of *Infected* travels through a geographic area of *Susceptible*. This will be shown by sending a pulse from the *Infected* group into a group of *Susceptible*. A travelling wave solution can be described as follows,

$$I(x, t) = I(z), \quad S(x, t) = S(z), \quad R(x, t) = R(z), \quad z = x - ct, \quad (2.6)$$

The value c describes the wave speed. This represents a wave of constant shape that travels in the positive x -direction. Eq.(2.6) can be inserted into Eq.(2.5). This result in the ordinary system Eq.(2.7)

$$\begin{aligned} S'' + cS' - IS &= 0, \\ I'' + cI' + I(S - \lambda) &= 0 \\ R'' + cR + I\lambda &= 0 \end{aligned} \quad (2.7)$$

This makes an eigenvalue problem. The value of λ needs to stay in a range where $c > 0$ is fulfilled. The values S , I and R have to stay nonnegative. This leads to

$$\begin{aligned} 0 \leq S(-\infty) < S(\infty) = 1 \\ I(-\infty) = I(\infty) = 0, \\ 1 \geq R(-\infty) \geq R(\infty) = 0 \end{aligned} \quad (2.8)$$

An epidemic wave can be seen in Fig.(2.1). The value of λ is set to 0.5. The initial value of *Susceptible* is 1 for the area and the *Removed* is set to 0. The *Infected* group has a Gauss curve around 0 at initial time. In the four subplots in Fig.(2.1), the epidemic wave travels towards the other side. The value z , which is defined in Eq.(2.6), is used to plot the travelling wave measured at a specific point, in this case $x = 15$. This travelling wave is shown in figure(2.2).

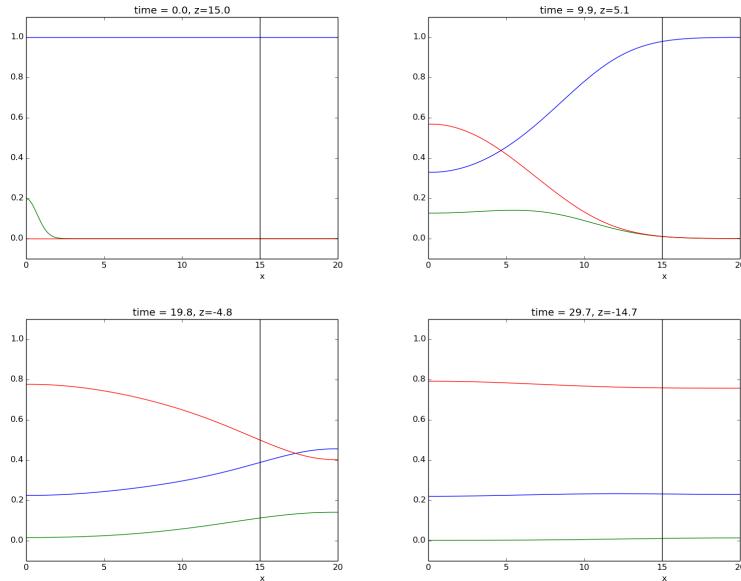


Figure 2.1: The system (2.7). A gaussian curve with height 0.2 placed on the left side. This causes an epidemic wave controlled by the parameter $\lambda = 0.5$. The size is measured at point $x = 15$ and can be seen in figure (2.2).

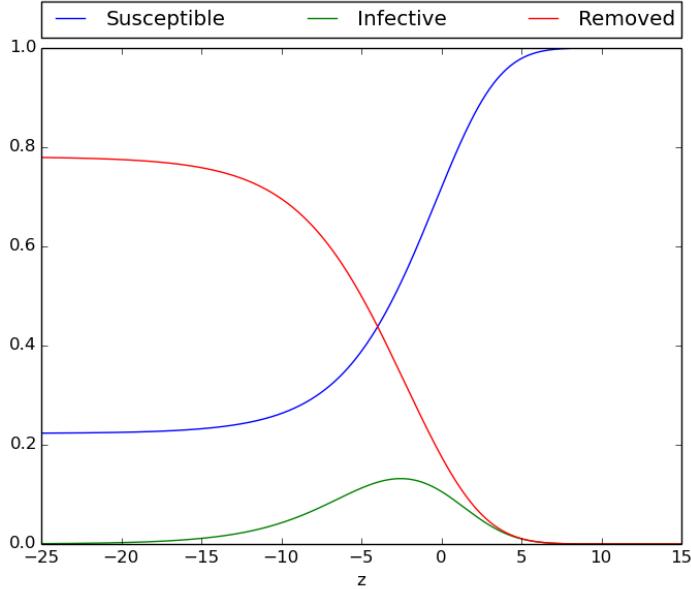


Figure 2.2: This shows the travelling wave measures at $x = 15$ in figure(2.1)

The *Infective* group in Eq.(2.7) can be linearised when $z \rightarrow \infty$. This leads to $S \rightarrow 1$ and $I \rightarrow 0$. The result then become

$$I'' + cI' + I(S - \lambda) \approx 0 \quad (2.9)$$

This can be found by

$$I(z) \propto \exp \left[(-c \pm c^2 - 4(1 - \lambda)^{1/2})z/2 \right] \quad (2.10)$$

Since it is required that $I(z) \rightarrow 0$ and $I(z) > 0$, oscillations around 0 must be prevented. If a travelling wave exists, $\lambda < 1$. If $\lambda > 1$, no travelling wave will exist. Then the disease will die out. The terms defined in Eq.(2.4) will give the threshold conditions,

$$\lambda = \frac{a}{rS_0} < 1 \quad (2.12)$$

This is the same value that was given for the ODE model in the previously chapter.

2.1.2 Verifying the solution

To verify the implementation of the solution, a couple of tests can be done one the system in Eq.(2.5). The system will be tested with a constant solution and a manufactured solution.

2.1.3 Constant solution

A constant solution uses preproduced constant values for the concentrations S , I and R . These can be replaced by $S = C_s, I = C_i, R = C_r$. The value of C_i can only be 0 in Eq.(2.5). This results in a poor test where several bugs can escape. The system can be expanded by adding a term βR to the *Susceptible* group and subtracting the same term from the *Removed* group. Then all three values can be tested. The system will then look like this:

$$\begin{aligned}\frac{\partial S}{\partial t} &= -IS + \beta R + \frac{\partial^2 S}{\partial x^2}, \\ \frac{\partial I}{\partial t} &= IS - \lambda I + \frac{\partial^2 I}{\partial x^2}, \\ \frac{\partial R}{\partial t} &= \lambda I - \beta R + \frac{\partial^2 R}{\partial x^2},\end{aligned}\tag{2.13}$$

By deriving Eq.(2.15), the following system Eq.(2.14) has to be solved

$$\begin{aligned}C_i C_s &= \beta C_r \\ C_i C_s &= \lambda C_i \\ \lambda C_i &= -\beta C_r\end{aligned}\tag{2.14}$$

The values β and λ are based on the constants C_s, C_i, C_r , which can be chosen freely. Here they are set to $C_s = 1.2, C_i = 0.8, C_r = 0.6$. This results in $\lambda = C_s = 1.2$ and $\beta = \frac{C_s C_i}{C_r} = 1.6$. A test is made in python and can be seen below.

```
def test_constant_solution():
    """
    Test problem where u=u_const is the exact solution, to be
    reproduced (to machine precision) by any relevant method.
    """
    def exact_solution(t):
        return C_s,C_i,C_r

    def lam(t,x):
        return C_s

    def beta(t,x):
        return (C_s*C_i)/float(C_r)

    #Constant values
    C_s = 1.2
    C_i = 0.8
    C_r = 0.6

    #lam = C_s
    #beta = (lam*C_i)/float(C_r)

    T = 2; Nt = 200
    X = 20; Nx = 40
    S_1 = np.ones(Nx+3)*C_s
    I_1 = np.ones(Nx+3)*C_i
    R_1 = np.ones(Nx+3)*C_r
```

```
t,x,S,I,R = simple_PDE(T,Nx,Nt,X,lambda,beta,S_1,I_1,R_1)
S_e,I_e,R_e = exact_solution(t)
difference = abs(S_e - S).max() # max deviation
tol = 1E-14
assert difference < tol

difference = abs(I_e - I).max() # max deviation
tol = 1E-14
assert difference < tol

difference = abs(R_e - R).max() # max deviation
tol = 1E-14
assert difference < tol
```

The test was run with no error, and the three constant values were produced correctly. This test is not good enough to qualify the program alone, however an error here would result in a large error in the program.

2.1.4 Manufactured solution

By constructing a function to each equation in the Eq.(2.5), a manufactured solution can be created. Here S, I and R are pre produced. The system will be

$$\begin{aligned}\frac{\partial S}{\partial t} &= -IS + \frac{\partial^2 S}{\partial x^2} + f(x, t), \\ \frac{\partial I}{\partial t} &= IS - \lambda I + \frac{\partial^2 I}{\partial x^2} + g(x, t), \\ \frac{\partial R}{\partial t} &= \lambda I + \frac{\partial^2 R}{\partial x^2} + h(x, t),\end{aligned}\tag{2.15}$$

where f, g and h are functions to achieve the expected results for S, I and R . In this case the functions will be:

$$\begin{aligned}f(x, t) &= \frac{\partial S}{\partial t} + IS - \frac{\partial^2 S}{\partial x^2} \\ g(x, t) &= \frac{\partial I}{\partial t} - IS + \lambda I - \frac{\partial^2 I}{\partial x^2} \\ h(x, t) &= \frac{\partial R}{\partial t} - \lambda I - \frac{\partial^2 R}{\partial x^2},\end{aligned}\tag{2.16}$$

When choosing the expected function for the groups, it is important that the boundary conditions from Eq.(2.2) is fulfilled.

$$u_x(0, t) = u_x(X, t) = 0\tag{2.17}$$

The quantities have been set to:

$$\begin{aligned}S(x, t) &= \cos\left(\frac{\pi}{X}x\right)t \\ I(x, t) &= \cos\left(\frac{\pi}{X}x\right)t \\ R(x, t) &= \cos\left(\frac{\pi}{X}x\right)t\end{aligned}\tag{2.18}$$

Now `sympy` can be used to do the calculations for the three functions f , g and h . The program can be seen in the Appendix. This results in the following equations seen in Eq.(2.19)

$$\begin{aligned} f(x, t) &= (t^2 \cos(\frac{\pi}{X}x) + \frac{\pi^2}{X}t + 1) \cos(\frac{\pi}{X}x) \\ g(x, t) &= (\lambda t - t^2 \cos(\frac{\pi}{X}x) + \frac{\pi^2}{X}t + 1) \cos(\frac{\pi}{X}x) \\ h(x, t) &= (-\lambda t + \frac{\pi^2}{X}t + 1) \cos(\frac{\pi}{X}x) \end{aligned} \quad (2.19)$$

A similar test made for the constant solution can be used here. While the constant test expected a difference on machine precision, this is not the case here. In this test, an expected convergence rate can be measured. The following manufactured test will be

```
def test_manufactured_solution(T,Nt,X,Nx):
    """
    Test problem where u=c*t+I is the exact solution, to be
    reproduced (to machine precision) by any relevant method.
    """

    def exact_solution_S(t,x):
        return np.cos(np.pi*x)*t

    def exact_solution_I(t,x):
        return np.cos(np.pi*x)*t

    def exact_solution_R(t,x):
        return np.cos(np.pi*x)*t

    def beta(t,x):
        return exact_solution_S(t,x)*exact_solution_I(t,x)/exact_solution_R(t,x)
    lam = 1
    def f(t,x):
        return (t**2*np.cos(np.pi*x) + np.pi**2*t + 1)*np.cos(np.pi*x)

    def g(t,x):
        return (lam*t - t**2*np.cos(np.pi*x) + np.pi**2*t + 1)*np.cos(np.pi*x)

    def h(t,x):
        return (-lam*t + np.pi**2*t + 1)*np.cos(np.pi*x)

    dx = X/float(Nx)
    dt = T/float(Nt)
    S_1 = exact_solution_S(0,np.linspace(0-dx,X+dx,Nx+3))
    I_1 = exact_solution_I(0,np.linspace(0-dx,X+dx,Nx+3))
    R_1 = exact_solution_R(0,np.linspace(0-dx,X+dx,Nx+3))

    t,x,S,I,R = simple_PDE(T,Nx,Nt,X,lambda,beta,S_1,I_1,R_1,f,g,h)
    S_e = exact_solution_S(t[-1],x)
    I_e = exact_solution_I(t[-1],x)
    R_e = exact_solution_R(t[-1],x)

    difference_S = abs(S_e - S).max() # max deviation

    #for i in range(4):
```

```

#      print "n",i,"S_e",exact_solution_S(t[i],x)

#print "S",S
#t_tot = np.sum(t[:-1])
#print "t_tot",t_tot
#difference_exp = t_tot*dt*np.cos(x*np.pi)*((2*(np.cos(np.pi*dx)-1))/dx**2+np.pi**2)
#print "diff_exp", (abs(difference_exp)).max()
#print "diff",difference_S
#tol = 1E-14
#assert difference < tol

difference_I = abs(I_e - I).max() # max deviation
#print "diff",difference_I
#tol = 1E-14
#assert difference < tol

difference_R = abs(R_e - R).max() # max deviation
#print "diff",difference_R
#tol = 1E-14
#assert difference < tol
return difference_S,difference_I,difference_R

```

2.1.5 Convergence rate

The program can be controlled by checking the convergence rate. The error term for this equation can be described as

$$\epsilon = C_x \Delta x^2 + C_t \Delta t \quad (2.20)$$

With Eq.(2.20), the expected convergence rate can be found for both Δx and Δt . To be able to separate the Δs , the other value has to be close to eliminated. To study the value Δx , the following term has to be fulfilled: $\Delta t \ll \Delta x$. This will lead to $C_t \Delta t \approx 0$, and the error term for Δx can be found. The opposite thing can be done for Δt . A table for the error is produced for different values for $\Delta t = 0.05$ and $\Delta x = 0.1$.

	Δx	$\frac{\Delta x}{2}$	$\frac{\Delta x}{4}$	$\frac{\Delta x}{8}$
Δt	9.8E-3	-	-	-
$\frac{\Delta t}{4}$	9.9E-3	2.5E-3	-	-
$\frac{\Delta t}{8}$	9.9E-3	2.5E-3	6.1E-4	-
$\frac{\Delta t}{16}$	9.9E-3	2.5E-3	6.1E-4	1.5E-4

2.1.6 The spatial error

The Tab.(2.1.5) gives information about the error when Δt and Δx are reduced. By studying the row where $\Delta t/16$, the $C_t \Delta t$ can be seen as close to negligible in Eq.(2.20). The error can be expressed

$$\epsilon \propto \Delta x^r \quad (2.21)$$

The value is expected to be $r = 2$, since Crank Nicolson is used in the spatial discretization. By comparing the error for different Δx , the convergence rate, r ,

can be expressed,

$$r_{12} \simeq \frac{\log(\epsilon_1/\epsilon_2)}{\log(\Delta x_1/\Delta x_2)} \quad (2.22)$$

Since the table above has four different error values, these can be used to give three different convergence rates. $\Delta x_1 = \Delta x$, $\Delta x_2 = \Delta x/2$ The same notation has been used for the different error values, ϵ .

	ϵ_1/ϵ_2	ϵ_2/ϵ_3	ϵ_3/ϵ_4
r	2.0056	2.0014	2.0004

Here the rate goes towards 2, and a 2.order convergence rate seems to be fulfilled.

2.1.7 The temporal error

The temporal error is hard to find since the *Stability criteria* expects Δt to fulfill the criteria in Eq.(2.23) to avoid oscillations.

$$\Delta t \leq \frac{\Delta x^2}{2} \quad (2.23)$$

As a result, $\Delta x \ll \Delta t$ is impossible, because this only leads to an unstable solution. By looking at the column for $\frac{\Delta x}{8}$, the only stable solution is for $\frac{\Delta t}{16}$. Therefore the technique used for the spatial error cannot be used here. By studying the diagonal numbers in the table, the expected convergence rate is fulfilled for both Δx , which gives $r = 2$ and for Δt , which gives $r = 1$

2.1.8 Travelling wave in 2D

The Eq.(2.5) can be discretized for a 2D area. This is more realistic when simulating a geographic spread of an epidemic disease. The non dimensional system can be discretized with Forward Euler in time and Crank Nicolson in space

$$\begin{aligned} \frac{S_{i,j}^{n+1} - S_{i,j}^n}{\Delta t} &= -I_{i,j}^n S_{i,j}^n + \left(\frac{S_{i-1,j}^n - 2S_{i,j}^n + S_{i+1,j}^n}{\Delta x^2} + \frac{S_{i,j-1}^n - 2S_{i,j}^n + S_{i,j+1}^n}{\Delta y^2} \right) \\ \frac{I_{i,j}^{n+1} - I_{i,j}^n}{\Delta t} &= I_{i,j}^n S_{i,j}^n - \lambda I_{i,j}^n + \left(\frac{I_{i-1,j}^n - 2I_{i,j}^n + I_{i+1,j}^n}{\Delta x^2} + \frac{I_{i,j-1}^n - 2I_{i,j}^n + I_{i,j+1}^n}{\Delta y^2} \right) \\ \frac{R_{i,j}^{n+1} - R_{i,j}^n}{\Delta t} &= \lambda I_{i,j}^n + \left(\frac{R_{i-1,j}^n - 2R_{i,j}^n + R_{i+1,j}^n}{\Delta x^2} + \frac{R_{i,j-1}^n - 2R_{i,j}^n + R_{i,j+1}^n}{\Delta y^2} \right) \end{aligned} \quad (2.24)$$

The known values can be placed on the right side. The system will then be

$$\begin{aligned} S_{i,j}^{n+1} &= S_{i,j}^n + \Delta t \left(-I_{i,j}^n S_{i,j}^n + \left(\frac{S_{i-1,j}^n - 2S_{i,j}^n + S_{i+1,j}^n}{\Delta x^2} + \frac{S_{i,j-1}^n - 2S_{i,j}^n + S_{i,j+1}^n}{\Delta y^2} \right) \right) \\ I_{i,j}^{n+1} &= I_{i,j}^n + \Delta t \left(I_{i,j}^n S_{i,j}^n - \lambda I_{i,j}^n + \left(\frac{I_{i-1,j}^n - 2I_{i,j}^n + I_{i+1,j}^n}{\Delta x^2} + \frac{I_{i,j-1}^n - 2I_{i,j}^n + I_{i,j+1}^n}{\Delta y^2} \right) \right) \\ R_{i,j}^{n+1} &= R_{i,j}^n + \Delta t \left(\lambda I_{i,j}^n + \left(\frac{R_{i-1,j}^n - 2R_{i,j}^n + R_{i+1,j}^n}{\Delta x^2} + \frac{R_{i,j-1}^n - 2R_{i,j}^n + R_{i,j+1}^n}{\Delta y^2} \right) \right) \end{aligned} \quad (2.25)$$

This results in an explicit system, which is easy to code. It consists of known values on the right side and only one unknown value on the left side.

2.1.9 A gaussian wave

In the PDE system for the 1D equation, a Gaussian quantity of infected humans was placed on the left side in the initial value. This resulted in a wave of *Infected* spreading along the x-axis. A similar procedure can be done for the 2D simulation. A couple of simulations have been produced for the 2D system. The first simulation is calculated with a Gaussian function along the points(0,y) for the *Infected* group at initial time. The second simulation has placed the Gaussian function at point(0,0) for the *Infected* group at initial value. Both simulations can be seen in the Appendix.

The size of the epidemic wave can be measured and compared by studying the travelling wave at a certain point. In these two 2D simulations in Fig.(2.3), the wave are measured in the point (15,15), while the travelling wave in the 1D simulation was measured at point(15).

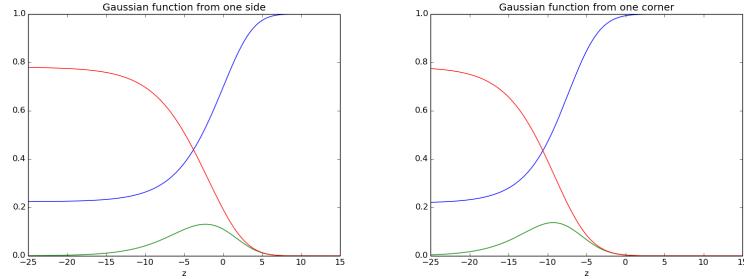


Figure 2.3: Travelling wave measured at point (15,15) with two different initial values for the *Infected* group. I:The initial value is set as a Gaussian line along (0,y).II:The initial value is set as a Gaussian point (0,0).

The shapes of the two travelling waves in Fig.(2.3) are similar. The only difference is the time when the wave occurs. The plot for 1D wave in Fig.(2.2) has the same shape. With a closer study, the area under the function can be measured in all three cases. The result can be seen in Tab.(2.1.9)

1D wave	2D wave line	2D wave point
1.43	1.43	1.43

The area in all three simulations approach the same area when Δt and Δx are reduced. The size and shape will not change by expanding the system from 1D to 2D. However, by studying Fig.(2.3), one can see that the wave occurs at different times. This is caused by the distance from the start position for the Gaussian wave. The first subplot that starts with a Gaussian function along the $x = 0$ axis gets a wave from the *Infected* group that flows along the x axis. This can be seen as a wave on the beach. Everyone that have the same distance from the ocean will be hit simultaneously. The travelling wave for the 1D simulation and the first subplot occurs at the same time, because they are measured at the same distance from the starting point. The last plot is also measured at (15,15), but occurs later. Since the wave starts at point (0,0), the distance to (15,15) is 21.21. This means that the wave will reach the point 6.21 time steps later. This is a reasonable conclusion based on the plot.

Change in initial flow. By increasing the initial wave of the *Infected* group, the initial value of *Infected* can be studied. The simulation is run with the same parameters as for the three simulations above and the only difference is the initial value for the *Infected* group. The Gaussian wave of *Infected* is placed at point(0,0) as for subplot II in Fig.(2.3). The simulation can be seen in Fig.(2.4).

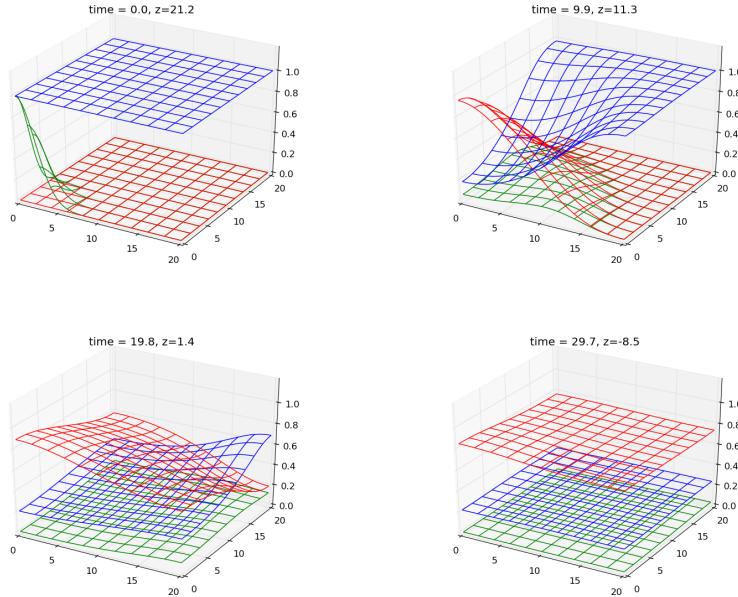


Figure 2.4: A major flow of *Infected* spread outwards in the field. After a certain time, the wave has passed the area and the number in each group stabilized.

The size and shape can be compared by measuring the travelling wave at point(15,15). The travelling wave for this simulation can be seen in Fig.(2.5) and the area for the travelling wave is measured to 1.43, which is similar to the three other simulations.

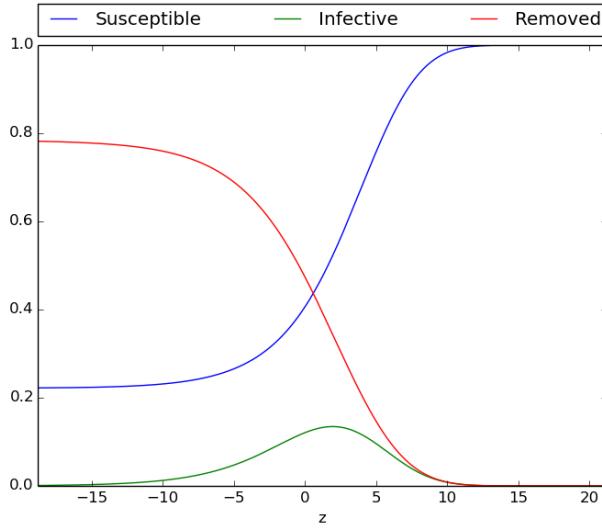


Figure 2.5: The travelling wave with a major increase of *Infected* at the initial time.

The size of the travelling wave will not be affected by changing the value of the *Infected* group. However there is a difference in the time when the wave occurs. In the simulation where the initial value is higher, the travelling wave reaches the measuring point (15,15) earlier. This can be explained by the idea of a ball dropped from a large height. If the ball is released or thrown to the ground, it will only affect the acceleration of the ball, not the terminal velocity. After a certain time the released ball and the thrown ball will reach the same maximum speed. This is the case for the speed of the travelling wave.

Change in lambda. The one thing that affects the speed and size, is the λ variable in the PDE system(2.5). This λ is a combination of a , which controls deaths among the *Infected* group, r , which controls the number from the *Susceptible* group that gets infected in meetings between the *Infected* and *Susceptible* groups. The last parameter in λ is the concentration of Susceptible, S_0 . By changing this parameter, the travelling wave will change in both size and shape. In Fig.(2.6), the simulation is run with four different values of λ .

To understand the results in Fig.(2.6), the λ function can be studied,

$$\lambda = \frac{a}{rS_0}, \quad (2.26)$$

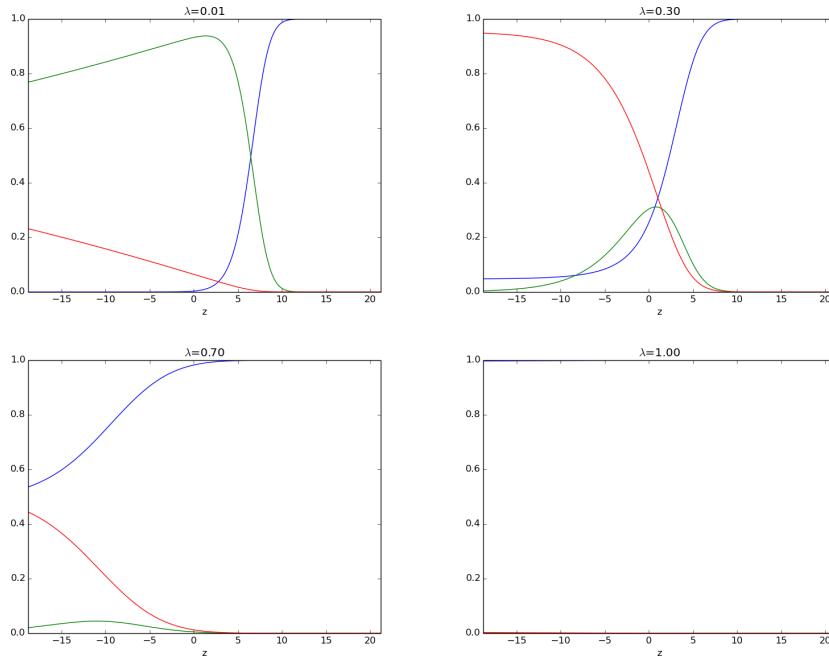


Figure 2.6: The travelling wave simulated with λ values in the range of 0.01 to 1.

A major and aggressive travelling wave is caused when $\lambda \rightarrow 0$. In Fig.(2.6), λ is run with value 0.01 in the first subplot. This results in a travelling wave of *Infected* that eradicates the *Susceptible* group in a short time. The wave starts decreasing when all *Susceptible* are infected. By looking at Eq.(2.26), one can see that a small value is caused by a small a compared to r and S_0 . If a is low, this results in few deaths/immune in the *Infected* group. This means that the *Infected* group will grow and be able to infect even more humans from the *Susceptible* group. The same thing will happen if r is large. A result of a large r will be an aggressive disease that infects a major part of the population. The same result will happen if S_0 is large. Then there are several possible humans to infect. Therefore an outburst of a disease is more critical in a crowded city than in the wilderness, far from other humans.

If λ increases above 1, the disease will not be able to spread. The number of *Infected* will decrease, since the number in the *Removed* group caused by the *Infected* group is higher than the amount of *Infected* humans from the *Susceptible* group. After a certain time, the number of *Infected* will die out. If λ stays at 1, the number of *Infected* will be equal the whole time.

2.2 English Boarding School

An example from an English boarding school was presented in the previously chapter *Epidemic models*. This example was based on the book from J.D Murray [?], and was modeled for an ODE system. A similar result should appear for the PDE system with the same parameter values and a uniform distribution of the groups. The school had 763 students, and one of the students brought a disease back to the school. The following numbers were used for the ODE system in chapter one. $N = 763, S_0 = 762, I_0 = 1, R_0 = 0, \rho = 202$ and $r = 2.18 \cdot 10^{-3}$.

The first simulation is produced with uniform distributed concentration. This is done to verify the implementation. A person is defined as one cubic. The total volume of the whole group is spread over the area. The area is set to be 100 m x 100 m, which results in an average height of 1/10000 m per person. This is done to get a uniformed distribution. This would of course be more difficult in real life, particularly if the person would be alive. Since the *Infected* group only consists of one person, the total height will be 0.0001 for the whole area. The *Susceptible* group consists of 762 students and the total height at each point will be 0.0762. The simulation can be seen in the Appendix.

The results from subplot I in Fig.(2.7) are equal to the results from the ODE system modeled in the previously chapter. This can be seen in Tab.(2.2). This is as expected, since the diffusion term is negligible in this system. The simulation results in a group of separate ODE systems modeled over an area.

	ODE system	PDE uniform dist	PDE center	PDE corner
<hr/>				
5 Days	—	—	—	—
Susceptible	444.62	444.62	748.03	757.33
Infective	209.56	209.56	7.36	2.35
Removed	108.82	108.82	7.60	3.32
<hr/>				
10 Days	—	—	—	—
Susceptible	37.59	37.59	697.71	743.58
Infective	117.59	117.59	24.43	6.66
Removed	607.82	607.82	40.86	12.76
<hr/>				
15 Days	—	—	—	—
Susceptible	21.09	21.09	597.01	717.02
Infective	17.30	17.30	46.96	12.37
Removed	724.62	724.62	119.03	33.61

2.2.1 Introducing a Gaussian distribution of infected

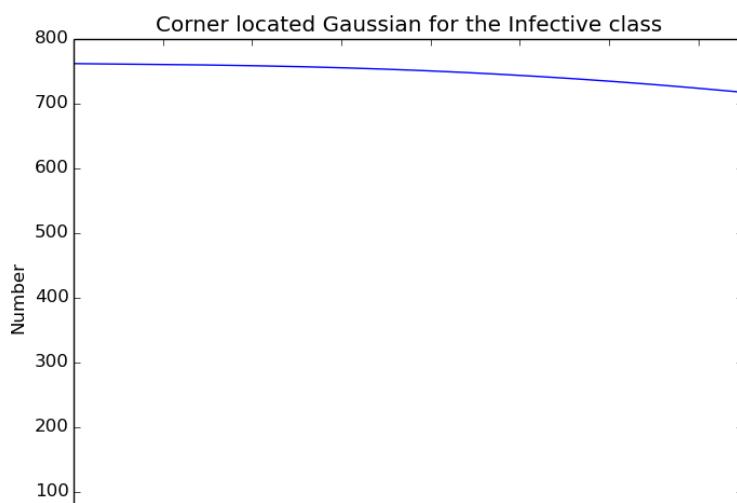
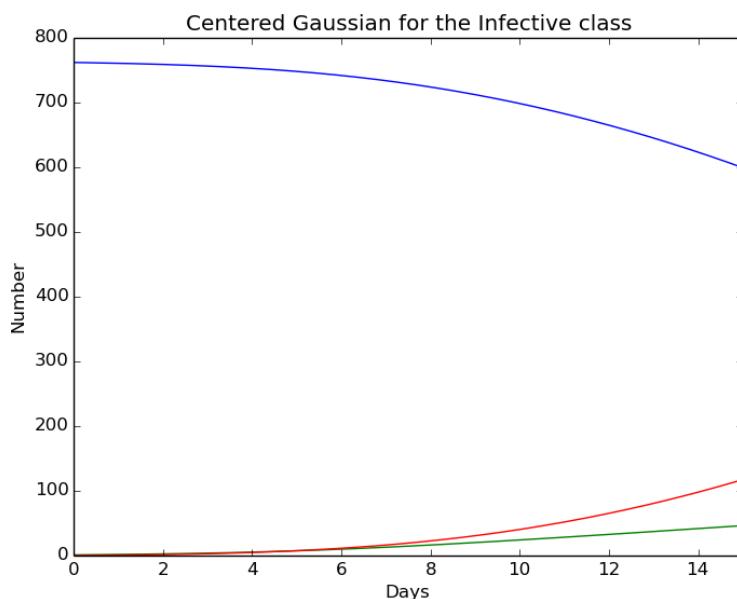
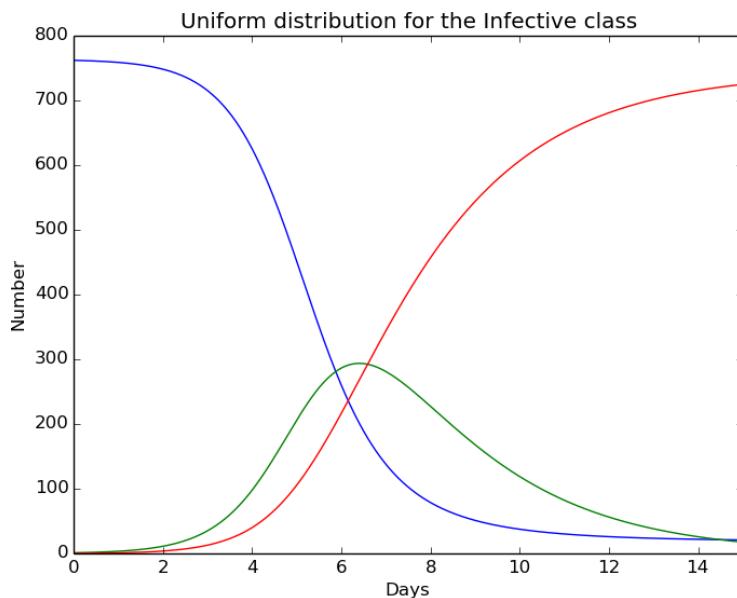
An assumption one can make is that a person is not able to be evenly distributed over an area. In this example, with only one infected student at initial time, the chance of being infected increases the closer the *Susceptible* group gets the infected student. The student is represented by a Gaussian function in the

middle of the school yard, to see if the position affects the result. The height is set to 1 and the volume of the Gauss function is set to 1 cubic. The simulation can be seen in Fig.(2.8) and the total amount of students in each group can be seen in Fig.(2.7).

The results from the uniform distributed and Gaussian distributed simulations show various results. The initial position of the *Infected* group is the difference between the simulations. This has a major impact. Since the only ones that can be infected by the Gaussian distribution are the students close to the infected student, this restricts the spread of the epidemic. The chance of getting infected in this area is higher. The Fig.(2.8) shows that the amount of *Infective* quickly grows in the center, where the infected was placed. Subplot IV in Fig.(2.8) shows that the amount from the *Removed* group in the center is close to the maximum of the initial value of *Susceptible*, while the students along the boundary of the schoolyard seem to be unaffected after 15 days. This simulation shows that the position of the *Infected* group has a major role in the simulation.

The position of the *Infected* group, here as a Gaussian function, also affects the outcome. Subplot III in Fig.(2.7) describes a simulation where the Gaussian function is placed in the corner with position(0,0). The total volume of the function is increased to 4 since only a quarter of the function is placed in the area. Tab.(2.2) shows that the total number of *Infected* is lower than for the centered placed Gaussian function. The infected student is only able to spread the disease to a quarter of the population compared to the infected student in the center. The simulation can be seen in the Appendix.

If the simulations are run for a long time, the difference between each group will decrease. After 100 days there will be about 18 students in the *Susceptible* group in the uniform distributed simulation, compared to 25 students in both of the Gaussian simulations. The simulations can be seen in the Appendix.



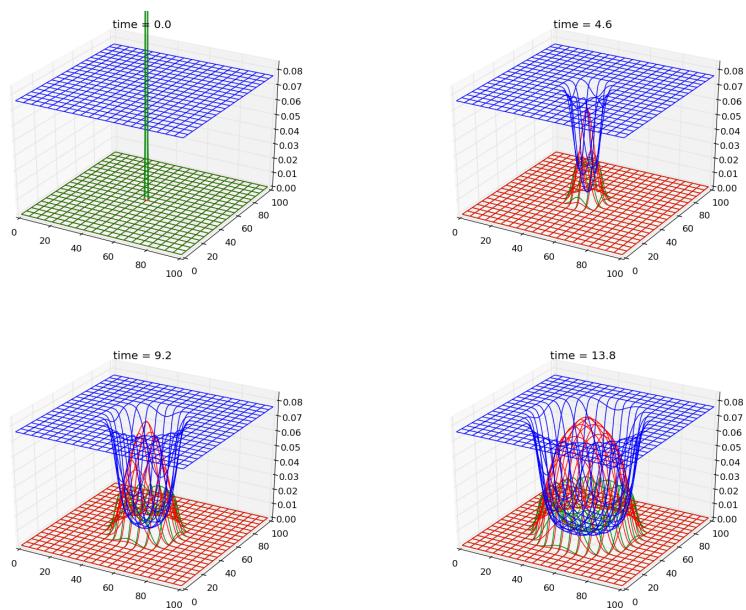


Figure 2.8: The infected student is placed in the center as a Gaussian function at initial time. The height of the Gaussian function is set to 1 m and volume is set to 1 cubic.

2.3 Zombiefication

The previous chapter studied an ODE system designed to calculate the number in the four groups: *Susceptible*, *Infected*, *Zombie* and *Removed* during the five first episodes in the TV series *Walking Dead* Ref.[?]. The model was based on the model from Langtangen, Mardal and Røtnes Ref.[?], with an extra term in the *counter attack* phase. The ODE system from the chapter *Epidemic models* can be expanded with a diffusion term in each group to make a PDE system. This can be seen in Eq.(2.27)

$$\begin{aligned}\frac{\partial S}{\partial t} &= \Sigma - (\beta + \mu\omega(t))SZ - \delta_S S + D_s \nabla^2 S \\ \frac{\partial I}{\partial t} &= (\beta + \mu\omega(t))SZ - \varrho I - \delta_I I + D_i \nabla^2 I \\ \frac{\partial Z}{\partial t} &= \varrho I - (\alpha + \omega(t))SZ + \zeta R + D_z \nabla^2 Z \\ \frac{\partial R}{\partial t} &= \delta_S S + \delta_I I - \zeta R + (\alpha + \omega(t))SZ + D_r \nabla^2 R\end{aligned}\tag{2.27}$$

The Eq.(2.27) can be solved numerically by discretization. Forward Euler is used for the time derivative and Crank Nicolson for the space derivative. This is solved with the same technique as for the SIR model(2.24). The system can be seen in Eq.(2.28)

$$\begin{aligned}\frac{S_{i,j}^{n+1} - S_{i,j}^n}{\Delta t} &= \Sigma - (\beta + \mu\omega(t))S_{i,j}^n Z_{i,j}^n - \delta_S S_{i,j}^n \\ &\quad + D_s \left(\frac{S_{i-1,j}^n - 2S_{i,j}^n + S_{i+1,j}^n}{\Delta x^2} + \frac{S_{i,j-1}^n - 2S_{i,j}^n + S_{i,j+1}^n}{\Delta y^2} \right) \\ \frac{I_{i,j}^{n+1} - I_{i,j}^n}{\Delta t} &= (\beta + \mu\omega(t))S_{i,j}^n Z_{i,j}^n - \varrho I_{i,j}^n - \delta_I I_{i,j}^n \\ &\quad + D_i \left(\frac{I_{i-1,j}^n - 2I_{i,j}^n + I_{i+1,j}^n}{\Delta x^2} + \frac{I_{i,j-1}^n - 2I_{i,j}^n + I_{i,j+1}^n}{\Delta y^2} \right) \\ \frac{Z_{i,j}^{n+1} - Z_{i,j}^n}{\Delta t} &= \varrho I_{i,j}^n - (\alpha + \omega(t))S_{i,j}^n Z_{i,j}^n + \zeta R_{i,j}^n \\ &\quad + D_z \left(\frac{Z_{i-1,j}^n - 2Z_{i,j}^n + Z_{i+1,j}^n}{\Delta x^2} + \frac{Z_{i,j-1}^n - 2Z_{i,j}^n + Z_{i,j+1}^n}{\Delta y^2} \right) \\ \frac{R_{i,j}^{n+1} - R_{i,j}^n}{\Delta t} &= \delta_S S_{i,j}^n + \delta_I I_{i,j}^n - \zeta R_{i,j}^n + (\alpha + \omega(t))S_{i,j}^n Z_{i,j}^n \\ &\quad + D_r \left(\frac{R_{i-1,j}^n - 2R_{i,j}^n + R_{i+1,j}^n}{\Delta x^2} + \frac{R_{i,j-1}^n - 2R_{i,j}^n + R_{i,j+1}^n}{\Delta y^2} \right)\end{aligned}\tag{2.28}$$

By setting the unknown to the left, the following system (2.29) can be solved:

$$\begin{aligned}
 S_{i,j}^{n+1} &= S_{i,j}^n + \Delta t \left(\Sigma - (\beta + \mu\omega(t))S_{i,j}^n Z_{i,j}^n - \delta_S S_{i,j}^n \right. \\
 &\quad \left. + D_s \left(\frac{S_{i-1,j}^n - 2S_{i,j}^n + S_{i+1,j}^n}{\Delta x^2} + \frac{S_{i,j-1}^n - 2S_{i,j}^n + S_{i,j+1}^n}{\Delta y^2} \right) \right) \\
 I_{i,j}^{n+1} &= I_{i,j}^n + \Delta t \left((\beta + \mu\omega(t))S_{i,j}^n Z_{i,j}^n - \varrho I_{i,j}^n - \delta_I I_{i,j}^n \right. \\
 &\quad \left. + D_i \left(\frac{I_{i-1,j}^n - 2I_{i,j}^n + I_{i+1,j}^n}{\Delta x^2} + \frac{I_{i,j-1}^n - 2I_{i,j}^n + I_{i,j+1}^n}{\Delta y^2} \right) \right) \\
 Z_{i,j}^{n+1} &= Z_{i,j}^n + \Delta t \left(\varrho I_{i,j}^n - (\alpha + \omega(t))S_{i,j}^n Z_{i,j}^n + \zeta R_{i,j}^n \right. \\
 &\quad \left. + D_z \left(\frac{Z_{i-1,j}^n - 2Z_{i,j}^n + Z_{i+1,j}^n}{\Delta x^2} + \frac{Z_{i,j-1}^n - 2Z_{i,j}^n + Z_{i,j+1}^n}{\Delta y^2} \right) \right) \\
 R_{i,j}^{n+1} &= R_{i,j}^n + \Delta t \left(\delta_S S_{i,j}^n + \delta_I I_{i,j}^n - \zeta R_{i,j}^n + (\alpha + \omega(t))S_{i,j}^n Z_{i,j}^n \right. \\
 &\quad \left. + D_r \left(\frac{R_{i-1,j}^n - 2R_{i,j}^n + R_{i+1,j}^n}{\Delta x^2} + \frac{R_{i,j-1}^n - 2R_{i,j}^n + R_{i,j+1}^n}{\Delta y^2} \right) \right)
 \end{aligned} \tag{2.29}$$

A simulation with uniform distributed groups can be done to verify the implementation of the system. The result is expected to be similar to the ODE system in the previous chapter. A zombie attack can be separated into three different phases, based on the paper from Langtangen, Mardal and Røtnes [?]. The first phase is short, and it is called the *Initial phase*. The humans are unfamiliar with the disease in this phase and are as a consequence quite naive to the disease. This result in a high chance of getting infected. The next phase is called the *Hysterical phase*. The humans are now more familiar with the situation and try to avoid the *Infected* group. This result in a lower chance of getting infected. The last phase, which happens at the same time as the *Hysterical phase*, is the *Counter attack*. This phase is often initiated when humans are attacked by zombies. The following parameters that were used for simulating the first episodes of *Walking Dead* will be used here. These can be seen in Tab.(2.3). By computing the system for all three phases, the value in each phase can be compared to the ones from the ODE system. This will give an indication of whether the discretization is done correct.

parameter	Initial phase	hysterical phase	counter attack
β	0.01155	0.000011	0.00011
ϱ	1.37	1.5	1.5
α	0.00044	0.000208	0.000208
a	0	0	0.0073
σ	0	0	0.005
μ	0	0	0.14

The simulation in Fig.(2.9) seems to match the results from the ODE system. A closer check can be done by comparing the groups in each phase. This result can be seen in Tab.(2.3)

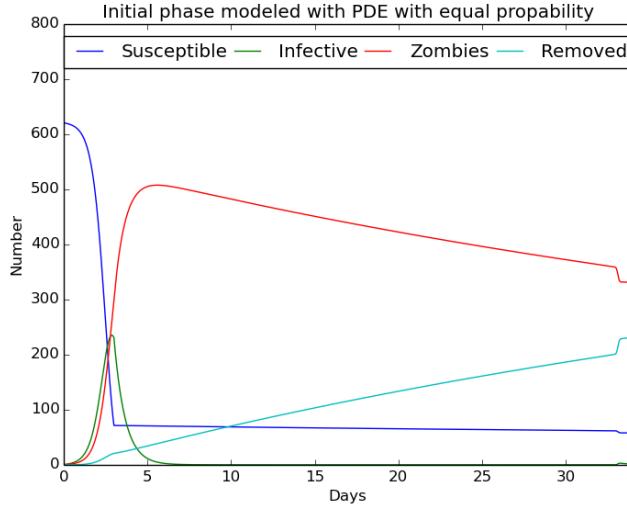


Figure 2.9: The Eq.(2.29) modeled with uniformed distributed groups. Initial values $S_0 = 621$, $I_0 = 0$, $Z_0 = 0$, $R_0 = 0$ with parameters from (2.3).

The initial values for the four groups are set to $S_0 = 621$, $I_0 = 0$, $Z_0 = 1$ and $R_0 = 0$ in all simulations. The values in Tab.(2.3) are measured at the final time for each phase. The *Initial phase* lasts for three days and the values are measured at *time* = 3. The *Hysterical phase* is a continuous phase, and will not stop until an eventual eradication. The values are therefore given before the *Counter attack* at *time* = 33. The *Counter attack* lasts for some hours, and is measured at *time* = 34, which is a day after the attack. The value of Δt is set to 1E-3.

	ODE system	PDE uniform dist	PDE gauss center
<hr/>			
Initial phase			
Susceptible	71.3	71.3	81.12
Infected	230.8	230.8	210.94
Zombie	298.9	298.9	310.11
Removed	21.0	21.0	20.60
<hr/>			
Hysterical phase			
Susceptible	61.6	61.6	70.55
Infected	0.3	0.3	0.34
Zombie	358.6	355.6	334.33
Removed	201.5	201.5	217.56
<hr/>			
Counter attack			
Susceptible	57.8	57.8	66.50
Infected	1.2	1.2	1.23
Zombie	331.8	331.8	305.86
Removed	231.3	231.3	249.19

These results shows that the PDE system gives the same results as the ODE system.

2.3.1 The position of the *Infected* group

In the previous section, *English boarding school*, the location of the *Infected* group was proven to have a major influence on the result. However here the *Susceptible* group was uniformly distributed over the schoolyard. The number in each group, based on the study of *Walking Dead*, was seen in three different locations in the TV series. By only studying the TV series, it is hard to decide the geographical distance between these three locations. Therefore they have been placed with a certain distance from each other. The following simulations are done on a grid with size(40 x 40) with the following positions for the towns: Small town(6,6) with size 21, middle town(12,25) with size 200 and large town(25,12) with size 400. Since these values were based on the humans and zombies seen in the series, these can be scaled up by 1000 to correspond a realistic population. A large town can be seen as an area with a population of 400 000. The length can be measured in kilometres. Then the distance between the middle and large town will be 18.38 km. Compared to the distance between Oslo and Bærum(Sandvika) which is 15 km, the simulation can be seen as a rough estimate of the area around Oslo, if the fjord is frozen and the mobility for the area is similar.

The diffusion term describes the diffusion for each group. This can be seen as the speed towards equilibrium for each group. If the diffusion constant is large, the flow towards equilibrium will go faster. The values have been set as follows.

- $D_s = 1$, The *Susceptible* has a basic moving speed. The other groups are based on the speed of a healthy human.

- $D_i = 0.5$, The *Infected* are often injured caused by recent fights. This affects their mobility.
- $D_z = 0.9$, The average moving speed for a zombie. There are zombies with the mobility of a human, but there are also zombies that drag themselves forward with only one arm. This results in a major difference in speed and therefore a lower average speed for the zombies.
- $D_r = 0$, The removed are here seen as dead, therefore there is no mobility.

The parameters from Tab.(2.3) will be used here, and the three phases will be modeled as shown for the uniformed distributed PDE system. The values will be used for three different simulations with the similar initial value for the different groups. The position of initial values can be seen in Fig.(2.10) and are based on the data given for each town above. The difference in the three simulations will be the position of the zombie at initial time. The zombie will be placed in center of the small, middle and large town.

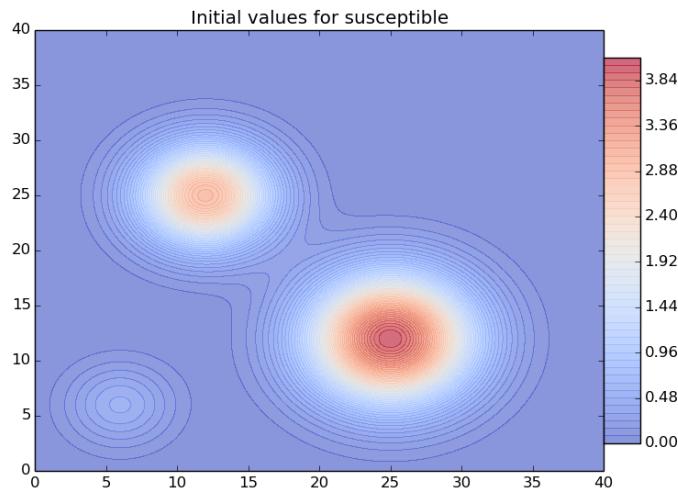


Figure 2.10: The initial value for the *Susceptible* group for three simulations. Small group(6,6) with volume 21, middle group(12,25) with volume 200 and large group(25,12) with volume 400. All three groups are build up with a Gaussian function.

Fig.(2.11) shows the simulation where the zombie is placed in the large town. The simulations of the small and middle town can be seen in the Appendix. The four subplots are from the different phases that arise during a zombie attack. The different groups have the same color as introduced in Fig.(2.9). It is difficult

to separate the three groups *Infected*, *Zombie* and *Removed*, since they all have a low value at initial time. The development of the amount can easier be seen in the Fig.(2.12), which also shows the results from the small and middle town. Since the amount of *Susceptible* is quite low in the small town where the zombie arises, the disease is not able to infect too many before the society has moved to the next phase, assuming that the broadcasting about the disease works okay for the first days. This results in an eradication of the disease in about a month. The table(2.3) shows that the number of zombies decreases towards zero after a month.

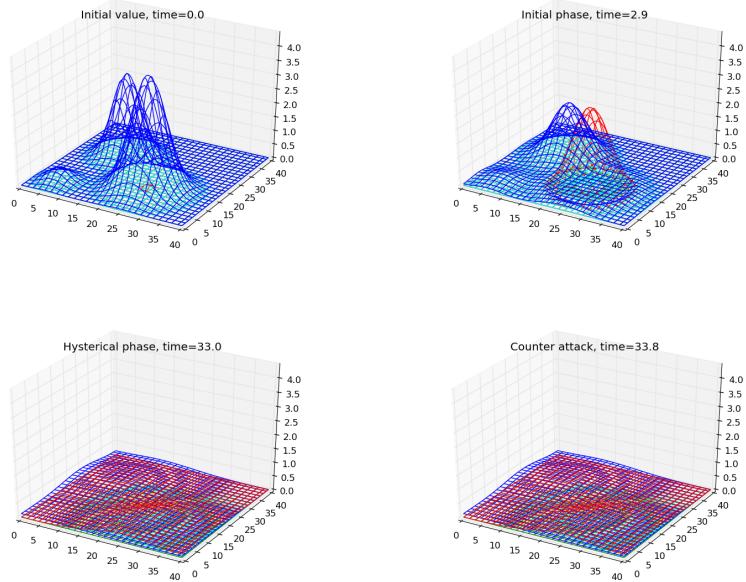


Figure 2.11: *Walking Dead* simulated with the zombie at initial time in the large town. Subplots shown at each phase.

By placing the zombie in the middle town, the amount in the *Zombie* group increases to a higher level. The amount can be seen as subplot II in Fig.(2.12). The damages are higher, and after a month the total population of *Susceptible* is reduced to 427. The last calculation done for the large town in Fig.(2.11) shows major damages. Here the amount in the *Zombie* group increases above the number of *Susceptible*. The *Infected* group also increases to above 100 after a couple of days in the initial phase. This can be explained by the high number of meetings between susceptible and zombies. By studying the subplot II in Fig.(2.11), the zombies are grouped in the large town, while the middle and small town mostly consist of *Susceptible*. By counting the loss of *Susceptible* during the first phase, the Tab.(2.3.1) shows that this amount corresponds with the size of the towns where the zombie was placed, given by the number 17,188 and 362 in the small,middle and large group. The percent is highest in the middle town with 94 percent. The percent in the large and small group are 90 and 81. The

simulation in the middle town has the highest percent because the *Susceptible* group in the large town also gets infected. These towns are coupled together, and the epidemic disease is able to spread. The small town, which can be seen as Nesodden(a peninsula outside Oslo), has the lowest percent. The reason is that the disease is not able to spread to the other towns, and therefore is causes less damage.

The results from the uniformed distributed simulation is still much higher for the *Zombie* group the in these three simulations. This shows that using the parameters from the ODE system in a geographical area makes little sense. A realistic assumption is that a zombie is restricted to a given area, and therefore the parameters will not be equal for all. The chance of getting infected is much higher if a person from the *Susceptible* group is close to an infected. There is also a greater chance of getting infected if the *Susceptible* group has a high density.

	Small town	Middle town	Large
Initial phase	—	—	—
Susceptible	603.74	433.22	259.20
Infected	2.96	25.51	50.94
Zombie	13.79	155.27	297.24
Removed	0.66	7.16	13.78
Hysterical phase	—	—	—
Susceptible	604.42	429.35	251.14
Infected	0.03	0.18	0.35
Zombie	6.25	87.31	178.45
Removed	12.14	106.00	192.90
Counter attack	—	—	—
Susceptible	604.21	427.45	247.33
Infected	0.08	0.59	1.17
Zombie	4.49	73.70	151.44
Removed	14.11	121.16	222.96

2.3.2 Free areas for the *Susceptible* group

To model a realistic zombie attack, humans ability to think logically is crucial in the fight. The mobility was presented as a factor in the previous section. Another important skill that the *Susceptible* group holds, is the ability to decide the safety of an area. In the TV series *Walking Dead*, the humans build barricades to keep the zombies outside. This gives the *Susceptible* group free areas where they can stay. This idea can be transferred to the PDE system by rewriting the Eq.(2.27) with spatial dependent diffusion terms. The diffusion constant D_u is now replaced with a diffusion function $\gamma_u(x)$ for $u = S, I, Z, R$, which is spatial discretized. Since a diffusion equation always goes towards equilibrium, this rewriting will only slow down/stop the selected group to diffuse into an area. In

this case it will stop the *Zombie* group from diffusing into the buildings.

$$\begin{aligned}\frac{\partial S}{\partial t} &= \Sigma - (\beta + \mu\omega(t))SZ - \delta_S S + \nabla(\gamma_S(x)\nabla S) \\ \frac{\partial I}{\partial t} &= (\beta + \mu\omega(t))SZ - \varrho I - D_i \delta_I I + \nabla(\gamma_I(x)\nabla I) \\ \frac{\partial Z}{\partial t} &= \varrho I - (\alpha + \omega(t))SZ + \zeta R + \nabla(\gamma_Z(x)\nabla Z) \\ \frac{\partial R}{\partial t} &= \delta_S S + \delta_I I - \zeta R + (\alpha + \omega(t))SZ + \nabla(\gamma_R(x)\nabla R)\end{aligned}\tag{2.30}$$

The diffusion term is the difference between this system and Eq.(2.27). The discretization can be shown for a general γ . This will be similar for all groups. A Crank Nicolson discretization is used in space.

$$\begin{aligned}&= \nabla(\gamma(x)\nabla S) \\ &= (\gamma(x)S_x)_x + (\gamma(x)S_y)_y \\ &= \left(\gamma(x) \frac{S_{i+1/2,j}^n - S_{i-1/2,j}^n}{\Delta x} \right)_x + \left(\gamma(x) \frac{S_{i,j+1/2}^n - S_{i,j-1/2}^n}{\Delta y} \right)_y \\ &= \left(\frac{\gamma(x_{i+1/2,j})(S_{i+1,j}^n - S_{i,j}^n) - \gamma(x_{i-1/2,j})(S_{i,j}^n - S_{i-1,j}^n)}{\Delta x^2} \right) \\ &\quad + \left(\frac{\gamma(x_{i,j+1/2})(S_{i,j+1}^n - S_{i,j}^n) - \gamma(x_{i,j-1/2})(S_{i,j}^n - S_{i,j-1}^n)}{\Delta y^2} \right)\end{aligned}\tag{2.31}$$

Since the calculations are based on spatial points, the values inside the function of γ need to be adjusted. This can be done by the use of an arithmetic mean, which can be seen in Eq.(2.32). The notation $q_{i+1/2}$ is a simplification of the function $q(x_{i+1/2})$ with $x_{i+1/2} = x_i + 1/2\Delta x$

$$q_{i+1/2} \approx \frac{1}{2}(q_i + q_{i+1})\tag{2.32}$$

This arithmetic mean can be inserted for all γ 's in the system. The system can be expressed:

$$\begin{aligned}
 S_{i,j}^{n+1} &= S_{i,j}^n + \Delta t \left(\Sigma - (\beta + \mu\omega(t))S_{i,j}^n Z_{i,j}^n - \delta_S S_{i,j}^n \right. \\
 &\quad \left. + \frac{1}{2\Delta x^2} (\gamma_S(x_{i-1,j})(S_{i-1,j}^n - S_{i,j}^n) + \gamma_S(x_{i,j})(S_{i-1,j}^n - 2S_{i,j}^n + S_{i+1,j}^n) + \gamma_S(x_{i+1,j})(-S_{i,j}^n + S_{i+1,j}^n)) \right. \\
 &\quad \left. + \frac{1}{2\Delta y^2} (\gamma_S(x_{i,j-1})(S_{i,j-1}^n - S_{i,j}^n) + \gamma_S(x_{i,j})(S_{i,j-1}^n - 2S_{i,j}^n + S_{i,j+1}^n) + \gamma_S(x_{i,j+1})(-S_{i,j}^n + S_{i,j+1}^n)) \right) \\
 I_{i,j}^{n+1} &= I_{i,j}^n + \Delta t \left((\beta + \mu\omega(t))S_{i,j}^n Z_{i,j}^n - \varrho I_{i,j}^n - \delta_I I_{i,j}^n \right. \\
 &\quad \left. + \frac{1}{2\Delta x^2} (\gamma_I(x_{i-1,j})(I_{i-1,j}^n - I_{i,j}^n) + \gamma_I(x_{i,j})(I_{i-1,j}^n - 2I_{i,j}^n + I_{i+1,j}^n) + \gamma_I(x_{i+1,j})(-I_{i,j}^n + I_{i+1,j}^n)) \right. \\
 &\quad \left. + \frac{1}{2\Delta y^2} (\gamma_I(x_{i,j-1})(I_{i,j-1}^n - I_{i,j}^n) + \gamma_I(x_{i,j})(I_{i,j-1}^n - 2I_{i,j}^n + I_{i,j+1}^n) + \gamma_I(x_{i,j+1})(-I_{i,j}^n + I_{i,j+1}^n)) \right) \\
 Z_{i,j}^{n+1} &= Z_{i,j}^n + \Delta t \left(\varrho I_{i,j}^n - (\alpha + \omega(t))S_{i,j}^n Z_{i,j}^n + \zeta R_{i,j}^n \right. \\
 &\quad \left. + \frac{1}{2\Delta x^2} (\gamma_Z(x_{i-1,j})(Z_{i-1,j}^n - Z_{i,j}^n) + \gamma_Z(x_{i,j})(Z_{i-1,j}^n - 2Z_{i,j}^n + Z_{i+1,j}^n) + \gamma_Z(x_{i+1,j})(-Z_{i,j}^n + Z_{i+1,j}^n)) \right. \\
 &\quad \left. + \frac{1}{2\Delta y^2} (\gamma_Z(x_{i,j-1})(Z_{i,j-1}^n - Z_{i,j}^n) + \gamma_Z(x_{i,j})(Z_{i,j-1}^n - 2Z_{i,j}^n + Z_{i,j+1}^n) + \gamma_Z(x_{i,j+1})(-Z_{i,j}^n + Z_{i,j+1}^n)) \right) \\
 R_{i,j}^{n+1} &= R_{i,j}^n + \Delta t (\delta_S S_{i,j}^n + \delta_I I_{i,j}^n - \zeta R_{i,j}^n + (\alpha + \omega(t))S_{i,j}^n Z_{i,j}^n)
 \end{aligned} \tag{2.33}$$

The diffusion term for the *Removed* group is taken away, since dead people are not able to move. This system looks quite messy, but it is straight forward to calculate. All values on the right side are known values and the system is easy to solve. Now every point will be controlled by the diffusion constants given in $\gamma(x)$. This makes it easier to control the flow in each group. With a high diffusion constant, the diffusion will spread fast. When the diffusion constant goes towards zero, the flow will decrease towards zero flow. This will result in a set of ODE systems modeled for each point.

2.3.3 Ten minutes at Frederikkeplassen

Frederikkeplassen at the University of Oslo is a possible area for an upcoming zombie attack. This simulation will try to model a ten minute sequence with the diffusion parameter added in this section. Since students often learn and interact fast, they will only use three minutes before they realize the danger and transitions into the *Hysterical phase*. A map of Frederikkeplassen is used to define the safe and critical areas. The buildings are set as areas where only the *Susceptible* are allowed to move. This is done by setting the diffusion constant to zero for the *Zombie* and *Infected* groups. Since the buildings are safe spots for the *Susceptible*, an idea would be to express this in the diffusion term by forcing the *Susceptible* for other areas into the buildings. This is more difficult, since the concentrations in each group wants to diffuse towards equilibrium. A way to delay this process is by setting the diffusion constant to be low in the

buildings and high outside. This will result in a fast diffusion in the open areas and a slow diffusion inside the buildings.

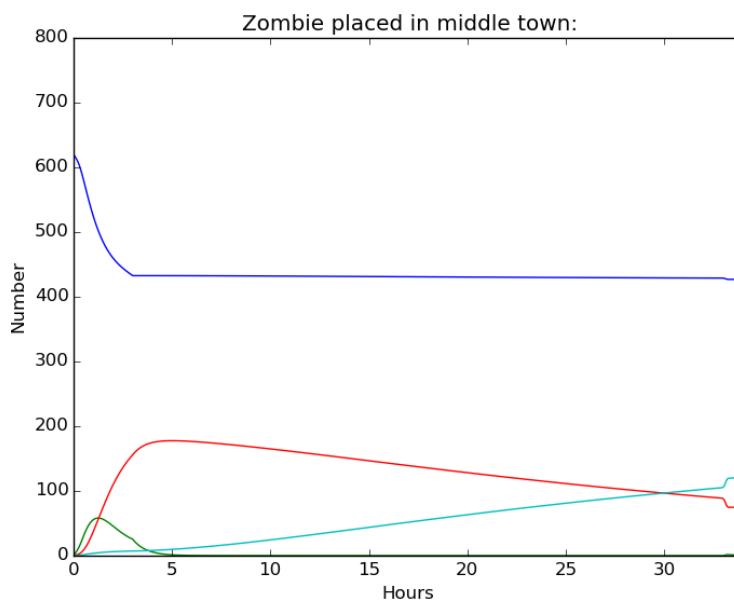
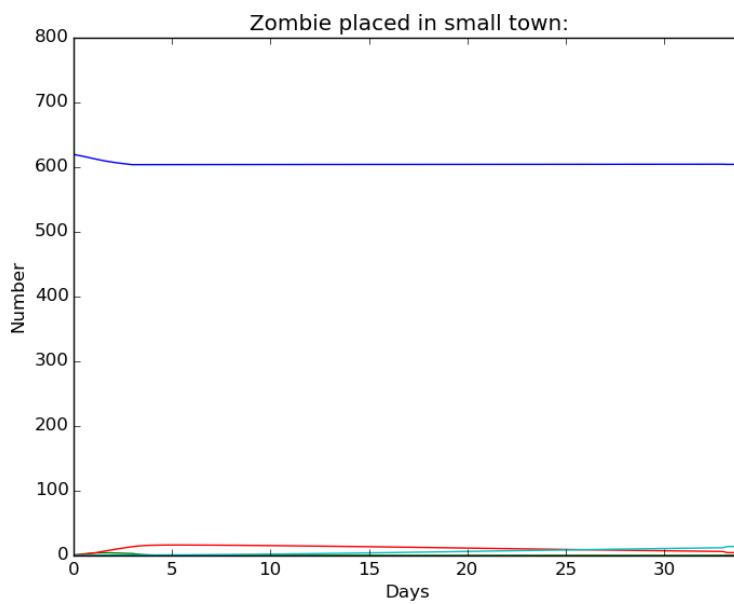
Two simulations have been done at Frederikkeplassen. The amount in each group can be seen in Fig.(2.13). The first simulation has a solution based on the ODE system, with uniformed distributed groups, equal diffusion constants and no free areas for the *Susceptible*. The second simulation is modeled with three groups of *Susceptible*, as in the previous section. The small group with 21 students is placed at point(4,4), the middle group with 200 students is placed at point(15,8) and the large group with 400 students is placed at point(8,13). The zombie is placed at point(8,10). The $\gamma(x)$ is set to zero in the buildings for the *Zombie* and *Infected* group, and one in the rest of the area. For the *Susceptible* group, $\gamma(x)$ is set to 0.1 in the buildings, which causes slow diffusion. In the outside areas, $\gamma(x)$ is set to 5 for the *Susceptible*. The desired result is to push them into the buildings, but this will only happen if there is a lower concentration inside the buildings. Therefore this will not reflect a realistic flow of a *Susceptible* population. This simulation can be seen in Fig.(2.14)

The results in Tab.(2.3.3) shows that the three first minutes are crucial. The number after three minutes shows that only 72 humans survived the attack in the uniformed solution, compared to 252 in the free areas. Even more survived in the simulation with the random placement. This will be discussed in the next section. The number in the *Zombie* group is quite similar in *Uniformed distribution* and *Free areas* measured at $t = 3$. However at $t = 7$ the difference is major. This can be explained by looking at Fig.(2.14) and the building with the middle group placed inside. When the zombie starts attacking at $t = 0$, the large group is exposed. This group is placed close to the zombie and the position is in an open area. The zombie can attack right away and the number in the *Infected* and *Zombie* groups increases fast. In the two first minutes, a major part of the large group is infected and the *Zombie* group starts to spread. After 2-3 minutes, the group has reached the buildings with the middle group. Here the diffusion is set to 0, and the spread of zombies stop. Since the diffusion variable for the *Susceptible* is quite low inside the buildings, it takes time before the group diffuses. Maybe the right diffusion value along the buildings would be 0, to avoid any leakage. This would again cause problem for the diffusion of *Susceptible* into the buildings. It is also reasonable to think that the *Susceptible* group needs to diffuse after a certain time. The lack of supplies would force them out.

	Uniform distribution	Free areas	Random placement
3 Minutes			
Susceptible	72.23	252.72	524.77
Infected	229.65	75.69	26.07
Zombie	296.67	276.55	66.51
Removed	20.84	13.94	3.66
7 Minutes			
Susceptible	70.78	251.35	524.23
Infected	0.83	0.51	0.20
Zombie	498.72	325.54	81.88
Removed	49.12	41.26	14.80
10 Minutes			
Susceptible	69.69	249.84	523.61
Infected	0.25	0.38	0.16
Zombie	479.00	295.71	69.67
Removed	70.55	72.36	27.77

2.3.4 Random distribution

This random distribution of *Susceptible* and *Zombie* is simulated with the idea of the next chapter. The volume of the *Susceptible* group is randomly placed over the area. The volume is similar to the two other simulations in Tab.(2.3.3). The zombie is described with a Gaussian function at a random position. The simulation has the same $\gamma(x)$ as the simulation of the free areas, but the result is different. This simulation was done a couple of times, and resulted in different solutions each time. The position of the zombie had a major influence on the result. If the position of the zombie was in the center, the amount of *Infected* and *Zombie* was larger than if the position of the zombie was in the corner. A specific random simulation cannot be compared with the two other simulations, but the range of the solutions be used. The result shows that the number of zombies at *time = 3* never exceeded the result from the free areas. A random position of people will never reach the amount when the *Zombie* group is uniformly distributed.



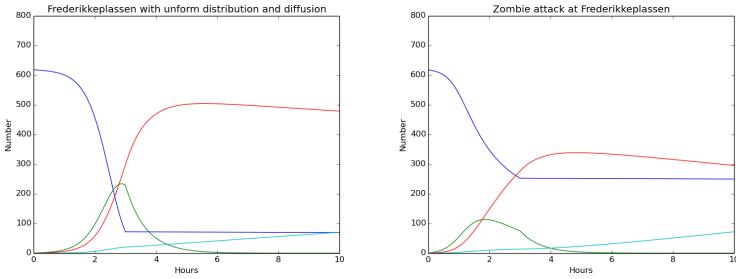


Figure 2.13: The amount in each group for two simulations of Frederikkeplassen modeled same parameters for 10 minutes/hours. a)Plot with uniformed distributed groups and same diffusion constants for all groups. b)Plot based on figure(2.14) with different initial values for each group.

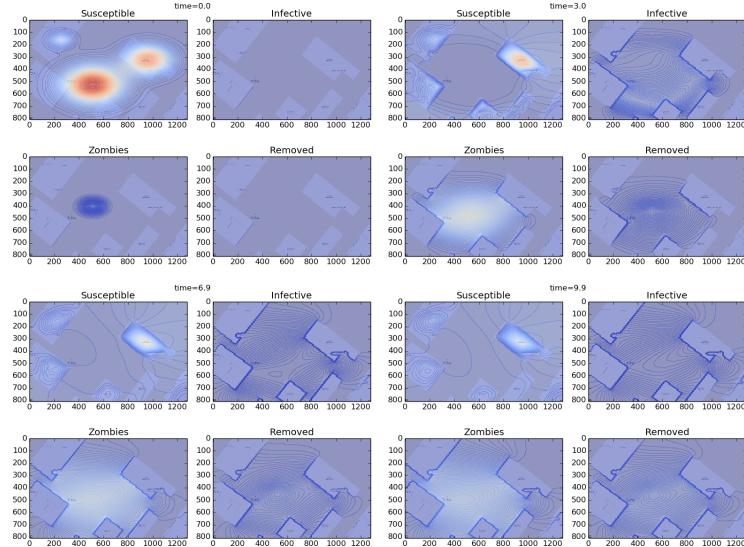


Figure 2.14: Frederikkeplassen modeled with free areas for the *Susceptible* group. The diffusion function $\gamma(x)$ is set to zero for the *Zombie* and *Infected* group in the buildings. The zombie at initial time is placed in the center of Frederikkeplassen

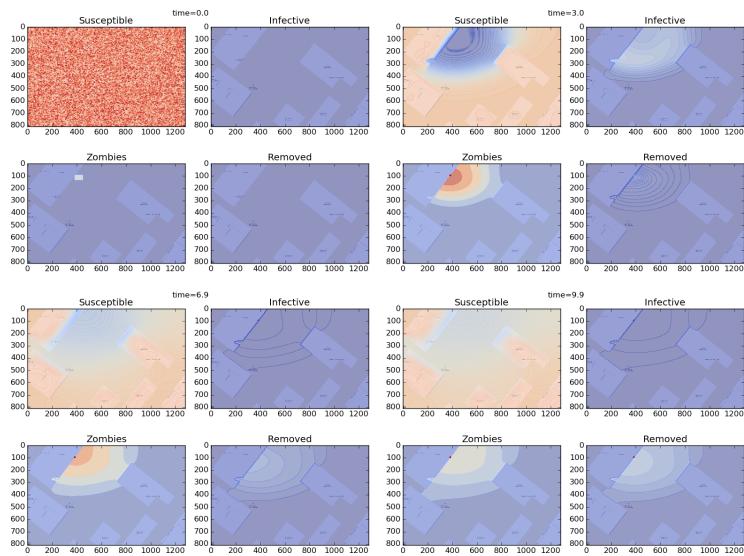


Figure 2.15: Random position of the *Susceptible* and *Zombie* group. Simulated with free areas inside the buildings.

2.4 Appendix

2.4.1 Sympy to find manufactured solution

```
>>> from sympy import *
>>> x,t,lam = symbols('x t lam')
>>> def s_simple(x,t):
...     return cos(pi*x)*t
...
>>> def i_simple(x,t):
...     return cos(pi*x)*t
...
>>> def r_simple(x,t):
...     return cos(pi*x)*t
...
>>> for x_point in 0,1:
...     print "s_x(%s,t): " , % x_point,
>>> for x_point in 0,1:
...     print "s_x(%s,t): " % x_point,
...     print diff(s_simple(x,t),x).subs(x,x_point).simplify()
...     print "i_x(%s,t): " % x_point,
...     print diff(i_simple(x,t),x).subs(x,x_point).simplify()
...     print "r_x(%s,t): " % x_point,
...     print diff(r_simple(x,t),x).subs(x,x_point).simplify()
...
s_x(0,t): 0
i_x(0,t): 0
r_x(0,t): 0
s_x(1,t): 0
i_x(1,t): 0
r_x(1,t): 0
>>> s = s_simple(x,t)
>>> i = i_simple(x,t)
>>> r = r_simple(x,t)
>>> f = diff(s,t)+i*s-diff(diff(s,x),x)
>>> print f.simplify()
(t**2*cos(pi*x) + pi**2*t + 1)*cos(pi*x)
>>> g = diff(i,t)-i*s+lam*i-diff(diff(i,x),x)
>>> print g.simplify()
(lam*t - t**2*cos(pi*x) + pi**2*t + 1)*cos(pi*x)
>>> h = diff(r,t)-lam*i-diff(diff(r,x),x)
>>> print h.simplify()
(-lam*t + pi**2*t + 1)*cos(pi*x)
```

2.4.2 Discretization

$$\begin{aligned} \frac{dS}{dt} &= \Sigma - \beta SZ - \delta_S S + \nabla(\gamma_s(x)\nabla S) \\ \frac{dI}{dt} &= \beta SZ - \varrho I - \delta_I I + \nabla(\gamma_I(x)\nabla I) \\ \frac{dZ}{dt} &= \varrho I - (\alpha + \omega(t))SZ + \zeta R + \nabla(\gamma_Z(x)\nabla Z) \\ \frac{dR}{dt} &= \delta_S S + \delta_I I - \zeta R + (\alpha + \omega(t))SZ \end{aligned} \tag{2.34}$$

The calculations will be shown for the diffusion part in the first equation. This idea will be used for the whole system

$$\begin{aligned} \frac{dS}{dt} &= \nabla(\gamma_s(x) \nabla S) \\ \frac{S_{i,j}^{n+1} - S_{i,j}^n}{\Delta t} &= \left(\gamma(x_{i+1/2,j}) \frac{S_{i-1,j}^n - 2S_{i,j}^n + S_{i+1,j}^n}{\Delta x} + \frac{S_{i,j-1}^n - 2S_{i,j}^n + S_{i,j+1}^n}{\Delta y} \right) \end{aligned} \quad (2.35)$$

2.4.3 2D Gaussian function from x=0

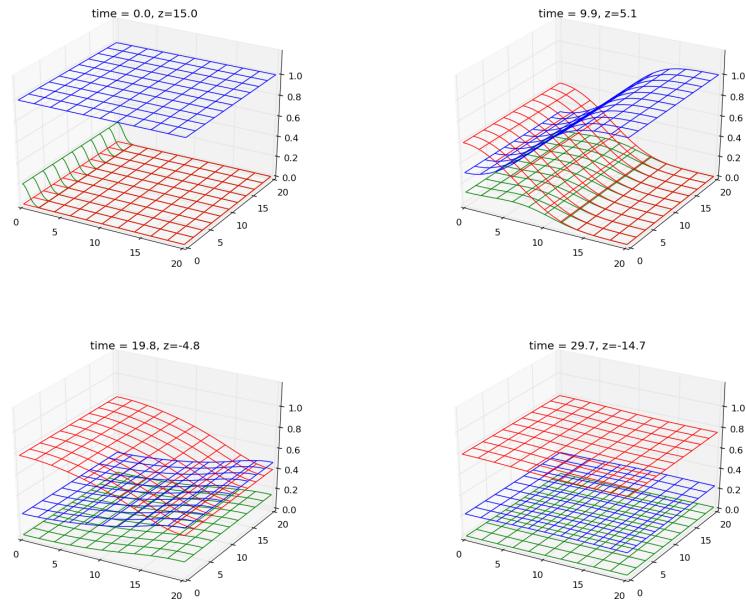


Figure 2.16: The PDE system (2.5) simulated for a 2D system with $\lambda = 0.5$. A gauss wave from $x = 0$.

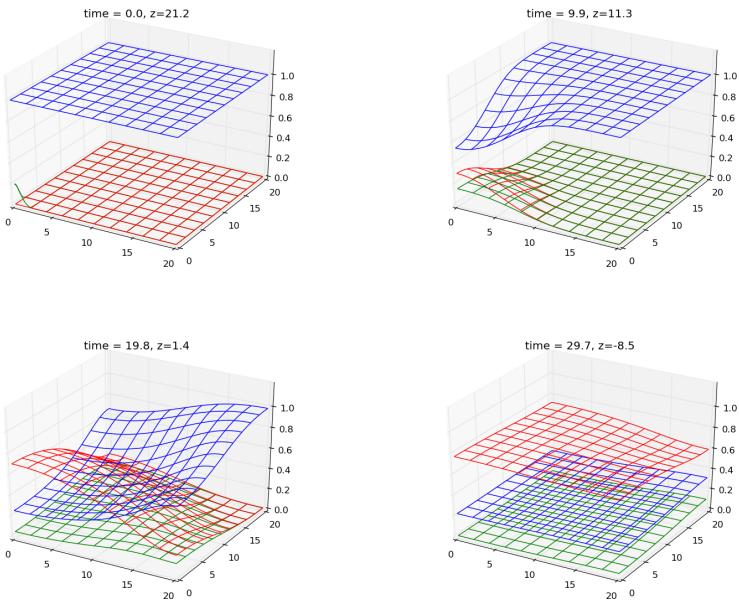
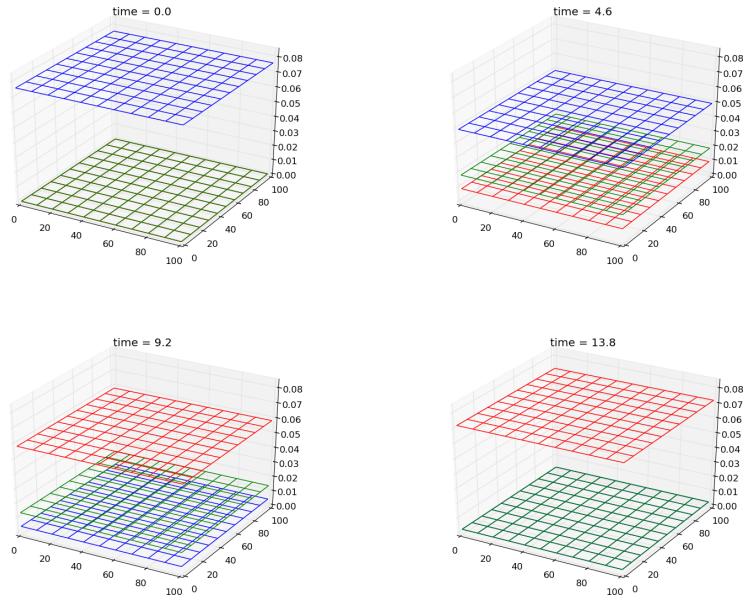
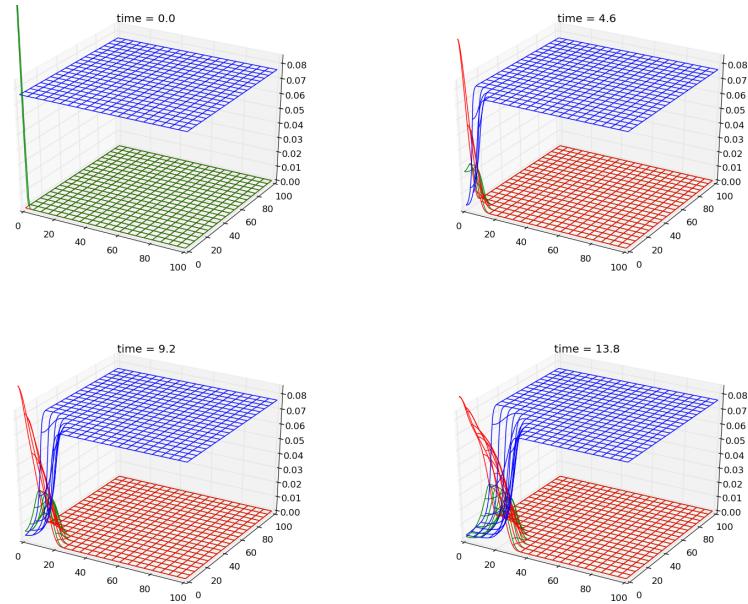


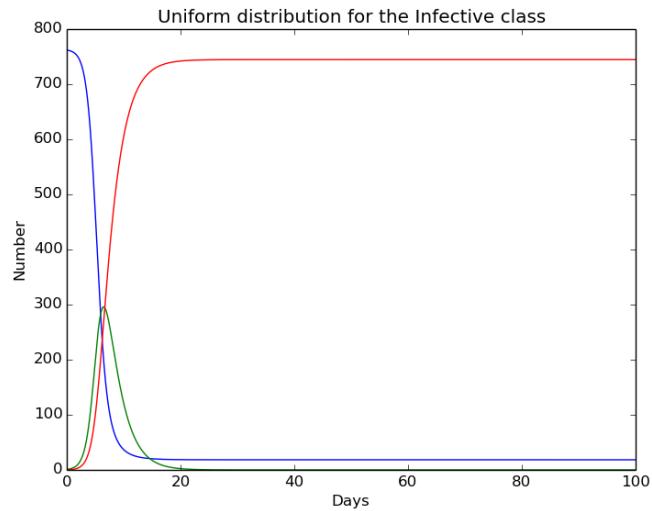
Figure 2.17: A gaussian function from $x = 0, y = 0$ based on the PDE system (2.5) with $\lambda = 0.5$

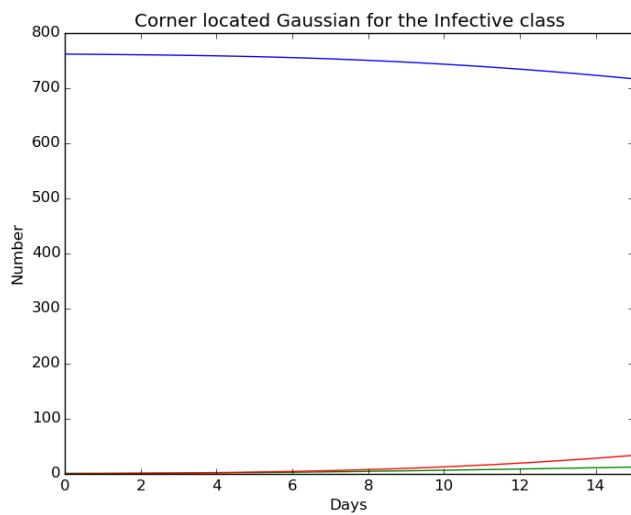
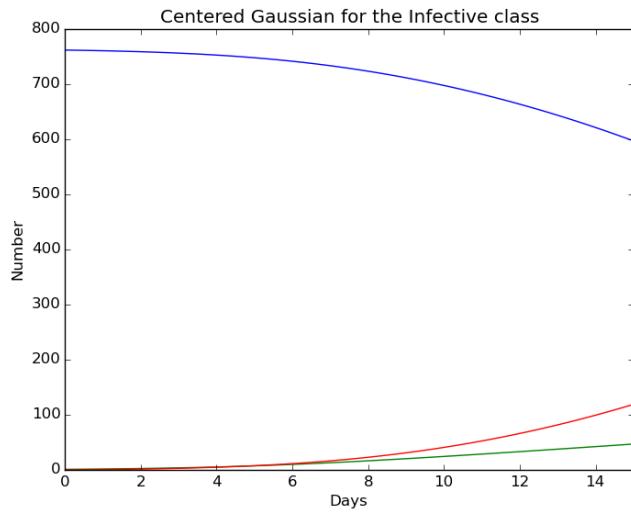
2.4.4 2D Gaussian function from $x=0, y=0$ **2.4.5 2D Gaussian function from $x=0, y=0$ with higher initial value****2.4.6 English Boarding School**

Gaussian from the corner.



A long simulation on 100 Days.





2.4.7 Zombiefication

Verify the initial phase.

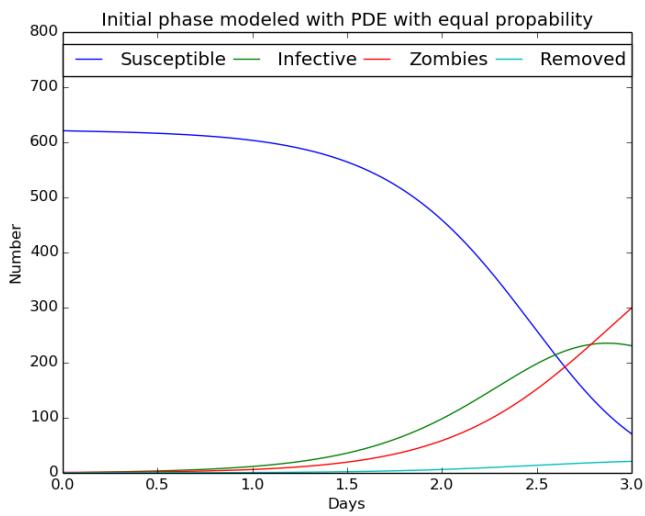
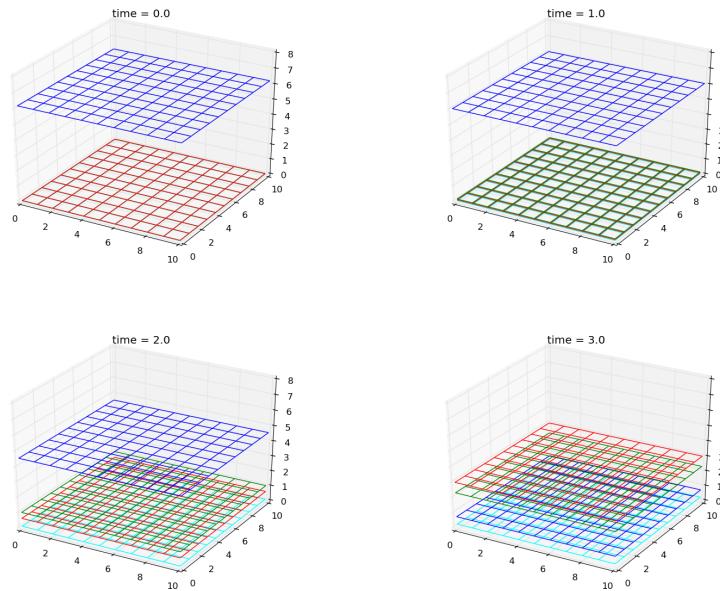
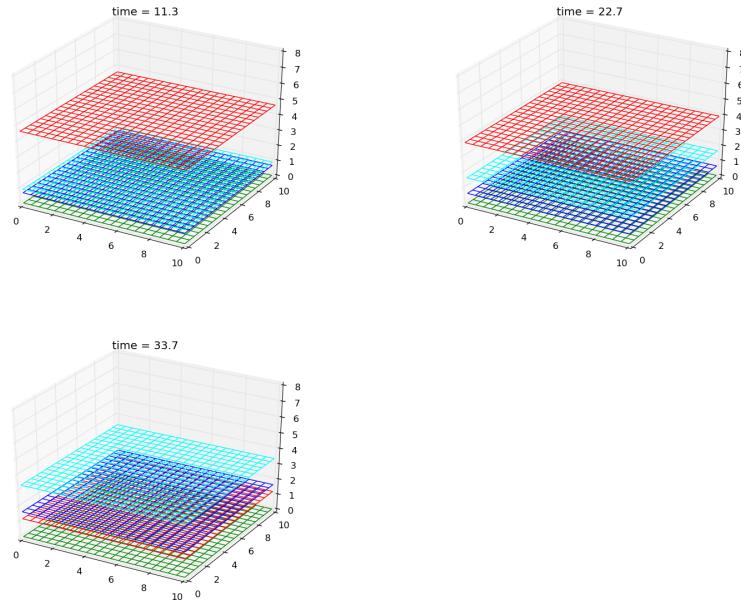
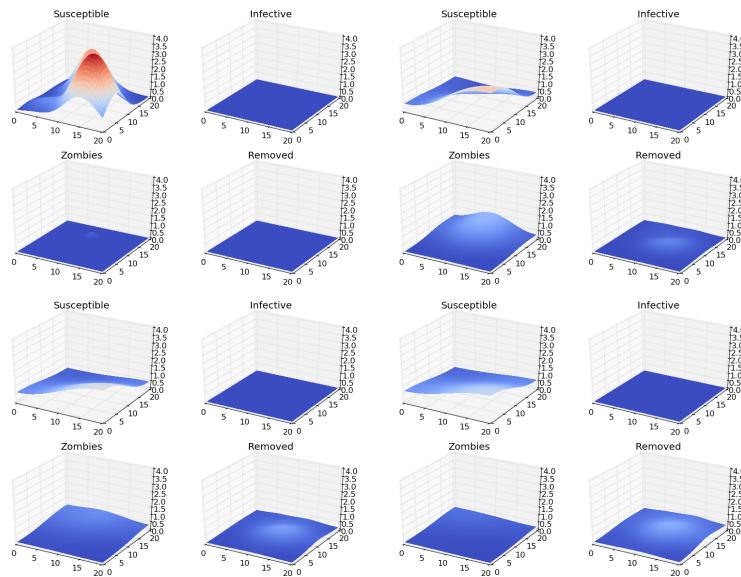


Figure 2.18: Creates the same results as for the ODE system

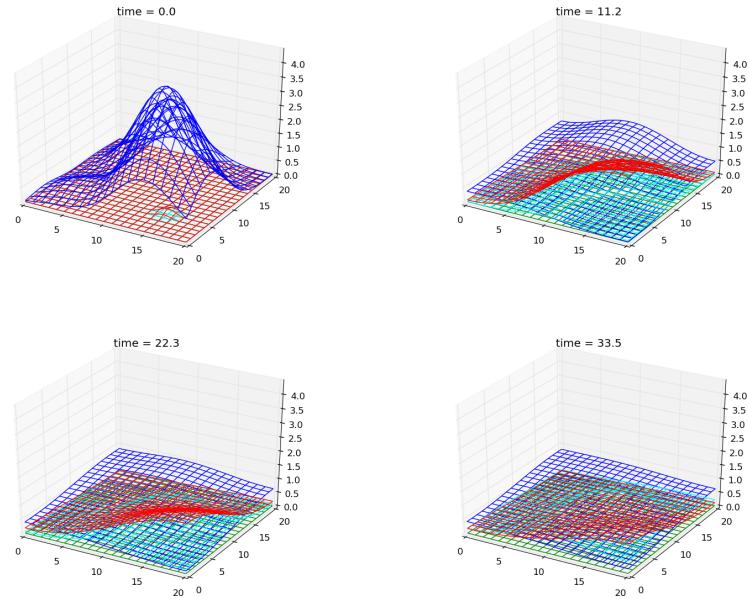
Three phases.



middle town.



large town.



Chapter 3

Random walk

The last chapter will study a third way to model epidemic disease. This will be done by using random walk. This technique quite different from the other models presented earlier, by using Monte Carlo simulations and probabilities instead of differential equations, which have been in focus earlier. The first section will be a study of Monte Carlo methods and Random walk based on the paper from M.H. Jensen Ref. [?]. The next sections will use the parameters from *English Boarding School* and *Walking Dead* to see if a Random walk system can expand the knowledge about epidemics. The model will be compared to the ODE system and PDE system from the previous chapters.

3.1 Monte Carlo methods

Techniques from Monte Carlo are widely used in several fields as chemistry, physics, medicine, biology and in finance Ref. [?]. These numerical methods can be seen in general terms as statistical simulations methods, which use random numbers to perform the simulations. The Metropolis algorithm is a central algorithm in this field, and is considered as one of the top ten algorithms during the last century Ref. [?]. A Monte Carlo strategy require four terms to be understood to use this method. These are:

- Random variable
- probability distribution functions (PDF)
- moments of a PDE
- the pertinent variance σ^2

3.1.1 Random variable

Random variable can be seen as stochastic variable, where the outcome cannot be presumed. Examples as tossing dice, flipping coins or gambling are based on this

principle. Although the outcome is unknown, knowledge about the probability and the range can be studied. The numbers in the *domain* for two dice are

$$\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

with the corresponding *probabilities* are

$$\{1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1\} \frac{1}{36}$$

By throwing two dice once, there is no guarantee that the result will be 7, though this has the highest probability. But by repeating this operation, the distribution would reflect the *probabilities* above. A stochastic variable can either be discrete or continuous, but will in both cases be denoted as capital letters, X, Y . A discrete example is the example above, where the domain is given with exact values, $x_1, x_2, x_3, \dots, x_n$. The continuous case can be seen as the probability in a given area. An example can be the distance from a dart to the center, after a random throw at a dartboard.

When using a computer to produce random number, they will in reality be pseudo random, since they have to be based on an algorithm. Therefore the choice of this algorithm is important. In this chapter, a generator of Marsaglia Ref. [?]

3.1.2 probability distribution functions (PDF)

The PDF is a function $p(x)$ on the domain that gives the probability or relative frequency for a outcome. In the discrete case, the function can be seen as

$$p(x) = Prob(X = x) \quad (3.1)$$

The PDF in the continuous is not able to directly depict the actual probability. The probability is instead defined as the density around x with an infinitesimal interval. This can therefore be seen as an integral, since it is the density of the probability rather than the probability Ref.[?]. This can be defined.

$$Prob(a \leq X \leq b) = \int_a^b p(x)dx \quad (3.2)$$

And by quoting M.H. Jensen *Qualitatively speaking, a stochastic variable represents the values of numbers chosen as if by chance from some specified PDF so that the selection of a large set of these numbers reproduces this PDF*. Ref.[?]. This sum up the relation between random variables and PDF. If this is not fulfilled, the group of stochastic variable does not fulfill the criteria for random numbers.

CDF- *cumulative probability distribution function*

There are two properties that the PDF must fulfill. The first one is the size of $p(x)$. This has to be in the interval $0 \geq p(x) \geq 1$, since the probability cannot

be negative or larger than 1 for an event to happen. The sum of all events has to be 1, both for discrete and continuous PDFs, and can be seen as follows

$$\sum_{x_i \in \mathbb{D}} \int_{x \in \mathbb{D}} p(x) dx = 1 \quad (3.3)$$

There are several distributions that are essential when looking at continuous PDFs. The two ones that will be used in this chapter are the uniform distribution.

$$p(x) = \frac{1}{b-a} \theta(x-a)\theta(b-x) \quad (3.4)$$

with:

$$\begin{aligned} \theta(x) &= 0, & x < 0 \\ \theta(x) &= 1, & x \geq 0 \end{aligned} \quad (3.5)$$

This distribution is natural to use, when a group of humans shall be evenly placed over an area. When comparing the ODE system from the first chapter and the uniformed distributed PDE system in previous chapter, Eq. (3.4) is natural to use. To get a correct estimate, it is important that the set of random numbers is large enough. Gaussian distribution is the second one, this is often called normal distribution and can be seen in Eq.(3.6)

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (3.6)$$

This will give the same distribution as the Gaussian function used in the previous chapter.

3.1.3 moments of a PDF

By define $h(x)$ as a arbitrary function, the *expectation value* can be written

$$\langle h \rangle_X \equiv \int h(x)p(x)dx \quad (3.7)$$

Here defined on the domain of the stochastic variable X with PDE $p(x)$. A more general way to write the expectation is by adding a power of, n , to the equation. This can now be seen as the *moments*. The n -th moment is defined

$$\langle x^n \rangle \equiv \int x^n p(x)dx \quad (3.8)$$

The value of n can be sat to zero. This result in $\langle 1 \rangle$ and creates a normalization condition of p . The first order is called *mean* and are often defined with a μ .

$$\langle x \rangle = \mu \equiv \int xp(x)dx \quad (3.9)$$

This represents the average value of PDF and is often called the expectation value of p Ref.[?].

3.1.4 The pertinent variance σ^2

Central moments is a special case of moments defined as

$$\langle (x - \langle x \rangle)^n \rangle \equiv \int (x - \langle x \rangle)^n p(x) dx \quad (3.10)$$

The first two central-moments are trivial and only result in 1 and 0, respectively for $n = 0$ and $n = 1$. But the second central-moment is more interesting to study. This is denoted as σ_X^2 or $\text{Var}(X)$, called the variance. This can be shown.

$$\sigma_X^2 = \langle x^2 \rangle - \langle x \rangle^2 \quad (3.11)$$

The square root of the variance, $\text{sigma} = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$ is called *standard deviation*. This is the deviation from the mean of PDF, and can be seen as the spread around the mean of PDF.

3.2 Random walks

The previous section explained the technical aspects of Monte Carlo, as the principle and algorithms for producing random numbers. This section will study the importance of a proper selection of variables and importance sampling. The challenge when model a Monte Carlo simulation, is the appropriate selection of random states. It is important that this match the probability distribution, PDF. This will be done through a Markov process, which is a random walk with a selected probability for making a move. A good reason to choose Markov process, is that this will reach equilibrium state, after certain number of simulations. This can first be shown for a simple diffusion equation, which can be expanded to the simple PDE system used in the previous chapter.

3.2.1 Diffusion equation and random walks

The British Botanist R. Brown developed a theory by studying how pollen dispersed in water. This idea is called Brownian motion and can be used to describe a group of particles spreading. The function $w(x, t)dx$ can be defined as the probability of finding a given number of particles in an interval with size dx . This function is the *PDF* as explained in the section above. The flux of particles that passing a point x , can be described by $j(x, t)$. This flow is proportionally with the gradient of the PDF, $j(x, t) = -D \frac{\partial w(x, t)}{\partial x}$ (3.12) D is here the diffusion constant. This example can be seen as a closed experiment, where the concentration is conserved. The relation between the flux and the *PDF* can also be expressed

$$\frac{\partial j(x, t)}{\partial x} = -\frac{\partial w(x, t)}{\partial t} \quad (3.13)$$

And a diffusion equation can be expressed based on these two equations.

$$\frac{\partial w(x, t)}{\partial t} = D \frac{\partial^2 w(x, t)}{\partial x^2} \quad (3.14)$$

The expectation value and the variance, σ^2 , can be studied. The expectation value for the function $f(x, t)$ can be defined as

$$\langle f(x, t) \rangle = \int_{-\infty}^{\infty} f(x, t) w(x, t) dx \quad (3.15)$$

The demands defined for the PDF in the previous section have to be fulfilled for $w(x, t)$. The normalization condition which is defined for this PDF require constraints for equation.

$$w(x = \pm\infty, t) = 0, \quad \frac{\partial^n w(x, t)}{\partial x^n}|_{x=\pm\infty} = 0 \quad (3.16)$$

This can be used to study the time derivative of the expectation value. The diffusion equation derived earlier can be used here.

$$\frac{\partial \langle f(x, t) \rangle}{\partial t} = \int_{-\infty}^{\infty} f(x, t) \frac{\partial w(x, t)}{\partial t} dx = D \int_{-\infty}^{\infty} f(x, t) \frac{\partial^2 w(x, t)}{\partial x^2} dx \quad (3.17)$$

By using integration by parts on the right side, the equation can be expressed,

$$\frac{\partial \langle f(x, t) \rangle}{\partial t} = D f(x, t) \frac{\partial w(x, t)}{\partial x}|_{x=\pm\infty} - 2D \int_{-\infty}^{\infty} \frac{\partial w(x, t)}{\partial x} dx \quad (3.18)$$

By using the demands from (3.16), the $\langle f(x, t) \rangle$ is independent of time.

$$\frac{\partial \langle f(x, t) \rangle}{\partial t} = 0 \quad (3.19)$$

Implying that if a simulation of a random walk with equal probability of jumping in each direction, will end up with the probability distribution centered around the initial position. This may not be the case for the variance. By using (3.18), the $\langle x^2 \rangle$ can be found.

$$\frac{\partial \langle x^2 \rangle}{\partial t} = D x^2 \frac{\partial w(x, t)}{\partial x}|_{x=\pm\infty} - 2D \int_{-\infty}^{\infty} x \frac{\partial w(x, t)}{\partial x} dx \quad (3.20)$$

Here, integration by parts can be used. This results in

$$\frac{\partial \langle x^2 \rangle}{\partial t} = D x w(x, t)|_{x=\pm\infty} + 2D \int_{-\infty}^{\infty} w(x, t) dx = 2D \quad (3.21)$$

this leads to

$$\langle x^2 \rangle = 2Dt \quad (3.22)$$

which gives the variance

$$\langle x^2 \rangle - \langle x \rangle^2 = 2Dt \quad (3.23)$$

And the square root of the variance can be expressed,

$$\sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{2Dt} \quad (3.24)$$

By comparing this with the displacement of a free particle, which moves with the function $x(t) = vt$ from the initial point, will the diffusion process move with $\sqrt{\langle x^2 \rangle - \langle x \rangle^2} \propto \sqrt{t}$. This can be used to describe a random walk, and could say that a random walker escapes much more slowly than a free particle from the initial time. This can be seen in an example.

3.2.2 Random walker

Now the random walker can be introduced in 1D. This can either jump to the left or the right with a length $\Delta x = l$. It is equal probability for both directions. $L = R = 1/2$. Then the average displacement will be

$$\langle x(n) \rangle = \sum_i^n \Delta x_i = 0, \quad \Delta x_i = \pm l, \quad (3.25)$$

after n jumps. The variance can be found by first finding $\langle x(n)^2 \rangle$.

$$\langle x(n)^2 \rangle = \left(\sum_i^n \Delta x_i \right) \left(\sum_j^n \Delta x_j \right) = \sum_i^n \Delta x_i^2 + \sum_i^n \Delta x_i \Delta x_j = l^2 n \quad (3.26)$$

The last term here will disappear after enough steps.

$$\sum_i^n \Delta x_i \Delta x_j = 0 \quad (3.27)$$

This gives the variance

$$\langle x(n)^2 \rangle - \langle x(n) \rangle^2 = l^2 n \quad (3.28)$$

Now this variance from a random walker can be coupled with the variance from the diffusion equation in the section above. By setting $n = t/\Delta t$. The random walker gets the following variance

$$\langle x(n)^2 \rangle - \langle x(n) \rangle^2 = l^2 \frac{t}{\Delta t} \quad (3.29)$$

Then the diffusion constant in (3.23) can be replaced by

$$D = \frac{l^2}{\Delta t} \quad (3.30)$$

And the variance between these can be compared.

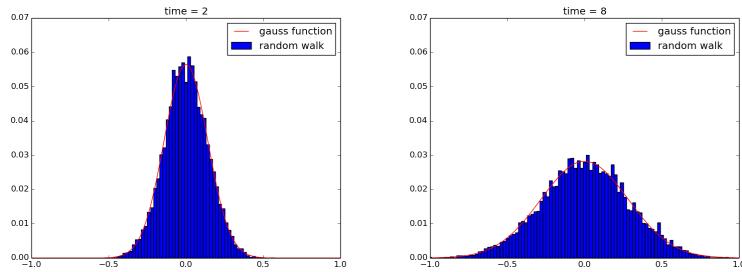


Figure 3.1: 10000 random walkers placed at $x=0$ at $t=0$. Every step has the length $\Delta x = 0.01$ and with a random step every $\Delta t = 0.01$.

The standard deviation can be found for this simulation. This is given by the square root of the variance

$$\sigma = \sqrt{l^2 \frac{t}{\Delta t}} \quad (3.31)$$

To study if the average random walker develops as expected, a table can with the outputs can be added.

	time=2	time=8
average displacement	0.000662	-0.002416
standard deviation(SD)	0.1414	0.2828
percent inside SD	70.96 percent inside gauss	68.26 height

The average displacement and the standard deviation shows that a group of random walkes spread similar as a standard diffusion function. This result in a major group of random walkers. In the simulation in Fig.(3.1), 10000 random walkers are used. By increasing the amount om random walkers, the precision will be better.

3.3 English Boarding School

This example has been common for all three systems, and will be used in this chapter. The chance of getting infected requires a meeting between infected person and a susceptible person..A random walker will after enough steps cover the whole area. A simulation is done for a student with a random position at initial time. 1000 random steps are performed every day, which results in a step every 90 seconds. The step length is set to 5.7024 m, and is based on the average distance a person walks every day. The simulation is performed for 15 days, which results in 15000 random steps. The size of the schoolyard is sat to 100 m x 100 m, and the disease can spread within a distance of 5 meters.

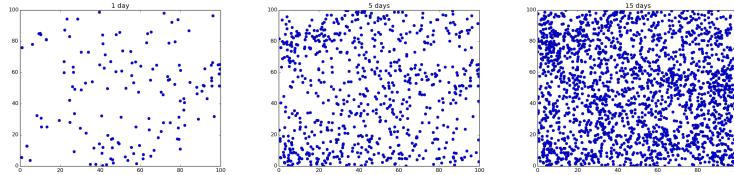


Figure 3.2: The positions a random walker has covered in 1,5 and 15 days. A random step with lenght 3.96 m is performed every minute. The positions are plotted for every ten minutes.

Fig.(3.2) shows that a random walker will be distributed over the area after enough steps. The students in the school are divided into three groups. The first group consists of susceptible students, and this group is at risk of getting infected. This group is described by S . The second group consists of infected

students. The group is described by I . The last group consists of students who are immune to the disease. This group is described by R . The total number of students is $N = 763$. The initial values are: $S_0 = 762$, $I_0 = 1$ and $R_0 = 0$. There are two parameters that are used in the simulation. The first parameter r , describes the gain of I from the susceptible group, S . This is rate proportional to the number of susceptible and infective and is given by rSI . The second parameter a describes the rate of removal of infective to the removed group, R . These two parameters are given by $r = 2.18 \cdot 10^{-3}$ and $a = 0.44036$. The simulation is based on the ODE system given in *Epidemic models* :

$$\begin{aligned}\frac{dS}{dt} &= -rSI \\ \frac{dI}{dt} &= rSI - aI \\ \frac{dR}{dt} &= aI\end{aligned}\tag{3.32}$$

The parameters r and a must be adapted to the random walk simulation. The parameter r is used in

$$rSI\tag{3.33}$$

is based on all possible combinations of S and I . This will not be the case in a random simulation. The meetings in a random simulation depends on the number of random walkers and the possibility for a meeting. If the possibility is small, the susceptible student has to be close to an infected student to be infected. The number of meetings during one time step can be used to adjust the parameter. The following term has to be fulfilled:

$$r_{random} NoM_0 = rS_0 I_0\tag{3.34}$$

Here NoM_0 is a constant value and represents number of meetings between the susceptible group and the infected group at initial time. This can be found by a numerical simulation of the random walkers. The number of meetings for the infected student during one day is simulated for 1000 rounds, the average value is $NoM = 1905.223$. Now Eq.(3.35) can be rewritten and r_{random} can be expressed by known values:

$$r_{random} = \frac{rS_0 I_0}{NoM_0}\tag{3.35}$$

The parameter r_{random} is now used to calculate the risk of getting infected in a meeting between a susceptible student and an infected student. The value of a also has to be adjusted. This parameter is only affected by the time. The random simulation is performed for 1000 random steps during one day. This means that there will be 1000 chances of being immune to the disease during 1 day, since a random number will be drawn every random step. A person can only get immune one time, therefore it is more interesting to see if a person stays infected during 1 day. All other results will cause an immune student. The chance of staying

infected during one day is $In = 1 - a$, which results in $In = 0.55964$. The chance of staying infected for 1000 time steps during one day demands the following percent for each time step: $In^{\frac{1}{1000}} = 0.9994197$. Since all other solutions gives an immune student, the percent can be set to: $a_{random} = 1 - In^{\frac{1}{1000}} = 0.00058$. A series of simulations can be done with these parameters. The average solution will be plotted against the solution from the ODE system given in Eq.(3.32). The standard deviation for the random simulations is given.

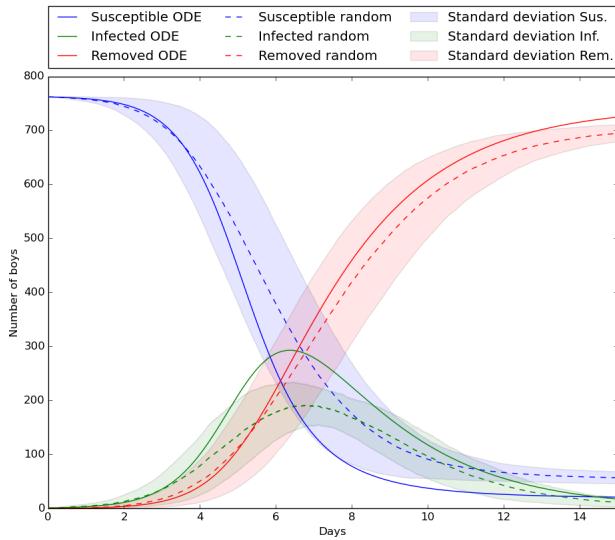


Figure 3.3: Random walk compared to an ODE simulation of Eq.(3.32). The random function is shown with a dashed line, with the standard deviation shown as the colored area around the dashed line. The random function is based on the average of a series of simulations.

3.3.1 A Lower average value

By studying the results in Fig.(3.3), one can see that the values of the random walk simulation is lower than for the ODE. The difference is low in the beginning of the simulation. This changes when the *Infected* group increases. This results in an increase of the combination between the *Susceptible* group and the *Infected* group. In an ODE simulation, this will lead to a rapid increase of the *Infected* group. Since the transformation for a student from the *Susceptible* group to the *Infected* group only requires one successful meeting, where successful is seen as the transmission of the disease. There will be no difference in the result if the transmission of the disease happens once or several times during one time step. This is reasonable to be the case when the amount of *Infected* increase.

3.3.2 Threshold value

In the previous chapters, the threshold value was found for the epidemic systems. The reproduction rate could be used to check if the disease would develop into an epidemic disease. The reproduction rate can be seen in Eq.(3.36)

$$R_0 = \frac{rS_0}{a} \quad (3.36)$$

If $R_0 > 1$ was fulfilled, an epidemic situation would occur. The result will be $R_0 = 3.77$ with the parameters from the ODE simulation. When simulating the random walk, only 56 percent of the simulations resulted in an epidemic disease. 44 percent resulted in a transmission of the infected student to the *Removed* group, before the student was able to infect other students. This simulations were performed on a small group, and the results may differ in larger groups. A removal rate above one will not necessarily lead to an epidemic disease, if the group is small enough.

3.4 Zombiefication

The ODE system given in chapter *Epidemic models* will be used for this simulation. This can be seen in Eq.(3.37). The parameters has to be adjusted for this simulation, similar as shown for the English boarding school. Frederikkepllassen at Blindern will be used as the area where the simulations will be done. The area is estimated to be 100m x 100m and the disease will be able to spread if the distance is closer than 4 meters. There will be done three different simulations in this section, where the human intelligence will be taken into account. The total time of the simulations will be set to 10 minutes, and there will be performed 100 random walks each minute.

$$\begin{aligned} \frac{dS}{dt} &= \Sigma - (\beta + \mu\omega(t))SZ - \delta_S S \\ \frac{dI}{dt} &= (\beta + \mu\omega(t))SZ - \varrho I - \delta_I I \\ \frac{dZ}{dt} &= \varrho I - (\alpha + \omega(t))SZ + \zeta R \\ \frac{dR}{dt} &= \delta_S S + \delta_I I - \zeta R + (\alpha + \omega(t))SZ \end{aligned} \quad (3.37)$$

As similar for the English boarding school, the parameters in Eq.(3.37) have to be adjusted. The parameters from the two first phases in *Walking Dead* will be used. These can be seen in the table below. The number of meetings per minute is sat to $NoM = 98.64$, based on the average of 300 timesteps. This is used to find the value of β and α , similar to the method shown in the previous section. The value for ϱ has been adjusted by first finding the average incubation time for the *Infected* group. This has been done by setting $1/\varrho$. The average incubation time is 0.72 minute. This is really fast, and is not a realistic

number, based on the TV Series *Walking Dead*. By numerical simulations, the value of ϱ_r can be set to 0.0137 for each simulation. This will result in a average incubation time of 0.72 minute. The same can be done for the value in the hysterical phase.

parameter	Initial phase	hysterical phase
β	0.01155	0.000011
β_r	0.07271	0.000693
ϱ	1.37	1.5
ϱ_r	0.0137	0.015
α	0.00044	0.000208
α_r	0.00277	0.001309

These parameter values are used for all simulations. This section will focus on the ability to transfer the human intelligence to a epidemic model. The previous chapter *Geographic models* introduced the spatial effect and the ability for humans to seek safe areas. This chapter will introduce different conditions for the interactions based on the phase that the groups are in. The section will use the area used in *Ten minutes on Frederikkepllassen* in *Geographic models*.

3.4.1 Random walk

Random walk will be the first condition the different groups will hold. This result in a smooth distribution of the whole group, and the spatial of a group of humans will work as for a diffusion equation. The simulation will be done for ten minutes at Frederikkepllassen. To get a better picture of the simulation. The walkers are represented with different images. These four figures can be seen in Fig.(3.4).



Figure 3.4: Figures used in the simulation. All seen from the sky. I: The walkers in the *Susceptible* group can be seen as humans with green sweater. II: The walkers in the *Infected* group can be seen with a red and green sweater, with one arm in front. III: The walkers in the *Zombie* group can be seen with a white sweater and both arms in front. IV: the walkers in the *Removed* group can be seen as a tombstone.

The walkers are at initial time placed out on Frederikkepllassen. The initial positions of the walkers can be seen in subplot I in Fig.(3.5).

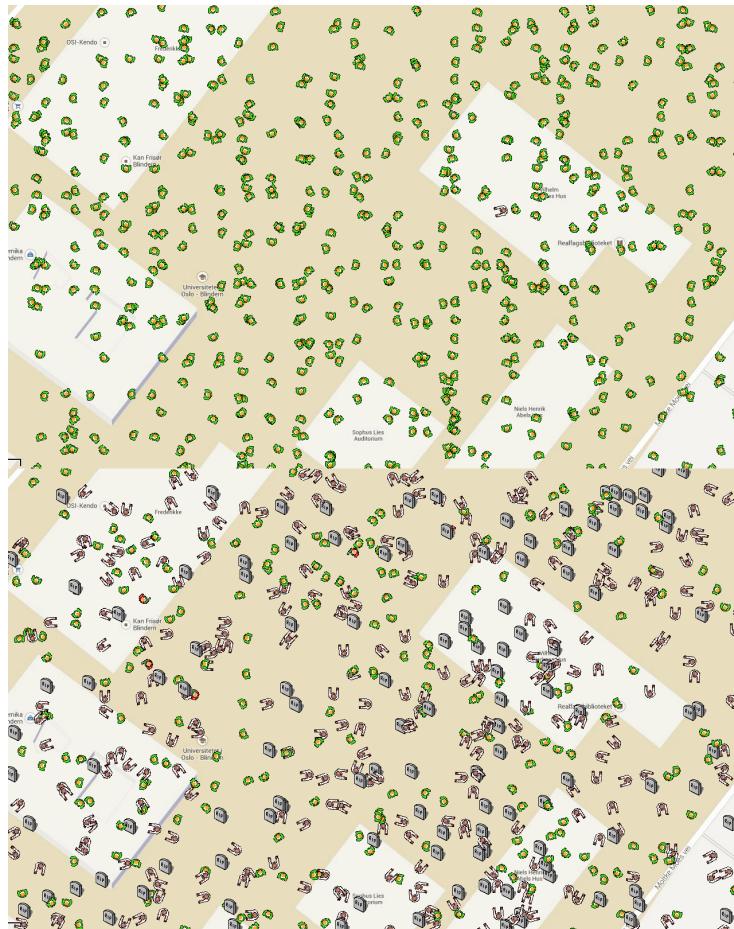
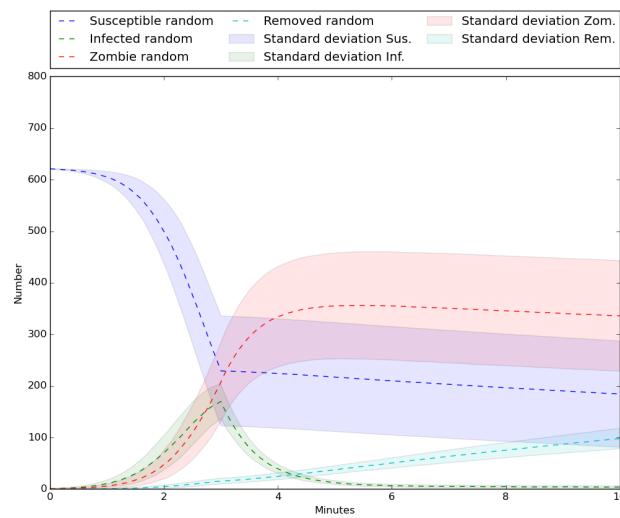


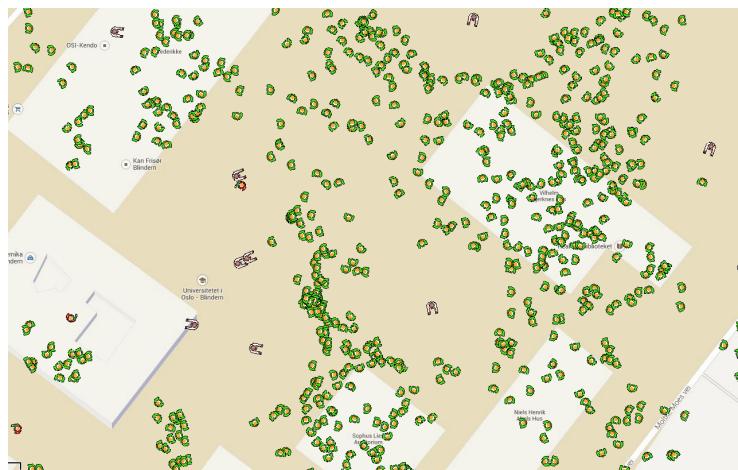
Figure 3.5:

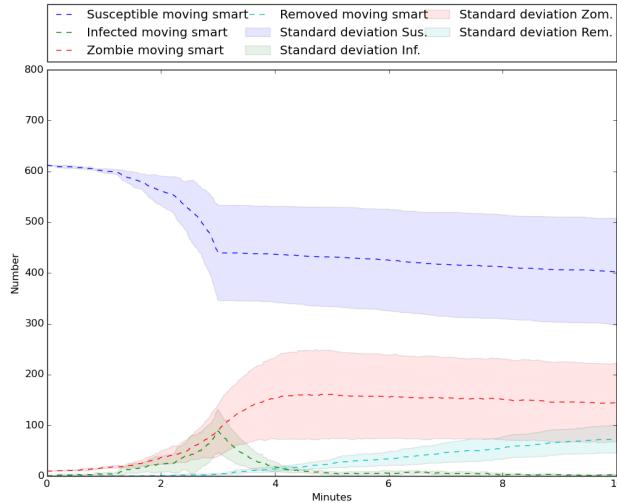


	Uniform distribution PDE	Uniform distribution Random walk
3 Minutes		
Susceptible	72.23	229.05
Infected	229.65	169.73
Zombie	296.67	207.83
Removed	20.84	15.38
7 Minutes		
Susceptible	70.78	203.07
Infected	0.83	4.95
Zombie	498.72	350.62
Removed	49.12	63.35
10 Minutes		
Susceptible	69.69	184.2
Infected	0.25	3.94
Zombie	479.00	335.99
Removed	70.55	97.87

3.4.2 Moving smart

Next phase is based on the pattern that would be naturally to

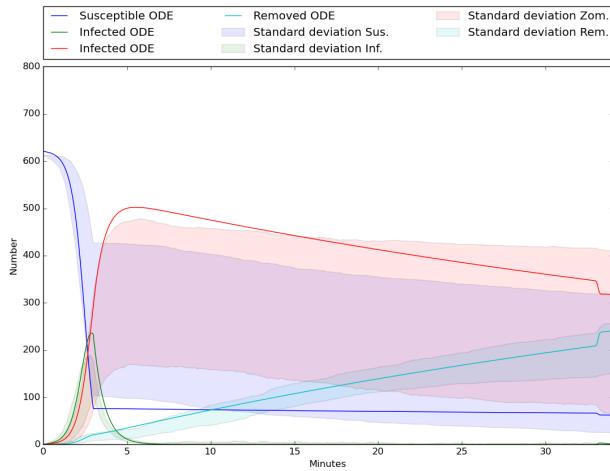




	Random walk mode	Moving smart mode
3 Minutes		
Susceptible	229.05	439.67
Infected	169.73	88.67
Zombie	207.83	89.67
Removed	15.38	4.0
7 Minutes		
Susceptible	203.07	417.0
Infected	4.95	5.0
Zombie	350.62	153.66
Removed	63.35	46.33
10 Minutes		
Susceptible	184.2	402.0
Infected	3.94	2.67
Zombie	335.99	144.67
Removed	97.87	72.67

3.4.3 Three phases in Walking Dead

A natural idea is to use different modus in different phases. The first phase, *Initial phase* will include a random walk mode. Here



	ODE	Random walk
<hr/>		
Initial phase		
Susceptible	71	265
Infected	231	139
Zombie	299	192
Removed	21	16
<hr/>		
Hysterical phase		
Susceptible	62	172
Infected	0	1
Zombie	359	250
Removed	202	189
<hr/>		
Counter attack		
Susceptible	58	171
Infected	1	1
Zombie	332	238
Removed	231	203

3.5 Discussion and conclusion

This is left to write

Bibliography

- [1] F. Darabont and G. A. Hurd. Walking dead: Season 1, 2010-2011.
- [2] Morten Hjorth-Jensen. Computational physics. *Lecture notes*, 2011.
- [3] Valerie Isham. Mathematical modelling of the transmission dynamics of hiv infection and aids: a review. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 5–49, 1988.
- [4] Hans Petter Langtangen, Kent-Andre Mardal, and Pål Røtnes. *Escaping the Zombie Threat by Mathematics*, chapter 3.6, pages x–y. University of Chicago Press, 2013.
- [5] Donna L.Hoyert and Jiaquan Xu. Deaths:preliminary data for 2011. Technical report, Division of Vital Statistics, 2012.
- [6] George Marsaglia. Xorshift rngs. *Journal of Statistical Software*, 8(14):1–6, 2003.
- [7] Joyce A. Martin, Brady E. Hamilton, Michelle J.K. Osterman, Sally C. Curtin, and T.J. Mathews. Births: Final data from 2012. Technical report, Division of Vital Statistics, 2013.
- [8] Philip Munz, Ioan Hudea, Joe Imad, and Robert J. Smith. When zombies attack!: Mathematical modelling of an outbreak of zombie infection. *Infectious Disease Modelling Research Progress*, 4:133–150, 2009.
- [9] J.D. Murray. *Mathematical Biology: I. an Introduction*. Interdisciplinary Applied Mathematics. Springer, 2002.
- [10] J.D. Murray. *Mathematical Biology II: Spatial Models and Biomedical Applications*. Intercisciplinary Applied Mathematics: Mathematical Biology. Springer, 2003.
- [11] Alan S. Perelson and Patrick W. Nelson. Mathematical analysis of hiv-1 dynamics in vivo. *SIAM review*, 41(1):3–44, 1999.
- [12] Yanni Xiao, Sanyi Tang, Yicang Zhou, Robert J. Smith, Jianhong Wu, and Ning Wang. Predicting the hiv/aids epidemic and measuring the effect of mobility in mainland china. *Journal of Object Technology*, 317:271–285, 2013.

