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

31/10/2014 Version 1.1

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 meta-population 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' [38]), we estimate the global extinction rate which represents the global persistence of disease in the meta-population. We find how bigger the forcing phase heterogeneity becomes, and how larger the global persistence gets.

Keywords: SEIR model, Markov chain, Monte-Carlo simulation methods, spatial synchronization, disease persistence, meta-population.

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, 37, 39]. 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) [42] 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]. It is the reason that the ecological theory of metapopualtion can point out that the persistence of a population depends on the dynamics of migrations between the different sub-populations of the metapopualtion [19, 20, 13, 31]. The "colonization" caused by migrations brings infection to an uninfected subpopulation and we call infected individuals colonizers.

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, 41]. 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, 34]. 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 [40], 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 [35, 36] 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 struc-

ture and the stochastic version of the model. Finally, we introduce the characterization of global persistence in the metapopulation based on the measles characteristics.

2 MATERIAL AND METHODS

2.1 MATERIAL

2.1.1 Deterministic model for many subpopulations

The standard SEIR model (susceptible-exposed-infective-recovered) has been strongly developed for the dynamics of directly infectious disease [3]. For disease-based metapopulation models, we give here a suitable new version of the SEIR equation that would be as follows:

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]:

$$\frac{dS_i}{dt} = \mu N_i - \lambda_i S_i - \mu S_i \quad (1)$$

$$\frac{dE_i}{dt} = \lambda_i S_i - \mu E_i - \sigma E_i \quad (2)$$

$$\frac{dI_i}{dt} = \sigma E_i - \mu I_i - \gamma I_i \quad (3)$$

$$\frac{dR_i}{dt} = \gamma I_i - \mu R_i \quad (4)$$

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. In a case the infectious contact rate is constant, the equilibrium values of the variables S, E, I and R can be expressed analytically (see appendix). 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] :**

$$\lambda_i = \sum_j \rho_{ij}\kappa_j \log\left[1 - \sum_{k=1}^{M}\left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right] \quad (5)$$

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 C_j comes from 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

$$\lambda_i = -\kappa_i \log(1 - \frac{I_i}{N_i} \times c_{ii}) \tag{6}$$

Consider that 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 :

$$\kappa_i(t) = \kappa_{i0} \left[1 + \kappa_{i1} \cos\left(\frac{2\pi t}{T} + \varphi_i\right)\right] \tag{7}$$

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.

2.1.2 Stochastic model for many subpopulations

In order to study the persistence of the disease, we must consider a stochastic version of the model [28, 32, 33]. We use for that a population-based time-to-next-event model based on Gillespie's algorithm [14]. Table 1 lists all the events of the model, occurring in subpopulation i.

Table 1 – Events of the stochastic version of the model of equations 1-4, 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	$I_i \leftarrow I_i - 1$
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 $R_i \leftarrow R_i + 1$

2.1.3 Spatial structures

A metapopulation is 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 order to identify clearly the causes of observed phenomena, these two aspects will be modeled distinctly. In this article, our null model (model 0) will be a metapopulation without any explicit spatial distance (all the subpopulations are at the same distance from each other) and where all the subpopulations 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:

$$\rho_{ij} = \rho, \qquad 0 \leqslant \rho \leqslant 1, \qquad \forall i, \forall j. \tag{8}$$

The structure of this metapopulation is thus characterized by 3 parameters: (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 between these subpopulations.

2.2 METHOD

2.2.1 Global persistence in a metapopulation

In order to examine questions of interaction between disease transmissibility and phase of seasonal forcing, we start in this section by studing the stochastic SEIR model in a metapopualtion of n subpopulations. For this meta-population, we observe the disease extinction in time due to spatial synchrony/asynchrony that are influenced by phase difference in seasonal forcing. To create the phase difference, we change the value of the forcing phase for each city. In this experience, we use a parameter φ_{max} in radian that runs in the interval from zero to π. With each value of φ_{max}, based on n the number of subpopulations in the metapopulation, we divide the interval [0, φ_{max}] into a set of (n-1) equal samples, so the value of the forcing phase of the i^{th} subpopulation is correspondent to i^{th} value in the set. We call φ_{max} asynchrony parameter.

The persistence of disease in the metapopulation was characterized by fitting an exponential survival model [8, 30] on data simulated by the 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. For our metapopulation of n subpopulations, to do so we run first m independent simulations of our stochastic model. We calculate then 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 persistence rates χ:

$$M(t) = \exp(-\chi t) \tag{9}$$

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

To assess global extinction rate as well as global persistence level in metapopulation, here, we use the parametric survival model for the exponential distribution (R package '*survival*' [38]). Due to that, we can capture one of the most important features of stochastic systems in spatial structure : its global persistence of disease.

2.2.2 Characterization of synchrony

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

$$\delta_{ij} = |\varphi_i - \varphi_j| \bmod 2\pi \tag{10}$$

where φ_i and φ_j are the phases of the contact rates (equation 7) 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

$$\xi_{ij} = \left|1 - \frac{\delta_{ij}}{\pi}\right|. \tag{11}$$

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

$$<\delta_{ij}> = <\delta_{ji}> = 2\varphi_{\max} \sum_{k=1}^{n-1} \frac{(n-k)k}{(n-1)n^2} = \frac{n+1}{3n}\varphi_{\max} \tag{12}$$

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

$$<\xi_{ij}> = <\xi_{ji}> = 1 - \frac{n+1}{3n}\frac{\varphi_{\max}}{\pi} \tag{13}$$

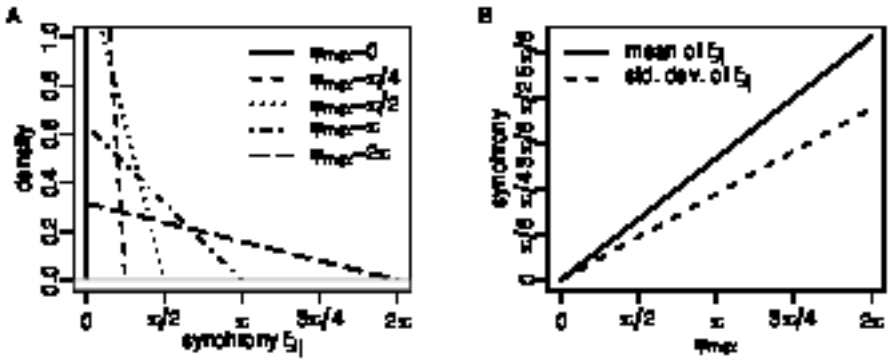

Figure 1 – Synchrony in the case of model 0. (A) distribution of synchrony ξ_{ij} for various values of φ_{max}. (B) mean and standard deviation of the distribution of ξ_{ij} as functions of $\varphi_{\max}$.

and thus

$$\lim_{n\to\infty} <\xi_{ij}> = 1 - \frac{\varphi_{\max}}{3\pi} \tag{14}$$

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 1.

2.3 Plan of experience

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 four big concerns that we must verify.

2.3.1 Quantifying disease persistence in the simplest metapopulation

Firstly, we are interested in the correlation between global disease persistence time when the forcing phase of the infection rate β in each subpopulation alters and the recolonization in metapopulation. In order to simplify this concern, we start with the metapopulation of two subpopulations. To have the two subpopulations in synchrony, we choose $\varphi_{max}= 0$. It means that the forcing phase of $subpopulation_1$ and $subpopulation_2$ are the same phase, $\varphi_1 = \varphi_2 = 0$. Thereby, the fluctuations of the contact rate $\beta(t)_1$ and $\beta(t)_2$ are in the same annual sinusoidal form. In contrast, to have the two subpopulations in asynchrony, we create the phase difference or the phase lag between the subpopulations. It means that we have created a metapopulation in heterogeneity. Here, we choose $\varphi_{max}= \pi/2$ and $\varphi_{max}= \pi$. The phase difference between the forcing phase of $subpopulation_1$ and $subpopulation_2$, for $\varphi_{max}= \pi/2$, is $\varphi_1 = 0$ and $\varphi_2 = \pi/2$, then for $\varphi_{max}= \pi$, is $\varphi = 0$ and $\varphi_2 = \pi$, so the fluctuation of $\beta(t)_1$ is phase lag for $\beta(t)_2$'s.

Based on φ_{max}, we are successfully buiding a complex disease metapopulation mode.

2.3.2 Quantifying global persistence and asynchrony level φ_{max}

We will exploit more strongly the role of asynchrony φ_{max} for global disease persistence time in metapopulation. φ_{max} is an important parameter that we use to break fixed points as well as first fixed points at begining moments between subpopulations. Our goal is to examine global persistence for each value of φ_{max} in the metapopulation. Simply, we base on survival regression model for global peristence curve at each different value φ_{max} to estimate its global extinction rate. We start also with the metapopualtion of n subpopulations given and φ_{max} from 0 to π. It means that we increase level of phase difference from the $subpopulation_1$ to $subpopulation_n$. We deploy m the number of different simulations. Then, we use survival analysis to quantify persistence level at each value of φ_{max} with confidence interval to 95%.

2.3.3 Influence of other parameters on global disease persistence

This last experience is how the population size of subpopulation, the coupling strength between subpopulations and the number of subpopulations in the metapopulation affect the disease persistence as well as the relation between the global disease persistence and the asynchrony parameter φ_{max}. In order to present the effect of the third parameter on the correlation of the two first parameters (global persistence level and φ_{max}), we will use a gradient coefficient for the trajectory of the correlation between global persistence and φ_{max}.

2.3.4 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 [**?**, 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 :

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

parameter	**description**	**value**
μ	birth and death rate per day	$1/(70*365)$
β_0	mean value of the infection rate β per day	$1250/365$
β_1	amplitude of the infection rate β	0.15
γ	recovery rate per day	$1/8$
σ	average exposed duration per day	$1/5$
ρ	coupling rate	$10^{-4} - 1.0$
φ_{max}	synchrony parameter in radian	$0 - \pi$
N	population size of subpopulation (individual)	$5000 - 500,000$
n	number of subpopulation	$2 - 30$
t_{max}	simulation time (year)	50

Following the table in detail about the convenient values of parameters, we will use them through 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 Quantifying disease persistence in the simplest metapopulation

We start quantifying global disease persistence in the metapopulation of two subpopulations by examining degree of synchronization between infected individuals in two connected subpopulations. This is the simplest metapopulation, for which we can view easily the influence of the level of synchrony on the disease persistence and recolonization of disease in the metapopulation over time due to the asynchrony parameter φ_{max}.

We chose here the metapopulation of two subpopulations, $N_1 = N_2 = 300,000$, the rate of coupling $\rho = 0.01$, the number of simulation repeats $m = 100$, and $\varphi_{max} = \{0, \pi/2, \pi\}$. Now, we have 100 metapopulations, we gather the first time where the metapopualtion gets global extinction. We have three Kaplan-Meier survival curves for each value of φ_{max} as figure 2.

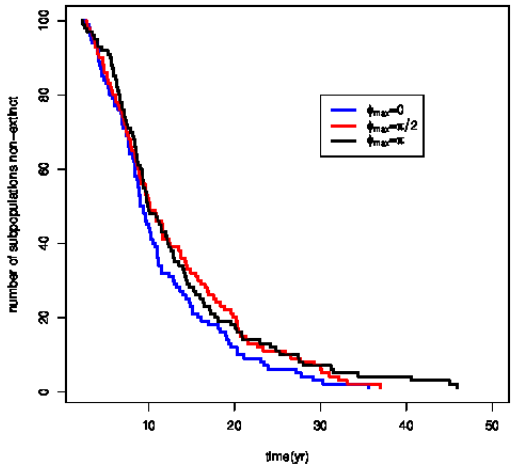

Figure 2 – Kaplan-Meier survival curves for disease persistence after 100 different simulations of $\varphi_{max} = 0$, $\varphi_{max} = \pi/2$ and $\varphi_{max} = \pi$. The blue survival curve for $\varphi_{max} = 0$, the red survival curve for $\varphi_{max} = \pi/2$ and the black curve for $\varphi_{max} = \pi$. The disease persistence time of $\varphi_{max} = 0$ is the shortest. This of $\varphi_{max} = \pi$ is the longest.

The phase of forcing of the $subpopulation_1$ is always fixed at 0, but that of the $subpopulation_2$ increases from 0 to π. It means that, in the first experience, $\varphi_{max} = 0$, the two subpopulations are in synchrony with all beginning conditions. The disease persistence time is the shortest. Then, the two subpopulations become asynchronous when

$\varphi_{max} = \pi/2$ or π. The symmetry of fixed points is just broken at the starting moment. It is the reason why the level of synchrony of the metapopulation decreases. Additionally, we find that the value of the asynchrony parameter φ_{max} change according to increasing tendency, the phase difference between the fluctuations of the two infection rates β augments also, and the global persistence time in the metapopulation thus goes up. When $\varphi_{max} = \pi$, the two subpopulations are in antiphase. The balance points at the begining moment of $subpopulaiton_1$ are inverse to that of $subpopulation_2$, it accounts for why the two subpopulations take a long duration to obtain the balance state. This is the most difficult case to find global extinction, the persistence time is the longest. Moreover, theses results are explicated by recolonization between the two subpopulations. When the disease no longer exists in one subpopulation at the moment t. If at this moment, the neighbouring subpopulation obtains also the extinction, so we have a global extinction in the metapopulation and the disease is entirely extinct. However, in a metapopulation, due to different factors for migration between subpopulations, we hardly see extinction at the same time in all subpopulations in first duration. The infected in other subpopulation migrates to the extinct subpopulation, so the disease is active. It is the reason why the disease exists for long term. In short, the level of synchrony between subpopulation is stronger, metapopualtion is easier to find global extinction. Make all subpopulations synchronize is the easiest way at which disease goes to extinct.

Additionally, we have done also experiences for local fluctuations of subpopulations. Here, local fluctuations is called local dynamics or local noises. It is an important factor that can increase local extinction numbers but make global extinction rate reduce as the following figure 3.

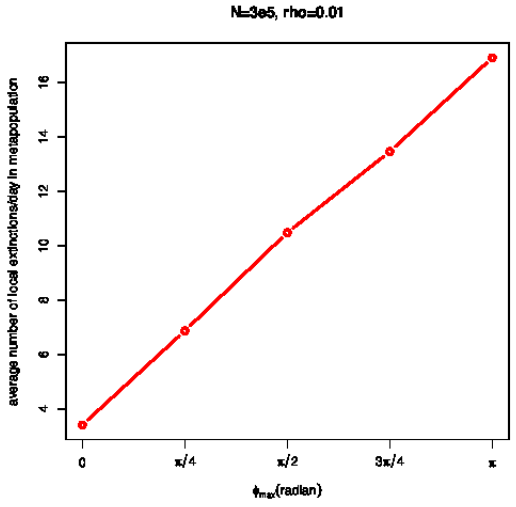

Figure 3 – Average number of local extinctions in the metapopulation of two subpopulations and the level of synchrony after 100 different simulations. In this figure, both the level of asynchrony increases from 0 to π and the number of local extinctions augments.

As illustrated above, global persistence rate is inversely proportional to global extinction rate. The global persistence

rate increases, it is synonymous with a reduction of global extinction rate. Thus, from the figure **3**, the number of local extinctions in the metapopulaiton scales inversely global extinction rates. Because of the local noises, subpopulations become extinct but are always recolonized in short time. Local correlation between subpopulations is soul of synchrony in metapopulation. We use local correlation to measure level of synchrony and through that we estimate probabilities of global extinction.

3.2 Quantifying global persistence and asynchrony level φ_{max}

As mentioned about, we use survival regression model for global peristence curve to estimate its global extinction rate and to quantify persistence level for each value φ_{max} with confidence interval 95%. The result below is pointed for the metapopulation of eight subpopulations, the population sise of each subpopulation $N = 3 \times 10^5$ and the coupling rate $\rho = 0.1$ as following figure 4.

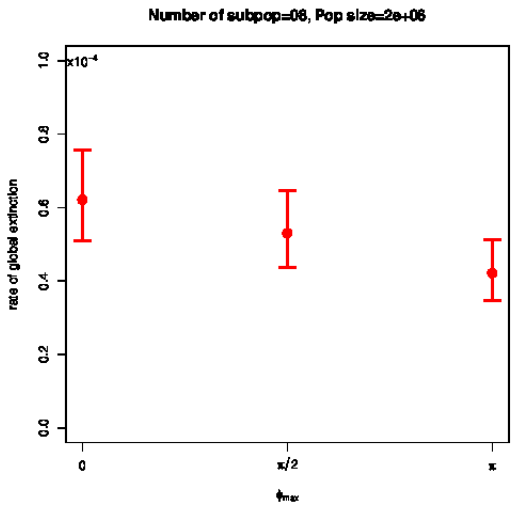

Figure 4 – Estimated level of global disease persistence in the metapopulation of the eight subpopulations after 100 different simulations $N = 3 \times 10^5$, coupling rate $\rho = 0.1$. Here, with 95% confidence interval, blue lines are confidence intervals for persistence rate of each value of φ_{max}. This intervals are limited by lower and upper confidence limits.

This figure 4 shows to us that the amplitude of the confidence intervals for each value of φ_{max} are quite far to each other. Furthermore, it rises robustly when φ_{max} runs from 0 to π. The phase difference strongly influences disease persistence time as well as global disease persistence level. Figure 4 indicates the trend of level of persistence with increasing level of asynchrony. The asynchrony between subpopulations is the main reason why the infectious disease has never been extinct.

In order to press the strong dependence of the global disease persistence to the asynchrony level parameter ϕ_{max}, here we introduce the results of the metapopulations of many different subpopulations. First, we study levels of disease persistence in metapopulations of six, ten and twenty subpopulations with population size of each subpopulation $N =$

3×10^5 and the coupling rate $\rho = 0.01$. We obtain also the same result as the metapopulation of eight subpopulations (figure 5), when the level of asynchrony φ_{max} increases, so the persistence of disease augments. On the other hand, the probability of disease persistence also alters with the different numbers of subpopulation. In this result (figure 5), the global persistence of the six subpopulations is minimum and that of the twenty subpopulations is maximum.

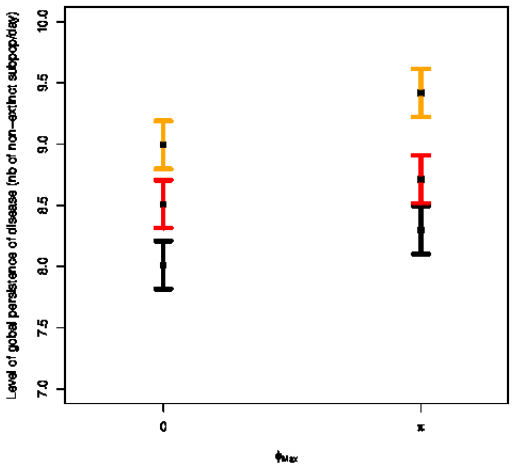

Figure 5 – Estimated level of persistence in the metapopulation of six, ten and twenty subpopulations after 100 different simulations. Here, with 95% confidence interval, black, red, and orange lines are for six, ten and twenty subpopulations, respectively. These intervals are limited by lower and upper confidence limits.

3.3 Influence of other parameters on global disease persistence

3.3.1 Number of subpopulation in a metapopulation

In this part, we developed the relation between the disease persistence and the number of subpopulations in a metapopulation. We performed simulations with metapopulations from three to 30 subpopulations, population sizes of each subpopulation $N = 10^5$ and $\varphi_{max} = \pi$. The result as following figure **6**.

The result exhibits to us that the number of subpopulations strongly has an influence for the disease persistence time. The global persistence level of an infectious disease in a metapopulation increases when the number of subpopulations in this metapopulation increases. As this is a coupling model, there are an interaction among subpopulations. The number of subpopulation increases, influence degree of neighbouring subpopulation on $subpopulation_i$ is more and more complex and strong. The probability of disease recolonization of $subpopulation_i$ also rises, thereby the disease persistence increases.

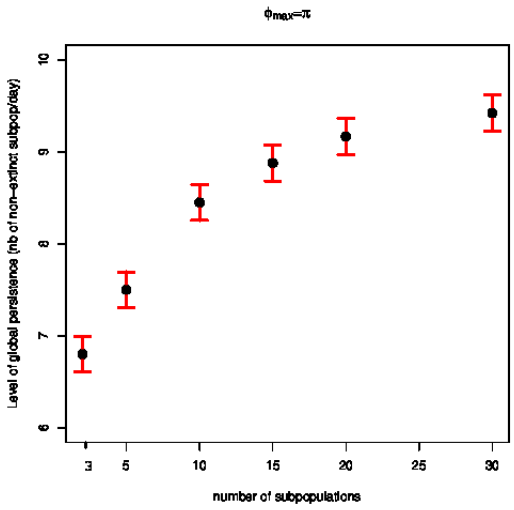

Figure 6 – Level of disease persistence in the metapopulation of many subpopulations with $\varphi_{max} = \pi$ and $\rho = 0.001$. The number of subpopulation is from three to 30.

3.3.2 Population size of each subpopulation in a metapopulation

In the metapopulation of 06 subpopulations, we implemented experiences with different population sizes of subpopulation. We performed 100 different simulations for the metapopulation of 06 subpopulations in which all subpopulations has the same population size N. We set N from 5×10^3 to 5×10^5 individuals. The result (figure 7) affirms that the population size influences strongly the global persistence of an infectious disease in a metapopulation. The number of individuals in a subpopulation grows, so the global persistence of disease increases also.

From the result (figure 7), the population size and the global persistence of disease are directly proportional. This is effect of the inscrease of population size that is reason for that the number of infectives present augments, and therefore both the demographic stochasticities and probability of global extinction decrease.

3.3.3 Coupling rate

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 metapopualtion of five subpopulations with the population size $N = 10^5$ 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 gradient coefficient for the relation between persistence level of disease and level of asynchrony φ_{max}, in condition φ_{max} belonging to the set $\{0, \pi/2, \pi\}$. Now, we have result as following figure 7.

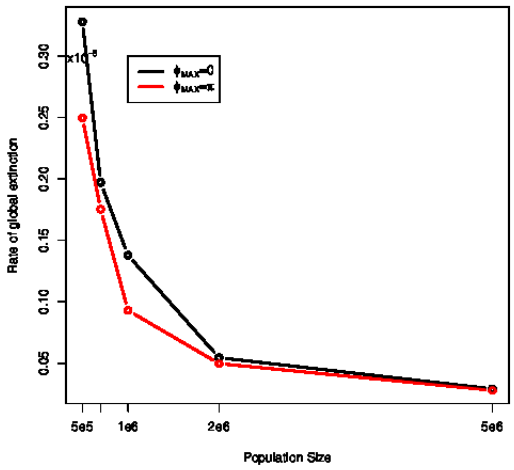

Figure 7 – Relation between level of global disease persistence and the population size of each subpopulation. With 95% confidence interval, the red points are lower limites and the green points are upper limites. Here, the number of subpopulation is six, the coupling rate ρ is 0.1, the level of asynchrony φ_{max} is $\{0\}$ and the population size of each subpopulaiton N is in $\{5 \times 10^3$,10×10^3,100×10^3, 300×10^3,$500 \times 10^3\}$.

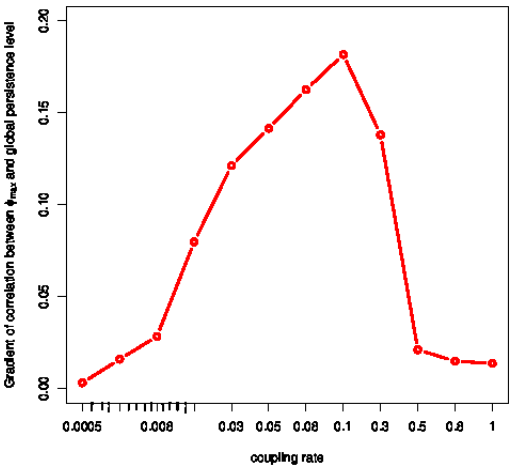

Figure 8 – Correlation between coupling rate and gradient of level of synchrony and persistence rate in the metapopulation of five subpopulation. Here, the coupling rate ρ is in $\{0.0005$,0.001,0.008,0.01,0.03,0.05,0.08,0.1,0.3,0.5,0.8,$1\}$, the level of synchrony φ_{max} in $\{0$,$\pi/2$,$\pi\}$ and the population size of each subpopulaiton N=1e5.

When the coupling rate is small from 0.0005 to 0.008, the gradient of the correlation between φ_{max} and the global persistence increases very slowly. However, the persistence level augments in a sudden way when the coupling rate changes from 0.01 to 0.1. Lastly, the gradient strongly decreases when the coupling rate is so robust from 0.3 to 1.0. Based on this figure, the global disease persistence in a metapopulation is one humped function for the coupling rate. The medium coupling rate (from 0.01 to 0.1) maximizes disease persistence in metapopulation. As in the case of the small and average 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. However, this trend of global persistence with increasing coupling rate, is not right any more when the dispersal rate is strong. The duration of persistence falls, because the metapopulation has tendency to become one big population. In this case, the phase difference or the recolonization among subpopulations are no longer significant.

4 Discussion and Conclusion

We successfully have built a version for the susceptible-infectied-recovered stochastic metapopulation model. 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) [35]. 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. 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 persistence 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.

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 2). When the phase difference between oscillations increases, the desynchronizing effect on population dynamics of the subpopulations augments. This enhances disease persistence time though the global extinction rate is reduced. Moreover, as the result above (figure **3**), in the local shape, our result, along with those of Bolker (1995) [3] and Heino (1997) [23], stress the number of local extinctions being inversely proportionnel with the global extinction rate in a metapopulation. When the level of synchrony is at a reduction and the global persistence time gets an increase, the global extinction rate of metapopulation goes down and the number of local extinction goes up. Due to the result about local extinction, we also affirm that 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 persistence 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 persistence of the disease in the metapopulation is not only a exponential survival function over time, but also a humped function for the coupling rate. In addition, the global disease persistence in metapopulation is maximum 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) [41] 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. 1, 2.1.1

[2] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992. 2.1.1

[3] B. Bolker and B. Grenfell. Space, persistence and dynamics of measles epidemics. *The Royal Society*, 348:309–320, 1995. 1, 2.1.1, 2.1.3, 2, 4

[4] B. M. Bolker and B. T. Grenfell. Chaos and biological complexity in measles dynamics. *Proc Biol Sci*, 251(1330):75–81, Jan 1993. 4

[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. 1

[6] M. Choisy and P. Rohani. Changing spatial epidemiology of pertussis in continental usa. *Proc Biol Sci*, 279(1747):4574–4581, Nov 2012. 1, 2

[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. 1

[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. 1, 2.2.1, 2

[9] G. S. Cumming. The impacts of low-head dams on fish species richness in wisconsin, usa. *Ecological Applications*, 14:1495–1506., 2004. 1

[10] S. Dey and A. Joshi. Stability via asynchrony in drosophila metapopulations with low migration rates. *Science*, 312(5772):434–436, Apr 2006. 1

[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. 1, 4

[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. 1

[13] Ilkka Hanski; Oscar E Gaggiotti. *Ecology, Genetics and evolution of metapopulations*. 2004. 1, 2.1.3

[14] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977. (document), 2.1.2, 2.3.4

[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. 1

[16] B.T. Grenfell, B. M. Bolker, and A. Klegzkowski. Seasonality and extinction in chaotic metapopulation. *The royal society*, 259:97–103, 1995. 2.1.1

[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. 1

[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. 1, 2.2.1, 2.3.4

[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. 1

[20] I. Hanski. Metapopulation dynamics. *Nature*, 396, 1998. 1

[21] D. He and L. Stone. Spatio-temporal synchronization of recurrent epidemics. *Proc Biol Sci*, 270(1523):1519–1526, Jul 2003. 4

[22] healthmap.org. Measles reemerges in vietnam, 2014. 1

[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. 1, 4

[24] M. Holyoak and S. P. Lawler. Persistence of an extinction-prone predator-prey interaction through metapopulation dynamics. *Ecology*, pages 1867–1879, 1996. 4

[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. 1

[26] C. B. Huffaker. Experimental studies on predation: dispersion factors and predator-prey oscillations. *Hilgardia*, 27:343–383, 1958. 4

[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. 2

[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. 2.1.2, 2.2.1, 2, 4

[29] M. J. Keeling and P. Rohani. *Modeling Infectious Diseases in humans and animals*. Princeton University Press, 2008. 1, 2.1.1, 2, 4, 6.3

[30] David G. Kleinbaum. *Survival analysis*. 2005. 2.2.1

[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. 1, 2.1.3

[32] A. L. Lloyd. Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theor Popul Biol*, 60(1):59–71, Aug 2001. 2.1.2

[33] E. Renshaw. *Modelling biological populations in space and time*, volume 11. Cambridge University Press, 1993. 2.1.2

[34] P. Rohani, D. J. Earn, and B. T. Grenfell. Opposite patterns of synchrony in sympatric disease metapopulations. *Science*, 286(5441):968–971, Oct 1999. 1

[35] 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. 1, 4

[36] G. Rozhnova1, A. Nunes, and A. J. McKane. Impact of commuting on disease persistence in heterogeneous metapopulations. 2013. 1

[37] 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. 1

[38] T. M. Therneau. *A Package for Survival Analysis in S*, 2014. R package version 2.37-7. (document), 2.2.1

[39] 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. 1

[40] WHO. Reported measles cases with onset date from oct 2013 to mar 2014, 2014. 1

[41] 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. 1, 4

[42] 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. 1, 6.3.1

5 Appendix : equilibrium values of the system 1–4

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

$$\frac{dS_i}{dt} = \mu N_i - \lambda_i S_i - \mu S_i \tag{15}$$
$$\frac{dE_i}{dt} = \lambda_i S_i - \mu E_i - \sigma E_i \tag{16}$$
$$\frac{dI_i}{dt} = \sigma E_i - \mu I_i - \gamma I_i \tag{17}$$
$$\frac{dR_i}{dt} = \gamma I_i - \mu R_i \tag{18}$$

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:

$$S_i^* = N_i \frac{(\gamma+\mu)(\sigma+\mu)}{\beta\sigma} \tag{19}$$
$$E_i^* = N_i \mu \left(\frac{1}{\sigma+\mu} - \frac{\gamma+\mu}{\beta\sigma} \right) \tag{20}$$
$$I_i^* = N_i \mu \frac{\beta\sigma - (\sigma+\mu)(\gamma+\mu)}{\beta(\sigma+\mu)(\gamma+\mu)} \tag{21}$$
$$R_i^* = N_i - S_i^* - E_i^* - I_i^* \tag{22}$$

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

$$S_i^* = N_i \frac{1}{R_0} \tag{23}$$
$$E_i^* = N_i \frac{\mu\sigma}{R_0} (R_0 - 1) \tag{24}$$
$$I_i^* = N_i \frac{\mu}{\beta} (R_0 - 1) \tag{25}$$
$$R_i^* = N_i - S_i^* - E_i^* - I_i^* \tag{26}$$

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

$(S_i^*, E_i^*, I_i^*, R_i^*) = (N_i \frac{1}{R_0},\, N_i \frac{\mu\sigma}{R_0}\,(R_0 - 1),\, N_i \frac{\mu}{\beta}(R_0 - 1),\, N_i(1 - \frac{1}{R_0} - \frac{\mu\sigma}{R_0}\,(R_0 - 1) - \frac{\mu}{\beta}(R_0 - 1))$.

6 Appendix: derivation of the equation 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 1. *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 2. *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.

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

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

- 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...

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

Overall, we have the formule of the infection force as follows:

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

Thus, in the forcing case, the contact rate is in the sinusoidal form as follows :

$$\kappa_i = \kappa_{i0}(1 + \kappa_{i1} cos(\frac{2\pi t}{T} + \phi_i)$$

Appendix: supplementary materials

Here, we present some results of the simulation with the conditions as the same for all initial values of variables, for all parameters of the metapopulation of two subpopulations without the synchrony parameter φ_{max}. We let φ_{max} belong to the set $\{0, \pi/2, \pi\}$. The results are shown in Figure 9, 10 and 11.

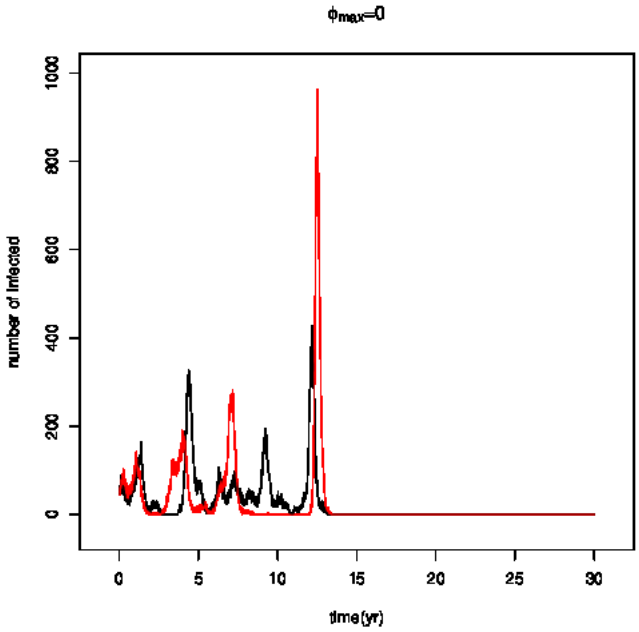

Figure 9 – Metapopulation of two subpopulations in phase.
Two stochastic SEIR metapopulation model with $N_1 = N_2 = 3 \times 10^5$, $\rho = 0.001$, $\varphi_{max} = 0$ and simulation time = 30 years. The black line for $subpopulation_1$ and the red line for $subpopulation_2$. Because we start doing simulation with the same all beginning conditions, so in the first period, the two subpopulations similarly fluctuate. Because of high level of synchrony, global extinction occurs before 10th year.

With $\varphi_{max} = 0$, the two subpopulation are synchronous about all starting conditions (values of all parameters, of all variables), so we gain the synchrony at the starting steps of simulation (fig 9). However, here is stochastic model, so the similar fluctuations are broken, the decorrelation between two dynamics rises. This bring a closeness to disease dynamics in reality. The metapopulation dynamics runs from global synchronization to desynchronization and the global synchronization in subpopulation decays over time because of coupling strength of community structure [42]. About fifth year, the $subpopulation_1$ gains local extinction. But, disease comes back after a short period of time by transfer of infected individuals from the $subpopulation_2$. However, due to the high degree of synchrony between the two subpopulations, the metapopulation obtains a global extinction. Disease can not return into this metapopulation.

Next, with $\varphi_{max} = \pi/2$, the two subpopulation are out of phase (figure 10). We give the same values for all parameters, for all variables, but the different forcing phases, so we gain the different contact rates. The phase difference occurs, the synchrony structure is broken from the start moment and the number of infected is governed

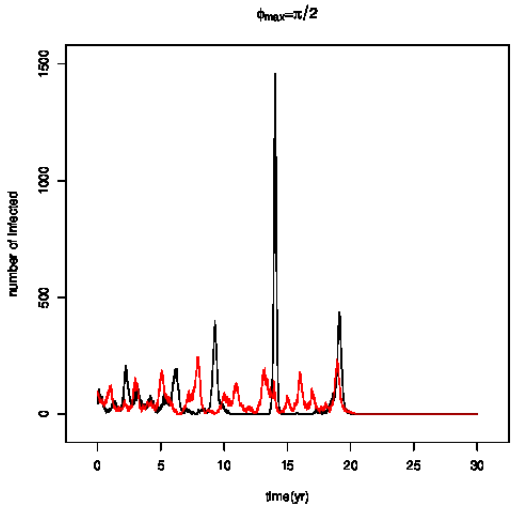

Figure 10 – Metapopulation of two subpopulations in phase quadrature.
Two stochastic SEIR metapopulation model with the same value for all parameters and all variables, except $\varphi_{max} = \pi/2$. The black line for $subpopulation_1$ and the red line for $subpopulation_2$. The start points of the two wavelettes are out of phase and all peaks are also always out of phase with each other. The degree of synchrony is worse than with $\varphi_{max} = 0$. So, the time of disease persistence is longer.

by each other. In this case, the $subpopulation_1$gains two times local extinction. But, this subpopulation is reinfected because disease is still availlable in the $subpopulation_2$. The metapopulation gets global extinction after 15^{th} year. The time of disease persistence is longer than the case $\varphi_{max} = 0$ since the high degree of synchrony between the two subpopulations with $\varphi_{max} = \pi/2$ is smaller than with $\varphi_{max} = 0$.

Lastly, in the figure 11, we have to say that this is the worst, global extinction have occurred latest although the first subpopulation has gotten many times local extinction. Here, the level of synchrony is the lowest, the phase difference is 180 degrees (π radian). The time of persistence is the largest. However, the phase difference with 180 degrees is step by step damped. There are some intervals where the distance between the number of the infected of the two subpopulations is very close to each other. So, we can see the small phase differences between peaks of the fluctuations. The more the phase difference increases, the more the disease persistence time increases.

In conclusion, local extinctions are desynchronizing. This reduces synchrony among subpopulations and promotes recolonization, and thus causes disease persistence in long term. In the other hand, view the figures, we can find multiyear oscilations in measles metapopulation.

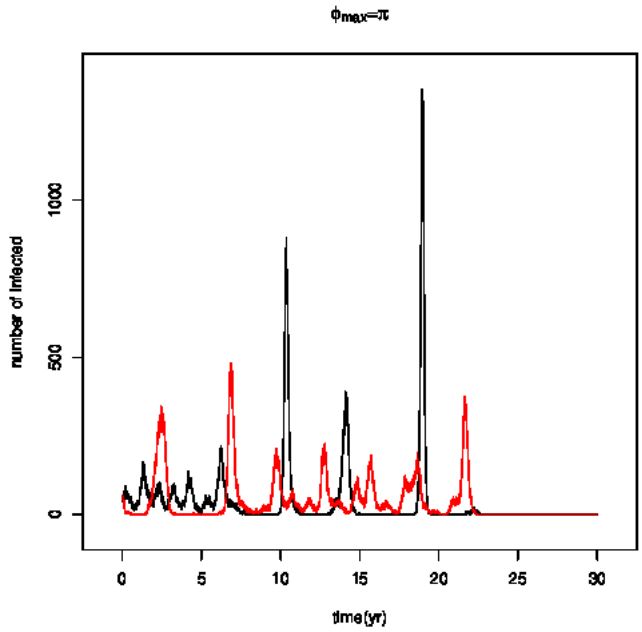

Figure 11 – Metapopulation of two subpopulations in antiphase.
Two stochastic SEIR metapopulation model with the same value for all parameters and all variables, except $\varphi_{max} = \pi$. The black line for $subpopulation_1$ and the red line for $subpopulation_2$. The start points of the two wavelets are in antiphase and all peaks are also always out of phase with each other. This is the worst case, the metapopulation finds global extinction late.