# FRTN30 Network Dynamics: Hand-In 4

*Alfred Bornefalk, al1718bo-s@student.lu.se, 2022-05-17*
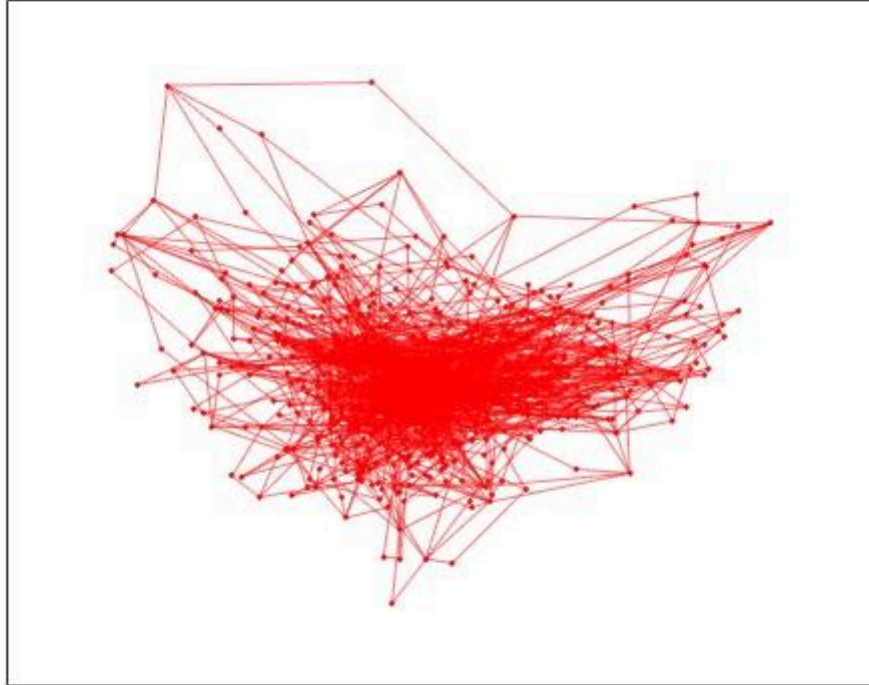
# Table of Contents

# 1. Preliminary Parts

In this home assignment, we will ultimately simulate the influenza H1N1 2009 pandemic in Sweden, with the goal of learning the network-structure characteristics and disease-dynamics parameters of this very pandemic. The task is divided into the four parts described in the table of contents. Accordingly, we will begin by doing the two preliminary parts; simulating an epidemic on a given graph, and generating a random graph with preferential attachment.

## 1.1. Epidemic on a Known Graph

In this subsection, we will simulate an SIR epidemic model on a symmetric $k$-regular graph $G = (V, \varepsilon)$ with $|V| = 500$ nodes, and $k = 4$. Additionally, we let $\beta =. 3$ and $\rho =. 7$. Defining one week as being one unit of time, the epidemic will be simulated for 15 weeks. The initial configuration consists of 10 infected nodes, selected randomly from $V$.

After following the probability algorithm driving the epidemic described in the assignment, and performing $N = 100$ simulations, the average number of newly infected individuals each week, and the average total number of susceptible, infected, and recovered individuals weekly, was calculated. The respective plots are shown in figure 1 and figure 2 below.
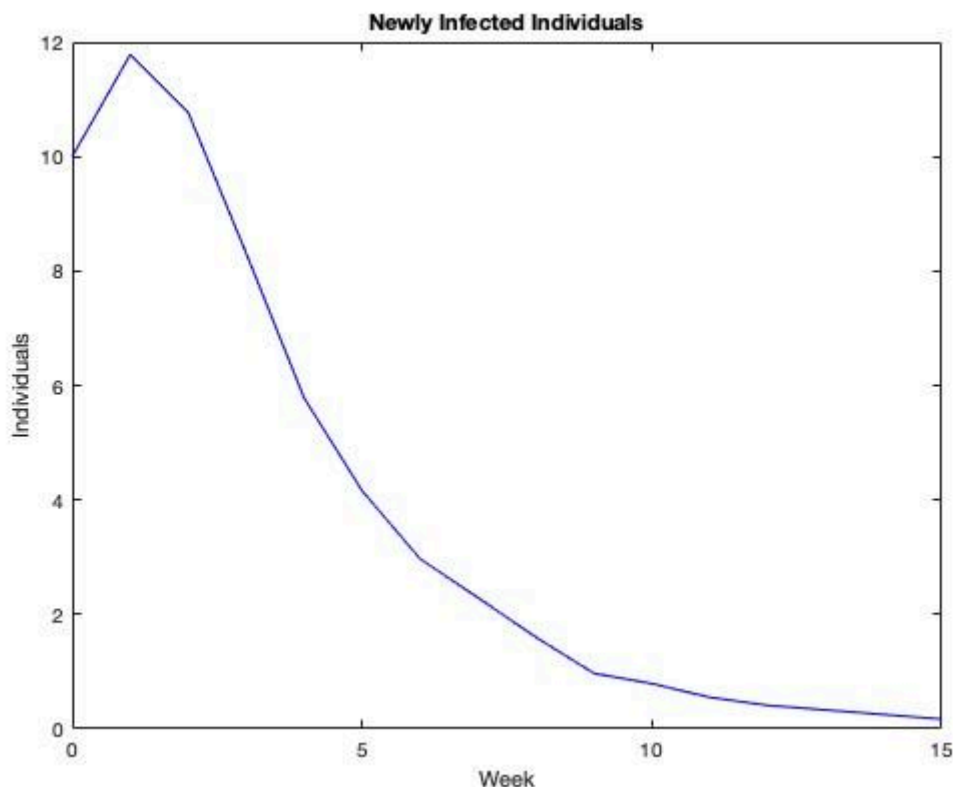


*Figure 1: The average number of newly infected individuals in the 4-regular graph, with $|V| = 500$, $k = 4$, $\beta =. 3$, and $\rho =. 7$, each week when performing $N = 100$ simulations.*
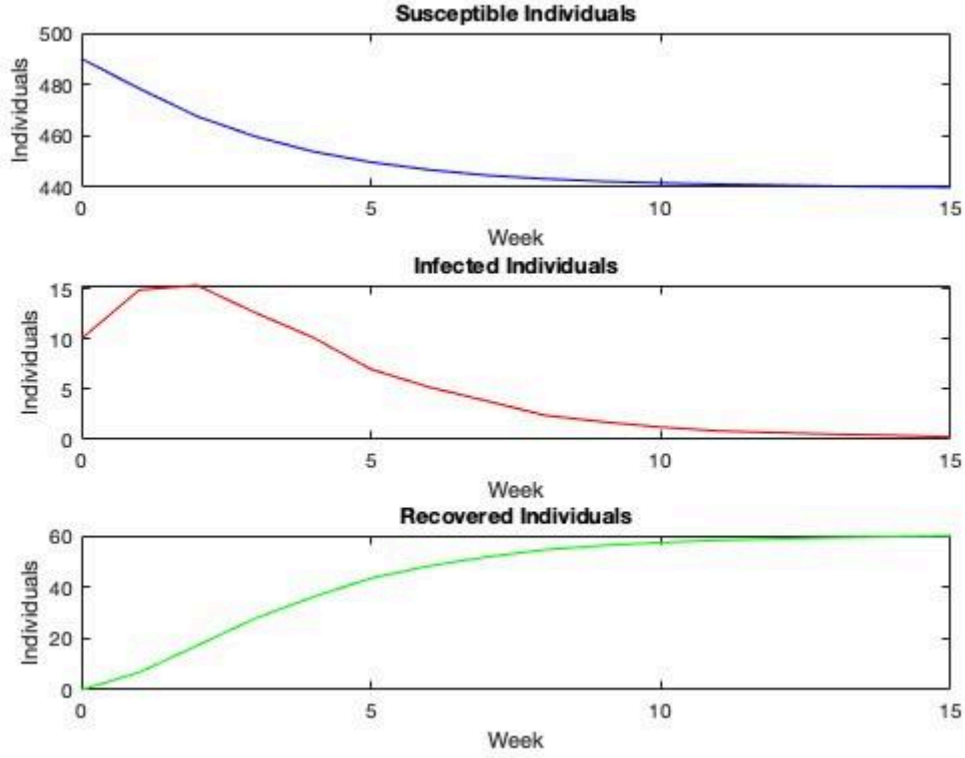
*Figure 2: The average total number of susceptible (blue), infected (red), and recovered (green)*
*individuals in the 4-regular graph, with $|V| = 500$, $k = 4$, $\beta = .3$, and $\rho = .7$, each week when*
*performing $N = 100$ simulations.*

Observing figure 1, we can see that in the first week, 11 nodes (on average) become infected.
Thereafter, this number decreases monotonically towards 0. Furthermore, figure 2 showcases
that the average number of infected nodes (or individuals) increases the first two weeks.
Then, this metric drops towards 0 as well as time (the weeks) moves on. The explanation
behind this situation can, at least partly, be found in the subgraph of recovered individuals.
Here, we see that the infected individuals recover, and by construction, they cannot become
infected again from this point. Letting $w$ denote the current week, and introducing

$$\bar{I}_{new}(w), \bar{S}(w), \bar{I}(w), \bar{R}(w)$$

as the average number of newly infected individuals, the average number of susceptible
individuals, the average number of infected individuals, and the average number of recovered
individuals at week $w$, respectively, we can distinguish from the plots that these numbers are
more or less in line with those given to us in the problem, meaningly

$$\bar{I}_{new}(w) \approx .2, \bar{S}(w) \approx 440, \bar{I}(w) \approx .3, \bar{R}(w) \approx 59.5$$

Hence, the implementation procedure can be viewed as verified.

## 1.2. Generating a Random Graph

Now, we will turn our focus towards generating a random graph according to the preferential attachment model described in the problem. More specifically, our goal is to, by using said preferential attachment, generate a random graph of a large size with average degree $k \epsilon Z^+$. With the requirement being *at least* 900 nodes, we will work with $n = 1,000$. In this process, we let the initial graph $G_1 = (V_1, \varepsilon_1)$ be a complete graph with $|V_1| = k_0 = k + 1$ nodes. Using our random graph generation-algorithm for $k_1 = 2$, $k_2 = 3$, $k_3 = 4.6$, $k_4 = 6$, and $k_5 = 10$, the average degrees of the respective graphs can be seen in table 1 below.

*Table 1: The set of $k_i$'s compared to the respective average degree of the randomly generated graph.*

| $k$ | Average Degree |
|-----|----------------|
| 2 | 2 |
| 3 | 3.012 |
| 4.6 | 4.588 |
| 6 | 6 |
| 10 | 10 |

As can be seen in table 1 above, the average degree of the graph corresponding to the choice of $k$ is (seemingly) working with perfect precision for the choice of even integers (the tested choices were, of course, 2, 6, and 10). In the case of odd integers (i.e. 3) and non-integer values (4.6), there are slight fluctuations around the sought value. Since the implementation is based on stochastic assignment, the average degrees values also vary from simulation to simulation. If adding a large number of nodes, one may expect that these will, on average, be equal to the input $k$. Thus, the results are deemed satisfying.

# 2. Simulating a Pandemic Without Vaccination

Moving on, we are going to use the graph generation mechanism above and then simulate an epidemic on it. Again, the disease-propagation model is the discrete-time version of the SIR epidemic model used previously. Utilizing these methodologies, our aim is to generate a preferential attachment random graph $G = (V, \varepsilon)$, with $|V| = 500$ nodes, with an average degree of $k = 6$. We let $\beta = .3$ and $\rho = .7$, and we will once again simulate the epidemic for 15 weeks. Once more, 10 infected nodes are chosen at random in the start. Figure 3 and figure 4 below displays the average number of newly infected individuals each week, and the average total number of susceptible, infected, and recovered individuals weekly, respectively.
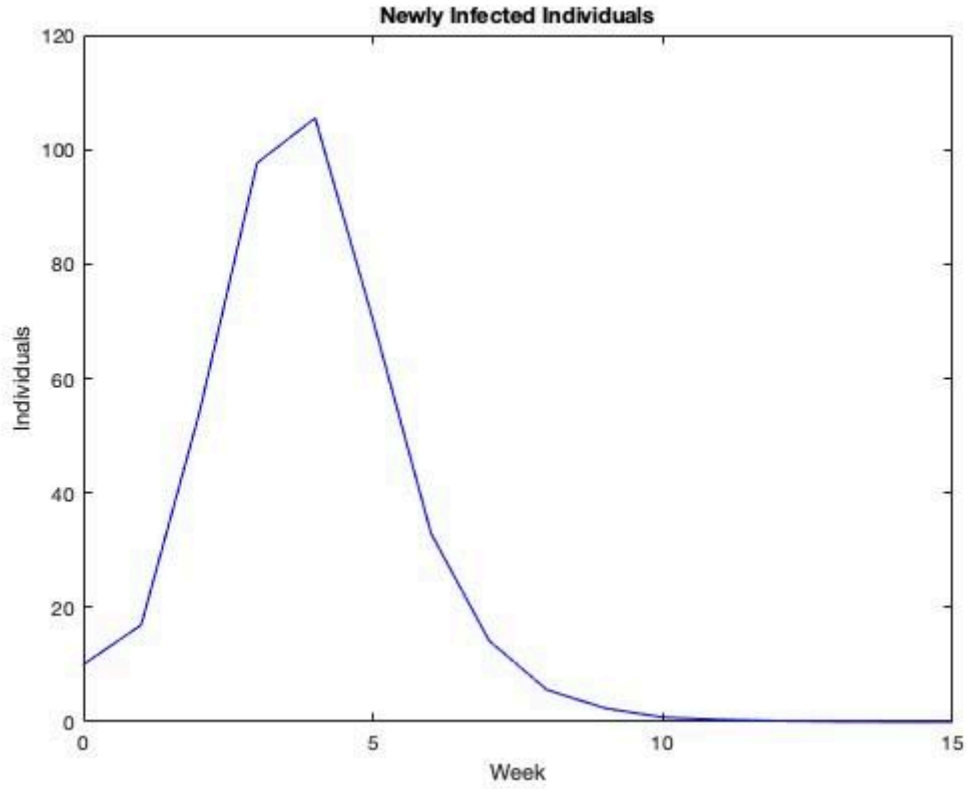
*Figure 3: The average number of newly infected individuals in the preferential attachment graph, with $|V| = 500$, $k = 6$, $\beta = .3$, and $\rho = .7$, each week when performing $N = 100$ simulations.*
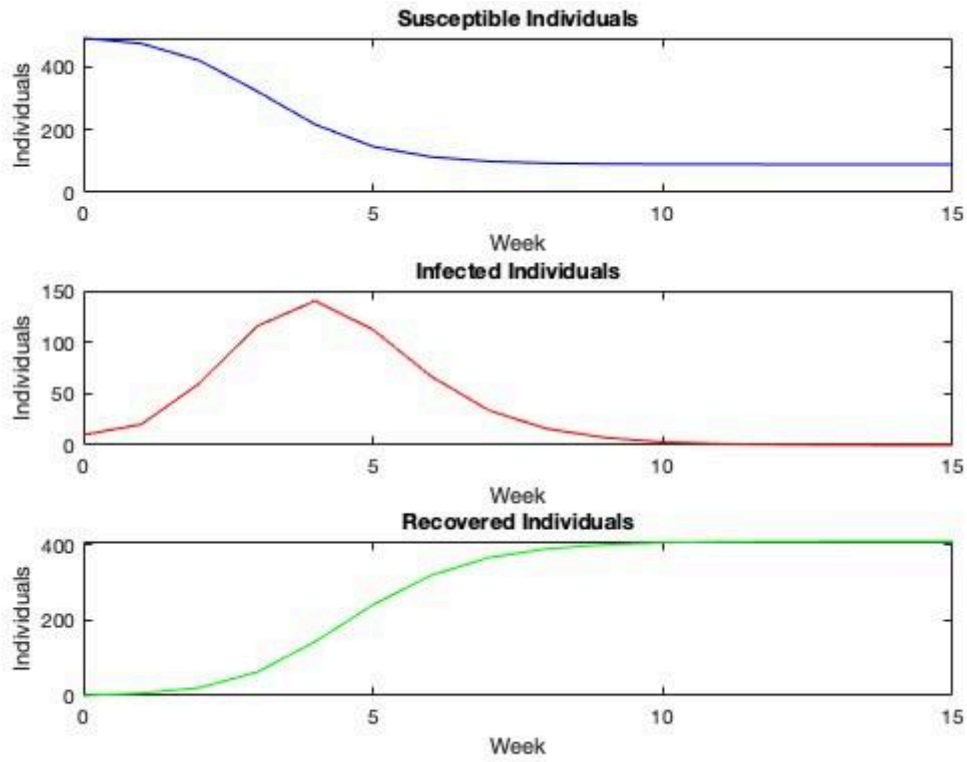


*Figure 4: The average total number of susceptible (blue), infected (red), and recovered (green) individuals in the preferential attachment graph, with $|V| = 500$, $k = 6$, $\beta = .3$, and $\rho = .7$, each week when performing $N = 100$ simulations.*

Looking at figure 3, we can observe quite stark differences from figure 1. Now, since the network is more connected compared to the previous section, the virus also spreads more effectively. The average number of newly infected individuals this time around increases the first four week, to thereafter drop to a level close to zero around the tenth week. Having this in mind whilst looking at figure 4, it seems wise to conclude that this is due to the fact that, by construction, no one dies in our network, and the recovered individuals turn immune to the virus. Thus, after ten weeks, the infection potential is almost eradicated. According to the subplot of the average total number of recovered individuals in figure 4, roughly 80% of the population becomes infected when there is no vaccine.

# 3. Simulating a Pandemic With Vaccination

In this section, we will essentially do the exact same thing as above, with the difference being that we try to take some action to slow down the epidemic through vaccination. We modify our model in such a way that *once a person is vaccinated, it cannot be infected*. The vector Vacc($w$), describing the total fraction of the population that has received vaccination by each week $w$, is given in the problem. One caveat is that vaccinations will be distributed uniformly at random to the entire population that has not yet received vaccination. Consequently, an infected, or recovered, individual may receive vaccination as well. Figure 5 and figure 6 below showcases the simulation results in similar manner as the previous sections.
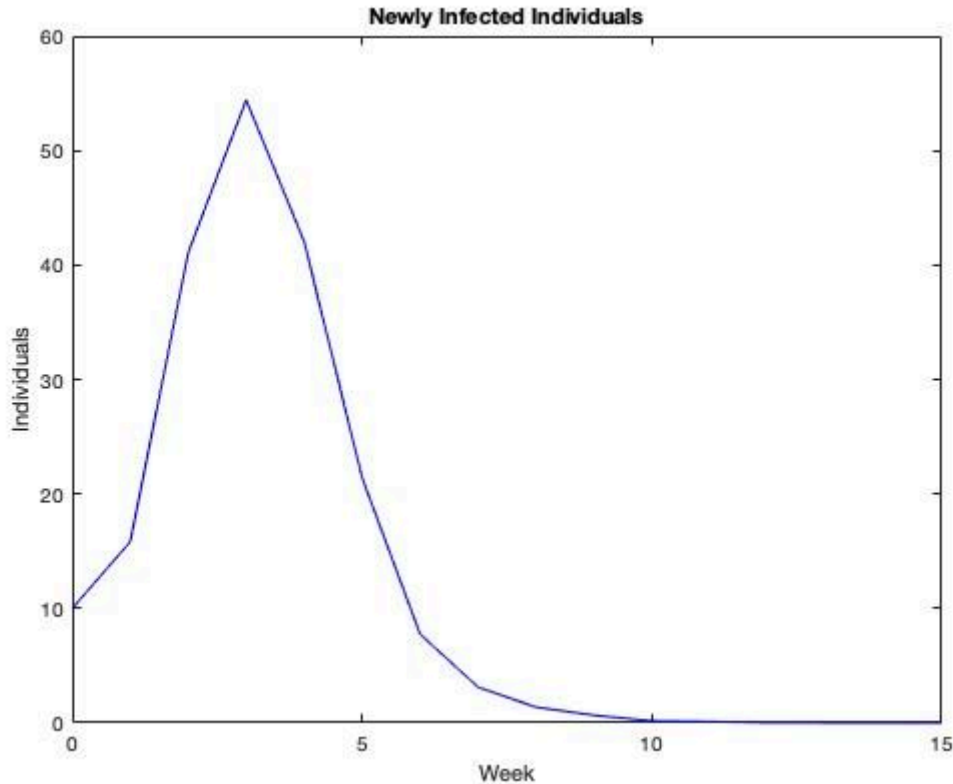


*Figure 5: The average number of newly infected individuals in the preferential attachment graph, with $|V| = 500$, $k = 6$, $\beta = .3$, $\rho = .7$, and Vacc($w$), each week when performing $N = 100$ simulations.*
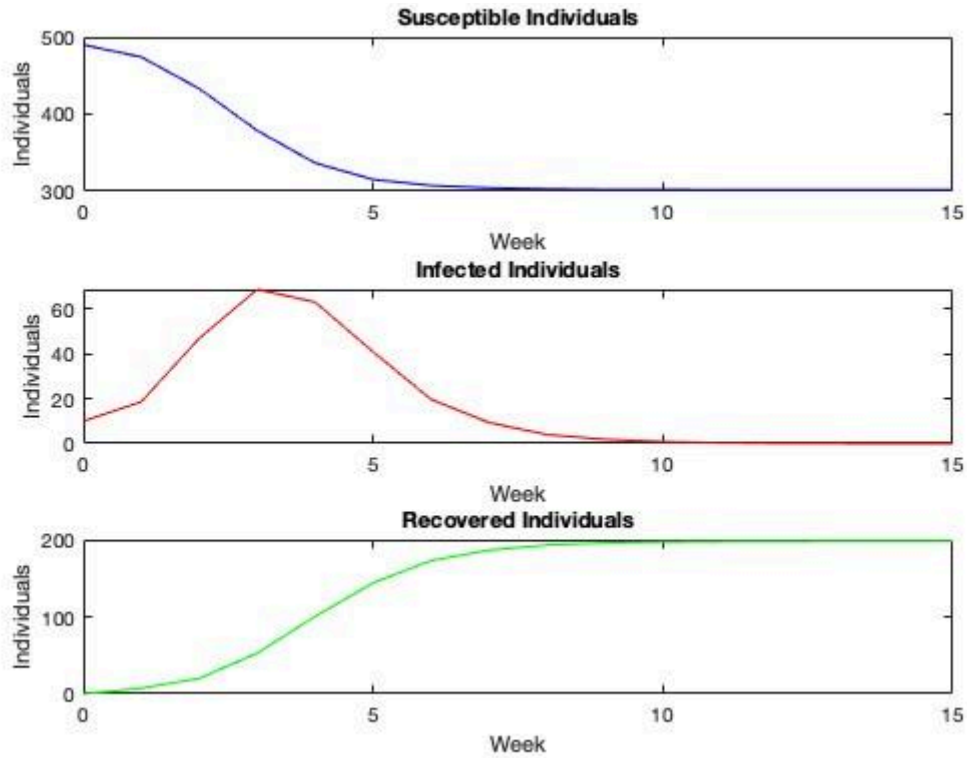
*Figure 6: The average total number of susceptible (blue), infected (red), and recovered (green) individuals in the preferential attachment graph, with $|V| = 500$, $k = 6$, $\beta = .3$, $\rho = .7$, and Vacc ($w$), each week when performing $N = 100$ simulations.*

Comparing figure 5 with figure 3, as well as figure 6 with figure 4, it is evident that the vaccination yielded the desired result in the simulations. The shape of the curves are highly similar, but the scaling factor has shifted quite dramatically. In the case of no vaccinations, the maximum number of newly infected individuals on average in one week was in excess of 100 (figure 3), whereas vaccinations brought this number down to circa 55 (figure 5). It can also be seen in figure 6 that the fraction of the population becoming infected with the slowing down measure decreased rather significantly, from roughly 80% (figure 4) to around 40%.

# 4. The H1N1 Pandemic in Sweden 2009

In the final, and ultimate part, of this assignment, we will use all of the previous parts to estimate the social structure of the Swedish population and the disease-spread parameters during the H1N1 pandemic in 2009. The simulation will run between week 42, 2009 and week 5, 2010. During these weeks, the total fraction of the population that had received vaccination is given by the vector Vacc($w$) in the problem. Furthermore, working with $n = |V| = 934$ for efficiency purposes, the actual number of newly infected individuals each week is represented by the vector $I_0(w)$; its values are also described in the problem.

To find the set of parameters $k$, $\beta$, and $\rho$ that best matches the real pandemic (i.e., the set of the parameters that minimizes the root mean-square error), we will perform a gradient-based

search over the parameter space; this algorithm is given in the problem description. The initial guesses of the average degree $k$ and the disease-spread parameters $\beta$ and $\rho$, as well as the values for their respective deltas, were run with the values suggested. Table 2 showcases, in order, the ten smallest values of the root mean-square error (RMSE) after performing multiple rounds of simulations with the algorithm.

*Table 2: The ten lowest values of RMSE based on the gradient-based search over the parameter space of the H1N1 pandemic dataset, presented in ascending order.*

| Rank | $k$ | $\beta$ | $\rho$ | RMSE |
|------|-----|---------|--------|------|
| 1 | 6.7 | .3 | .7 | 3.7903 |
| 2 | 7.7 | .3 | .7 | 3.915 |
| 3 | 5.6 | .4 | .7 | 4.5236 |
| 4 | 10 | .2 | .6 | 4.5558 |
| 5 | 5.6 | .4 | .6 | 4.6573 |
| 6 | 6.7 | .3 | .8 | 4.6615 |
| 7 | 6 | .3 | .5 | 4.7904 |
| 8 | 5.8 | .3 | .5 | 4.8859 |
| 9 | 6.8 | .3 | .5 | 4.9038 |
| 10 | 9 | .2 | .6 | 4.9314 |

As can be seen in table 2, obtaining a RMSE close to 0, or at least a value that stood out amongst the others, was difficult. The results varied between 3.8 and 5.8 in the different runs, and among the four best (smallest) scores, $k$ varies between 5.6 and 10. Since the element of randomness plays a role as well, new simulations would likely provide somewhat different results. With the information at hand however, it is concluded that the best estimates of the parameters in question are the ones associated with the smallest RMSE, i.e.

$$k = 6.7, \beta =.3, \rho =.7$$

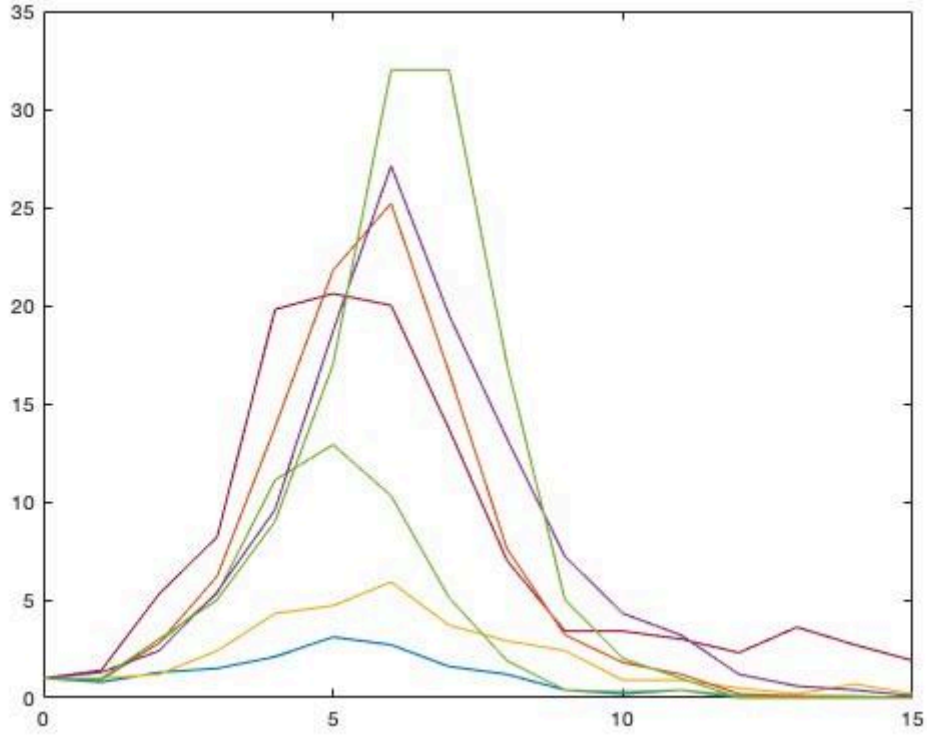The process of obtaining the above estimates are visualized in figure 7 below.

*Figure 7: The process of the algorithm finding the best parameter set $k = 6.7$, $\beta =.3$, and $\rho =.7$ through minimizing the RMSE. The green plot represents the true scaled-down value of newly infected individuals weekly in Sweden during the H1N1 pandemic, whereas the other plots represent more and more optimized fits regarding the average number of newly infected individuals in the simulated Swedish social structure, with the blue plot having the largest RMSE and the purple plot having the smallest RMSE. The x-axis represents the weeks, and the y-axis the number of infected individuals, with $|V| = 500$, and Vacc(w), when performing $N = 10$ simulations.*

From figure 7, we can see the optimization process that led to the best parameter estimates. The final fit (the purple plot) can be deemed a relatively good resemblance of the actual (scaled-down) values, as shown by the green plot. Figure 8 highlights these plots, whereas figure 9 uses the parameter estimates to generate the average total number of susceptible, infected, and recovered individuals weekly when choosing $N = 100$.
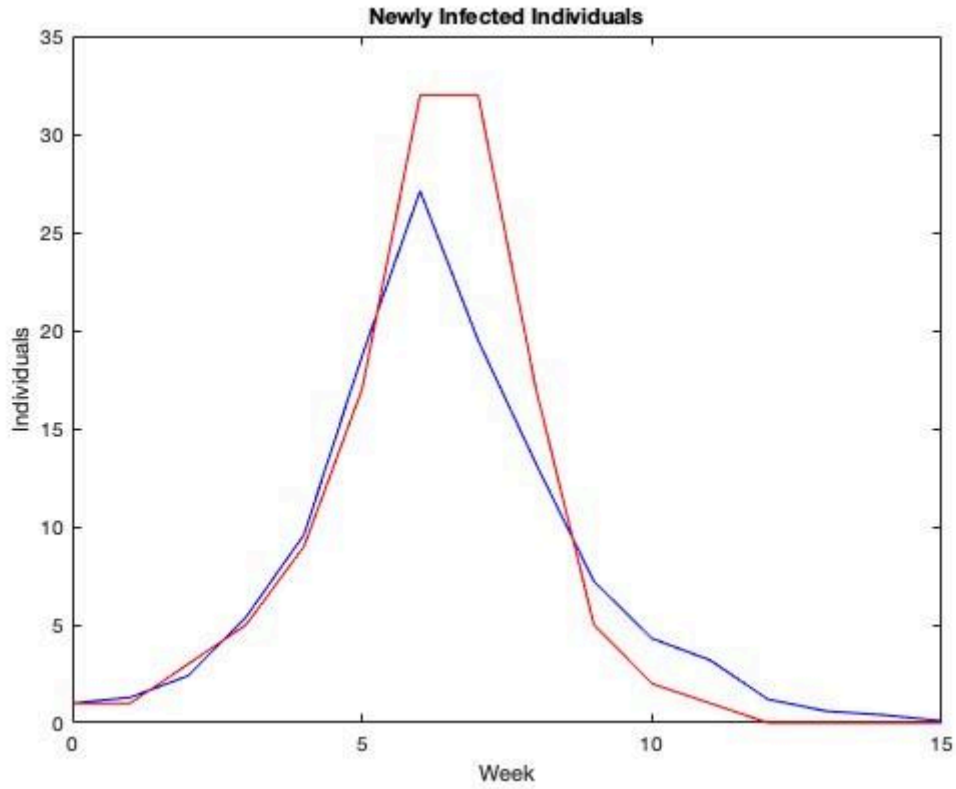
*Figure 8: The average number of newly infected individuals in the simulated Swedish social structure (blue), with $|V| = 934$, $k = 6.7$, $\beta = .3$, $\rho = .7$, and $Vacc(w)$, each week when performing $N = 10$ simulations, compared to the true scaled-down value of newly infected individuals weekly in Sweden during the H1N1 pandemic (red).*
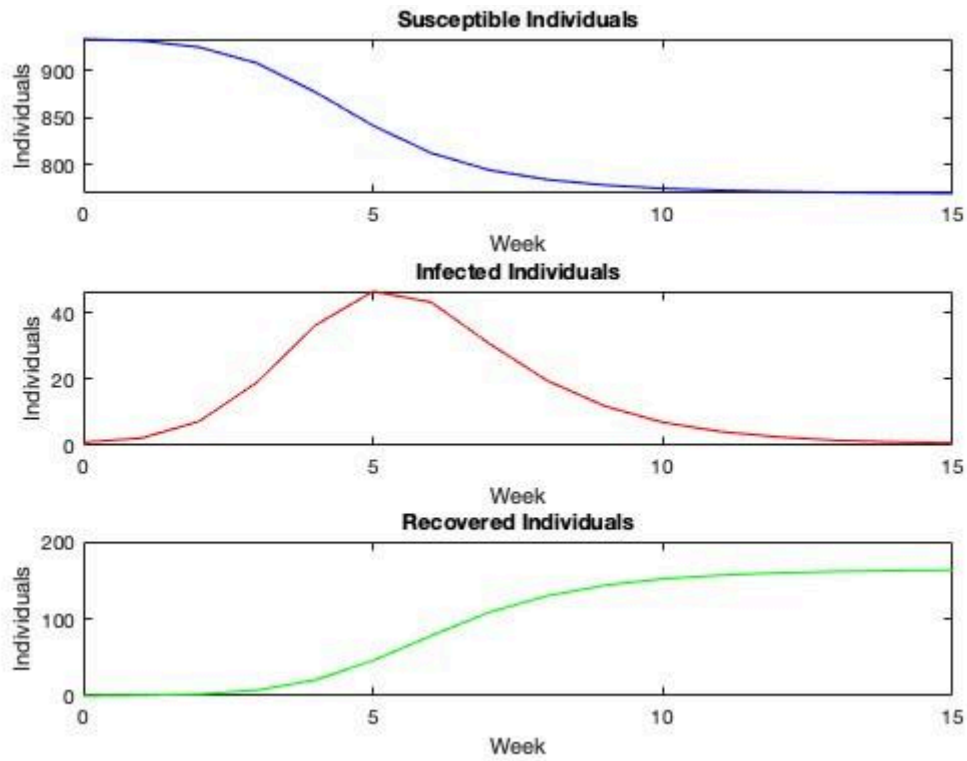
As mentioned above, figure 8 is an extract of figure 7, with the difference being that only the best fit and the actual values are presented. Moving on to figure 9, we can observe that the average total number of infected individuals peaks five weeks after the start of the simulation process, corresponding to week 47 in 2009. From that point, the number drops monotonically, to the point of the virus being almost eradicated by simulation week 15 (week 5 in 2010). We know from the problem description that about 1.5 million people out of a total of 9.34 million were infected with H1N1; this corresponds to a total infection rate of approximately 16%. From figure 9, dividing the average total number of recovered individuals in week 15 (163) with the size of the simulated Swedish social structure (934), we get a simulated total rate of infection of circa 17%. This further suggests that our modeling choice is a reasonable one, and with our findings, we could in a future study experiment with the vaccination vector to see how sensitive our results are to the rate of vaccination. Indeed, investigations like this one should be of interest to scientists and politicians alike; as we have learned during the past few years, another pandemic could always be around the corner.