Modélisation, Simulation multi-niveau pour l'optimisation de politiques de vaccination

18 Aout 2014

September 8, 2014

Thesis direction: Pr. Yann Chevaleyre (PR, Université Paris 13) Dr. Jean-Daniel Zucker (DR1, UMI UMMISCO, IFI-MSI, IRD)
Co-supervision: Dr. Marc Choisy (CR1, MIVEGEC, NIHE, IRD) Dr. Vu Dinh Thiem (NIHE)

1 Introduction

This thesis gives birth in the context where many recent events occur in the world, for epidemiologie : SRAS in 2003, avian influenza in 2004 or swine flu in 2009, for natural disaster : earthquake - Haiti, Chile, China-, tsunamis - Indian Ocean in 2004, France in 2010, Japan in 2011-floods,..., storm insects (cloud of locust), in particular, infectious disease Ebola being making the entier world worry. These events have revealed important risks of environment. The rapid spread of the flu virus recently illustrates how it is important to anticipate the epidemiological phenomena to better manage when they occur. Moreover, for vaccination policies, the mass policy (vaccinate the maximum number of children before certain age) is the oldest (started from the 1950s in the rich countries) and is now the most used. The problem of this vaccination policy is too expensive, ineffective and quite impossible to implement in poor countries, especially in Africa as both financial and logistical problems. (e.g. the WHO project of extinction measles in Vietnam before 2012 is failed). Therefore, OPTIMIZE the vaccination policies in Artificial Intelligence in order that these policies may become more effective, less expensive, and take into account the spatial dimension, is the goal of this thesis.

To carry out this goal, the first task is to make a state of the art covering the different epidemic models being used in the field of epidemiological modeling. In the second step, we will focus on design or expand a representation language that allows a modelisator to express the goals of model simulations as well as constraints related to them. In the third step, we will set disease population in spatial context, then study its local/global disease persistence, disease periods and the influence of space on the disease persistence. Finally, we will perform vaccination policies in a metapopulation and optimize them.

2 Epidemic models

2.1 State of the art

We know that, a mathematical model uses the language of mathematics in order that the description of a system becomes more refined and more precise. In epidemiology, this model permits us to translate between behavior at various scales, or extrapolate from a known set of conditions to another. More robustly, the mathematical models allow us to predict the epidemic dynamics of population-level, state of disease under individual as well as under population, spread of disease and impact of vaccination on the spread of disease infection. These mathematical models in the epidemiologic field are called the epidemic models.

2.1.1 Types of epidemic models

There are two types of epidemic models : deterministic and stochastic

Deterministic

The deterministic model is described by ordinary differential equations. It is used in large populations as in the case of infectious disease such as tuberculosis, mealses, dengue,etc. Here, individuals in the population are assigned to differents subgroups. Each subgroup represents a specific stage of the epidemic. The transition rates from one subgroup to another are derivative equations. The population size in a compartment is differentiable over time and the epidemic process is deterministic.

Stochastic

"All the disease models that have been considered so far are deterministic" is the words of Keeling(2008). It means that we have fixed "clockwork" processes with the same starting conditions, exactly the same trajectory is always observed. It isn't right for dynamics of real pathogens in the real-world. We can not observe exactly the same people becoming infected at exactly the same times. Therefore, the stochastic model comes into being. "Stochastic" here means having a random variable. A stochastic model is a tool for estimating probability distributions of potential outcomes by allowing for random variation in one or more inputs over time. Or, the stochastic models can be an approximation or mimicry of this random or probabilistic element. This model has a important role in special cases, such as when the number of infectious individuals is rather small, when global extinction is probable to occur, when the population size is also small and when an infection has just invaded or is in the trough phase of the epidemic cycle.

Terminology

Here, we give some common terms used in epidemic models.

- M : passively immune infants
- S : susceptibles
- E : exposed individuals in the latent period
- I : infectives
- R : removed with immunity
- β : contact rate such that a susceptible becomes an infectived
- λ : infection rate
- μ : average death rate and at the same time, average birth rate
- σ : average rate of the latent period
- γ : average rate of the infectious period
- N : population size

2.1.2 Epidemic models

At present, there are many epidemic models. We can divide these models into two base types, the type of three base elements and the type of more than three base elements.

Epidemic models with three base elements:

We have the SIR model, the SIR model with births and deaths, the SIS model with births and deaths, and the SIRS model. The first simplest model is the SIR model created by W. O. Kermack and A. G. McKendrick in 1927. The model is fixed with three elements: susceptible S(t), infected I(t) and removed R(t). These three elements are correspondant to its three classes. From the simplest SIR model, in order to accord each infectious disease and real property of disease, scientists have modified it, make it more multiforme. For example, for SIS models, the individuals recover with no immunity to the disease, that is, inviduals are immediately susceptible once they

have recovered. This SIS models are used in sexually transmitted diseases like gonorrhoea because once recovered, the host is once again susceptible to infection.

Epidemic models with more than three base elements:

Here, we also have a lot of epidemic models with more than three base elements, to accord many different disease types, such as the SEIS model, SEIR model, MSIR model, MSEIR model and MSEIRS model. However, in shape of this thesis, we concentrate on the SEIR model (as the figure 1)that consists with many currently infectious diseases in the world. Diseases have a latent or exposed phase, during which the individual is said to be infected but not infectious.

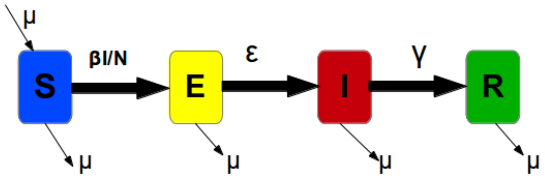

Figure 1: SEIR model

In this model, the host population (N) is divided into four classes : susceptible S(t), exposed E(t), infected I(t) and recovered R(t). We have :

$N = S(t) + E(t) + I(t) + R(t)$

- Classe S(t) : contains the number of individuals not yet with the disease at time t, or those susceptible to the disease.
- Classe E(t) : contains the number of individuals who are in the exposed or latent period of the disease.
- Classe I(t) : contains the number of individuals who have been infected with the disease and are capable of spreading the disease to those in the susceptible category.
- Classe R(t) : contains the number of individuals who have been infected and then removed from the disease, either due to immunization or due to death. Individuals of this classe are not able to be infected again or to transmit the disease infection to others.

2.2 Demographic stochasticity

Demographic stochasticity is still called event-driven stochasticity, because it is built upon events. This stochasticity is considered as fluctuation in population process that arise from random differences among individuals. Therefore, demographic stochasticity becomes a common tool of scientists to epidemics model due to its mechanistic approach to including randomness and the individual nature of its formulation. The number of infectious, susceptible, exposed and recovered individuals in a population must here be integers.

In short, in demographic stochasticity, there are two elements in common, one is event and the other is random number. The first algorithm "First Reaction Method" is born in 1976 by Gillespie. Then, based on this first algorithm and these two key elements of demographic stochasticity, many scientists improved the first method, and created many better algorithms for stochastic simulations. But typically, there are only two main types of algorithms, exact algorithms and approximative algorithms. The typical algorithm that most practitioners use, among exact algorithms, is the algorithm “Direct Method” of Gillespie(1977) improved from the first algorithm “First Reaction Method”, and among approximative algorithms, is the “tau-leaping” algorithm. With these two types of algorithms, each has its private advantages and its private disadvantages. They have a common point. Their processes are in continuous-time Markov process for which the transition rates are constants, isn’t a function of time. The future state of the process, is only conditional

on the present state, but independent of the past. In constract, privately, for the exact algorithm, its advantage give us a really exact approach of simulating population-based time-to-event through two step with many iterations of 1) searching the time of next event by an exponentially distributed function and 2) searching the nature of next event. This process is repeated to iterate the model. In addition to the exact algorithms, Gibson & Bruck (2000) modified the first reaction method and created the Next Reaction method that substantially more challenging to program but is significantly faster than even the method when there are a large number of different event types. The Direct, First Reaction and Next Reaction methods are all exact stochastic approaches of the underlying ODEs. For these exact methods, on the one hand, noise in simulation only affects the probabilities associated with fates of individuals and the updating of each consecutive event is independent – there is no assumption concerning environmental stochasticity. On the other hand, these exact solutions become too slow and impractical when any one transition rate is large, when there is a big number of subpopulations or one a big number of event in a metapopualtion. It is the reason for that, approximate models are born instead of the exact stochastic methods. Gillespie (2001) has proposed a new method that decreases the simulation accuracy, but increases simulation speed. This is the "$\tau - leap$ method" known as an approximate method reduces the number of iterations by treating transition rates as constant over time periods for which this approximation leads to little error [9]. View the figure 2.

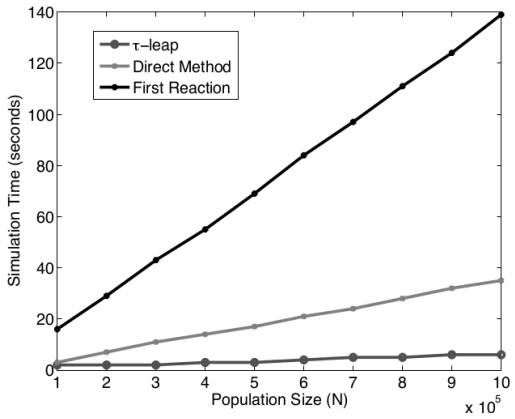

Figure 2: Figure is extrait from Keling (2008) The time (in seconds) to simulate 1000 years of SEIR epi- demics ($\mu = 0.02$ per year, $1/\sigma = 8$ days, $1/\gamma = 10$ days) on a 3.4GHz Pentium PC. For Gillespie's Direct (light grey line) and First Reaction (black line) methods, as the population size increases so does the simulation time, which, for very large populations, could become prohibitive. In contrast, the "$\tau - leap$ method (dark grey) is fast and largely unaffected by population size [9].

2.2.1 Disease transmission types in epidemic models

We know that, for each disease, knowing well its transmission methods is important for implementing proper infection control measures and prevention campaigns with large scale. The disease transmission methods depend on the characteristics of each disease and the nature of the microorganism that causes it. Here, we will point out some main types of transmission. For some infectious disease, its transmission method is complex and variant, such as anthrax. The anthrax can be spread through direct contact to a cut on the skin or through airborne spores which are inhaled.

Transmission by direct contact

This transmission requires a close contact between an infected person and a susceptible person, such as touching an infected individual, kissing, sexual contact, contact with oral secreations, or contact

with body lesions. Therefore, disease usually occurs between members of the same household or close friends and family.

Transmission by indirect contact

The indirect transmission has an intermediate factor where a susceptible person can be infected by a contaminated surface. The touch surfaces are usually in the life : door knobs, door handles; tables, beds, chairs; washroom surfaces; cups, dishes, trays; medical instruments; children's toy, etc. So, in order to reduce the number of infected individuals, we always make frequent touch surfaces be properly disinfected.

Transmission by droplet contact

These droplets are infected droplets that we usually see on the eye, nose and mouth. Droplets containing microorganisms can spread diease when an infected person coughs, sneezes and talks or through medical procedures. Droplets are too large to be airborne for long periods of time, and quickly settle out of air. We know measles and SARS are two exemples of transmission with droplet transmission. However, to reduce the infection, we can take personal protective barriers such as face masks and googgles to prevent disease transmission.

Transmission by airborne

This is quite dangerous cas where droplets are dust particles containing microorganisms. These organisms can remain suspended in air for long periods of time. And we become infected when we inhale these organisms. Fortunately, only a limited number of diseases are capable of airborne transmission, such as tuberculosis, chickenpox and measles.

Fecal-oral transmission

Fecal-oral transmission is usually associated with organisms that infect the digestive system, through ingestion of contaminated food and water. Fecal-oral transmission can be reduced by

- Proper storage of food at proper temperatures
- Thorough cooking of food
- Frequent and thorough handwashing, especially after washroom use
- Adequate sewage treatment and water filtration/chlorination systems
- Disinfection of frequent touch surfaces to prevent indirect contact transmission
- Increased public awareness of proper hygiene and food handling

Vector-borne transmission

Vectors are animals that are capable of transmitting diseases, such as flies, mites, fleas, ticks, rats, dogs, and in particular, mosquito. Since vectors are mobile, they increase the transmission range of a disease. So, changes in vector behaviour will affect the transmission pattern of a disease. It is important to study the behaviour of the vector as well as the disease-causing microorganism in order to establish a proper method of disease prevention.

Others

In addition to the transmission due to the characteristics of each disease and the nature of the microorganism, we also study demographic factors as the following:

- Age-structured populations
- Variable infectivity
- Distributions that are spatially non-uniform
- Diseases caused by macroparasites
- Acquired immunity through vaccination

3 Approaches in use

In the shape of this thesis, we use the most general SEIR model for infectious diseases, feature as diseases with transmission by airborne, such as meales that the entier world in the beginning of the 2014 was quite interrested. The World Health Organization (WHO) had officially to state 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 [13], in particular in southeast Asia particularly, in Vietnam [7]. Moreover, in this thesis, we will introduce the SEIR model for many subpopulations in a metapopulation.

3.1 Deterministic model for many cities

3.1.1 Ordinary differential equations for the deterministic model of many cities

The standard SEIR model (susceptible-exposed-infective-recovered) has been strongly developped 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]:

$$\frac{dS_i}{dt} = \mu N_i - \lambda_i S_i - \mu S_i \tag{1}$$

$$\frac{dE_i}{dt} = \lambda_i S_i - \mu E_i - \sigma E_i \tag{2}$$

$$\frac{dI_i}{dt} = \sigma E_i - \mu I_i - \gamma I_i \tag{3}$$

$$\frac{dR_i}{dt} = \gamma I_i - \mu R_i \tag{4}$$

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.

3.1.2 Formula for force of infection

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 [9].

$$\lambda_i = \beta_i \frac{I_i}{N_i} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \rho_{ij} \left[\frac{\beta_i[(1-\varepsilon_{ij})I_j - I_i]}{N_i} + \frac{\varepsilon_{ij}\beta_j I_j}{N_j} \right] \tag{5}$$

where β_i is the contact rate in population i and $\rho_{ij} = \rho_{ji}$ ($0 \leqslant \rho_{ij} \leqslant 1$ and $\rho_{ii} = 1$) is the coupling between subpopulations i and j. Among the infections caused by contacts with infected from other subpopulations, $\varepsilon_{ij} = \varepsilon_{ji}$ ($0 \leqslant \varepsilon_{ij} \leqslant 1$) is the proportion of infections due to susceptible individuals visiting other populations as opposed to infected individuals from other populations visiting the focal population. 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

$$\lambda_i = \beta_i \frac{I_i}{N_i}. \tag{6}$$

Consider that the contact rate β_i is seasonally forced [1] and seasonality is an annually periodic function of time [5]. As a result,

$$\beta_i(t) = b_0 \left[1 + b_1 \cos\left(\frac{2\pi t}{T} + \varphi_i \right) \right] \tag{7}$$

where t is the time, b_0 and b_1 are the mean value and amplitude of the infection rate β at which susceptible individuals become infected, T and φ_i are the period and the phase of the forcing. With the annual sinusoidal form of the infection rate, we really have the sinusoidally forced SEIR metapopulation model.

3.1.3 Equilibrium values of the system

In a case the infectious contact rate is constant, the equilibrium values of the variables S, E, I and R can be expressed analytically as follows. We know that, in simulation, 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 1 - 4) 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:

$$S_i^* = N_i \frac{(\gamma+\mu)(\sigma+\mu)}{\beta\sigma} \quad (8)$$

$$E_i^* = N_i\mu\left(\frac{1}{\sigma+\mu} - \frac{\gamma+\mu}{\beta\sigma}\right) \quad (9)$$

$$I_i^* = N_i\mu\frac{\beta\sigma - (\sigma+\mu)(\gamma+\mu)}{\beta(\sigma+\mu)(\gamma+\mu)} \quad (10)$$

$$R_i^* = N_i - S_i^* - E_i^* - I_i^* \quad (11)$$

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

$$S_i^* = N_i \frac{1}{R_0} \quad (12)$$

$$E_i^* = N_i \frac{\mu\sigma}{R_0}(R_0 - 1) \quad (13)$$

$$I_i^* = N_i \frac{\mu}{\beta}(R_0 - 1) \quad (14)$$

$$R_i^* = N_i - S_i^* - E_i^* - I_i^* \quad (15)$$

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

3.2 Stochastic model for many subpopulation in a metapopulation

3.2.1 Stochastic model for many subpopulations in a metapopulation

In order to study the persistence of the disease, we must consider a stochastic version of the model [8, 10, 11]. We use for that a population-based time-to-next-event model based on Gillespie's algorithm [4]. Table 1 lists all the events of the model, occurring in subpopulation i.

3.2.2 Exact method for many subpopulations in a metapopulation

As pointed out above, the direct method is usually more computationally efficient. This is the exact stochastic method more commonly used of the underlying ODEs. However, this method is only proposed for one population, so here we modify a little the direct method in order to accord with a metapopulation of multi-subpopulation. We call it "the direct method of multi-subpopulation". View the figure 3.

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 $R_i \leftarrow R_i + 1$

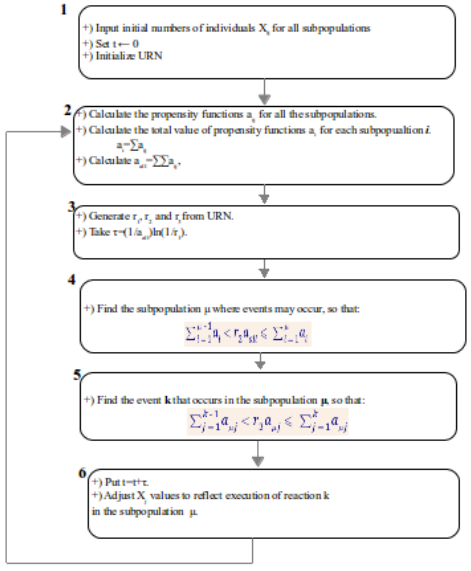

Figure 3: Improving the exact algorithm of Gillespie (1977) for a metapopulation of multi-subpopulation.

This multi-subpopulation method add two steps into the direct method, these are steps 3 and 4. We generate one more new random number, and then we convert population rates into probabilities, finally this new random number is used to select one of the subpopulation in the metapopulation of n subpopulations.

3.3 Result

3.3.1 Package "dizzys" for stochastic SEIR metapopulation model

At present, we built successfully a package named "dizzys". This package in R implements both the exact and approximate methods for the deterministic/stochastic SEIR/SIR models by integrating the R package and the C++ implementation. We use C++ to perform the algorithms, and use R to create interfaces. Hence, this new integration is faster than any pure R implementation.

Simulation 1:

Simulating the deterministic and stochastic SEIR model for one seul population. View the figure 4.

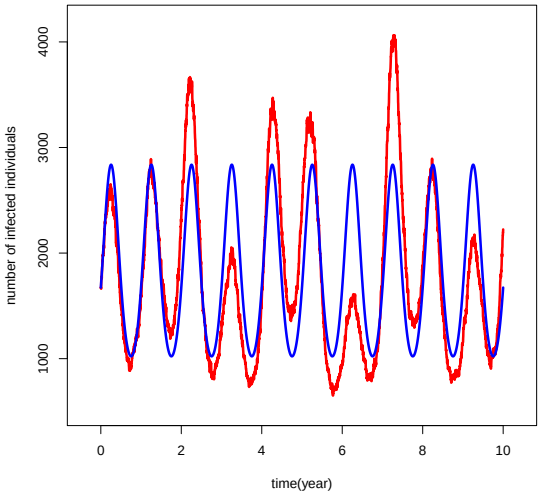

Figure 4: Deterministic/stochastic SEIR model for one seul population with N=10^7, duration=10year. The blue fluctuation for the deterministic model and the red for the stochastic model.

Simulation 2:

Simulating three subpopulations with the different population sizes, then continuing or redoing this simulation with other values of parameter, we can do it by exploiting the 'simul' function in the package. View the figure 5.

Simulation 3:

The SEIR stochastic model using "adaptive tau-leaping" algorithm by exploiting the function "simul" in the package. To do this algorithm, we only make the parameter $method = "adaptivetau"$. We can compare the result of the direct algorithm with the result of the adaptivetau algorithm as follows. View the figure 6.

Simulation 4:

Plot in 3D, view the figure 7.

3.3.2 Global persistence in a metapopulation

Here, we start exploiting the package "dizzys". In order to illustrate the interaction between disease transmissibility and phase of seasonal forcing, we start in this section by studing the stochastic

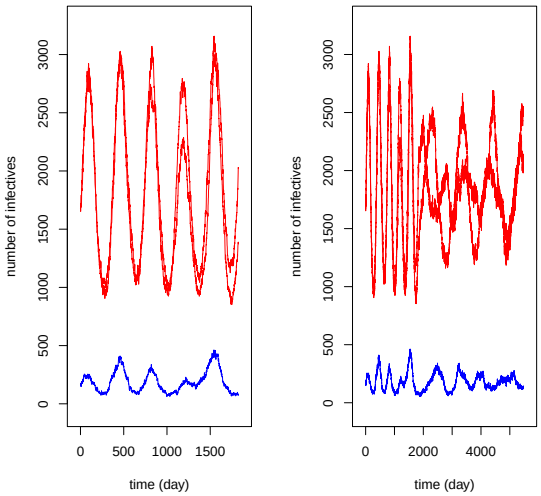

Figure 5: Continuing or redoing this simulation. The start seir object with type="stochastic", duration=5*365, mu=1/(70*365), beta0=1000/365, beta1=.1, sigma=1/8, gamma=1/5, nbVilles=3, N=c(10^7,10^6), phi=0, duration=5*365; and the last seir object with type="stoch", continue=T, append=T, beta1=0.0, phi=pi, duration=10*365. The two red fluctuations for the two first subpopulations with N=10^7 and the blue for the third subpopulation with N=10^6

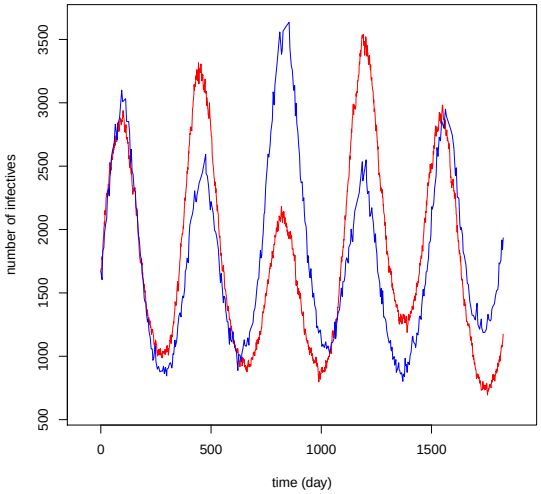

Figure 6: Compare the result of the direct algorithm with the result of the adaptivetau algorithm. The red fluctuation for the direct method and the blue for the "τ - leap" method.

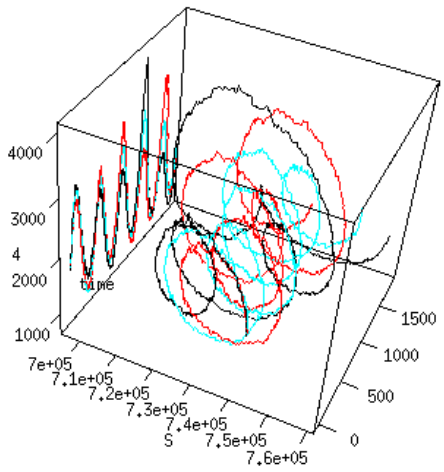

Figure 7: Plot in 3D of the stochastic SEIR model with the three subpopulations in the metapopulation. We view the relation between the number of susceptible individuals, the number of infected individuals and the time. Then, we can also view the projection of these fluctions on the correspondent surface.

SEIR model in a metapopualtion of n subpopulations. For this meta-population, we observe the disease extinction in time due to spatial synchrony or spatial 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} city is correspondent to i^{th} value in the set. We call φ_{max} synchrony parameter.

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 allowed to draw Kaplan-Meier survival curves from which we estimated the persistence rates χ:

$$M(t) = \exp(-\chi t) \tag{16}$$

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 estimate the persistence rate, here, we use the parametric survival model for the exponential distribution (R package '*survival*' [12]). For that, we find the persistence rate of the number of metapopulations in which the disease is extinct in time. Based on this persistence rate, we study the relation between the global persistence and its average synchrony due to the phase of the forcing.

Global persistence and time We fixe here the metapopulation of two subpopulations, $N_1 = N_2 = 300,000$, the rate of coupling $\rho = 0.01$, the simulation time 50 years, the number of simulations $m = 100$, 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 8.

The phase of forcing of the $subpopulation_1$ is always fixed 0, but this 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

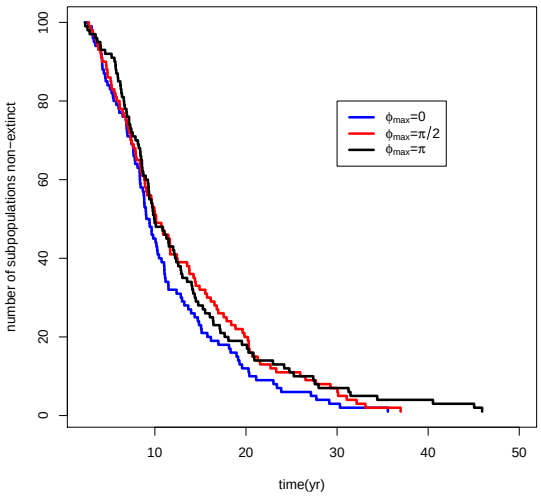

Figure 8: Kaplan-Meier survival curves for disease persistence after 100 different simulation Disease persistence time after 100 simulations of $\varphi_{max} = 0$, $\varphi_{max} = \pi/2$ and $\varphi_{max} = \pi$. The blue survival curve for $\varphi_{max} = 0$, the red survival curve for $\varphi_{max} = \pi/2$ and the black curve for $\varphi_{max} = \pi$. The persistence time of $\varphi_{max} = 0$ is the shortest. The persistence time of $\varphi_{max} = \pi$ is the longest.

two subpopulations become asynchrony when $\varphi_{max} = \pi/2$ or π. The symmetry of fixed points is just broken at the starting moment. It is reason for that the level of synchrony of the metapopulation decreases. Additionally, we find that the value of the synchrony parameter φ_{max} change according to increasing tendency, the phase difference between the fluctuations of the two infection rates β increases also, and the global persistence time in the metapopulation thus augments. When $\varphi_{max} = \pi$, the two subpopulations are in antiphase. This is the most difficult case to find global extinction, the persistence time is the longest. 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.

Global persistence rate and φ_{max} In this part, the Kaplan-Meier survival curves of the global disease persistence time in a metapopulation for each value φ_{max} is considered as a parameter that is called global persistence rate of that φ_{max}. We used survival functions to estimate this global persistence rate with confidence interval 95%. The result is shown as following figure 9.

This figure 9 points out to us that the amplitude of the confidence intervals for each value of φ_{max} are quite close to each other. However, it gradually rises when φ_{max} runs from 0 to π. The phase difference strongly influences disease persistence time, at the same time, global disease persistence rate. The asynchrony between subpopulaitons is the main reason to respond very familiar question why has the infectious disease been never extinct.

One more factor that was pointed in this introduction part is coupling strength between subpopulations. We can say that this coupling parameter can symbolize migration strength. The disease transmission speed fast augments when coupling rate goes up in metapopulation. Similar to that, global disease persistence increases also. Here, we permit coupling rate change from weak to strong in a metapopualtion of five subpopulations with $N = 10^5$ for each subpopulation. The dispersal rate ρ is divised into three intervals. These are low, intermediate and high coupling rate intervals. In each interval, we choose 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 estimated persistence rates and values φ_{max}, in condition φ_{max} belonging

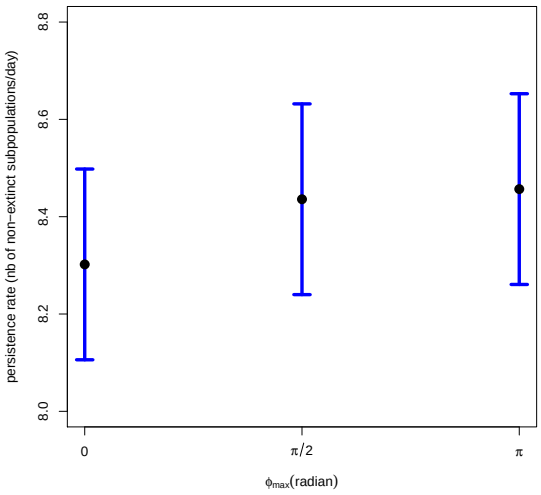

Figure 9: Estimated rate of global disease persistence in the metapopulation of the two subpopulations after 100 different simulations N=3e5, coupling rate $\rho = 0.01$. Here, with 95% confidence interval, blue lines are confidence intervals for persistence rate of each value of φ_{max}. This intervals are limited by lower and upper confidence limits.

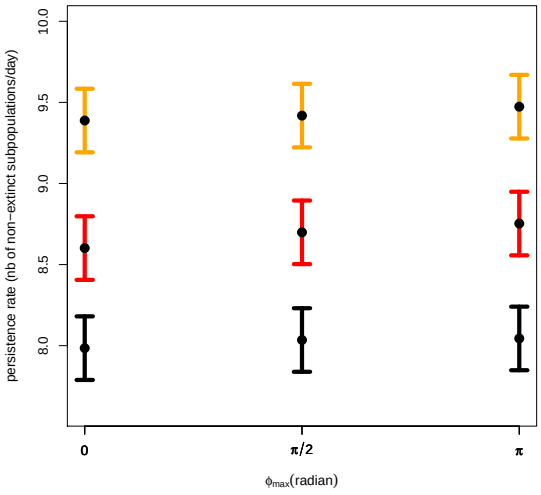

Figure 10: Estimated persistence rate in the metapopulation of three, five and ten subpopulations after 100 different simulations. Here, with 95% confidence interval, black, red, and orange lines are for three, five and ten subpopulations, respectively. These intervals are limited by lower and upper confidence limits.

to the set $\{0, \pi/2, \pi\}$. Now, we had curves presenting the correlation between coupling rate and gradient of φ_{max} and persistence rate (view the following figure 11).

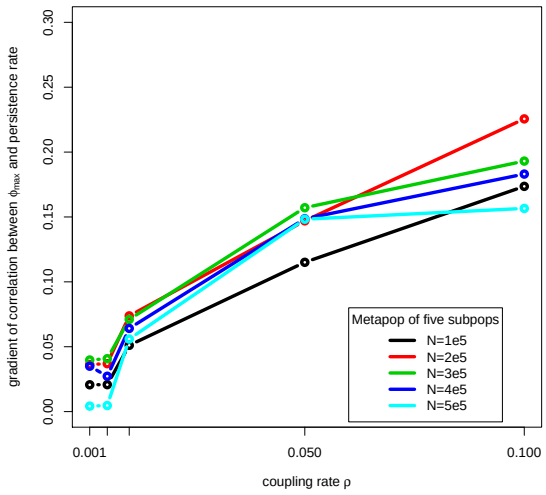

Figure 11: Correlation between coupling rate and gradient of level of synchrony and persistence rate in the metapopulation of five subpopulation. Here, the coupling rate ρ is in $\{0.001, 0.005, 0.01, 0.05, 0.1\}$, the level of synchrony φ_{max} in $\{0,\pi/2 ,\pi\}$ and the population size of each subpopulaiton N from$100,000$ to $500,000$.

Due to this result (figure 11), we show that the coupling rate between subpopulation augments, thus we get an increase of gradient coefficient. It means that global disease persistence time in metapopulation strengthens for coupling rate between subpopulations though how many is population size of subpopulation with the interaction strength ρ from 10^{-3} to 0.1 [9].

Moreover, because of dispersal rate ρ, the persistence time of metapopulation is considered as a function of dispersal rate. In order to affirm this supposition, here, we permit coupling rate change from weak to strong in a metapopualtion of five subpopulations with $N = 10^5$ for each subpopulation. The dispersal rate ρ is divised into three intervals. These are low, intermediate and high coupling rate intervals. In each interval, we choose some coupling rates that highlight the coupling strength among subpopulations in a metapopulation. We have the result as follows (figure 12).

When the coupling rate is small from 0.001 to 0.005, the gradient of correlation between φ_{max} and persistence rate increases very slowly. However, the persistence rate augments in a sudden way when the coupling rate changes from 0.005 to 0.1. Lastly, the gradient strongly decreases when the coupling rate is so robust from 0.5 to 1. Based on this figure, the disease persistence is one humped function for the coupling rate and the coupling rate is . The medium coupling rate (from 0.005 to 0.1) maximizes disease persistence in metapopulation.

4 Conclusion and perspective

We successfully built a version for the susceptible-infectied-recovered stochastic metapopulation model. The infection rate λ_i for $subpopulation_i$ portrayed all effects inside as well as outside on disease tranmission chain between individuals in the same subpopulation or in other subpopulations. Moreover, our metapopulation model becames more detailed when we brought seasonality in metapopulation model to create periodic transmission in year that here highlights seasonal change as well as school period of children. We have metapopulation model with different contact rates.

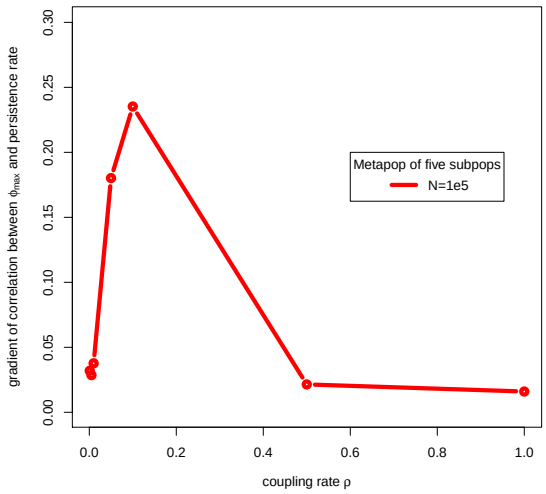

Figure 12: Correlation between coupling rate and gradient of level of synchrony and persistence rate in the metapopulation of five subpopulation. Here, the coupling rate ρ is in {0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1}, the level of synchrony φ_{max} in $\{0,\pi/2\ ,\pi\}$ and the population size of each subpopulaiton N=1e5.

This is more complex model than any used metapopulation model. We sketched successfully in-phase and sometime out-of-phase ("antiphase") models across suburbs of He's 2003 [6].

In order to continue my thesis, we will exploit the global persistence and the disease transimission among subpopulations under a new shape of metapopualtion in epidemiology. It is the gravity model in epidemiology. This model is a very simple model, and can be used to explain the spatial epidemiologic dynamics. We says it simply that the strength of coupling between any two populations depends on size of these two populations and the distance between them (so this model is quite similar to Newton's law of universal gravitation, the attraction between two bodies depends on the masses of the bodies and the distance between the bodies). In epidemiology, we will call the "mass" the population size. Then, we set a gravity metapopulation model with the different positions and the different population sizes of subpopulations. Finally, we will exploit the disease transmission in the gravity metapopulation model.

One more main work in my thesis, it is the optimization of the vaccination policies. We set it in the last part of my thesis, because first of all, we must deeply exploit the actual nature of the infectious disease transmission in the different disease model, then we are going to have a good vaccination policy for a metapopulation

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

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

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

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

[8] 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.

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

[10] 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.

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

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

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

[1] http://microbiology.mtsinai.on.ca/faq/transmission.shtml
[2] http://en.wikipedia.org/wiki/Epidemic_model