Scientific Computation Project 4

Matthew Cowley, CID 1059624

April 24, 2020

1

The article is very appropriate for current times as it is about modeling epidemics, predicting where they originate from and how they potentially spread. Understanding these processes allows for better practical responses to epidemics, such as Covid 19 for example.

The article considers the spread of an infection on a collection of M coupled populations. Its starts by introducing a standard SIR model (1), similar to the example covered in lectures, to define the spread of an infection within these local populations. The inter-population spread is dependent on the number of people traveling between them; this is defined by the rate equation (2). The rates in (2) are defined using direct air traffic between local populations.

Models (1) and (2) are then combined, along with the addition of a sigmoid function to give (3), assuming that the total flux in and out of a population is proportional to the local population size. A key aspect of the model (3) is matrix P, which is passenger flux.It is akin to the Random walks system studied in CW2. But for this problem it is more appropriate to investigate other centrality's, rather than degree centrality; perhaps the weighted betweenness centrality or Percolation centrality.

As (3) is a coupled PDE system it would perhaps be interesting to investigate for chaotic behaviour, in a similar way to CW3. Seeing what initial conditions cause such behavior and to what degree.

In the article the model (3) is used to simulate an outbreak originating in Hong Kong and comparing results to SARS (2003) and H1N1 influenza (2009) epidemics. It compares results using standard distance in Fig 1, vs arrival times. There is no clear correlation or wavelike behavior. The article then goes on to define a new distance in (4) and then an effective distance (5), which is essentially a shortest path. To calculate this it would be appropriate to use a algorithm similar to Dijkstra's.

This effective distance gives far better results with strong correlations in Fig 2, showing wavelike behavior from the origin location in simulations. It then goes on to analyse results produced by the GLEAM model.

The article claims that it can compute relative arrival times with limited data using (6), which is often a major problem with alternative epidemic models. This would make it very appropriate in the study of Covid 19. (6) requires heavily on P, which is why a study of it, using the centrality's already stated,is potentially insightful.

The article shows how the epidemic origin is correlated to the effective distance, with the actual origin have good concentricity and a wavelike structure, seen in fig 4. It then shows how the origin can be calculated with only a snapshot of data using a method akin to characteristic roughness, giving fairly reasonable results, for small time lengths.

The article would perhaps be more complete if it provided examples or cases where their model does not perform well, as it seems to pick and choose it's simulations.

If the model in the article is correct it would suggest an appropriate response to Covid19 would be to shut down all forms of transport and halt travel, then the virus would spread just like a normal wave, in a way like the black as discussed at the start of an article, given it is not too late.

2

The Bar chart below shows that the graph is very well connected, aside from a few nodes. It also shows that there is a exponential increase in the average total weight for higher degree nodes i.e higher node degree produces more flux through a node.
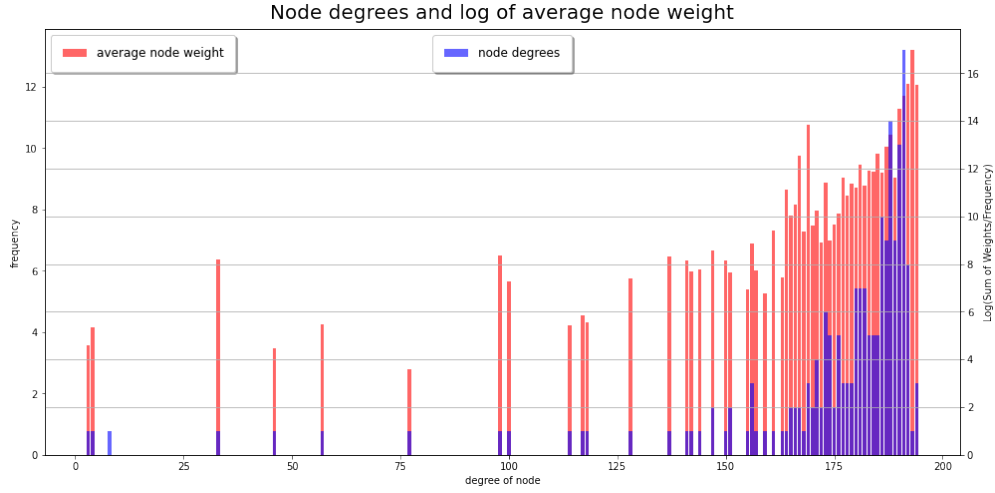

Figure 1: Graph shows node degrees and log of the average total weight for a nodes of that degree, the average weight for Andorra (degree=8), has been omitted as it skewed the axis (it value was approximately 0.006, giving a negative value of around 5)
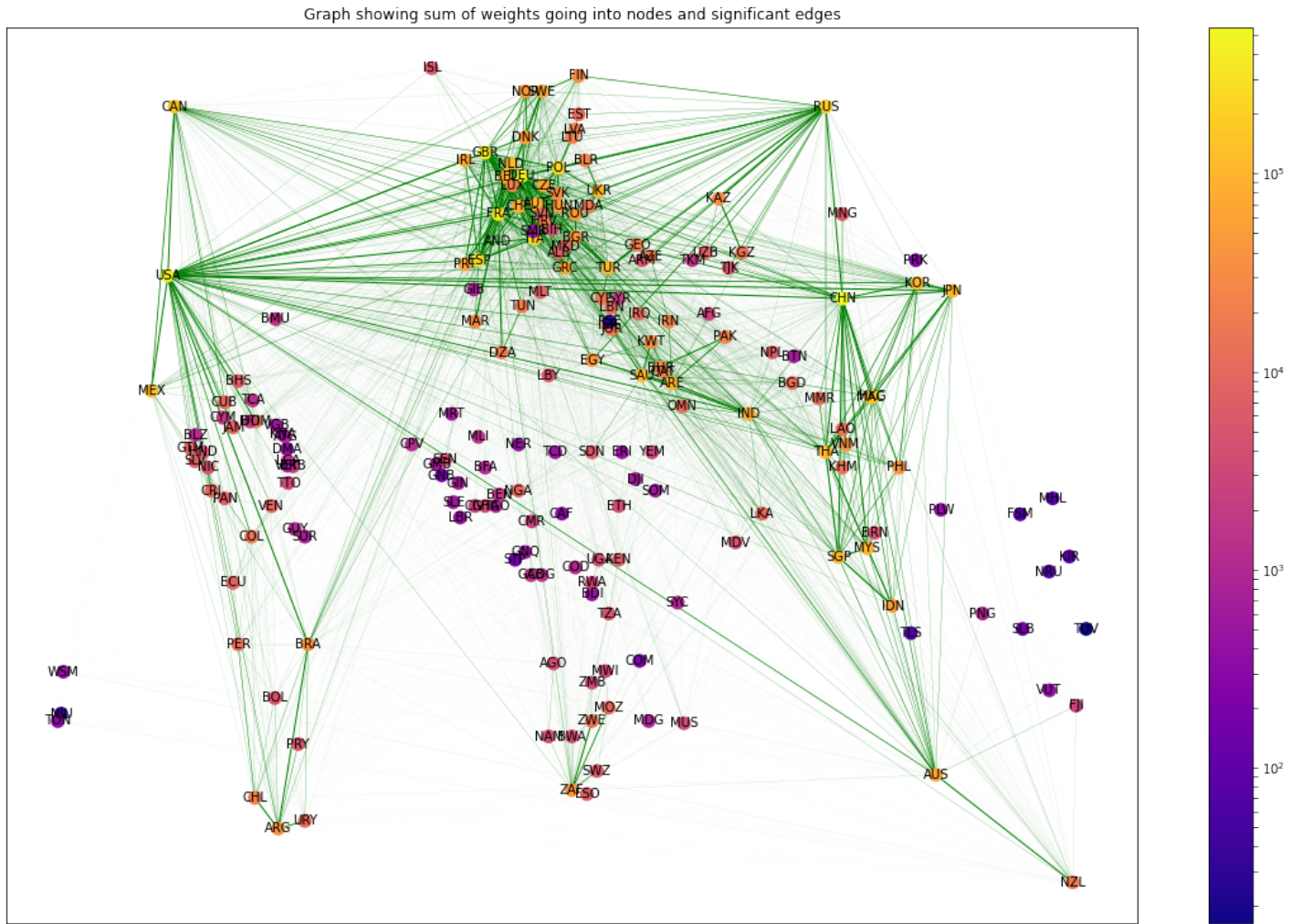
Figure 2: Graph total weight of a node with log scaled colour bar, again the weight for Andorra (degree=8), has been omitted as it skewed the axis (it value was approximately 0.006, giving a negative value of around 5), it shows significant edges with a larger flux representing larger width. The positions of the nodes are represented by the average Latitude and Longitude of a country.

Clearly there is high mobility in Europe,USA and China. It can be noted that the wealthier nations have higher mobility. This becomes even more clear when the total flux is divided by the degree of a node: I have not included this graph due to figure limits. Generally the small island nations have very low flux e.g Tonga.
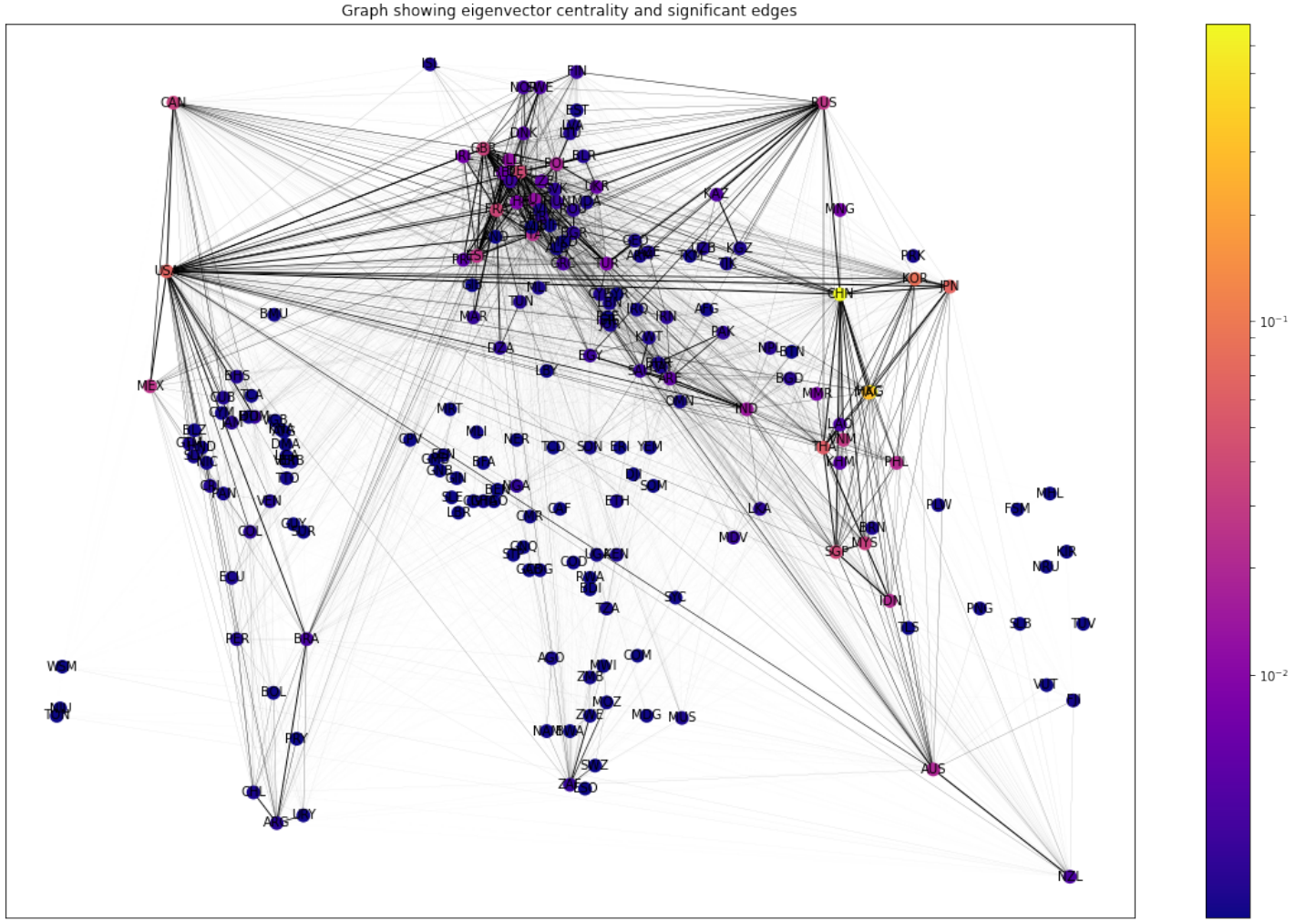
Figure 3: Graph showing Eigenvector centrality (calulated using weights of the graph), significant edges with a larger flux are represented with larger edge width. The positions of the nodes are represented by the average Latitude and Longitude of a country.

The Eigenvector centrality of a node is the proportional to its neighbors importance. This shows where virus will likely spread to and if they do would, then the best place to spread a virus. This suggests that the best place to start a epidemic would be China, which is interesting.

3

I have implemented the model (3) by creating a function RHS which calculates the right hand side of (3). This produces 2*N numpy array, representing the change in the coupled system given prevous values, for all nodes in G. This is used as input for scipys odeint, which calls RHS solving over a desired time frame and initial conditions.

The cost is mainly constructing RHS which gets called by odeint every time step and where improvements in efficiency can be made. I made RHS to be efficient, by precomputing as many variables as possible and vectorising, along with in general reducing the number of operations needed e.g in the sum of I only calculate $P_{mn}s_m$ as $P_{mn}s_n = s_n$. Everytime RHS is called it has roughly $(15 + \eta)N + 2N^2$ operations to perform.

The biggest cost generally of RHS is the matrix multiplication of $P_m n$, the overall cost is roughly $(15 + \eta)N + 2N^2$ which has complexity $O(N^2)$ if η is not greater than N.

4

I have investigated how the geographic and effective distance are connected to simulation results, in a similar way to the paper. I have also investigated how initial outbreak location effects simulation, along with parameter values. It would have be good to simulate an actual outbreak and compare results, seeing if outbreak location could be calculated using a snapshot of such data. But this is not possible, as it would require an extensive amount of work and more importantly data. In general we are interested in j_n to see how an infection spreads.

Achieving an epidemic spread of the model BH requires a fine balance of its parameters($x_0, \hat{j_0}, \alpha, \beta, \gamma =$ average mobility , ϵ = local invasion threshold, η = gain parameter), for example $\alpha, \beta, \hat{j_0}$

- β is rate of recovery, α/β is reproduction ratio. Increasing β decreases the length of epidemic, Increasing α/β also decreases the time length of epidemic and maximum concentration of infection. α and β have to be balanced so that infection spreads globally.

- increasing strength of the initial outbreak $\hat{j_0}$ generally decreases length of epidemic (it spreads faster). Not strong enough and it doesn't spread.
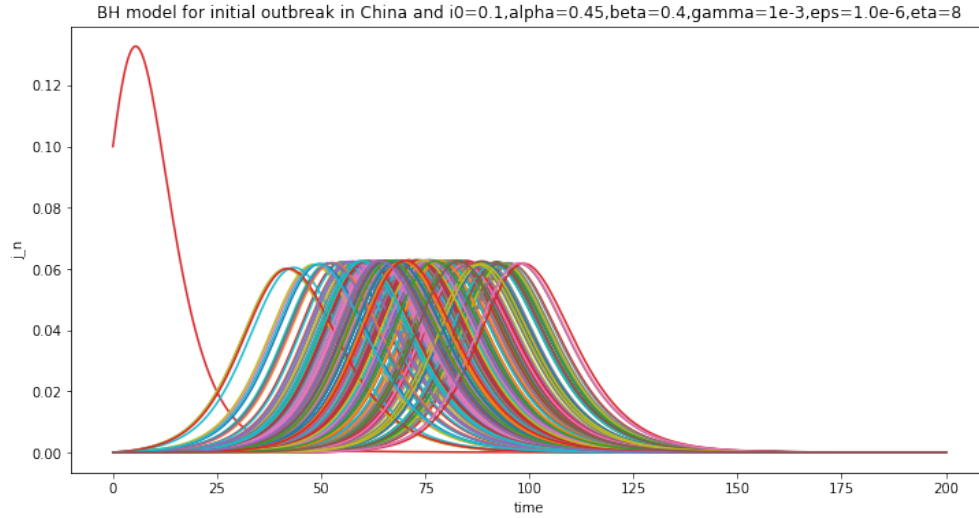

Figure 4: How local populations are infected with an outbreak originating in china

Figures 4 and 5 show the differences between an initial outbreak in China compared to Peru. Interestingly the outbreak is spread over a larger period of time for the outbreak in Peru. Which intuitively makes sense and agrees with the literature, as China is the central hub shown by figures in question 2, so it spread very quickly. Where the initial outbreak does effect results a fair amount, for example it was near impossible to simulate an epidemic starting from Andorra.

The results also show that in population spreading is the same irrelevant of outbreak location.
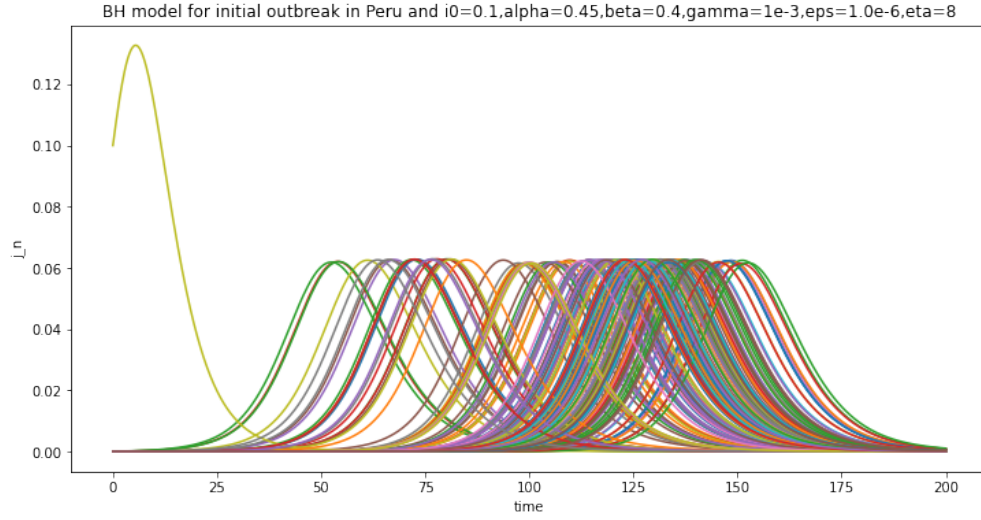
Figure 5: How local populations are infected with an outbreak originating in Peru

The figure below shows how the Euclidean distance of country's to initial outbreak country correlate with simulation results for China. The contour plot can be used to see when the infection first enters a node aswell its overall behavior. As can be seen there is a general correlation with distance.
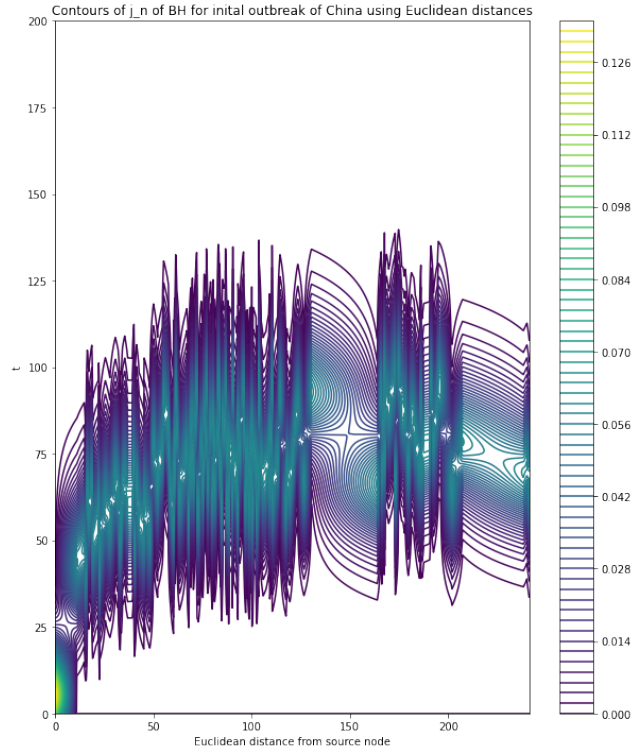

Figure 6: How infection spread with regards to Euclidean distance calculated from average Latitude and Longitude of a country, same parameters as figure 4.

Having calculated the effective distance in equation (5) using the Floyd Warshall Algorithm, I see how this correlated to simulation result instead of Euclidean distance.
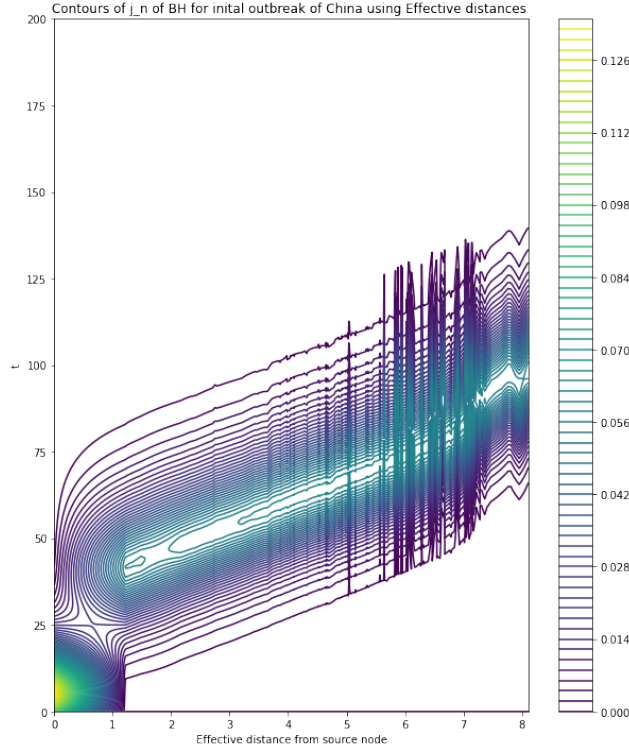

Figure 7: How infection spread with regards to Effective distance, same parameters as figure 4.

For the effective distance there is very clear correlation with not only the initial entry time of the infection to a node but also the way it spreads through a node. The jutts for larger effective distance in the figure 6 are caused by small unconnected nodes like Andorra. So the idea of BH of effective distance being correlated to outbreak, holds true for this simulation.

For consistency I compare results in the same way for an initial outbreak in Peru, shown on the next page. It produces similar results, but is not as clearly correlated for large effective distances (more than 8).

The main idea of the effective distance of BH seems to be apply to the simulations here, however it would need to be compared to data from a real outbreak to see if it truly agrees. Below I have provided some general comments and issues with BH model, with regards to this problem (country to country spreading):

- It assumes that local populations being proportional to transit values, which is not true generally.

- modeling by country is ineffective as it needs to be via populations centres of roughly same size on a scale where the diffusion model is appropriate. For example comparing Andorra and China on the same level is wrong.

- It might not be appropriate to model for a major epidemic like CoVid19 as global mobility is significantly lower than averages of past years, due to lock downs, so the data F would need to be recalculated.

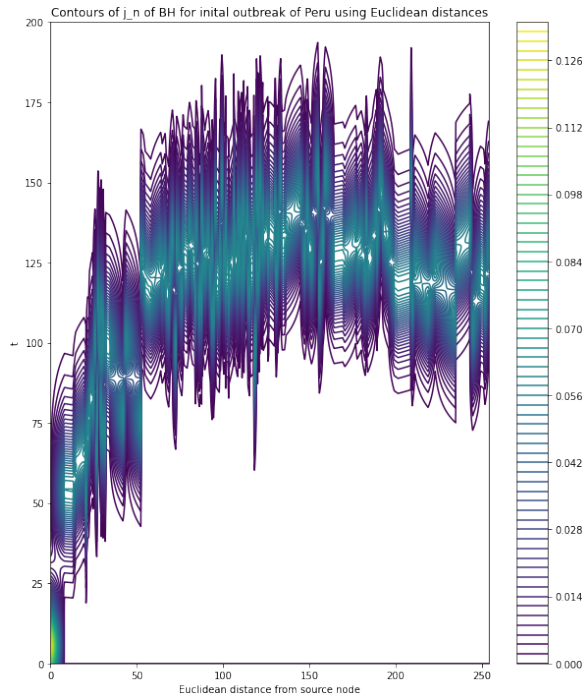- could maybe take the seasonal effects of mobility into account.

Figure 8: How infection spread with regards to Euclidean distance calculated from average Latitude and Longitude of a country, same parameters as figure 4.
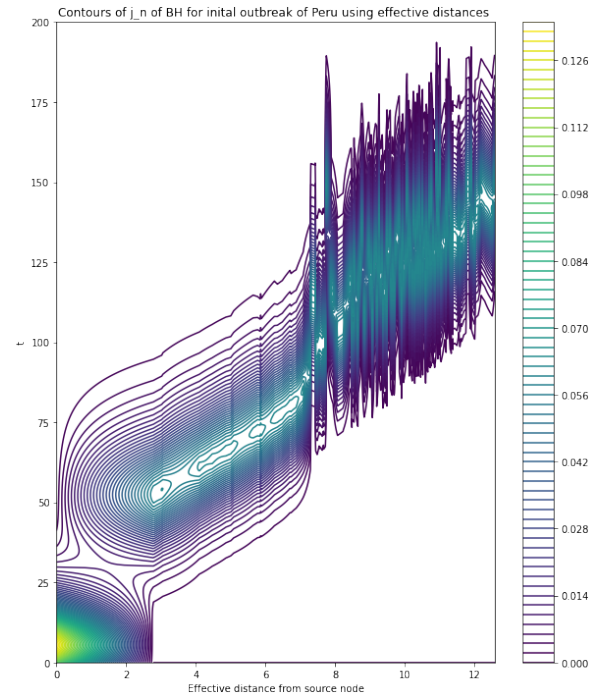

Figure 9: How infection spread with regards to Effective distance, same parameters as figure 5.