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

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

25/2/2016

Abstract

Global persistence of infectious diseases is a big problem for epidemiologists. Studies have showed that there are a lot of reasons to answer why many communicable diseases still exist and have been developed in more dangerous form. The asynchrony and the recolonization among subpopulations are two key reasons pointed out. However, why these are the asynchrony and the recolonozations in a metapopulation is still an open question. Here we study the combined effects of forcing phase heterogeneity in the seasonally forced contact rate on global persistence of disease. We carry out an exploitation of stochastic dynamics in a susceptible-exposed-infectious-recovered (SEIR) model of the spread of infectious diseases in a metapopulation of n subpopulations. Starting with continuous-time Markov description of the model of deterministic equation, the direct method of Gillespie(1977) [12] in the class of Monte-Carlo simulation methods allows us to simulate exactly the spread of disease with the SEIR model. Our finding of the exploitation of stochastic dynamics points out that the persistence of the disease in the metapopulation is characterized as an exponential survival model on data simulated by the stochastic model. Using a parametric survival model for an exponential distribution (R package 'survival' [27]), we estimate the global extinction rate which represents the global persistence of disease in the meta-population. Besides, we estimate the locale extinction rate and the recolonization rate in metapopulation by using the Poisson process. We find how bigger the forcing phase heterogeneity becomes, and how smaller the local extinction rate gets.

1 INTRODUCTION

The persistence effect, whereby local population extinction is pointed by movement of disease in space, is an important mechanism for preventing complete extinction of a regional population [7, 13, 26, 28]. The exact form of disease movement depends on a number of local factors (demographic including population size [16], growth rate and death rate [5], sociological such as school period of children, work tendency from rural to urban, environmental and climatic comprising seasonal variations in seasonality [6, 14], temperature

and rainfall, immunological for diseases, etc...) as well as the connections between the different populations (i.e. spatial structure) [31] such as distance [16], coupling rate, number of individuals between populations, etc....Thus, the problem finding the form of disease movement in space is still open and quite complex.

The characteristic of the human infectious diseases is that they can be spread from one person to another via indirect or direct contacts, and can move in space, from one place to another. Thus, disease risk occurs in worldwide sales if we don't control disease transmission. Modelling studies on the spatial distribution and spread of infectious diseases in coupled populations are becoming increasingly detailed and are giving more exact predictions. For human infectious disease, metapopulation model is the simplest model that is a set of subpopulations with mutual interaction [22] here a subpopulation can only go extinct locally and be recolonized by another after it is emptied by extinction [3, 17, 22]. A metapopulation is also a population of populations (subpopulations). Such a structure implies an heterogeneity in the sense that the probability of contact (or contact rate) between individuals from a same subpopulation is higher than the probability of contact between individuals of different subpopulations [11]. Such heterogeneity is actually the result of the interaction between two phenomena that are often difficult to disentangle in nature. The first one relates to the granularity of the metapopulation (as rendered by the number of and sizes of subpopulations) and the second one relates to the isolation between subpopulations (as can be rendered, among others, by physical distances separating each pair of subpopulations). Moreover, according to the findings of Benjamin Bolker (1995) [2], there is no coexistence between periodicity and disease persistence in non-spatial measles models, and spatial structure is an important factor to both enhance persistence and create new types of dynamic behaviour.

The disease persistence capability in a metapopulation depends positively on level of synchrony/asynchrony between subpopulations [16, 19]. Many studies showed that the synchronization of epidemics in all asynchronous subpopulations causes the recolonization of diseases for locally extinct subpopulation [3, 16, 17, 19, 30]. So, the recolonization becomes the main reason for which disease persistence exists. The disease always appears in metapopulation if and only if there is at least one non-extinct subpopulation. In 1996**,** in order to explain why measles persists after lot of vaccination policies, Bolker [3] used the measles data before and after vaccination from 1964 to 1988 in England and Wales. Vaccination has broken high synchrony between the UK cities in prevaccination era, and at the same time, causes decorrelation and enhances global persistence of the infection, because of decorrelating factors of vaccination such as the starting moment of vaccination policies, number of susceptibles vaccinated and interaction between vaccination policies [3, 23]. A decrease in correlation between subpopulations may make a metapopulation more difficult to eradicate infectious diseases [3, 9]. Moreover, the level of synchrony between the subpopulations is strongly governed by the migration rate and distances between them [19]. In our modern world, the distance problem isn't large anymore for individuals who want to travel. There are a lot of cities very remote, but very connected,

and thus very synchronous as in the USA [4]. In contrast, migration among subpopulations has become a big problem. The disease synchrony speed within a metapopulation can strongly increase when migration rate there is strong [21]. The migration rates are directly proportional to the amount of variation in metapopulation size, but inversely to the amount of variation in subpopulation size, over time [8, 14]. So, migration is key to the recolonization of empty subpopulation and simultaneously increases the degree of synchrony between subpopulations in spatially structured metapopulation. However, the vast majority of infectious diseases control policies that are applied in the world are still based on rationales that do not consider the local extinction/recolonization dynamics. This is maybe a reason why measles persists around the world, despite highly local vaccine coverages [5]. For example, in the start of the 2014, the World Health Organization (WHO) had officially stated the global measles epidemic outbreak. In the first three months of the year 2014, there were about 56,000 cases of measle infections in 75 countries [29], including countries in south-east Asia and most particularly, Vietnam [18]. We discovered measles persistence in the world for many years without extinction, from one nation to another as well as from cities to other cities. Though the moment of disease outbreaks in each region differs. For neighbouring regions with disease persistence, there is a time-lag differs between disease outbreak. This is explained in sociology by difference in culture, in geographic condition and more particularly seasonality.

Seasonality has been one in rather robust ingredients influencing the disease persistence process. Seasonal changes can alter migration tendency between urban and rural areas [10], also residence time of hosts, vectors and pathogens. Seasonal variation can thus determine population size, migration and interaction capabilities and particularly infection rate at which susceptible individuals become infected [1, 20]. Hence, infectious disease outbreak occurs due to this infection rate. However, finding a clear mechanism of seasonal forcing for modelling is a very difficult work because of unidentified formula for seasonal forcing [1, 10]. For indirectly transmission diseases such as water-born and vector-born, finding seasonality characteristics is less a problem, however, in direct contrast to transmission diseases such as measles. Seasonal forcing in a metapopulation is influenced by weather and climate in region, school schedule of children, and rural-urban migration in countries [2, 5, 10, 15]. In these factors, the seasonal aggregation of children in primary schools affects clearly the infection rate in metapopulation. The infection rate decreases due to children holidays but is inverse when the children come back to school [5]. So**,** exploring the influence of seasonality for the infection rate β in simulation has been developed in many previous years. If the infection rates given are the same in all subpopulations, so this metapopualtion model is a rather simple model [24, 25] and the symmetry of the fixed points among subpopulations will not be broken. In contrast, if the infection rate is different in all subpopulations, so we have a more complex oscilation metapopulation model, but close to the oscilations in reality. Thus, realizing the oscilations of infectious diseases in life within metapopulation simulation models to estimate global disease persistence time has become a large problem and the infectious disease eradication

has become our aim [9].

Here we propose a simulation study to quantify the effect of synchrony on the persistence of infectious diseases. We use stochastic simulations for infectious diseases in a metapopualtion, then we consider different spatial structures from the simplest to more complex. These forcings can reflect local demographic, sociological, environmental, or climatic factors. The level of synchrony is computed from the phases of forcing in the different subpopulations and persistence is quantified by using statistical tools from survival analysis. Here, we are concerned about measles and simultaneously use the parameter values from articles and measles reports. As the persistence of measles has been largely studied in the literature and its still unexplained while global persistence is a growing concern for WHO and public health authorities around the world [6].

To do this, we first build the deterministic model for a metapopualtion. Then, we describe the spatial structure for the SEIR metapopulation model. Finally, we introduce the characterization of mass extinction, locale extinction and recolonization in the metapopulation based on the measles characteristics.

2 METHODS

3 RESULT

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

[3] B. M. Bolker and B. T. Grenfell. Impact of vaccination on the spatial correlation and persistence of measles dynamics. *Proc Natl Acad Sci U S A*, 93(22):12648–12653, Oct 1996.

[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 and B. T. Grenfell. Seasonality and the persistence and invasion of measles. *Proc Biol Sci*, 274(1614):1133–1141, May 2007.

[6] 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.

[7] G. S. Cumming. The impacts of low-head dams on fish species richness in wisconsin, usa. *Ecological Applications*, 14:1495–1506., 2004.

[8] S. Dey and A. Joshi. Stability via asynchrony in drosophila metapopulations with low migration rates. *Science*, 312(5772):434–436, Apr 2006.

[9] D. J. Earn, P. Rohani, and B. T. Grenfell. Persistence, chaos and synchrony in ecology and epidemiology. *Proc Biol Sci*, 265(1390):7–10, Jan 1998.

[10] M. J. Ferrari, A. Djibo, R. F. Grais, N. Bharti, B. T. Grenfell, and O. N. Bjornstad. Rural-urban gradient in seasonal forcing of measles transmission in niger. *Proc Biol Sci*, 277(1695):2775–2782, Sep 2010.

[11] Ilkka Hanski; Oscar E Gaggiotti. *Ecology, Genetics and evolution of metapopulations*. 2004.

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

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

[14] B. D. Griffen and J. M. Drake. Environment, but not migration rate, influences extinction risk in experimental metapopulations. *Proc Biol Sci*, 276(1677):4363–4371, Dec 2009.

[15] 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.

[16] T. J. Hagenaars, C. A. Donnelly, and N. M. Ferguson. Spatial heterogeneity and the persistence of infectious diseases. *J Theor Biol*, 229(3):349–359, Aug 2004.

[17] I. Hanski. Metapopulation dynamics. *Nature*, 396, 1998.

[18] healthmap.org. Measles reemerges in vietnam, 2014.

[19] M. Heino, V. Kaitala, E. Ranta, and J. Lindstrom. Synchronous dynamics and rates of extinction in spatially structured populations. *The Royal Society*, 264:481–486, 1997.

[20] P. R. Hosseini, A. A. Dhondt, and A. Dobson. Seasonality and wildlife disease: how seasonal birth, aggregation and variation in immunity affect the dynamics of mycoplasma gallisepticum in house finches. *Proc Biol Sci*, 271(1557):2569–2577, Dec 2004.

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

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

[23] P. Rohani, D. J. Earn, and B. T. Grenfell. Opposite patterns of synchrony in sympatric disease metapopulations. *Science*, 286(5441):968–971, Oct 1999.

[24] G. Rozhnova, A. Nunes, and A. J. McKane. Phase lag in epidemics on a network of cities. *Phys Rev E Stat Nonlin Soft Matter Phys*, 85(5 Pt 1):051912, May 2012.

[25] G. Rozhnova1, A. Nunes, and A. J. McKane. Impact of commuting on disease persistence in heterogeneous metapopulations. 2013.

[26] D. J. Smith, A. S. Lapedes, and J. C. de Jong. Mapping the antigenic and genetic evolution of influenza virus. *Science 305, 371 (2004);*, 2004.

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

[28] C. Viboud, O. N. Bjørnstad, D. L. Smith, L. Simonsen, M. A. Miller, and B. T. Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. *science*, 312(5772):447–451, 2006.

[29] WHO. Reported measles cases with onset date from oct 2013 to mar 2014, 2014.

[30] G. Yaari, Y. Ben-Zion, N. M. Shnerb, and D. A. Vasseur. Consistent scaling of persistence time in metapopulations. *Ecology*, 93(5):1214–1227, May 2012.

[31] G. Yan, Z. Q. Fu, J. Ren, and W. X. Wang. Collective synchronization induced by epidemic dynamics on complex networks with communities. *Phys Rev E Stat Nonlin Soft Matter Phys*, 75(1 Pt 2):016108, Jan 2007.

5 Appendix : equilibrium values of the system ??–??

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 ??

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 $\rho_{i,j} = Pr_{x\sim\mathcal{X}_i}[x \in V_{j,t}]$, 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 [21], 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:*

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

So we have:

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

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.

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

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

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

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 $t + \delta t$:

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

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 [21]. 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 :

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

And for the "Keeling & Rohani" Interpretation :

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

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 **??**) 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

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

Consider that in the metapopulation the phases φ_i of the contact rates in the n subpopulations are evenly

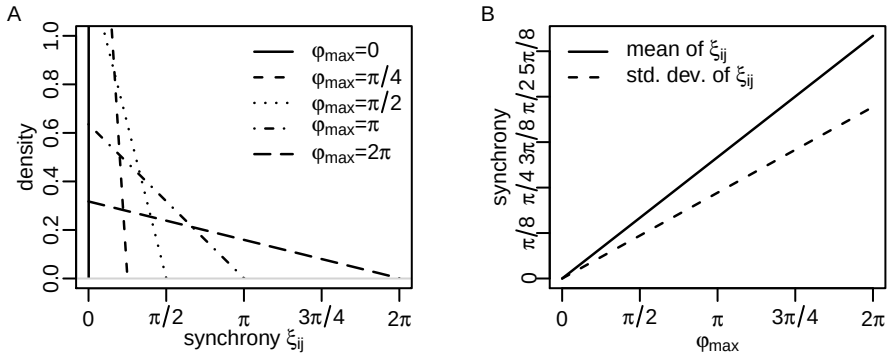

Figure 7.1: Synchrony in the case of model 0. (A) distribution of synchrony ξ_{ij} for various values of φ_{max}. (B) mean and standard deviation of the distribution of ξ_{ij} as functions of $\varphi_{\max}$.

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.