MODÉLISATION, SIMULATION MULTI-NIVEAU POUR L'OPTIMISATION DE POLITIQUES DE VACCINATION

TRAN THI CAM GIANG, JEAN DANIEL ZUCKER, MARC CHOISY, YANN CHEVALEYRE

Un plan pour ton mémoire de thèse pourrait être:

1. State of the art:

1.1. Epidemiology (and monitoring).

1.1.1. *Epidemiology.* As we know that public health problems are one of the emerging troubles in the entire world. They directly influence human heath, the health of one person, the health of a community. In particular, any news about infectious diseases for children has always been a subject of concern to parents as well as everyone. Hence, in the world, a discipline "epidemiology" has risen to study the factors, causes, and effects of infectious diseases.

This thesis is proposed in a context in which many public health serious events have occurred in the world : SRAS in 2003, avian influenza in 2004 or swine flu in 2009, etc. In particular, at the start of 2014, the World Health Organization (WHO) officially stated 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 [**?**], particularly in southeast Asia and in Vietnam [**?**]. This has pointed out the important role of the epidemiological phenomena anticipation when diseases occur. Many studies proposed by the WHO, the Pasteur Institute and the Inserm in the field of "environmental security" try to understand disease phenomena and spread of disease over a territory, to better manage when diseases occur. These researches consist of mathematical or statistical studies via surveillance networks [**?**]. This is one of the axes of the UMMISCO laboratory's research themes (IRD UMI 209).

1.1.2. *Control.* As we know, pathogenic microorganisms such as bacteria, viruses, parasites or fungi are key factors causing infectious diseases. The diseases can be spread directly or indirectly from one person to another, through a mediate environment or contaminated tools. As far as directly infectious diseases are concerned, meaninf diseases directly transmitted from one person to another, we have some normal policies to prevent the spread of diseases such as vaccines, anti-viral medications, and quarantine. In this thesis, we focus on vaccines in the human community. A vaccine is understood as a biological preparation that provides active acquired immunity to a particular disease for our body. After having been vaccinated, we transport microorganisms in weakened or killed form of the microbe into our body. The body's immune system produces the right antibodies to recognize the germs as a threat, destroy them and keep a record of them. Because of that, when the disease occurs, our immune system can recognize and destroy with a better chance

Date: 03/04/2015.

Question	Answer	Why?
Are vaccines safe?	YES	Vaccines are generally q
Are there vaccines for all infectious?	NO	For example: deng
Are all vaccines free?	NO	Funding problem
Are all people vaccinated before a requested age for each disease?	NO	Funding/geographic/cultura

TABLE 1. Vaccine state

of success any of these germs that it later encounters. The administration of vaccines is called vaccination. Vaccination has greatly helped human beings. The vaccination of influenza, Human Papillomavirus (HPV) and chicken pox have been particularly appreciated. Smallpox is a particular example. This disease was filled people with terror during the closing years of the 18th century. Smallpox killed an estimated 400,000 Europeans annually and among the people that luckily survived, a third had been blinded by the disease. However, the World Health Organization (WHO) offcially stated the eradication of smallpox in 2011 [**?**, **?**, **?**]. In addition, many infectious diseases are clearly restricted such as influenza, polio, measles and tetanus from much of the world. Thus, one big question proposed is why many infectious diseases still exist in the world though we have produced vaccines for most infectious diseases. In order to answer this question, first of all, we have to answer to some following small questions :

With the four answers above, we can say that the human still faces up to infectious diseases. A thorough knowledge of the disease is essential in order to implement large-scale proper infection control measures and prevention campaigns. Granted that the disease transmission methods depend on the characteristics of each disease and the nature of the microorganism that causes it. In this thesis, we will investigate popular infectious diseases with 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 with oral saliva, or contact with body lesions. Therefore, these diseases usually occur between members of the same household or close friends and family. In particular, this thesis will mostly focus on measles. Because measles is a highly contagious, serious disease caused by a virus. It is a typical infectiuous disease with direct transmission. In 1980, approximate 2.6 milion people was killed each year before we had the widespread vaccination policies. It spreads very fast by coughing and sneezing in human communities via close interpersonal contact or direct contact with secretions. Its main symptoms consist of high fever, cough, runny nose and red eyes. These first symptoms usually take from 10 to 12 days after exposure to an infectious person, and lasts 4 to 7 days [**?**]. In fact, now there is no proper treatment for measles to totally prevent the spread of measles except routine measles vaccination policy for children. According to the report by the World Health Organization (WHO), since 2002 measles was eradicated from U.S. However, today measles vaccination has not been extensively popularized in the entire world. Beside the obtained results, for example, in 2013, there was about 84% of the world's children having received one dose of measles vaccine, and during 2000-2013, measles vaccination prevented an estimated 15.6 million deaths; we have had to face upabout 145700 measles deaths globally- estimated 400 deaths every day or 16 deaths every hour in 2013. Measles becomes one of the leading causes of death among

young children in the world, although now we are having a big stock of safe and readily available measles vaccines.

Mass policy (or the routine measles vaccination policy for measlses) that vaccinates the maximum number of children before a certain age, is the oldest (started from the 1950s in the rich countries) and is now the most used. The policy has obtained clear results : a clear decrease of the incidence in most countries. However, the problem of this vaccination policy is too expensive, really ineffective and quite impossible to implement in poor countries, especially in Africa because of both financial and logistical problems. (e.g. the WHO project "Extended Program on Immunization" in Vietnam for the measles extinction before 2012 failed [**?**]). In addition, when a vaccination policy is performed in a country, there is only one policy deployed, but in modeling, we can realize many policies and assess their results.

In short, measles is still a common and often fatal disease in the world. We still very much need to model the transmission dynamics of measles and investigate the effect of vaccination on the spread of measles in the entire world. More largely, we need to give new optimal vaccination policies in artificial intelligence in order that these policies may become more effective, less expensive, and take into account the spatial dimension for all popular infectious diseases.

1.2. **dynamiques/structures spatiales (théorie métapopulations, réseaux, etc. . .)**

- For directly transmitted infectious diseases by virus and bacteria, susceptible individuals are not only infected by infected individuals in the same location, but also by other infected individuals due to the movement of individuals between populated regions. This is one very important part in the domain studying the geographical spread of infectious diseases. We care for host population characteristics, then characteristics of spatial spread of an infectious disease among populations. Through these characteristics, we find optimal policies to minimize the number of infected individuals in a community. In fact, there are many studies about the interactions among populations. However, we can divide the spatial structure of populations into two main levels: "inter-city level" and "intra-city level". At the inter-city level (or called "micro-level"), we use differential equations to control its models. At the "intra-city level" (also called "macro-level") in which we provide connections between the populations, simulate the intra-city traffic. We consider the effect of travel through the connections between population regions as a means of spreading a virus [**?**].
- We have two basic models considered in the "macro-level", the model has no explicit movement of individuals and the models describes enough travels and movements of individuals among populations and even takes into account the resident population as well as the current population of individuals [**?**]. A population may be simplified as a city, community, or some other geographical region. Population travel (e.g. among animals and among people by foot, birds, mosquitoes and in particular, people travel by air from one city to another), is the main reason why diseases can spread quickly among very distant cities such as SARS disease in 2003. Therefore, the term "metapopulation" arrived in the ecological literature in 1969 by Levins [**?**, **?**]. A metapopulation is a population of a set of

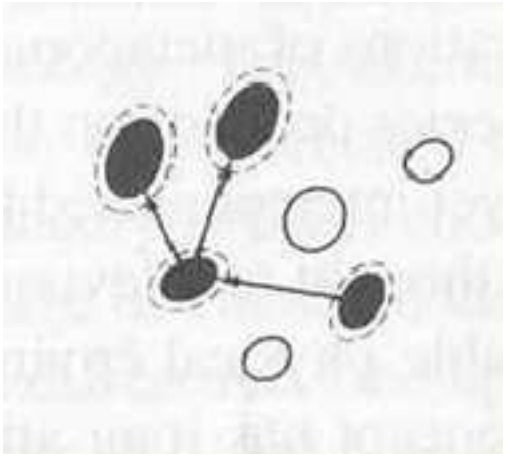

FIGURE 1.1. Classic Levins Metapopulation Model [?]

spatially discrete local populations (or subpopulations in short) with mutual interaction [?]. In the metapopulation in which a subpopulation can only go extinct locally and be recolonized by another after it is emptied by extinction [?, ?, ?] and migration between subpopulaitons is significantly restricted. In a metapopulation, if recolonization rates are smaller than extinction rates, then total extinction of all local population will easily be reached. The persistence time of the metapopulation is measured as the time until all subpopulations go extinct. According to Harrison (1991) [?] there are four types of spatially dynamic populations : classic Levins metapopulation, mainland-island metapopualtion, patchy population and non-equilibrium populations.

- The first metapopulation model was proposed in 1969 by Levins. It is called the classic Levins Metapopulation [?]. Wilson in 1980 [?] stated that in this classic model "A nexus of patches, each patch winking into life as a population colonizes it, and winking out again as extinction occurs."
  All subpopulations in this classic model are relatively small. The levels of interaction among individuals within a subpopulation is much higher than between subpopulations.
- The second model is the mainland-island metapopulation in which there are some small "island" subpopulations within dispersal distance of a much larger "mainland" subpopulation.
  It is evident that smaller subpopulations have a high probability of local extinction, but the mainland population will hardly become extinct. The migration from the mainland to the islands is independent of the islands white or filled, but is propagated for the connected islands. Therefore, if the mainland population has a low individual density and there is no immigration, then population growth rate is positive. Inversely, if island populations are in the same conditions as the mainland, then its population growth rate is negative. Thus, the islands would go down to extinction if there are no imemigrants.

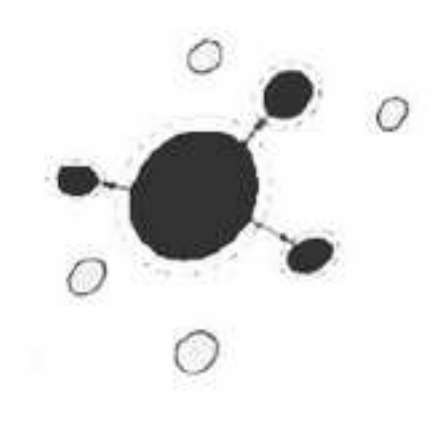

FIGURE 1.2. Mainland-Island Metapopulation [?]

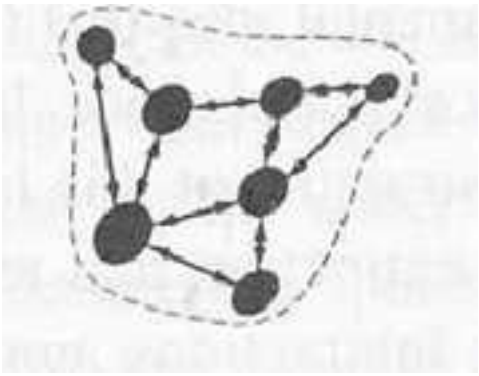

FIGURE 1.3. Patchy population [?]

- The third model is patchy population. The local populations exist in a big habitat population and the dispersal rate between subpopulations is high.
  Here we can find that the population structure is grouped and the interaction among them is frequent. However, this model is not referred as a concept for metapopulation and most researchers do not consider this a meta-population either.
- The final model is the non-equilibrium population. The local populations are patches, its local extinctions are much greater than its recolonisation.
  It is obvious that white patches are rarely or never recolonized. Therefore, this model is not considered as a functional metapopulation. We can find this model in forested agricultural fields.

We already have four metapopulation models. In order to model the metapopulations mentioned above, we have three main model to implement : spatially-implicit model, spatially-explicit model and spatially-realistic model. For the first model,

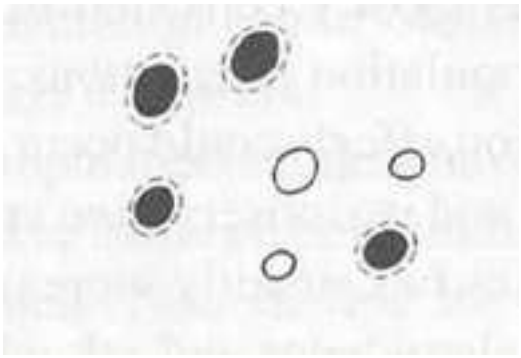

FIGURE 1.4. Non-equilibrium population [?]

this is the type of model used in Levins (1969) [?] in which supposing that all local populations are connected with each other and they have independent local fluctuations. At any one time, we save track of the proportion of local populations and we do not take care the distance between them and the population size of each subpopulation. This model are mathematically and conceptually easy to implement. But this model can only answer some metapopulation problems because it ignores so many variables of a metapopulation. This model should be used for metapopulation close to a steady state.

For the second model, the spatially-explicit model is more complex than the first model. Subpopulations may be filled or vacant. Local populations only have interactions with the nearest neighbors. Subpopulations are organized as cells on a grid and migration among them depends on population density. We also only consider presence or absence of a species in each subpopulation. The advantage of this model is easy to model because of same local behaviors from subpopulation to subpopulation. However, we cannot simply describe the state of the metapopulation through filled subpopulations. Finally, the spatially-realistic model uses GIS to realize attributes, geometric coordinates, etc ... to a metapopulation. The first author using this model is Hanski in 1994 [?]. His model was defined as the incidence function (IF) model. This model is more realistic, and we can estimate quantitative predictions about metapopulation fluctuation. However, in fact, this model is very complicated, and many geographic data have to be estimated. Hence, the metapopulation concept start to no longer exist.

In the scope of this thesis, we focus on a metapopulation model that is result of combination between the spatially-explicit model and the patchy population. In general, this a simple spatial model, but is one of the most applicable model to descrire spread of diseases in human communities. This metapopulation consists of distinct "subpopulation, each of which fluctuates independently, together with interaction limited by a coupling parameter ρ. These subpopulations may be filled or empty and contact with any neighbours.

1.3. **Epidemiologic models.** It is known that, there are many current models that are used to model complex systems in nature, in ecology system and in epidemiology. Mathematical models in epidemiology are a typical exemple. These

models permit us to present behavior of diseases and disease process in mathematics. However, explaining the transmission of infectious diseases is a difficult problem for an epidemiologist. Because there are many different interacting factors causing the outbreak of diseases such as the environment, the climate, the geography, the culture,...Hence, the role of the epidemiologist is how to model the characteristics and the transmission process of an infectious disease. Researchers have proposed compartmental models in epidemiology by dividing the population into "compartments" that illustrate health states of human through individuals. These compartmental models are called the epidemic models too. The first benefit of these models is to model the transmission process of a communicable disease through compartments. Then, we can predict the properties of the disease dynamics such as the estimated number of infected individual, the time of persistence of disease, further that where and when we can implement vaccination policies to have both a minimum number of vaccined individuals and the minimum number of infected individuals in a given population. Let image that now in your country, there is an infectious disease as measles, a baby can be infected. According to the process of infection of disease, firstly this baby was born, he is fine and he is not infected yet by the measles but he may be infected in the future. We say that he belongs to the susceptible group (in short, S). Then, his mother takes him to a supermaket, there he see so many people, he is really infected through any way. He starts having a high fever, he may have to pass this state from 3 days to 5 days. In this period, he is really infected but he cannot infecte others. We say that he belong to the exposed group (in short, E). After that, he start decreasing the temperature, but at the same time, he begins having red rashs on the back of the ears, after a few hours, on the head, on the neck and finally most of the body. This period appears from five to eight days after the exposed step. This duration is very sensible. The baby is completely infected and he can infecte others if they see him. He belongs to the infected group (in short, I). Finally, he passes to the final period, he comes back good state. We say that he belongs to the recovered group with immunity (in short, R).

Around these four main health groups presenting the process of infection propagation in community, there are many epidemic models proposed. We give here the development of epidemic models by focusing on acute infections, assuming the pathogen causes illness for a periods of time followed by (typically lifelong) immunity. The first simplest model is the S-I-R model created by W. O. Kermack and A. G. McKendrick in 1927. The authors categorized hosts within groups as described above **S**usceptible (if not yet exposed to the pathogen), **I**nfected (if currenly infected by the pathogen) and **R**ecovered (if they have successfully cleared the infection). From the simplest SIR model, in order to accord each infectious disease and real property of disease, scientists have modified it, made it different multiforme. However, in shape of this thesis, we concentrate on the SEIR model (as the figure 1.5) that fit many currently infectious diseases in the world. Each patient must pass four health steps : susceptible stage, incubation stage, infectious stage and recovered stage.

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 :

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

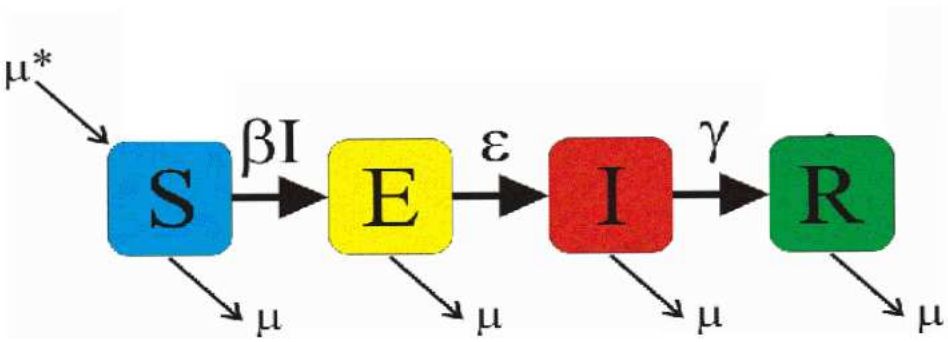

FIGURE 1.5. SEIR model

- 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.

The conceptual descriptions of the model can be represented by a flow diagram above. The flow diagram for the SEIR model uses arrows to present the movement between the S and I classes, the E and I classes and the I and R classses. Here, individuals are born susceptible, die at a rate μ, become infected with the force of infection λ that is a function among the contact rate β, the number of infected invidual I and the population size N, infectious after a latency period of an average duration of $1/\sigma$ and recover at the rate γ.

The SEIR model is investigated by ordinary differential equations (ODE) that are deterministic [**?**]. The value of variable states is only determined by parameters in the model and by sets of previous states of these variables. Moreover, the epidemic models are often proposed for one single population [**?**]. In the scope of this thesis, we propose a deterministic model for many subpopulations in a metapopulation. The standard SEIR model (susceptible-exposed-infective-recovered) has been strongly developed for the dynamics of directly infectious disease [**?**]. 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 [**?**]:

$$\frac{dS_i}{dt} = \mu N_i - \lambda_i S_i - \mu S_i \tag{1.1}$$

$$\frac{dE_i}{dt} = \lambda_i S_i - \mu E_i - \sigma E_i \tag{1.2}$$

$$\frac{dI_i}{dt} = \sigma E_i - \mu I_i - \gamma I_i \tag{1.3}$$

$$\frac{dR_i}{dt} = \gamma I_i - \mu R_i \tag{1.4}$$

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 [**?**] **:**

$$\lambda_i = \sum_j \rho_{ij}\kappa_j \log\left[1 - \sum_{k=1}^{M}\left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk}\right)\right] \tag{1.5}$$

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 subpopulation 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

$$\lambda_i = -\kappa_i \log(1 - \frac{I_i}{N_i} \times c_{ii}) \tag{1.6}$$

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

$$\kappa_i(t) = \kappa_{i0}\left[1 + \kappa_{i1}\cos\left(\frac{2\pi t}{T} + \varphi_i\right)\right] \tag{1.7}$$

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 subpopulation i per unit of time, T and φ_i are the period and the phase of the forcing. With the annual sinusoidal form of the average contact rate, we really have the sinusoidally forced SEIR metapopulation model.

In detail, the deterministic model performs the same way for a given set of initial conditions. It doesn't have randomness, dynamics, and don't present dynamic of

diseases in nature. Thus, stochastic models have been proposed. A stochastic model is always more realistic than a deterministic one. These models have stochastic and variable states are not described by unique values, but by probability distributions. It is why we will use the stochastic models to predict extinction propability of disease in spatial context[**?**].