

PANDEMICS

Year 2017-18

School: I.S.I.S.S. "M.Casagrande", Pieve di Soligo, Treviso - Italy

Students:	Beatrice Gatti	(III)
	Anna De Biasi	(III)
	Erica Piccin	(III)
	Marco Micheletto	(III)
	Yui Man Kwan	(III)
	Camilla Viviani	(IV)
	Leonardo Breda	(IV)
	Andrea Munarin	(IV)
	Simone Boscaratto	(V)
	Bianca Della Libera	(V)
	Francesco Luigi De Faveri	(V)

Teachers: Fabio Breda
Francesco Maria Cardano
Francesco Zampieri

Researcher: Alberto Zanardo, University of Padova - Italy

Presentation of the research topic: PANDEMICS

A disease is spreading. At every moment for each patient occurs that: a patient could infect another entity with a certain probability; a patient could heal with a certain probability; a patient could die with a certain probability. What could happen considering different probabilities? Is it better to vaccinate a part of the population or to isolate the population in small groups?

Brief presentation of the conjectures and results obtained:

We attempted a statistical approach, developing a $n \times n$ cells matrix that represents the population, with the disease spreading by one person and infecting the still healthy adjacent cells. Our simulation uses the three probabilities as input parameters and returns the percentages of dead, healed and not even infected people, as well as the duration of the pandemic. We checked the behaviour of the simulation in relation to real diseases and looked for mathematical relations for the output data. We finally developed the algorithm to take into account the possibility of vaccination and to simulate people movement.

Contents

1 Approach to the problem	3
2 Simulation tools	3
2.1 C++ Algorithm	3
2.2 Java Algorithm	4
3 Simulated models	6
3.1 The simulation	8
4 Data analysis, static case	9
4.1 Results	10
5 More realistic models	13
5.1 Vaccination and herd immunity	13
5.2 People movement and transmission ways	13
Conclusions	14
Webography	16

1 Approach to the problem

Due to the huge dimensionality of the problem's parameters, that in our opinion would have prevented the success of an analytical approach, a numerical/statistical method has been followed. In fact we developed Java and C++ tools both to simulate several disease models/scenarios and to retrieve and statistically treat the output data.

Another important point of our strategy has been to adopt a matrix point of view of the situation. The structure and the behaviour of this matrix changed during the study of the problem. In fact we started with a static approach where every cell represented a person who can't make any movement and than we switched to a more realistic simulation, adding the possibility of movement and vaccines which inhibit the spread of the disease.

Our Java application works with the input parameters mentioned in the problem statement: the infection, healing and death probabilities. We also decided to develop our simulation based on some real epidemics and to compare our results with the ones expected from the aforesaid diseases. So we extracted some data for medical literature and adapted them to the ones requested in the problem statement.

2 Simulation tools

Our simulation has been developed firstly using a C++ algorithm, and then we switched to a Java application in order to use a graphical user interface (GUI).

2.1 C++ Algorithm

This application has been used to give us a quick idea of the behaviour and the spreading of the infection in a static model in which all the people are in contact each other (that will be thereafter referred to as the *rock concert model*). In this version of the program the basic logic of transmission of the disease has been sketched out.

Our first concern was to define a metric on the matrix: we decided to consider all the 8 cells neighboring to a chosen one as adjacent, and capable of infecting it as a consequence.

We also decided to let the disease spread from a single cell (the *zero patient*) located in the center of the matrix. In our opinion this choice does not involve a significant loss of generality, because in the *real case* the disease would be able to spread in a *world* large enough not to depend in a significant way on the starting position.

We finally agreed in fixing some additional rules regarding the behaviour of the simulation:

- healed people are immune from the disease and can no longer spread it;
- the illness cannot even be spread by dead people;
- the epidemic is thus to be considered exhausted when there are no longer any infected people in the matrix.

The better to follow the spread of the disease, we also decided to adopt a graphical approach according to the following numerical and chromatic codes:

State	Number	Color
Healty	0	White
Infected	1	Yellow
Healed	2	Green
Dead	3	Red

At each time step (arbitrarily assumed equal to one day) the status of every cell is updated according to the input probabilities. The algorithm proceeds evaluating the status of each cell of the matrix scanning it by row and putting the updated values on a *transport* matrix. Every status change is evaluated generating a random number which is then compared with the corresponding probability.

The possible states are checked in the following order and with the corresponding outcomes:

- if the cell's status $s = 3$ or $s = 2$ (the patient is dead or has already healed) the value remains the same;
- if the cell's status $s = 1$ the patient is diseased and it can die ($s' = 3$), heal ($s' = 2$) or stay infected ($s' = 1$) according to the corresponding input probabilities;
- if the cell's status $s = 0$ (the person is still healthy) it can be infected depending on the corresponding probability and the number of neighboring patient.

In the last case, in order to greatly reduce the number of random extractions (and thus improve the run time as a consequence) we searched a way to evaluate a total infection probability dependent on the number of adjacent patients. We figured out that this probability can be evaluated as complementary to that of resisting the attempts of infection by all the neighboring patients, as shown in figure (1).

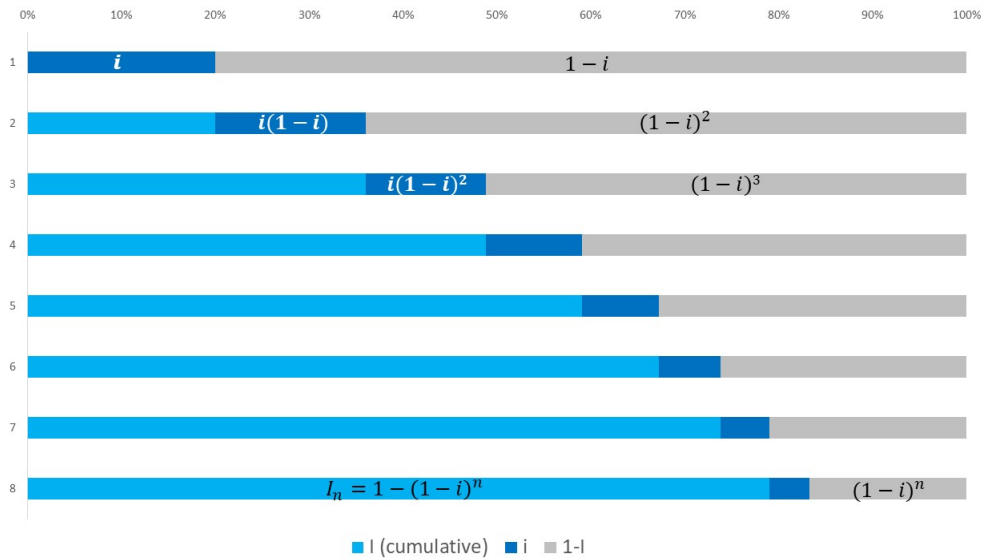


Figure 1: Cumulative probability of contracting the disease as a function of the number of neighboring patients and of the input infectivity. This probability can be evaluated as complementary to that of resisting the attempts of infection by all the adjacent patients.

Since the process is essentially binomial in nature, it's in fact difficult to determine a direct formula for the cumulative probability as the sum of gradually smaller contributions, while it's way easier to determine the overall probability of not being infected after n attempts, which turns out to be:

$$H_n = (1 - i)^n;$$

thus resulting in a total infection probability given by:

$$I_n = 1 - (1 - i)^n. \quad (1)$$

2.2 Java Algorithm

In order to simulate more complex scenarios, a Java algorithm that included the following features has then been developed:

- the possibility to take into account the effects of vaccination with variable coverage;
- the movement of the people with different means of transportation;

- the dimension of the matrix and its population;
- a random starting position for the *zero patient*;
- the existence of various transmission routes corresponding to different infection radii;
- a Graphical User Interface (GUI) the better to handle all the aforesaid options.

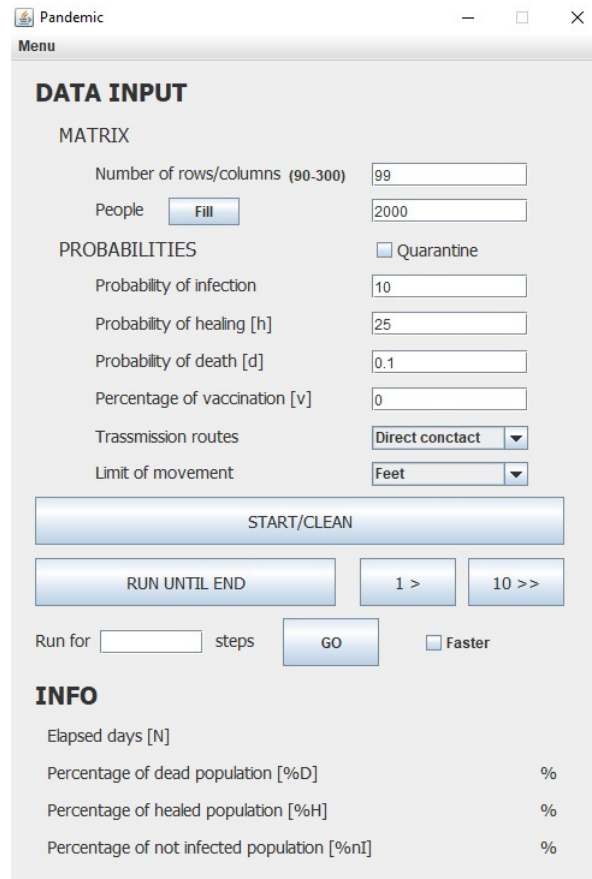


Figure 2: The Graphical User Interface of the Java app.

3 Simulated models

First and foremost, we had to establish some values of the probabilities to run the simulation. Initially we defined the geometric locus in which the three probabilities had to stay in:

- The probability of infection i can assume values from 0 to 100% ;
- The sum of the probabilities of healing and death can assume values from 0 to 100% .

These conditions define a triangular prism, in which every point correspond to a triplet of probabilities that can :

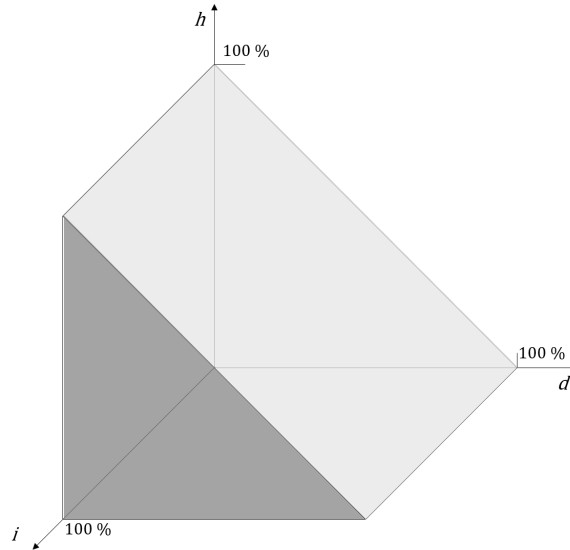


Figure 3: The triangular prism that is the *locus* of the parameters space.

We estimated some probabilities to run the program taking some data from medical literature. In particular, we analysed the basic reproductive ratio R_0 , the average course of illness N and the mortality rate D of seven real diseases, given that:

- Basic reproductive ratio R_0 : average number of new cases that an ill patient generates during the course of illness (considered as the whole duration of the disease);
- Course of illness N : average number of days that a disease lasts for, after which a patient can be considered eventually healed or died;
- Mortality rate D : percentage of infected people that are dead after the course of the disease.

We summed up the data in a table:

Disease	R_0	N	$D(\%)$
Ebola	2	30	50
SARS	3.5	45	15
Pertussis	14.5	80	0.35
Diphtheria	6.5	12	15
Measles	15	20	0.35
Mumps	5.5	20	0.01
Flu	1.65	7	0.03

These data have then been used to calculate the three probabilities needed:

- Probability of infection i : considered R_0 as the ratio between the number of new cases that a patient generates during the course of the disease N and the maximum number of tries this could have attempted over the same time, it can be expressed as:

$$R_0 = N \cdot i$$

In particular, in the static framework we have that the infection could spread the eight adjacent cells, so the probability i has to be multiplied by 8:

$$R_0 = 8N \cdot i \tag{2}$$

Thus, it follows that:

$$i = \frac{R_0}{8N}$$

- Probability of death d : the mortality rate D can be considered as the cumulative probability of being dead after N days of the disease, it can be applied the same reasoning used to calculate I_n given i and n (1), changing the number of infected adjacent cells n with the course of the disease N :

$$D = 1 - (1 - d)^N$$

It follows that:

$$d = \sqrt[N]{1 - D}$$

- Probability of healing h : considered that the probability of being healed at the end of the course of the disease is complementary to the probability of being dead, it follows that:

$$H = 1 - (1 - h)^N = 1 - D \Rightarrow h = 1 - \sqrt[N]{1 - H} = 1 - \sqrt[N]{D} \tag{3}$$

The three probabilities has been calculated for the seven diseases:

Disease	$i(\%)$	$d(\%)$	$h(\%)$
Ebola	0.8	2.28	2
SARS	1.0	0.36	4
Pertussis	2.3	0.004	7
Diphtheria	6.8	1.35	15
Measles	9.4	0.02	25
Mumps	3.4	0.0005	37
Flu	2.9	0.004	69

Some arbitrary values of the probabilities have then been established, covering with a greater density the low value zone, that is the more interesting one according to the data shown above. To do it properly we ordered the diseases by each probability, obtaining graphs in which the horizontal axis represents a cardinal number associated with the disease and the vertical one represents the corresponding probability, and fitting the medical data with exponential relations.

The described method is shown in the graph 4 for the healing probability, while the overall results for the three probability parameters are summarized in the following table.

$i(\%)$	$d(\%)$	$h(\%)$
0.8	0.0004	2
1.2	0.002	4
1.8	0.008	7
2.7	0.035	13
4.1	0.15	22
6.3	0.66	39
9.4	2.9	70
	12.5	

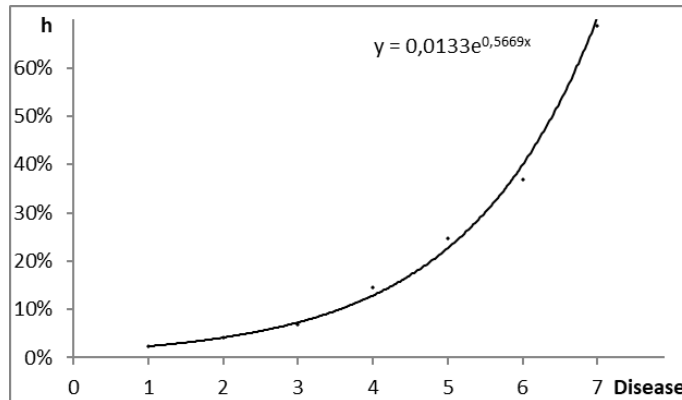


Figure 4: The exponential fit used to obtain input values of the healing probability for our simulations.

3.1 The simulation

To extract the data needed to do statistics on the evolution of the disease, each model has been run 100 times over a 501×501 cells matrix; the odd number chosen is big enough to let the disease spread almost freely for many steps and permits to let it start from the very center of the matrix.

For each run are obtained:

- Percentage of not even infected population $NI(\%)$;
- Percentage of the population dead of illness $D(\%)$;
- Percentage of the population healed from the disease $H(\%)$;
- Number of steps (cycles ran) needed for stability N .

The average value and the corresponding standard deviation σ of these four parameters has been calculated for each model.

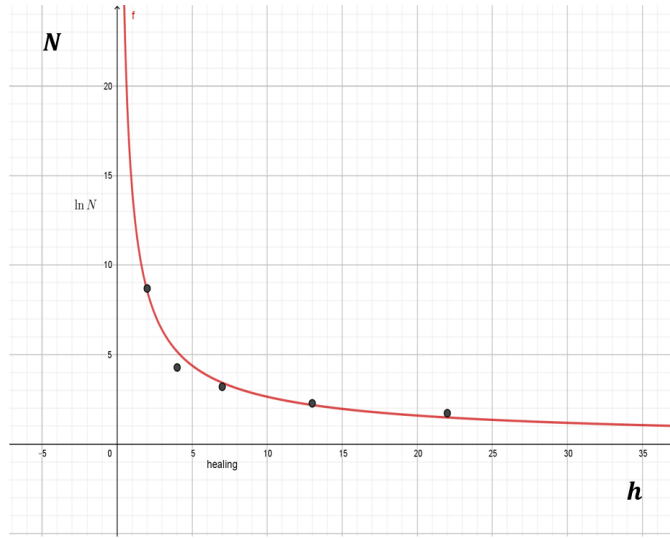


Figure 6: Power law model for the relation between the logarithm of the duration of the disease and the probability of healing

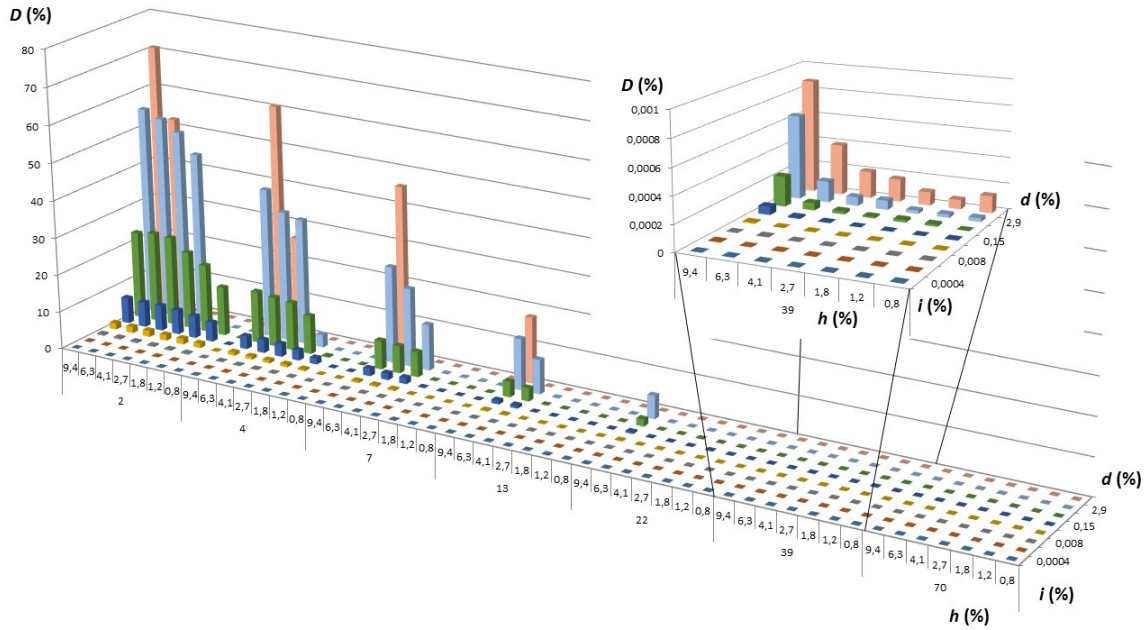


Figure 7: The mortality rate as a function of the three probabilities. A zoom on the areas where the data of the mortality rate are apparently zero shows they are just smaller in order of magnitude.

4.1 Results

The data obtained from the simulations were checked with those collected from some real diseases which had similar probabilities: measles and ebola. As measles had slightly different probabilities of death and healing from the one used in the simulations, the two mortality rates obtained are different and in particular the one calculated from the simulation is five time bigger than the measles one, but is still acceptable. Considering ebola, the two mortality rates are more similar but there are some differences in this case too,

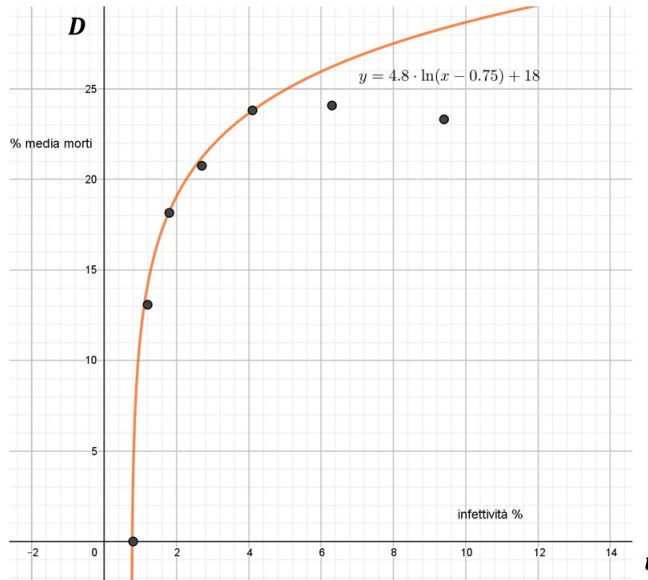


Figure 8: Logarithmic model for the relation between the mortality rate and the probability of infection, both in percentage

due to the same motive as before.

MEASLES

- Known data
 - Basic Reproductive Ratio $R_0 = 15$
 - Mortality Rate $\mathbf{D = 0,035\%}$
 - Duration $R = 20$ days
- Estimated parameters:

$$i = 9,4\%, \quad d = 0,02\%, \quad h = 25\%$$
- The most similar simulated model:

$$i = 9,4\%, \quad d = 0,035\%, \quad h = 22\%$$
- Estimated mortality rate

D = 0,157%

EBOLA

- Known data
 - Basic Reproductive Ratio $R_0 = 2$
 - Mortality Rate $\mathbf{D = 50\%}$
 - Duration $R = 30$ days
- Estimated parameters:

$$i = 0,8\%, \quad d = 2,28\%, \quad h = 2\%$$
- The most similar simulated model:

$$i = 0,8\%, \quad d = 2,9\%, \quad h = 2\%$$
- Estimated mortality rate

D = 67,5%

As first result it has been proven that, if the probability of infection is high enough, the ratio between the percentage of healed patients H and the percentage of dead patients D is approximately the same between the probability of healing and the probability of death;

$$\frac{H}{D} \simeq \frac{h}{d}$$

It has been also proven that decreasing the probability of infection this relation is not true any more.

Using the mortality rate graph it has been possible to study the change of this parameter in different situations, in connection with the change of the probabilities and two different kind of results have been obtained. It has been observed that, fixing the probabilities of infection and death and increasing the healing one, the mortality rate has a monotonic decrease.

Instead, if the probability of infection and healing are fixed, it has been proven that the biggest value of the mortality rate doesn't always correspond to the biggest value of the probability of death: this happens because if the probability of death was very high many patients would die and the disease would not even have the time to spread.

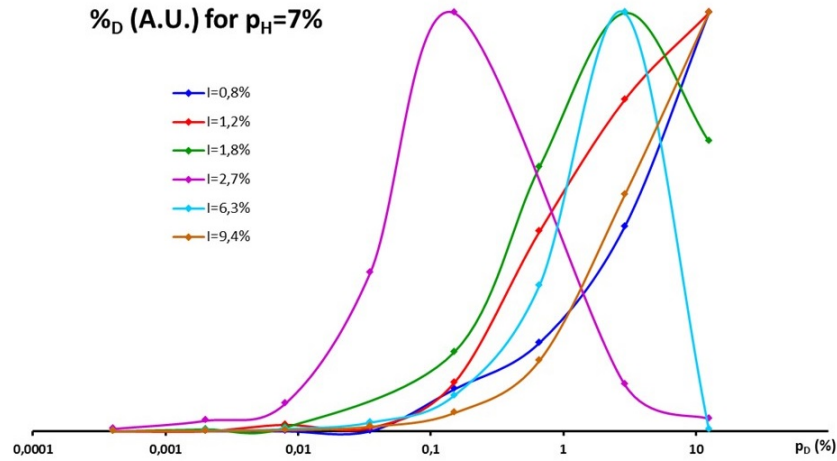


Figure 9: Trend of the mortality rate in function of the probability of death and the probability of infection.

This graph shows different developments of the mortality rate compared to the probabilities of infection. Different colours correspond to different curves: it can be seen that if the probability of infection is a low value, the mortality rate has a monotonic increasing; but as the probability grows, the mortality rate changes its trend and, in particular, has its biggest value when the probability of death is not the biggest. Instead, when the biggest value of the probability of infection is reached, the mortality rate has a monotonic increase, again.

5 More realistic models

In order to make the simulation more realistic, models describing the effect of vaccinations and movement have finally been added.

5.1 Vaccination and herd immunity

For simulating the effect of a vaccination we can choose a $\%_V$ (vaccination percentage), inserted in the input data, which describe the percentage of vaccinated people inside the matrix, allocated with a random distribution. To introduce the vaccines in the simulation we considered the cell representing the person vaccinated with the number 4 and to represent it in the graphic interface it was colored with the blue color. Therefore, the vaccinated cell can not be contagious and at the end of the run all the vaccinated people were considered healthy. Before we started the simulation, we focused on the *Herd Immunity Threshold* (HIT), that is the percentage of the population that must be vaccinated so that the disease does not spread:

$$HIT = 1 - \frac{1}{R_0}$$

We simulated a *measles-like* disease ($i = 10\%$, $d = 0,1\%$ e $h = 25\%$) with increasing percentages of immunized people (0 to 99%) each run 100 times. We attempted to provide both a semi-theoretical and a data estimation of HIT.

Starting with the theoretical estimation, we extracted N solving the following system by means of numerical methods:

$$\begin{cases} d = 1 - \sqrt[N]{1 - D}; \\ h = 1 - \sqrt[N]{D}. \end{cases}$$

Then we found R_0 from the relation (2), thus obtaining:

$$HIT = 1 - \frac{1}{R_0} = 0,92 = 92\%.$$

Considering the data estimation of HIT, using the $0\%_V$ files data, we calculated

$$D = \frac{\%dead}{1 - \%notinfected},$$

and inverted the relation (3), thus obtaining:

$$N = \frac{\ln D}{\ln(1 - h)}.$$

We then found R_0 as shown above, finally obtaining

$$HIT = 1 - \frac{1}{R_0} = 0,9566 = 95,66\%.$$

Considering the great number of uncertainties inherent in the simulated framework, the two estimated percentages can be considered compatible with each other and with the value reported in the medical literature (92 – 95%).

5.2 People movement and transmission ways

In order to simulate a more realistic behaviour, the algorithm has been modified to include the movement of people and different kinds of transmission modes. To obtain this new algorithm we had to add a new value to the matrix: -1 . This value corresponds to the empty cell, necessary to allow the movement of people.

As mentioned before, we found it useful to define an additional parameter: the range of movement. This parameter determines the maximum number of cells that a person can travel in one step. The new position is determined by a random number for the x axis and a random number for the y axis ranging from - movement radius to + movement radius.

The movement rays we have defined are:

- static;
- walking (max 100 cells per axis);
- bike (max 200 cells per axis);
- car (max 400 cells per axis).

The random is an integer generated between the values just described. Being in the random the possibility of obtaining the value 0, people have the possibility to remain still. Another condition that we used when we decided to implement the movement is the fact that each person can not move on to another.

The other parameter added is the radius of infection: in this way it is not enough to check only the 8 adjacent cells to determine the probability of being infected, but the $(N \cdot 2 + 1)^2 - 1$ cells around the person forming a new control square.

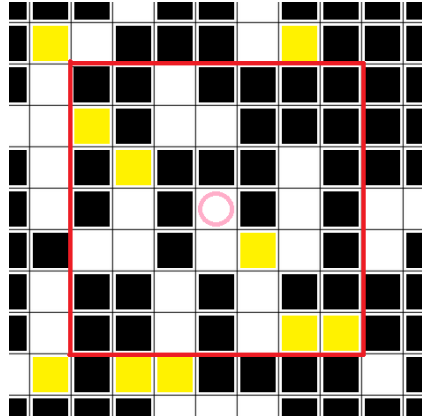


Figure 10: An example for trasmission routes.

For example, if the radius of infection is 3, as we can see in the image, the controlled cells are the 48 around. In this case the infected people in the delimited region are 5, so the probability of the central cell to be infected corresponds to:

$$I_n = 1 - (1 - i)^5$$

We defined different rays of infection:

- direct Infection (r=1 cells);
- droplets (r=5 cells);
- air (r=10 cells);
- vector (r=20 cells).

After defining the movement, we proceeded to collect the data to find the HIT in a more realistic model. Using again our *measles-like* disease, vaccination rates that focused around the hypothetical HIT (from 80% to 99%), the maximum range of motion (Car) and the biggest radius of infection (Vector) we obtained this graph, representing the number of uninfected people compared to the vaccination rate:

It is found that HIT in the static model is quite similar to the one evaluated in this model. In fact, the HIT calculated from this model, as can be seen from the graph, is equivalent to 92%.

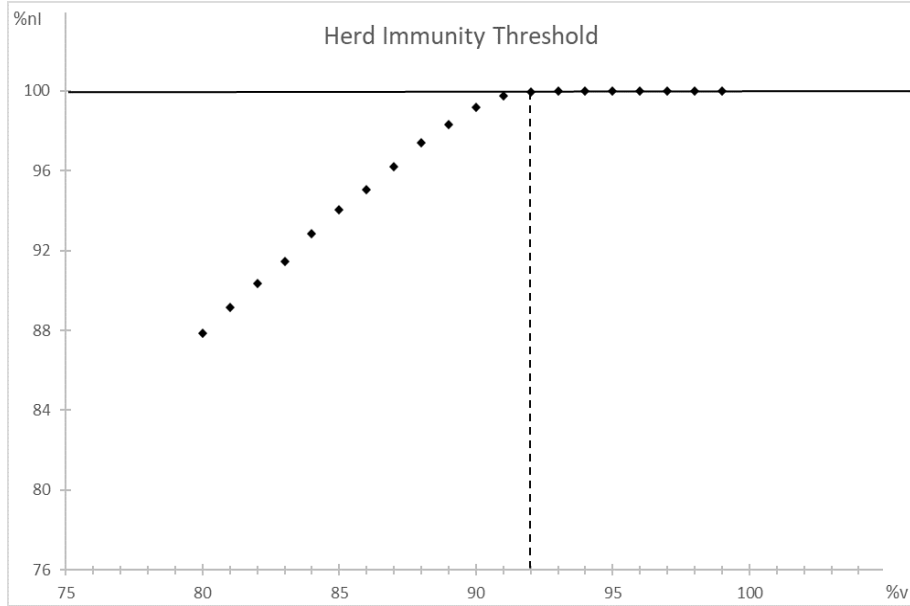


Figure 11: HIT for the model with movement

Conclusions

In this work, a numerical and statistical approach has been used to simulate the spread of various diseases on a two-dimensional matrix. Using C++ and Java languages, various algorithms have been developed to simulate the behaviour on a static framework first, and then extend the study to a more general case including vaccinations, people movement, and different transmission ways.

The problem statement required to consider three different probabilities: infection, healing and death ones. We decided to chose their values by comparison with data related to real diseases and checked that our simulation provided results in accordance with what was expected for the same epidemics.

Regarding the obtained results it is also essential to highlight that, while some output distributions (6) can be fitted with certain relations (4), on the other hand a precise mathematical law describing the spread of the disease can't easily be defined in our two-dimensional model. This is shown for example in graph 9, where it can be noticed that the number of dead people isn't always associated with the maximum death probability, but it is determined by conditions of unstable equilibrium in which the infectivity of the disease plays a decisive role too.

In the last part of this work, a more general model including vaccinations, people movement, and different transmission ways, has finally been developed. Vaccinations have been firstly applied on the static framework, estimating the *Herd Immunity Threshold* for a *measles-like* disease both in fully numerical way and semi-theoretically, and verifying that the value obtained was compatible with that reported in the medical literature. Finally a fully dynamical framework was implemented, including various forms of people movement and transmission ways, verifying in this case too that the HIT obtained was compatible with the values previously estimated and reported in the medical literature.

Webography

1. https://en.wikipedia.org/wiki/Herd_immunity
2. https://en.wikipedia.org/wiki/Ebola_virus_disease
3. https://en.wikipedia.org/wiki/Severe_acute_respiratory_syndrome
4. https://en.wikipedia.org/wiki/Whooping_cough
5. <https://en.wikipedia.org/wiki/Diphtheria>
6. <https://en.wikipedia.org/wiki/Measles>
7. <https://en.wikipedia.org/wiki/Mumps>
8. <https://en.wikipedia.org/wiki/Influenza>