Global disease transmission

What if SARS would have been more lethal?

Erik Lagerstedt

June 1, 2011

Bachelor thesis

Department of
Astronomy and Theoretical Physics
Lund University

Abstract

Global disease transmission is an important area, since knowledge of the time of the onset of an epidemic in a region might help authorities to save many lives. In this thesis I have used a deterministic model with eight parameters, of which two were time dependent, for spreading of the SARS disease. My model is an extension of a model by Colizza and co-workers, where I have added a Heaviside step function, that prevent latent people to interact if they are to few. I have also added a "repulsion factor", that reduces the traffic to an infected city. In my model I used data on how 3,147 of the airports of the world are connected. The strength of the connections were not given, so they were estimated, using the population sizes of the cities. The corresponding rate equations was solved by computer programs, written by me in Java, one of which displayed the result on a map. A major finding of this thesis is that vaccination can be used to prevent an epidemic from fully break out, and the doses are preferably concentrated to well connected regions, although the effect is not dramatical. This is demonstrated by comparing the effect of vaccinating London, with the effect of vaccinating Tianjin, and with the effect of vaccinating New York. I also demonstrate that the latency of a disease is an important factor for mortality of a disease.

Acknowledgements

I would like to thank my supervisor Tobias Ambjörnsson for all the help and support he has provided. I would also like to thank Robert Guimerà for giving me access to the air traffic connectivity matrix and the airport data. Finally I would like to thank Dirk Brockmann for his insightful feedback and helpful ideas during his visit to Lund May 2011.

Contents

1 Introduction **4**

2 Methods **5**
- 2.1 Models . . . 5
  - 2.1.1 SIS . . . 5
  - 2.1.2 SIR . . . 6
  - 2.1.3 Colizza's SARS model . . . 7
  - 2.1.4 My SARS model . . . 10
  - 2.1.5 Fractions of individuals . . . 11
  - 2.1.6 The repulsion factor . . . 11
  - 2.1.7 Network connection strength . . . 12
  - 2.1.8 Population division . . . 14
  - 2.1.9 Time dependent parameters . . . 14
- 2.2 Program . . . 15
  - 2.2.1 Two city model . . . 15
  - 2.2.2 Graphic output . . . 17
  - 2.2.3 Maps and coordinates . . . 18
  - 2.2.4 Data generator . . . 18

3 Results **19**
- 3.1 Simulations with Colizza's parameters . . . 19
- 3.2 Adjusting to the measured data . . . 19
- 3.3 Vaccination . . . 21
- 3.4 Exploring the phase space . . . 22

4 Discussion **25**
- 4.1 Stochastic models versus deterministic models . . . 25
- 4.2 Psychology . . . 25
- 4.3 Alternative models for connection weights . . . 26

5 Conclusions **27**

6 Outlook **28**

A Supplementary information **31**
- A.1 Analytic solution to equation (32) . . . 31
- A.2 Derivation of equation (32) . . . 32

1 Introduction

By understanding the mechanisms of diseases spreading, countermeasures can be optimised so that epidemics can be prevented, and the number of effected people reduced. The epidemics problem can be divided into two parts, the behaviour in a population and the interaction between populations[1]. The last century bullet trains and air-planes have made it possible to travel so far and fast, that it is appropriate to look at the entire world population as one big population, with the regions close to the stations and airports as subpopulations. A few research groups have looked into this aspect of epidemic modelling, both to analyse specific cases of epidemics[2] and to be able to predict future epidemics[1].

What is required of a disease to make big enough impact to be considered an epidemic, and is it possible to prevent an epidemic if it appears? Is it possible to find the strength of the individual couplings between cities (i.e. number of persons travelling per day) given only connections and data from known diseases? These are examples of questions that I attempt to answer in this thesis. The first of the two questions are probably the most interesting, since it might be a support to people who try to prioritise among limited resources and counteractions, in case of possible epidemics. Among the possible counteractions are to vaccinate parts of the population, or quarantine some cities or regions.

In this thesis I will look into the global transmission of Severe Acute Respiratory Syndrome (SARS), which is a disease caused by the corona-virus SARS-CoV[3]. The first cases of SARS was found in Guangdong Province, China, in November 2002, but were only identified as such in April 2003[4]. On February 21 an infected medical doctor carried it to Hong Kong, where it spread to others[4]. During the spring of 2003 SARS turned into a global epidemic, with totally 8,096 confirmed cases, of which 774 were fatal[5].

This thesis is organised as follows: in section 2.1 different models, used for simulating spreading of diseases, are presented, starting with the simplest and after adding some complexity finally arrives at the model that I used. This model include the travel between 3,147 cities, inhabited by totally 998,328,995 people, and I will try to capture some of the psychological effects (such as changes in travelling patterns) caused by an epidemic. Previous studies usually ignore this effect, so it is still a rather unexplored, yet crucial, aspect of disease dynamics. In section 2.2 the programs (whose code can be found in the appendix) I used will be introduced. Section 3 will contain the results I have produced, and they will be discussed in section 4. Section 4 will also contain some discussion on choices for parameters and models, that I have made, and some alternative methods. My main conclusions will be given in section 5, and in section 6 I will briefly mention things that might be interesting to look at in the future.

2 Methods

2.1 Models

Disease transmission can be modelled as a stochastic process, for example by using a Gillespie algorithm[6], where the probability for every event are decided, and with some help from randomly generated numbers the next time step is calculated. In this thesis, however, deterministic rate equations will be discussed, that is, only mean values of population sizes are considered. The population will be divided into different categories, and the mean number of members of every category will change with a certain rate.

2.1.1 SIS

The simplest model for disease spreading is the SI model[7], in which S is the number of susceptible members of the population and I is the number of infected members of the population, i.e. this model has two categories. The susceptible part of the population will, through encounters with infected people, gradually turn into infected, and will then stay that way, unless a feedback term is added, turning the SI model to a SIS model. This later model is described by the following differential equations:

$$\partial_t S = -\beta \frac{SI}{N} + \alpha I \tag{1}$$

$$\partial_t I = \beta \frac{SI}{N} - \alpha I \tag{2}$$

where $N = S + I$ is the total population size. From equations (1) and (2) we see that

$$\partial_t N = \partial_t S + \partial_t I = 0 \tag{3}$$

that is, the total population stays constant. Equations (1) and (2) describe how the number of people in the different categories changes with time. The first term on the right-hand sides remove people from the category denoted S, and add them to the category denoted I, thus it describes people getting infected. The more susceptible people present, the more people will get infected, if there are any infectious people. A higher concentration of infectious people in the population should make it more probable to meet someone infectious, and thus be infected. This is the reason for the I/N-factor. β is a rate constant, with dimension inverse time, characterising the disease. The second terms are the feedback terms that removes people from the category denoted I, and add them back to the category denoted S, that is, people recover from the disease. Since people recover by themselves, independent of the other people in the population, these terms are only proportional to the number of infected. The rate constant α, characterising the disease, is the inverse of the average recovery time. For a disease from which it is impossible to recover, $\alpha = 0$, we have the SI model. It can be seen as the disease has an infinite recovery time.

From an SIS model, it is easy to calculate the "basic reproductive ratio", R_0, which is the average number of secondary cases produced by an average

infectious individual in a totally susceptible population, and it is defined by equation (4)[7].

$$R_0 = \frac{\beta}{\alpha} \tag{4}$$

If R_0 is greater than one, then each infected person will infect more than one susceptible person, and the disease will spread. If, on the other hand, R_0 is less than one, the disease will die out, since less new infected will be generated each generation. With R_0 it is possible to calculate the fraction of the population that will be infected at equilibrium in the SIS model[8]:

$$\frac{I_E}{N} = 1 - \frac{1}{R_0} \tag{5}$$

To consider two separate, yet interacting, populations, a "travel term" is added[8], which gives the following differential equations for the rate of change of the n:th population:

$$\partial_t S_n = -\beta \frac{S_n I_n}{N_n} + \alpha I_n + \sum_{m \neq n} [\omega_{nm} S_m - \omega_{mn} S_n] \tag{6}$$

$$\partial_t I_n = \beta \frac{S_n I_n}{N_n} - \alpha I_n + \sum_{m \neq n} [\omega_{nm} I_m - \omega_{mn} I_n] \tag{7}$$

where ω_{mn} determines the strength of the connection between the populations m and n, so for example $\omega_{mn} S_n$ is how many in the susceptible part of population n that travels to population m per unit time. The sum is over all populations, and by imposing detailed balance[8] (which means that $\omega_{nm} N_m = \omega_{mn} N_n$ for all populations), equation (3) will still be valid, and the total populations will remain constant.

2.1.2 SIR

In slightly more advanced models, a third population category is introduced. Usually this category is denoted R (recovered) and is the number of people in the population who have gained immunity against the disease. The total population is now $N = S + I + R$. The flow will usually be from S to I to R, but to consider vaccination of the population, a rate from S directly to R can be introduced. R is usually a dead end, that is, there is no flow from R, but it could be interesting to consider a flow from R if one would consider a disease which immunity wears off after some time (e.g. hepatitis[9]) or it might be a simple way to model a disease with high evolution rate (e.g. influenza[10]). An example of a set of differential equations describing this model is the following:

$$\partial_t S = -\beta \frac{SI}{N} - VS + \varrho R \tag{8}$$

$$\partial_t I = \beta \frac{SI}{N} - \alpha I \tag{9}$$

$$\partial_t R = VS + \alpha I - \varrho R \tag{10}$$

In this case the number of vaccinated per unit time is the constant V times the number of susceptible, and ϱ is the rate at which R turns back to S. To obtain the basic SIR model, V and ϱ are set to zero. Travelling between cities can be added to equations (8)-(10) in the same manner as in the SIS model. By adding up equations (8)-(10) we see that the total population is constant, that is,

$$\partial_t N = \partial_t S + \partial_t I + \partial_t R = 0 \tag{11}$$

also for this model.

2.1.3 Colizza's SARS model

When choosing a model to use, it is good to know what disease it is suppose to describe. Knowledge of the disease will make it possible to make good choices of population categories, and transitions between them. I chose to look at Severe Acute Respiratory Syndrome (SARS), since there is good data on the epidemic[5], and because similar studies have been made before[1, 2].

SARS is caused by corona-virus SARS-CoV[3]. The entire genome of SARS-CoV has been sequenced, showing that it was not closely related to any other known corona-virus at the time of the outbreak[3]. Recently however, similar viruses have been found in animals such as civet cats, thus it is possible that SARS-CoV was an animal virus that adapted to humans[3].

SARS has two phases[3]. The first (acute) phase lasts for about 10 days, and is characterised by diffuse alveolar damage (DAD) which means that the alveoli in the lungs (the end of the respiratory tree, where the actual gas exchange is performed) are flooded with transudate (a water solution, rich in salts but low in protein), which triggers hyaline (≈cartilage) membrane formation[3]. In the late part of this phase antibody start to kill infected cells[3]. The second phase is characterised by increasing fibrosis (scar tissue) in the lungs and the state is no longer acute, but it can still be lethal[3]. All this can be seen in histological pictures, that is, sliced tissue viewed through a microscope[3]. It can be hard to diagnose SARS, since it not only infects the respiratory systems, but can also cause diarrhoea and infect the kidneys among other things[3].

To make a realistic model for SARS seven population categories have to be introduced[2], as was done by Colizza *et al.* The categories are susceptible (S), latent (L), infectious (I), hospitalised and recovering (H_R), hospitalised and dying (H_D), recovered (R) and dead (D) and the flow between the categories is graphically interpreted in figure 1. The latent (L) are the number of people in the population that have been exposed to the disease, but have yet to show symptoms. They act much as the susceptible; they are able to travel, and they cannot infect others. The hospitalised categories (H_R and H_D) are the number of people in the population that have been taken care of, and are being partially isolated from the rest of the population. The reason for two different hospitalised categories is that the recovering spend, on average, less time in hospital than the dying (that is $\mu_R^{-1} < \mu_D^{-1}$, see below). The following differential equations describes the model:

$$\partial_t S = -\beta \frac{SI}{N} - r_\beta \beta \frac{SH_R}{N} - r_\beta \beta \frac{SH_D}{N} \tag{12}$$

$$\partial_t L = \beta \frac{SI}{N} + r_\beta \beta \frac{SH_R}{N} + r_\beta \beta \frac{SH_D}{N} - \varepsilon L \tag{13}$$

$$\partial_t I = \varepsilon L - d\mu I - (1-d)\mu I \tag{14}$$

$$\partial_t H_R = (1-d)\mu I - \mu_R H_R \tag{15}$$

$$\partial_t H_D = d\mu I - \mu_D H_D \tag{16}$$

$$\partial_t R = \mu_R H_R \tag{17}$$

$$\partial_t D = \mu_D H_D \tag{18}$$

The total population is $N = S+L+I+H_R+H_D+R+D$ which, in combination with equations (12)-(18) gives constant population:

$$\partial_t N = \partial_t S + \partial_t L + \partial_t I + \partial_t H_R + \partial_t H_D + \partial_t R + \partial_t D = 0 \tag{19}$$

Equation (12) describes how the number of susceptible changes with time. Since those who recover are immune, there is no positive term (unless susceptible people are allowed to travel in from different populations), and thus will the number of members in this category steadily decrease. Each of the three terms represents the effect from the respective infectious category, I, H_R and H_D, and r_β is a factor that reduces the infectivity of the hospitalised part of the population. Equation (13) describes how the number of latent changes with time. The three positive terms are the people that disappeared from equation (12), and the negative term is the number of people that develop symptoms, thus turn into I. The constant ε is the inverse of the average time that an individual spend in this category. Equation (14) describes how the number of infectious changes with time. The first term represents those L who developed symptoms. When someone is identified as infectious (this take on average μ^{-1} days), this person will be taken to hospital, which is represented by the two negative terms. Since d is the case fatality rate, the fraction $d\mu I$ of the infectious will be classified as hospitalised and dying, the rest of the hospitalised ($(1-d)\mu I$) will be recovering. This can be seen in equations (15) and (16), which describes how the number of hospitalised and recovering respectively hospitalised and dying is changing with time. The negative terms are those who recovers (recovery takes on average μ_R^{-1} days) respectively dies (the average time from hospitalisation till death is μ_D^{-1} days). Equations (17) and (18) simply keeps track of those who recovers respectively dies, and they will steadily increase.

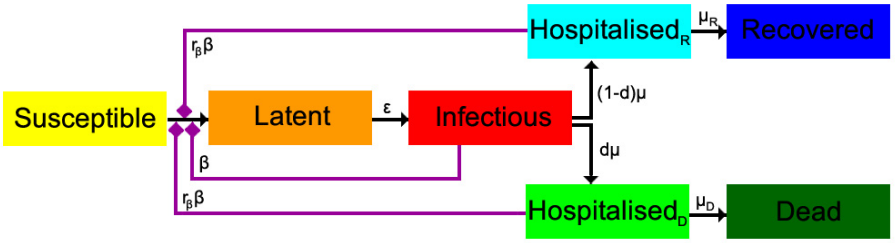


Figure 1: A graphic interpretation of Colizza's SARS model (equations (12)-(18)) in one city. The symbol by each black arrow represents the rate for the respective "reaction", for example the number of recovered people will each time step increase with the number of recovering people at hospital times μ_R. The purple arrows with square arrow heads indicates that the fraction of the population in the category from which the arrow originates will effect the rate at which the arrow points. For example the infected people will make $S \cdot \beta I/N$ susceptible people latent each time step.

Since the dimension of all the rate constants are inverse time, it is possible to get the rates from the existing data simply by inverting the average time that is spent in each category. The values of the parameters used by Colizza can be found in table 1.

Table 1: Values of the parameters used by Colizza *et al.*[2] in their model of transmission of SARS. The rate of transmission (β) and the number of initial latent individuals ($L(t=0)$), were found by fitting the model output to data. The rest of the parameters were found by analysing the data provided by WHO[5]. The step function $s_f(t)$ is a scaling factor that changes β with time, so that the actual value is $\beta s_f(t)$ at a given time t. This should be seen as the effect of control measures, and will be discussed in section 4.2.

Parameter	**Description**	**Value**
β	Rate of transmission (days^{-1})	0.57
$L(t=0)$	Number of initial latent individuals	10
$s_f(t)$	Scaling factor for β: t_1 = 21 February-20 March t_2 = 21 March-9 April t_3 = 10 April-11 July	 1.00 0.37 0.06
r_β	Relative infectiousness of hospitalised patients	0.2
ε^{-1}	Average latency period (days)	4.6
$\mu^{-1}(t)$	Average time before admission (days): t_1 = 21 February-25 March t_2 = 26 March-1 April t_3 = 2 April-11 July	 4.84 3.83 3.67
μ_R^{-1}	Average period from admission to recovery (days)	23.5
μ_D^{-1}	Average period from admission to death (days)	35.9
d	Case fatality rate	0.2

Hospitalisation is important in this model since the people in that category are much less contagious than those in the infected category (adjusted in the model by the constant r_β), and the interpretation of that should be that those

in hospital are being isolated and treated. The latent category is important since those in it will be treated as if they are susceptible and unexposed, which makes it unlikely to succeed with a quarantine. In this model, the recovered are immune and will never be infected or infectious again.

To this model Colizza *et al.* added the possibility to travel, in a similar way as in section 2.1.1 and section 2.1.2. Since Colizza's model is stochastic, the probability p_{mn} to travel is used rather than the rate ω_{mn} that I used. When the possibility of travelling is added, it is reasonable to only let some of the population categories be able to travel. How this is done in Colizza's model is not clearly described in their article[2], so in my model (described further in section 2.1.4) I chose to let the the susceptible, latent and recovered people travel. It is obvious why the dead and hospitalised don't travel, and the reason why the infected don't travel can be both because the individual are too tired or feel too miserable to travel, or because authorities don't let them travel, as a way to restrict the spreading.

2.1.4 My SARS model

The model I used for modelling SARS is very similar to Colizza's model. There are three major differences though. The first is that my model is fully deterministic, whereas Colizza's model is stochastic. The other two differences are both in the travelling part. Firstly I only considered the air traffic, and not trains, boats or any other way of transportation, and secondly I introduced a repulsion factor, as a way to satisfy the detailed balance premiss[8], thereby making sure that the total number of inhabitants of each city is constant:

$$\frac{\partial N_n}{\partial t} = 0 \tag{20}$$

Adding travelling to Colizza's model is done in the same way it was done in section 2.1.1 and section 2.1.2. This time however, a "repulsion factor" Γ_n, defined in equation (21), is multiplied to each of the travel-terms.

$$\Gamma_n = \left(\frac{I_n + H_{Rn} + H_{Dn} + D_n}{S_n + L_n + R_n} + 1 \right) \tag{21}$$

Only the people classified as S, L or R will be able to travel, so the differential equations for this model will thus be:

$$\partial_t S_n = -\beta \frac{S_n I_n}{N_n} - r_\beta \beta \frac{S_n H_{Rn}}{N_n} - r_\beta \beta \frac{S_n H_{Dn}}{N_n} + \sum_{m \neq n} [\omega_{nm} S_m \Gamma_m - \omega_{mn} S_n \Gamma_n] \tag{22}$$

$$\partial_t L_n = \beta \frac{S_n I_n}{N_n} + r_\beta \beta \frac{S_n H_{Rn}}{N_n} + r_\beta \beta \frac{S_n H_{Dn}}{N_n} - \varepsilon L_n + \sum_{m \neq n} [\omega_{nm} L_m \Gamma_m - \omega_{mn} L_n \Gamma_n] \tag{23}$$

$$\partial_t I_n = \varepsilon L_n - d\mu I_n - (1-d)\mu I_n \tag{24}$$

$$\partial_t H_{Rn} = (1-d)\mu I_n - \mu_R H_{Rn} \tag{25}$$

$$\partial_t H_{Dn} = d\mu I_n - \mu_D H_{Dn} \tag{26}$$

$$\partial_t R_n = \mu_R H_{Rn} + \sum_{m \neq n} \left[\omega_{nm} R_m \Gamma_m - \omega_{mn} R_n \Gamma_n \right] \tag{27}$$

$$\partial_t D_n = \mu_D H_{Dn} \tag{28}$$

The repulsion factor Γ_n is intended to be interpreted as an unease from travelling to a contaminated area. If a large part of the population is infected, hospitalised or dead, it will probably scare off travellers from going there. The repulsion factor is thoroughly described and discussed in section 2.1.6.

2.1.5 Fractions of individuals

Since I use rate equations to model disease transmission, it is rather cumbersome to look at populations as integers. Even though it is quite counter-intuitive I allow parts of people to travel and change category. This shouldn't be taken as body parts roll off on their own, but rather as the mean number of people averaged over many stochastic simulations.

One actual problem is that all the cities are connected in some way, so after very few time steps, every city will be infected, even if the number of infected people in the connected cities are a fraction of an individual. For most of the cities the number of infected will seem insignificantly small, but since they are present, the transmission process can start within that population. This phenomenon will lead to a faster epidemic, which makes it more dangerous, since there will be less time to prepare counter measures. A simple but effective way to fix this is to introduce a Heaviside step function[1] for the latent category:

$$\Theta(L - L_c) = \begin{cases} 1 & \text{if } L > L_c \\ 0 & \text{if } L < L_c \end{cases} \tag{29}$$

The Heaviside step function will deny the latent people in a population to interact, unless they are more than a certain critical number, L_c, which can be seen as that there are no latent people until they are sufficiently many. I used $L_c = 1$ throughout this study. The reason for choosing the latent people to have a theta-function is that they are a bottleneck category, as seen in figure 1. Since the infectious don't travel, the only way to get infectious persons in a population is when latent changes category. Without any infectious people, the only way to get latent people in the city is for them to travel in from other cities. Thus will the disease not break out unless enough latent people travel to the city and start to turn infectious.

2.1.6 The repulsion factor

To make the discussion a little easier to follow I will introduce two acronyms: SLR and HDI. SLR stands for "Susceptible, Latent and Recovered" and represents the part of the population that is allowed and capable of travelling. HDI stands for "Hospitalised, Dead and Infectious" and represents the part of the population that isn't allowed or is incapable to travel.

The repulsion factor Γ_n (see equation (21)) was introduced to keep the detailed balance condition fulfilled, and thereby making sure that N_n (the total

[1] I am grateful to Dirk Brockmann for pointing this out.

number of people in city n) is kept constant. The reason that it is needed is that only people in three of the seven categories are allowed to travel. Because of that, in the unmodified model the number of people travelling from the city will decrease with an increasing number of HDI. The cities with healthy populations will not be affected by this, so the number of people travelling from the infected city will be less than the number of people travelling to the infected city, which will increase its population. This phenomenon isn't very intuitive, it is more likely that people will avoid going to a contaminated area. The problem could be solved by recalculating the coupling strengths every time step, using only the SLR part rather than the entire population. One problem with this is that it will take long time to do, and it has to be done many times.

Trying to interpret the repulsion factor, it can be considered as information available to the inhabitants of the world. When a big part of the population in a city is HDI it will be big, noticeable news compared to if only a small part of the population is HDI. That makes a greater infection more known, thus it effects the travel pattern more. It can also be seen as a journey to a city with a bigger fraction HDI will be associated with a greater risk than a journey to a city with a smaller fraction HDI. Both these ways of looking at it will result in people avoiding infected cities.

There might be problems with devastating round off errors when SLR goes to zero (division with zero in the first term of equation (21)) if a numerical approach is used to solve the problem. Analytically this isn't a problem since the repulsion factor is multiplied with ether S_n, L_n or R_n, which will be positive, but also go to zero. Since $S_n \leq (S_n + R_n + L_n)$, $L \leq (S_n + L_n + R_n)$ and $R_n \leq (S_n + L_n + R_n)$ the ratios of S_n and $(S_n + L_n + R_n)$, R_n and $(S_n + L_n + R_n)$ and L_n and $(S_n + L_n + R_n)$ respectively can never be larger than one, that is, the numerator goes to zero faster than the denominator, which makes it impossible to go to a city with no SLR.

2.1.7 Network connection strength

The connections between the cities used in my SARS model are the same as those used by Guimerà *et al.* in their analysis of the worldwide air transportation network[11]. Their data contains the 531,574 passenger flights operating in the time period November 1, 2000, to November 7, 2000, between 3,883 cities distributed according to table 2[11]. I didn't have the data for all the population sizes for all those cities, so my model only used 3,147 cities (see section 2.1.8).

Table 2: The number of cities, and their spatial distribution, used by Guimerà *et al.*[11] in their analysis of the worldwide air transportation network.

Region	**No. of cities**
Africa	364
Asia and Middle East	719
Europe	691
Latin America	523
North America	1,064
Oceania	522
Total	3,883

From that data 27,051 city pairs could be found[11], however those connections are either 1 or 0, that is, the information they hold is only whether there is a connection or not, independently of how strong the connection is, i.e. how many people is travelling between the the city in a certain time. To calculate the strength of the connections I made three assumptions:

- The total population (including the dead) in each city is constant over time
- The number of people travelling to large cities is greater than the number of people travelling to small cities
- The number of people travelling from large cities is greater than the number of people travelling from small cities

The easiest way to satisfy these assumptions is by assuming that the weights to use in my model are those in equation (30), which has the consequence that the relation in equation (31) is satisfied. In these equations N is the total population, and the indices are the cities that are considered. The weights ω_{mn} are actually the fraction of the population in city n that is travelling to city m each day. The number of people travelling from city n to city m each day is thus the product $\omega_{mn}N_n$, so equation (31) is the required detailed balance condition[8].

$$\omega_{mn} = CN_m \tag{30}$$

$$\omega_{mn}N_n = \omega_{nm}N_m \tag{31}$$

The constant C (with unit per day) in equation (30) should be the same for every pair of cities, since the weights were assumed to only be depending of the sizes of the cities. Statistics on the yearly number of people passing through an airport is often available on the respective airport's web page, so by choosing airports with only one connected airport, it is possible to find the strength of that connection. To find a good C I chose ten pairs of cities from ten different countries, and calculated the their respective C from the weights using equation (30) (see table (3)). The different values of C is probably biased, since all the cities chosen to be m are big and well connected, whereas all the cities chosen to be n are small and have only one connection (the one to m). The average value of C is $\bar{C} = 3.99 \cdot 10^{-8}$ per day.

Table 3: List of the city pairs and the weights that were used to calculate C. The distances d between the cities are also listed. The weights are the fraction of the population in city n that is travelling to city m each day.

City$_m$	City$_n$	Country	d (km)	ω_{mn}	C_{mn}
Ankara	Batman	Turkey	763	$1.31 \cdot 10^{-3}$	$3.84 \cdot 10^{-10}$
Ottawa	Kitchener	Canada	428	$7.52 \cdot 10^{-4}$	$3.20 \cdot 10^{-9}$
Jamestown	Devil's Lake	USA	133	$1.27 \cdot 10^{-3}$	$8.18 \cdot 10^{-8}$
Tallinn	Kärdla	Estonia	109	$7.66 \cdot 10^{-3}$	$1.91 \cdot 10^{-8}$
Rio de Janeiro	Campos	Brazil	250	$7.19 \cdot 10^{-5}$	$6.50 \cdot 10^{-12}$
Naha	Kumejima	Japan	86	$2.17 \cdot 10^{-2}$	$1.81 \cdot 10^{-7}$
Kota Kinabalu	Lahad Datu	Malaysia	270	$8.21 \cdot 10^{-3}$	$5.66 \cdot 10^{-8}$
Patna	Ranci	India	255	$6.47 \cdot 10^{-4}$	$4.70 \cdot 10^{-10}$
Cali	Guapi	Colombia	200	$7.03 \cdot 10^{-3}$	$4.14 \cdot 10^{-9}$
Perth	Newman	Australia	1020	$7.29 \cdot 10^{-2}$	$5.22 \cdot 10^{-8}$

2.1.8 Population division

The number of inhabitants in 2,445 of the 3,147 cities that are connected in my model, were extracted from a database[12]. The rest of the population sizes were found on various places on the Internet. The most important tool to find them was *wikipedia.org*, which links to relevant local census reports, but *wolframalpha.com* was also used for some cities. Information on which region and country the cities belong to as well as the coordinates, IATA-name and ICAO-name of the associated airport were stored. The regional division is according to the 18 WHO influenza transmission zones[13]. In total 998,328,995 people are accounted for in my model, which is about 14% of the entire world population[14].

One problem for theoretical epidemiology is to correctly identify separate communities[8]. When counting the population of a city there might, for example, be problems deciding how much of the suburb that should be taken into account. Sometimes it might be more convenient to place an airport in a smaller city close to a big city, than to place it in the big city. Should the sum of the two city's populations be used in this case? Since the only way to travel between cities in my model is by air, it is probably better to use overestimations, where a better model would have used more cities which would have been connected with trains, cars and maybe more means of transportation.

2.1.9 Time dependent parameters

Two of the parameters in the model (namely $s_f(t)$ (scaling factor for β) and $\mu^{-1}(t)$ (average time before admission in days)) changes with time. In table 1 the values for them can be found, and these are based on data from Hong Kong. In my model I assumed that these step functions, with the same intervals, are the same everywhere, but with different starting dates. I introduced one of each function for each WHO region[13], and chose the starting date, as the first day that the total number of people categorised as infectious in the respective region where more than one.

2.2 Program

Initially I used the Euler method to solve the rate equations presented in section 2.1, since it is fast and easy to implement. After some test, I felt that I could afford to use a more complex and accurate method, so I changed to the 4:th order Runge-Kutta method[15]. Whenever I made a major change to my algorithm, I started with the Euler method, so that I would have something to compare with.

2.2.1 Two city model

Before simulating a real disease spreading in the entire world, I tried an SIS model in two cities. This way a lot of non-intuitive phenomenon can be cleared up, so that the number of numerical problems that have to be solved in the full model can be minimised. It is also easy to solve the long time-limit for the rate equations analytically, making it possible to tell whether a numerical solution is good or not. In this model equation (6) and equation (7) are the appropriate ones to use, however they can be written as one equation by introducing $j_n = I_n/N_n$, that is, the fraction of the population that is infected. In equation (32) I used that $(I_n + S_n)/N_n = 1$, see appendix for full derivation (there is a typo in equation (32) in the review article of Brockmann *et al.*[8]).

$$\partial_t j_n = \beta j_n(1 - j_n) - \alpha j_n + \sum_{m \neq n} \omega_{mn} \left(j_m - j_n\right) \tag{32}$$

The cities I chose in this model was the two Turkish cities Ankara (population[12]: 3,764,000) and Batman (population[12]: 293,024). Among the reasons that I chose them of all cities is:

- The airport of Batman has only one connection
- The airport of Ankara has many connections
- Both cities are quite large
- The ratio of the populations is about 0.1

Since Batman only have one connection it is easy to find the number of people using that connection, thus the connection strength. If it is interesting to add a few more cities to the model, it can be done since Ankara is fairly well connected. The population difference is good, because it might lead to a difference in j_n, and if the cities would have had the same population sizes, it would be impossible to see that phenomenon. Since the difference isn't extreme, and the cities are large, the risk for numerical errors should be low.

After simulating a scenario where one person in Batman got infected with a disease with the properties $\beta = 0.04$ per day and $\alpha = 0.01$ per day, it turned out that the same result was reached (to at least five decimal places) when equation (32) was solved, as when equations (6) and (7) were solved (I used 0.01 days as step size, and fourth order Runge-Kutta algorithm in both cases). The results can be seen in table 4, and the analytic derivation can be found in the appendix. Figure 2 shows how the number of people in the different categories in the

different cities changes with time. Just as expected, the number of susceptible in each city will decrease, and the number of infected will increase. After about 700 days equilibrium is reached in both cities, and unless the parameters of the disease is changed, the system will stay in this equilibrium. From figure 2 (and by looking at the actual numbers at the end of the simulation) we see that the total population in each city is constant, which is as it should be.

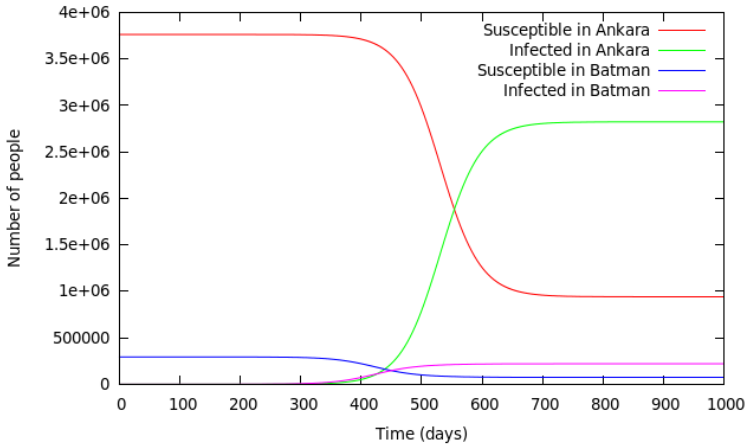


Figure 2: The development of the health situation in the two connected cities Ankara and Batman. At time zero one person (in Batman) were infected with a disease characterised by $\beta = 0.04$ per day and $\alpha = 0.01$ per day. The fourth order Runge-Kutta algorithm was used on equation (6) and equation (7) with the step size 0.01 days. After about 700 days, equilibrium is reached in both cities.

Since it is an SIS model, equation (5) can be used to calculate the fraction of the population that will be infected at equilibrium. To get R_0 equation (4) is used:

$$R_0 = \frac{\beta}{\alpha} = \frac{0.04}{0.01} = 4 \tag{33}$$

Inserting $R_0 = 4$ in equation (5) gives $j = 0.75$, so it looks as if Ankara and Batman reaches the same equilibrium as they would have if their connection would have been broken just after the disease had spread to both the cities.

Table 4: The different values for the fraction of the population that were infected at equilibrium in a SIS model. To calculate them, the method in the first column was on the equations in the second column. The values of the parameters were $\beta = 0.04$ per day, $\alpha = 0.01$ per day. For the numerical results $\Delta t = 0.01$ days and fourth order Runge-Kutta was used. Initially all the 3,764,000 inhabitants of Ankara were susceptible, and one of the 293,024 inhabitants of Batman were infected, the rest were susceptible.

Approach	**Equation**	j_{Ankara}	j_{Batman}
Numerical	(32)	0.75000	0.75000
Numerical	(6) and (7)	0.75000	0.75000
Analytic	(32)	0.75000	0.75000

2.2.2 Graphic output

To make it easy to get an overview of the problem, I wrote a Java program with graphic output. Figure 3 is a screen shot of the program after simulating SARS, using Colizza's parameters, step size 0.01 days, and no Heaviside step function or scaling factor for β. Each dot in the big map (equirectangular projection) corresponds to a city, and the colour of the dots changes with the health situation in the respective city (see table 5). By clicking on the dots, it is possible to get detailed information on the specific city, or infect it with the disease. The small map (also equirectangular projection) is divided according to the 18 WHO influenza transmission zones[13]. Just like the dots, the zones changes colour with the regional health situation, but it uses a different scale (see table 5).

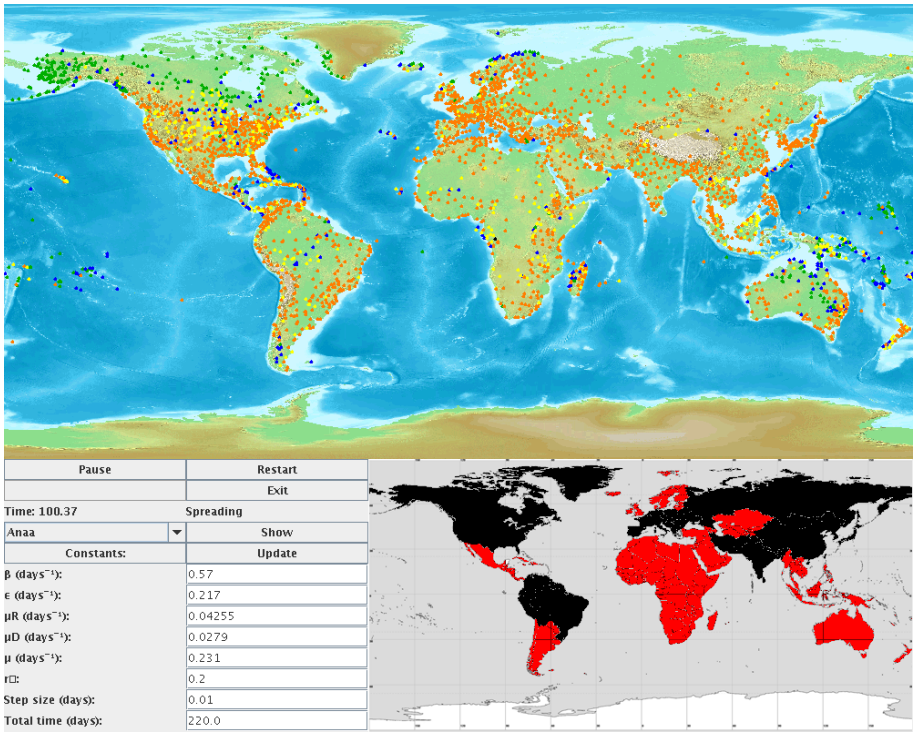


Figure 3: A screen shot of my graphic program. Each dot in the big map (totally 3147 dots) represents a city, and the colour represents the health situation in that city (see table 5). The small map has been divided into the 18 WHO influenza transmission zones[13], and the colour of a region depends on the health situation in all the cities in that region (see table 5). The value of the parameters that are used are those in the lower left corner, and are those used by Colizza *et al.* (see table 1). The screen shot was taken 100 days into the simulation which took, on a desktop computer, about fifteen hours to simulate.

Table 5: Legend for the graphic program. The left part is for the large map, and the right part is for the small map. The lower part of the legend for the large map is prioritised over the upper part, that is, if someone dies in a city with 100 inhabitants (the rest is susceptible), the dot representing that city will immediately go from green to black (and not yellow).

Large map		**Small map**	
Circumstance	**Colour**	**# regional deaths**	**Colour**
< 1 infectious and < 1 dead	Green	$0 - 1$	Green
≥ 1 infectious and < 1 dead	Blue	$1 - 10$	Turquoise
≥ 1 dead	Yellow	$10 - 100$	Blue
$> 0.1\%$ of population dead	Orange	$100 - 1,000$	Yellow
$> 0.5\%$ of population dead	Red	$1,000 - 10,000$	Orange
$> 1.0\%$ of population dead	Black	$10,000 - 100,000$	Red
		$> 100,000$	Black

2.2.3 Maps and coordinates

Both maps that I use in my program with graphic output I found on the Internet. The URI for the large map is *http://upload.wikimedia.org/wikipedia/commons/c/cf/WorldMap-A_non-Frame.png*, and the URI for the small map is *http://maps-world.cn/map/world_map_gif.gif*. I modified the small map, by removing the text. Both maps uses equirectangular map projection, and the coordinates that I got for the airports were measured in arc-seconds from the Prime Meridian and the Equator, with positive direction east and north. Equirectangular map projection uses the geographical coordinates as Cartesian coordinates, and get in that way a rectangular map[16]. The poles will however be very stretched compared to the equator. In my program $(0, 0)$ is in the top left corner and the positive direction is down and right. The large map is 1000 pixels wide and 500 pixels high, so the transformations and rescaling I had to do are those in equations (34) and (35).

$$x = \left(\frac{\lambda}{3600} + 180\right) \cdot a \tag{34}$$

$$y = \left(90 - \frac{\phi}{3600}\right) \cdot a \tag{35}$$

In equations (34) and (35) (x, y) are the Cartesian coordinates, and (λ, ϕ) are the geographical coordinates (longitude, latitude). The division with 3600 is to get the longitudes and latitudes from arc-seconds to degrees. Adding 180 to this result respective subtracting the result from 90 is to get the coordinates in the interval $0 \leq x \leq 360$ and $0 \leq y \leq 180$. By choosing $a = 2.777$ (dimension length per angle), the new intervals will be $0 \leq x \leq 1000$ respectively $0 \leq y \leq 500$, and they will fit on the map.

2.2.4 Data generator

I also wrote a program optimised for generating data. It uses the same model as the graphic program, i.e. solves equations (22)-(28) using fourth order of the Runge-Kutta algorithm, but generates a file as output. The idea is that the data generating program should be used to find interesting scenarios, which then can

be simulated in the graphic program. The generated data is qualitatively better, and more suited for analysis, but it is much easier to get an overview with the graphic output.

3 Results

3.1 Simulations with Colizza's parameters

When I used the parameters from table 1, and step size 0.01 days, in my model, the number of people categorised as latent in Hong Kong (where the disease was initiated) would decrease dramatically, and before the end of the first day, less than one person was categorised as latent in Hong Kong. A small fraction of a person had become infectious, but the majority had travelled to other cities. Due to this, the latent people spread around the world, but since they were so few initially, the number of latent people in each city, would always be less than one. Since the Heaviside step function (see equation (29)) made it impossible for a latent person to turn infectious if the total number of latent people in that population was less than one, there wouldn't be anyone turning infectious, and the disease would get extinct before breaking out.

Using my model, without the Heaviside step function, to simulate SARS with the parameters from table 1, and with step size 0.01 days, will give very different results compared to when the Heaviside step function was present. In this case an epidemic breaks out, and the number of total deaths in Hong Kong is about 30 people. According to the measured data[5] 298 people died from SARS in Hong Kong, which is roughly ten times the result of the simulation.

Simulating SARS with the parameters from table 1, and with step size 0.01 days, but with neither the Heaviside step function, nor scaling factor for β (that is, the behaviour of people will not change with time), will give an immense epidemic, and 220 days after the ten latent people were placed in Hong Kong only about four million people is still susceptible in the world. That is less than half a percent of the people in my model, and can be compared to the data for the epidemic of 2003[5] after which more than 99.999% were susceptible. After that comparison it is obvious that these model parameters does not describe the real epidemic. Simulating this latter scenario in my program with graphic output will, for day 100 (after about 15 hours of simulating on a desktop computer), give the output that is shown in figure 3.

3.2 Adjusting to the measured data

According to Colizza *et al.*[2], their value of β and initial number of latent people were found by simulations and curve fitting. Since the other values of the parameters (see table 1) were extracted from measured data, it should be the β and initial number of latent people that I should change to try to get better agreement between my model's output and the measured data[5].

Choosing $\beta = 0.8$ per day and $L(t = 0) = 100$ will give the behaviour shown by the curve titled "Simulation 1" in figure 4 for Hong Kong. The shape

of the curve, representing the number of dead in the simulation, looks much like the shape of the measured data. The slope of the simulated data should be a little steeper, and that can be achieved by choosing a larger β, however when the slope starts to decrease, it does that to slowly, so the total number of death will be to much in the end. To get a better behaviour when the the slope starts to decrease, the disease should pass faster, that is, the time from infection till recovery or death should be shorter. This is a bit problematic, since it is extracted from measured data.

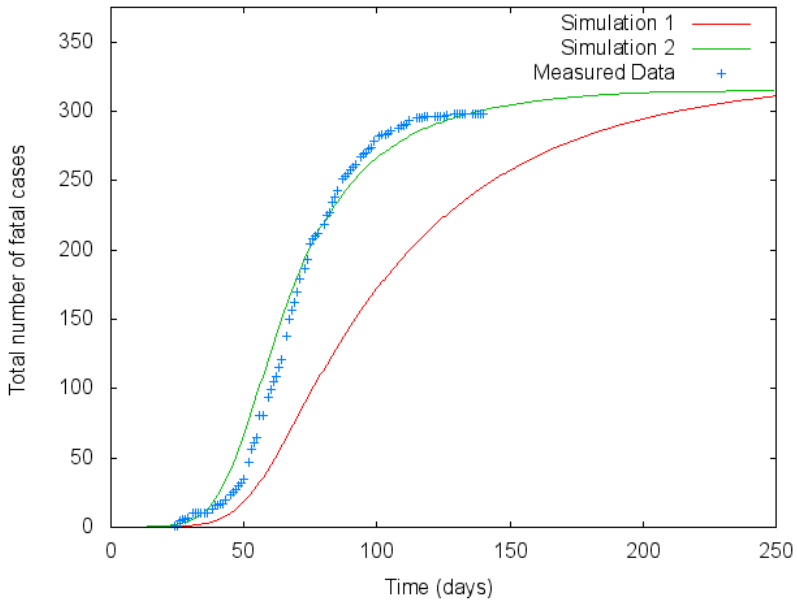


Figure 4: The points represents the measured data[5] for SARS in Hong Kong from March 17, 2003, to July 11, 2003. The curves represents the result from simulations with my program. In "Simulation 1" the parameters found in table 1 was used, except that I instead used $\beta = 0.8$ per day and $L(t = 0) = 100$. In "Simulation 2" the parameters found in table 1 was used, except that I instead used $\beta = 0.73$ per day, $L(t = 0) = 100$, $\varepsilon = 0.35$ per day, $\mu_D = 0.048$ per day, $s_f(t_2) = 0.25$ and $s_f(t_3) = 0.03$. The step size in the simulations was 0.01 days.

Even if it looks acceptable in Hong Kong (data: 299 fatal cases, simulation: 311 fatal cases), the number of deaths in other cities in the simulation is not that good. In Singapore and Taiwan there were 180 and 173 fatal cases respectively, and according to the measured data it should be 33 fatal cases in Singapore and 37 fatal cases in Taiwan[5]. Totally in the world 42,797 died in my simulation, to be compared to the 774 deaths according to the measured data[5].

To get a better curve for the fatal cases in Hong Kong, I try to change some of the parameters that were extracted from the measured data. By changing ε (the inverse average latency period) I can effect the slope. A larger ε will lead to a steeper slope, just like a larger β, however a larger ε will shorten the latency period, which will reduce the number of latent (in opposite to β), and thus will it effect its surroundings a little less. The curve titled "Simulation 2" in figure 4, uses Colizza's parameters (see table 1) with the following exceptions: $\beta = 0.73$

per day, $L(t = 0) = 100$, $\varepsilon = 0.35$ per day, $\mu_D = 0.048$ per day, $s_f(t_2) = 0.25$ and $s_f(t_3) = 0.03$. It is the scaling factor for β, and the inverse average period from admission to death that governs how fast equilibrium is reached.

With these values the effect on the world is still too much to be realistic (the number of fatal chases in Hong Kong is this time 315). In Singapore and Taiwan there were, with the new parameters, 181 and 157 fatal cases respectively (should be 33 in Singapore and 37 in Taiwan[5]). Since Taiwan lies closer to Hong Kong than Singapore does, and there is more fatal cases in Singapore than in Taiwan in my simulation, and the opposite in the measured data, there are reasons to suspect that there is some kind of distance dependence of the weights (the number of inhabitants in the two different areas are within the same order of magnitude), but since it is only one case, there might be many other reasons. The reason for the similar result in Singapore, but so different in Taiwan, might have something to with the fact that Hong Kong and Taiwan lies in the same region (east Asia), and that Singapore lies in another one (south-east Asia). The regions are important for the time dependent parameters, of which one ($s_f(t)$) was changed in the second simulation.

Totally 44,339 people died in the world in this simulation, to be compared to the 774 deaths according to the measured data[5]. The reason that the impact on the surrounding world is bigger in this simulation (where ε was larger) is probably because of the changes in $s_f(t)$ and μ_D.

3.3 Vaccination

In this section I used the adjusted parameters and simulated SARS. The difference from the last section ("Adjusting to the measured data") is that some people will be vaccinated before the epidemic, that is, they will be initiated as members of the R-category. I compare seven different cases; no one is vaccinated, everyone in London is vaccinated (7,172,091 people), the same amount of vaccine (7,172,091 doses) is distributed evenly across the world, the same amount of vaccine (7,172,091 doses) is used in the Chinese city Tianjin (7,499,181 inhabitants), the same amount of vaccine (7,172,091 doses) is used in the American city New York (8,008,278 inhabitants), everyone in Tianjin is vaccinated (7,499,181 people) and finally everyone in New York is vaccinated (8,008,278 people). The reason that I chose London is because it is (with is 239 connections) one of the most connected cities in the world. The reason for Tianjin is that it is almost the same size as London, but it is much less connected (it is only (directly) connected to 25 other cities). Another interesting feature for Tianjin is that it lies within the same region (east Asia) as Hong Kong, in which the disease is initiated. The reason for New York is again that it is almost the same size as London, but it is less connected (it is only (directly) connected to 177 other cities). Moreover, it is not in the same region as Hong Kong, so it might be interesting to compare the result from New York, with the result from Tianjin.

Table 6: The result of different strategies for vaccination. The numbers in the table are from equilibrium, 300 days after initiation (step size 0.01 days).

Doses	Distributed in	Recovered globally	Deaths globally
0	The world	174,836	44,365
7,172,091	The world	7,336,322	41,672
7,172,091	London	7,762,944	41,356
7,172,091	Tianjin	7,326,981	41,386
7,499,181	Tianjin	7,653,164	41,251
7,172,091	New York	7,576,533	41,455
8,008,278	New York	8,439,560	41,133

From table 6 we see that it is better to concentrate the operation to a concentrated area, rather than spreading the doses evenly around the world. The reason for that, I think, is because vaccination of a key location like London, will close a pathway for the disease, and thus delay the disease transmission. The case with least fatal cases were when entire New York was vaccinated. That is probably because 836,187 more doses of vaccine was used, but it might also be because, those in the city that were not vaccinated made a big enough path for the disease. The same phenomenon can be seen in Tianjin.

When equal amounts of vaccine are compared, the least number of fatal cases will be achieved if London is vaccinated, and the least effective way to vaccinate is to spread the vaccine evenly across the world. Even though New York is more connected than Tianjin it is more efficient to vaccinate Tianjin in this case. I think that is because Tianjin is in the same region as Hong Kong, but it might also be since there is less susceptible in Tianjin after vaccination compared to New York after vaccination.

3.4 Exploring the phase space

In my model, there are eight parameters that governs the disease dynamics in each population. Since only three of the seven categories are allowed to travel, the dynamics of the individual populations will effect the global disease transmission. Apart from that, it is the connections and their weights that governs the global disease transmission. Since I wasn't able to get the real weights, I tried to use the simplest possible model to get them (equations (30) and (31)), since it was hard to get better results with other models (see section 4.3).

The number of initial latent individuals ($L(t = 0)$)

With a larger number of initially latent individuals, the total number of recovered and dead will increase. This is because the disease will be more established when counter-measures are initiated (the time dependent step functions).

Rate of transmission (β)

A larger β will increase the slope seen in figure 4. This will lead to more dead and recovered in the end.

Scaling factor for β ($s_f(t)$)

The scaling factor for β is a way to capture how people react when there is an epidemic. Intuitively this will decrease with time, but there is of course the possibility that people react in a "wrong" way, and their new behaviour makes it easier for the disease to spread. If $s_f(t)$ is low enough, the number of latent people will start to decrease, otherwise will a low $s_f(t)$ just delay the disease, and the reason for a decrease in number of latent people is then because of few susceptible people to infect. The limit that separates the two behaviours is R_0.

Average time before admission ($\mu^{-1}(t)$)

By increasing $\mu(t)$ the average time before admission to hospital is shortened, that is, the time as infectious is shortened. That will ultimately lead to less fatal cases, since there were less time for the infectious to spread the disease.

Relative infectiousness of hospitalised patients (r_β)

When $r_\beta = 1$ there will be no difference in infectivity between before and after hospitalisation. It is intuitive to think that $r_\beta < 1$, but one plausible scenario when this isn't true is if there is a central hospital, that treats people from a large and sparsely populated area. A larger r_β will increase the slope seen in figure 4.

If $r_\beta = 0$, that is, hospitalised people will not be contagious at all, the only category that can infect susceptible people is the infectious people. In that case it will be meaningful to talk about R_0, since β is, on average, the number of people that an infected individual will infect per day, in an otherwise totally susceptible population. The α in this case is $\mu(t)$, since it is the average number of infectious people that is hospitalised every day. If infectious people are hospitalised faster than they infect the susceptible, the disease will die, otherwise there will be an epidemic. If $r_\beta \neq 0$ there will be some corrections, but the main point is still valid.

Average latency period (ε^{-1})

A larger ε will increase the slope seen in figure 4. This will lead to more dead and recovered in the end. The difference between the effects of a change in β and the effects of a change in ε is the contact with other cities. β governs the transition from susceptible to latent, and people from both of these categories are allowed to travel. ε governs the transition from latent to infectious, and only the people from the latent category are allowed to travel. This means that ε will change the number of SLR, and β will not, that is, if two simulations are made, each simulation with different β and ε, but in such way that the number of deaths in Hong Kong is the same in both cases, then the number of deaths will have changed globally.

If the sum $\varepsilon + \beta$ is kept constant, and SARS is simulated with different ε and β, the behaviour in figure 5 will be found. Each point is the result of a simulation with step size 0.01 days, $\varepsilon + \beta = 1.10$, $L(t = 0) = 100$, $\mu_D = 0.048$

per day, $s_f(t_2) = 0.25$ and $s_f(t_3) = 0.03$. The other parameters are those from table 1, and the measurement is done 300 days after initiation.

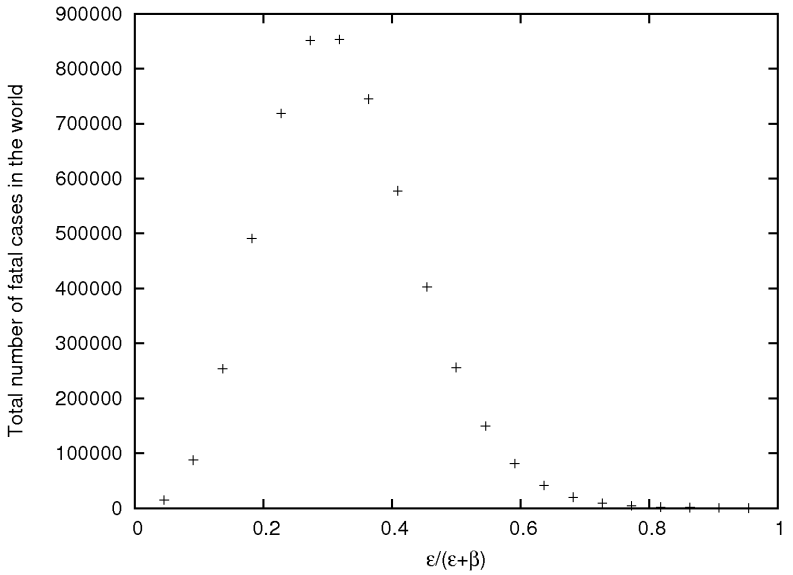


Figure 5: Each data point represents the number of fatal cases in the world after 300 days, for different values of ε. Each simulation is done with step size 0.01 days, and $\varepsilon + \beta = 1.10$ per day in all cases. The other parameters are those from table 1, except that $L(t = 0) = 100$, $\mu_D = 0.048$ per day, $s_f(t_2) = 0.25$ and $s_f(t_3) = 0.03$.

A small ε means a very long latency period, and with the the assumption that $\varepsilon + \beta = 1.10$, the β will be large, and thus will the disease be very contagious. If the latency period is too long, the epidemic will start very slow, and there will be long time to prepare (the clock for $\mu^{-1}(t)$ and $s_f(t)$ will start long before the epidemic gets serious). A large ε means a short latency period, and small β (low infectivity). If the latency period is too short the disease will have problem spreading to other cities, since it does that with the latent people. A low β will make the epidemic slow, since the increase in the number of infectious each generation is low. With low enough β, each infected person will infect less than one person, and thus will the disease die. The number of fatal cases is peaking at $\varepsilon \approx 0.42 \cdot \beta$.

Average period from admission to death (μ_D^{-1})

A small μ_D will mean a long average period from admission to death. The effect (both locally and globally) is that the number of effected will increase, since a long time in hospital before death will mean a long time exposing others. The time till equilibrium will be longer with a smaller μ_D.

Average period from admission to recovery (μ_R^{-1})

The response from the system to a change in μ_R is the same as to a change in μ_D, that is, low μ_R will lead to larger but slower epidemic.

Case fatality rate (d)

With higher case fatality rate the number of deaths will in general increase both locally and globally. The total number of affected might however be less. That is if there is big enough difference between those categorised as "hospitalised and dying" and those categorised as "hospitalised and recovering". There might also be a difference since a different d will give a different ratio between SLR and HDI.

4 Discussion

4.1 Stochastic models versus deterministic models

In this thesis I have only used purely deterministic equations. With enough people this should generate the same result as the average result from the stochastic model, that is, when there is enough people in the model, the behaviour of individuals should not effect the big picture. The problem with deterministic equations in my model is that there sometimes isn't enough people. Some cities only have population sizes of the order of magnitude of ten, so in these cities the fate of each individual will greatly effect the city mean.

More importantly though is in the initial phase of the epidemic. If, for example, one person is infected with a disease, then this person will, in the deterministic model, start to spread it in the area, and it may ultimately lead to a global epidemic. In the stochastic case this person might die before meting someone else, thus is there no opportunity for the disease to be transmitted, and the epidemic will not happen. In this example the behaviour of one person had great significance in the end, so a stochastic model is better in that sense. On the other hand, with a deterministic model, the same results will be reached every time, whereas the stochastic model will reach different results with different seeds for the random number generator. The stochastic simulation has thus to be done many times to get a statistically reliable result. With this approach the deterministic model is better, since it is much faster computationally.

4.2 Psychology

The next step for modelling transmission of diseases is to look at the change of behaviour of people. At present it is (in models) assumed that the travel habits of people doesn't change when an epidemic is on the rise. Intuitively one would think that people would avoid infected cities (which may lead to faster extinction of the disease, since there will be a lack of hosts), and even try to leave their city if it is infected (which may lead to faster spreading, since latent people will unknowingly carry the disease to uninfected cities). My repulsion factor takes this into account to a certain degree, but it was primarily introduced as a way to satisfy the detailed balance condition.

A way to handle the change of behaviour is to change the parameter β with time. It may be interpreted as people avoid contact with other people, they start to wear masks, wash their hands more often and are more careful in general. It may seem a little blunt, but even though this model can be made better by

investigating the behaviour more thoroughly, it is important to remember that the more complexity that is added to the model, the less general it will become, and it will be harder to get an overview and to make predictions. Without noticing it, properties that belong to a specific disease, or abiotic effects such as change of weather, might be added to the model of general behaviour of people. Influenza, for example, shows very strong seasonal variations in the temperate areas of the earth[17]. These effects are important as well, but to be able to understand them, they should be treated separately.

Another way to model the psychological effects is by using an agent-based modelling (ABM) approach[18]. In this model groups of people (each group may contain only one person), called agents, will be given different properties. These properties can for example be what condition the agents are in (S, L, I, etc.), how inclined they are to travel, how social they are (how much people they spend time with), in what condition their immune system is in, etc. This way the transmission of diseases can be studied with very high resolution, and with a practically arbitrary number of parameters, the result from a simulation can be almost arbitrarily close to measured data. As mentioned before, a problem with to many parameters is that it is hard to explore the entire phase space, thus making it very hard to fully understand the dynamics of the system, or even to identify key features.

4.3 Alternative models for connection weights

The strength of the weights that I used to find an average value of C were very diverse (see table 3), which implies that using the average is an oversimplification. One reasonable assumption is that the weights are dependent, in some way, of the distance between the cities. A shorter trip should be more probable than a long trip. It has been suggested[19] that the connection strength should decrease exponentially with the distance, resulting in equation (36) in which the r is a constant.

$$\omega_{mn} = CN_m e^{-\frac{d_{mn}}{r}} \tag{36}$$

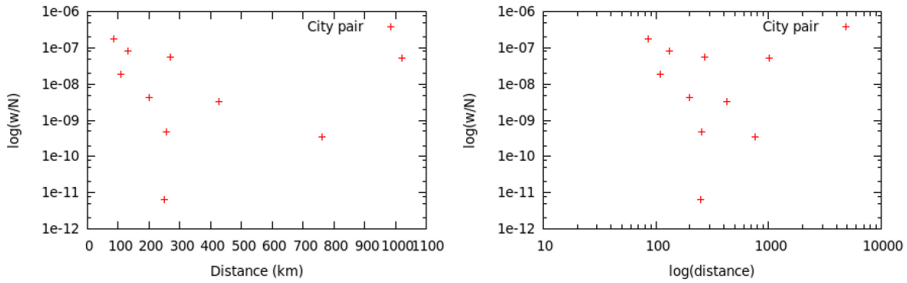
The group that suggested this relation found that it only is good if $d < 300$ km, but they included other kinds of transportation, as well as air traffic, in their analysis. By plotting $\log(\omega_{mn}/N_m)$ against d_{mn} using the data in table 3, it is possible to see if equation (36) might work. The result (see figure 6a) isn't clear though, the Australian value $(5.22 \cdot 10^{-8}; 1020)$ is far too strong relative to the distance, and the Brazilian value $(6.50 \cdot 10^{-12}; 250)$ looks a bit too weak relative to the distance. These two values might however be some kind of exceptions or extreme cases, so more points should be used to get a conclusive answer.

It has also been suggested that some power law, like the one in equation (37), should be used to calculate the couplings[8].

$$\omega_{mn} = CN_m d_{mn}^{-(2+\mu)} \tag{37}$$

By plotting $\log(\omega_{mn}/N_m)$ against $\log(d_{mn})$ it is possible to see if equation (37) might work. The points in this plot (see figure 6b) doesn't seems to lie on a line, which would speak against equation (37). Still the Australian point and the

Brazilian point stands out a bit, so it would be better with some more points, so that the theory can be dismissed with more confidence.



(a) Log-Normal **(b)** Log-Log

Figure 6: (**a**) The logarithmic weight per inhabitant plotted against the distance between the cities presented in table 3. If a linear approximation could be made, it would imply that the number of people travelling between two cities will decrease exponentially with the distance. (**b**) The same as in (**a**), but this time the distance is plotted on a logarithmic scale as well. If a linear approximation could be made in this case, it would imply that the number of people travelling between two cities will decrease with the distance raised to some power.

Among other things that probably affect the weights, but are harder to put a number on, are politics, geography and climate. With politics I mean that the relationship between countries might have an effect on how people travel between them. For example, two major cities may lie close to each other, but if they happen to lie in separate countries, with hostile attitude to each other, there will be little traffic going on between them. By geography I mean that some places might be hard to get to without using air plains. In some places the probability for attractive weather might be high, which is why those places might have strong connections despite the fact that they are small and remote. In this case the climate is an important factor.

Even if the, from data, measured couplings were used, there would still be some problems in a disease transition model. That is because some people only make stopovers in an airport, and within hours they are on a new flight. They will contribute to the weight, but they will never be a part of the disease dynamics of that city. To get a good result, it would be better to treat them as they went directly to their final destination. This is of cause very hard to do, and it will create some problems of its own.

If only the air traffic is considered the model won't be perfect, since there are more means of transport, which may be faster, or more effective in short distances.

5 Conclusions

The conclusions are that the air traffic couplings are not only dependent on the city sizes, but there is no easy connection between the strength of the couplings and the distance between the cities, or any other variable that I have found. To minimise an epidemic, the scaling factor for β is an important tool. This can

actually be altered in a case of a real epidemic. Vaccination is another way to reduce the impact of an epidemic, and it pays off to find important areas, and focus the resources on that.

If SARS would have been more lethal, that is, the case fatality rate would be higher, more people would have died, but if by "more lethal" one means the time from contamination till death, then it might actually be less, since there will be to little time fore the disease to be transmitted. For example will a very short latency period lead to few fatal cases in the world, since the disease will have problem to spread between cities. A very long latency period will, on the other hand, lead to a very slow transmission of the epidemic. In both of these cases there will be much time for the population to change behaviour, so that the disease dies. This can be seen in figure 5.

6 Outlook

There is more to be done when it comes to exploring the phase space, for example, it might be interesting to see more extensive studies on how the number of fatal cases in the world is dependent on the different parameters. It might also be interesting to see what happens if β would be dependent on the number of infectious rather than dependent on the time since the first infectious was detected. That is because people might be more keen to avoid contact with others, wash their hands thoroughly and more frequent etc., if they know that there are many infected people in the area, rather than if they know that there were a confirmed case a while back. After that, it might be possible to find better values for the parameters and the critical value for the Heaviside step function, to get a better agreement with the measured data. The natural next step after that is to introduce stochasticity in the model. Other things that needs to be taken into consideration is for example weather, seasons and school holidays.

References

[1] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. *PNAS*, 101(42):15124–15129, 2004.

[2] Vittoria Colizza, Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. Predictability and epidemic pathways in global outbreaks of infectious diseases: the sars case study. *BMC Medicine*, 5(1):34, 2007.

[3] John Nicholls, Xiao-Ping Dong, Gu Jiang, and Malik Peiris. Sars: clinical virology and pathogenesis. *Respirology*, (8):6–8, 2003.

[4] Severe acute respiratory syndrome (sars): Status of the outbreak and lessons for the immediate future, 2003. Available at http://www.who.int/csr/sars/resources/en/.

[5] Data on the SARS outbreak are available from WHO at http://www.who.int/csr/sars/country/en/.

[6] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58(1):35–55, 2007.

[7] Matt J. Keeling. Mathematical epidemiology. Available at http://www.warwick.ac.uk/∼masfz/Teaching_index.html.

[8] Dirk Brockmann, Vincent David, and Alejandro Morales Gallardo. Human mobility and spatial disease dynamics. Diffusion Fundamentals III, Leipziger Universitätsverlag, 2009.

[9] Harunor Rashid and Shafquat Mohammed Rafic. Hepatitis b vaccination in bangladesh: a suggestion based on current evidence. *Hepatitis Monthly*, 1(6):41–44, 2006.

[10] Alan J. Hay, Victoria Gregory, Alan R. Douglas, and Yi Pu Lin. The evolution of human influenza viruses. *Philosophical Transactions: Biological Science*, 356(1416):1861–1870, 2001.

[11] R. Guimerà, S. Mossa, A. Turtschi, and L.A.N. Amiral. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences of the United States of America*, 102(22):7794–7799, 2005.

[12] Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IPFRI), the World Bank; and Centro Internacional de Agricultura Tropical (CIAT), 2004. Global Rural-Urban Mapping Project (GRUMP): Settlement Points. Palisades, NY: CIESIN, Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw. (2011-02-24).

[13] Influenza transmission zones. Available at http://www.who.int/csr/disease/swineflu/transmission_zones/en/. (2011-04-26).

[14] U.S. Census Bureau. World popclock projection, May 2011. Available at http://www.census.gov/ipc/www/popclockworld.html.

[15] Nicholas J. Giordano and Hisao Nakanishi. *Computational Physics*. Pearson Prentice Hall, second edition, 2006. ISBN: 0-13-146990-8.

[16] Yoshiaki Ohsawa, Hirofumi Kageya, and Takeshi Koshizuka. Map projection errors in the weber problem. *Journal of Geographical Systems*, 4(3):325–342, 2002.

[17] Ian G. Barr, John McCauley, Nancy Cox, Rod Daniels, Othmar G. Engelhardt, Keiji Fukuda, and et al. Epidemiological, antigenic and genetic characteristics of seasonal influenza a(h1n1), a(h3n2) and b influenza viruses: Basis for the who recommendation on the composition of influenza vaccines for use in the 20092010 northern hemisphere season. *Vaccine*, 28(5):1156–1167, 2010.

[18] Eliot R. Smith and Frederica R. Conrey. Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, 11(1):87–104, 2007.

[19] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J. Ramasco, and Alessandro Vespignani. Multiscale mobility networks and the large scale spreading of infectious diseases. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51):21484–21489, 2009.

A Supplementary information

A.1 Analytic solution to equation (32)

For the case of only two connected cities, the stationary state (long-time limit) of equation (32) will, by setting the left-hand sides to zero, become:

$$\left.\begin{array}{l}\beta j_1(1-j_1)-\alpha j_1+\omega_{21}j_2-\omega_{21}j_1=0\\ \beta j_2(1-j_2)-\alpha j_2+\omega_{12}j_1-\omega_{12}j_2=0\end{array}\right\} \tag{38}$$

From the first equation of equation system (38) j_2 is isolated by dividing everything with ω_{21}, and then by moving every term, except the third, to the right hand side, which gives:

$$j_2 = j_1\left(\frac{\alpha+\omega_{21}-\beta}{\omega_{21}}+j_1\frac{\beta}{\omega_{21}}\right) = j_1\left(c+\frac{\beta}{\omega_{21}}j_1\right) \tag{39}$$

The second equation of equation system (38) can be written

$$0 = \beta j_2-\beta j_2^2-\alpha j_2+\omega_{12}j_1-\omega_{12}j_2 \tag{40}$$

and inserting the result from equation (39) gives

$$0 = \beta c j_1+\frac{\beta^2}{\omega_{21}}j_1^2-\beta\left(cj_1+\frac{\beta}{\omega_{21}}j_1^2\right)^2-\alpha c j_1-\frac{\beta\alpha}{\omega_{21}}j_1^2+\omega_{12}j_1-\omega_{12}cj_1-\frac{\omega_{12}\beta j_1^2}{\omega_{21}} \tag{41}$$

so that finally

$$j_1\left(\beta c-\alpha c+\omega_{12}-\omega_{12}c+j_1\left(\frac{\beta^2}{\omega_{21}}-\beta c^2-\frac{\beta\alpha}{\omega_{21}}-\frac{\omega_{12}\beta}{\omega_{21}}+\right.\right.$$
$$\left.\left.+j_1\left(-\frac{2c\beta^2}{\omega_{21}}+j_1\left(-\frac{\beta^3}{\omega_{21}^2}\right)\right)\right)\right)=0 \tag{42}$$

Inserting equation (42) with $\beta = 0.04$ per day, $\alpha = 0.01$ per day, $\omega_{12} = 0.001$ per day and $\omega_{21} = 7.784909670563230\cdot 10^{-5}$ per day in the Matlab R2009a function "solve" gives the following solutions:

$$\left\{\begin{array}{l}j_{1,1}=-0.001881347557858\\ j_{1,2}=0\\ j_{1,3}=0.747988892722544\\ j_{1,4}=0.750000000000032\end{array}\right. \tag{43}$$

Inserting the different j_1 in equation (39) will give the following j_2:

$$\left\{\begin{array}{l}j_{2,1}=0.724935120140279\\ j_{2,2}=0\\ j_{2,3}=-0.024935120152573\\ j_{2,4}=0.750000000012353\end{array}\right. \tag{44}$$

Since j is the fraction of the population that is infected, a necessary assumption is that $0 \leq j \leq 1$. The only pairs that meet that requirements is $(j_{1,2} = 0; j_{2,2} = 0)$ and $(j_{1,4} = 0.750000000000032; j_{2,4} = 0.750000000012353)$. The first pair corresponds to the repulsive node before the epidemic breaks out, and the second pair corresponds to the attractive node when equilibrium is reached.

A.2 Derivation of equation (32)

Since $j_n = I_n/N_n$, equation (32) can be derived by dividing equation (7) with N_n, which gives

$$\frac{\partial_t I_n}{N_n} = \beta \frac{S_n I_n}{N_n N_n} - \alpha \frac{I_n}{N_n} + \sum_{m \neq n} \left[\omega_{nm} \frac{I_m}{N_n} - \omega_{mn} \frac{I_n}{N_n} \right] \tag{45}$$

Rewriting it using $S_n = (N_n - I_n)$ and $1 = N_m/N_m$ gives

$$\frac{\partial_t I_n}{N_n} = \beta \frac{(N_n - I_n) I_n}{N_n N_n} - \alpha \frac{I_n}{N_n} + \sum_{m \neq n} \left[\omega_{nm} \frac{I_m N_m}{N_n N_m} - \omega_{mn} \frac{I_n}{N_n} \right] \tag{46}$$

and with $j_n = I_n/N_n$ it is

$$\partial_t j_n = \beta(1 - j_n) j_n - \alpha j_n + \sum_{m \neq n} \left[\omega_{nm} j_m \frac{N_m}{N_n} - \omega_{mn} j_n \right] \tag{47}$$

According to the detailed balance condition (see equation (31)) we have $\omega_{mn} = \omega_{nm} \frac{N_m}{N_n}$, and with that we finally arrive at equation (32):

$$\partial_t j_n = \beta(1 - j_n) j_n - \alpha j_n + \sum_{m \neq n} \omega_{mn} \left(j_m - j_n \right) \tag{48}$$