

For the stochastic compartmental model: for each country/region we have the following equations (we will write down the stochastic version which we solve below).

$$\frac{dS_i}{dt} = -\lambda_i S_i$$

$$\frac{dI_i}{dt} = \lambda_i S_i + \gamma_i I_i$$

$$\frac{dR_i}{dt} = \gamma_i I_i$$

where the force of infection  $\lambda_i$  is given by:

$$\lambda_i = \sum_j m_{ij} \frac{I_j}{N_i}$$

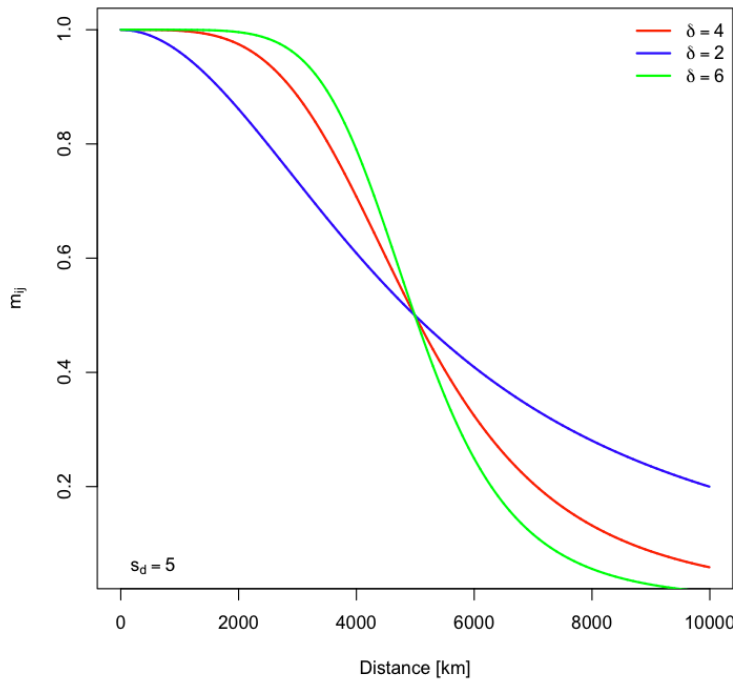
note that we have  $N_i$  in the denominator above which means that we have made an implicit assumption that transmission takes place in population  $i$ , presumably resulting from the movement of an infectious individual from population  $j$ . We could have made the other assumption: that transmission is due to individual from population  $i$  travelling to population  $j$  and picking up the infection during this travel time. This will require putting  $N_j$  in the denominator above.

The recovery rate,  $\gamma_i$ , can be different (or the same for all countries/regions).

The coefficients,  $m_{ij}$ , measure the strength of interaction between populations and we will assume that it has an off-set power-law dependence on the distance between the centroids of the populations ( $r_{ij}$ ):

$$m_{ij} = \frac{1}{1 + (s_d/r_{ij})^\delta}$$

The power determines the mixing between the populations and  $s_d$  is the distance (in km) below which the kernel function is saturated. We can decide not to have an off-set and we can start with  $\delta=-4$ . The matrix will be normalized by ensuring that the sum over rows equals 1. By looking at the median value of our distance matrix we can set the distance  $s_d$ . Below is a plot of  $m_{ij}$  with  $\delta=-2$  (blue), -4 (red) and -6 (green) and  $s_d=5$ km, seems that we may even want to set it to a smaller number (2km like in the Mills/Riley paper?)



We will be solving the stochastic version of the three equations for  $S_i$ ,  $I_i$  and  $R_i$

$$S_{i,t+1} = S_{i,t} - N_{i,S,t}^{inf}$$

$$I_{i,t+1} = I_{i,t} + N_{i,S,t}^{inf} - N_{i,I,t}^{rec}$$

$$R_{i,t+1} = R_{i,t} + N_{i,t}^{rec}$$

where we will calculate the probability of population  $i$  to be infected as:

$$p_i^{inf} = 1 - \exp(-\lambda_i \cdot \Delta t)$$

(see above for the definition of  $\lambda_i$ ), and the probability of this population to recover as:

$$p_i^{rec} = 1 - \exp(-\Delta t \cdot \gamma_i)$$

with  $\Delta t$  being the time step. Using these two probabilities we can calculate the number of infected/recovered for each country as:

$$N_{i,S,t}^{inf} = rbinom(1, S_i, p_i^{inf})$$

$$N_{i,I,t}^{rec} = rbinom(1, I_i, p_i^{rec})$$

I am using the term 'country' in a loose way here. It can mean country/state/region/country etc.