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) [14] 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' [36]), 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

News about infectious diseases has always been a subject of worry to parents as well as all. It has brought many problems to humain society. Recent works have shown that infectious diseases do spread in space [9, 15, 35, 37]. The exact form of disease movement depends on a number of local factors (demographic including population size [19], growth rate and death rate [7], sociological such as school period of children, work tendency from rural to urban, environmental and climatic comprising seasonal variations in seasonality [8, 17], temperature and rainfall, immunological for diseases, etc...) as well as the connections between the different populations (i.e. spatial structure) [40] such as distance [19], coupling rate, number of individuals between populations, etc....Hence, we focus on spread of disease in space by using variations in the seasonal aspects of subpopulations and after examine how synchrony could affect the persistence of infectious diseases in a metapopulation.

In modeling of ecological system, presenting interactions between humans, subpopulations, geographic conditions and the metapopulation model is a good choice. Metapopulation is a set of subpopulations with mutual interaction [31] here a subpopulation can only go extinct locally and be recolonized by another after it is emptied by extinction [5, 20, 31]. A metapopulation is also a population of populations (subpopulations). Such a structure implies an heterogeneity in the sense where 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 [13]. 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) [3], 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.

In addition, the disease persistence capability in a metapopulation depends positively on level of synchrony/asynchrony between subpopulations [19, 23]. Many studies showed that the synchronization of epidemics in all asynchronous

subpopulations causes the recolonization of diseases for locally extinct subpopulation [5, 19, 20, 23, 39]. 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 [5] 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 [5, 32]. A decrease in correlation between subpopulations may make a metapopulation more difficult to eradicate infectious diseases [5, 11]. In addition, the level of synchrony between the subpopulations is strongly governed by the migration rate and distances between them [23]. 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 [6]. 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 [29]. 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 [10, 17]. 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 [7]. 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 [38], including countries in south-east Asia and most particularly, Vietnam [22]. 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 [12], 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, 25]. 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, 12]. 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 [3, 7, 12, 18]. 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 [7]. 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 [33, 34] 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 [11].

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 [8].

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

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 [29] :

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

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 [31], 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 [16]. 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 $[0, \varphi_{max}]$ 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 [14] 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

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 $R_i \leftarrow R_i + 1$

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 [15]. Estimating the ability of the disease persistence is still a complex problem and 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*' [36]). 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 [14] for metapopulation model as described in the previous part. With the SEIR metapopulation model, measles is modeled [2, 18]. 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.

3 RESULT

3.1 Effect of the number of subpopulations in the coupled model

As shown in Fig 3.1, there is a good agreement between simulations and analytic results for the synchronization and phase lag between subpopulations in the metapopulation. We implemented the simulation for the SEIR metapopulation model of subpopulations coupled by individual movement with time-varying periodic contact rate.

The result shows that the number of subpopulations in the metapopulation strongly affects the extinction rates, increasing the number of subpopulations leads an increase of the rates of local extinction as well as global extinction. The extinction rates of the synchrony is always higher than that of the asynchrony in the independent metapopulation as well as in the coupled metapopulation. Moreover, the extinction rates of the metapopulation of independent subpopulations is larger than the one of the metapopulation of coupled subpopulations.

Table 2: Some Disease Parameter Values for Measles from the Literature [3, 6, 8, 27, 28, 29]

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

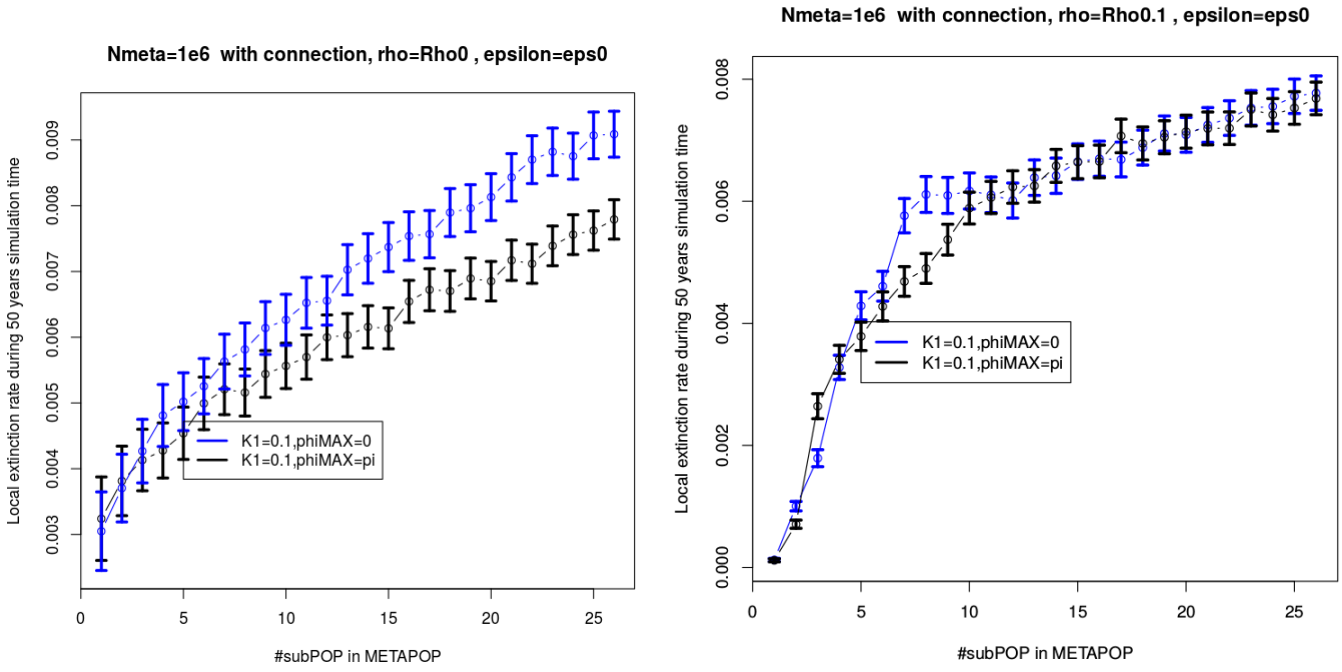

Figure 3.1: Influence of the number of subpopulations on the local extinction rate in the forced model of the coupled subpopulations and of the isolated subpopulations. The number of subpopulations n from 1 to 26, the coupling rate $\rho = 0.1$, the fixed metapopulation size $N = 1e6$, and $\varphi_{\max} = 0$ or π and the amplitude of the number of contacts $\kappa_1 = 0.1$. The results show that for the spatial stochastic model, the local extinction rate of the uncoupled model is higher than that of the coupled model.

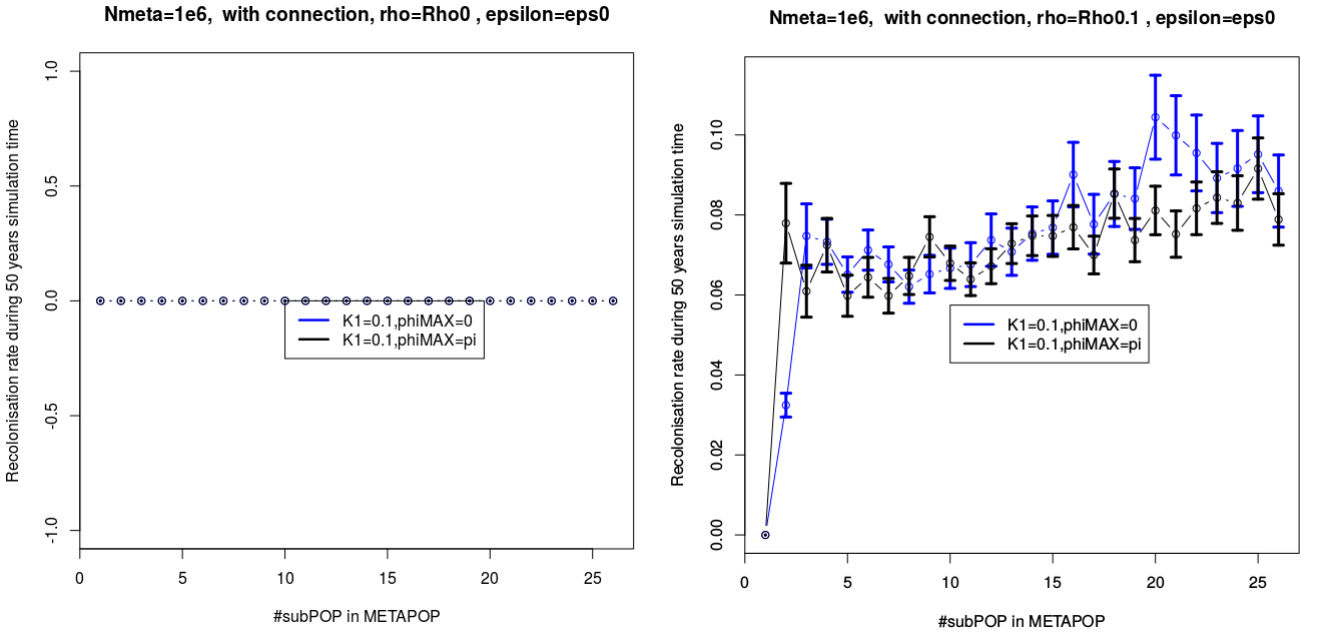

Figure 3.2: Influence of the number of subpopulations on the rate of recolonisation in the forced model of the coupled subpopulations and of the isolated subpopulations. The number of subpopulations n from 1 to 26, the coupling rate $\rho = 0.1$, the fixed metapopulation size $N = 1e6$, and $\varphi_{\max} = 0$ or π and the amplitude of the number of contacts $\kappa_1 = 0.1$. The results show that for the spatial stochastic model, the rate of recolonisation in the synchrony case is higher than that in the asynchrony case. Inversely, for the non-spatial stochastic model, no migration has no recolonisation, then the recolonisation rate of this uncoupled model is always zero.

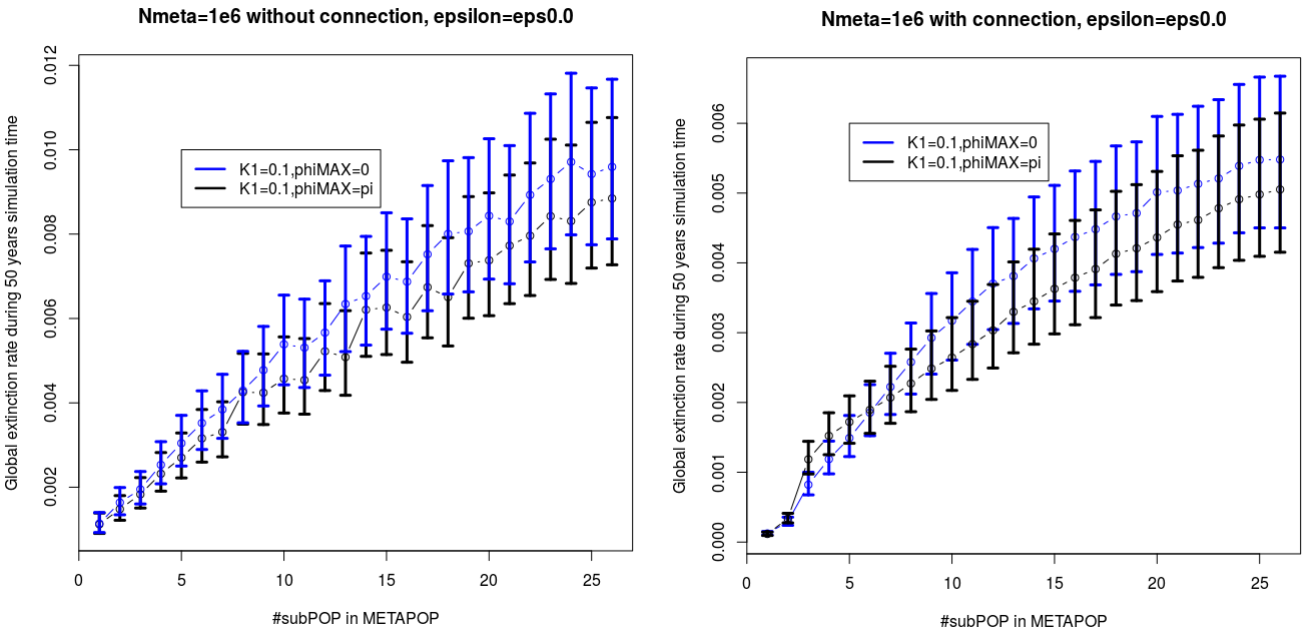

Figure 3.3: Quantifing the global extinction rate in the metapopulation. The number of subpopulation n from 1 to 26, the metapopulation size $N = 1e6$ and the coupling rate $\rho = 0.1$ and the amplitude of the number of contacts $\kappa_1 = 0.1$. As this result, the global extinction rate of the synchrony with $\varphi_{\max} = 0$ has the tendance to be larger than that of the asynchrony with $\varphi_{\max} = \pi$. Moreover, the increased number of subpopulation leads to strongly decline the subpopulation size. This drives quite the decrease of the global persitence time of the metapopualtion, so the mass extinction rate of the metapopulation increases significantly.

In order to investigate the dependence among subpopulations, we use a parameter ρ that is referred to the coupling interaction. Increasing the value of ρ means that we increase the strength of connection among subpopulations in a metapopulation. Besides, $\varphi_{\max}$ is an important parameter which clearly reflects the difference between subpopulations'phases. With the decreasing of $\varphi_{\max}$, we increase the similarity between sub-populations, but we limit the chances of re-colonizations.

In the model of isolated subpopulations no immigration, the subpopulations fluctuate totally independently. The rescue ability of disease among subpopulations is zero. It is why the probability of disease persistence is smaller than the case of the metapopulation of coupled subpopulations. Decreasing the probability of disease persistence leads an increase of the rates of global extinction as well as of local extinction. These extinction rates of the metapopulation without connection is greater than the metapopulation with connection. Due to no immigration among subpopulations, the rate of recolonization is zero in all cases, it means that the disease can not turn back when a subpopulation did find the local extinction. However, a model of isolated subpopulations no immigration is impossible in our world, so this result is only significant for theory, but non-significant for reality.

Moreover, by fitting the seasonality amplitude of the average contact rate κ_1 at which a susceptible will have when visiting city i per unit of time, so as to describe more exactly the variation in contact rate due to the opening and closing of schools, to term-time forcing and more further, to an age-structure. A forcing model will appear, when we make κ_1 be greater than zero, the average number of contacts per unit of time of each subpopulation will be a sinusoidal function. Besides, to describe the seasonal phase difference among subpopulations, we use the parameter $\varphi_{\max}$ to adjust the phase difference of disease between seasonal air-temperature variations in a metapopulation.

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.

Moreover, in the coupled model, when the number of subpopulations increase, we can see that the the local extinction rate of the synchrony goes closer to that of the asyschrony. Because, first we increase the number of subpopulation, we increase the similarity among subpopulations, we decrease the ability of the recolonization among subpopulation, and one more plus point, increasing the number of subpopulation in the case of the fixed metapopulation size makes the subpopulation size decrease.

3.2 Effect of the number of subpopulations in the model without the forcing

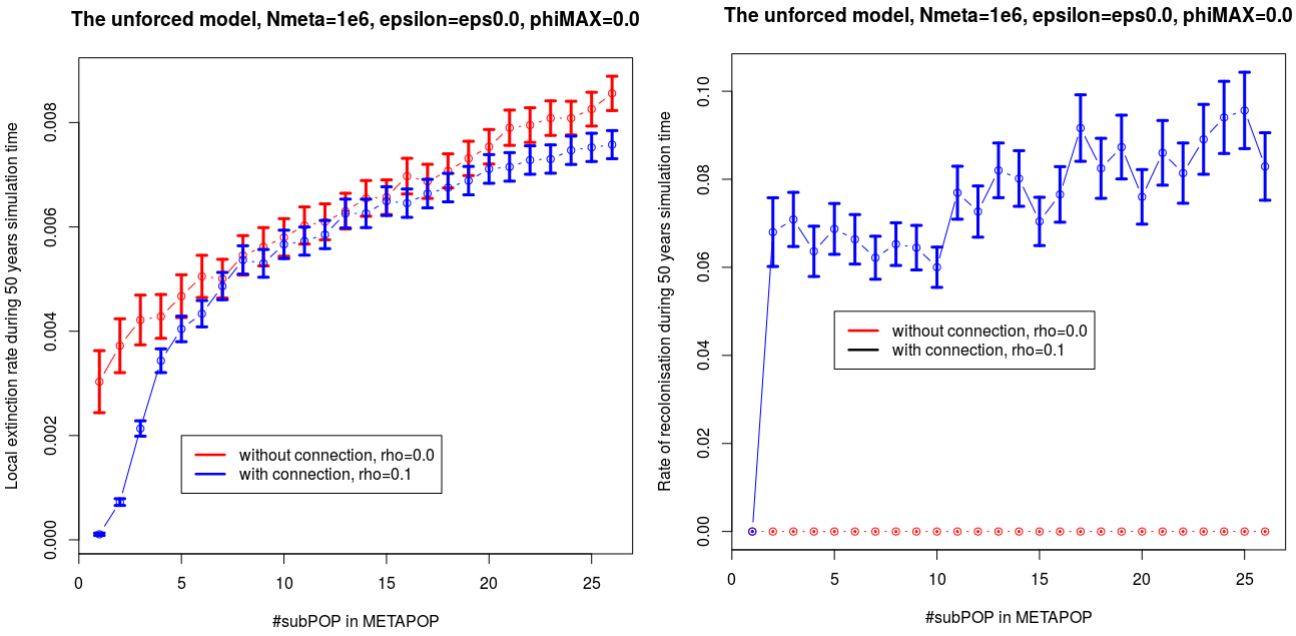

Figure 3.4: Influence of the number of subpopulations on the local extinction rate in the model without the forcing. The number of subpopulations n from 1 to 26, the coupling rate $\rho = 0.1$, the fixed metapopulation size $N = 1e6$, the amplitude of season $\kappa_1 = 0$ and $\varphi_{\max} = 0$. The results point out that in the unforced model, it means that the rate of the force of infection is constant over time, then the local extinction rate of the stochastic seasonally unforced model without connection is greater than that with connection. Besides, the rate of recolonisation in the model without connection is always equal to zero, and much smaller than that in the model with connection.

In this experience, we are interested in the influence of the number of subpopulations on the local extinction rate for the standard SEIR model without forcing. This simple metapopulation model just investigates subpopulations connected by individual movement without environmental transmission. In the non-seasonal model, the average number of contacts per unit of time of a susceptible when visiting another subpopulation is set as a constant. The strength of connection ρ is set to 0 for the model of isolated subpopulations (island model) and different from 0 for the model of coupled subpopulations (coupling model). Hence, in this case, the main factor affecting the extinction rate is the migration of individuals rather any environmental factor.

The result shown in the Fig 3.2 points out that in the isolation model with the fixed metapopulation size, the increased number of subpopulations leads to a decrease of the subpopulation size. This limits the ability of the rescue effect to ensure locally extinct subpopulations become recolonized. Hence the local extinction rate rises significantly, and is greater than that of the coupled model without forcing. Besides that, due to the disjunction among subpopulations in the metapopulation, a subpopulation will obtain the global extinction immediately after it get the first local extinction, so any there won't be any recolonization occurs in this island model. The recolonization rate is zero in all cases as shown in the Fig 3.2.

Inverse to the coupled model, the strength of interaction between the two subpopulations have clearly demonstrated its influence on the extinction ability in metapopulation. The curve of the coupled case will be increased notably when the number of subpopulation extends in the range from 1 to 26. As explained in above obtained

results of the Fig 3.1, although the number of subpopulation augments but is still small, the number of native individuals in a subpopulation is much larger than the number of tourists in any specific time. Thus, for the fixed metapopulation size, the increased number of subpopulations draws a decline of the subpopulation size. The disease of subpopulation easily get extinct, whenever the local extinction rate augments. However, when the number of subpopulation becomes a large number, then the subpopulation size becomes really small. The ability of disease persistence significantly decreases, then the rates of extinction augments. As pointed out in the Fig 3.2, the recolonization rate has the same form to the local extinction rate.

3.3 Influence of the coupling strength on the local extinction rate

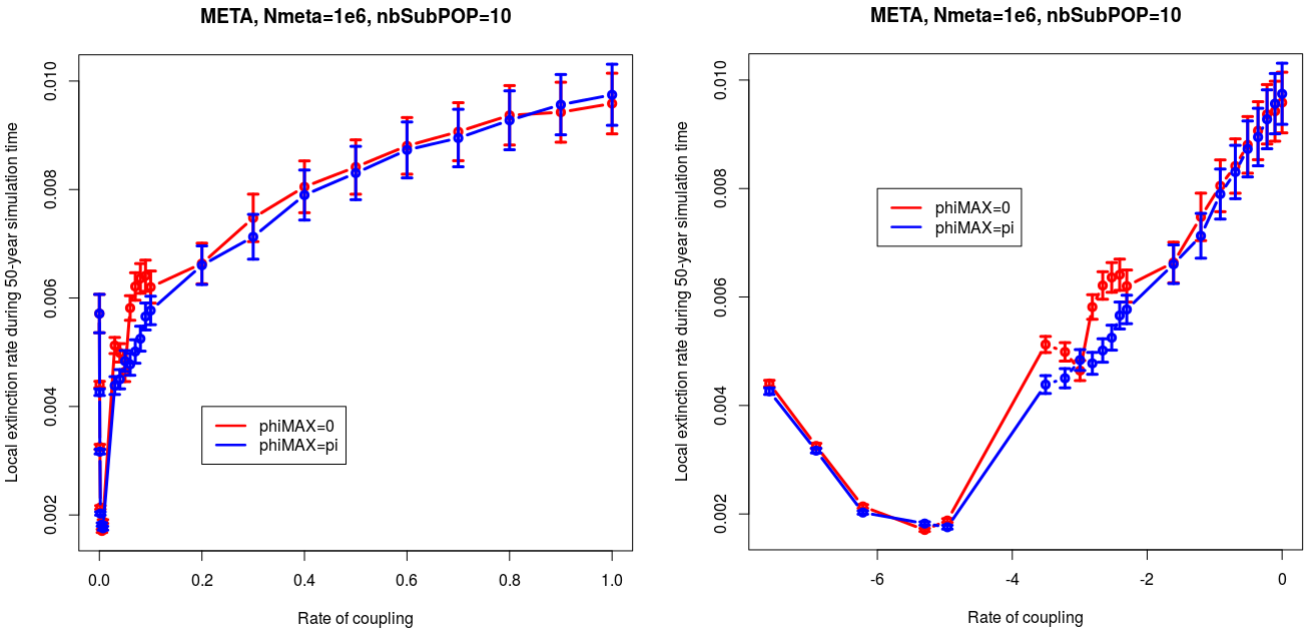

Figure 3.5: Influence of the coupling strength on the local extinction rate of the metapopulation of 10 subpopulations with the metapopulation size $N = 1e6$. The line that presents the relation between the coupling strength and the local extinction rate looks like an inverse bell. The extinction rate goes down to minimum when the streng of connection is medium.

One more factor that was pointed in the introduction part is coupling strength between subpopulations. 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 with the population size $N = 1e6$ for each subpopulation. 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. We have result as following figure Fig 3.5.

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 3.5, 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

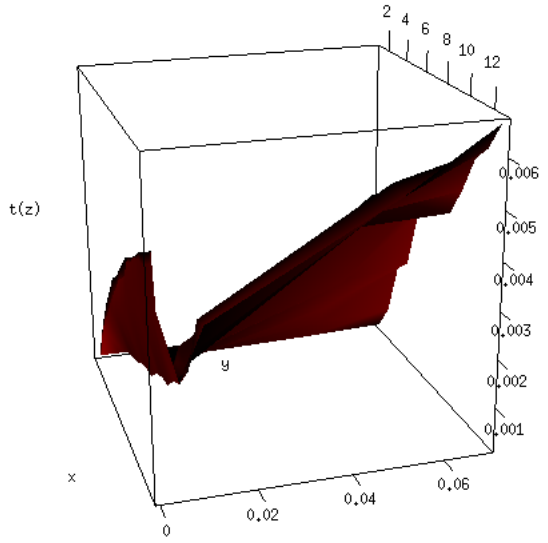

Figure 3.6: Three-dimensional relation among the number of subpopulation in a metapopulation, the estimated local extinction rate and the coupling strength. The three-dimensional result shows that the line of the gap between the coupling rate and the extinction rate is as a convex curve and obtains minimum when the coupling rate is medium. Moreover, increasing the number of subpopulations drives an increase of the extinction rate. The x-axis is the number of subpopulation, the y-axis is the coupling rate, and the z-axis is the estimated extinction rate.

rate rises.

4 DISCUSSION

We successfully have built a version for the susceptible-infected-recovered stochastic metapopulation model (subpopulations connected by individual movement), which describes both movement-based and environmental transmission. The infection rate λ_i for $subpopulation_i$ has portrayed all effects inside as well as outside of the disease tranmission chain between individuals in the same subpopulation or in other subpopulations. Moreover, our metapopulation model became more detailed when we brought seasonality in metapopulation model to create periodic transmission in year that highlighted seasonal changes as well as school period of children [4, 3, 11, 28]. We have metapopulation model with different contact rates for each subpopulation. This is a more complex model than any used metapopulation model. We have sketched successfully in-phase and sometime out-of-phase ("antiphase") models across suburbs of He's 2003 [21].

This complex metapopulation model is also an expected result of Rozhnova(2012) [33]. It's a good result for scientists wanting to use the SEIR metapopulation model for simulating dynamics of infectious diseases. Our results roughly support those of Rozhnova's 2012. The authors gave different values of the contact rates β of each subpopulation. However, the rates β here are fixed by constants and the number of subpopulations in experiences are maximum of three. We don't know why these authors made simulations with just three cities. We find that, this number of subpopulations is quite small and the obtained result is not enough strong to ocnfirm. Comparing this result with our's, in a coupling metapopulation, the degree of synchorny is maintained when the coupling rate between subpopulations is weak.

Moreover, our stochastic SEIR metapopulation model with subpopulations connected to each other, we have quantified disease extinction of seasonality as well as spatial synchrony. With our model, we can easily create level of seasonality in year and at the same time, phase difference in seasonality between subpopulations. It's the reason why we have model quite close to the metapopulation model in reality.

In addition, for approaches of measuring the probability of disease persistence in a metapopulation, as previous analyses, the ability of disease persistence was characterized by fitting an exponential survival model [8, 30] on a

data simulated by a stochastic model. To measure the persistence in ecology and epidemiology, so many methods we can use [8, 18, 28]. For example, Keeling et al.(2002) [28] gave two methods. One method was for an isolated metapopulation without migration by calculating the expected extinction time or the extinction rate during a given period. This was a theoretical measure as no real data exists to compare with model results. The other method for a population with migration was found by calculating the number or the total duration of extinctions. Then in 2010, "mean annual fade-out" and "fade-outs post epidemic" methods proposed by Conlan [8] were used to quantify persistence by basing on the proportion or on the frequency of zero reports in a given reporting interval. The above approaches have the same point that they are easily calculated in all models from simple to complex. But, the obtained results are not exact, because as we mentionned above, the curves of disease persistence are Kaplan-Meier survival curves that present the lifetime of every subpopulation, so the ability of disease persistence is estimated as a survival function and in fact, it was proved as an exponential survival model. Hence, in this paper, we have revealed a new method. We have mesured the rates via Kaplan-Meier survival curves and survival functions. Our obtained results held the true values of the rates with the confidence interval of 95%.

Due to the phase difference between infection coefficient β, we can change by an increase or a decrease in level of synchrony. We want to decrease level of synchrony, by simply increasing the phase difference between forcing phase coefficients in the formulas of contact rate β. Clearly, the level of synchrony between two subpopulations are the worst when the two fluctuations are in antiphase (as figure 3.1). When the phase difference between oscillations increases, the desynchronizing effect on population dynamics of the subpopulations augments. This declines the ability of disease extinction.

Moreover, as the result above (figure 3.4) and the global extinction rate (figure 3.3), our results, along with those of Bolker (1995) [3] and Heino (1997) [23], stress the asynchrony that is the main factor that reduces the probability of global extinction. Disease is always available in metapopulation if and only if at least one subpopulaiton is not extinct.

Our finding has specified the two main factors influencing the persistence ability of an infectious disease. One factor is transmission characteristics of the infectious disease and the other is interplay between mixing subpopulations in metapopulation. The interaction between the disease persistence and the spatial heterogeneity becomes a major key to unlock questions about infectious disease in epidemiology. This result takes a large part in epidemic disease persistence domain that has being exploited in scientific epidemically research works. We gained a robust understanding of how disease extinction is affected by local factors such as spatial heterogeneity, demographic asynchrony and seasonality, as well as mixing factors such as migration, disease transmission between hosts and pathogens. Lastly, we also highlighted recolonization effects. It is like rescue for disease. Because of connection between subpopulation, individuals can go everywhere. Subpopulation is quickly re-infected althought the disease has became extinct. Thus, the disease rescus makes local extintions difficult to extend into global extinctions.

As a matter of the fact of coupling strength among subpopulations in metapopulation, we proved that the extinction of the disease in the metapopulation is not only computed by a exponential survival function over time, but also a concave function for the coupling rate. In addition, the disease extinction in metapopulation is minimum when the coupling rate between subpopulations is just medium. This finding is similar to those of Huffaker(1958) [26], Holyoak and Lawler(1996) [24], and Yaari et al. (2012) [39] when they exhaustively explored the disease persistence behavior of many different metapopulation models. And our result one more time affirms that the disease persistence and the interaction in metapopulation models are significant when the interaction strength ρ is from 10^{-3} to 0.1 [29].

To summarize, we have built successfully a sinusoidally forced SEIR stochastic metapopulation model. This model is like a physical system of coupled oscillators. We have pointed out that spatial synchronization consistently and predictably makes extinction risk increase by using the model 0 where all the subpopulations have the same population size N and there is no explicit spatial distance. So, it's good for the future, we can continue this work with different population size of each subpopulation and different spatial distance between subpopulations and then, create synchronous metapopulations that optimize vaccination policies.

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] B. M. Bolker and B. T. Grenfell. Chaos and biological complexity in measles dynamics. *Proc Biol Sci*, 251(1330):75–81, Jan 1993.

[5] 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.

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

[7] 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.

[8] 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.

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

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

[11] 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.

[12] 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.

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

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

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

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

[17] 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.

[18] 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.

[19] 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.

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

[21] D. He and L. Stone. Spatio-temporal synchronization of recurrent epidemics. *Proc Biol Sci*, 270(1523):1519–1526, Jul 2003.

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

[23] 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.

[24] M. Holyoak and S. P. Lawler. Persistence of an extinction-prone predator-prey interaction through metapopulation dynamics. *Ecology*, pages 1867–1879, 1996.

[25] 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.

[26] C. B. Huffaker. Experimental studies on predation: dispersion factors and predator-prey oscillations. *Hilgardia*, 27:343–383, 1958.

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

[28] 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.

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

[30] David G. Kleinbaum. *Survival analysis*. 2005.

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

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

[33] 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.

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

[35] 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.

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

[37] 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.

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

[39] 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.

[40] 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 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:

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

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

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

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 [29], 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 $y_1^k \ldots y_{l_k}^k$ 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 $t+\delta t$:

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

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 [29]. 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 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

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

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

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

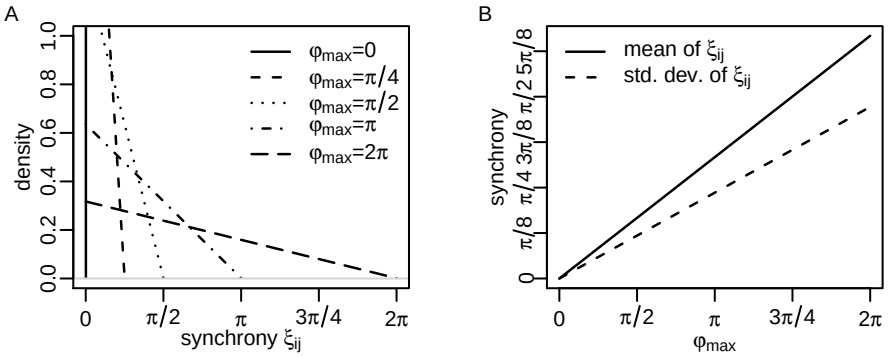

Figure 7.1: Synchrony in the case of model 0. (A) distribution of synchrony ξ_{ij} for various values of φ_{max}. (B) mean and standard deviation of the distribution of ξ_{ij} as functions of $\varphi_{\max}$.

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.