

Daily Mobility and the Spread of Communicable Infectious Diseases

Alexandre CEBEILLAC^{1,2} and Eric DAUDÉ¹

1 IDEES, CNRS, Normandie Université, Rouen, France

2 Institut Pasteur, Paris, France

1.1. Introduction

Communicable infectious diseases are caused by microorganisms, called pathogens (viruses, fungi, bacteria, parasites) that infect a living organism, called the host. Pathogens are transmitted from an infected host to a healthy host through various channels. A first category of transmission channels involves direct, physical contact (saliva, blood, sexual contact: AIDS, hepatitis B, syphilis, etc.) or the air or respiratory tract (cough: flu, tuberculosis, pertussis, nasopharyngitis, etc.). A second category involves an indirect mode of transmission, either by ingestion (food, water: hepatitis A, cholera, rotavirus, poliomyelitis) or through a vector (mosquito, tick, flea: dengue fever, Zika, Chikungunya, Lyme). Whether direct or indirect, this transmission pattern implies the proximity of hosts moving around the same geographic space. Mobility therefore becomes a key element for understanding the spread of epidemics.

For a color version of all the figures in this chapter, see www.iste.co.uk/vallee/everyday.zip.

Everyday Mobility and Health,
coordinated by Julie VALLÉE. © ISTE Ltd 2024.

An epidemic corresponds to a rapid increase in the number of cases of a disease in a local or national area – a pandemic is its global expression. Identifying the emergence of the pathogen responsible for a disease as well as its modes of propagation should enable a better understanding, or even anticipation, of its geographical diffusion. The example of the spread of Covid-19 after the first cases were identified in China evidences the weight of international (and then national) mobility in this propagation. At a local scale, in everyday spaces, the role and weight of mobility is more complex to disentangle, because mobility is concomitant with many other factors which, combined, lead to local disease diffusion mappings: local disease surveillance and control actions, characteristics of the environments visited, behavior of the populations at risk, etc. This complexity partly explains why local daily mobility has long remained understudied in retrospective analyses. In the majority of cases, individual exposure factors are solely deduced from the characteristics of the living space of sick people, a bias partly related to the recording of the case at their place of residence. However, this obstacle tends to be partially surmounted thanks to digital tools and data. As it modifies a static vision of the exposure to pathogens, daily mobility renews research on the spread of communicable diseases.

The first section of this chapter (section 1.2) will summarize the evolution of knowledge on the contagiousness of diseases and their modes of transmission in an epidemic context, emphasizing the link between daily mobility and epidemiology. The modeling of daily mobility will be addressed in section 1.3. We will discuss different modeling methods and the concepts used in an epidemic context. Special emphasis will be placed on the most recent classes of models, namely individual-based models (IBMs). These models integrate geographic information as well as social and epidemiological knowledge into a single formalism, enabling exploration of a wide variety of scenarios. Finally, a last section (section 1.4) will prompt a selective literature review, dedicated to the data currently in use to better understand daily mobility and to calibrate increasingly detailed epidemiological models.

1.2. Analyzing the role of daily mobility in the spread of epidemics: concepts, studies and tools

Understanding the mechanisms responsible for the spread of an epidemic is a prerequisite for any action attempting to prevent it. Spatial epidemiology now plays a major role in this fight.

1.2.1. The path of daily mobility in the geographical reflection of infectious diseases

When it reached Europe in 1829, little medical controversy was stirred up in identifying Asiatic cholera morbus. The symptoms described by colonial doctors stationed in India matched the symptoms observed at Europe's doorstep: "rice water" stools, profuse vomiting of liquid matter, spasms, horrible cramps, imperceptible pulse, bluish tints in the extremities and around the lips (hence the name "Blue Death"), body cooling, collapse, etc. and, above all, a mortality rate exceeding 50% in many cities. Though the poison had a name and a geographical origin, it was still necessary to determine its medical origin, its modes of propagation, the means to protect against it and the treatment to eradicate it. These questions fueled part of the medical research of the 19th century when two currents of thought collided, the insufficiency of medical knowledge leading to different interpretations of identical statistics and cartographic documents (Eliot et al. 2012).

Doctors then became divided between two schools: contagionists and anti-contagionists. In the 19th century, the sense of the term "contagion" implied that the disease was transmitted from person to person, via contact. This excluded diseases transmitted by air or water from the list of contagious diseases. One of the measures against cholera for the doctors who classified it under contagious diseases involved the isolation of any infected person in view of protecting the population: quarantine. On the other hand, anti-contagionists believed that disease was transmitted via the interactions between human beings and the environment – chiefly as a result of miasmatic exhalations. On the one hand, favorable conditions to the development and propagation of epidemics could be environment-related (climate, atmospheric density, degree of humidity, prevailing winds, the term "malaria" being derived from the Latin *Mal'aria* – which means "bad air" – whereas the term "palu" means a swamp) and, on the other hand, they could be owing to the type of habitat (cleanliness of houses and clothes, as well as inhabitants' mores).

Overcoming this opposition between contagionists, anti-contagionists and supporters of the miasma theory, John Snow (1813–1858) shed new light on the understanding of the modes of propagation of cholera. As Europe experienced a new epidemic onslaught and the "Blue Death" spread across the cities and the countryside, Snow identified the waterborne nature of the "poison" that was wreaking havoc from Bengal to America, and ravaging Europe.

In his famous treatise *On the Mode of Communication of Cholera* published in 1855, supplemented with maps and statistics, the English doctor offered an original demonstration to support his hypotheses on the disease's mode of propagation (Snow 1855). The singularity of Snow's work was his inclusion on medical topography maps – which occasionally recorded deaths and people contaminated by cholera – of water points by neighborhood, notably that of Broad Street in London. He proved that most of the people who died of cholera lived near dirty water points.

Apart from drawing up these maps, Snow sought to understand why certain houses in the neighborhood – for instance, the workhouse on Poland Street or the brewery on Broad Street – were spared the disease. It turned out that the former had its own water point, whereas the workers of the latter simply did not drink water because they were entitled to rations of malt¹ liquor. Another interesting fact: these surveys revealed that the deaths of certain persons living far from Broad Street were due to the fact that they preferred to drink water from this point, considering it less dirty than the source near their home².

When he analyzed the mortality data by gender in London in 1848, he also realized that the epidemic outbreak affected men more than women, but in time, those numbers tended to equilibrium. The interpretation – paraphrased in modern terms – is related to gender mobility trends; men having a greater ability to disperse across the city because of their work, whereas women generally stayed at home. Men had a wider variety of sources to meet their water needs and therefore, a higher risk of contamination during an epidemic outbreak. At a second stage, when cholera spread throughout the city, men and women became equally at-risk³ populations. Snow also noted that valets and

1. "The men are allowed a certain quantity of malt liquor, and Mr. Huggins believes they do not drink water at all; and he is quite certain that the workmen never obtained water from the pump in the street" (Snow 1855).

2. "It requires to be stated that the water of the pump in Marlborough Street, at the end of Carnaby Street, was so impure that many people avoided using it. And I found that the persons who died near this pump in the beginning of September, had water from the Broad Street pump" (Snow 1855).

3. "The greater part of the female population remain almost constantly at home, and take their meals at home, whilst a considerable number of the men move about in following their occupations, and take both food and drink at a variety of places; consequently, in the early part of an epidemic, when the disease only exists in a few spots, the male part of the population is most liable to come within the operation of the morbid poison; but at a later period of the epidemic, when the cholera is more generally diffused, it may reach those who stay at home as readily as those who move about" (Snow 1855).

servants were the least affected by the disease because these workers used to live at the home of their “masters” in wealthy neighborhoods with better sanitary infrastructure. In this sense, they were affected by cholera less than their employers, whose daily routine was marked by numerous visits across the city.

The waterborne origin of cholera was thus established, as well as the first reflections on the role of mobility and the living spaces of populations, their exposure to pathogens and the spread of the epidemic. Based on these first results on cholera, Snow came to the conclusion that plague and yellow fever spread in the same way⁴. It only took a few more decades for research in epidemiology to show that both cases were vector-borne diseases. It is likely that the co-occurrence of mosquitoes with water points and the swarming of rats in slums could have misled him.

At the beginning of the 20th century, while the pathogens responsible for the main deadly diseases of the time and their mode of transmission were known, understanding the emergence and diffusion of new pathogens in epidemic form represented a challenge. This was the case of AIDS in the 1980s, of dengue in the 1990s and of SARS-CoV2 after 2019. Even if the main mechanisms contributing to the transmission and propagation of communicable diseases were known, their combinations in heterogeneous territories and within host populations with varied susceptibilities was quite complex. This required acknowledging the environmental and social criteria in which pathogens develop, the socio-demographic characteristics and behavior of hosts (in particular daily mobility and interactions) as factors of direct or indirect exposure to pathogens.

1.2.2. *The role of mobility in the spread of epidemics*

The spatial diffusion of an infectious and communicable disease implies that a pathogen multiplies and spreads from one host to another by means of a propagation channel (e.g. by airborne transmission or an animal vector). It is necessary to identify the actors involved in the transmission cycle (the pathogen’s potential hosts and transmission channels) as well as the processes interacting within the cycle, namely the pathogen’s propagation mechanisms, which lead to host contamination (Langlois and Daudé 2007).

4. In the section “Evidence that plague, yellow fever, dysentery and typhoid fever are communicated in the same way as cholera”.

The *contamination process* acts on the hosts: it represents the invasion and development of the pathogen within the host. Associated with infection, this process can alter the host's health condition after the amount of pathogens present reaches a certain threshold. Different health conditions can be defined for the host, such as:

- contamination, moment of the pathogen invasion;
- infection, when the pathogen develops within the host;
- viremia or contagiousness; the period during which the host emits a sufficient amount of pathogens to contaminate another susceptible person (who has not developed immune defenses capable of fighting the infection).

These health conditions may or may not be associated with symptoms, and the infection may end in partial or total recovery, with or without immunity to the pathogen, or result in death. This process can be globally represented by the number of hosts contaminated by the pathogen and the number required to qualify the disease as an epidemic. In fact, a certain density threshold of pathogens must be reached within an individual to trigger the disease, and reach a sufficient volume of infected individuals for an epidemic to break out.

The *propagation process* represents the pathogen's "behavior" emitted from each infected host. The pathogen's survival outside the host and before the invasion of a new host depends on environmental conditions. The influenza A virus, for example, survives between one and two days on hard surfaces and less than 12 hours on fabric. We can globally represent the propagation process by the number of active pathogens in the propagation channel (in this case, the environment). When this number is very low, particularly during the pre-epidemic phase, epidemic success is a stochastic process which depends on social and environmental conditions: high concentrations of people, absence of barrier measures, unfavorable environmental conditions. This is when highly targeted control measures can prove effective for preventing the spread of the disease. Pathogen viability emitted from the infected host also depends on the behavior of its possible vector. For instance, dengue viruses transmitted by mosquitoes can pose a risk to humans throughout the infected mosquito's life cycle. When assessing epidemic risk, the number of pathogens can then be confused with the population of infected vectors.

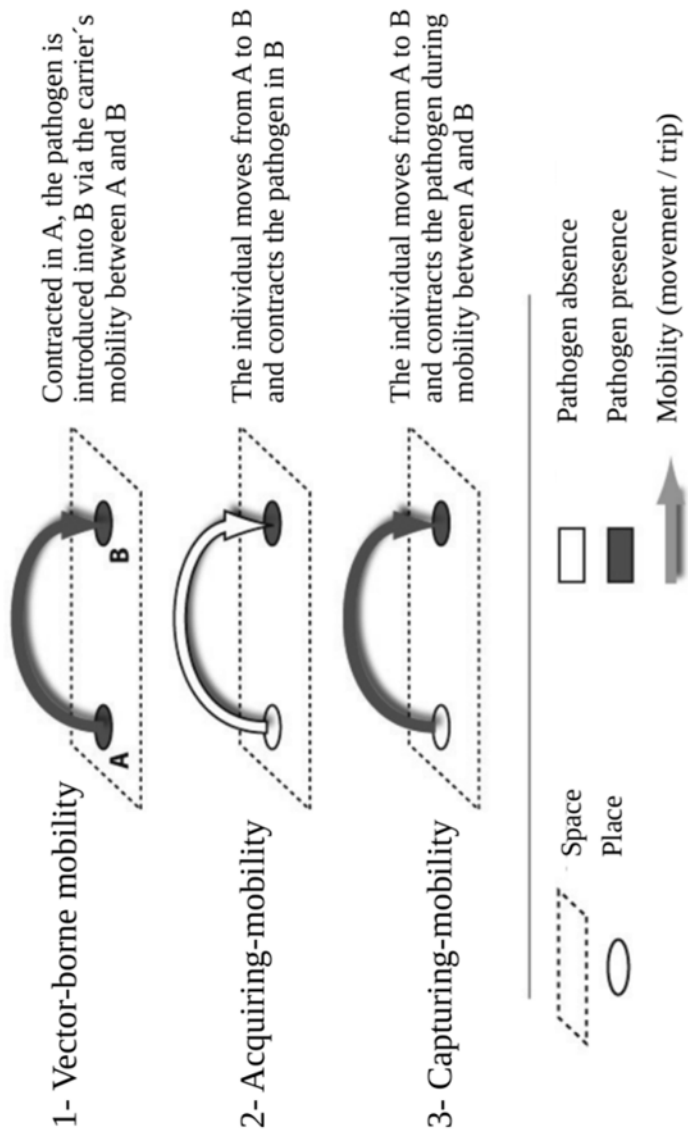


Figure 1.1. Mobility and types of pathogen transmission (source: Daudé and Eliot (2005))

By multiplying and differentiating the possible areas of contamination and propagation, the daily mobility of populations contributes to the pathogen's spatial diffusion. This offers risk reduction strategies: ceasing mobility reduces the number, nature and places for interactions between hosts and can stop – if not slow down – the spread of the disease (the intention behind quarantines and lockdowns). Knowledge about host mobility can be useful for predicting the next contamination places or for tracing contamination sources a posteriori, in view of discovering the index case (the first contaminated person). Finally, epidemiological studies in places in which populations interact (markets, schools, places of work, etc.) or in places in which hosts and vectors co-exist (gardens, parks, public transport, etc.) can be useful for identifying potential sites for the emergence of pathogens or contamination and for reducing risk at the source (school closures, vector control in gardens).

Mobility acquires a different status depending on the mechanism underlying the transmission from pathogen to host or vector. On the one hand, mobility can be an intermediary in the transmission of the infectious agent from place to place. The agent is transmitted from one place to another from an infection source (outbreak). Here, mobility is a vector of contamination (Figure 1.1, (1) “**vector-borne mobility**”). This is the case, for example, when a contagious person struck down with influenza introduces the virus into the workplace, by sharing a meal in a communal room with several susceptible colleagues. On the other hand, mobility can lead to situations featuring exposure to pathogens. In this case, contamination is the result of mobility that exposes the host through the places they frequent (Figure 1.1, (2) “**acquiring-mobility**”). Here, it is the susceptible colleague who visits the communal room at the workplace who is exposed to the pathogen, whereas at home, there was potentially no level of risk. Movement is “at the origin” of the agent's contamination. Lastly, mobility can constitute a contamination “channel”. It is during transportation that the host is contaminated by a pathogen and introduces it into the place of destination, in which case we speak of “**capturing-mobility**” (Figure 1.1, (3)). This is the moment when public transport becomes a key contamination place among traveling hosts.

These three basic forms of contamination are exclusive in the propagation of a pathogen at the individual-host level, but not at the group scale, and even less so at the scale of a whole population. This makes it all the more complex to identify the most vulnerable places exposed to contamination, the riskiest journeys, the duration and number of contacts most likely to

spread a given disease during trips. Field studies are a first attempt to appraise the role of these different forms of contamination underlying pathogen diffusion.

1.2.3. Case studies: epidemic dynamics based on the location of proven cases

One major difficulty when trying to assess the role and weight of daily mobility in the spread of an epidemic lies in identifying the circumstances, place and time of contamination. In the majority of cases, public health data and epidemiological surveys only identify sick people at their place of residence. Trying to reproduce the spread of the disease from clusters of patients can be misleading – a cluster of residents in a neighborhood does not necessarily convey that there is a high risk of residential contamination. By the same token, an absence of clusters based on the place of residence does not necessarily mean that another type of cluster would not have emerged if the cases had been located otherwise (e.g. at the workplace or a recreational area).

1.2.3.1. Looking for contact cases

To overcome this limitation, further research may be conducted in order to collect ad hoc data and try to reconstruct the person's itinerary a few days or weeks prior to the illness. The search for contact cases begins with a questionnaire administered to the hospitalized or sick person to accurately trace their movements after the appearance of the first symptoms, or a few days earlier, depending on viremia. Family members, acquaintances and colleagues are also questioned in order to identify the nature and duration of contacts with the sick person.

This method varies depending on the pathogen considered, as well as its contagiousness: a highly contagious virus will require extending the field of research to the greatest number of potential contact cases, with all the inaccuracies this entails. Asking a person how long they spent near another person and at what distance at a supermarket queue systematically induces a bias which over-represents risk. Applications for detecting proximity contacts using Bluetooth technology embedded in mobile phones were developed and widely distributed during the Covid-19 pandemic to partially address this issue. Nevertheless, there is currently little feedback from experience or studies clearly defining the contributions and limitations of those methods.

1.2.3.2. *Spatio-temporal cluster analysis*

In addition to the monitoring of individual trajectories, spatio-temporal cluster analysis assumes that detected cases are related both in time (duration of the contagion period and exposure duration) and in space (co-presence of people during contamination). But given the fact that determining the place of contamination is difficult, these studies are often conducted taking into account the places of residence of infected people. This boundary is even more significant for vector-borne diseases, as it dissociates the co-presence of healthy and infected hosts in the contamination/propagation process. An infected host living in district A can contaminate a vector during their morning visit to a park in district B. A few days later, this same vector can contaminate a host in district B (but be living in neighborhood C) during their nightly jog. Assessing the share of contamination that occurs while traveling outside the place of residence is a real methodological challenge. Phylogeography, which studies the distribution of the different genetic lineages of a virus, is an approach that partly makes it possible to overcome this obstacle.

Viruses are organisms that tend to mutate relatively quickly, and slight mutations accompany variations in infections and transmissions. The more differences there are between two viruses affecting two people, the smaller the chances for these individuals to belong to the same contamination chain. Here lies the origin of phylogeography, a science that studies the dynamics of viruses and analyzes the evolution of their genome, producing new variants. This approach is widely used at a global scale, particularly in the context of influenza, in order to estimate which strain will be dominant and which vaccine should be recommended.

This method was used by Salje et al. (2017) to determine the contamination chains of dengue fever in Bangkok. To do this, they first developed a phylogenetic tree including 640 genetic sequences of dengue viruses, drawn from blood samples of children having suffered from dengue fever between 1994 and 2010, and whose home address was known. Two individuals belong to the same transmission chain when the differences between the genomes of the same virus are relatively small. They found that 60% of the cases in the same chain had homes less than 200 m apart, which underlines the importance of residential neighborhoods in the spread of the epidemic. Despite this, we should observe that the subjects studied by Salje et al. (2012, 2017) were children, whose mobility is narrower than that of

adults, often limited to the neighborhood and to schools, in most cases, located near their homes.

In another study and with another method, Vazquez-Prokopec et al. (2010) studied the spatial distribution of homes of confirmed dengue fever cases during the 2003 epidemic, which affected the city of Cairns in Australia. It was the first major epidemic, which means that most patients had no prior experience of the virus (naïve population), and were therefore more susceptible to infection. Out of 383 confirmed cases, the team evaluated the aggregation and/or dispersal levels in spatial point patterns based on distance-methods and Ripley's K (1977). They showed that in 63% of cases, residences were located less than 800 m from the index case (the first confirmed case of dengue fever). They also applied a spatio-temporal partition method for identifying cases both close in time (less than 20 days) and in space (within 100 m of places of residence), which resulted in 18 independent clusters. Finally, the order of appearance of these clusters revealed that the disease had spread both in the direction of prevailing winds (which can facilitate the movement of mosquitoes) and along the axes of communication.

According to these two illustrative studies, while nearly 60% of the dengue cases studied presented a high probability of local contamination – thus confirming the usefulness of knowing the patients' place of residence – nearly half of them could not be associated with nearby cases, nor had themselves acted as the source of local contamination. In the absence of data on mobility in retrospective studies, different modeling-based methods have been developed to ascertain the role of mobility in epidemic propagation.

1.3. Modeling epidemics and daily mobility

The spread of an epidemic can be modeled as an expansion process in time and space. This process is activated by the circulation of individuals and/or groups who move differently depending on their socio-demographic profile and who transmit a pathogen following different modalities.

Geographic research on epidemics has identified three propagation modes:

- The first mode is based on an ordered sequence of places (**hierarchical diffusion**).

- The second one is based on the proximity between places (diffusion by **spatial contiguity**).
- The third one is based on the relocation possibilities of a phenomenon (**jump diffusion**).

However old, these works have remained largely confined to the disciplinary universe of geography and have not been mobilized by other disciplines studying the spread of epidemics.

1.3.1. Mobility, largely absent from the first mathematical models

A classical modeling approach in epidemiology assigns different clinical states to a host population and models the transition of individuals from compartment to compartment (Kermack and McKendrick 1927). The first clinical state concerns a fraction of the total population that is susceptible (S) to becoming contaminated. Part of this group may be contaminated and join the infected (I) compartment. After a certain period of time, infected people are cured and may develop (or not) immunity against the pathogen (R, standing for *recovered*). It is usually described in the form of stocks and their transitions: $S \Rightarrow I \Rightarrow R$ model, where $S + I + R$ equals the total population.

The SIR model is formalized as three differential equations, which describe the evolution of each of the S, I and R clinical states over time and according to an infection force β and a recovery rate γ :

$$\frac{ds}{dt} = -i\beta SI \quad [1.1]$$

$$\frac{dl}{dt} = i\beta SI - \gamma I \quad [1.2]$$

$$\frac{dR}{dt} = \gamma I \quad [1.3]$$

During each time step, the number of newly infected people is described as the product of the capacity β (or force of infection) of an infected (I) person to transmit a disease to a susceptible (S) person, depending on the number of possible i interactions between the two subgroups, S and I. After a certain period, the infected people eventually recover (R) following recovery rate γ – which is inversely proportional to the infection's duration – and acquire

immunity. The choice, number and order of the compartments depend on the epidemic studied. For example, clinical state E (Exposed) could be added, characterizing the period after which the host has been infected, has developed a viral load but is not yet contagious. In that case, we speak of the SEIR (Susceptible–Exposed–Infected–Recovered) model. This class of model describes pathogen transmission among the hosts of the same population, while other models make it possible to introduce inter-species interactions.

Ronald Ross was the first to propose a mathematical transmission model for malaria based on a system with two differential equations taking human hosts and vectors into account (Ross 1911a, 1911b). It primarily revealed that below a threshold value of the mosquito population, the malaria epidemic should disappear. The classic deterministic models used in the context of vector-borne diseases thus take up the bases of the SIR or SEIR models and add new compartments describing the different states of the mosquito, their growth and especially their interactions with humans (Derouich et al. 2003). These differential equation systems are solved mathematically and show great sensitivity to the initial parameters. In the case of a highly contagious disease, a high force of infection will lead to rapid contamination of the whole population, whereas a long period of infection implies that more people are sick at the same time. We should also bear in mind that these approaches only model epidemics at a given place, without taking into account human mobility between territories. Nevertheless, the need to integrate statistics on mobility and migration into these models⁵ quickly emerges (Waite 1910). However the difficulty in obtaining mobility data and the limitations of computing power/capacity at the beginning of the 20th century did not make such integration possible. Thus, the work of Ross and McKendrick – the basis for modern mathematical epidemiology – lies on the postulate of a total intermixing of the population, where all individuals are likely to interact, excluding socio-spatial structures and individual mobility behaviors.

1.3.2. The first steps of human mobility in the modeling of epidemics

While conducting epidemiological research on AIDS, Peter Gould (1993) clearly showed the limitations of these models, which were misdirected,

5. “As emigration and immigration vary considerably in different localities, and in the same locality at different times, their influence on the malaria rate cannot be satisfactorily dealt with except in particular cases where the necessary statistics are available; neither would results in general terms be of much practical use.”

partial and based on a series of population dynamic models that acknowledged the temporal dimension but not the spatial one. Since the end of the 1960s, so-called metapopulation models have integrated mobility, primarily as population fluxes between the cities and across the country (Baroyan and Rvachev 1967). In the context of epidemics, these are more or less complex models (SIR, SEIR, etc.) that are able to take into account subgroups with different serological characteristics as well as the commuting of individuals between different geographical areas (Sattenspiel and Dietz 1995; Arino and van den Driessche 2003). In concrete terms, the SIR model is replicated according to the number of cities in the urban system. The population volumes for each city are then used for calibrating each of these sub-models (Figure 1.2).

Interaction rules are integrated into the equations to convey the arrival to/departure from a city towards another city. Inter-city population movements among the infected host compartments are explored in depth. These journeys, processed in the form of mobility flow (or travel probabilities), are generally represented by origin/destination matrices. These metapopulation models, still widely used today, provide powerful simulations at the macroscale because national (motorway flow) and international transport (air flow) are the main modes of propagation of a pathogen at this scale (Sattenspiel and Dietz 1995; Arino 2005). Like any model, they require good estimates of the flow matrices between the different zones, and depend on good quality mobility data (Tizzoni et al. 2014).

Simplified versions, with a smaller number of differential equations, have also been developed to reflect the arrival/departure of populations from a territory. The dengue epidemic that took place in Madeira in 2012 was simulated using air traffic towards the island as a proxy for the importation of the epidemic (Lourenço and Recker 2014). Another study adopted a twofold approach: first, an epidemic was simulated in the area of Karachi (Pakistan) – where dengue fever is endemic. This stage made it possible to estimate the number of infected persons. They then introduced infected people into other areas in the north of the country (where dengue fever is seasonal), matching the likelihood of interactions between different country sub-districts (Wesolowski et al. 2015).

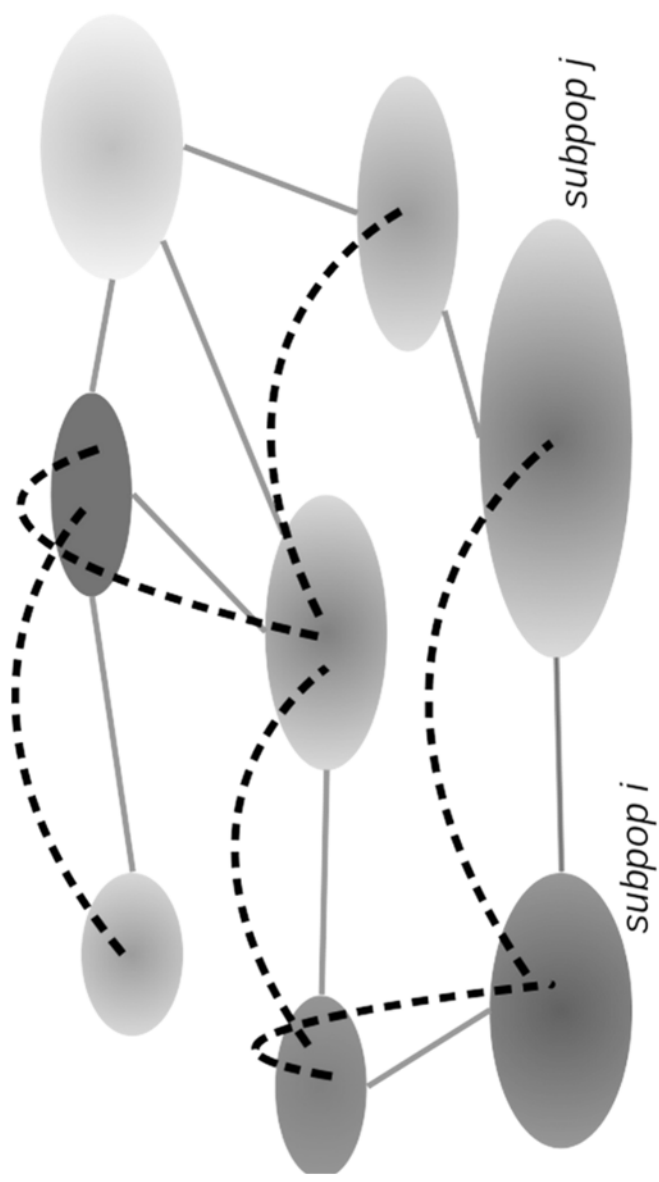


Figure 1.2. Illustration of the principle of spatialized compartmental models, with transport infrastructure (gray links) and inter-city (ellipses) flow (black dotted lines) (inspired by Pastor-Satorras et al. (2015))

There are solid arguments for choosing differential equation modeling:

- Firstly, it offers the possibility of understanding the global properties of a system through a small number of macroscopic variables (parsimony principle).

- Secondly, all other things being equal, it helps estimate the maximum number of infected persons during an episode if current trends continue.

- Thirdly, it helps to identify threshold effects in the dynamics of a phenomenon. Threshold effects are particularly interesting when several strategies for fighting an epidemic are in competition and when striving to find the path offering the most effect and the fewest constraints.

Yet, these models are hampered by significant barriers to understanding epidemics. First and foremost, the thorough and homogeneous mixing of the population, which results in a lack of consideration for social and spatial structures. This leads to an extremely simplified vision of the characteristics and behaviors of individuals in the population. Even if it is possible to multiply the compartments to represent the different population categories, and to replicate these compartments to signify different territories, calibrating all the parameters that govern these differential equation systems quickly becomes a challenge.

A second limitation of these models: they cannot track isolated individuals and tend to overestimate the propagation phase and epidemic peaks (Ajelli et al. 2010). While these models may be effective in the middle of an epidemic, they struggle to represent and simulate the emergence of epidemics, at the moment when “everything is at stake” by only relying on stochastic interactions between a few localized hosts. They also have difficulty in explaining the forms of diffusion, largely dependent on the spatial structures that existed prior to the epidemic. A necessary complement would require studying the ordinary dynamics of the city and the spaces involving population mobility, alongside inter-host contagion mechanisms, to better understand the geography and local spread of communicable diseases.

1.3.3. *Deciphering activity spaces and daily mobilities*

Most individuals travel on a daily basis to fulfill their obligations. They do so regularly (work, education, purchase of goods and services) or occasionally (health, administration). They also travel to fulfill their life plans (socialization, leisure), and all of these journeys are subject to a series

of constraints (accessibility, alternative choices in more or less constrained schedules) and opportunities (concentration of activities made possible by urban densification). Depending on the nature of the places frequented, these mobilities can increase or decrease the likelihood of exposure to pathogens.

1.3.3.1. *Complex host–pathogen system*

When stepping foot in different territories, individuals can be exposed to risky environments (Figure 1.1). The pathogenic areas described are in line with the work on pathogen complexes by Sorre (1933). The endemicity of dengue fever in Indian cities is related to environmental conditions (monsoon period, favorable temperatures, availability of breeding sites), letting mosquitoes survive and spread the virus within the population residing within and nearby that pathogenic area. But the environment frequented by individuals is not the only contributing factor to the geography of diseases. Their origin also lies in social factors such as lifestyle or work, individual and collective behavior, resource levels, etc. Disease is thereby considered as the by-product of multiple interacting factors. Through his analysis of the host–pathogen system, Picheral (1983) conceptualized the intertwining of the links between factors provoking, conditioning and predisposing exposure to certain diseases. Health risks are explored as a complex articulation of factors within a complex host–pathogen system.

The concept of *complex pathogenic system* proposes a representation of the global dynamics of an epidemic, pinpointing its elementary entities and interactions. These entities are the pathogens and the prerequisites favoring their development, the possible hosts with their own characteristics and behaviors, and finally, the environmental and societal conditions that evolve over time and are differentiated in space.

The study of a complex pathogenic system is based on the analysis of the different interactions between the system's entities, which act at different scales (or levels), as illustrated by the example of dengue fever (Figure 1.3). Even if a person's vulnerability to a pathogen depends on individual risk factors (age, risky practices, etc.), it also depends on macroscopic context effects (herd immunity, population density). Environmental risk can be measured along different scales, at the level of the district (with the presence of potential mosquito breeding sites) or at the macroscopic level such as meteorological conditions, which influence the vector population dynamics (Daudé et al. 2015).

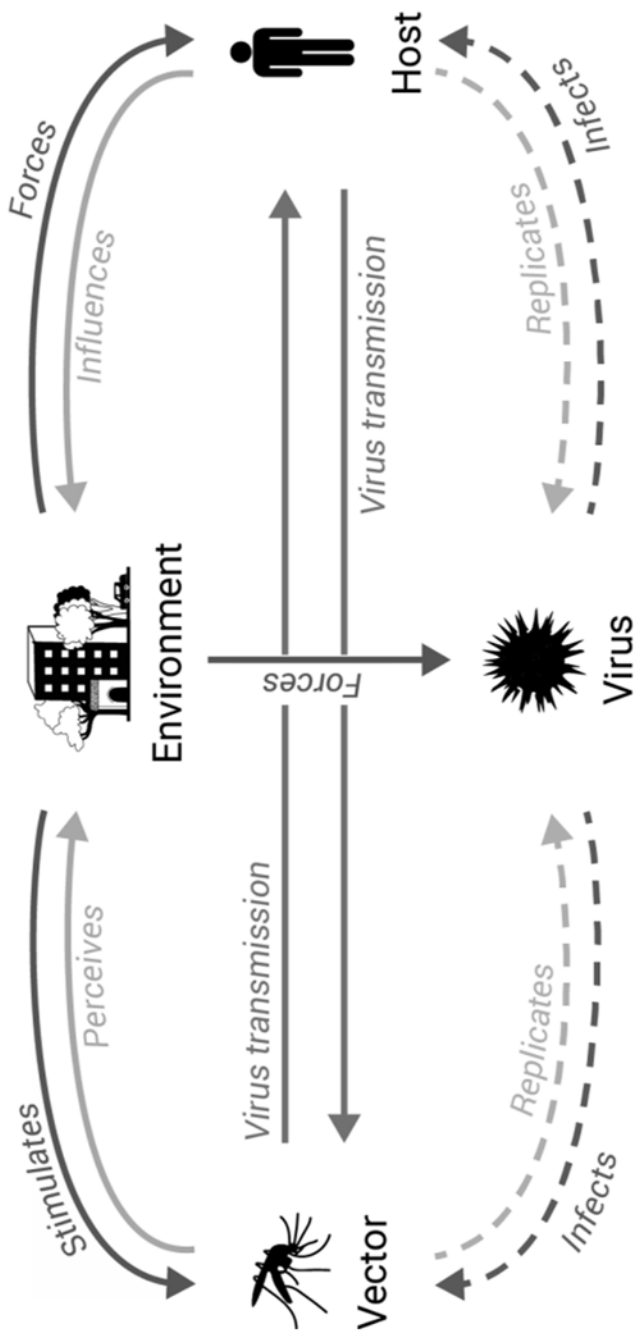


Figure 1.3. Complex host-pathogen system of dengue fever (source: Misslin and Daudé (2017))

As mobility also occurs at different spatial scales, it can favor the contact between hosts and vectors in high-risk environments. While the mosquito is the viral vector in its living environment (limited to a few hundred meters from its place of emergence, see Ordóñez-Gonzalez et al. (2001)), humans are the main virus carriers over longer distances: at urban (Stoddard et al. 2013), national (Cummings et al. 2004) or worldwide scales (Tatem et al. 2006). The integration of daily mobility then becomes key to understanding the differentiated geography of communicable infectious disease epidemics, in particular at the micro- (neighborhood/district) and the mesoscale (city).

1.3.3.2. *Time-geography and activity space*

The activity space is defined as the portion of the urban space individuals frequent in the course of their daily activities (see Chapter 2). These are all the places frequented by an individual in relation to different activities (work, leisure, etc.). Activities are defined in time (by an ordered sequence of events, as in a schedule) and in space (places in which such activities take place) (Hägerstrand 1970). Depending on their level of spatio-temporal flexibility and the individual's degree of freedom, activities can be classified into two main categories. On the one hand, fixed (or constrained) activities are conducted in predefined places, where individuals do not have much leeway as to their execution time (such as sleeping, working, studying, etc.). On the other hand, flexible activities are less restricted in time and/or space, and require less planning (e.g. impromptu shopping, outings, etc.).

This conceptualization endows each individual with a spatialized time diary, portraying more or less regular events. The study of these activity spaces brings to light the differences between individuals, both in terms of the distances traveled and the number and location of activities carried out throughout the space. They also reveal the spatio-temporal co-presence of individuals in the city. When coupled with socio-economic and demographic information, it is possible to observe mobility and to quantify its potential according to socio-economic status, age, social network size, type of residential area, etc. The differentials in access to basic resources such as education places for the youngest or health systems for the poorest can also be assessed through this approach.

Figure 1.4, inspired by the work of Hägerstrand, shows the activity space for four persons belonging to the same household, performing different activities throughout a day. Each of these activities is associated with different health risks depending on the type of place frequented, the people

gathered and the time of day. As the activity space associates an individual with the places frequented, it is well suited to estimate a person's potential exposure to various pathogens (Perchoux et al. 2013). In absolute terms, by cross-referencing all the spatialized time diaries of individuals in a population, and by knowing the incidence rate of a communicable pathogen within that same population, we could calculate the probability of there being one infected person every n people present at a place, and therefore predict the contamination risk knowing the number of people present in the same place at the same time (Stoddard et al. 2009; Vallée 2017).

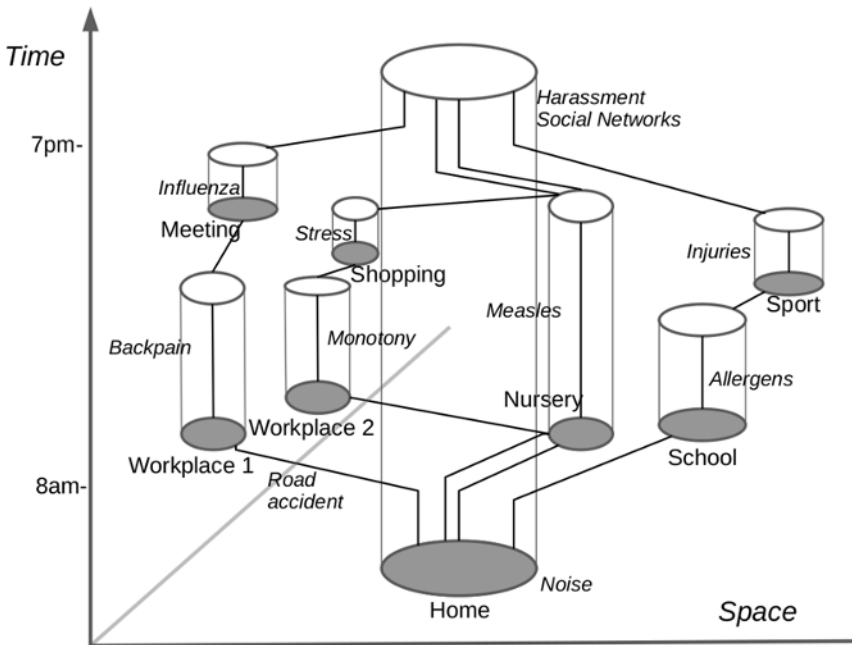


Figure 1.4. *Representation of the activity space within a health context (source: Schaerström (1996))*

Taking into account activity spaces and individual daily mobility at differentiated geographical spaces (population density, land use or an ecological niche favorable to the presence of vectors), is the basis for empirical and methodological advances in epidemiology (see Chapter 2) and public health because it feeds the reflection on the production of social inequalities in health (see Chapter 3).

1.3.3.3. The contribution of agent-based modeling to the study of the role of daily mobility in the spread of communicable diseases

Together with the theories of complex systems (Morin and Lemoigne 1999) that fuel reflections on complex pathogenic systems, modeling methods have been developed to overcome the obstacles of mathematical models. The challenge is then to model distributed and spatialized processes. In this context, the agent-based model (ABM) has been the main methodological framework used so far (Treuil et al. 2008). These computer models simulate the interactions/behaviors between entities – called agents – in order to better understand the functioning of a system and how it may evolve. Multi-agent systems model and simulate dynamic phenomena in environments in which the individual components are highly heterogeneous, and make it possible to add qualitative and quantitative rules at different hierarchical levels (Daudé 2004; Sanders 2006). These systems can also make it possible to take into account emerging collective structures stemming from the individual behavior of agents. More and more agent-based work in epidemiology is looking into the question of individual mobility, facilitated by the explosion of available data for the calibration and evaluation of these models (see section 1.4). Mobile synthetic populations in a virtual city can simulate activities in various city areas, at different times, following diverse movement rules and tendencies, as well as assess the location/duration of exposure to communicable diseases.

1.3.4. EpiSim: a concrete example of epidemic modeling in urban areas acknowledging daily mobility

One of the first dynamic models of a communicable disease epidemic represented by an ABM was applied to influenza in six counties in southern California (Eubank et al. 2004). In this model (called “EpiSim”), 18.8 million individuals were generated, living and moving in a space totaling 2.3 million places. These places set up a physical space (home, school, building, etc.), and were connected to one another via a road network graph. Individuals were partitioned into 6.3 million households, whose characteristic distribution was representative of the demographic data of the neighborhoods in the six counties. In addition to belonging to a household, each individual could belong to one or more groups (study group, working group, colleagues, etc.), which facilitated contact opportunities among individuals. In the EpiSim model, each individual had a daily schedule, listing activities with a start and end time, and a venue. These schedules

were designed on the basis of a survey conducted on a population sample composed of 60,000 households (National Household Transportation Survey, NHTS). The model made sure that the number of affected individuals in the different locations matched the occupancy statistics for such locations and that the commuting distances from home to work, and between intermediate locations (neither to nor from home), corresponded to those presented by the NHTS.

From this model, different epidemic events were simulated to suit different intervention scenarios: quarantine strategies, vaccination scenarios, mask wearing, avoidance behaviors, school closures, etc. By being modeled at the individual level, the disease and its progression within the population could be followed by tracing the interpersonal contacts at the origin of the contamination and the pathogen's propagation by the individual. In concrete terms, between the moment an individual arrived at a place and left it, the model recorded all of the contact times with other individuals and the subject's relational network dynamic graph. The model was thus able to calculate the disease transmission probability based on the contact time between the infected and the healthy. The likelihood of an individual becoming contaminated during an activity was a function of the cumulative time spent with contaminated individuals. A baseline transmission rate (per contact/hour) defines the risk of transmission.

To go further in inter-individual differentiation, EpiSim models the viral risk of transmission as being equal to the viral transmission rate multiplied by a susceptibility potential and an infection potential. The susceptibility potential, specific to each individual, depends on age, gender and medical history. It is attributed to each agent at the start of the simulation and does not vary. The infection potential, also specific to each individual, depends on the infection stage in the contaminated person, on the treatments administered and evolves throughout the simulation.

Modeling at the individual level, grouped per household, social groups and localizable during the simulation makes it easier to conduct statistical analyses at different spatio-temporal or socio-demographic scales. The distribution per number of daily contacts/per agent is useful to precisely track those who had the most contacts, and thus identify potential "super spreaders". The distribution per contact duration within the population and for each type of activity is used for measuring the dose-response effect and associates it with different types of places or activities. These possibilities offer solutions for many methodological limitations related to the spatial and

temporal context in epidemiology. Thanks to them, it is now possible to identify the spatio-temporal co-presences of individuals as well as the contamination contexts of a pathogen that is directly communicable from person to person, or indirectly transmitted via a vector.

For an entirely susceptible population of agents, the introduction of an influenza virus into the EpiSim model offers some lessons regarding the simulation zone. Due to their stochastic character at the early stages of the epidemic, simulation results may vary. This is because the low number of cases at the onset can cause the emergence of the epidemic to fail, slow it down or accelerate it, depending on the location and the first households affected. In a scenario in which an influenza virus is freshly introduced into a totally healthy population, the number of cases at the end of the epidemic reaches approximately 8,650,000 cases, with 46% of the population infected, among which 30% are asymptomatic. Due to the conservation of individual footprints of activities during contamination, the simulations show that the fraction of new infections at home increases from 50% at the onset of the epidemic to nearly 65% by the end of the epidemic. This rises from 10 to 15% for professional activity places, and decreases from 25% to 3% in schools.

Detailed analysis of the attack rate (the number of new cases divided by the size of the susceptible population) also provides accurate information on the progress of the influenza virus. The work of Eubank et al. (2004) showed a strong correlation between the attack rate and the average size of households and the school population share. The attack rate was higher in neighborhoods accommodating large families, where the number of school-attending children is higher. While nearly 60% of the variation in the attack rate by sector was negatively correlated with the per capita income, neither the population density, nor the population per sector were correlated with the attack rate. This last result clearly shows the importance of taking into account individual activities and household composition.

EpiSim illustrates the benefit of an individual-based approach for studying communicable diseases. This approach comprises a variety of risk scenarios, acknowledging the vulnerability of people, population fluxes, responses obtained and timelines. Many scenarios have been tested in EpiSim, providing useful decision-making support for decision-makers: “laissez-faire” to seek herd immunity, opting for an egalitarian or equitable vaccination strategy based on random selection, or for a solution relying on the properties of social networks, for example, by favoring the vaccination of people who have a higher number of contacts than average. As the model is

able to generate dynamic networks of contacts by monitoring the agents who are in the same place at the same time, experts can evaluate to what extent the vaccination of one agent, and all of the person's contacts, can be effective at the global level.

Since the first pioneering model, the use of this type of approach has become widespread to other urban areas. But an important obstacle remains: the availability and acquisition of suitable data, in particular, for building activity schedules.

1.4. The importance of mobility data in the analysis and propagation modeling of urban epidemics

There are several methods for acquiring daily mobility data, each with its own advantages and disadvantages regarding the collection method or the type of data collected.

1.4.1. Addressing mobility by collecting *in situ* data, or active data

Until recently, few solutions were available for analyzing daily mobility. It was necessary to draw on institutional data from a large sample of individuals, data had to take mobility into account, be recent and available in the study area, and a quantitative approach was required. In this context, we are referring to home–work data from censuses or origin–destination surveys (Commenges 2013), or so-called “time use” surveys. For example, household travel surveys⁶ in France examine several thousand individuals within the same urban area and provide highly detailed information on the activities performed, the commuting times and the modes of transport used on a certain day of the week. These studies are considered reliable thanks to their proven sampling methodology and contain socio-economic and demographic information. They often serve as the basis for evaluating the quality of other data sets (e.g. Calabrese et al. (2011a); Lenormand et al. (2014)). In the absence of this type of study, or the deliberate choice to carry out a more qualitative or hybrid survey, it is always possible to collect data by conducting individual interviews, either directly or remotely (online or telephone questionnaires).

6. Now called “CEREMA-certified mobility surveys” (EMC²).

For example, Mossong et al. (2008) conducted research by administering a questionnaire to over 7,000 people from eight European Union countries. Not only did they collect information on age groups/gender, but also on the 97,000 persons that the sample members had been in contact with (involving physical contact or not). They also recorded the type of place in which such contact had taken place, as well as its duration. These data were then used to calibrate an epidemic mathematical model used for exploring a variety of scenarios. Thanks to this model, experts can define under which conditions the youngest would be likely to record the highest incidence rates in the case of an emerging disease when the entire population is naïve, and to gauge the effects of specific measures for these age groups.

For the last 20 years, many studies have been based on GPS footprints collected from consenting respondents (see Shen and Stopher (2014) for a literature review or Chapter 2 of this book). For example, based on GPS data elicited from some inhabitants in Iquitos (Peru), Stoddard et al. (2009) were able to reconstruct the activity spaces of individuals, which were then compared to the density of mosquitoes responsible for dengue fever. They then calculated an individual risk level according to the activities performed, mosquito density, and exposure time/duration. In this specific case, places of residence recorded the highest risk levels. After pursuing their study via the contact tracing method (which reflects social ties), researchers concluded that commuting between these places significantly contributed to the local propagation of dengue fever (Stoddard et al. 2013).

When conducted among individuals, these types of GPS-based studies can collect highly accurate geographical information. However, they cannot be applied to too large samples, due to cost reasons (in time and/or labor), nor over too long periods, because they are fairly intrusive and constitute a logistical constraint for the respondent.

1.4.2. Approaching mobility through data collected ex situ, or passive data – digital footprints and online data

Since the end of the 2000s, technological development in the fields of communications and computing has been characterized by the widespread use of mobile phones and, more recently, social networks. This has been facilitated by improved access to communication networks (better speed, lower prices, technological breakthroughs in mobile phones, etc.). These online and roaming behaviors generate massive data and (sometimes

geolocated) digital footprints. As such, they can be used for analyzing individual or collective mobility.

1.4.2.1. *Studying mobility using data from mobile telephony*

A mobile phone that is switched on will continually seek to connect and identify itself with the cell phone tower – base transceiver station (BTS) – offering the best signal (hypothetically the closest one) in order to be reachable or to be able to make a call at any time. The telephone's location can therefore be associated with that of the BTS, which is known. In theory, it is possible to trace a telephone (and its carrier) over time, with a spatial accuracy that depends on the network connectivity of the BTS (Figure 1.5(a)). In France, this spatial accuracy ranges from a few hundred meters in densely populated urban areas to a few kilometers in rural areas. Spatio-temporal information on connections represents a considerable amount of data, not necessarily stored by operators. Furthermore, they are difficult to access for research purposes because operators do not wish to compete with offers they can market themselves. Despite this, other, more “degraded” data can also be accessed. Call detail reports (CDR), which are recorded for practical reasons (invoicing and service continuity) do not offer complete tracking in space/time because the telephone's location can only be tracked when a telephone call (or message) is transmitted or received (Figure 1.5(b)). However, CDR data have the advantage of providing social network analysis based on the number of phone calls between two people. These CDR data have been used for research purposes since the mid-2000s, particularly in the context of human mobility flow studies.

At the urban scale, data from mobile telephony has revealed the high level of spatio-temporal regularity of human mobility, which can be explained by the fact that people spend most of their time in a few places (González et al. 2008). The same team also studied the entropic nature of the mobility of 50,000 users whose regular phone calls sufficed to sow continuous footprints over time. They concluded that 93% of journeys could be predicted, and that this rate was constant in all the population, regardless of the distances typically traveled (Song et al. 2010b). This result should nonetheless be put into perspective. In fact, in cities, individuals often spend most of their time between their home and their workplace, at relatively regular times. There is a high probability that an individual will be present at the same place between two contiguous time periods, especially at night or during working hours. It is also necessary to take into account the level of spatial accuracy, of at least hundreds of meters.

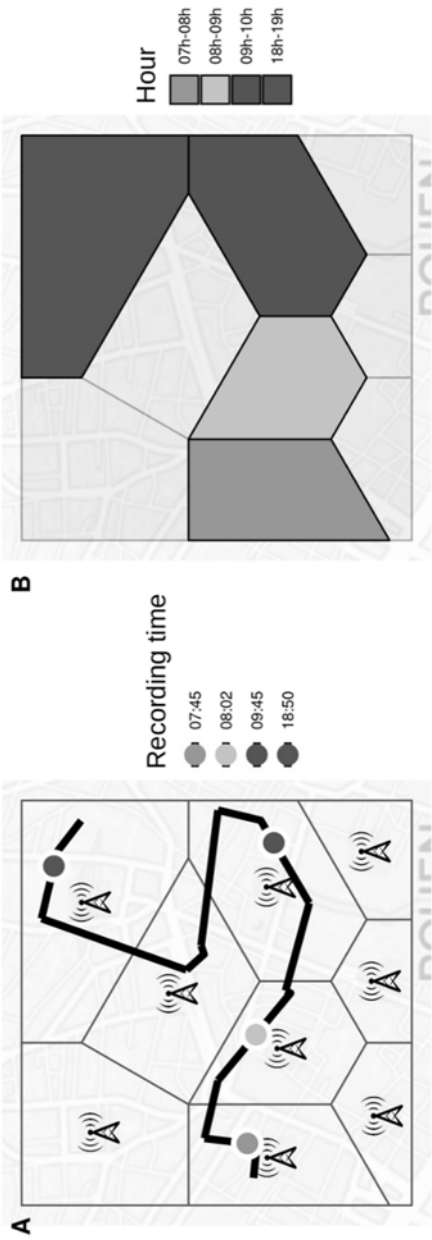


Figure 1.5. Location of a telephone near relay antennas (source: Montjoye et al. (2013)). a) Exact telephone location; b) Voronoi polygons where the phone has been active and where the user will therefore be located according to CDRs

In line with the work by González et al. (2008), a mobility model has been proposed: the exploration and preferential return (EPR) model. According to this, when one agent decides to change location, they can either visit a new location (with a probability that decreases with the number of locations already visited), or return to an already visited location (Song et al. 2010a). This relatively simple formalization made it possible to reproduce two observed pieces of information at the same time: the increase in the number of places frequented by an individual over time, and the probability distribution of their presence at each place.

Many studies on mobility have been conducted using these telephone data sets in different contexts, either to assess the sample's representativeness, to search for mobility laws, to analyze spatial interactions or to establish comparisons with classic mobility models (gravity, radiative, etc.). See Blondel et al. (2015) for a literature review on the analysis of mobility based on these data and Barbosa et al. (2018) for a more specific review of mobility models inspired by CDRs.

1.4.2.1.1. Telephone footprints in the context of infectious disease outbreaks

A growing number of studies on the relationship between mobility and epidemic propagation relies on telephone footprints. For example, regarding the 2014 Ebola outbreak, Wesolowski et al. (2014) pointed out that rapid and coordinated access to telephone data offered real-time monitoring of epidemic propagation. Data on population dynamics are also essential in the case of malaria at a regional scale (Wesolowski et al. 2012; zu Erbach-Schoenberg et al. 2016), and as for dengue, human commuting plays an important role in epidemic propagation, in addition to the hyper-localized movements of mosquitoes. Despite the fact that gaps in epidemiological data make it difficult to model epidemic propagation (Buckee et al. 2013), they do offer an estimate of the main flow responsible for importing the disease and high-risk areas, information that can be used to improve the monitoring system (Buckee et al. 2017).

1.4.2.1.2. Telephone footprints in the context of epidemics in urban areas

Frias-Martinez et al. (2011) built an agent-based mobility model applied to H1N1 influenza, which takes into account mobility and social networks. Each agent is assigned a probability of being in a cell (a Voronoï polygon corresponding to the spatial influence of the BTS) at a given time t ,

a variable that depends on the day of the week/weekend, all of which is based on location history. The social network aspect is determined by the number of calls between two people, especially when they are around the same BTS, which generally pre-announces an imminent physical encounter (Calabrese et al. 2011b). Frias-Martinez et al. (2011) constrained their model to the following consideration: there was a higher probability of contact between an infected person and a healthy person when the individuals had already contacted each other on the phone and were in the same area. The authors then applied their model to H1N1 cases recorded in 2009 in a city in Mexico. They estimated that the travel restrictions imposed by the government might have reduced the peak of infected people by 10%.

Massaro et al. (2019) studied the influence of mobility on the dynamics of the dengue epidemic in Singapore between 2013 and 2014. To do so, they built daily mobility models, which they then associated with compartmental models describing the interactions between humans and mosquitoes. Drawing on telephone data from 2.3 million people (about one-fourth of the total population), they built an agent-based model reflecting home–work commuting. The place of residence and the main place of activity were associated with the antenna where the person had made or received the most calls or messages at night and during weekdays. They then built three other ABMs (though more related to origin–destination matrices). One of them included places of residence based on CDRs but randomly distributed places of main activity. Another one involved defined primary activity locations in such a way that overall commuting distances followed a truncated Levy distribution, as suggested by González et al. (2008). Finally, in the third model, the places of residence corresponded to those from the census, but workplaces were estimated from a radiation model (Simini et al. 2012). In all cases, individuals move in 320 m² cells associated with a stochastic compartmental model of dengue. The development of mosquitoes and larvae (larvae mortality, incubation rate, etc.) in each cell of the grid solely depended on temperature. Other parameters, such as the transition rates between E and I, and I and R, were fixed. The number of mosquitoes per human and the number of bites were estimated by calibrating the mobility models in such a way that the correlation coefficients between the simulated and observed dengue cases are closest to the city’s scale. Following Lourenço and Recker (2014), they applied a differential equation model for each of the compartments (vectors and hosts), for each cell and each time lapse. In order to be able to follow the serology of each agent, they estimated the new clinical state by performing an independent draw according to a binomial distribution where probability P depended on the total number of

people having changed their clinical state in that cell at time t . The results of the models were then compared to the dengue cases recorded by the Singaporean monitoring system.

Despite being hindered by some limitations (the entire population is susceptible, only one dengue serotype is taken into account, the infected agents do not change their mobility practices, which are themselves limited to commuting, etc.), the study nonetheless showed that the models inspired by CDRs offer the closest results to observations, both in time and in space, that the diffusion processes of dengue fever in urban areas depend on mobility and are in no way random.

1.4.2.2. *Digital footprints and online data*

The arrival of increasingly sophisticated and Internet-connected telephones goes hand in hand with the emergence of social media platforms. These enable users to interact with the world, and – more or less consciously – to share information about the places visited and the times of such visits: these data are known as geolocated digital footprints. For example, X (formerly Twitter), the popular social network, allows members to share their precise location (GPS coordinates from their phone), by adding content to their message (or tweet). Some social networks have quite large penetration rates at a global scale (e.g. X, Facebook, Instagram, etc.), even if their use varies depending on user age, socio-professional category and geographical area (Facebook is not used in China, whereas Russians prefer Vkontakte).

In the context of research on mobility, geolocated tweets initially made it possible to understand part of the flow at a regional and a global scale (Hawelka et al. 2014; Jurdak et al. 2015). In urban areas, they also showed different mobility potentials matching different social categories in Chicago (Luo et al. 2016) and Delhi (Cebeillac and Rault 2016). By cross-checking these geolocated tweets with land use layers from Google's Points Of Interest (or POIs), individual activity spaces were created, studied and compared with the spatio-temporal discontinuities of dengue fever cases observed in Bangkok (Cebeillac et al. 2017, 2018).

More recently, Data for Good at Meta⁷ has provided data on flow that are aggregated every eight hours into relatively large grids covering several kilometers or at a municipal scale, depending on data sets (Galeazzi et al. 2021). These data on the movement of individuals aggregated in time and

7. See: <https://dataforgood.facebook.com/>.

space revealed the effect of lockdown on local and regional mobility during the Covid-19 pandemic, and favored the study of correlations between the two at these scales (Shepherd et al. 2021). Google and Apple have followed the same path and made aggregated and pre-processed data available, providing information on the activities of users who have activated geolocation⁸ on their mobile devices. Nevertheless, as these data are aggregated, they do not make it possible to understand the direct links between individual mobility behavior and pathogen contamination/propagation processes for now.

In line with work that associates areas frequented by individuals and types of activities present in these areas, Chang et al. (2021) have proposed studying the infection places by coupling a SEIR model to geolocated digital footprints. These provide the precise geolocation of nearly 100 million Americans in 10 of the largest cities in the country. They have been offered as part of a cooperation agreement with the company Safegraph, so as to better understand Covid-19. The company collects and cross-references individual and geolocated data with partners and third-party⁹ applications, without further information on their origin. Thanks to them, experts can estimate the area of residence as well as the types of places visited (POIs), and the median duration of a visit. Inspired by these data, researchers applied a network SEIR model in which contamination could take place at home or at a POI. The infection rates calibrated by their simulations varied depending on POI size, median visit duration and density of susceptible persons. They were able to show that certain POIs play a more important role in the contamination chain, including restaurants, gyms, bars and cafes – we should nonetheless note that education areas and workplaces were not part of the study. For example, according to their simulations for the city of Chicago, 10% of the POIs represented 85% of the infections. They also detected various vulnerabilities relating to the income level of the area of residence. The mobility of people residing in low-income areas declined less than that of persons living in richer areas, probably because teleworking practices are less widespread or even feasible in the professional categories occupied by the least privileged social classes. Another interesting observation: the types of places that this subgroup frequented were more at risk, with higher population densities, longer exposure durations and higher visit frequencies.

8. See: <https://covid19.apple.com/mobility> and <https://www.google.com/covid19/mobility/>.

9. See: <https://www.safegraph.com/privacy-policy>.

A wide variety of mobility data can be used, either collected directly from respondents or from digital footprints. Their level of representativeness and accuracy (social, spatial, temporal or semantic) conditions their use, both from the point of view of the phenomenon's description and its prospective modeling.

1.5. Conclusion

Daily mobility is complex, specific to each individual and variable in time and space. Although it plays a key role in the diffusion process of communicable infectious diseases at all scales, its consideration has only been gradual. First, this can be explained by the complexity of mobility behavior, whose relation to pathogens is difficult to ascertain. On the other hand, the pathogen systems are also complex in themselves. It also took time for the relevant concepts developed in different disciplines to be intersected (geography, computer science, epidemiology, etc.). Technical limitations have to be mentioned, in particular those concerning computing capacities/time, as well as the availability of suitable data.

Although the use of massive data from mobile telephony or geolocated digital footprints is disconnected from their original goal and de facto raises ethical issues in relation to surveillance and privacy, these massive data nonetheless make it possible to better understand daily mobility and to shed light on the diffusion processes of diseases at different scales. Their use in the context of dynamic, distributed and spatialized models is likely to persist over time and undergo significant developments. Among these, a reflection on the association with other databases and the limitations this poses will have to be encouraged. In fact, these telephone data or data from social networks are data without context: no access is granted to socio-demographic information on the profile of individuals, the reasons for traveling, emotional states, the culture of risk or the acceptance of instructions. However, these elements are just as important as mobility when assessing the vulnerability of individuals, and populations, to a certain pathogen.

1.6. References

- Ajelli, M., Gonçalves, B., Balcan, D., Colizza, V., Hu, H., Ramasco, J.J., Merler, S., Vespignani, A. (2010). Comparing large-scale computational approaches to epidemic modeling: Agent-based versus structured metapopulation models. *BMC Infectious Diseases*, 10(1), 190.
- Arino, J. (2005). A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology*, 22(2), 129–142.
- Arino, J. and van den Driessche, P. (2003). A multi-city epidemic model. *Mathematical Population Studies*, 10(3), 175–193.
- Barbosa, H., Barthelemy, M., Ghoshal, G., James, C.R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J.J., Simini, F., Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports*, 734, 1–74.
- Baroyan, O.V. and Rvachev, L.A. (1967). Deterministic models of epidemics for a territory with a transport network. *Cybernetics*, 3(3), 67–74.
- Blondel, V.D., Decuyper, A., Krings, G. (2015). A survey of results on mobile phone datasets analysis. *EPJ Data Science*, 4(1) [Online]. Available at: <http://www.epjdatascience.com/content/4/1/10>.
- Buckee, C.O., Wesolowski, A., Eagle, N.N., Hansen, E., Snow, R.W. (2013). Mobile phones and malaria: Modeling human and parasite travel. *Travel Medicine and Infectious Disease*, 11(1), 15–22.
- Buckee, C.O., Tatem, A.J., Metcalf, C.J.E. (2017). Seasonal population movements and the surveillance and control of infectious diseases. *Trends in Parasitology*, 33(1), 10–20.
- Calabrese, F., Di Lorenzo, G., Liu, L., Ratti, C. (2011a). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing*, 10(4), 0036–44.
- Calabrese, F., Smoreda, Z., Blondel, V.D., Ratti, C. (2011b). Interplay between telecommunications and face-to-face interactions: A study using mobile phone data. *PLoS ONE*, 6(7), e20814.
- Cebeillac, A. and Rault, Y.-M. (2016). Contribution of geotagged Twitter data in the study of a social group’s activity space. The case of the upper middle class in Delhi, India. *Netcom. Réseaux, communication et territoires*, 231–248.

- Cebeillac, A., Huraux, T., Daudé, É. (2017). Where? When? And how often? What can we learn about daily urban mobilities from Twitter data and google map in Bangkok (Thailand) and what are the perspectives for Dengue studies? *Réseaux, communication et territoires*, 31-3/4, 283–308. doi: 10.4000/netcom.2725.
- Cebeillac, A., Daudé, É., Vaguet, A. (2018). Spatial discontinuities, health and mobility. *Revue Internationale de Géomatique*, 28(4), 389–407.
- Chang, S., Pierson, E., Koh, P.W., Gerardin, J., Redbird, B., Grusky, D., Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840), 82–87.
- Commenges, H. (2013). L'invention de la mobilité quotidienne. Aspects performatifs des instruments de la socio-économie des transports. PhD Thesis, Université Paris-Diderot-Paris VII, Paris [Online]. Available at: <https://tel.archives-ouvertes.fr/tel-00923682/>.
- Cummings, D.A.T., Irizarry, R.A., Huang, N.E., Endy, T.P., Nisalak, A., Ungchusak, K., Burke, D.S. (2004). Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature*, 427(6972), 344–347. doi: 10.1038/nature02225.
- Daudé, É. (2004). Apports de la simulation multi-agents à l'étude des processus de diffusion. *Cybergeog: European Journal of Geography*. doi: 10.4000/cybergeog.3835.
- Daudé, É. and Éliot, E. (2005). Exploration de l'effet des types de mobilités sur la diffusion des épidémies. *Actes du colloque des Septièmes Rencontres ThéoQuant* [Online]. Available at: <https://thema.univ-fcomte.fr/theoq/pdf/2005/TQ2005%20ARTICLE%2021.pdf>.
- Daudé, É., Vaguet, A., Paul, R. (2015). La dengue, maladie complexe. *Natures Sciences Sociétés*, 23(4), 331–342.
- Derouich, M., Twizell, E.H., Boutayeb, A. (2003). A model of dengue fever. *BioMedical Engineering* [Online]. Available at: <http://dspace.brunel.ac.uk/handle/2438/1692>.
- Eliot, E., Daudé, E., Bonnet, E. (2012). Interpréter les épidémies du passé : l'exemple de l'épidémie du choléra-morbus en Normandie en 1832. *Geoconfluences* [Online]. Available at: <https://halshs.archives-ouvertes.fr/halshs-01082682>.

- zu Erbach-Schoenberg, E., Alegana, V.A., Sorichetta, A., Linard, C., Lourenço, C., Ruktanonchai, N.W., Graue, B., Bird, T.J., Pezzulo, C., Wesolowski, A. (2016). Dynamic denominators: The impact of seasonally varying population numbers on disease incidence estimates. *Population Health Metrics*, 14(1) [Online]. Available at: <http://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-016-0106-0>.
- Eubank, S., Guclu, H., Anil Kumar, V.S., Marathe, M.V., Srinivasan, A., Toroczkai, Z., Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988), 180–184.
- Frias-Martinez, E., Williamson, G., Frias-Martinez, V. (2011). An agent-based model of epidemic spread using human mobility and social network information. *Proceedings of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, 57–64 [Online]. Available at: <http://ieeexplore.ieee.org/abstract/document/6113095/>.
- Galeazzi, A., Cinelli, M., Bonaccorsi, G., Pierri, F., Schmidt, A.L., Scala, A., Pammolli, F., Quattrocioni, W. (2021). Human mobility in response to COVID-19 in France, Italy and UK. *Scientific Reports*, 11(13141) [Online]. Available at: <https://www.nature.com/articles/s41598-021-92399-2>.
- González, M.C., Hidalgo, C.A., Barabási, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196), 779–782.
- Gould, P. (1993). *The Slow Plague: A Geography of the AIDS Pandemic*. Blackwell Publishers, Hoboken.
- Hägerstrand, T. (1970). What about people in Regional Science? *Papers of Regional Science Association*, 24, 7–21.
- Hawelka, B., Sitko, I., Beinart, E., Sobolevsky, S., Kazakopoulos, P., Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260–271.
- Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., Newth, D. (2015). Understanding human mobility from Twitter. *PLoS ONE*, 10(7), e0131469.
- Kermack, W.O. and McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 115(772), 700–721.

- Langlois, P. and Daudé, É. (2007). Concepts et modélisations de la diffusion géographique. *Cybergeo*, 364, 23 [Online]. Available at: <http://193.55.107.45/articles/364res.htm>.
- Lenormand, M., Picornell, M., Cant-Ros, O.G., Tugores, A., Louail, T., Herranz, R., Barthelemy, M., Frias-Martinez, E., Ramasco, J.J. (2014). Cross-checking different sources of mobility information. *PLoS ONE*, 9(8), e105184.
- Lourenço, J. and Recker, M. (2014). The 2012 Madeira dengue outbreak: Epidemiological determinants and future epidemic potential. *PLoS Neglected Tropical Diseases*, 8(8), e3083.
- Luo, F., Cao, G., Mulligan, K., Li, X. (2016). Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago. *Applied Geography*, 70, 11–25.
- Massaro, E., Kondor, D., Ratti, C. (2019). Assessing the interplay between human mobility and mosquito borne diseases in urban environments. *Scientific Reports*, 9(1) [Online]. Available at: <http://www.nature.com/articles/s41598-019-53127-z>.
- Misslin, R. and Daudé, É. (2017). An environmental suitability index based on the ecological constraints of *Aedes aegypti*, vector of dengue and Zika virus. *RIG, International Review of Geomatic*, 27(4), 481–501. doi: 10.3166/rig.2017.00044.
- de Montjoye, Y.-A., Hidalgo, C.A., Verleysen, M., Blondel, V.D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3 [Online]. Available at: <http://www.nature.com/articles/srep01376>.
- Morin, E. and Lemoigne, J.-L. (1999). *L'intelligence de la complexité*. L'Harmattan, Paris.
- Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G.S., Wallinga, J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, 5(3), e74.
- Ordóñez-Gonzalez, J.G., Mercado-Hernandez, R., Flores-Suarez, A.E., Fernández-Salas, I. (2001). The use of sticky ovitraps to estimate dispersal of *Aedes aegypti* in northeastern Mexico. *Journal of the American Mosquito Control Association*, 17(2), 93–97 [Online]. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11480827>.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3), 925–979.

- Perchoux, C., Chaix, B., Cummins, S., Kestens, Y. (2013). Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place*, 21, 86–93.
- Picheral, H. (1983). Complexes et systèmes pathogènes : approche géographique. De l'épidémiologie à la géographie humaine. *Travaux et documents de géographie tropicale*, 5–22.
- Ripley, K. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society*, 39(2), 172–212.
- Ross, R. (1911a). Some quantitative studies in epidemiology. *Nature*, 87, 466–467.
- Ross, R. (1911b). *The Prevention of Malaria*. John Murray, London.
- Salje, H., Lessler, J., Endy, T.P., Curriero, F.V., Gibbons, R.V., Nisalak, A., Nimmannitya, S., Kalayanarooj, S., Jarman, R.G., Cummings, D.A. et al. (2012). Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. *Proceedings of the National Academy of Sciences*, 109(24), 9535–9538.
- Salje, H., Lessler, J., Maljkovic Berry, I., Melendrez, M.C., Endy, T., Kalayanarooj, S., A-Nuegoonpipat, A., Chanama, S., Sangkijporn, S., Cummings, D.A. et al. (2017). Dengue diversity across spatial and temporal scales: Local structure and the effect of host population size. *Science*, 355(6331), 1302–1306.
- Sanders, L. (2006). Les modèles agents en géographie urbaine. In *Modélisation et simulation multi-agents : applications pour les Sciences de l'Homme et de la Société*, Amblard, F. and Phan, D. (eds). Hermes-Lavoisier, Paris.
- Sattenspiel, L. and Dietz, K. (1995). A structured epidemic model incorporating geographic mobility among regions. *Mathematical Biosciences*, 128(1/2), 71–91.
- Schaerström, A. (1996). Pathogenic paths? A time geographical approach in medical geography. PhD Thesis, Lund University, Lund.
- Shen, L. and Stopher, P.R. (2014). Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*, 34(3), 316–334.
- Shepherd, H.E.R., Atherden, F.S., Chan, H.M.T., Loveridge, A., Tatem, A.J. (2021). Domestic and international mobility trends in the United Kingdom during the COVID-19 pandemic: An analysis of Facebook data. *International Journal of Health Geographics*, 20, 46. doi: 10.1186/s12942-021-00299-5.

- Simini, F., González, M.C., Maritan, A., Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, 484(7392), 96–100.
- Snow, J. (1855). *On the Mode of Communication of Cholera*, 2nd edition. John Churchill, London [Online]. Available at: <http://www.ph.ucla.edu/epi/snow/snowbook.html>.
- Song, C., Koren, T., Wang, P., Barabási, A.-L. (2010a). Modelling the scaling properties of human mobility. *Nature Physics*, 6(10), 818–823.
- Song, C., Qu, Z., Blumm, N., Barabasi, A.-L. (2010b). Limits of predictability in human mobility. *Science*, 327(5968), 1018–1021.
- Sorre, M. (1933). Complexes pathogènes et géographie médicale. *Annales de Géographie*, 42(235), 1–18.
- Stoddard, S.T., Morrison, A.C., Vazquez-Prokopec, G.M., Paz Soldan, V., Kochel, T.J., Kitron, U., Elder, J.P., Scott, T.W. (2009). The role of human movement in the transmission of vector-borne pathogens. *PLoS Neglected Tropical Diseases*, 3(7), e481.
- Stoddard, S.T., Forshey, B.M., Morrison, A.C., Paz-Soldan, V.A., Vazquez-Prokopec, G.M., Astete, H., Reiner, R.C., Jr, Vilcarromero, S., Elder, J.P., Scott, T. W. et al. (2013). House-to-house human movement drives dengue virus transmission. *Proceedings of the National Academy of Sciences*, 110(3), 994–999.
- Tatem, A.J., Rogers, D.J., Hay, S.I. (2006). Global transport networks and infectious disease spread. *Advances in Parasitology*, 62, 293–343 [Online]. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0065308X0562009X>.
- Tizzoni, M., Bajardi, P., Decuyper, A., Kon Kam King, G., Schneider, C.M., Blondel, V., Smoreda, Z., González, M.C., Colizza, V. (2014). On the use of human mobility proxies for modeling epidemics. *PLoS Computational Biology*, 10(7), e1003716.
- Treuil, J.-P., Drogoul, A., Zucker, J.-D. (2008). *Modélisation et simulation à base d'agents : exemples commentés, outils informatiques et questions théoriques*. Dunod, Paris.
- Vallée, J. (2017). Challenges in targeting areas for public action. Target areas at the right place and at the right time. *Journal of Epidemiology and Community Health*, 71(10), 945–946.

- Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P., Ritchie, S.A. (2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. *PLoS Neglected Tropical Diseases*, 4(12), e920.
- Waite, H. (1910). Mosquitoes and malaria. A study of the relation between the number of mosquitoes in a locality and the malaria rate. *Biometrika*, 7(4), 421–436.
- Wesolowski, A., Eagle, N., Tatem, A.J., Smith, D.L., Noor, A.M., Snow, R.W., Buckee, C.O. (2012). Quantifying the impact of human mobility on malaria. *Science*, 338, 267–270.
- Wesolowski, A., Buckee, C.O., Bengtsson, L., Wetter, E., Lu, X., Tatem, A.J. (2014). Commentary: Containing the Ebola outbreak – The potential and challenge of mobile network data. *PLoS Currents Outbreaks*, 6.
- Wesolowski, A., Qureshi, T., Boni, M.F., Sundsøy, P.R., Johansson, M.A., Rasheed, S.B., Engø-Monsen, K., Buckee, C.O. (2015). Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proceedings of the National Academy of Sciences*, 112(38), 11887–11892.

