Quantifying the effect of CCS on the persistence of infectious diseases in a metapopulation

Tran Thi Cam Giang, Marc Choisy, Jean-Daniel Zucker, Yann Chevaleyre

03/5/2016

Abstract

CCS =Critical community size

The Critical Community Size (CCS) is defined as the smallest population size that does not suffer disease extinction. For measles the CCS has been found to be remarkably consistent, estimated at between 300,000 and 500,000.

1 INTRODUCTION

.....

2 METHODS

2.1 Metapopulation model

Consider a metapopulation of n sub-populations. In a subpopulation i of size N_i, disease dynamics can be deterministically described by the following set of differential equations [2]. The formulation of the metapopulation model with environmental and movement-based transmission is given by the deterministic system of equations as follows :

$$\frac{dS_i}{dt} = \mu N_i - \lambda_i S_i - \mu S_i \tag{2.1}$$

$$\frac{dE_i}{dt} = \lambda_i S_i - \mu E_i - \sigma E_i \tag{2.2}$$

$$\frac{dI_i}{dt} = \sigma E_i - \mu I_i - \gamma I_i \tag{2.3}$$

$$\frac{dR_i}{dt} = \gamma I_i - \mu R_i \tag{2.4}$$

where S_i, E_i, I_i et R_i are the numbers of susceptible, exposed, infectious and recovered in this sub-population i respectively. Individuals are born susceptible, die at a rate μ, become infected with the force of infection λ_i, infectious after a latency period of an average duration of $1/\sigma$ and recover at the rate γ. The force of infection depends not only on the total population size N_i and the number of infected I_i in subpopulation i, but also in other sub-populations [12] :

$$\lambda_i(t) = \sum_j \rho_{ij}\kappa_j \log\left[1 - \sum_{k=1}^{M}\left(\frac{|I_k(t)|}{N_k(t)} \times c_{ik} \times \xi_{jk}\right)\right] \tag{2.5}$$

where $c_{i,k}$ $(0 \leqslant c_{ij} \leqslant 1)$ is the probability that a susceptible individual native from i being in contact with another infected individual native from k gets infected. ξ_{jk} $(0 \leqslant \xi_{ij} \leqslant 1)$ refers to the probability that an individual y meeting x in the subpopulation C_j comes from the subpopulation C_k. κ_j is the average number of contacts per unit of time a susceptible will have when visiting city j. $\rho_{i,j}$ $(0 \leqslant \rho_{ij} \leqslant 1)$ is denoted as the probability that an individual from subpopulation i visits subpopulation j, of course, $\sum_{j=1}^{M} \rho_{ij} = 1$. See appendix for detail on the construction

of this equation. We can verify that in the limit case on one single subpopulation in the metapopulation ($i = j$ and $n = 1$) we have

$$\lambda_i = -\kappa_i \log(1 - \frac{I_i}{N_i} \times c_{ii}) \tag{2.6}$$

Describing the strength of connection ρ in a metapopulation (subpopulations connected by individual movement) is quite complex. In this work, we will study the null model (model 0), which is a metapopulation without any explicit spatial distance (all the subpopulations are at the same distance from each other) and where all the metapopulation have the same population size N. Like the original Levins's model [13], this model considers that all the subpopulations are at equal distance from each other:

$$\rho_{ij} = \rho, \qquad 0 \leqslant \rho \leqslant 1, \qquad \forall i, \forall j. \tag{2.7}$$

Hence, we will focus on studing three key parameters that characterize the metapopulation: (i) the number n of sub-populations, (ii) the population size N ($N_i = N$, $\forall i$) of all these subpopulations and, (iii) the coupling (or distance) ρ_{ij} between two subpopulations i and j that refers to the the probability that an individual from subpopulation i visits subpopulation j.

2.2 Environmental forcing

The average number of contacts per unit of time κ_i is seasonally forced [1] and seasonality is an annually periodic function of time [8]. As a result, for the subpopulation i :

$$\kappa_i(t) = \kappa_{i0} \left[1 + \kappa_{i1} \cos\left(\frac{2\pi t}{T} + \varphi_i\right)\right] \tag{2.8}$$

where t is the time, κ_{i0} and κ_{i1} are the mean value and amplitude of the average contact rate κ_i at which a susceptible will have when visiting city i per unit of time, T and φ_i are the period and the phase of the forcing. With the annual sinusoidal form of the average contact rate, we really have the sinusoidally forced SEIR metapopulation model.

In our metapopulation model, we define a new parameter $\varphi_{\max}$, which typifies the synchrony of the metapopulation. A given synchrony value $\varphi_{\max}$ in radian that is in the interval from zero to π, we divide the interval [0, φ_{max}] into a set of $(n-1)$ equal samples for the metapopulation of n the number of subpopulations. Hence, the value of the forcing phase of the i^{th} subpopulation is correspondent to i^{th} value in the set.

2.3 Measuring persistence (local and global)

The Stochastic Simulation Algorithms (SSA) uses Monte Carlo (MC) methods to study the evolution process of disease in continuous time by solving the corresponding stochastic differential equations. In our work, we focus on studying demographic and environmental stochasticities in epidemic models. Demographic stochasticity is considered as fluctuation in population processes that are based the random nature of events at the level of the individual. Each event is related to one baseline probability fixed, individuals are presented in differing fates due to chance. Besides, the number of infectious, susceptible, exposed and recovered individuals is now required to be an integer. Modeling approaches that incorporate demographic stochasticity are called event-driven methods. These methods require explicit consideration of events. The first approach published by Daniel T.Gillespie in 1976 [6] is an exact stochastic simulation approach for chemical kinetics. The Gillespie stochastic simulation algorithm (SSA) has become the standard procedure of the discrete-event modelling by taking proper value of the available randomness in such a system. The methods modelling the event-driven model demands explicit presentation of events. For the standard SEIR model, we have to consider the nine events that can occur, each causing the numbers in the relative groups to go up or down by one. Table 2.3 lists all the events of the model, occurring in subpopulation i of a metapopulation:

To initialize the variables S, E, I, R for subpopulations in a metapopulation, we fixe the metapopulation size N. We compute the equilibrium values of these variables by using the equilibrium equations of the system (in appendix). We simulate the fluctuation of a single population with the size N and the simulation time of 100 years, to obtain stationary regime for disease dynamic. At the 100^{th} point, we get the stationary values of variables. Now, with n subpopulations in the metapopulation, the initial values of variables for each subpopulation are identical and equal to the stationary values divided by n. Thereby, we build successfully the stationary metapopulation just at the initial step of simulation.

Previous analyses have demonstrated that there are recurrent waves of extinction and re-colonization in spatial dynamics within subpopulations [7]. Estimating the ability of the disease persistence is still a complex problem and

Table 1: Events of the stochastic version of the model of equations 2.1-2.4, occuring in subpopulation i.

Events	Rates	Transitions
birth	μN_i	$S_i \leftarrow S_i + 1$ and $N_i \leftarrow N_i + 1$
death of a susceptible	μS_i	$S_i \leftarrow S_i - 1$
death of an exposed	μE_i	$E_i \leftarrow E_i - 1$
death of an infected	μI_i	$I_i \leftarrow I_i - 1$
death of an immune	μR_i	$R_i \leftarrow R_i - 1$
infection	$\lambda_i S_i$	$S_i \leftarrow S_i - 1$ and $E_i \leftarrow E_i + 1$
becoming infectious	σE_i	$E_i \leftarrow E_i - 1$ and $I_i \leftarrow I_i + 1$
recovery	γI_i	$I_i \leftarrow I_i - 1$ and $R_i \leftarrow R_i + 1$

there is no exact formule to calculate this ability. Such, estimating global persistence as well as local, recolonization, and the link among local persistence, local extinction, recolonization is the focus of this paper. We will reveal methods how we can measure the rate of the local/global extinction and of the recolonization.

First, in order to measure global extinction rate, we simulate a metapopulation of n subpopulations. We gain n independent fluctuations of our stochastic model. Then, we calculate the average metapopulation size by summing subpopulations at each sample time and averaging across the entire time series for each metapopulation. Lastly, we record the dates t of global disease extinction in all these m metapopulations. These dates allow to draw Kaplan-Meier survival curves from which we estimate the global extinction rates χ:

$$M(t) = \exp(-\chi t) \tag{2.9}$$

where $M(t)$ $(0 \leqslant M(t) \leqslant m)$ is the number of metapopulations in which the disease is not extinct at time t.

After that, we use the parametric survival model for the exponential distribution (R package '*survival*' [14]). Due to that, we can capture one of the most important features of stochastic systems in spatial structure : its global extinction characteristics of disease.

Besides, to compute local extinction rate that illustrates the probability of local extinction event in the duration of fluctuations in incidence. We save up all disease duration in all subpopulations of the metapopulation. We consider the data array like a Poisson process and we estimate local extinction rate from this data.

Finally, we also save up durations where there is no disease in all subpopulations of the metapopulation. We have a data set like a Poisson process and we assess recolonization rate for the metapopulation.

2.4 Plan of experiment

In this section, we will describe our plan of experience to quantifying the effect of synchrony on the persistence of infectious diseases in a metapopulation. We have three big concerns that we must verify.

- (1) relations between rates: local extinction rate, global extinction rate and recolonization rate.
- (2) effects of the population parameters (such as the number of subpopulations n and the coupling rateρ) and of the environmental parameters (such as the amplitude of the average contact rate$\kappa_1$and the synchrony parameter φ_{max}).
  - the number of subpopulations n : this parameter plays an important role in metapopulation. For a given metapopulation size, this number is inversely scaled the subpopulation size.
  - The coupling rate ρ: it illustrates the strength of connextion among subpopulations in a metapopulation. When $\rho = 0$, the subpopulations are totally independent. But with $\rho = 1$, the subpopulations are an unified population. The metapopulation becomes a big population.
  - The amplitude of the average contact rateκ_1: this parameter introduces the influence of the season on the disease persistence of the metapopulation. When $\kappa_1 = 0$, the metapopulation is in static state, any environnemental factor affects dynamics of metapopulation.
  - The synchrony parameter φ_{max}: φ_{max} is an important parameter that we use to break fixed points as well as first fixed points at begining moments between subpopulations. The effect of increasing φ_{max} drives the synchronization between subpopulations down.

2.4.1 Stochastic metapopulation simulations

In order to run simulations, we use the same values of all parameters for all subpopulations. We use the Gillespie's direct algorithm [6] for metapopulation model as described in the previous part. With the SEIR metapopulation model, measles is modeled [2, 9]. Moreover, in this work, we use also the values of parameters for the measles to do experiences. We have a table of the convenient values for parameters of measles as follows :

Following the table in detail about the convenient values of parameters, we will use them throughout all simulations. We start doing a simulation from a initial random number. Then, we aggregate the daily data (number of individuals in the suceptible, exposed, infected and recovered groups) into one-day intervals, and use this as the time step in the model.

Table 2: Some Disease Parameter Values for Measles from the Literature [3, 4, 5, 10, 11, 12]

parameter	**description**	**value**	**unit**
μ	birth and death rate per day	$1/(70*365)$	1/(people*day)
κ_0	mean value of the number of contacts κ per unit of time a susceptible will have when visiting one city	$[20, 150]$	people/day
κ_1	amplitude of the number of contacts κ per unit of time	$[0.01, 0.1]$	
γ	recovery rate per day	$1/8$	1/(people*day)
σ	average exposed duration per day	$1/5$	1/day
ρ	coupling rate (ρ_{ij} the probability that an individual from subpopulation i visits subpopulation j)	$[0, 1]$	
φ_{max}	synchrony parameter in radian	$[0, \pi]$	radian
N	population size of subpopulation	$5000 - 1000,000$	people
n	number of subpopulation	$[1, 30]$	number of subpopulation
t_{max}	simulation time	50	day

3 RESULT

3.1 Plotting extinction rates against population size

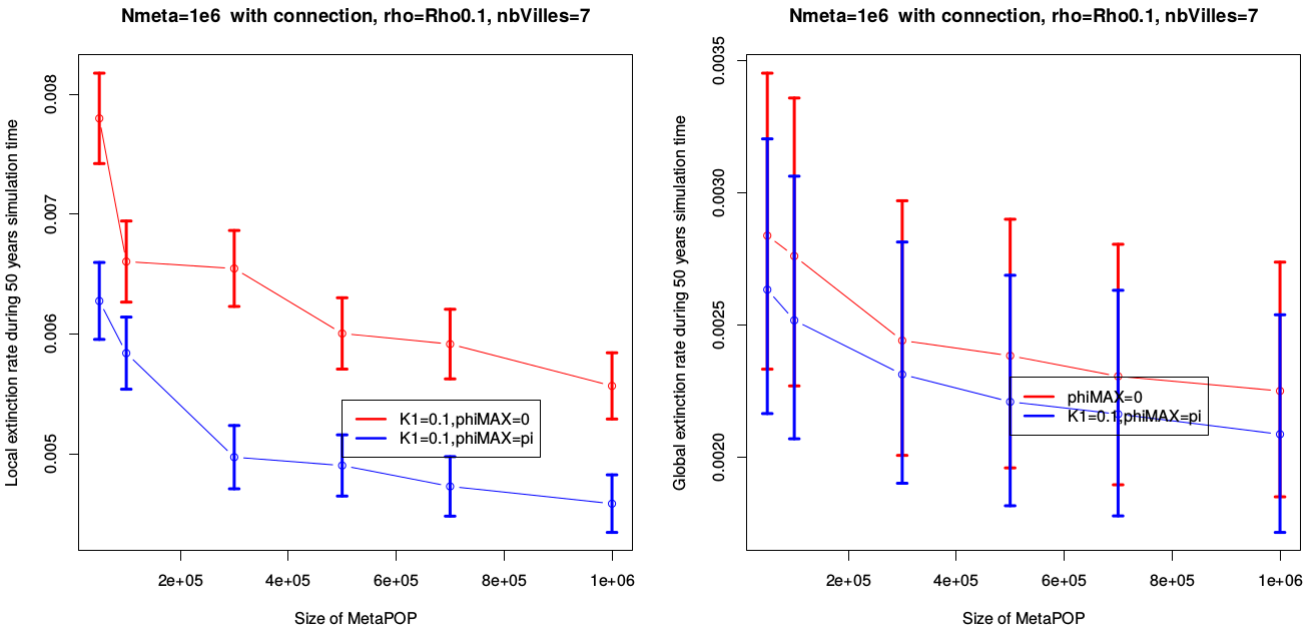

Figure 3.1: Influence of the size of metapopulation on the local extinction rate in the stochastic seasonally forced and spatially structured model. The number of subpopulations n is fixed 7, the coupling rate $\rho = 0.1$, the metapopulation size N from $5e+04$ to $1e+06$, the amplitude of season $\kappa_1 = 0.1$ and $\varphi_{\max}$ is 0 or π. The results point out that the extinction rates of the synchrony is higher than those of the asynchrony for all sizes of metapopulation. Besides, increasing the size of metapopulation leads the decrease of the extinction rates.

In the stochastic seasonally forced and spatially structured model, to obtain the synchrony in the metapopulation, we set $\varphi_{\max} = 0$. The subpopulations are grouped into pairs based on their interactions, and particularly their disease phases are in-phase. All the subpopulations are in synchrony. So the ability of rescue effect among subpopulations are small, they simply get extinct. Consequently, the local extinction rate when $\varphi_{\max} = 0$, is higher than when $\varphi_{\max}$ is different from zero. On the other hand, the asynchrony is presented by $\varphi_{\max} \neq 0$. The phase variation in disease dynamics of the coupled subpopulations are out of phase, the synchrony is interrupted. With $\varphi_{\max} = \pi$, the decreased similarity between sub-populations leads an increase of the chance of recolonization among sub-populations. A sub-population has even reached the local extinction, but the disease comes easily back due to the recolonization among sub-populations. Hence, the rates of local extinction as well as of global extinction in the case $\varphi_{\max} \neq 0$ is lower. Briefly, the extinction rates of the synchrony is better than that of the asynchrony. Besides, the increased size of metapopulation draws the increase of the size of subpopulation. Thus, the figure 3.1 reveals that the extinction rates reduces with increasing population size.

To enrich results about extinction rates against population size, we investigated this relation in the stochastic seasonally forced and isolated model. In this model, the subpopulation are independant, uncoupled. There is no immigration among subpopulations. It's the reason for that the rate of recolonization is always equal to zero. Then, we set the model in seasonal variation. We let the model in synchrony with $\varphi_{\max} = 0$ and in asynchrony with $\varphi_{\max} = \pi$. We obtained the same results with the seasonally forced and coupled model. The extinction rates of the synchrony is greater than those of the asynchrony. Then, the increased metapopulation size leads to a decrease of the extinction rates.

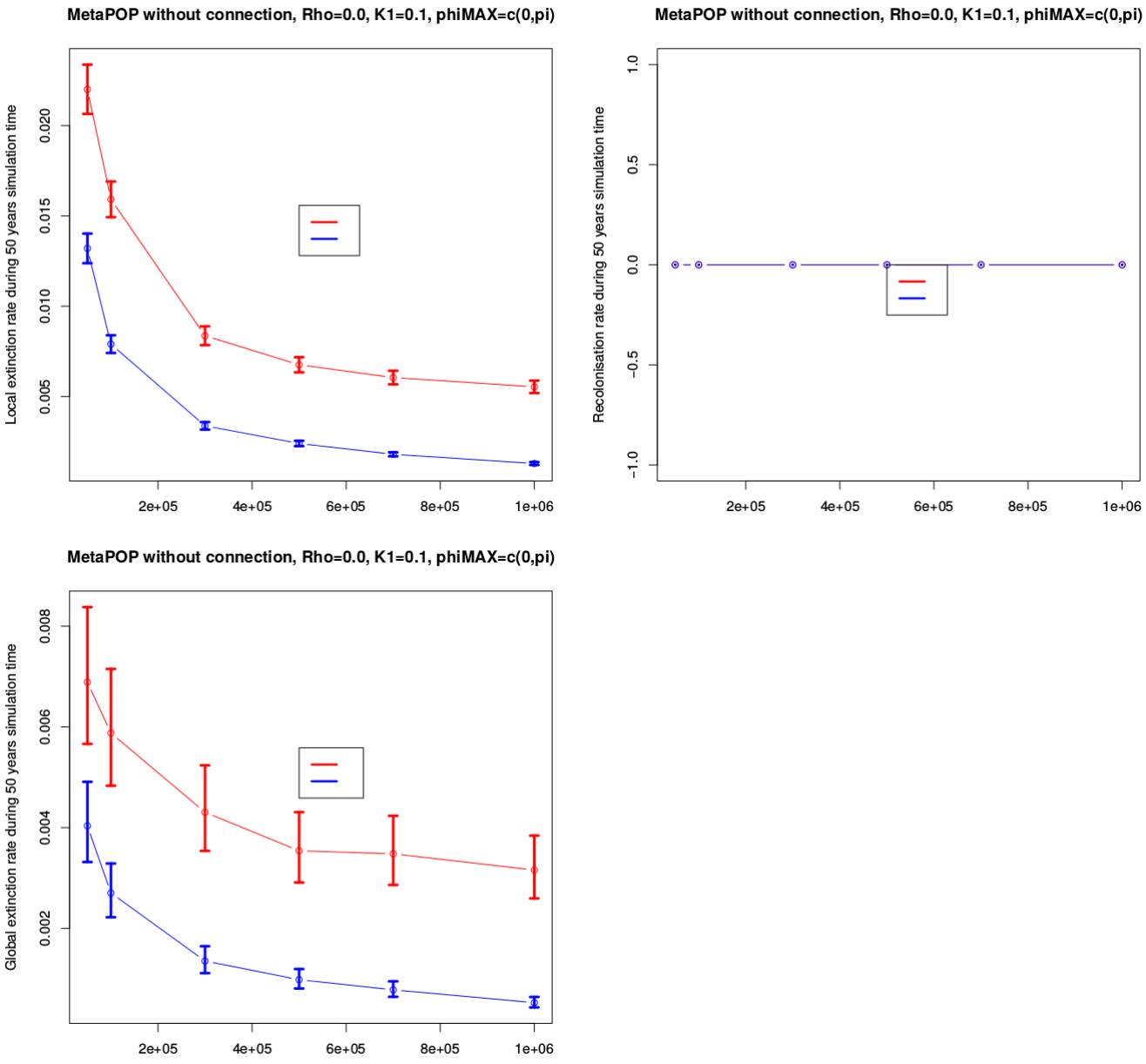

Figure 3.2: Plotting extinction rates against population size in the stochastic seasonally forced and isolated model. The number of subpopulations n is fixed 10, the coupling rate $\rho = 0.0$, the metapopulation size N from $5e+04$ to $1e+06$, the amplitude of season $\kappa_1 = 0.1$ and $\varphi_{\max}$ is 0 or π. The x-axis is the metapopulation size. The red line is for the forced model of uncoupled subpopulations with $\varphi_{\max} = 0$ and the blue line is for the same model with $\varphi_{\max} = \pi$.

3.2 Extinction rates in the coupled model and the uncoupled model

In order to study the gap of extinction rates in the coupled model as well as the uncoupled model, we plotted the extinction rates against the metapopulation size with the subpopulation number ($n = 3$).

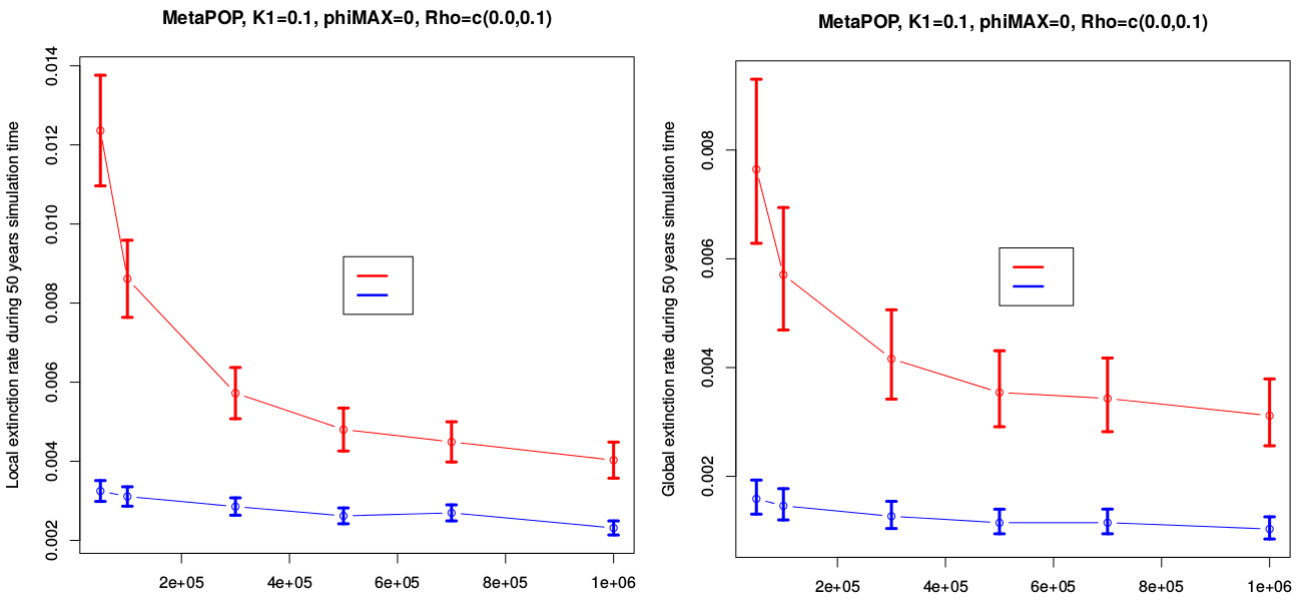

Figure 3.3: Plotting extinction rates against population size in the coupled model and the uncoupled model. The number of subpopulations n is fixed 3, the coupling rate for the uncoupled model $\rho = 0.0$, and for the coupled model $\rho = 0.1$, the metapopulation size N from $5e+04$ to $1e+06$, the amplitude of season $\kappa_1 = 0.1$ and $\varphi_{\max}$ is 0. The x-axis is the metapopulation size. The red line is for the forced model of uncoupled subpopulations with $\varphi_{\max} = 0$ and the blue line is for the forced model of uncoupled subpopulations with $\varphi_{\max} = \pi$.

The figure 3.3 reveals that the extinction rates of the isolated model is higher than those of the coupled model. Because, in the isolated model, the subpopulations are independent, no exchange, no immigration. Then, there is no occurred recolinisation in this model. Inversly for the coupled model, due to the connection among subpopulations in the metapopulation, there are imigrations among them, then the disease is again saved in subpopulations that have gone locally extinct. Besides, increasing the metapopulation size draws an increase of the subpopulation size. Hence, the extinction rates have the tendency to reduce when the metapopulation size rises.

3.3 Plotting the metapopulation size and the coupling rate against the estimated extinction rate

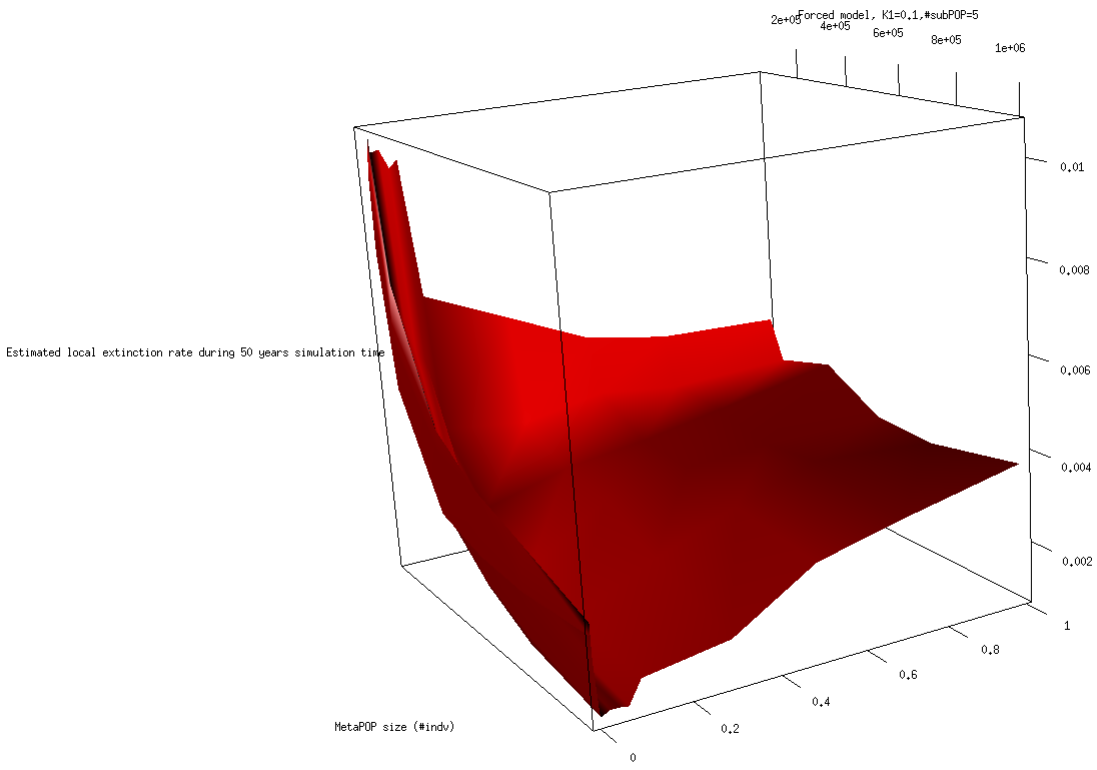

Figure 3.4: Three-dimensional relation among the metapopulation size, the estimated local extinction rate and the coupling strength. The x-axis is the metapopulation size, the y-axis is the coupling rate, and the z-axis is the estimated extinction rate. The three-dimensional result shows that the line plotting the coupling rate against the extinction rate is as a convex curve and obtains minimum when the coupling rate is medium. Moreover, increasing the metapopulation size drives an increase of the subpopulation size and a decrease of the extinction rate.

Figure 3.4 pointed out that the line plotting the coupling rate against the local extinction rate estimated is as a convex curve and obtains minimum when the coupling rate is medium.

Here, the coupling rate or the dispersal rate ρ can be considered as migration strength. The disease transmission speed grows fast when coupling rate goes up in metapopulations. Similar to that, global disease persistence surges also. In this part, we permit coupling rate change from weak to strong in a metapopulation of five subpopulations. The dispersal rate ρ is divided into three intervals. These are low, intermediate and high coupling rate intervals. In each interval, we chose some coupling rates that highlight the coupling strength among subpopulations in a metapopulation. With each value of coupling rate, we estimated local extinction rate that presents the extinction probability of disease in a subpopulation. When the coupling rate is small from 0.0 to 0.001, the locale extinction rate significantly decreases. However, this rate is minimum when the coupling rate has medium values from 0.001 to 0.01. Lastly, the extinction rate augments back when the coupling rate is very strong from 0.01 to 1.0. As shown in the Fig 3.4, the local extinction rate in a metapopulation is one humped function for the coupling rate. The medium coupling rate (from 0.01 to 0.1) minimises the extinction rate of disease in metapopulation. Because in the case of the small and medium coupling rates, the coupling rate and the speed of migration among subpopulations are directly proportional. The dispersal speed increases. Thereby the local recolonization speed rises, the duration of persistence grows, the local extinction rate goes down. However, this trend of local extinction with decreasing coupling rate, is not right any more when the dispersal rate is strong. The metapopulation has tendency to become one big population. In this case, the phase difference or the recolonization among subpopulations are no longer significant. Hence, the local extinction rate goes up. Besides, increasing the metapopulation size draws an increase of the subpopulation size, then the local extinction rate goes significantly down.

In order to enrich our results, here we present a three-dimensional relation among the metapopulation size, the number of infected persons and the coupling strength (Fig 3.5). Differ from the result in Fig 3.4, the line of the number of infected looks like a humpbacked curve. This number is maximum when the coupling strength is medium. Besides, the number is directly scaled to the metapopulation size.

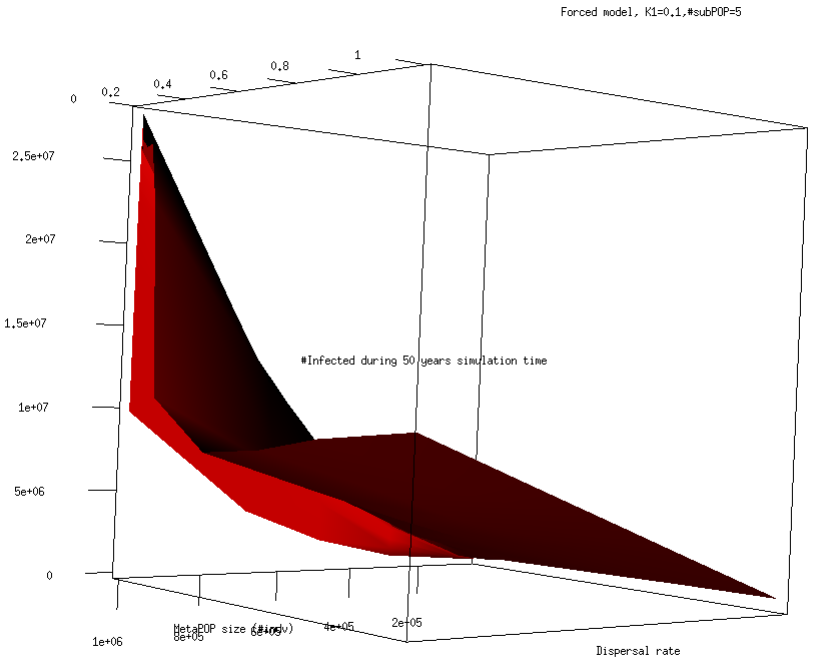

Figure 3.5: Three-dimensional relation among the metapopulation size, the number of infected persons and the coupling strength. The x-axis is the metapopulation size, the y-axis is the coupling rate, and the z-axis is the number of infected persons . The three-dimensional result shows that the line plotting the coupling rate against the number of infected is as a humpbacked curve and obtains maximum when the coupling rate is medium. Moreover, increasing the metapopulation size draws an increase of the subpopulation size as well as of the number of infected.

3.4 Plotting the metapopulation size and the number of subPOP against the estimated extinction rate

An other result, we describe a three-dimensional relation among the metapopulation size, the number of subpopulations and the local extinction rate estimated. The figure 3.4 shows that the local extinction rate decreases with increasing population size. Inversely, the extinction rate increases with increasing number of subpopulation.

It is obvious that with the fixed subpopulation number, we increase the metapopulation size, so we increase the subpopulation size. Hence, the time of disease persistence augments, so the local extinction rate goes down.

Inversely, with the fixed metapopulation size, we increase the number of subpopulations, it means that we decrease the subpopulation size. Hence, the persistence time significantly declines, so the extinction rate goes up.

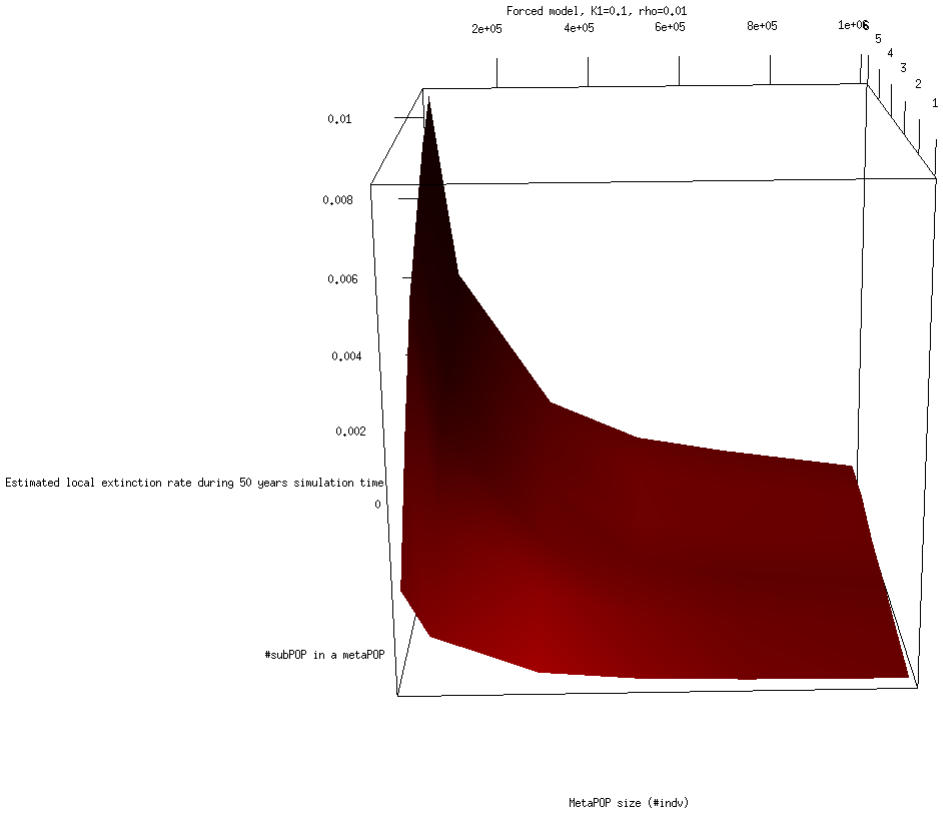

Figure 3.6: Three-dimensional relation among the metapopulation size, the number of subpopulations and the local extinction rate estimated. The x-axis is the metapopulation size, the y-axis is the number of subpopulation, and the z-axis is the local extinction rate estimated. The result is pointed out in the forced model with $\kappa_1 = 0.1$ and $\rho = 0.1$.

4 DISCUSSION

....

References

[1] S. Altizer, A. Dobson, P. Hosseini, P. Hudson, M. Pascual, and P. Rohani. Seasonality and the dynamics of infectious diseases. *Ecol Lett*, 9(4):467–484, Apr 2006.

[2] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992.

[3] B. Bolker and B. Grenfell. Space, persistence and dynamics of measles epidemics. *The Royal Society*, 348:309–320, 1995.

[4] M. Choisy and P. Rohani. Changing spatial epidemiology of pertussis in continental usa. *Proc Biol Sci*, 279(1747):4574–4581, Nov 2012.

[5] A. J. K. Conlan, P. Rohani, A. L. Lloyd, M. Keeling, and B. T. Grenfell. Resolving the impact of waiting time distributions on the persistence of measles. *J R Soc Interface*, 7(45):623–640, Apr 2010.

[6] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[7] B. T. Grenfell, O. N. Bjørnstad, and J. Kappey. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414(6865):716–723, Dec 2001.

[8] B.T. Grenfell, B. M. Bolker, and A. Klegzkowski. Seasonality and extinction in chaotic metapopulation. *The royal society*, 259:97–103, 1995.

[9] C. E. Gunning and H. J. Wearing. Probabilistic measures of persistence and extinction in measles (meta)populations. *Ecol Lett*, 16(8):985–994, Aug 2013.

[10] M. J. Keeling and B. T. Grenfell. Disease extinction and community size: modeling the persistence of measles. *Science*, 275(5296):65–67, Jan 1997.

[11] M. J. Keeling and B. T. Grenfell. Understanding the persistence of measles: reconciling theory, simulation and observation. *Proc Biol Sci*, 269(1489):335–343, Feb 2002.

[12] M. J. Keeling and P. Rohani. *Modeling Infectious Diseases in humans and animals*. Princeton University Press, 2008.

[13] R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15:237–240, 1969.

[14] T. M. Therneau. *A Package for Survival Analysis in S*, 2014. R package version 2.37-7.

5 Appendix : equilibrium values of the system 2.1–2.4

We start with ordinary differential equations for a $subpopulation_i$ in a metapopulation as follows:

$$\frac{dS_i}{dt} = \mu N_i - \lambda_i S_i - \mu S_i \tag{5.1}$$
$$\frac{dE_i}{dt} = \lambda_i S_i - \mu E_i - \sigma E_i \tag{5.2}$$
$$\frac{dI_i}{dt} = \sigma E_i - \mu I_i - \gamma I_i \tag{5.3}$$
$$\frac{dR_i}{dt} = \gamma I_i - \mu R_i \tag{5.4}$$

In simulation, we know that the equilibrium state allow a disease to persist in a population for a long time. So, an infectious disease in the $subpopulation_i$ is available in long term this system is at equilibrium. It means that at which $\frac{dS_i}{dt} = \frac{dE_i}{dt} = \frac{dI_i}{dt} = \frac{dR_i}{dt} = 0$ (*). Thus, we let all equations (equations 15 - 18) in the system be equal to zero, then calculate the values of the variables (now denoted by S_i^*, E_i^*, I_i^* , and R_i^*) that satisfy this condition (*). We have these values as folllows:

$$S_i^* = N_i \frac{(\gamma+\mu)(\sigma+\mu)}{\beta\sigma} \tag{5.5}$$
$$E_i^* = N_i\mu\left(\frac{1}{\sigma+\mu} - \frac{\gamma+\mu}{\beta\sigma}\right) \tag{5.6}$$
$$I_i^* = N_i\mu\frac{\beta\sigma - (\sigma+\mu)(\gamma+\mu)}{\beta(\sigma+\mu)(\gamma+\mu)} \tag{5.7}$$
$$R_i^* = N_i - S_i^* - E_i^* - I_i^* \tag{5.8}$$

Here, if we set $R_0 = \frac{\beta\sigma}{(\gamma+\mu)(\sigma+\mu)}$, so we have

$$S_i^* = N_i\frac{1}{R_0} \tag{5.9}$$
$$E_i^* = N_i\frac{\mu\sigma}{R_0}(R_0 - 1) \tag{5.10}$$
$$I_i^* = N_i\frac{\mu}{\beta}(R_0 - 1) \tag{5.11}$$
$$R_i^* = N_i - S_i^* - E_i^* - I_i^* \tag{5.12}$$

One nomal conditions for all population availabes is that the equilibrium values cannot be negative. Therefore, an infectious disease is available in the *subpopulation*$_i$ if $R_0 > 1$. Now, the endemic equilibrium in the system is given by $(S_i^*, E_i^*, I_i^*, R_i^*) = (N_i \frac{1}{R_0}, N_i \frac{\mu\sigma}{R_0}(R_0 - 1), N_i \frac{\mu}{\beta}(R_0 - 1), N_i(1 - \frac{1}{R_0} - \frac{\mu\sigma}{R_0}(R_0 - 1) - \frac{\mu}{\beta}(R_0 - 1))$.

5.1 Stationary distribution in metapopulation

Here we show some assumptions for the stationary distribution model as follows :

- Assumption 1. For each city V_i, there exists a markov chain M_i describing where (i.e. in which city) individuals native from V_i travel at each time step.
- Assumption 2. Each M_i has a stationary distribution $\rho(M_i)$.
- Assumption 3. At time t=0, each agent is located in a city randomly drawn from $\rho(M_i)$.

When we consider a simplified model in which the dynamics of the agents is stationary: each agent native from V_i no more follows a markov chain, but is relocated at each time step on a city randomly drawn from $\rho(M_i)$.

Then, under assumptions 1,2,3, at any time t, when the total number of agents grows to infinity, the size of the populations under the markovian dynamics converges towards the size of the populations under stationary dynamics.

Hence, any statistics computed on the densities of agents from the same population in various cities will not distinguish the markovian from the stationary dynamics.

Based on this conclusion, we will deploy a stationary distribution in a metapopulation. First of all, we choose a population size N for the metapopulation. Then, we compute the After that, the transition matrix converge towards a stationary distribution matrix. Finally, we apply the stationary distribution matrix in the metapopulation of n subpopulations

6 Appendix: derivation of the equation 2.5

Here, we will point out that the contact rate β is a function of the average contact number per unit of time and the probability of successful disease transmission following a contact.

Definition 1. During the small time interval δt, each individual native of the city i visits one single city j (with the probability ρ_{ij}) and will see in average κ_j individuals. These individuals come from all the cities.

6.1 Notation :

Here, we present list of sets and events describing the state of the system at time t :

- C_i is the set of all individuals born in subpopulation i.
- $V_{i,t}$ is the set of all individuals physically located in subpopulation i from time t to time $t + \delta t$. This includes foreigners traveling in subpopulation i at time t, and all natives from subpopulation i which are not traveling abroad at time t.
- S_t, E_t, I_t, R_t are the sets of all individuals respectively susceptible, exposed, infected and recovered at time t. Note that these set include individuals from all subpopulations.
- $S_{i,t}, E_{i,t}, I_{i,t}, R_{i,t}$ are the same sets, restricted to natives of subpopulation i. So formally, $S_{i,t} = S_t \cap C_i$, $E_{i,t} = E_t \cap C_i$, $I_{i,t} = I_t \cap C_i$, and $R_{i,t} = R_t \cap C_i$.
- $Transmit(y, x)$ is an event indicating that individual x gets infected by individual y which was already infected
- $c_{i,k}$ is the probability that a susceptible individual native from i being in contact with another infected individual native from k gets infected.
- κ_j is the average number of contacts per unit of time a susceptible will have when visiting city j.
- ξ_{jk} refers to the probability that an individual y meeting x in C_j comes from C_k.
- $\rho_{i,j}$, the probability that an individual from subpopulation i visits subpopulation j. Of course, $\sum_{j=1}^{M} \rho_{ij} = 1$.

Proposition 2. *The coefficient κ should also depend on i, because an individual native from city i meets more people in his own city than abroad ($\kappa_{i,i} > \kappa_{i,j}$).*

6.2 The background

One general question is always posed "how does the population of exposed individuals of subpopulation i evolve ?". For the sake of simplicity, in the process of transmission of the SEIR model, we focus on the incidence and we assume for now that the latent period and the recovery rate, repectively $\mu = \sigma = 0$. Thus, we write a probabilistic formulation of $\frac{dE_i}{dt}$. Assuming the time is discrete, we have $\frac{dE_i}{dt} \approx \mathbb{E}\left[E_{i,t+1} \setminus E_{i,t}\right]$. Then,

$$\begin{aligned}
\mathbb{E}\left[E_{i,t+1} \setminus E_{i,t}\right] &= \mathbb{E}\left[E_{i,t+1} \cap S_{i,t}\right] \\
&= \sum_{x \in C_i} Pr\left[x \in E_{t+1} \wedge x \in S_t\right] \\
&= \sum_{x \in C_i} Pr\left[x \in S_t\right] * Pr\left[x \in E_{t+1} \mid x \in S_t\right] \\
&= Pr_{x \sim \mathcal{X}_i}\left[x \in E_{t+1} \mid x \in S_t\right] * \sum_{x \in C_i} Pr\left[x \in S_t\right] \\
&= |S_{i,t}| \times Pr_{x \sim \mathcal{X}_i}\left[x \in E_{t+1} \mid x \in S_t\right]
\end{aligned}$$

Assume there are M cities. An individual x of the subpopulation i may be visiting another subpopulation, or staying in its own subpopulation. Applying the law of total probabilities, we get:

$$\begin{aligned}
Pr_{x \sim \mathcal{X}_i}\left[x \in E_{t+dt} \mid x \in S_t\right] &= \sum_{j=1}^{M} Pr_{x \sim \mathcal{X}_i}\left[x \in E_{t+dt} \wedge x \in V_{j,t} \mid x \in S_t\right] \\
&= \sum_{j=1}^{M} Pr_{x \sim \mathcal{X}_i}\left[x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}\right] . Pr_{x \sim \mathcal{X}_i}\left[x \in V_{j,t}\right] \\
&\quad \sum_{j=1}^{M} Pr_{x \sim \mathcal{X}_i}\left[x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}\right] \times \rho_{ij}
\end{aligned}$$

Where $\rho_{i,j} = Pr_{x \sim \mathcal{X}_i}\left[x \in V_{j,t}\right]$, the probability that an individual from subpopulation i visits subpopulation j. Of course, $\sum_{j=1}^{M} \rho_{ij} = 1$.

6.3 Study of case where agent x native from city i visits city j

Here, we look at the probability that a susceptible $x \sim \mathcal{X}_i$ visiting j gets infected or not after δt time steps. Let $\mathcal{Y}$ be the uniform distribution over $V_{j,t}$. The correct mathematical approach for this would be to assume that for each city k, the number of people native from k that we meet during δt follows a Poisson process. So both the number of people we meet and the number of infected people we meet during δt should be random variables.

In the approach described in [12], the authors did not do this. They assumed that both the number of people we meet and the number of infected people we meet *are fixed* (otherwise the maths they write would have been different). We will call this the "Keeling & Rohani" interpretation that we will present it in the following parts.

We introduce an alternative approximation, where we assume that the number κ of people we meet during δt is *fixed*, but each of these people has *some probability* to be infected. This is an *in-between interpretation*, easier than the Poisson process maths, but better than Keeling&Rohani's one. We will call this the "Yann-Giang" interpretation.

6.3.1 The "Yann-Giang" interpretation

Proposition 3. *Agent x meets* exactly κ_j *other individuals, and each of these individuals has a probability $\frac{|I_{k,t}|}{N_k}$ of being infected, where k is its native city. Let $y_1 \dots y_{\kappa_j}$ be the individuals that x meets. We get:*

$$\begin{aligned}
& Pr_{x \sim \mathcal{X}_i}\left[x \in S_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \\
= \; & Pr_{x \sim \mathcal{X}_i, y_1 \dots, y_{\kappa_j} \sim \mathcal{Y}}\left[\bigwedge_{p=1}^{\kappa_j} \neg\left(y_p \in I_t \wedge Transmit(y_p, x)\right) \mid x \in S_t \wedge x \in V_{j,t}\right]
\end{aligned}$$

So we have:

$$\begin{aligned} & Pr_{x\sim\mathcal{X}_i}\left[x \in S_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \\ = \;& Pr_{x\sim\mathcal{X}_i,y\sim\mathcal{Y}}\left[\neg\left(y \in I_t \wedge Transmit(y,x)\right) \mid x \in S_t \wedge x \in V_{j,t}\right]^{\kappa_j \delta t} \end{aligned}$$

Moreover, we have:

- the probability so that a susceptible individual x is infected by an infected individual y :

$$\begin{aligned} & Pr_{x\sim\mathcal{X}_i,y\sim\mathcal{Y}}\left[y \in I_t \wedge Transmit(y,x) \mid x \in S_t \wedge x \in V_{j,t}\right] \\ = \;& \sum_{k=1}^{M} Pr_{x\sim\mathcal{X}_i,y\sim\mathcal{Y}}\left[y \in I_t \wedge Transmit(y,x) \mid x \in S_t \wedge x \in V_{j,t} \wedge y \in C_k\right].Pr_{y\sim\mathcal{Y}}\left(y \in C_k\right) \\ = \;& \sum_{k=1}^{M} \{Pr_{x\sim\mathcal{X}_i,y\sim\mathcal{X}_k}\left[y \in I_t \mid x \in S_t \wedge x \in V_{j,t}\right] \\ & \quad \times Pr_{x\sim\mathcal{X}_i,y\sim\mathcal{X}_k}\left[Transmit(y,x) \mid y \in I_t \wedge x \in S_t \wedge x \in V_{j,t} \wedge y \in C_k\right] \times Pr_{y\sim\mathcal{Y}}\left(y \in C_k\right)\} \\ = \;& \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right) \end{aligned}$$

$\xi_{jk} = \frac{N_k \rho_{kj}}{\sum_{v=1}^{M} N_v \rho_{vj}}$ refers to the probability that an individual y meeting x in C_j comes from C_k.

- hence, the probability so that a susceptible individual x is not infected by an infected individual y :

$$1 - \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)$$

- thereby, the probability so that a susceptible individual x is not infected after κ_j contacts per unit time δt.

$$\left[1 - \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right]^{\kappa_j \delta t}$$

- thus, the probability so that a susceptible individual x becomes infected after κ_j contacts per unit time δt.

$$Pr_{x\sim\mathcal{X}_i}\left[x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \quad = \quad \left[1 - \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right]^{\kappa_j \delta t}$$

We now apply the *log* approximation which consists in approximating $1-(1-u)^v$ by $v\log(1-u)$:

$$Pr_{x\sim\mathcal{X}_i}\left[x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \quad = \quad -\kappa_j \delta t \log\left[1 - \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right]$$

So, the transmission rate per susceptible individual is as follows :

$$\frac{dPr_{x\sim\mathcal{X}_i}\left[x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}\right]}{dt} \simeq -\kappa_j \log\left[1 - \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right]$$

In fact, we use the parameter λ to present this quantity, and it is denoted as the "force of infection" :

$$\lambda_i = \sum_j \rho_{ij} \kappa_j \log\left[1 - \sum_{k=1}^{M} \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right]$$

If there is only one city i, then

$$\lambda_i = \kappa_j log(1 - \frac{|I_i|}{N_i} \times c_{ii})$$

6.3.2 "Keeling & Rohani" Interpretation

Proposition 4. *Agent x meets* exactly $\kappa_j \delta t \xi_{jk} \frac{|I_{k,t}|}{N_k}$ *other infected individuals native from city k.*

Let $l_k = \kappa_j \delta t \xi_{jk} \frac{|I_{k,t}|}{N_k}$.

Let $y_1^k \dots y_{l_k}^k$ be the infected individuals native from k that our individual x meets between t and $t + \delta t$.

We have the probability so that a susceptible individual x is not infected after having seen l_k individuals between t and $t + \delta t$:

$$\begin{aligned}
& Pr_{x\sim\mathcal{X}_i}\left[x \in S_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \\
= \; & Pr_{x\sim\mathcal{X}_i}\left[\bigwedge_{\substack{k=1\dots M \\ p=1\dots l_k}} \neg\left(Transmit(y_p^k, x)\right) \mid x \in S_t \wedge x \in V_{j,t}\right] \\
= \; & \prod_{k=1}^{M} Pr_{x\sim\mathcal{X}_i}\left[\bigwedge_{p=1\dots l_k} \neg\left(Transmit(y_p^k, x)\right) \mid x \in S_t \wedge x \in V_{j,t}\right] \\
= \; & \prod_{k=1}^{M} (1 - c_{ik})^{\kappa_j \delta t \xi_{jk} \frac{|I_{k,t}|}{N_k}}
\end{aligned}$$

Then, we plug this back into the previous formula, and we get:

$$Pr_{x\sim\mathcal{X}_i}\left[x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \quad = \quad 1 - \prod_{k=1}^{M} (1 - c_{ik})^{\kappa_j \xi_{jk} \frac{|I_{k,t}|}{N_k} \delta t}$$

The first order approximation of $1 - \prod_{k=1}^{M}(1 - c_{ik})^{v_k}$ is $\sum_{k=1}^{M} -v_k \log(1 - c_{ik})$. Applying this approximation here, we get:

$$Pr_{x\sim\mathcal{X}_i}\left[x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}\right] \quad \simeq \quad \delta t \sum_{k=1}^{M} \left(-\kappa_j \xi_{jk} \frac{|I_{k,t}|}{N_k} \log\left(1 - c_{ik}\right)\right)$$

Define $\beta_{ijk} = -\kappa_j \log\left(1 - c_{ik}\right)$, let δt converge to zero, and we get:

$$\frac{dPr_{x\sim\mathcal{X}_i}\left[x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}\right]}{dt} \simeq \sum_{k=1}^{M} \left(\xi_{jk} \frac{|I_{k,t}|}{N_k} \beta_{ijk}\right)$$

If there is only one city i, then we fall back to the formula of [12]. We have :

$$\beta_i = -\kappa_i \log\left(1 - c_i\right)$$

$$\frac{d}{dt}\mathbb{E}\left[|E_{i,t+dt} - E_{i,t}|\right] \simeq -\left|S_{i,t}\right| \left(\frac{|I_i|}{N_i}\beta_i\right)$$

and the force of infection as follows :

$$\lambda_i = \beta_i \frac{|I_i|}{N_i}$$

6.4 Final Formula

We simply have to plug in the probability ρ_{ij} that i visits j.

We get, for the "Yann-Giang" interpretation :

$$\frac{d}{dt}\mathbb{E}\left[|E_{i,t+dt} - E_{i,t}|\right] \simeq -|S_{i,t}| \sum_j \rho_{ij}\kappa_j \log\left[1 - \sum_{k=1}^{M}\left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right]$$

And for the "Keeling & Rohani" Interpretation :

$$\frac{d}{dt}\mathbb{E}\left[|E_{i,t+dt} - E_{i,t}|\right] \simeq -|S_{i,t}| \sum_j \rho_{ij} \sum_{k=1}^{M}\left(\xi_{jk}\frac{|I_{k,t}|}{N_k}\beta_{ijk}\right)$$

7 Appendix : Characterization of synchrony

Call $\delta_{ij} = \delta_{ji}$ $(0 \leqslant \delta_{ij} < 2\pi)$ the phase difference between subpopulations i and j :

$$\delta_{ij} = |\varphi_i - \varphi_j| \bmod 2\pi \tag{7.1}$$

where φ_i and φ_j are the phases of the contact rates (equation 2.8) in subpopulations i et j. Populations i and j are perfectly in phase if $\delta_{ij} = \delta_{ji} = 0$ or 2π and in opposition of phase if $\delta_{ij} = \delta_{ji} = \pi$. We can thus define the degree of synchrony $\xi_{ij} = \xi_{ji}$ $(0 \leqslant \xi_{ij} \leqslant 1)$ between populations i and j as

$$\xi_{ij} = 1 - \frac{|\delta_{ij}|}{\pi}. \tag{7.2}$$

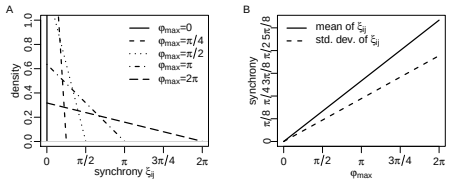

Figure 7.1: Synchrony in the case of model 0. (A) distribution of synchrony ξ_{ij} for various values of φ_{max}. (B) mean and standard deviation of the distribution of ξ_{ij} as functions of $\varphi_{\max}$.

Consider that in the metapopulation the phases φ_i of the contact rates in the n subpopulations are evenly distributed between 0 and $\varphi_{\max}$ $(0 \leqslant \varphi_{\max} \leqslant \pi)$. We can express the mean of the pairwise phase differences $\delta_{ij} = \delta_{ji}$ as

$$<\delta_{ij}> = <\delta_{ji}> = 2\varphi_{\max} \sum_{k=1}^{n-1} \frac{(n-k)k}{(n-1)n^2} = \frac{n+1}{3n}\varphi_{\max} \tag{7.3}$$

and thus the mean of the synchronies $\xi_{ij} = \xi_{ji}$ as

$$<\xi_{ij}> = <\xi_{ji}> = 1 - \frac{n+1}{3n}\frac{\varphi_{\max}}{\pi} \tag{7.4}$$

and thus

$$\lim_{n\to\infty} <\xi_{ij}> = 1 - \frac{\varphi_{\max}}{3\pi} \tag{7.5}$$

This last result shows that, for a high enough number n of subpopulations, the mean value of the ξ_{ij} does not depend on the number of subpopulation.

The values of φ_i are chosen so that they are uniformly distributed between $\varphi_{\min} = 0$ and $\varphi_{\max}$. The distribution of ξ_{ij} doesn't depend on n the number of subpopulation, but only depends $\varphi_{\max}$ and may be is characterized by one single parameter (we choose the average value of all ξ_{ij}), view figure 7.1.