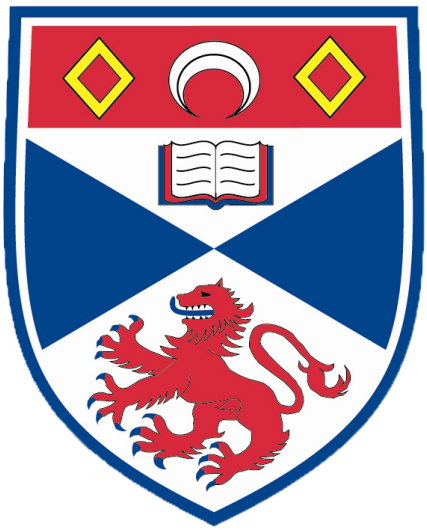
A New Ebola Virus Disease Compartmental Model for Predicting Success of Safe Funerals



University
of
St Andrews

Department of Computer Science and Mathematics
CS4796: Joint Project

Author: Tatiana Tay
Supervisors: Prof. Simon Dobson, Prof. Mark Chaplain

April 4, 2016

Abstract

In the recent years, Ebola virus disease has become a serious threat to people in West African countries. However, the only epidemiological action taken to improve the situation was an attempt to eliminate traditional African funerals where mourners could get bursts of infection from a recently dead person. To measure the effectiveness of this step the new compartmental model was presented together with the software that runs simulations on networks. The model was called SEIDR and it incorporates burial ceremonies into the SEIR model that is traditionally applied when studying Ebola virus disease. Various disease parameters were considered and the results obtained show that the rate at which the infection spreads indeed increases significantly with the transition from SEIR to SEIDR model which justifies the introduced practise of safe funerals in West African countries. However, the same results might not hold in the future if the contact rate between infected and susceptible people grows or the disease mutates in such a way that the transmission probability increases.

Declaration

I declare that the material submitted for assessment is my own work except where credit is explicitly given to others by citation or acknowledgement. This work was performed during the current academic year except where otherwise stated.
The main text of this project report is 11777 words long, including project specification and plan.
In submitting this project report to the University of St Andrews, I give permission for it to be made available for use in accordance with the regulations of the University Library. I also give permission for the title and abstract to be published and for copies of the report to be made and supplied at cost to any bona fide library or research worker, and to be made available on the World Wide Web. I retain the copyright in this work.
If there is a strong case for the protection of confidential data, the parts of the declaration giving permission for its use and publication may be omitted by prior permission of the Honours Coordinator.

Contents

Chapter 1

Introduction

Ebola virus disease was introduced in 1976 during two synchronous outbreaks in Sudan and Democratic Republic of the Congo. The most recent one was registered in March 2014 and is considered to be the largest flare since it first appeared. During that outbreak, more people became ill and died than in all the others combined. The disease also spread across countries as it originated in Guinea and propagated to Sierra Leone, Liberia, Nigeria and Senegal (World Health Organization, 2016b).
The disease caused by Ebola virus is transmitted via close contact "(through broken skin or mucous membranes) with the blood, secretions, organs or other bodily fluids of infected people, and with surfaces and materials (e.g. bedding, clothing) contaminated with these fluids". The Airborne transmission of the virus does not occur (World Health Organization, 2016b).
The definitive diagnosis of Ebola virus disease can only be made in the laboratory on the basis of a number of different tests. If the results of those tests are positive, patients are referred to the confirmed cases. Otherwise, if it is not possible to take samples for research individuals are regarded as either probable cases if they were examined by clinical specialist or suspected cases if they had a sudden body temperature rise and they showed, at least, three symptoms that are associated with the disease or if they had an unexplained bleeding or sudden inexplicable death. Furthermore, anyone who had an unexpected rise of temperature who also had a contact with an alive or dead individual whose diagnosis was confirmed or considered probable or with an animal which has the disease would also be referred to suspected cases.
The World Health Organisation is working hard to identify the cases early and provide proper treatment to those who need it. They invest people and resources to various programs including social education and outreach programs to ensure better practises with regard to the disease. However, the situation is not getting much better as new cases appear.
The basic case reproduction number R_0 shows how infectious a disease is and for the Ebola virus disease it is not very high, which is why it was surprising that the disease spread so quickly in some West African countries. Therefore, the hypothesis was suggested that the reason for such rapid spread might lie in traditional African burial ceremonies. A human is most infectious once passed away and all the people who gather together at funerals are at very high risk especially since they initiate close contact with the dead body. Furthermore, when investigating what the people on the ground did it was found that they provided a huge amount of support for people who were infected. But the only thing that they did epidemiologically was that they stopped funerals.
The hypothesis is looked into using, firstly, mathematical modelling by adding another compartment to the traditionally used SEIR model that represents individuals who recently died. At the second stage, the elaboration of the hypothesis is tested. The idea is that there is this social network of people, somebody dies and when they die the network rewires, so lots of their friends connect to them at their funeral and they are highly contagious and all those people get bursts of potential infection from the recently dead person and then the network rewires back out again. So at this stage, the aim is to verify using simulations whether this network rewiring would change the rate at which the disease progresses. Finally, the experiments using simulation software are conducted and obtained results analysed and discussed.

Chapter 2

Ethics

The project does not raise any ethical issues. People were not surveyed, observed or interviewed, neither were they asked to test the software. Furthermore, there was no data collection and the data used was taken from available for public researches.
Despite the fact that some of the parameters used in simulations to model the spread of the Ebola virus disease were taken from the research that focused on the situation in Liberia, the software cannot be used to assess the dynamic of the disease spread there since the parameters were only used as a starting point, the observations and estimates in the research were based on the data collected in 2014 and might not correspond to the current situation there and finally there are many assumptions that underlie the simulation process.

Chapter 3

Context Survey

Mathematical Epidemiology

"Epidemiology is the study of the distribution and determinants of health-related states or events (including disease), and the application of this study to the control of diseases and other health problems". Epidemiological research can be divided into a descriptive and analytical examination. The former is used to assess the spreading among populations at a particular time and location while the latter is applied to determine what causes the emergence of a disease or a health problem (World Health Organization, 2016a). Some sources also distinguish the third group in epidemiology which involves experimental studies. Its aim is to identify risk factors and eliminate or reduce the possibility of disease occurrence by managing those factors. Experimental epidemiology unlike descriptive or analytical is not limited to observations only but also involves manipulation of influence factors (Brachman, 1996). However, in order to carry out those experiments, researchers need to provide evidence that the approach they are planning to proceed with will lead to positive results. Otherwise running the experiments can turn out to be a waste of time, resources and money. Therefore, researchers often try and predict the outcomes before actually initiating trials in which case mathematical epidemiology could provide a good theoretical basis.
A widely used tool that is applied in mathematical epidemiology to describe the process of disease progression is compartmental models. These models represent epidemiological process as a varying number of individuals in one of the discrete states. Nevertheless, as the number of factors that need to be taken into account grows, so does the number of states and corresponding parameters which significantly complicates models. Thus, compartmental models that are used follow relatively simple rules of individual behavior and are forced to make a number of assumptions.
Analysis of existing deterministic and stochastic models shows that they most often include the following states: 'Susceptible' (S), 'Infected' (I) and 'Recovered' or 'Removed' (R). Sometimes when more details need to be depicted more states are added, for example, the carrier state or the exposed state when a latent period is taken into account. The epidemiological situation is then described by a system of equations which depending on the type of model defines the transition dynamic of individuals from one state to another. The main parameters in the system of equations are the contact rate between infected and susceptible individuals, the rate of recovery and in some cases the rate of immunity loss as they determine epidemic evolution (Grishunina et al., 2014).
It is important to note that it is rarely possible to solve the system of equations analytically which complicates the analysis of how model parameters affect the epidemiological picture (Grishunina et al., 2014).

Simulations on Networks

One of the main assumptions that underlines classical mathematical models is that the considered population is fully mixed which means that every individual meets every other individual or in other words every possible interaction that could happen does happen. This could be plausible when a small number of individuals is viewed, but is not realistic for large populations. Thus, another approach was also examined where interactions between individuals are described using networks.
In the network version, a social network is built where nodes represent people and edges reflect contacts between those people. Therefore, the population is no longer fully mixed as interactions are only allowed between neighbouring nodes (Wittena, Poultera, 2005).
However, another problem arises when replacing mathematical models with networks. The population

structure is very complex and capturing it in details is time-consuming and impractical. Depending on what the network is used for, the structure of it would vary (Bansal, Grenfell, Meyers, 2007). For example graph or network structures would be different if a sexually-transmitted disease or an airborne disease are modelled. Thus, simplified or idealised versions needed to be created that would capture the important properties and leave the rest out (Keeling, Eames, 2005).

Complex Networks in Epidemiology

There are several graphs that are generally used when modelling epidemics on networks. All of them are simplified but have some useful properties. Their main distinction lies in the geographical distribution of individuals (nodes) and in the way graphs are constructed. These main features are described for the most common graphs in epidemiology.

Random Network

One of the oldest and most used random graphs is the Erdős–Rényi (ER) model.
Suppose a set $V = \{1, 2, ..., n\}$ is given, whose elements are called vertices. This set is then used to build a random graph on. Directed graphs, graphs with multiple edges or loops are not considered here, therefore, a random network can have at most C_n^2 edges. Any two vertices i and j are connected with probability $p \in [0..1]$ regardless of the rest C_{n-1}^2 pairs of vertices. In other words, the standard Bernoulli scheme is applied and a set of edges E is produced. As a result, an Erdős–Rényi random graph $G = (V_n, E)$ is created with a roughly Poisson degree distribution (Keeling1, Eames, 2005).
In the real world, however, this degree distribution would be different as some people are likely to have many contacts while others would have significantly fewer interactions. Furthermore, these graphs lack in clustering which means that they do not take into account communities in a society. These are just some of the reasons why this type of graph and random networks, in general, are not very realistic when population structure is modelled (Newman, 2002b).

Lattice Graph

A general case of a lattice graph is a network whose nodes have x and y coordinates in diapasons $[1..n]$ and $[1..m]$ respectively and a unit distance between them.
Unlike random graphs this type of network is highly clustered and all interaction only happen locally. Hence, a disease does not travel far away from sources of infection quickly. Moreover, every node in this graph except those that are located on sides has the same number of neighbours (Keeling1, Eames, 2005).

Small World Network

Small world graph is also known as Watts–Strogatz model has the following property. If two random nodes are taken they are likely not to be neighbours, but one can easily reach another with a small number of steps through other nodes. Therefore, a small world graph is defined as a network whose characteristic distance L (the number of steps required to make to reach one node from another) between two randomly chosen nodes is proportional to the logarithm of N, where N is the total number of nodes in the graph (Telesford et al., 2011).
This graph also has a high level of clusterisation as it tends to have cliques or almost-cliques. Another property of this network is that there are nodes of high degree that play a linking part between other nodes. This means that once infection occurs it does not stay local, but can move to a distant part of the network which considerably changes how far and how quickly disease spreads from for example if simulated on a lattice graph (Keeling1, Eames, 2005).

Scale-free Graph

In this graph new nodes are added according to preferential attachment mechanism proposed by Albert-László Barabási and Réka Albert. At every step a new node is added and the probability that it will be connected to the existing node i is proportional to the degree of i. As a result node degrees in this network are distributed according to the power law. This means that the proportion of nodes in the network that have k connections is $P(k) \sim ck^{-\gamma}$ for large k, where c is a normalisation parameter and γ is a network characteristic which usually lies in the diapason $(2, 3)$ (Bernovskiy, Kuzyurin,2012).
This degree distribution is more realistic since there are nodes with a high degree surrounded by those with a power degree and this agrees with a friendship paradox, which states that a person is likely to be friends with someone who has more friends, recognised by Scott Feld.

Summary

All of these graphs have some kind of deviation from a real world population structure. Nevertheless, they have useful properties which allow these graphs to be used as an idealised version of population structure and their continuous application proves that despite some unrealistic assumptions these networks are useful for modelling epidemics.

Static and Dynamic Graphs

Networks in epidemiology are also distinguished by their persistence. They can be static, which means that once set up, the structure does not change during the simulation process, or they can be adaptive, in which case the network changes over time as some nodes that were initially linked are separated or nodes that were not neighbours can become connected (Bansal et al., 2010). The example of changing over time would be if an individual becomes infected some of the neighbours would immediately want to stop seeing that person and so they will disconnect. The network is then rewired accordingly which reduces the connectivity of infected individuals which could consequently change the way the epidemic spreads through the network (Bansal et al., 2010).
Depending on the model used in a simulation the scenarios of how a graph is rewired can vary. Suppose a disease that has a long incubation period is considered. Another state is added to the model and during this state an individual is infectious but does not yet show any symptoms. Thus, the network might stay the same once a person becomes infected, but is in that new state, and only rewires once that individual proceeds to the infected state. The number of neighbours would then drop dramatically, but some of them could have already been infected by then. Thus, isolation procedures, in this case, might not be as effective as when dealing with the disease where the incubation period is shorter than the latent period and using rewiring on networks might help to determine that.

Ebola Virus Disease

General Information

A lot of epidemiological research has been done with respect to Ebola Virus Disease (EVD). The disease previously known as Ebola hemorrhagic fever (EHF) is serious and often lethal. People can become infected from wild animals and can further spread the infection from person to person (World Health Organization, 2016b).
Those who were infected cannot spread the disease during the incubation period which generally lasts from two to twenty-one days (World Health Organization, 2016b). Moreover, the incubation period for Ebola virus disease is shorter than the latent period (Dobson, 2015), which means that there is a period of time between the symptoms onset and the start of infectiousness which makes early isolation practise effective (World Health Organization, 2016b).
People with the highest risk of infection are considered to be hospital workers and mourners. The former are in close contact with patients and the frequent failure to comply with health standards puts them in danger. The latter following old traditions and carrying out burial ceremonies initiate direct contact with the dead body which is still highly contagious (World Health Organization, 2016b).

Previous Works of Ebola Virus Disease Modelling

A number of models were constructed that were focusing on different aspects of Ebola virus Disease. M.I. Meltzer et al. from "National Center for Emerging and Zoonotic Infectious Diseases" were using their model to predict a future number of infected people and how prevention measures could decrease that number (Meltzer at al., 2014) while A. Rachah and D.F.M. Torres in their research concentrated on one type of control measure specifically vaccination and how it would affect the basic reproduction number (Rachah1, Torres, 2015). Furthermore, D. Chowell et al. looked into how early detection of the virus could change the rate of the disease spread (2015). These works show that early prevention measures and a proper health-care delivery system would significantly improve the situation with EVD and decrease the rapidness of the spread.
M.A. Kiskowski in her research went into a different direction. She introduced a network which differs from those traditionally used, with a purpose of explaining why different regions demonstrate different rates of Ebola virus disease spread (2014). The attention was drawn to the importance of community mixing and how a network structure affects the results of modelling. Many other research papers were examined, but no work was found that would solely focus on the effect of safe burial practices.

Chapter 4

Mathematical Modelling and Software Development Process

Planning and Agile Development

The work on this project was split into two parts. During the first part, a lot of background reading was done to see what other researchers were focusing on when analysing Ebola virus disease. Useful methods and topics were noted and a deeper understanding of what was required for this project developed. After that, a general plan was made in which it was decided to first construct required mathematical models and analyse them and then proceed to software development. In the second part, the plan was set in motion.
Two compartmental models were looked into starting from the simpler one, a general model for which already existed and did not require much modification, and moving on to a slightly more complicated model, the structure of which was designed by the author to serve the needs of the project. The models were described by differential equations which allowed plotting those equations with different parameters to see the disease dynamic over time.
As with mathematical models where the author did not have to start from scratch but made use of previous researches, the starting point for the network simulations was provided by the supervisor which was used to incrementally develop the software.
Once the plan was outlined the author had a set of requirements that needed to be fulfilled. Therefore, it was decided to use an agile approach with a technique similar to Scrum. It was not exactly Scrum since it was a one member project and thus there were no stand-up meetings except meetings with the supervisors. However, the overall task was split into smaller parts that were put into product backlog with different priorities and those parts were assigned to be completed in sprints.
Agile development was chosen as it was very appropriate for the structure of models. The two models that were considered have many similarities and once one was completed it could have been reused to construct the second one by making several changes. Furthermore, since the second model is built on top of the first one the testing of the latter was required before proceeding to the next one as if there were any undesired behaviour it would affect another model as well. Thus, the incremental approach was more natural than for example waterfall where testing happens after the implementation part is finished.

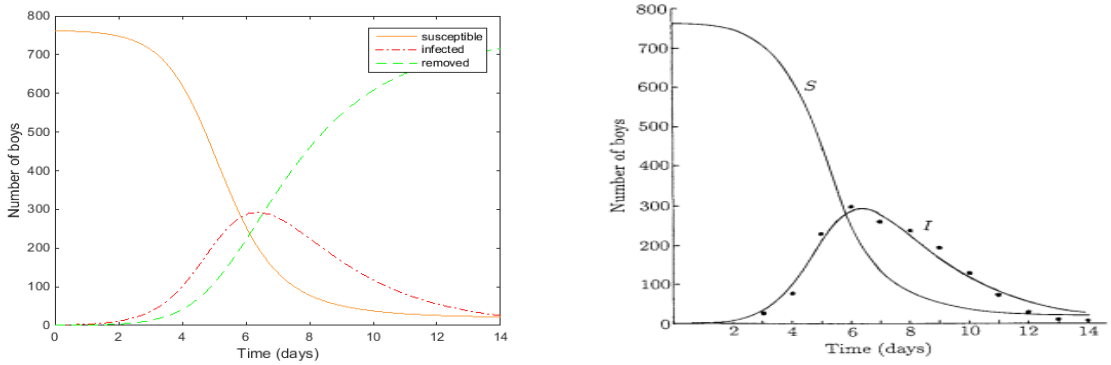
Version Control

Version control is used a lot when working in teams to synchronise code between team members and keep track of who did what and when. Then if something goes wrong it is easy to find out what caused it and return to an earlier version if necessary. Although this project was not done in a team it was a large project and keeping the history of software development to be able to backtrack in case a piece of code gets broken was important. Git with GitHub repository was chosen for version control since it is one of the most well-known code management systems and it was used many times in previous projects without any problems occurring.
Moreover, the work was done using lab machines and personal laptop and the repository served as a link between those. Getting the most up-to-date version on the current working machine was just a matter of several commands.
Furthermore, although the report was written using sharelatex online service which meant that hard-drive failure of the personal laptop would not cause the loss of the report, it was decided to store versions of

the report in the repository as well to see the progress and be extra safe.

Matlab

Since it is a joint project and although Python programming language is widely used among mathematicians, Matlab is not less common if not more and since the author did not have much experience with it before it was decided it was important to learn at least basic functionality. Therefore, Matlab was chosen for analysis of mathematical models and plotting solutions
Getting familiar with the Matlab tools was done with the help of the supervisor and by starting to plot simple differential equations that represented SIR models from existing research. Thus, once plotted the results could have been compared with the plots in the research to make sure that the resulting plot was correct. Hence, the graph below on the left obtained using Matlab corresponds to the one on the right taken from Mathematical Biology: I. An Introduction,Third Edition by J.D. Murray.



(a) Constrcted using Matlab

(b) Source: J.D Murray, 2002, p.326

Figure 4.1: "Influenza epidemic data"; "parameter values N = 763, S0 = 762, I = 1, ρ = 202, r = 2.18 · 10-3/day." Here r corresponds to β and calculated value for γ is 0.44036.

Ipython and Jupyter Notebook

IPython is a replacement to the standard Python shell that offers an advanced list of features that makes working with it more effective. Jupyter notebook, originally developed as Ipython notebook, is an environment where various data types, from bits of code and text to equations, graphs, and many other types, can be combined into one document. The notebook is represented as a JSON object and every section is placed into a separate cell and every one of them can be executed separately from another (Pérez, Granger, 2007). In this project it was particularly useful during the development stage it allowed seeing intermediate results. Thus, instead of running the whole program every time something is changed, only the cells that follow the modified bit were executed. Moreover, work with graphs and analysis of results required visualisation which were easily produced and embedded in a notebook file. Nevertheless, it was much more efficient to use python script files when running experiments since intermediate steps were no longer needed and it was easier to start the scripts from the command line, which was important when automating the process. Therefore, the built-in Jupyter Notebook feature was used to export python script files and those were used at the final stage of the project.

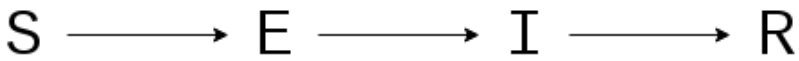
Chapter 5

Methodology

Models used in epidemiology are often simplified and are based on many assumptions. It is important to carefully choose the set up to capture a close to the real dynamic of disease progression.
When modelling Ebola virus disease most researchers consider SEIR compartmental model as the period between getting infected and being infectious is quite significant to be ignored (Volz,Pond, 2014). Thus, this model was also chosen by the author to serve as a contrasting model.
In order to confirm or refute the proposed hypothesis the comparisons between SEIR and SEIDR compartmental models needed to be made. Therefore, both of these were examined in the context of a homogeneous and heterogeneous population (in terms of a number of connections an individual has).
In the fully-mixed population that is investigated in mathematical modelling every individual is considered to be the same as every other, the infectiousness period is the same for each of them and they can infect any individual with the same probability. Thus, the two models can be described by the following schemes and differential equations.

Mathematical Models

SEIR Model



SEIR is an abstract model of how a disease spreads. It says that every individuals is in either of the three compartments: susceptible (S), exposed(E), infected (I) or removed (recovered in some sources) (R) (Shang, 2012). Susceptible means a person can get the disease and after the exposed stage become infected, exposed means an individual was infected, but is not yet infectious, infected means that a person has the disease and can pass it to other people and removed means they no longer have the disease and they are immune to it or that they died and were buried at which point they can no longer infect other people. This mathematical model can be described by the following equations.

$$\frac{dS}{dt} = -(\beta \cdot S \cdot I) \tag{5.1}$$

$$\frac{dE}{dt} = (\beta \cdot S \cdot I) - (\eta \cdot E) \tag{5.2}$$

$$\frac{dI}{dt} = (\eta \cdot E) - (\gamma \cdot I) \tag{5.3}$$

$$\frac{dR}{dt} = (\gamma \cdot I) \tag{5.4}$$

$$S(0) = S_0 > 0, \quad E(0) = 0, \quad I(0) = I_0 > 0, \quad R(0) = 0 \tag{5.5}$$

$$N = S + E + I + R \tag{5.6}$$

$$\beta = contact\_rate\_i \cdot transmission\_probability\_i \tag{5.7}$$

$$\eta = \frac{1}{latent\_period} \tag{5.8}$$

$$\gamma = \frac{1}{infectious\_period} \tag{5.9}$$

$$N - whole\_population$$

$$S - susceptible$$

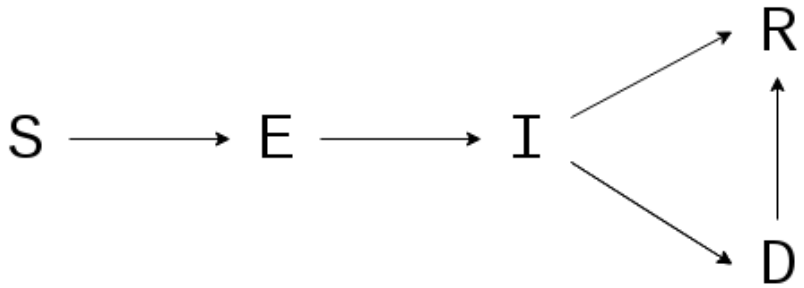
$$E - exposed$$

$$I - infectious$$

$$R - removed$$

What happens with these classical models is that they assume that the population is closed (equation 5.6) and everybody starts in the susceptible compartment apart from a small fraction of people who are infected (equation 5.5). The number of people in the susceptible compartment decreases by the infectiousness number β multiplied by the number of Is and Ss that meet. The idea is that all the infected meet all the susceptible at a specified contact rate and out of those interactions some fraction β of those susceptible becomes exposed. At the same time the number of exposed people increases with the same rate as the number of susceptible people declines. However, it also decreases with the rate η as latent period ends and people become infectious. Moreover, the number of infected people increases by the number of individuals who move from the exposed to infected compartments and decreases by some factor γ as they recover.
This is a very general and abstract model and there already exist many variations of it from the more general models, e.g. SIR (Susceptible-Infected-Recovered) or SIS (Susceptible-Infected-Susceptible) to the more specific ones such as MSIR model where M stands for the maternally-derived immunity and many others. In case of the Ebola virus disease when people die they do not stop being infectious. On the contrary, they get to the peak of their infectiousness and therefore they cannot immediately be moved from the infected to removed compartments. Moreover, the contact rate between newly dead people and those in susceptible compartment is different from the contact rate between people in the infected and susceptible compartments. Thus, the new model was suggested called SEIDR.

SEIDR Model



SEIDR model introduces another compartment of newly dead people and is represented by the following differential equations.

$$\frac{dS}{dt} = -(\beta \cdot S \cdot I) - (\delta \cdot S \cdot D) \tag{5.10}$$

$$\frac{dE}{dt} = (\beta \cdot S \cdot I) + (\delta \cdot S \cdot D) - (\eta \cdot E) \tag{5.11}$$

$$\frac{dI}{dt} = (\eta \cdot E) - (\gamma \cdot I) \tag{5.12}$$

$$\frac{dD}{dt} = ((1-\epsilon) \cdot \gamma \cdot I) - (\zeta \cdot D) \tag{5.13}$$

$$\frac{dR}{dt} = (\epsilon \cdot \gamma \cdot I) + (\zeta \cdot D) \tag{5.14}$$

$$S(0) = S_0 > 0, \quad E(0) = 0, \quad I(0) = I_0 > 0, \quad D(0) = 0 \quad R(0) = 0 \tag{5.15}$$

$$N = S + E + I + D + R \tag{5.16}$$

$$\delta = contact\_rate\_d \cdot transmission\_probability\_d \tag{5.17}$$

$$\zeta = \frac{1}{burial\_period} \tag{5.18}$$

$$\epsilon = probability\_of\_recovery \tag{5.19}$$

$$D - newly\_dead$$

In this new model, some of the features are the same as in the SEIR model, but several new are added. Initially, it is again assumed that most people are in the susceptible compartment apart from a small part of the population which is infected. However, people in the susceptible compartment can now become infected by newly dead people. Therefore, the number of susceptible people further decreases by the fraction δ multiplied by the amount of susceptible and newly dead people who meet and the number of exposed people increases by the same amount. Furthermore, since there are now two compartments to which people can move once they leave infected compartment, the new variable ϵ is added which represents the probability of recovery. Once in the newly dead compartment, people leave it and move to the removed compartment with a specified rate ζ as they get buried and can no longer infect anyone.

Analysis of Equations

The systems of equations are quite complicated to be solved analytically. However, some conclusion can be made by analysing them.
Since number of people cannot be negative, $S > 0$, $E > 0$, $I > 0$, $R > 0$ and additionally for SEIDR $D > 0$. Furthermore, all the parameters $\beta, \gamma, \eta, \delta, \epsilon, \zeta$ are all positive. Therefore, the following can be inferred for the SEIR model.

$$\frac{dS}{dt} < 0 \quad \forall t \tag{5.20}$$

$$\frac{dR}{dt} > 0 \quad \forall t \tag{5.21}$$

$$\frac{dE}{dt} + \frac{dI}{dt} = (\beta \cdot S \cdot I) - (\gamma \cdot I) \tag{5.22}$$

In the equation 5.22, $\frac{dE}{dt}$ and $\frac{dI}{dt}$ are looked into together as the changes in the infected fraction of population is of interest here and despite not being yet infectious $\frac{dE}{dt}$ also represents changes in the number of individuals who were previously infected. Since $I > 0$ equation 5.22 also implies the following.

$$\frac{dE}{dt} + \frac{dI}{dt} > 0 \Leftrightarrow S > \frac{\gamma}{\beta} \tag{5.23}$$

This gives the basic reproduction number R_0. This number represents the average number of people who can be directly infected by the individual with the disease during the contagious period when that individual is found in the wholly susceptible population. R_0 is affected by the number of susceptible people in the total population, contact rate between infected and susceptible people, transmission probability and the infectious period of those with a disease. This number determines whether the epidemic occurs. Hence, if susceptible part of the population is larger than the fraction $\frac{\gamma}{\beta}$, a disease will be spreading till this is no longer true. Once the susceptible proportion drops the epidemic will start dying out.
However, when considering SEIDR model there are more variables that need to be taken into account. There is another compartment of newly dead people and they can also pass the infection on to others. Thus, while equations 5.20, 5.21 also hold for SEIDR, $\frac{dE}{dt}$ and $\frac{dI}{dt}$ now have to be looked into together with $\frac{dD}{dt}$. This gives us the following.

$$\frac{dE}{dt} + \frac{dI}{dt} + \frac{dD}{dt} = (\beta \cdot S \cdot I) + (\delta \cdot S \cdot D) - (\gamma \cdot I) - ((1-\epsilon) \cdot \gamma \cdot I) - (\zeta \cdot D) = I \cdot (\beta \cdot S - \epsilon \cdot \gamma) + D \cdot (\delta \cdot S - \zeta) \tag{5.24}$$

This also gives us the basic reproduction number, however in this case it is more complicated to analyse it since there are several scenarios to consider. Firstly, $\frac{dE}{dt} + \frac{dI}{dt} + \frac{dD}{dt} > 0$ if $S > \frac{\gamma \cdot \epsilon}{\beta}$ and $S > \frac{\zeta}{\delta}$ and similarly $\frac{dE}{dt} + \frac{dI}{dt} + \frac{dD}{dt} < 0$ if $S < \frac{\gamma \cdot \epsilon}{\beta}$ and $S < \frac{\zeta}{\delta}$. However, if only one of the conditions is satisfied the number of people with a disease will either increase or decrease depending on many factors.
In order to see some evidence that these calculations would indeed work with the systems of differential equations, Matlab was used to plot the changes in time. To start off some of the parameters were taken from the data collected by Z. Xia et al. in the in the research "Modeling the transmission dynamics of Ebola virus disease in Liberia" (2015). However, since differential equations describe fully-mixed

population which does not correspond to the real population structure (which was used in the research to collect data) the β parameter had to be reduced. The resulting plot for the SEIR model is the one below.

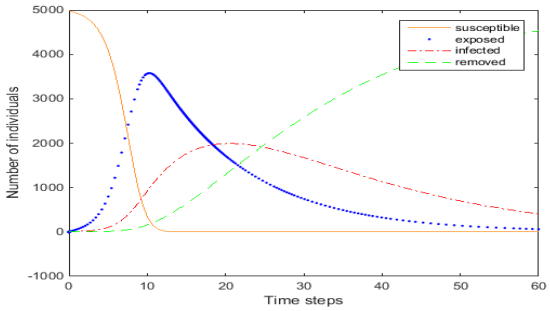
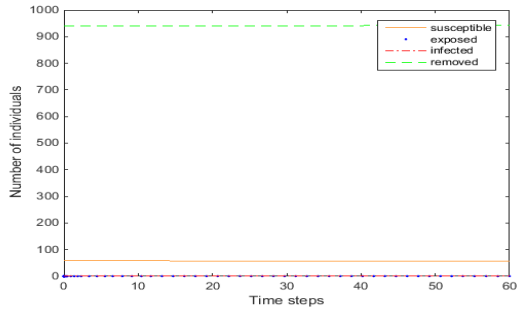
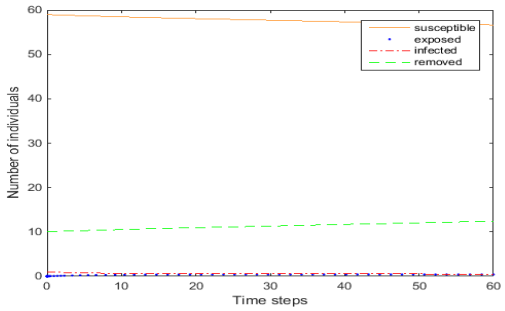


Figure 5.1: Ebola virus disease. SEIR model. $S_0 > \frac{\gamma}{\beta}$

The parameters used to plot the graphs are $\beta = = 0.001151$, $\gamma = 0.06851662$, $\eta = 0.083333$, $S_0 = 4999$, $E_0 = 0$, $I_0 = 1$, $R_0 = 0$. Therefore, $\frac{\gamma}{\beta} \approx 60$, which means that the number of susceptible individuals needs to drop below the value of this fraction for an epidemic to start dying out. Using Matlab again two arrays that contain numbers of individuals in exposed and infected compartment are added together and the peak point is found. It was also calculated that this point located where $t \approx 15$ and where the number of susceptible individuals falls from approximately 71 to 56.
Consider another experiment where most of the population are immune to the disease and are in the removed compartment or if simply a smaller population size is considered and $S_0 < \frac{\gamma}{\beta}$. No epidemic occurs and this is visible in the graphs below.



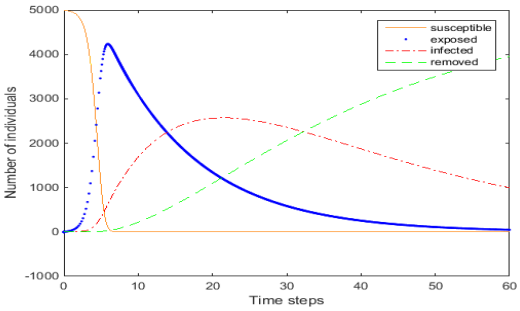
(a) $S_0 = 59$, $E_0 = 0$, $I_0 = 1$, $R_0 = 940$



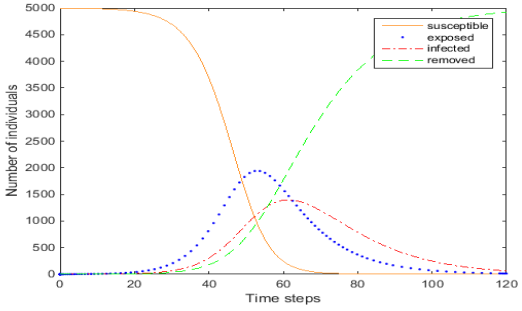
(b) $S_0 = 59$, $E_0 = 0$, $I_0 = 1$, $R_0 = 10$

Figure 5.2: Ebola virus disease. SEIR model. $S_0 < \frac{\gamma}{\beta}$

Other experiments also showed that changing values of β, γ alter the threshold value of S. When it is reached, this causes the epidemic to gradually stop, the graphs below show scenarios with other parameter values. However, this still corresponds to the basic reproduction number, which means that it was presented correctly.



(a) $\beta = = 0.005151$, $\gamma = 0.03851662$, $\eta = 0.083333$, $\frac{\gamma}{\beta} \approx 7$



(b) $\beta = = 0.000151$, $\gamma = 0.08851662$, $\eta = 0.083333$, $\frac{\gamma}{\beta} \approx 586$

Figure 5.3: Ebola virus disease. SEIR model. $S_0 > \frac{\gamma}{\beta}$

Following similar process SEIDR model was examined. For the same values of β, γ, η but with the new compartment which also contains contagious humans it is noticeable on the graph below that the slope which shows the changes in the number of susceptible people is much steeper which means that the disease progresses quicker. The new parameters are $\delta = 0.189$, $\zeta = 0.5$, $\epsilon = 0.2$, $d_0 = 0$.

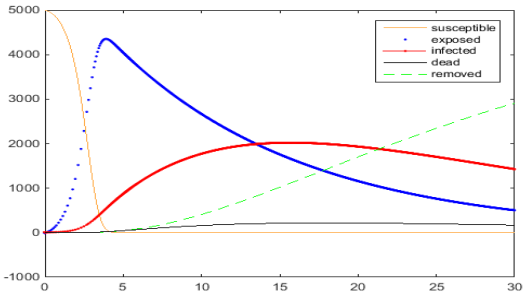
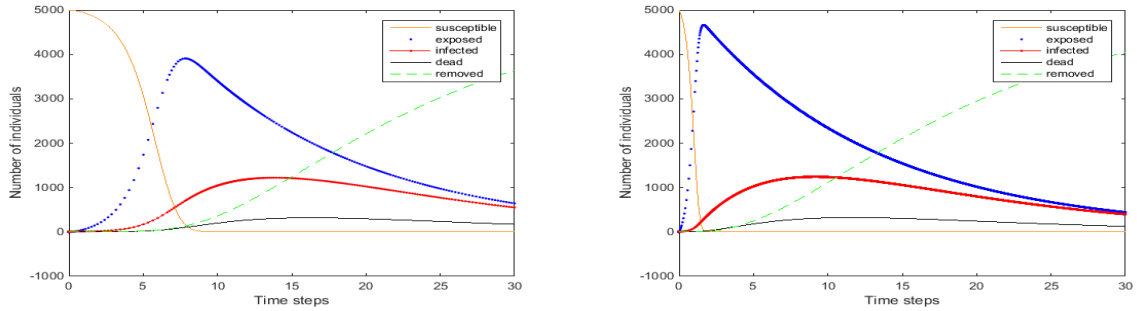


Figure 5.4: Ebola virus disease. SEIDR model. $S_0 > \frac{\gamma \cdot \epsilon}{\beta}$ and $S_0 > \frac{\zeta}{\delta}$

The threshold values are $\frac{\gamma \cdot \epsilon}{\beta} \approx 11$ and $\frac{\zeta}{\delta} \approx 2$. It was discussed earlier that the number of people in the S compartment should go below either both of these values or otherwise the number of infected and dead individuals should be also taken into account. In the graph above the peak of the sum of people in the compartments E,I and D is reached when $S \approx 3$ which means that it only satisfies the first condition, therefore, the whole equation 5.24 is looked into. When computed at the point right after the peak the result of this equation was negative as the left side of it started to decrease and outweighed the right side. Again by varying the parameters the changes in the dynamic can be seen below.

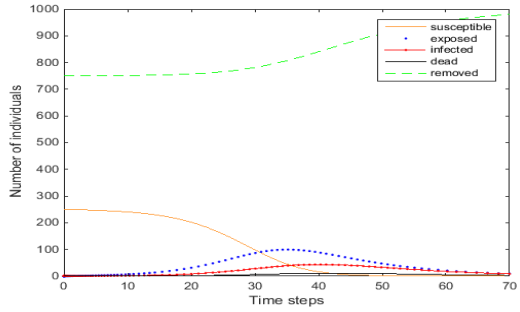


(a) $\beta = = 0.001151$, $\gamma = 0.16851662$, $\eta = 0.083333$, $\delta = 0.0189$ $\epsilon = 0.2$, $\zeta = 0.5$, $\frac{\gamma \cdot \epsilon}{\beta} \approx 292$ and $\frac{\zeta}{\delta} \approx 26$

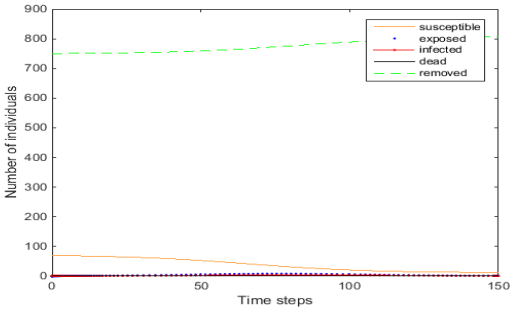
(b) $\beta = = 0.01151$, $\gamma = 0.16851662$, $\eta = 0.083333$, $\delta = 0.489$ $\epsilon = 0.2$, $\zeta = 0.5$, $\frac{\gamma \cdot \epsilon}{\beta} \approx 3$ and $\frac{\zeta}{\delta} \approx 1$

Figure 5.5: Ebola virus disease. SEIDR model. $S_0 > \frac{\gamma \cdot \epsilon}{\beta}$ and $S_0 > \frac{\zeta}{\delta}$

Now if the SEIDR model with the parameters $\beta = = 0.001151$, $\gamma = 0.16851662$, $\eta = 0.083333$, $\delta = 0.0189$ $\epsilon = 0.2$, $\zeta = 0.5$ is considered the values of fractions are $\frac{\gamma \cdot \epsilon}{\beta} \approx 292$ and $\frac{\zeta}{\delta} \approx 26$. If the initial value $S_0 = 250$, although it is smaller than $\frac{\gamma \cdot \epsilon}{\beta}$ it is not enough to make the equation 5.24 go below zero and the fact that the epidemic occurs is shown on the graph (a) below. However, if $S_0 = 70$ is taken by the end of the disease spread the number of susceptible individuals does not drop to zero (graph (b) below).



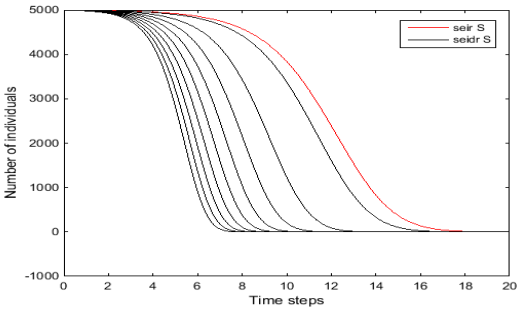
(a) $S_0 = 250$



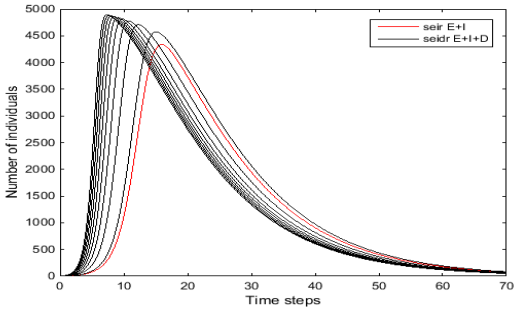
(b) $S_0 = 70$

Figure 5.6: Ebola virus disease. SEIDR model. $S_0 < \frac{\gamma \cdot \epsilon}{\beta}$ and $S_0 > \frac{\zeta}{\delta}$. $\beta = = 0.01151$, $\gamma = 0.16851662$, $\eta = 0.083333$, $\delta = 0.489$ $\epsilon = 0.2$, $\zeta = 0.5$.

Finally, a comparison between two models with respect to the changes in the number of susceptible individuals and those who have the disease in the system is made while parameters that are only specified for the SEIDR model were varied. The two graphs below show the difference in the dynamic for SEIR and SEIDR models when the δ parameter, which describes the contact rate between dead people and susceptible part of the population and the probability of transmission between those, was changed. The higher infectiousness of dead humans, the steeper the curves for the SEIDR model become. It was also noted that with the SEIDR model the epidemic does not last longer but affects a larger number of people quicker.



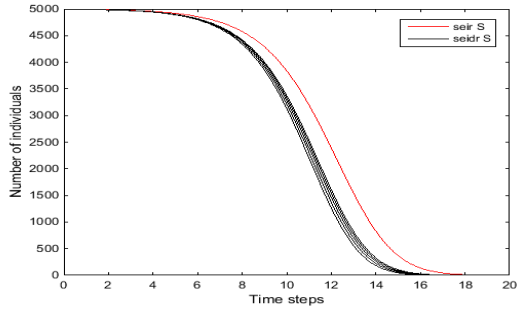
(a) Changes in the susceptible fraction of population.



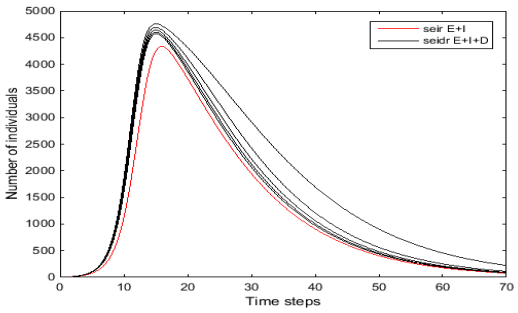
(b) Changes in the population of people who have the disease.

Figure 5.7: Ebola virus disease. SEIR and SEIDR model comparison when δ value is varied.

The next parameter that was changed was ζ, which describes the rate of dead people being buried. The graphs below show that with β and δ values being equal the ζ parameter has less effect on how fast the number of susceptible humans declines and people who have the disease grows. The same also applies to cases where epsilon value, which denotes the probability of recover, is altered.

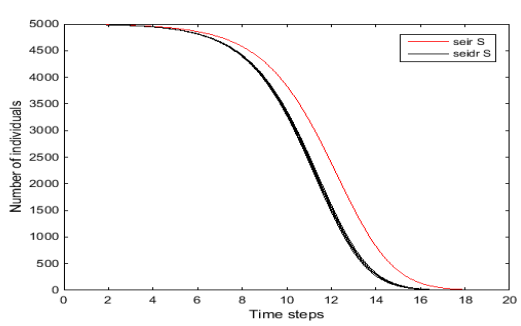


(a) Changes in the susceptible fraction of population.

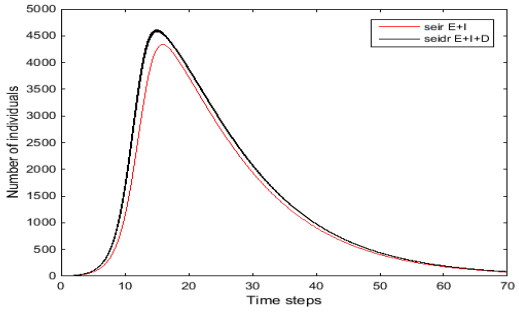


(b) Changes in the population of people who have the disease.

Figure 5.8: Ebola virus disease. SEIR and SEIDR model comparison when ζ value is varied.



(a) Changes in the susceptible fraction of population.



(b) Changes in the population of people who have the disease.

Figure 5.9: Ebola virus disease. SEIR and SEIDR model comparison when ϵ value is varied.

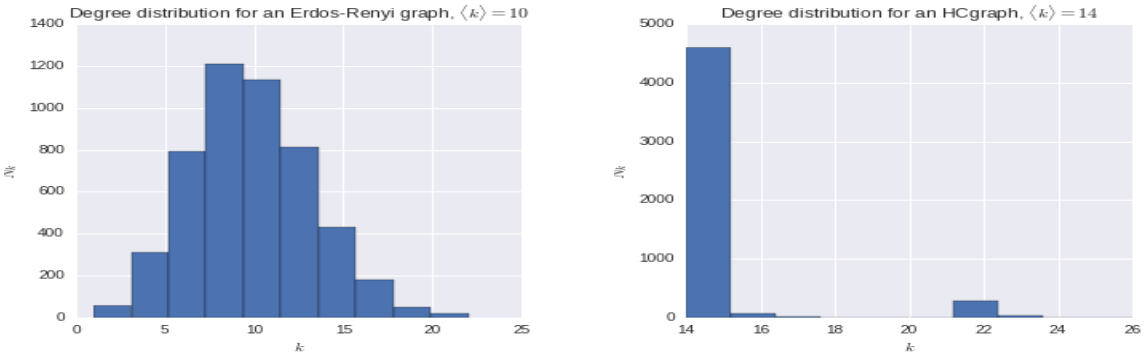
To summarise, the main difference in the dynamic of the disease spread when SEIDR model is used compared to the SEIR model is that the epidemic propagates faster and takes more time to die out. This happens as the new compartment of dead people who are still contagious is introduced together with appropriate parameters. It was also found that the δ parameter appeared to influence the difference in the dynamic the most.
Going back to the setup part from the assumptions mentioned earlier the author also wanted to add that in reality infectious probability between pairs vary. However, M.E.J. Newman has proven that although the probabilities between pairs vary, it does not matter. In fact, the only thing that matters is the mean value. M.E.J. Newman states that "in the population as a whole the disease will propagate as if all transmission probabilities were equal to T", where T is "the "transmissibility" of the disease" (Newman, 2002a). Nevertheless, there are a few assumptions that make the results obtained from mathematical modelling less accurate and although analysing these mathematical models provides with an interesting insight simulations on networks should give a better view of how the disease progresses.

Simulations on Networks

Graph Choice

Running experiments on just one graph type would not be sufficient since a network structure significantly affects the spread of the disease. Therefore, the Erdős–Rényi random graph and the custom made *HCgraph* (household-communities-graph) were chosen as two contrasting graph models.
The main features of the first one as described earlier is that it has a binomial degree distribution which becomes Poisson as n tends to infinity, where n is the order of the graph and the connections between nodes are not limited to local area (Keeling1, Eames, 2005). Meanwhile, the second graph model has not been discussed yet, therefore, the attention will be focused on it.
The idea behind *HCgraph* was taken from the research carried out by M. A. Kiskowski where she used her graph to distinguish between close contact and casual contact while modelling the spread of the Ebola virus disease. In her research, the probability of disease transmission varied when considering close and casual contacts. Thus, the graph consisted of households where nodes were in close contact and communities where the connections between individuals were considered casual. The household then lied within communities and some households could belong to several of them at the same time. The term household did not necessarily include family members only but was used as a general term for everybody in the close contact. Hence, it could also mean friends, hospital workers, mourners in the funerals, etc. The research showed that such network structure with different community mixing parameters closely follows the population structure in some of the West-African countries, the ones that were looked into. Therefore, some concepts for the *HCgraph* were taken from the graph in the research. However, not all features were suitable which is why several modifications were made.
As in the research by M. A. Kiskowski in HCgraph households are represented as fully connected subgraphs, nevertheless since the differences between close and casual contacts are not considered here nodes in the communities are not all connected, but there are some nodes that link households together. The same nodes are chosen to connect different households to ensure that there are some nodes with a high degree to escape having on average the same degree distribution for all nodes in the graph. Thus, when adding households together the node with the highest so far degree is chosen as this must be the node selected before to link previous households. The parameters that have to be specified to create the graph are the household size, the community size and the number of communities.

As a result, the graph has some similarities with the small-world graph model however, there are no large distance connections, but some locally "popular" nodes exist instead.
The reason why this graph was chosen is that there are fewer choices to be made when rewiring the network. It made sense to connect randomly chosen people in the same community to the dead node and since there are some nodes which belong to the households that lie in several communities the number of connected nodes naturally varies resulting in a larger number of additional connects to nodes of high degree.
To sum up, the graphs chosen have some useful distinct features which ensure that the research is not biased towards the use of one specific graph that would show the desired results. The main differences between two graphs lie in their degree distribution which can be observed in the histograms below and the range of distances in connections as for the random graph the distances vary while for the second one all connections are local, which means that the disease is more likely to spread in a wavy manner.



(a) Erdős–Rényi graph with 0.002 parameter.

(b) HCgraph with parameters 15, 10, 37.

Figure 5.10: Degree distribution of ≈ 5000 nodes.

Running Simulations

Once the graphs were chosen and implemented, the models were developed according to the idea that lies behind differential equations, except that the population considered was no longer fully-mixed, but described by the network structures. Then in order to compare the models the simulations needed to be executed. Running simulations once with different parameters was not an option since transitions between compartments depend on probabilities which means that every run is likely to produce slightly different results. Therefore, after running some experiments the population size and the number of runs were chosen based on the variance in the results and the time left before all the data had to be gathered. It was found that the time spent on one run is mostly affected by the amount of rewiring done. Thus, the time varied from around one to five minutes when the graph consisted of five thousand nodes.

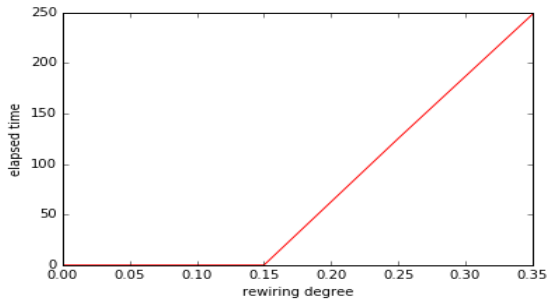


Figure 5.11: Time dependence on rewiring degree.

It was also observed that approximately fifty runs was enough to minimise the variance in the results obtained when the same size of the network was considered. Thus, running an experiment with one set of parameters would take from fifty minutes to more than four hours. Furthermore, experiments had to be run with different sets of parameters and on different graphs which meant that those hours would transform to days. Hence, it was decided that although the size of the population where the epidemic

spreads in real life is much larger, it is enough to compare the models.
Before any actual data was gathered the models had to be tested to make sure that they operate in an expected manner. Thus, some unit tests were written. However, the tests could spot weird behaviour, but were not always perfect at providing hints about what was causing it. Furthermore, debugging process when working with graphs is quite complicated and, in this case, the easiest way to find where something goes wrong is by visualising the simulation process. Lattice graph was used for this as all the connections between nodes are clearly visible and if for example one node became infected when none of the neighbours could actually transmit the disease it was easy to notice and deal with the problem. A segment of what the visualised dynamic looks like is displayed below.

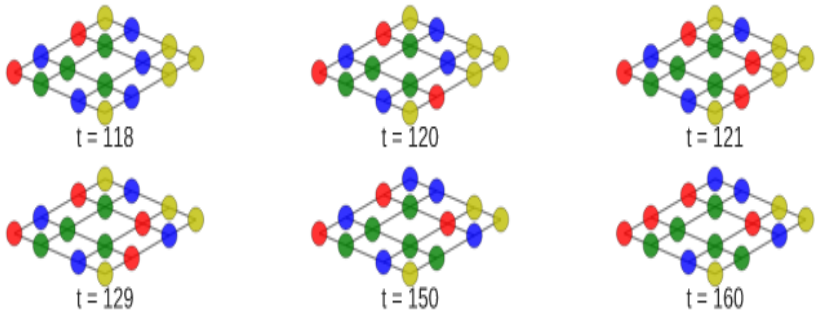


Figure 5.12: Visualisation bit of simulation on the Lattice graph

After that when the author was convinced that the models were implemented correctly, the last thing that needed to be decided on was the parameter values. Therefore, finding parameters that would ensure that the models closely describe the dynamic of the Ebola virus disease became the next aim. In most of the previous researches even if the parameters were used in modelling only results were reported. Nevertheless, one paper was found that provided a full table of data observed or estimated that was gathered during the outbreak in Liberia []. Most of the parameters were used as a starting point and those that were particularly interesting for model comparisons were later varied. The chosen distinctive features are the following:

- model type
- graph type
- rewiring degree
- δ value
- β value

Furthermore, the attributes selected to be recorded are listed below:

- graph parameters
- graph order
- percentage of initially infected population
- disease parameters (β, γ, η with SEIR and SEIDR and δ, ζ, ϵ with SEIDR only)
- maximum and mean outbreak sizes and outbreak proportion
- number of disease transmissions from infected and from dead individuals with SEIDR model
- elapsed time
- number of events (transitions from one compartment to another)
- total number of time steps
- number of time steps when events happened
- number of events at each time step
- size of compartments at every time step with events

- rewiring degree

In the end, everything was ready to start experiments and although the process of running them was as automated as possible, it still took more than two weeks to get all the data required.
The amount of data collected was substantial, however gathering the data was not as hard as understanding what it implicates. In order to assist with that, the data was averaged out among the runs with the same parameters, visualised and then compared.

Chapter 6

Implementation

The first thing that had to be considered when starting to work on the simulation on networks part of the project was the programming language. Python was the first choice for this as it is extensively used in the field and provides a high level of support when working with graphs. Moreover, some of the code was provided as a starting point and it was written using Python 2 and thus, it was decided to stick to it.

Libraries Used

One of the essential packages used was the **networkX** one. This library was written in Python and is used to work with graphs and other network structures. This is a free software distributed under the new BSD license. The main library capabilities include classes for working with simple, focused and weighted graphs, the ability to save/load graphs to/from the most popular formats of graph file storage, built-in procedures for creating basic types of graphs, the accessibility of such graph characteristics as degrees of vertices, graph height, weight, diameter, radius, center and so on. There are the main ones that were applied, but there are much more ways to make use of this package. Furthermore, the level of performance is very high which is very important for this research as a lot of graph manipulations were required. It is claimed that the library can easily manage graphs with up to ten million nodes with a hundred million edges between them since it is based on the dictionary of dictionaries data structure where memory is effectively used.
Other libraries were used less extensively but also played an important role in the process. During the simulation process, every decision on the change of node's dynamic state was based on a certain probability. Therefore, it was necessary to generate random numbers for which **numpy.random** was used.
Furthermore, working with graphs without seeing what is happening makes it complicated to spot any undesired behaviour, therefore, some visualisation tools were required. Plots for analysis and visualisation were created with the use of **matplotlib.pyplot** library and the final animation of the disease spread on the graph was made with the **animation** tool from the **matplotlib** package.

HCgraph

Before simulations could be done graphs which would be used to run them on needed to be prepared and unlike the Erdős–Rényi graph which exists in the networkX package, the *HCgraph* had to be implemented. It was made as a subclass of the *networkx.Graph* class so that all the methods on graphs could also be used on *HCgraph*. It is constructed by creating communities and combining them by connecting some of the nodes within them. In order to build a community all households within it are first created by constructing fully connected graphs of *household_size*, combining them into one bigger network using built-in into networkX library *disjoint_union* function and linking them afterwards.
All the methods within the class except one are only used while creating the graph. The one method that is used later during the simulation helps to determine which community/communities a node belongs to. Using the *household_size* and *community_size* the method estimates the interval(s) in which nodes in the same community/communities lie. This method is used during the rewiring process to connect nodes to the one that moved to the dead compartment.

Provided Code

Some of the code that was used in the project was not written by the author. In fact classes named *GraphWithDynamics* and *GraphWithSynchronousDynamics* were left almost untouched. The first modification made was the change of the superclass of the *GraphWithDynamics* from *nx.Graph* (*networkx.Graph*) to *HCgraph* in order to be able to access some of methods from the superclass during rewiring if *HCgraph* is used in a simulation. Since *GraphWithSynchronousDynamics* is a subclass of *GraphWithDynamics*, the inheritance structure for it was also automatically changed. The second one was the addition of the populations method to the *GraphWithDynamics* which will be described in the next subsection.
The understanding of the structure and functionality of these classes is important, therefore the short description of key methods is provided.

GraphWithDynamics

GraphWithDynamics is mostly just a structure which describes what methods need to be implemented by subclasses. A few of them have a non-empty body and those of interest are *skeletonise* and *populations* method. The former takes out all the edges that were not used to transmit a disease and returns the resulted network. This method is called after the simulation has finished and the returned network is used to find the outbreak size. The latter counts a number of individuals in each compartmental state and is called during those time steps when at least one event has occurred. An event here is considered a transition of an individual from one compartment to another.

GraphWithSynchronousDynamics

The method of a particular interest here is the *dynamics* method. What it does is it calls the *dynamics_step* method at every time step till equilibrium is achieved. A dynamic step and equilibrium state are model specific and will be described in the 'Modified Implementation of Models' section.

Modified Implementation of Models

The code provided captured the dynamic of the SIR compartmental model and required changes to fit the SEIR and SEIDR models. Firstly, apart from a list of infected that were stored in the initial *SIRSynchronousDynamic* class provided, a list of exposed nodes needed to be kept in both *SEIRSynchronousDynamic* and *SEIDRSynchronousDynamic* classes and a list of dead nodes in the second one. These lists make calling specific methods in the dynamic step easier as those methods should only be called on nodes in corresponding states, e.g. *end_latent* method is only used with exposed nodes. Furthermore, a list of possible dynamic states of a node was extended by adding missing compartments, a list that contains edges in the network at a particular time step as well as the dictionary that uses time steps as keys and has edges at all time steps were added to be used later for visualisation and a dictionary *populations* that stores a number of nodes in each state for each time step with events together with counters of numbers of exposed nodes that got infection from either infected or dead nodes were added to gather data for experiments. Finally, a few new methods were created and existing modified. The table below describes what each method does for both of the models.

Methods	*SEIRSynchronousDynamic*	*SEIDRSynchronousDynamic*
__init__	Parameters for *HCgraph* are set if specified as well as the disease parameters. The naming conventions for the disease parameters are kept the same as in the mathematical modelling for convenience	Same as in *SEIRSynchronousDynamic* except that a rewire degree can be set otherwise it defaults to zero. The rewire degree specifies the percentage of the population that connects to an individual when they enter the dead state.

before	The *_exposed* and *_infected* lists are cleared, the counters of sources of exposed nodes are set to zero in case rerun happens from a dirty state, the network is seeded with initially infected nodes according to the *p_infected* fraction and the dynamical state of the rest of the nodes is set to susceptible and all edges are marked as unoccupied (not yet used for infection transmission).	Same as in *SEIRSynchronousDynamic* except the *_dead* list is also set to be empty.
rewire	-	The first *rewire* method used on *HC-graph* connects all the nodes in the same community as the dead node (that are not yet connected and whose dynamical state is not removed) to it. It mimics funerals when all friends and family gather together for burial ceremonies.
rewire1	-	This *rewire* method used on a random graph chooses nodes at random to connect to the dead node. The chosen nodes are again those that are not already connected, are not in the removed compartment and not the dead node itself. Since the edges are stored as tuples, both versions for example the i-th node and the dead node and the symmetric image of it need to be checked to satisfy the "not yet connected" condition above. The number of nodes that are connected depends on the *rewire_degree* which is passed when the simulation is started.
model	The method is only called on infected nodes and while going through the neighbours of these infected nodes the disease is transmitted to them with probability β making the nodes move to the exposed compartment. The infected nodes also recover with probability γ.	Same as in *SEIRSynchronousDynamic* apart from the recovering nodes part. The infected nodes leave the infected compartment with probability γ, however once this happens they either join the recovered compartment with probability ϵ or the dead one with probability $1 - \epsilon$. If the latter happens, one of the rewiring methods is called.
model1	-	Same idea as in the *model* method except now the method only works with dead nodes and the neighbours are infected with probability δ. Some of the dead nodes are removed with probability ζ.
end_latent	The exposed nodes move to the infected compartment with probability η.	Same as in *SEIRSynchronousDynamic*.

dynamics_step	All the methods that change the dynamic state of nodes are brought together to perform one step in the simulation. During this step firstly, the *model* method is called on every singe node in the *_infected* list after which the list is reconstructed since some of the nodes might have recovered. Secondly, the *end_latent* method is called on exposed nodes followed by the list reconstruction and finally the dynamic states of each node are stored for that particular time step.	Same as in *SEIRSynchronousDynamic* except there is an extra step between the first and the second steps where the *model1* method is called on each node in the *_dead* list and when the dynamic states of nodes are changed the list is reconstructed.
at_equilibrium	The simulation is at equilibrium if the *_infected* and *_exposed* lists are empty or if it reaches the time step limit. Thus, these conditions are checked.	Same as in *SEIRSynchronousDynamic* except the *_dead* list should also be empty.
dynamics	The *dynamics* method from the superclass is called and the simulation parameter and metrics are returned together with results.	Same as in *SEIRSynchronousDynamic*.

Table 6.1

Additionally, the code below shows the logic of the transition from S to E compartments with the D compartment as a source of infection. It does exactly what was described in the table 6.1 and other transition methods were implemented in a similar manner.

```
def model1(self, n):
    '''Apply the SEIDR dynamics to (dead) node n.
    n: the node
    returns: the number of changes made'''
    events = 0

    # infect susceptible neighbours with probability delta
    for (_, m, data) in self.edges_iter(n, data = True):
        if self.node[m][self.DYNAMICAL_STATE] == self.SUSCEPTIBLE:
            if numpy.random.random() <= self._delta:
                events = events + 1

                # infect the node
                self.node[m][self.DYNAMICAL_STATE] = self.EXPOSED
                self._exposed_from_dead = self._exposed_from_dead + 1
                self._exposed.insert(0, m)

                # label the edge we traversed as occupied
                data[self.OCCUPIED] = True

    # node moves to the R compartment with probability zeta
    if numpy.random.random() <= self._zeta:
        events = events + 1
        self.node[n][self.DYNAMICAL_STATE] = self.REMOVED

    return events
```

Auxiliary Functions

Two additional functions were created one of which is used for visualisation of the simulation and the second one prints the simulation results to files.
The *show_changes* function creates separate lists of nodes in every dynamic state for each time step where events occurred and then plots the nodes in different colours depending on their compartment. A small plot for every time step with changes is drawn which makes it easy to see the spread. Furthermore, since the edges for every time step during the simulation are recorded, the rewiring is also visible on these plots.
All the disease parameters, metrics, results obtained from simulation are recorded in files with "csv" extension for further analysis. The results obtained had to be split up between two files. It was more convenient to store the sizes of each compartment as well as the number of events happening for each time step separately as this data is more massive that all the parameters and it was also intended to be processed independently. Furthermore, one auxiliary function was provided that animated SIR dynamic of a disease progression. This function was modified for the models used in the research.

Testing

Although one of the easiest ways to see that nothing unexpected happens during the simulation process was by visualising every timestep where changes occurred, the official testing was required. Therefore, a set of python unit tests were implemented.
In these tests, the logic of transitions between compartments was incorporated and forced. For instance, the lists that contain nodes of specific compartments cannot overlap (e.g. the node cannot be in the exposed and infected compartments at the same time) or at any point during a simulation a node cannot move to the left in a model chain (S-E-I-R or S-E-I-D-R) in other words if for example a node reaches a compartment D it can never be in S, E or I compartment again, otherwise, an assertion fails and a provided message gives a hint about where the mistake might be.
Overall, the testing bit was very useful during the development process when modifications were made to check that the introduced changes did not alter the specified behaviour and to ensure that the simulations on SEIR and SEIDR models operate as expected.

Structure of the Software

Although SEIR model could have been implemented through SEIDR model by using the default values for additional disease parameters, it was decided to separate the two of them for a simpler process of data collection and code readability. Therefore, there is a notebook file which contains both of them and two of the auxiliary functions described and is called *TwoModels* and matlab files with corresponding names that were used to plot mathematical models. Moreover, two python files were exported from the notebook files to run the simulations and are called *SEIRmain.py* and *SEIDRmain.py* Furthermore, the implemented tests are in the *Tests* file and notebook files that were used to construct all the graph used in the report are called *Data Analysis* and *General Data Analysis* Finally, the script that was used to run all the experiments is in the *ExperimentRunner* folder together with all the data collected.

Chapter 7

Results

In this chapter, the implication of gathered data is discussed and visualisation of some features provided. The discussion is split into three parts where in the first and second parts the data collected with both models is analysed separately and in the third part it is compared. Furthermore, each of these sections is further divided into segments where results obtained with Erdős–Rényi graph and HCgraph are outlined and contrasted.
The analysis includes comparisons of outbreak sizes and the difference between the number of individuals who received the disease from infected and dead humans in SEIDR model, however, its main focus is the dynamic of the disease and thus, the information about the sizes of compartments at each time step is scrutinised.
The table below shows the changes in variables and population structure (graph types) that were made during data acquisition together with the line type used on graphs for those values. Moreover, throughout this chapter the compartments on graphs are represented by the following colours.

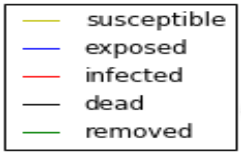


Figure 7.1: Compartments on graphs

SEIR Model

The dynamic of the disease on the SEIR model is represented by the changes in sizes of four compartments. To visualise this the following graph was constructed.

β	δ	*rewiring_degree*	*graph_type*	*line_type*
0.1151	0.289	0.0	Erdős–Rényi graph	—
0.2151	0.389	0.15	HCgraph	- . -
0.3151	0.489	0.25	–	. . .
–	0.589	0.35	–	- -

Table 7.1: Variables with corresponding line types on graphs

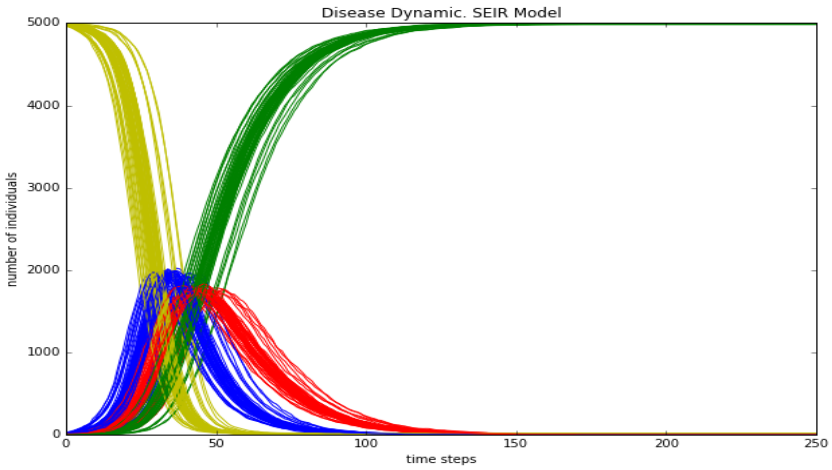


Figure 7.2: SEIR model. Visualisation of compartment sizes at every time step for fifty simulation runs.

This is just an example with specific parameters, however, it gives a general idea of the disease dynamic during multiple runs. Once the data was averaged across simulation runs every compartment could be represented by a single line on a graph as shown below.

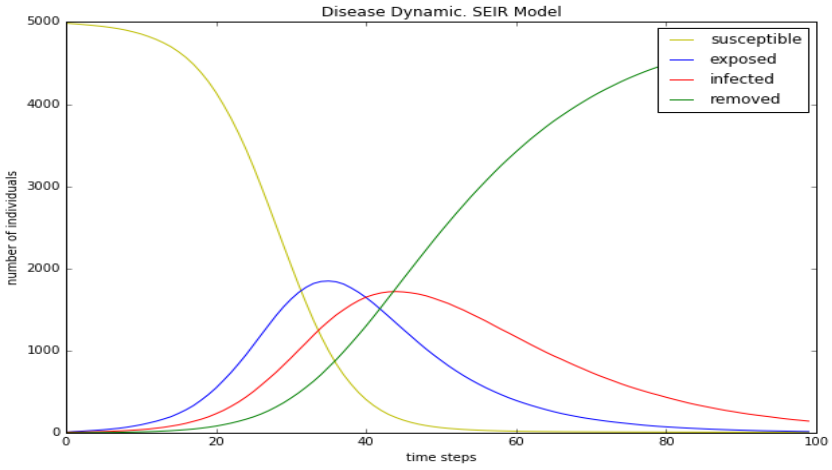


Figure 7.3: SEIR model. Visualisation of averaged compartment sizes at every time step.

Simulations with the same parameters were performed on two different graphs and according to the data, there is a significant difference in the time frame of the epidemic duration. The graph below shows that for all β values it took much more time for the epidemic to spread across the population when HCgraph rather than Erdős–Rényi network was chosen as an underlying population structure.

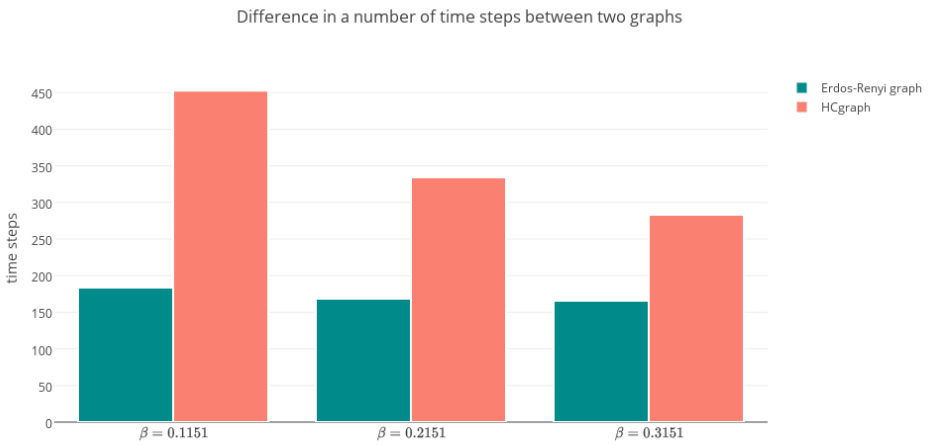


Figure 7.4: Total number of time steps for Erdős–Rényi network and HCgraph with different β values.

On the example where $\beta = 0.1151$ the averaged changes in the compartment sizes are plotted for Erdős–Rényi graph against HCgraph. This also serves as an evidence of a slower flow of disease spread on HCgraph.

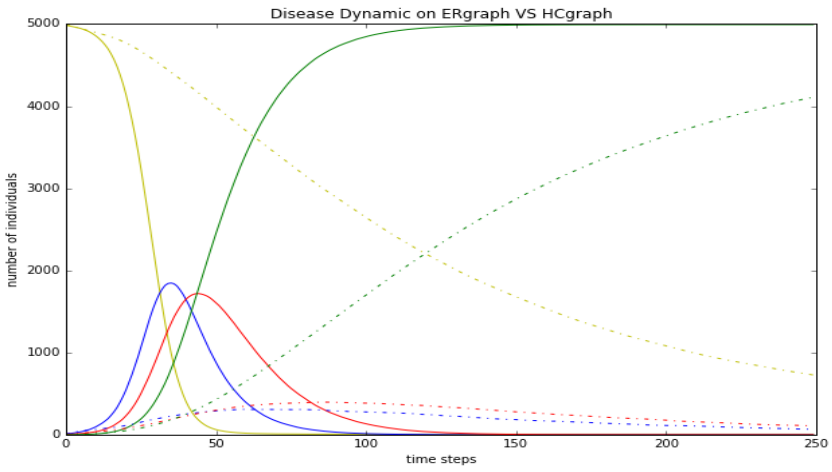


Figure 7.5: Disease dynamic on Erdős–Rényi graph and HCgraph with $\beta = 0.1151$

This demonstrates how the structure of a graph affects the dynamic in many of its aspects. It does not influence the total number of transitions between compartments (figure 7.6) since the epidemic in both cases involves approximately the same number of individuals and the same stages have to be passed to move from S to R compartments, nevertheless, the largest number of people who simultaneously have the disease is substantially different as well as the threshold value of S after which the epidemic starts to die out. On the graph above these variables are 3320.46 against 698.48 and 591.84 against 3039 for Erdős–Rényi graph and HCgraph respectively.

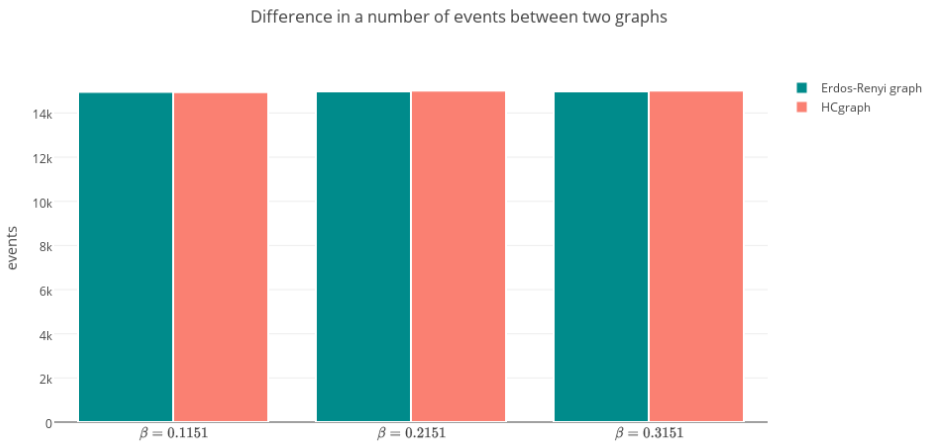


Figure 7.6: Total number of events for Erdős–Rényi network and HCgraph with different β values.

Moving on to the changes in disease parameters, the only one that was varied while simulating disease dynamic using SEIR model was β. Thus, the difference in the dynamic with simulations on both graphs is presented below.

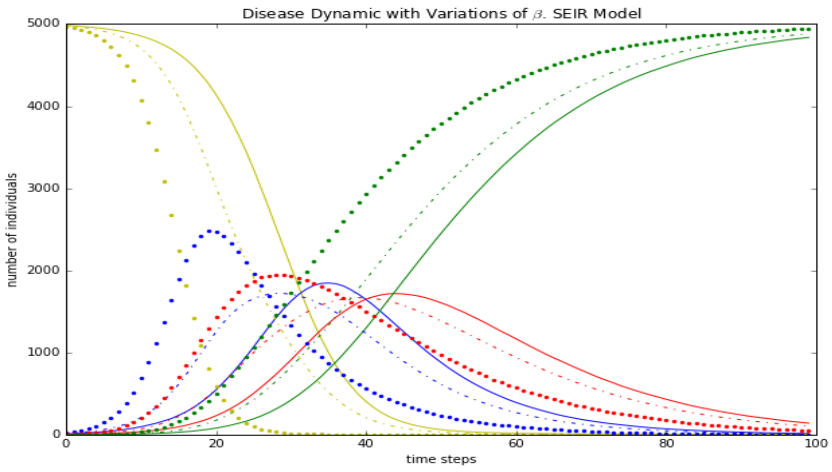


Figure 7.7: Disease dynamic on Erdős–Rényi graph with different β values

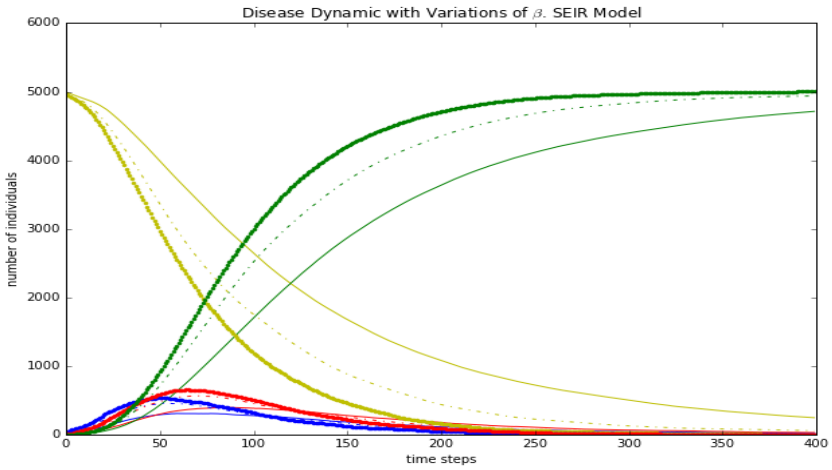


Figure 7.8: Disease dynamic on HCgraph with different β values

Naturally, the higher the value of β parameter gets, the sooner the epidemic reaches its peak and starts to gradually fade away. On Erdős–Rényi graph the time steps where the peak occurs are 38, 34 and 22 and on the HCgraph 84, 65 and 58. Nevertheless, unlike in the results from mathematical modelling, higher β value with the same γ value when running simulations on networks does not necessarily mean that the threshold value of S will be lower. In fact, when simulating the spread on HCgraph it increases from 698.48 to 1003.08 and then to 1163.12 for β values 0.1151, 0.2151 and 0.3151 respectively while on Erdős–Rényi graph that threshold first slightly decreases from 3320.46 to 3167.09 and then grows to 3991.86. This might be explained by the fact that when carrying out simulations on networks the increased rate of infection spread does not always appear through the enlarged threshold value of S, but an earlier peak of an outbreak. Moreover, there are many factors that need to be considered and although repeated runs should minimise variations in results the chosen number of those runs might not be sufficient for the number of random components present.

SEIDR Model

The dynamic of the disease on the SEIDR model is described by the changes happening in five compartments, thus, the graphs with all runs and with averaged results now have an extra data set as shown below. The example chosen has the following parameter values $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \delta = 0.489, \zeta = 0.5, \epsilon = 0.2, rewiring\_degree = 0.15$ all of which except the degree of rewiring correspond to the parameters determined in the research by Z. Xia et al (2015).

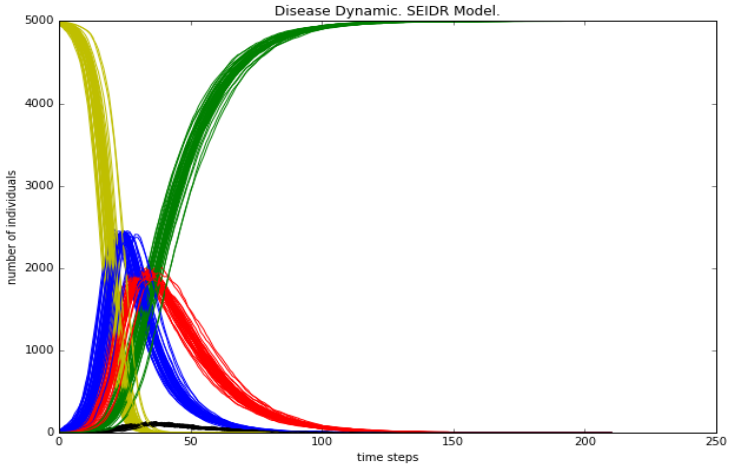


Figure 7.9: SEIDR model. Visualisation of compartment sizes at every time step for fifty simulation runs.

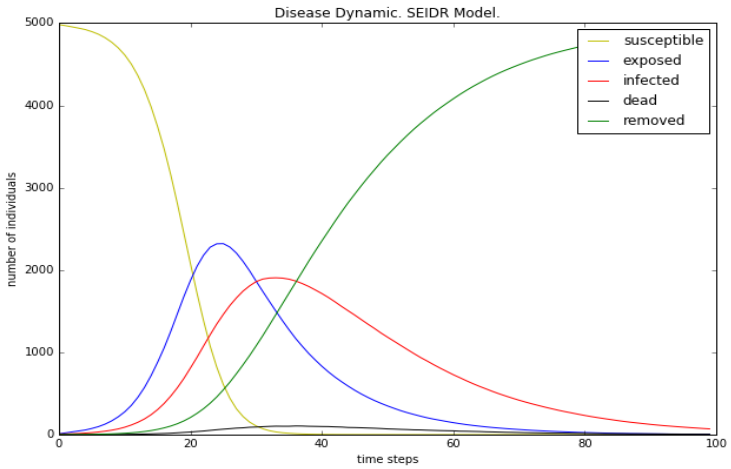


Figure 7.10: SEIDR model. Visualisation of averaged compartment sizes at every time step.

Following the same process as in the SEIR model description, the dynamic of the disease is compared between two graphs. With SEIDR model the time frame when the epidemic lasted was again longer when running on the HCgraph rather than on the Erdős–Rényi model with every combination of different values of β, δ and *rewiring_degree* which is especially noticeable when $rewiring\_degree = 0.0$ (figure 7.11).

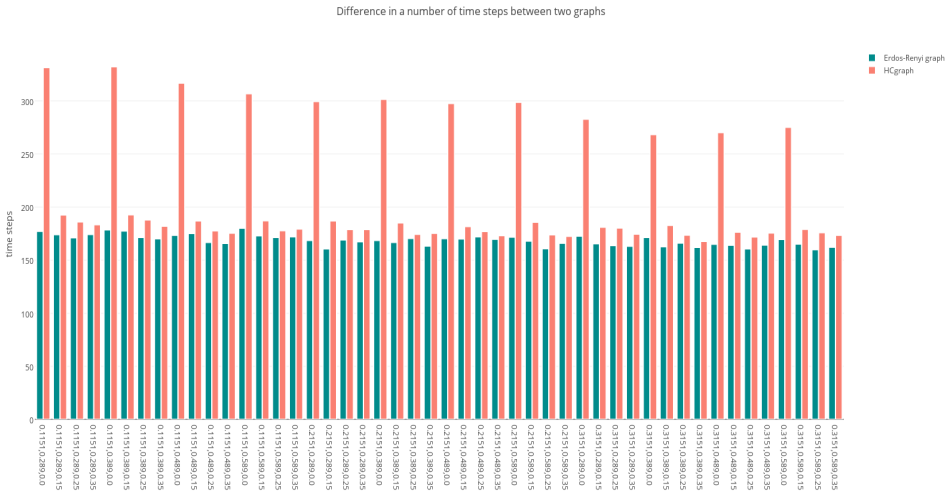
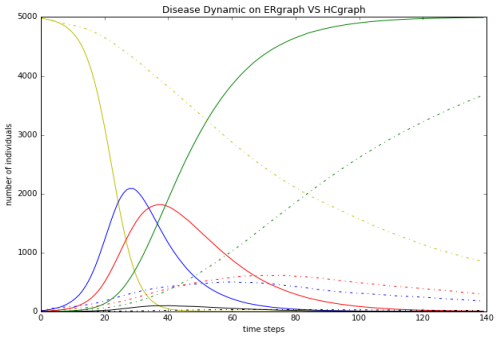
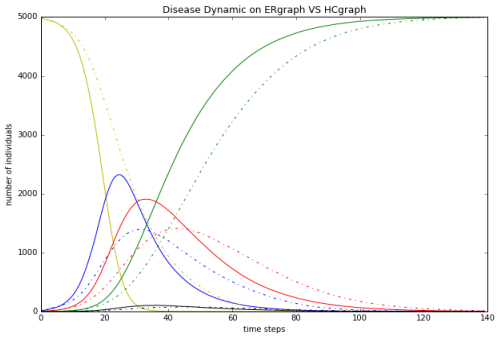


Figure 7.11: Total number of time steps for Erdős–Rényi network and HCgraph with different β, δ and *rewiring_degree* values.

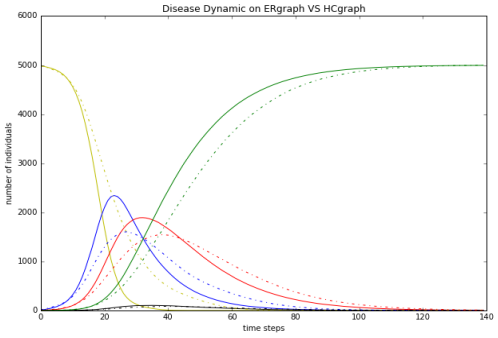
This might be explained by local connections on the HCgraph, which means that the disease has to distribute gradually through households and then communities while on the random graphs it can instantly travel long distances. Furthermore, the network rewiring speeds up the spread on HCgraph as it connects people from other households but the same community to a dead individual. The disease dynamic with $\beta = 0.1151, \delta = 0.489$ and different rewiring degrees is displayed below.



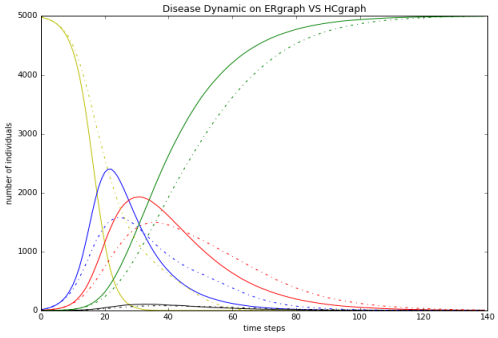
(a) $rewiring\_degree = 0.0$



(b) $rewiring\_degree = 0.15$



(c) $rewiring\_degree = 0.25$



(d) $rewiring\_degree = 0.35$

Figure 7.12: Disease dynamic on Erdős–Rényi graph and HCgraph with $\beta = 0.1151, \gamma = 0.06851662$, $\eta = 0.083333, \delta = 0.489, \zeta = 0.5$ and $\epsilon = 0.2$

The next step is contrasting results with a different combination of parameters. Although the data was analysed for each of them the outcomes were similar, therefore, representative examples were chosen to be displayed here and the rest can be found in the appropriate folders attached.
Changing the value of β shows the same trend as when SEIR model was considered, the higher the contact rate between infected and susceptible individuals together with the corresponding transmission probability, the faster the disease spreads (figures 7.13 and 7.14).

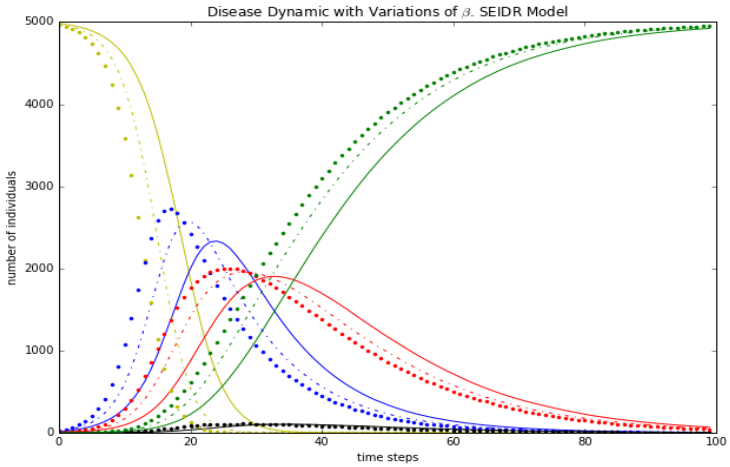


Figure 7.13: Disease dynamic on ERgraph with $\gamma = 0.06851662, \eta = 0.083333, \delta = 0.489$, $\zeta = 0.5, \epsilon = 0.2, rewiring\_degree = 0.15$ and different β values.

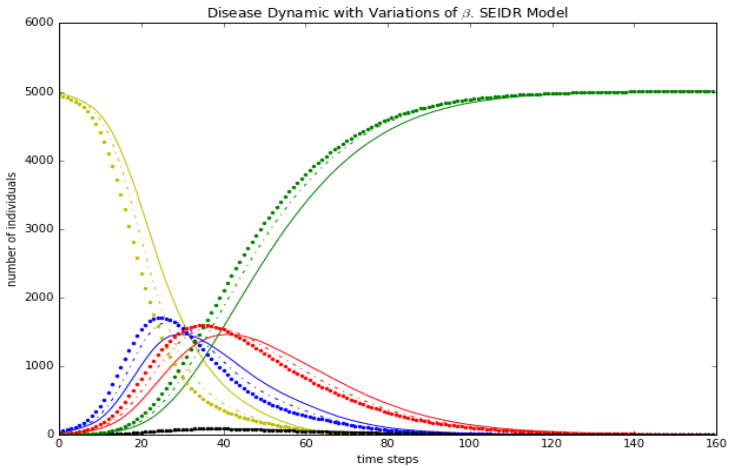


Figure 7.14: Disease dynamic on HCgraph with $\gamma = 0.06851662, \eta = 0.083333, \delta = 0.489$, $\zeta = 0.5, \epsilon = 0.2, rewiring\_degree = 0.15$ and different β values.

Same applies to the tendency that describes the impact of *rewiring_degree* on the speed of an epidemic. Moreover, the difference on the HCgraph is substantial when no rewiring is done compared to the degree of 0.15 and then it becomes moderate as the rewiring degree further increases(figure 7.16) while on the Erdős–Rényi graph changes happen more insensibly (figure 7.15). The explanation might lie as before in the locality of connections on the HCgraph which is decreased when rewiring is introduced.

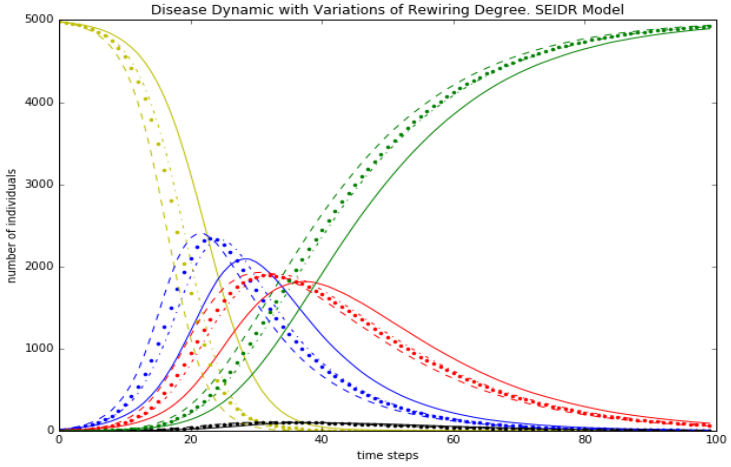


Figure 7.15: Disease dynamic on ERgraph with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \delta = 0.489$, $\zeta = 0.5, \epsilon = 0.2$ and different *rewiring_degree* values.

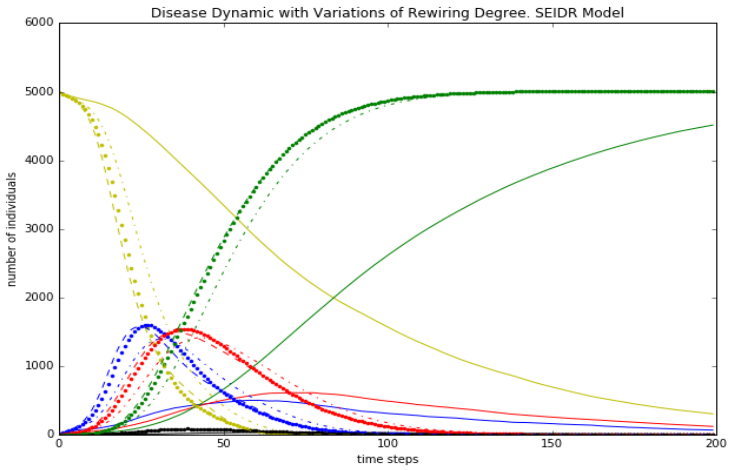


Figure 7.16: Disease dynamic on HCgraph with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \delta = 0.489$, $\zeta = 0.5, \epsilon = 0.2$ and different *rewiring_degree* values.

Furthermore, the effect of alterations of the δ variable is looked into. It was discovered that with all other parameters staying unchanged it produces the least impact (figures 7.17 and 7.18) especially when simulations were run on the HCgraph.

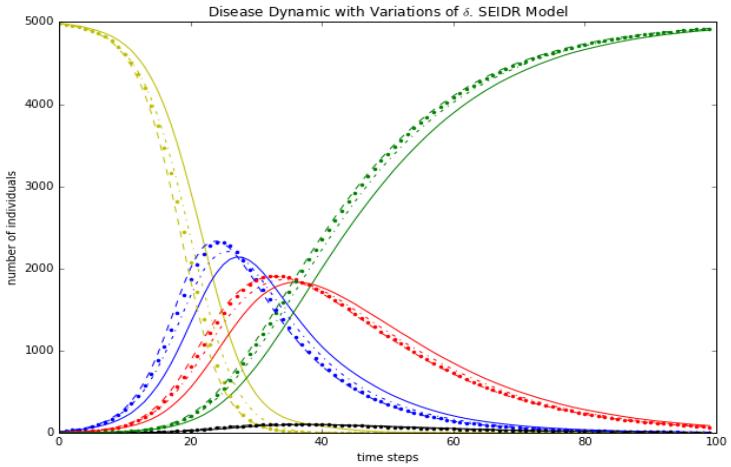


Figure 7.17: Disease dynamic on ERgraph with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5$, $\epsilon = 0.2, rewiring\_degree = 0.15$ and different δ values.

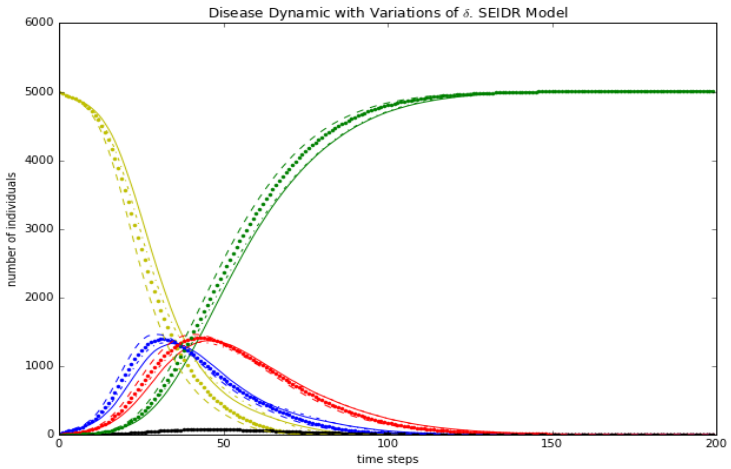


Figure 7.18: Disease dynamic on HCgraph with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5$, $\epsilon = 0.2, rewiring\_degree = 0.15$ and different δ values.

Finally, the difference in the number of infection transmissions from compartments I and D as sources is looked into. On both graphs the picture observed is similar. It is quite obvious that as the δ value and the value of the $rewiring_degree$ grow, the number of transmissions from people in the dead compartment increases. Moreover, as δ and $rewiring\_degree$ increase the gap in the number of virus transfers between the two compartments gets smaller. This is understandable since δ represents the contact rate between people in the dead and susceptible compartments times the probability of disease transmission between them and the $rewiring\_degree$ increases connectivity of a dead person.
Despite the fact that with all combination of parameters more people get the disease from infected individuals, the existence of D compartment still affects the overall dynamic significantly since this number of infection transfers from dead people does not replace the transmissions from the infected humans, but is added to it, thus, increasing the total number of transmissions.

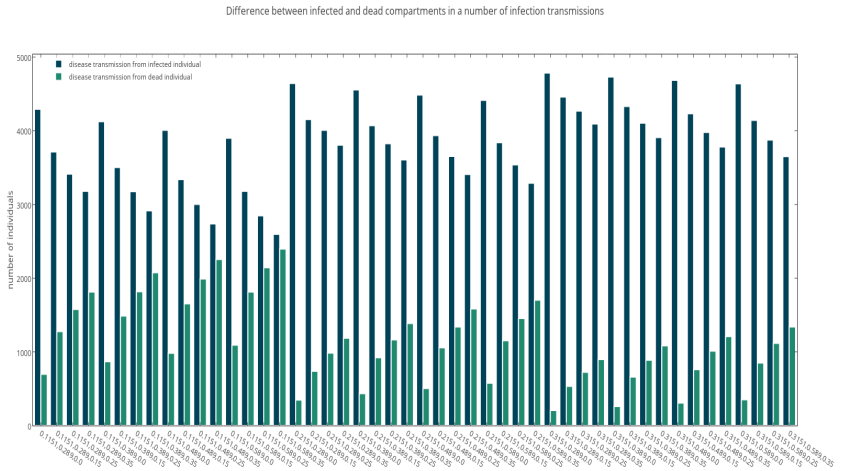


Figure 7.19: Number of infection transmissions from I VS D compartments on the ERgraph

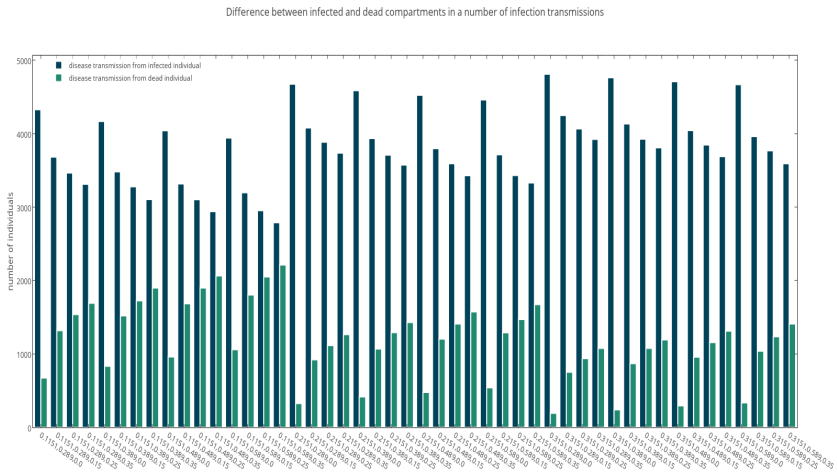


Figure 7.20: Number of infection transmissions from I VS D compartments on the HCgraph

SEIR and SEIDR Models

When the number of runs was chosen the decision was based on the variation in the changes of compartmental sizes. Therefore, the outbreak size was not taken into account. Once the data was fully analysed, it was found that the number of runs selected was not sufficient to minimise the variations in the results connected with outbreak size, thus, the averaged number produced could not be a good representative of the data set.

Further in this section, the comparisons between the SEIR and the SEIDR models will be made with all parameters being equal and except for the *rewiring_degree*, δ and β. It was found earlier using mathematical modelling that from the disease parameters that were added when the D compartment was introduced the value of δ made the biggest difference while changes in ϵ and ζ only slightly affected results. Therefore, the initial values of these variables that were taken from the Z. Xia et al. research stay the same (2015).

The most interesting cases here were discovered when the differences between the models were looked into when the β value was changed.

With initial $\beta = 0.1151$ from the research the disease spread quicker even when there was no rewiring occurring and with different δ values especially on the HCgraph (figures 7.23 and 7.24). Another evidence for this was produced when the peak of an outbreak was analysed. At that point, the number of people who are contagious is much smaller and the peak is reached later when SEIR rather than SEIDR model is used. The graphs below demonstrate this on the example with the following parameters $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \delta = 0.489, \zeta = 0.5, \epsilon = 0.2$ and different *rewiring_degree* and when δ value is changed similar results were found.

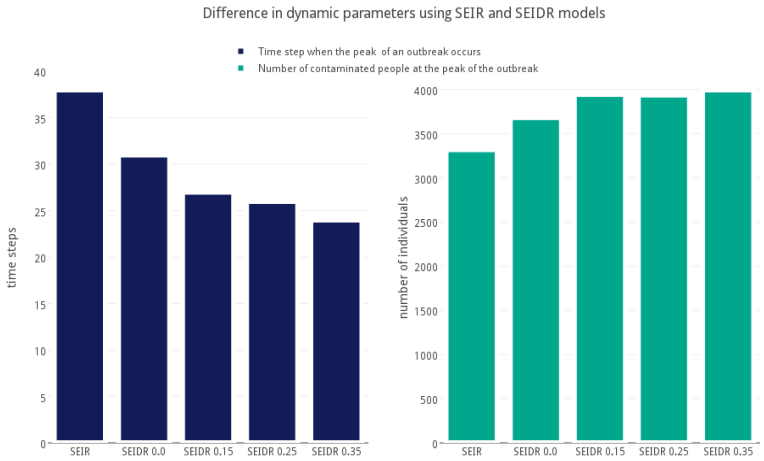


Figure 7.21: Disease dynamic on ERgraph with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \delta = 0.489$, $\zeta = 0.5, \epsilon = 0.2$ and different *rewiring_degree* values.

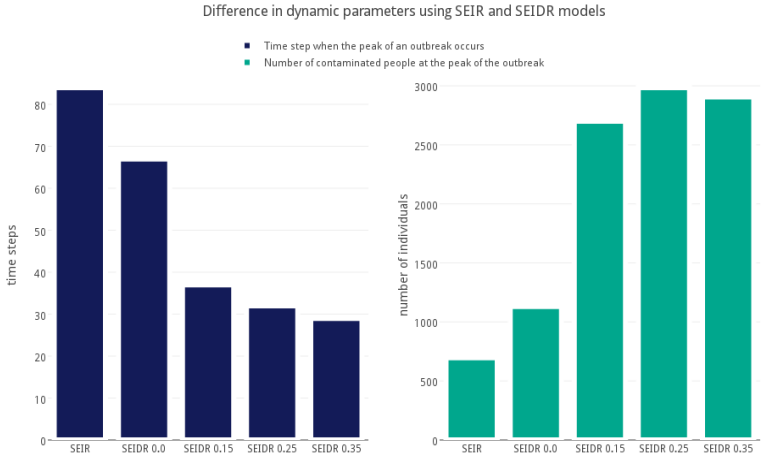
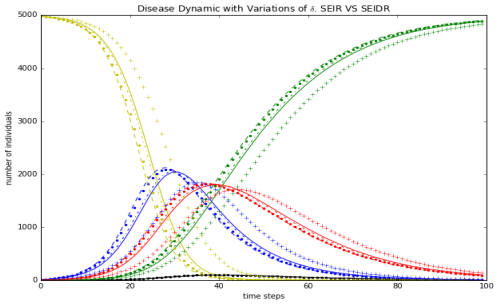
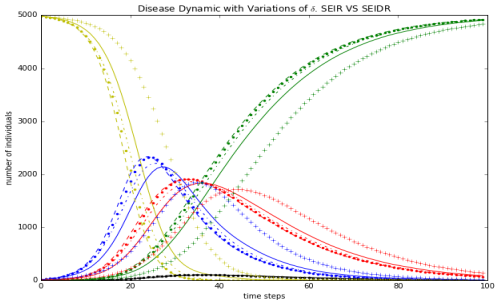


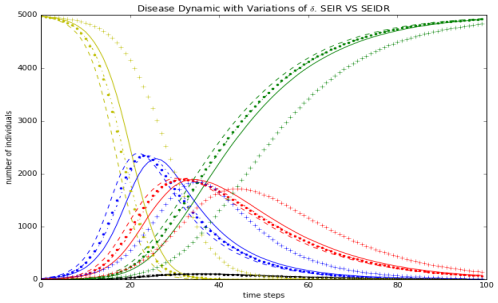
Figure 7.22: Disease dynamic on HCgraph with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \delta = 0.489$, $\zeta = 0.5, \epsilon = 0.2$ and different *rewiring_degree* values.



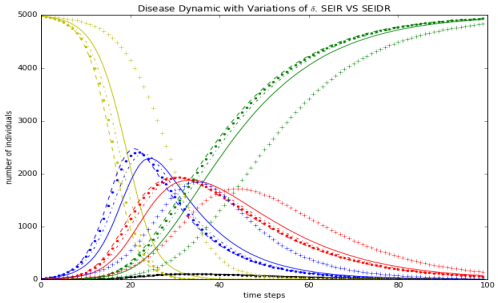
(a) $rewiring\_degree = 0.0$



(b) $rewiring\_degree = 0.15$

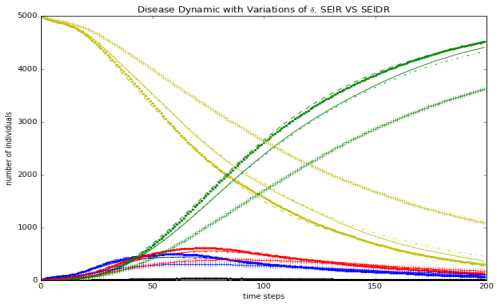


(c) $rewiring\_degree = 0.25$

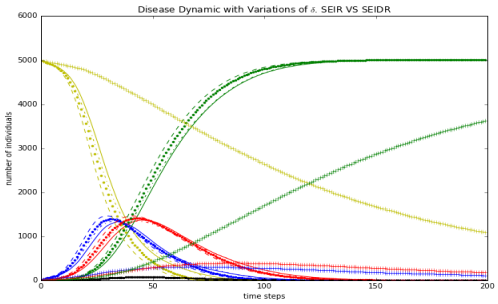


(d) $rewiring\_degree = 0.35$

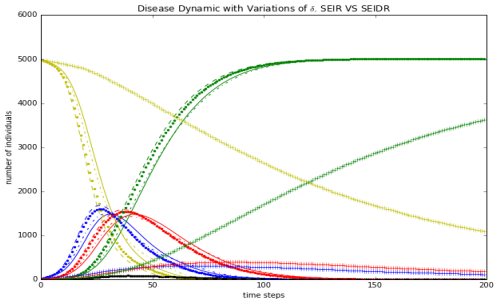
Figure 7.23: Disease dynamic on ERgraph using SEIR(+) and SEIDR (—, - . -, . . ., - -) models with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5, \epsilon = 0.2$ and different δ values.



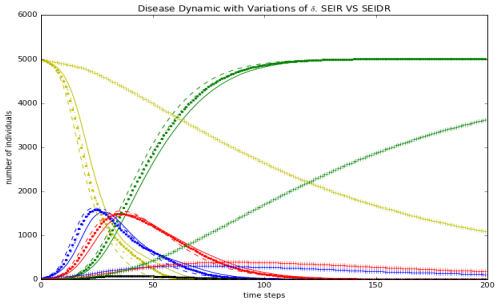
(a) $rewiring\_degree = 0.0$



(b) $rewiring\_degree = 0.15$



(c) $rewiring\_degree = 0.25$



(d) $rewiring\_degree = 0.35$

Figure 7.24: Disease dynamic on HCgraph using SEIR(+) and SEIDR (—, - . -, . . ., - -) models with $\beta = 0.1151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5, \epsilon = 0.2$ and different δ values.

Meanwhile, when β increases to 0.3151 the difference observed when simulations were run on the ERgraph is not that noticeable even with the highest value of $\delta = 0.589$ and the highest $rewiring\_degree = 0.35$. Furthermore, when no rewiring is done SEIR curves follow SEIDR curve with $\delta = 0.289$ almost exactly and there is just a small gap between curves with other δ values (figure 7.27) which means that the rate

of the disease spread is almost the same. Similar conclusions can be made when the peak values are investigated with various δ values and no rewiring. Distinctions between SEIR and SEIDR results are minor on the HCgraph (figure 7.26) and even less significant on the Erdős–Rényi network (figure 7.25). Nevertheless, on the HCgraph, when the rewiring is introduced, the spread speeds up significantly so that even higher value of β does not even curves out (figure 7.28).

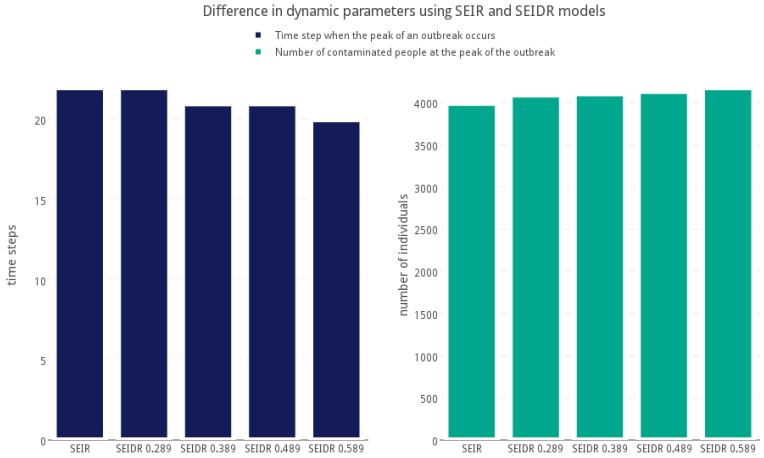


Figure 7.25: Disease dynamic on ERgraph with $\beta = 0.3151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5$, $\epsilon = 0.2, rewiring\_degree = 0.0$ and different δ values.

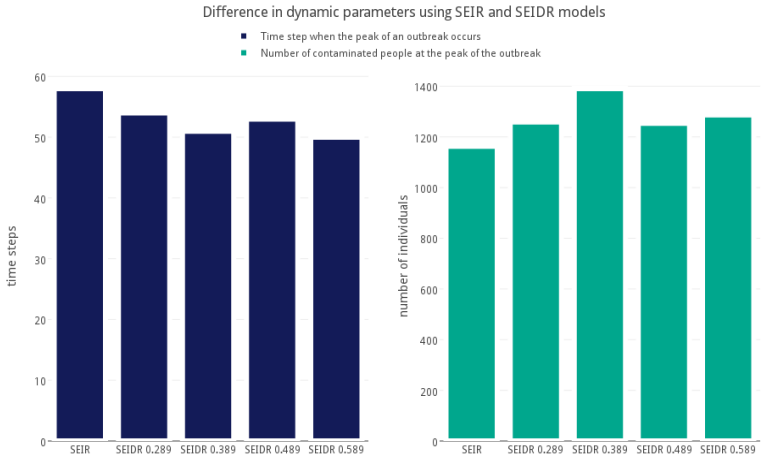
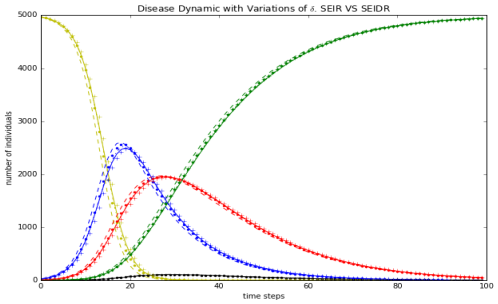
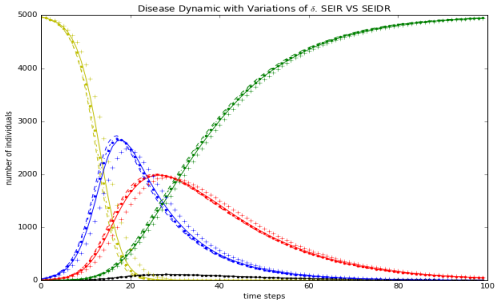


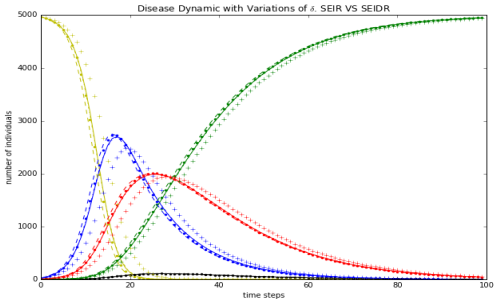
Figure 7.26: Disease dynamic on HCgraph with $\beta = 0.3151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5$, $\epsilon = 0.2, rewiring\_degree = 0.0$ and different δ values.



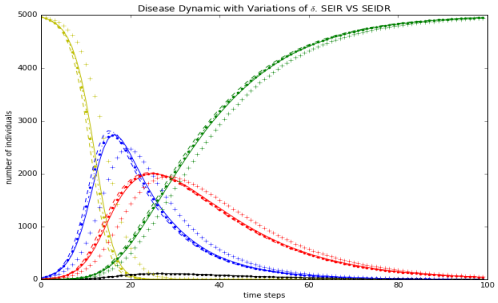
(a) $rewiring\_degree = 0.0$



(b) $rewiring\_degree = 0.15$

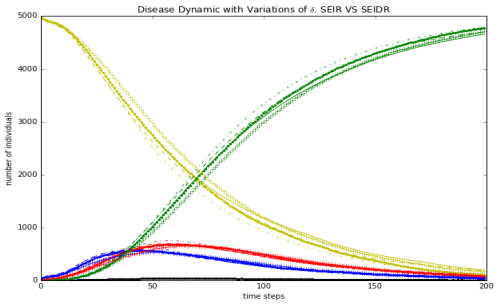


(c) $rewiring\_degree = 0.25$

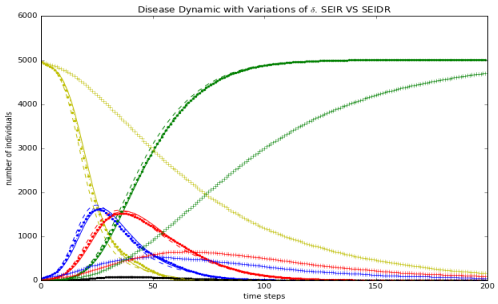


(d) $rewiring\_degree = 0.35$

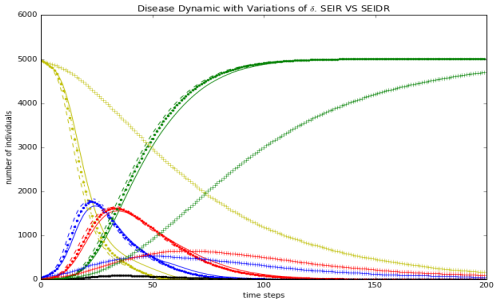
Figure 7.27: Disease dynamic on ERgraph using SEIR(+) and SEIDR (—, - . -, . . ., - -) models with $\beta = 0.3151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5, \epsilon = 0.2$ and different δ values.



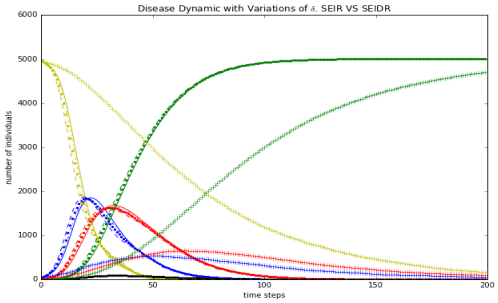
(a) $rewiring\_degree = 0.0$



(b) $rewiring\_degree = 0.15$



(c) $rewiring\_degree = 0.25$



(d) $rewiring\_degree = 0.35$

Figure 7.28: Disease dynamic on ERgraph using SEIR(+) and SEIDR (—, - . -, . . ., - -) models with $\beta = 0.3151, \gamma = 0.06851662, \eta = 0.083333, \zeta = 0.5, \epsilon = 0.2$ and different δ values.

Chapter 8

Discussion

The experiments carried out were based on a bounded set of parameters all of which were either directly taken from the previous research on Ebola virus disease or were slightly changed to minimise the amount of bias that essentially occurs when variables are chosen. Taking an entirely different set could provide with more information about models behaviour and confirm or disprove conclusions made.
Furthermore, the limited power of the machine used together with the fixed amount of time influenced the decision made about the number of repeated simulation runs which in its turn affected some of the results collected. As mentioned earlier, there was not enough data collected for some sets and further work is required to analyse other aspects of the disease spread.
Nevertheless, the results obtained do show that the rate at which the disease progresses is always higher for SEIR model than SEIDR model when network rewiring occurs and is generally higher even without rewiring if the value of β is not amplified. While the former observation was made based on the simulation part of the research, the latter one was noted in mathematical modelling as well. Thus, the main conclusion made while analysing mathematical models was that with a fully mixed population the D compartment speeds up the disease outbreak while on the networks an important observation was that the higher the rewiring degree is, the more humans get infection from dead people, the more rapidly the disease spreads which was demonstrated on the graphs for two models that showed the disease dynamic and the comparisons of how early the peak of an outbreak is reached and how many people have the disease at that point.
Since two network types that are not similar in their design were considered, it minimised the possibility that the population structure affected the behaviour of models. However, other network types could be used to become entirely convinced as it was discovered that graph features do influence the significance of the rate change. Incorporation of the new compartment together with rewiring to the simulation on the Erdős–Rényi model produced less effect than on the HCgraph. The reason for this might lie in the high prevalence of short distance connections, which makes the disease spread gradually until network rewiring takes place.
Moreover, every disease parameter considered had a different effect on the rate of the disease progression. Thus, the *rewiring_degree* produced a greater impact than δ, while significantly larger β value could outweight the all the rewiring done even with an even larger δ value. Nevertheless, such value of β is quite unrealistic at least for now for Ebola virus disease. Therefore, these result prove the suggested hypothesis. The introduction of the D compartment with rewiring process on networks does not change whether an epidemic occurs or not, since it is connected with the transmissibility, but it does change the rate at which the epidemic breaks out. Although what was found as the most important parameters for the contrast of the two models were considered, looking into how other parameters influence the dynamic might have provided with some interesting insight as well. Hence, those could be considered as a part of the improvement plan for this research.
Finally, although the number of humans who get the disease from the individuals in the infected compartment is mostly much larger than the number of those whose source of infection is people in the dead compartment, the additional flow of infection cannot be neglected as this is the reason why the epidemic speeds up.
To summarise, results obtained from mathematical modelling and simulations on networks showed that the rate at which the disease spreads does generally increase with the introduction of the new D compartment except for the occasion when β value is very high. For now this is not the case, which means that the elimination of traditional African funerals does in theory slow down the spread of an epidemic. However, there is no way to find out what value the parameter β will have next time the outbreak occurs and whether the results will hold in the future.

Bibliography

Bansal, S., Grenfell, B.T., L.A. Meyers, 2007. When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society*, [online] Available at: <http://rsif.royalsocietypublishing.org/content/4/16/879> [Accessed 15 December 2015].

Bansal, S., et al., 2010. The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, [online] Available at: <http://www.tandfonline.com/doi/full/10.1080/17513758.2010.503376> [Accessed 15 December 2015].

Bernovskiy, M.M., Kuzyurin, N.N., 2012. Random graphs, models and generators of scale-free graphs. *Russian Academy of Sciences Institute of System Programming*, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3604768/> [Accessed 4 March 2016].

Brachman, P.S., 1996. *Medical Microbiology.* [e-book] Galveston: University of Texas Medical Branch at Galveston. Available through: The National Center for Biotechnology Information website <http://www.ncbi.nlm.nih.gov/books/NBK7993/> [Accessed 4 March 2016]

Chowell, D., et al., 2015. Modelling the effect of early detection of Ebola. *The Lancet Infectious Diseases*, [online] Available at: <http://www.thelancet.com/journals/laninf/article/PIIS1473-3099%2814%2971084-9/fulltext> [Accessed 10 March 2016].

Dobson, S.A., 2015. *Complex Networks. Complex Processes.* [online] Available at: <http://www.simondobson.org/complex-networks-complex-processes/complex-networks-complex-processes.pdf> [Accessed 10 October 2015].

Grishunina, J.B. et al., 2014. Modelling of the epidemic situation taking into account external risks. *Epidemiology and Immunization N5*, [online] Available at: <https://www.hse.ru/pubs/share/direct/document/136317640> [Accessed 6 March 2016].

Keeling, M.J., Eames, K.T.D., 2005. Networks and epidemic models. *Journal of the Royal Society*, [online] Available at: <http://leonidzhukov.ru/hse/2013/socialnetworks/papers/rsif20050051.pdf> [Accessed 25 December 2015].

Kiskowski,M.A., 2014. A Three-Scale Network Model for the Early Growth Dynamics of 2014 West Africa Ebola Epidemic. *PLOS. Current Outbreaks*, [online] Available at: <http://currents.plos.org/outbreaks/article/a-three-scale-network-model-for-the-early-growth-dynamics-of-2014-west-africa-ebola-epidemic/> [Accessed 10 November 2015].

Meltzer, M.I., at al., 2014. Estimating the Future Number of Cases in the Ebola Epidemic — Liberia and Sierra Leone, 2014–2015. *Centers for Disease Control and Prevention*, [online] Available at: <http://www.cdc.gov/mmwr/preview/mmwrhtml/su6303a1.htm> [Accessed 12 January 2016].

Murray, J.D., 2002. *Mathematical Biology:I. An Introduction, Third Edition.* [online] Available at: <http://www.ift.unesp.br/users/mmenezes/mathbio.pdf> [Accessed 20 October 2015].

Newman, M.E.J., 2002a. Random Graphs as Models of Networks. *SFI Working Paper*, [online] Available at: <http://www.santafe.edu/media/workingpapers/02-02-005.pdf> [Accessed 12 January 2016].

Newman, M.E.J., 2002b. *The spread of epidemic disease on networks.* [online] Available at: <http://arxiv.org/pdf/cond-mat/0205009.pdf> [Accessed 10 November 2015].

Pérez,F., Granger, B.E., 2007. IPython: A System for Interactive Scientific Computing, *Computing in Science and Engineering*, [online] Available at: <http://ipython.org> [Accessed 20 March 2016].

Rachah, A., Torres, D.F.M., 2015. Mathematical Modelling, Simulation, and Optimal Control of the 2014 Ebola Outbreak in West Africa. *Discrete Dynamics in Nature and Society*, [online] Available at: <http://www.hindawi.com/journals/ddns/2015/842792/> [Accessed 15 January 2016].

Shang, Y., 2012. SEIR Epidemic Dynamics in Random Networks. *Epidemiology*, [online] Available at: <http://www.hindawi.com/journals/isrn/2013/345618/> [Accessed 4 March 2016].

Telesford, Q.K. et al., 2011. The Ubiquity of Small-World Networks. *Brain Connectivity*, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3604768/> [Accessed 13 January 2016].

Volz, E., Pond, S., 2014. Phylodynamic Analysis of Ebola Virus in the 2014 Sierra Leone Epidemic. *PLOS. Current Outbreaks*, [online] Available at: <http://currents.plos.org/outbreaks/article/phylodynamic-analysis-of-ebola-virus-in-the-2014-sierra-leone-epidemic/> [Accessed 10 January 2016].

Wittena, G., Poultera, G., 2005. Simulations of infectious diseases on networks. *Computers in Biology and Medicine*, [online] Available at: <http://www.leonidzhukov.net/hse/2015/networks/papers/witten-poulter-2006.pdf> [Accessed 24 November 2015].

World Health Organization, 2016a. *Epidemiology.* [online] Available at: <http://www.who.int/topics/epidemiology/en/> [Accessed 4 March 2016].

World Health Organization, 2016b. *Ebola virus disease.* [online] Available at: <http://www.who.int/topics/epidemiology/en/> [Accessed 4 March 2016].

Xia, Z. et al., 2015. Modeling the transmission dynamics of Ebola virus disease in Liberia. *Nature*, [online] Available at: <http://www.nature.com/articles/srep13857> [Accessed 10 January 2016].