Hydrological, Anthropogenic and Ecological Processes in Cholera Dynamics

THÈSE N^{O} 5802 (2013)

PRÉSENTÉE LE 26 JUIN 2013

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE D'ÉCOHYDROLOGIE
PROGRAMME DOCTORAL EN ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Lorenzo RIGHETTO

acceptée sur proposition du jury:

Prof. M. Parlange, président du jury
Prof. A. Rinaldo, Dr E. Bertuzzo, directeurs de thèse
Prof. M. Blokesch, rapporteur
Prof. M. Gatto, rapporteur
Prof. A. Pugliese, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2013

Time has told me
You're a rare rare find
A troubled cure
For a troubled mind.
And time has told me
Not to ask for more
Someday our ocean
Will find its shore.
— **Nick Drake**

Acknowledgements

First, I would like to thank my supervisor Prof. Andrea Rinaldo for the continuous, remarkable interest and enthusiasm he has been lavishing in cholera research. Thanks also for the share of responsibility that was granted to me and that has taught me a great deal.
A special thanks goes to Dr. Enrico Bertuzzo, my co-supervisor. Thank you for your precious help and constant presence all along my research work. I learned a lot from you.
A special thanks goes also to Prof. Renato Casagrandi, Dr. Lorenzo Mari and Prof. Marino Gatto for their constructive suggestions during my research. Your presence has been a constant point of reference for me.
Thanks to Dr. Sirajul Islam, Dr. Zahid Mahmud, Mr. Rokon Uz Zaman and all the people in Bangladesh who have been working with me, either in person or from the distance. Your support has been vital in our field work endeavour.
I would like to special thank my colleagues and friends Samir Suweis, Serena Ceola, Francesco Carrara, Andrea Giometto, and Ludovico Nicotina with whom I spent a lot of time and nice moments during my PhD. Thanks also to my colleagues Bettina Schaefli, Cara Tobin, Pierre Queloz and Anna Rothenbühler. I would like to special thank Bernard Sperandio for his technical support.
A remarkable thanks goes to my family, my parents Roberto and Olivia, my brother Filippo and my sister Lucia, whose love has been with me throughout these years. Thanks also to my grandparents, who have been always proud of me.
Finally, a special thanks is for Francesca. We have never stopped getting closer, while growing up together.

Lausanne, 11 Avril 2013 L. R.

Abstract

The present Thesis deals with understanding, measuring and modelling epidemic cholera. The relevance of the endeavour stems from the fact that mathematical epidemiology, properly guided by model-guided field validation, is a reliable and powerful tool to monitor and predict ongoing epidemics in time for action, and to save lives through evaluation of the effectiveness of mitigation policies or the deployment of medical staff and supplies. In recent years, waterborne diseases – and cholera in particular – have consolidated their role as a major threat for developing countries, where sanitation conditions are poor and the vulnerability to extreme events is highest. Several important features controlling the dynamics of occurrence and spreading of the disease in a region are studied in the present Thesis, from both theoretical and experimental perspective. Understanding the basic processes that regulate cholera infection is key to build reliable prediction tools. In this Thesis several driving mechanisms of cholera are investigated, with the objective of connecting together hydro-climatology, ecology and epidemiology in a comprehensive framework. First, the role of water volume fluctuations is analyzed partly through a bifurcation study in a mostly theoretical assessment. Such work is then particularized in a field campaign carried out in rural Bangladesh, where hydro-climatological variables and *Vibrio cholerae* concentrations have been monitored for more than a year in one of the ponds constituting the local water reservoir. Concomitantly, a procedure for the detection of *Vibrio cholerae*, based on flow cytometry, is tested in the field. Further emphasis is also given to the role of human mobility in disseminating the disease among different communities. In particular, the contribution of human mobility in the dispersal of vibrios along the hydrologic network is specifically analyzed. All the knowledge collected in these studies is then used to add essential details to a modeling framework that is applied to the dramatic case of the Haiti epidemic. A spatially explicit model, taking into account both hydrological transport and human mobility, is developed to simulate the spreading of the disease since its onset. It is also shown that the resurgence of the disease, coinciding with the rainy season of June-July 2011, can only be reproduced if hydrological forcings are considered. The framework is tested by forcing it with synthetic rainfall scenarios and projecting epidemiological outputs. It is shown that the model can quantify correctly the number of cases in a given time span, even when calibrated with limited information. This result allows then to use it as a tool to assess a priori the effectiveness of intervention policies, such as vaccination

and sanitation. The effect of these two is tested both in the short- and in the long-term, with different results. Such endeavour represents the ultimate goal of the work presented in this Thesis – albeit further effort is needed to link together public health management and mathematical epidemiology in this field.

Keywords: Waterborne diseases, mathematical epidemiology, hydro-climatology, periodically-forced models, bifurcation analysis, ecology of *Vibrio cholerae*, flow cytometry, human mobility models, spatially explicit models, hydrological networks, Haiti cholera epidemic, epidemic control.

Sommario

Questa tesi riguarda la comprensione, la misura e la modellizzazione di epidemie di colera. La rilevanza di questo lavoro è originata dal fatto che l'epidemiologia matematica, se opportunamente guidata da una validazione di campo indirizzata dal modello, è uno strumento affidabile e potente per monitorare e predire epidemie in corso in tempo per agire, e per salvare vite attraverso la valutazione dell'efficacia di politiche di mitigazione o per lo schieramento di staff medico e risorse. In anni recenti, le malattie idrotrasmesse – e il colera in particolare – hanno consolidato il loro ruolo come minaccia importante per i paesi in via di sviluppo, dove le condizioni di sanitazione sono carenti e la vulnerabilità agli eventi estremi è massima. Diverse caratteristiche importanti, che controllano le dinamiche di occorrenza e diffusione della malattia in una regione sono studiate nella presente Tesi, dal punto di vista sia teorico che sperimentale. Comprendere i processi base che regolano l'infezione del colera è fondamentale per costruire strumenti affidabili di predizione. In questa Tesi diversi meccanismi forzanti vengono investigati, con l'obiettivo di connettere tra di loro idro-climatologia, ecologia ed epidemiologia in una cornice concettuale unitaria. In primo luogo, il ruolo delle fluttuazioni del volume d'acqua è analizzato, in parte attraverso uno studio di biforcazione in un lavoro soprattutto teorico. Tale analisi è poi particolareggiata in una campagna di raccolta dati portata avanti nel Bangladesh rurale, dove variabili idro-climatologiche e le concentrazioni di *Vibrio cholerae* sono state monitorate per più di un anno in una delle pozze che costituiscono la riserva d'acqua locale. In concomitanza di questo lavoro, è stata testata una procedura per l'individuazione di *Vibrio cholerae*, basata sulla citometria a flusso. Ulteriore enfasi viene data, inoltre, al ruolo della mobilità umana nel disseminare la malattia tra diverse comunità. In particolare, si analizza sèecificamente il contributo della mobilità umana nel disperdere i vibrioni lungo la rete idrologica. Tutto il sapere raccolto in questi studi è poi utilizzato per aggiungere dettagli essenziali a una struttura modellistica che viene applicata al caso drammatico dell'epidemia di Haiti. Un modello spazialmente esplicito, che tiene conto sia il trasporto idrologico che la mobilità umana, è sviluppato per simulare la diffusione della malattia fin dalla sua apparizione. Si mostra inoltre che la risorgenza della malattia, coincidente con la stagione piovosa di Giugno-Luglio 2011, può essere riprodotta solo se le forzanti idro-climatologiche vengono prese in considerazione. Il modell è testato forzandolo con scenari sintetici di pioggia e proiettando le uscite epidemiologiche. Si mostra che il modello riesce

a quantificare correttamente il numero di casi in un intervallo temporale definito, anche se calibrato con informazioni limitate. Questo risultato permette quindi il suo utilizzo come strumento per valutare a priori l'efficacia di politiche di intervento quali vaccinazione e sanitazione. L'effetto delle due è testato sia nel breve che nel lungo periodo, con diversi risultati. Ciò rappresenta l'obiettivo finale del lavoro presente in questa Tesi – sebbene sforzi ulteriori siano necessari per collegare la gestione della salute pubblica con l'epidemiologia matematica in questo campo.

Keywords: Malattie idrotrasmesse, epidemiologia matematica, idro-climatologia, modelli forzati periodicamente, analisi di biforcazione, ecologia del *Vibrio cholerae*, citometria a flusso, modelli di mobilità umana, modelli spazialmente espliciti, reti idrologiche, epidemia di colera in Haiti, controllo delle epidemie.

Contents

List of Figures

List of Tables

1 Introduction

The World Health Organization (WHO) indicates diarrheal diseases – most of which are waterborne – as the fifth cause of death globally, and the second among children under five years of age (WHO, 2011). Even though eradicated from the developed part of the world, cholera, a human disease of the intestinal duct, still represents a major threat at global scale. Cholera can cause severe dehydration and the death of the host if not properly treated via rehydration. As a waterborne disease – caused by the bacterium *Vibrio cholerae* – cholera can strike violently countries where hygienic conditions are poor and water and food sanitation infrastructures are lacking. In recent years, namely since the start of the new millennium, cholera outbreaks have struck the poorest and most unstable parts of the world, where migrations of large numbers of persons are frequent and the vulnerability to extreme events (i.e. hurricanes) and calamities is highest. As a result, the pattern of cholera insurgence has shifted from historically endemic regions (e.g. the Bengal Bay area) to previously untouched – but highly troubled – countries, in Africa (see the Kwazulu-Natal epidemic in South Africa in 2000-2002 and the outbreaks in Congo and Sudan; Bertuzzo et al., 2008; Nkoko et al., 2011), and Central and South America, where the recent Haitian epidemic is still causing victims, having passed a total of 7,000 since its start in October 2010 (PAHO, 2010). In Figure 1.1 the temporal evolution of the spatial distribution of cases at global scale is shown. Notably, during 2011 the Haitian epidemic alone has doubled the global cholera incidence. On the other side, even though endemic countries no longer contribute substantially to the annual figures of cholera cases, they still represent a crucial field for the understanding of the mechanics of insurgence of the disease, especially in relation to hydro-climatological drivers and to the ecology of the survival of *V. cholerae* in the aquatic environment (Lipp et al., 2002; Emch et al., 2010).

Being a waterborne and, more generally, an environmentally mediated human disease, cholera dynamics are controlled by human behavior and other societal factors (e.g. the water sanitation system) and by environmental conditions (see Figure 1.2

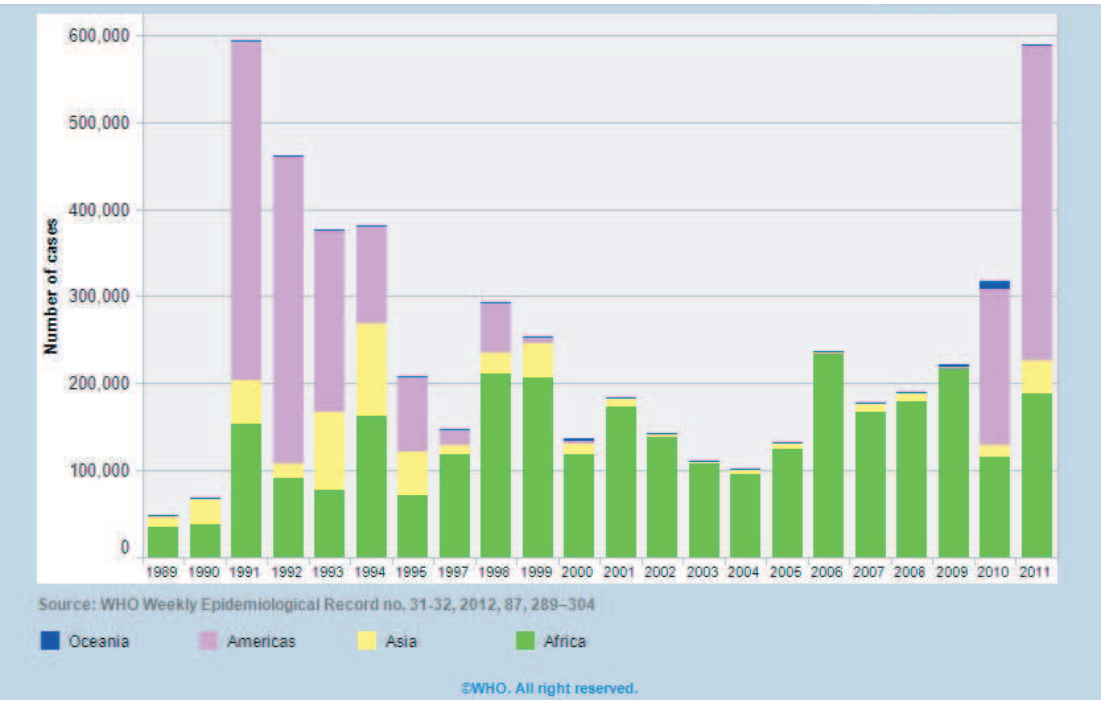

Figure 1.1: Distribution of reported cholera cases at continental scale from 1989 to 2011 (WHO, 2011).

for a summary of the principal dynamics). Societal factors indeed prevail when the epidemics first start in a region, causing abrupt outbreaks especially in populations which were never exposed to the disease. During these periods, the high prevalence of the disease (i.e. the fraction of infected individuals over the entire population at a given time) causes ingent quantities of excreted vibrios to contaminate the water reservoir, thus dominating the normal ecological dynamics of survival of *V. cholerae* in the environment (Faruque et al., 2005). Such dynamics become important if and when cholera is endemic in a region, that is when it has reached a stationary pattern of attack rate (i.e. the integral of reported cases in the span of a year). In such settings, a large share of the population has been already exposed to the disease and it is therefore immune. As a consequence, the disease incidence is low and, in turn, the excretion by infected individuals contributes only marginally to the overall abundance of *V. cholerae* populations in the environment. In this case environmental drivers can indeed play a role in regulating the seasonal re-insurgences of the disease (Lipp et al., 2002). Several environmental signatures have been linked to *V. cholerae* biological cycles, which in turn regulate cholera seasonality (Pascual et al., 2000; Lipp et al., 2002; Altizer et al., 2006; de Magny et al., 2008; Emch et al., 2008; Koelle, 2009). However, fewer studies have analyzed how particular meteorological conditions could influ-

ence the contraction of the disease, that is by worsening the sanitation conditions in a region, caused by floods and extreme precipitation events in general (Ruiz-Moreno et al., 2007; Akanda et al., 2009; Luque Fernandez et al., 2009; Hashizume et al., 2010).

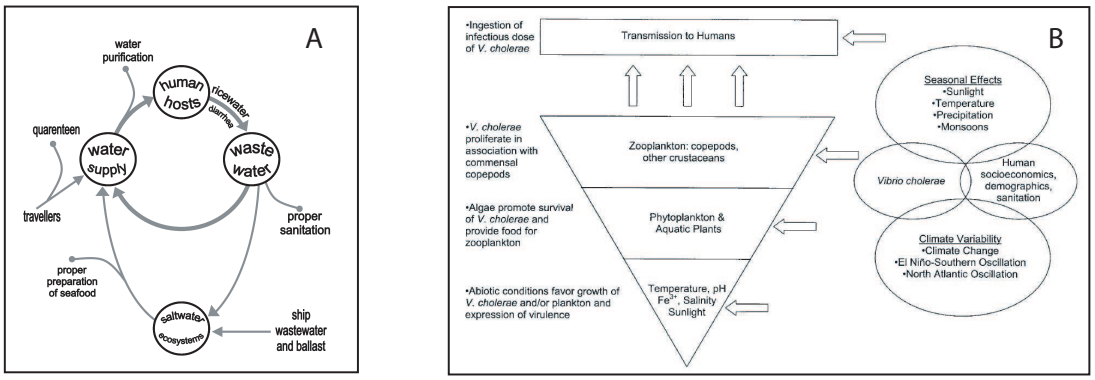

Figure 1.2: Panel A: Principal pathways leading to cholera infection in a community; panel B: overview of the factors involved in cholera transmission (Lipp et al., 2002).

A complete modeling framework has thus to include the mechanisms involved in both the onset and the spreading of the disease at all spatial scales. There is, hence, a need for simple yet rigorous, detailed yet synthetic models to describe the spatial and temporal evolution of the disease in a given geographical region. The research effort this thesis work has been part of deals with setting up a spatially explicit framework to describe the spreading of cholera over a network of human communities. Being a waterborne disease, cholera spreading is primarily mediated by the hydrologic network, where bacteria can be transported along the stream to reach other settlements. Human mobility does have a major role in the process, though, as a high fraction of people infected by cholera do not show any symptoms, and so they can move freely. The seminal work by Bertuzzo et al. (2008) first considered hydrologic transport of vibrios on a network of human communities to reproduce the spreading of the disease in space and time. This model was applied to the case study of the outbreak in the KwaZulu-Natal province of South Africa, showing a good characterization of the spatial pattern, besides the temporal one, which is the typical subject of the analyses of mathematical epidemiology. Later, Bertuzzo et al. (2010) have shown that, depending on the size of the system under study, spatially explicit schemes exhibit increasingly different patterns, as the timescale of disease propagation in space progressively overcomes the intrinsic dynamics of the disease at local scale.

In this context of high unpredictability, the planning of emergency interventions to mitigate outbreaks can highly profit from indications coming from mathematical models. These can help in understanding, predicting and controlling new outbreaks, provided that an active and reliable monitoring system is put into place. The application

of this framework to real case studies is of course mediated by specific corrections, connected, for instance, to the available data or to the hydroclimatology of the region.

The work presented in this Thesis is focused on the study of epidemic cholera, in particular through tools of mathematical epidemiology and its system theory declination. The introductory material is organized as follows. A thorough review of the literature and a brief walk-through of the steps taken by the research in this field since its inception are carried out in the introduction. Spatially explicit approaches are introduced later, with a view to the state of the art in the relevant mathematical epidemiology. Results of various applications, chiefly to the recent Haiti outbreak which proves of particular importance, are also presented, jointly with an extended discussion of the original contributions pertaining to the work of the present Thesis.

1.1 State of the art

1.1.1 Local, homogeneous epidemiological models

Mathematical epidemiology gained particular attention in the last century, as a way to quantify global patterns of infection in human populations, coupled with immunological and strictly biomedical development (Anderson and May, 1992). Hamer (1906) formulated the idea that the intensity of an epidemic depends on the rate of contact between susceptible and infected individuals, the so-called the "mass-action principle" - i.e. the spread of the infection depends on the number of susceptibles times the number of infectives in the system. Further on, Kermack and Kendrick (1927) theorized that the introduction of infected individuals would not give rise to an outbreak if the susceptible population were under a threshold number, or density. The analysis of the coupled dynamics of these two classes of individuals has been a key point of research in mathematical epidemiology, albeit mostly via probabilistic models. The rise of system theory led then to a progressively more detailed analysis of factors determining heterogeneity and periodicity in epidemiological patterns (Anderson and May, 1978; Becker, 1979; Anderson and May, 1979).

Most of the research was confined, though, to directly transmissable, air-borne diseases – such as measles and influenza – until the Mediterranean cholera pandemic in the early 70s generated the research by Capasso and Paveri-Fontana (1979), who formulated a dynamical system composed by two coupled equations, one for the infective population, one for the concentration of bacteria in the sea. This study enlightened the need for a description of the dynamics of the interaction between the human community and the environmental reservoir of bacteria and, as such, a waterborne infection of humans. In that study Capasso and Paveri-Fontana (1979) also postulated that the infection process would follow a nonlinear, saturating function. In

the following decades the disease remained mostly confined to the endemic areas of India and Bangladesh until early 90s, when outbreaks were reported first in Peru and then in other countries. The pandemic – the 7th in the history of cholera – persists nowadays.

Seminal work on the ecology of *V. cholerae* and of its association with more complex organisms, which lets bacteria survive for long periods in aquatic habitats, was published (Colwell et al., 1985; Islam et al., 1994). Little advance was made in the filed of mathematical epidemiology of cholera epidemics, until Codeço (2001) defined a model to study the conditions for endemicity in a system. Her model revolved around the standard Susceptible-Infected-Recovered compartments, with the addition of the dynamics of an aquatic population of *Vibrio cholerae*. It can be written as follows:

$$\begin{aligned} \frac{dS}{dt} &= \mu\,(H - S) - \beta \frac{B}{K + B}\, S \\ \frac{dI}{dt} &= \beta\, \frac{B}{K + B}\, S - (\mu + \alpha + \gamma)\, I \\ \frac{dR}{dt} &= \gamma\, I - \mu\, R \\ \frac{dB}{dt} &= \mu_B\, B + \frac{p}{W}\, I \end{aligned} \tag{1.1}$$

The first equation describes the dynamics of susceptibles in a community of size H. Susceptible individuals are born and die on average at rate μ. Newborn individuals are considered susceptible. Susceptible people become infected at a rate $\beta\, B/(K + B)$, where β is the rate of contact with contaminated water and $B/(K+B)$ is a logistic dose response curve that links the probability of becoming infected to the concentration of vibrios B in water. Infected people (whose dynamics is described by the second equation) die at a rate which is the sum of natural mortality μ and disease-caused mortality α, and recover with rate γ. The third equation describes the dynamics of the free-living infective propagules in the reservoir. Infected people contribute to the concentration of vibrios at a rate p/W, where p is the rate at which bacteria are produced by one infected person and W is the volume of the contaminated water body. The growth rate μ_B of the free-living bacteria in the water body is usually negative because bacteria mortality in natural environments exceeds reproduction. If μ_B were positive, the model would predict an exponential growth of the vibrios concentration and all the susceptible population would be affected by the disease. People recovered from cholera are considered permanently immunized and thus the Recovered equation is in fact uncoupled from the others.

Codeço (2001) showed that there exists a threshold value of system parameters under which no outbreak can emerge in the system. This model has, in fact, two equilibria, namely the disease-free equilibrium $\overline{X}_0 = [H\,0\,0\,0]^T$ (the symbol T stands for matrix transposition) and the non-trivial, endemic equilibrium $\overline{X}_+ = [\overline{S}_+\,\overline{I}_+\,\overline{R}_+\,\overline{B}_+]^T$, in which there is a convergence to a constant fraction of infectives in the total population. In particular, $\overline{X}_+$ is positive and the model converges to endemicity only when the basic reproductive number of cholera, R_0 (usually referred to as "the average number of secondary infections produced when one infected individual is introduced into a host population where everyone is susceptible and no control actions are in place", see Anderson and May, 1992), i.e.

$$R_0 = \frac{\beta p H}{K W (\mu + \alpha + \gamma)(\mu_B)}$$

is larger than unity. It can be shown that at equilibrium the prevalence $\overline{I}_+$ (the fraction of infected individuals) increases with R_0. When $R_0 < 1$, $\overline{X}_0$ is stable. Note that, at $R_0 = 1$, there is a transcritical bifurcation of the system: equilibria $\overline{X}_0$ and $\overline{X}_+$ coincide and exchange stability.

R_0 is directly proportional to the contact rate β and the excretion rate p, to the total size of the human community H and to the times of residence in the compartments of infectives $(\mu + \alpha + \gamma)^{-1}$ and bacteria $(\mu_B + \lambda)^{-1}$. On the other hand, it is inversely proportional to the critical dose K.

Later, several developments of the local model were devised in order to account for the hyper-infectivity of newly shed vibrios (Hartley et al., 2006) and for human-to-human contact (Tien and Earn, 2010).

1.1.2 Spatially explicit models for waterborne diseases

Standard, process-based epidemiological models lack, though, any description of the dynamics of transport and, consequently, of redistribution of pathogens between communities. These spatial dynamics become important, though, when the travel times of pathogens are comparable to the time scales of infection. In this case the spatial distribution of population and their possible heterogeneities, e.g. as regards the sanitation system, indeed play a major role in the spreading of the disease and cannot be neglected if one aims at mimicking correctly the propagation of cholera in a region.

In Bertuzzo et al. (2008) the transport of pathogens along the hydrological network were modeled with a biased random-walk process on an oriented graph (see also Johnson et al., 1995; Campos et al., 2006; Bertuzzo et al., 2007). Nodes represented suitable population centroids. The edges reproduced the connections between them,

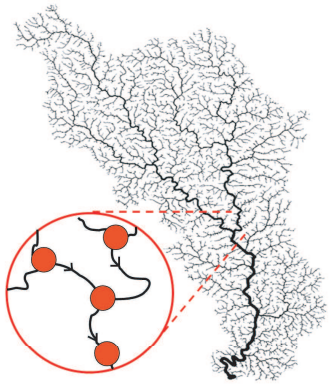

Figure 1.3: Schematic representation of the network model, where the nodes represent the human communities, symbolically located at the junction of river branches (the edges); arrows indicate that there is a downstream advection process ongoing in the transport along the network.

following the flow direction path imposed by the morphology of the region. An infectious propagule could move with some probability from a node to one of the adjacent nodes, which were all the nodes that are connected to it through an inward or outward edge. Bertuzzo et al. (2008) assigned to each edge of the graph an orientation according to the flow direction (Fig. 1.3). Once one applies the local epidemic model at each node of the network, one has $3N$ state variables S_i, I_i, B_i, where the subscript i identifies the nodes. Vibrios can be assumed to be removed at every node with a certain rate l (day^{-1}) and transported through the network following the transition probabilities P_{ij}. Then, the equations that describe the coupled process become:

$$\begin{aligned} \frac{dS_i}{dt} &= n\,(H_i - S_i) - \beta\,\frac{B_i}{K + B_i}\,S_i \\ \frac{dI_i}{dt} &= \beta\,\frac{B_i}{K + B_i}\,S_i - (r + m + n)\,I_i \\ \frac{dB_i}{dt} &= n_B\,B_i + \frac{p}{W_i}\,I_i - l\,B_i + \sum_{j=1}^{N} l\,P_{ji}\,B_j\,\frac{W_j}{W_i} \end{aligned} \tag{1.2}$$

for $i = 1, 2, \ldots, N$, where N is the total number of nodes. Note that all the parameters are node-independent except for the population size H_i and the water volume W_i. The latter represents the whole set of water supplies available for that community, not only the one provided by the river. The network acts as a link through which different sets of water supplies of different communities can be connected and contaminated. In order to further minimize the number of parameters, Bertuzzo et al. (2008) assumed that the water volume is a non-decreasing function of the population size: $W_i = f(H_i)$. Usually, $W_i = cH_i$. Different choices of the function f can lead to

different scenarios of the epidemic.

This framework allows to portray the spatial propagation of the disease, once an initial concentration of bacteria (or an initial number of infected individuals) is seeded in a fraction of the network. Note that, in this early case, all mechanisms of pathogen transport (especially the "active" ones, e.g. attachment to bigger organisms, human mobility) were summarized with a generic advection-diffusion scheme.

1.2 Thesis outline

The work described in this Thesis will deal with broadening and deepening the insight on the processes considered in the presented framework. In particular, it will focus on several open topics, from the explicit description of human mobility to the role of environmental (and hydro-climatological, in particular) drivers in determining epidemiological patterns in both endemic and epidemic contexts, with particular reference to the dramatic case of the Haiti epidemic. As part of a bigger European Research Council project, this Thesis will also present some field work which has been carried out in rural Bangladesh. This experimental endeavour will be devoted mostly to studies related to the ecology of *V. cholerae* and its implications on the modelling framework.

Given the state of the research on mathematical tools for the analysis of cholera epidemics at the time this Thesis work was initiated, several topics needed further study. Little emphasis was given, for instance, to the hydrological forcings that preside over the seasonality of the disease, which is controlled, intuitively, also by the fluctuations of the local water volume and, thus, by the periodical increases of bacterial concentration. Such concentration had been also observed to be controlled by ecological interactions, which are again studied here in a field setting, in order to correctly model how vibrios may survive and grow in the aquatic environment. Finally, a description of how human movement affects the spreading of the disease was lacking – or, at least, considered only surreptitiously – and here a first assessment of such process is presented. All these studies serve as a knowledge base for an application to the dramatic case of the Haiti epidemic, in which the developed framework has been tested as a monitoring and prediction tool. For this conclusive work, a more detailed description of human mobility and a dependance on hydrologic variables has been included.

In Chapter 2, fluctuations of water availability to the human community under study are explicitly accounted for. The seasonality of water input in the reservoir drives the variation of concentration of *V. cholerae*. Two compartments are added to the Susceptible-Infected-Bacteria model. First, the Recovered individuals, which, over many seasons, lose their immunity to the disease and replenish the Susceptible group. Second, the water volume of the reservoir, which determines bacterial dilution and,

consequently, the probability of contracting cholera by ingesting contaminated water. By forcing the model with a seasonally varying hydrologic input, simulations are compared to available data for various regions of the world characterized by different hydrological and epidemiological regimes. The model is shown to satisfactorily reproduce important characteristics of disease insurgence and long-term persistence. Using bifurcation analysis, different degrees of seasonality and values of the basic reproductive number are shown to influence the expected long-term epidemiological time series.

Moreover, the relationship between water volume fluctuations and cholera occurrence is tested by carrying out a field campaign in the area of Matlab in rural Bangladesh, where cholera is well-known to be endemic. In Chapter 3 the results of this campaign are shown. The water level in one of the ponds that constitute the common water reservoir for human settlements in that area, between May 2011 and December 2012, is monitored. Together with water volume, other relevant hydro-climatological variables were measured, such as rainfall and solar radiation. Concurrently, samples of the water were withdrawn biweekly from the pond and they were analyzed for phytoplankton and *V. cholerae* concentrations. At the time of sampling, the physicochemical characteristics of the pond water were also assessed. Vibrio populations have been reported to survive in the environment in conjunction with other microorganisms. The dynamics of the water volume in the reservoir are then reconstructed as a balance of input (from rainfall and drainage) and output (evaporation) terms. Finally, a process-based model of the ecology of the bacterium is proposed, in conjunction with the dynamics of phytoplankton and of the nutrient content of the water, which fluctuates according to the balance equation previously identified.

In Chapter 4, a method for the field validation of model parameters and, more generically, assumptions is tested. Presence of *Vibrio cholerae* serogroups O1 and O139 in the waters of the rural area of Matlab, Bangladesh was in fact investigated by quantitative measurements carried out using a portable flow cytometer. The relevance of this work relates to the testing of a field-adapted measurement protocol that might prove useful for cholera epidemic controls. Such nearly real-time measurement could be well suited to complement spatially explicit mathematical models thereby helping to predict cholera epidemics. Water samples were collected from different water bodies that constitute the hydrologic system of the region, a well-known endemic area for cholera. Water was retrieved from ponds, river waters, and irrigation canals during an inter-epidemic time period. Each sample was filtered and analyzed with a flow cytometer for a fast determination of *V. cholerae* cells contained in those environments. More specifically, samples were treated with O1- and O139-specific antibody, which allowed precise concentration measurements based on flow cytometry. The total living cell concentration was also measured, as well as environmental abiotic parameters, to test whether there is any correlation with the vibrio concentration.

In Chapter 5, an important extension of the spatially explicit and hydrology-driven, model of cholera by Bertuzzo et al. (2008) spreading is proposed. This reformulation concerns the role of different modelizations of human movement on the overall dynamics of cholera epidemics, which were previously expressed synthetically by bacterial diffusion. The underlying spatial framework is based on a reaction-diffusion scheme, which is applied, for simplicity, to ecological corridors modeled as one-dimensional lattices mimicking river reaches. In this geometry important features of disease propagation can be predicted, like the analytical evaluation of the spreading celerity of travelling fronts of epidemics. The explicit mechanism of human movement presented here leads to a comparison with previous approaches, revealing notable differences in how cholera spreads in space and time along the system. Studying human movement allows the analysis of specific quantities useful to policy makers and epidemiologists, such as the value of human diffusion which can generate upstream propagation of the disease. A robust relationship among key parameters, which links together diffusion mechanism, drift speed and resulting epidemic intensity, is also shown.

To summarize, several and diverse processes have been studied to widen the knowledge on the modeling of cholera spreading. In particular, the Thesis work has focused on the processes regulating the seasonality of the disease, including the effect of hydrological drivers and the role of ecological interactions of *V. cholerae* with other species. Moreover, further insight has been provided on the additional dissemination of the infection caused by human displacement along the spatial network. Such endeavour is capitalized in the last chapter where it is used for the most useful applications of this framework, that is for the simulation and the prediction of epidemic outbreaks of the disease.

In recent years, in fact, the attention of research on waterborne diseases has been gathered by the violent outbreak of cholera which has struck Haiti. The epidemic started in October 2010 and is still ongoing, having caused nearly $650,000$ reported cases and more than $8,000$ deaths until now. A joint collaboration of researchers from EPFL (including myself), Politecnico di Milano, Princeton University and Harvard University developed a spatially explicit framework to simulate and predict the course of the epidemic, starting from the first, abrupt prevalence peak that was observed in December 2010 to the new resurgence of the summer of 2011, which was triggered by extreme rainfall events (Figure 1.4).

Following the empirical evidence of a clear correlation between rainfall events and cholera resurgence that was observed in particular during the recent outbreak in Haiti, a spatially explicit model of epidemic cholera is re-examined in Chapter 6. Specifically, a multivariate Poisson rainfall generator, with parameters varying in space and time, is tested as a driver of enhanced disease transmission. These, in turn, are driven by rainfall washout of open-air defecation sites or cesspool overflows, hydrologic trans-

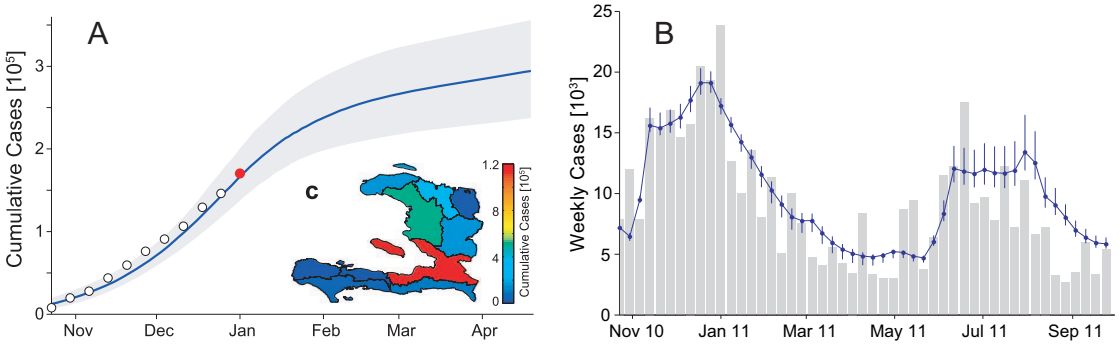

Figure 1.4: Panel A: Predicted (solid line) and reported (circles) cases at the scale of whole Haiti for the time span between October 2010 and April 2011. Data reported in red were not used for the calibration of the model. The grey shaded range shows the $5-95$th percentile range of the model prediction, while the inset depicts the spatial distribution of cumulative predicted cases up to May 1, 2011 (Bertuzzo et al., 2011); panel B: weekly reported (grey bars) and simulated (solid line) cases at the country scale between October 2010 and September 2011. Error bars highlight the range of uncertainty due to parameter estimation (Rinaldo et al., 2012).

port through waterways and by mobility of susceptible and infected individuals. An *a posteriori* analysis (from the beginning of the epidemic on October, 17, 2010 until December 2011) is performed to test the model reliability in predicting cholera cases and in testing control measures, involving vaccination and sanitation campaigns, for the ongoing epidemic. A multi-seasonal prediction of the course of the epidemic until December 2015 is run, then, to investigate conditions for further resurgences and endemicity of cholera in the region with a discussion on the policies which may bring to the eradication of the disease in Haiti.

2 The role of aquatic reservoir fluctuations in long-term cholera patterns

2.1 Introduction

In the context of cholera becoming an emerging threat for developing countries, mathematical modelling is becoming a subject of increasing epidemiological interest (Bertuzzo et al., 2008; Chao et al., 2011; Tuite et al., 2011; Mari et al., 2012a), both for acute epidemics, which tend to have self-sustained dynamics driven by exposure to excess concentrations of pathogens through dispersal by water pathways and human mobility, and for recurrent events in endemic regions, i.e. outbreaks due to hydrometeorological seasonality and to the ecology of the *Vibrio cholerae*. In particular, research is needed to better understand the mechanisms of its recurring insurgence in endemic regions, i.e. where outbreaks of the disease are cyclically observed. The role of many environmental drivers has been examined to date (see e.g. (Akanda et al., 2009; Colwell, 1996; Pascual et al., 2002)), but there is no clear indication of a robust, unambiguous correlation between one exogenous forcing factor and long-term cholera patterns. Viability of *Vibrio cholerae* in riverine and marine environments is obviously a key factor, which depends on many hydroclimatological variables, such as water temperature and precipitation or flooding indices as symptoms of increased exposures (Pascual et al., 2002; Bouma and Pascual, 2001; Altizer et al., 2006; Ruiz-Moreno et al., 2007). On the other hand, bacterial concentration in the local water reservoir, i.e. in the water volume that is commonly utilized by the local community for its daily needs, is directly related to the hydrological cycle. The dilution effect is the expected mechanism underlying the link between the seasonal fluctuations of water reservoir volumes and the insurgence of the disease – the intensity of droughts has, in fact, been found to be positively correlated with the intensity of the disease (Akanda et al., 2009). The aim is to systematically analyze via a newly proposed model the influence of different hydrological regimes on cholera long-term temporal patterns to understand whether this factor has indeed a key role in determining the observed periodicities and bursts of the disease in various geographic regions.

To focus on temporal fluctuations, a spatially implicit description is used, in order to discuss the different epidemiological patterns that can affect a single human community as a result of time-varying water availability. It is thus assumed that the time scale of spatial spread of the disease in the whole region considered (through dispersion in waterways and human mobility) is much shorter than the time scale of the local epidemics (see Bertuzzo et al. (2010) for details). Under this circumstance, spatially explicit and spatially implicit models basically show a very similar behavior, as individuals move rapidly inside the region without introducing large delays at the epidemic time scale. Spatially detailed models relating transport of bacteria along river networks and cholera insurgence were previously addressed in the literature (Bertuzzo et al., 2008, 2010; Righetto et al., 2011; Chao et al., 2011; Tuite et al., 2011), but the seasonal variations of water volumes in the river network were not considered therein. This is a factor that needs to be addressed, because there is ample evidence that the variable conditions of water supply are tightly linked to cholera temporal patterns (Akanda et al., 2009).

The core of this model is based on Codeco's Susceptible-Infected-Bacteria (SIB) compartmental model (Codeço, 2001), as described in the Introduction, suitably modified to explicitly incorporate temporary immunity of hosts and the water reservoir dynamics. In the original formulation of Codeço (2001) the host population was subdivided into susceptible and infected compartments alone. The disease was indirectly transmitted via ingestion of water contaminated by *Vibrio cholerae*. In order to study long-term (i.e. several years) cholera patterns, the process of immunity loss by individuals who recovered from the disease must be accounted for. Therefore, a third compartment is added in the population. Also, as first proposed by Pascual et al. (2002), here the variation over time of the volume of water available to the community is considered, expressed as a mass balance equation. The amount of water available in the system influences in a nonlinear way not only the concentration of bacteria, but also the probability of contracting the disease.

The specific aim is to investigate the role played by hydrological seasonality in the processes leading to cholera insurgence. The focus is not restricted only to areas of historically acknowledged endemicity of the disease (Pascual et al., 2002, 2000), but the analysis is extended to countries with sporadic cholera outbursts to understand whether hydrological seasonality alone can reproduce – at least qualitatively – some of the global patterns that have been identified at different latitudes (Lipp et al., 2002; Emch et al., 2008).

To investigate which causal factors are key for the insurgence of various epidemiological patterns, bifurcation analysis of nonlinear systems (Kuznetsov, 1995) is used. To this end, in fact, it is the most effective tool, because it permits to systematically classify the different possible behaviors of the model and to link them to particular model parametrisations. Since the majority of model parameters can be given reliable esti-

mates from the literature (Codeço, 2001; Hartley et al., 2006; Neilan et al., 2010) and do not display very pronounced geographical variations, it is possible to concentrate here on a few, significant quantities that are either characterized by uncertainty or do vary appreciably in different regions of the world. In particular, the basic reproductive number R_0, the variability of the hydrological cycle (also called degree of seasonality) and the water reservoir residence time have been singled out. Lastly, it is tested how different dynamics of loss of immunity, whose duration has also been questioned (see King et al. (2008)), may affect the behavior of the model.

2.2 The Model

To analyze the role played by the dynamics of the water reservoir in determining the epidemiology of cholera, the basic model by Codeço (2001) is reformulated as follows:

$$\begin{aligned}
\frac{dS}{dt} &= \mu\,(H-S) - \beta\frac{B}{kW+B}\,S + \rho\,R \\
\frac{dI}{dt} &= \beta\,\frac{B}{kW+B}\,S - (\mu+\alpha+\gamma)\,I \\
\frac{dR}{dt} &= \gamma\,I - (\rho+\mu)R \\
\frac{dB}{dt} &= -\mu_B\,B + p\,I - \lambda\,B \\
\frac{dW}{dt} &= q(t) - \lambda\,W
\end{aligned} \tag{2.1}$$

where: S, I and R are respectively the number of susceptible, infected and recovered individuals in the host community of size H; B is the total number of *Vibrio cholerae* bacteria, contained in the water reservoir of size W (thus bacterial concentration is B/W). The human population is assumed to be at demographic equilibrium in the absence of cholera, with μ being the natural background mortality rate and H the demographic equilibrium. Infection is regulated by the contact rate β and depends on the number of bacteria B through the logistic dose-response curve introduced by Codeço (2001), i.e. $B/(kW+B)$; the parameter k represents the concentration of bacteria that grants 50% probability for a susceptible of contracting the disease. Once infected, individuals can die either from natural causes (μ) or from cholera infection (α), or they can recover from the disease at a rate γ. As for *Vibrio cholerae* dynamics, it is assumed that bacteria would not persist in the natural environment, at least in appreciable quantities (μ_B being their mortality rate), in absence of contamination due to faecal excretion by infected individuals (occurring at rate p). Note that I includes asymptomatic individuals as well, whose excretion represents an important input to

Symbol	Description
β	rate of exposure to contaminated water (day^{-1})
μ	population natality and mortality rate (day^{-1})
p	per capita contamination rate ($cells\, day^{-1}\, infected^{-1}$)
H	human population size at demographic equilibrium
γ	recovery rate (day^{-1})
k	concentration of *V. cholerae* in water that yields 50% chance of being infected with Cholera ($cells/m^3$)
α	cholera mortality rate ($days^{-1}$)
μ_B	natural mortality rate of *V.cholerae* in the aquatic environment (day^{-1})
λ	rate of drainage from the water reservoir (day^{-1})
ρ	loss of immunity rate (day^{-1})

Table 2.1: Description of the parameters used in the model

V.cholerae concentration in the water (King et al., 2008; Kaper et al., 1995).

As anticipated in the Introduction, two model compartments have been added to the Codeco model Codeço (2001). The first is the Recovered class R. Infection from *Vibrio cholerae* does not confer total immunity from cholera. Therefore recovered individuals have an active role in the long-term epidemiology of the disease, because they replenish the susceptible compartment at a rate ρ, which is compatible with the decadal temporal scales of interest. Most studies indicate in fact a long-lasting immunity of about 5 years (Clemens, 1990; Koelle et al., 2005), even though a shorter immunity duration of 12 weeks has been advocated for in other studies (see e.g. King et al. (2008)). Second, as suggested in Pascual et al. (2002), a description of the hydrologic dynamics of the water reservoir volume W is explicitly incorporated. The water reservoir receives an input $q(t)$ and it is depleted at a rate λ, which is here assumed to be mainly due to drainage (i.e. evaporation is considered small with compare to drainage). The outflux of water from the water reservoir (λW) determines also an output of *Vibrio cholerae* proportional to the actual concentration B/W. The linear outflux model assumption is widely used in hydrology Brutsaert (2005); Rodriguez-Iturbe and Rinaldo (1997) either in conceptual models of the instantaneous unit hydrograph or in geomorphologic schemes of the hydrologic response (e.g. Rinaldo and Iturbe (1996)). The assumption yields stationary residence time distributions that allow to tackle via convolutions complex input-output sequences and are seen as a first-order approximation to a wide class of hydrological settings (e.g. Rinaldo et al. (2011)). The characteristic size of the water reservoir is defined as $\overline{W} = q_0/\lambda$, where q_0 is the time average

of $q(t)$ (i.e. $q_0 = 1/T \int_{t_0}^{t_0+T} q(t)\,dt$, for $T \gg 1/\lambda$). The interactions among the model compartments are illustrated in Figure 2.1, while a summary of the symbols used for parameters is in Table 2.1. To mathematically simplify the analysis, the model is reformulated in terms of dimensionless variables, i.e. $S \triangleq \frac{S}{H}$, $I \triangleq \frac{I}{H}$, $R \triangleq \frac{R}{H}$, $W \triangleq \frac{W}{\overline{W}}$ and $B \triangleq \frac{B}{k\overline{W}}$.

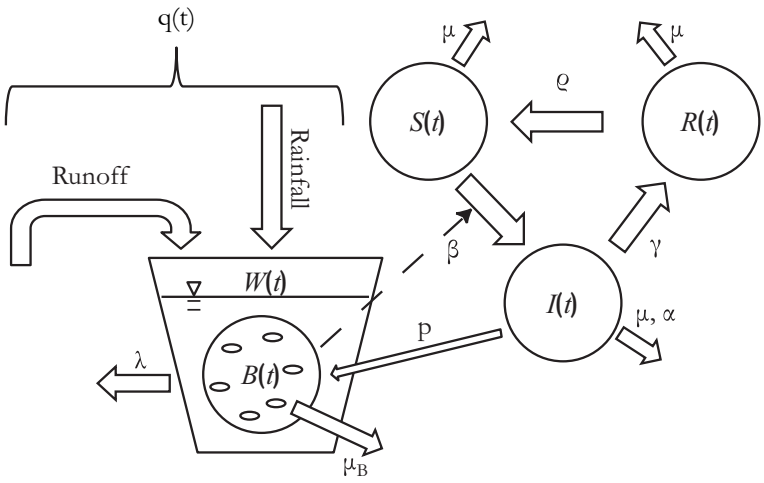

Figure 2.1: Schematic graph with the interactions among compartments of model (2) and the basic processes involved in the dynamics of the system.

The rescaled model therefore reads as:

$$\begin{aligned}
\frac{dS}{dt} &= \mu\,(1-S) - \beta\frac{B}{W+B}\,S + \rho\,R \\
\frac{dI}{dt} &= \beta\,\frac{B}{W+B}\,S - (\mu+\alpha+\gamma)\,I \\
\frac{dR}{dt} &= \gamma\,I - (\rho+\mu)R \\
\frac{dB}{dt} &= -\mu_B\,B + \frac{pH}{k\overline{W}}\,I - \lambda\,B \\
\frac{dW}{dt} &= Q(t) - \lambda\,W
\end{aligned} \tag{2.2}$$

where $Q(t) = q(t)/\overline{W}$. Consistently with Codeço (2001), model (2) – when $Q(t)$ is kept constant – has two epidemiological equilibria, namely the disease-free equilibrium $\overline{X}_0 = [1\,0\,0\,0\,1]^T$ (the symbol T stands for matrix transposition) and the non-trivial, endemic equilibrium $\overline{X}_+ = [S_+\,I_+\,R_+\,B_+\,1]^T$. In particular, X_+ is positive and the model converges to endemicity only when the basic reproductive number of cholera,

R_0 (Anderson and May, 1992), i.e.

$$R_0 = \frac{\beta p H / k\overline{W}}{(\mu+\alpha+\gamma)(\mu_B+\lambda)}$$

is larger than unity. It can be shown that at equilibrium the prevalence I_+ (the fraction of infected individuals) increases with R_0.When $R_0 < 1$, $\overline{X}_0$ is stable. Note that, at $R_0 = 1$, there is a so-called transcritical bifurcation of the system: equilibria $\overline{X}_0$ and X_+ coincide and exchange stability (Kuznetsov, 1995).

R_0 is directly proportional to the contact rate β and the excretion rate p, to the total size of the human community H and to the times of residence in the compartments of infectives $(\mu+\alpha+\gamma)^{-1}$ and bacteria $(\mu_B+\lambda)^{-1}$. On the other hand, it is inversely proportional to the critical dose k and to the characteristic water volume $\overline{W}$ (the dilution effect).

Model (2) has 8 parameters, each with a specific epidemiological or hydrological meaning. Ranges for the numerical values of the majority of these parameters can be found in the literature (see (Bertuzzo et al., 2008; Codeço, 2001) for references). As for the average transition time from recovered to susceptible state, it is assumed here that immunity is lost after 5 years (Koelle et al., 2005), so $\rho = 0.00055$ days^{-1}. This assumption will be relaxed in Section 4. Hydrologic parameters will be discussed in the next section.

2.3 Epidemiological patterns originated by the model

As outlined in the Introduction, the goal of the work presented in this Chapter is to investigate how epidemiological dynamics are connected to the hydrological regime. To this end an idealized seasonal pattern is introduced for the water input $Q(t)$, which qualitatively mimics the seasonality of river flow and/or rainfall. Such fluctuations represent the typical hydroclimatic conditions of many regions in the world.. In formulas, it is set:

$$Q = Q_0 \left[1 + \epsilon \cos\left(\frac{2\pi t}{365}\right)\right] \tag{2.3}$$

where Q_0 is the annual average normalized flow and t is measured in days. The parameter ϵ is a measure of the flow variability – the larger is ϵ, the larger is the ratio of the variance to the mean. This simplified formalism corresponds to a unimodal hydrological pattern, characterized by one annual peak occurring in the appropriate

season depending on the region. It is worth noting that not all regions of the world have a unimodal precipitation pattern. For instance some regions display bimodal patterns (Ruiz-Moreno et al., 2007).

A first question concerns whether the processes described by the seasonally forced model can reasonably reproduce, at least qualitatively, epidemiological behaviors observed in data. Unfortunately, available studies that point at the role of seasonality in cholera transmission (e.g. Pascual et al., 2002; Lipp et al., 2002; Emch et al., 2008) do not discuss in quantitative detail the interplay between hydrological drivers and cholera insurgence. In a recent and exhaustive review on seasonality of cholera, Emch et al. (2008) show data for the disease in areas of the world at different latitudes. Among many other examples of the same kind, cholera in Iran shows a pattern of insurgence that clearly suggests periodicity, with yearly outbreaks concentrating in October. Rainfall data from the Iran Meteorological Service are available online (Iran Meterological Service, 2011). By aggregating throughout the whole region data from rain gauges for the months between 1973 and 2004, a clearly unimodal pattern is obtained (Fig.2.2A, blue bars), with a wet season (late winter/early spring), followed by a dry summer. The red bars in the same Figure $2.2A$ show the sum of cholera outbreaks occurred in Iran month by month between 1974 and 2005, as reported in Emch et al. (2008). Note that an outbreak is defined therein as the presence of at least one reported cholera case in any calendar month, which cannot be referred to as an actual peak of prevalence, but simply testifies the occurrence of cholera in the region during the specific month. Although such an aggregated indicator must not be confused with disease prevalence, it is plausible to consider the number of outbreaks recorded in a month during the study period as a proxy for the infected in that month.

In panel B of Figure 2.2 the temporal pattern (red solid line) of prevalence, predicted after transient by model (2) when forced with a fluctuating water input $Q(t)$, as the light blue solid line, is shown. It is assumed here $\epsilon = 0.7143$, which corresponds approximately to the percent variation from the mean that is observed in rainfall data shown in panel A. For values of R_0 slightly larger than unity (assuming in this case $R_0 = 1.2$), the model exhibits a periodic behavior, with one yearly peak of cholera prevalence. The maximum number of outbreaks in Iran occurs three or four months after the lowest rainfall, which corresponds to the minimum bacterial dilution (Fig. 2.2A). Interestingly, the model reproduces correctly a similar time lag between the two patterns (Fig.2.2B). Delays of the same kind have been observed also in other regions, such as Central Amazonia (Pascual et al., 2002; Codeço, 2001). Such patterns are obtained for a reasonable value of the retention time of the water reservoir, equal to 50 days. This parameter is prominent in determining how much the minimum water level should lag with respect to the minimum water input, because it sets the response time of water reservoir volume with respect to the input. The delay between the minimum water volume and the epidemiological peak is then set by the basic reproductive number R_0, as it determines the threshold on susceptible population for

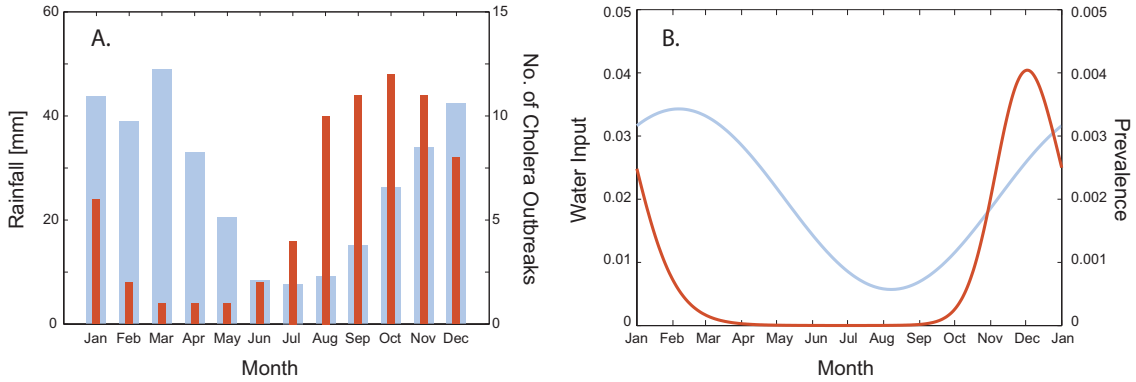

Figure 2.2: A) Data for the annual rainfall regime in Iran (light blue, left axis), obtained aggregating data from the Iran Meterological Service (2011) of Iranian stations (monthly averages from 1974 to 2003 and the number of cholera outbreaks in Iran from 1974 to 2005 (red; as in Emch et al. (2008)); B) A simulation with model (2) under seasonal forcing (see eq. (3)) and $R_0 = 1.2$, $\epsilon = 0.7143$, $Q_0 = 0.02\,days^{-1}$, $\lambda = 0.02\,days^{-1}$; $\rho = 0.00055\,day^{-1}$; other parameters as in Bertuzzo et al. (2008): $\mu = 0.00005\,days^{-1}$, $\beta = 1\,days^{-1}$, $\gamma = 0.2\,days^{-1}$, $\alpha = 0\,days^{-1}$, $\mu_B = 0.228\,days^{-1}$.

the epidemic to occur.

Other hydrological regimes, not displaying a single rain period per year, are worth being considered, to investigate whether the resulting epidemiological patterns can be mimicked by this model. Some studies (Pascual et al., 2002; Altizer et al., 2006; Ruiz-Moreno et al., 2007) point out that bimodal rainfall patterns throughout the year can cause dramatic changes in the epidemiological patterns. In the 'wet' regions of the Indian districts of Madras Presidency (the Administrative subdivision for British colonies), where two rain seasons are present (one around May, the other in October), cholera appears quite regularly every year. On the contrary, in Madras regions, where one rainfall peak per year is registered, disease outbreaks are recurrent but with irregular frequencies and amplitude.

The model is run by taking as water input the actual monthly averages of rainfall shown in Ruiz-Moreno et al. (2007) for the wet province of Trichinopoly (India), normalizing the quantities shown there so that:

$$Q(t) = Q_0 \frac{r_{Tr}(t)}{\max r_{Tr}(t)} \tag{2.4}$$

where $r_{Tr}(t)$ is the monthly averaged rainfall pattern as reported in Ruiz-Moreno et al. (2007).

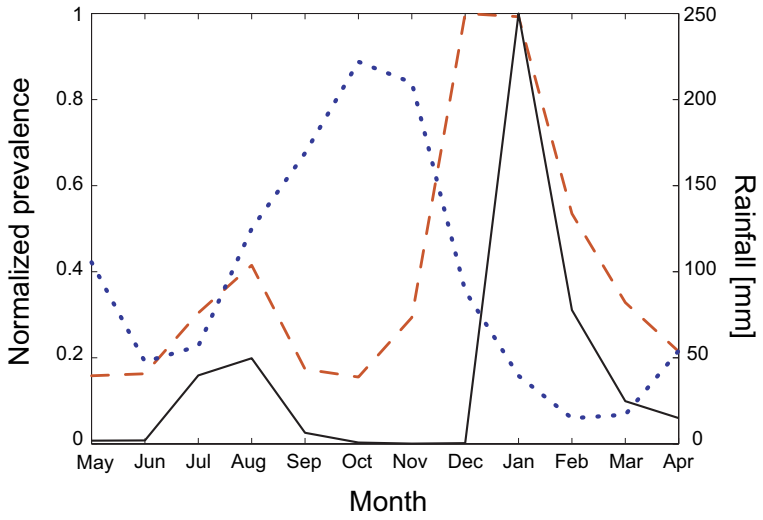

Figure 2.3: Comparison between data on monthly averages for cholera insurgence (yellow dashed line) and rainfall (blue dotted line) for the province of Trichinopoly, as shown in Ruiz-Moreno et al. (2007), and a model simulation (black solid line) with $R_0 = 7.5$, $\lambda = 0.2\,day^{-1}$; $\rho = 0.00055\,day^{-1}$ and water input given by equation (4); other parameters as in Bertuzzo et al. (2008): $\mu = 0.00005\,days^{-1}$, $\beta = 1\,days^{-1}$, $\gamma = 0.2\,days^{-1}$, $\alpha = 0\,days^{-1}$, $\mu_B = 0.228\,days^{-1}$.

Figure 2.3 shows $r_{Tr}(t)$ (blue dotted line) and cholera normalized prevalence $I_{Tr}(t)$ (yellow dashed line), also derived from data on cholera mortality shown in Ruiz-Moreno et al. (2007) as a proxy. Then the latter is compared to the long-term simulation (black solid curve) of the model (2), when forced assuming that the water input pattern follows each year the rainfall pattern depicted in Figure 2.3. Daily prevalences obtained from the model have been aggregated to monthly values and normalized. Note that, if the rainfall input repeated every year in the same way as in this case, the simulated asymptotic pattern would be periodic. This allows us to ascribe the dynamics shown in Figure 2.3 to the intrinsic dynamics of the model, which proves quite effective in grasping the overall mechanics of disease insurgence. In fact the winter peak occurs in January, when rainfall is decreasing, but before the rainfall minimum. The epidemiological peak drains the Susceptible basin; so, when conditions for disease outbreak (i.e., high vibrio concentration) occur again, near the second rainfall minimum (in June), the subsequent spring peak of cholera prevalence is approximately 60% lower. This result highlights the non-triviality of the interplay between human compartments, concentration of *Vibrio cholerae* and the fluctuations of the water reservoir, as no unique relationship between the water input timing and the timing of the epidemic can be identified.

Data on long-term dynamics of cholera insurgence show high variations in the magnitude of prevalence peaks (Ruiz-Moreno et al., 2007; King et al., 2008; Koelle et al.,

2005; de Magny et al., 2008; Glass et al., 1982; Huq et al., 2005; Islam et al., 2009; Pascual et al., 2008). Some studies have explicitly highlighted the large unpredictability of prevalence peaks (Pascual et al., 2002; Altizer et al., 2006). This might lead to the conclusion that a chaotic behavior can be characteristic of long-term cholera patterns.

2.4 Bifurcation analysis of the model

To understand the full complexity of seasonal epidemiological patterns, it is helpful to conduct a systematic analysis of the qualitative long-term behaviors of the model. The most effective way of cataloguing the functioning modes of nonlinear systems such as (2) is bifurcation analysis (Kuznetsov, 1995). By using software that implements continuation techniques (e.g. AUTO or MATCONT; Dhooge et al., 2003), it is possible to compute in relevant parameter spaces the bifurcation surfaces partitioning parametric regions where the model phase portraits are topologically equivalent. For example, parametric conditions for which the model exhibits one attracting limit cycle of period 1 year are separated by a surface from those where one single cycle of period 2 or a chaotic attractor emerges.

This kind of analysis is especially suited to the case at hand, because, as explained in the Introduction, uncertainties and variations are mainly limited to three parameters, two characterizing the hydrological regime (λ, ϵ) and one the epidemiological process (R_0). In particular, the amplitude of the forcing or degree of seasonality ϵ is representative of the climatic regime of a region and can widely vary throughout the world, while the basic reproductive number of the disease R_0 is very seldom estimated in the cholera literature. The retention time of the water reservoir is also an important uncertainty factor when describing the hydrological dynamics and thus governs the interactions between environment and vibrio concentration: for this reason the robustness of the analysis with respect to this parameter will be tested. Figure 2.4 shows the bifurcation diagram of the model in the parameter space (ϵ, R_0), obtained for a reservoir characterized by a retention time of $1/\lambda = 100$ days, together with typical temporal patterns of the disease obtained by the model. Each panel shows an example of the prevalence fluctuations corresponding to a different parametric region. To the far right the characteristic power spectra of the chaotic time series is also shown, depicted aside.

Here the structure of the bifurcation diagram shown in Figure 2.4 is detailed. At small ϵ's (region A), the system exhibits a stable one-year cycle. By varying the parameter values, such periodic attractor loses stability in favor of other attractors that undergo a series of bifurcations all depicted as red curves in Fig.2.4. The dashed red curve $f_1^{(1)}$ is a flip bifurcation which is subcritical within the interval (ZLV_1, ZLV_2) and supercritical elsewhere. Crossing the supercritical part of $f_1^{(1)}$ from left to right (regions G and I) causes the one-year cycle attractor to lose its stability in favor of a cycle with a

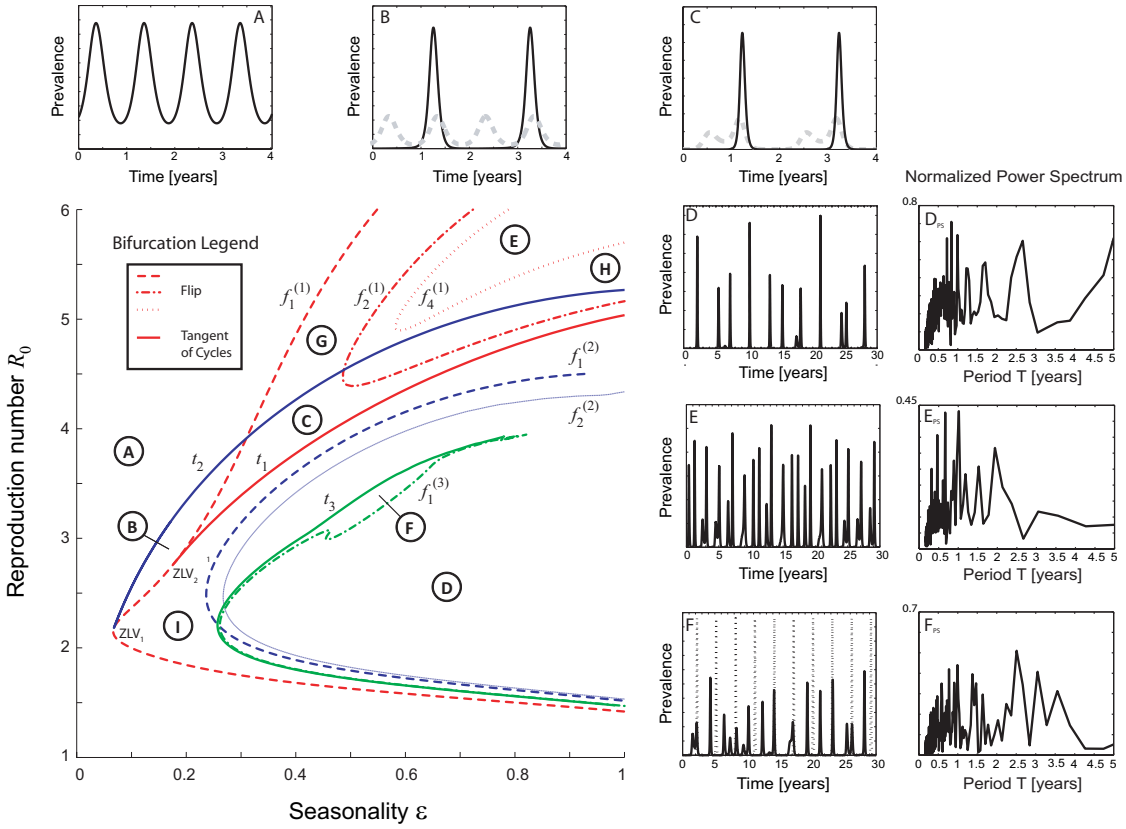

Figure 2.4: Bifurcation diagram of model (2) in the parameter space (ϵ, R_0 for $\lambda = Q_0 = 0.01\, days^{-1}$. Line types and symbols explained in the Supplementary Information. Unspecified parameters as in Fig. 2.2.

2-year period. From an epidemiological viewpoint, it is important to emphasize that the period doubling does not reflect an actual temporal displacement of prevalence peaks but a differentiation in their amplitude. In other words, close to the right of the supercritical part of $f_1^{(1)}$, there is a periodic attractor whose prevalence peaks every year, but with two slightly different peaks in odd and even years. The red dashed and dotted line $f_2^{(1)}$ is a supercritical flip bifurcation of the 2-year cycle just described. Inside the convex parametric region H formed by the curve (and close to it) a 4-year period cycle is stable. This 4-year cycle also loses its stability via period doubling at the red dotted flip bifurcation curve $f_4^{(1)}$. A cascade of flip bifurcations (Feigenbaum cascade) occurs nearby $f_4^{(1)}$ and a parametric region where model (2) shows chaotic behavior is entered (region E). The corresponding temporal patterns shown in panel 4E is indeed very resemblant to actual long-term time-series (as extensively discussed above). Whilst the amplitude and the timing of each peak is unpredictable for the chaotic attractors in this parametric region, the average return time of two prevalence peaks (i.e., of cholera outbreaks) remains rather close to approximately one year, yet

decreases with increasing R_0.

The solid red curve t_1 is a tangent of cycles bifurcation that involves the 2-year periodic attractor originated at the flip bifurcation $f_1^{(1)}$ and another unstable period 2 cycle. At t_1, these two cycles collide and disappear for parametric values at the right of it (region I).

There is a second important structure, depicted with blue curves, in the bifurcation diagram of Fig. 2.4. The blue solid line t_2 is a tangent bifurcation of cycles that identifies the formation of an attractor of period 2 years *sensu stricto*. In other words,this cycle has an epidemiological periodicity of 2 years, i.e. it has prevalence peaks that only appear every 2 years (solid time-series in panel 4B). The parametric region between the solid blue and red curves is therefore characterized by bistability, namely the system can alternatively converge to one cyclic solution of period 2 or another of either period 1 (region B) or 2 (with an interpeak period of 1 year though, region C). The period 2 attractor emerged at t_2 also undergoes a cascade of flip bifurcations. The blue dashed curve $f_1^{(2)}$ and the dashed and dotted $f_2^{(2)}$ curve are the first two flip bifurcations of a sequence leading to a region (labeled as D) where the system displays chaotic behavior. As noticed for the strange attractors of region E, chaotic regimes obtained in region D also maintain to some extent the characteristic periodicity of the original cycle that underwent the Feigenbaum cascade. The mean return time of aperiodic patterns in region D ranges from values around 2.36 years for $R_0 = 3$, to values of nearly 1.64 years for $R_0 = 4$.

Within the chaotic region D, other interesting structures emerge, although it is difficult to continue the bifurcations, for numerical reasons. As an example, a tangent bifurcation of period three cycles (solid green curve, t_3) and a supercritical flip bifurcation ($f_1^{(3)}$) are shown, where the period three cyclic attractor originated at t_3 doubles its period. In region F therefore bistability is found, which implies the coexistence of a chaotic attractor and a stable period 3 cycle (see panel F).The presence of regions in which there is bistability, namely the coexistence of periodic attractors of different characteristic period, is particularly underlined. Three main stable limit cycles, in fact, can be found indifferent regions of the parameter space under study, with a characteristic basic periodicity of 1, 2 and 3 years. The formation of the 2(3)-year cycle via a tangent of cycles bifurcation is marked by the blue(green) solid line, while the red solid line locates the disappearance of the 1-year cycle. In regions B and C, thus, bistability between the 1-year period – which, though, acquires biannual periodicity in region C – and the 2-year period attractors is encountered, whereas region D and F show coexistence between the 2-year period and the 3-year period attractors. This is a feature that holds important epidemiological significance, as simple variations in the values of state variables – say, in the number of vibrios – can strongly influence the general epidemiological pattern, shifting the system trajectories to another attractor. The formation of bistability regions was also observed in other bifurcation

diagrams of epidemiological models, chiefly in Kuznetsov and Piccardi (1994) and in Earn et al. (2000), for low-intermediate levels of the contact rate of the disease, which can be taken as a proxy of the basic reproductive number for comparison with these results. In fact, in this model too it is found that bistability is to be expected for low or intermediate values of R_0.

At increasing values of the degree of seasonality ϵ, the appearance of chaotic behavior is thus verified, occurring as periodic attractors progressively double their characteristic period via what is called a Feigenbaum cascade of flip bifurcations – detailed with dashed, dashed-dotted and dotted lines in Figure 2.4; colors refer to different attractors, as explained above. Two chaotic regions are found, originated by different periodic attractors. More specifically, in region E one may find chaotic patterns generated by the limit cycle of annual period, while in region D there exists bistability between the chaotic attractors generated by the biannual and the triannual limit cycles. The trajectories arising from such parametric conditions show non-trivial signatures, as shown in the far right panels D_{PS}, E_{PS} and F_{PS} of Fig. 2.4. There the normalized values of the Fast Fourier Transform (FFT) spectra of suitably long windows of model simulations of cholera incidence (aggregated on a monthly basis) are shown, for values of the parameters corresponding to the respective areas of the bifurcation diagram. The characteristic periodicity of the cycle from which each chaotic attractor was generated leaves its fingerprint, as one can detect spikes located at 1-year (E_{PS}), 2-year (F_{PS}) and 3-year (D_{PS}) periodicities. Interestingly, however, many other frequencies, greatly attenuated when trajectories are simply periodic, emerge as important from the spectral analysis of chaotic patterns. Biannual frequencies arise where a 1-year periodicity is preponderant (panel E_{PS}), while annual frequencies emerge when 2-year and 3-year peaks explain much of the fluctuations of incidence trajectories (panels D_{PS} and F_{PS}).

A comparison with the dominant frequencies of real long-term time series of cholera incidence can be useful to determine the significance of the above results. Weekly epidemiological reports on cholera cases, reported by WHO, have been taken and aggregated on a monthly basis (distributing cases between months accordingly). It is assumed that no reports for a particular nation and month correspond to zero cases. In Figure 2.5 it is shown, again, the normalized values of the FFT spectra of the time series of new cholera cases in three countries, that are chosen for their long and mostly complete records (approximately 35 years for Cameroon and Niger, 20 for the Philippines). Spectral analysis is restricted to 5 years' periods, as a significant analysis of lower frequencies would require longer time series. Moreover, lower frequencies can be controlled by interannual variation of climatic variable (e.g. ENSO for the pacific regions (Pascual et al., 2000; Rodo et al., 2002)). This model does not account for interannual forcings and thus it cannot consistently capture frequencies lower than 5 years of period. Despite these shortcomings, important similarities are found between the spectra obtained from chaotic model trajectories and the ones depicted in

Figure 2.5, not only qualitatively, but also quantitatively. Cameroon and Niger show, in fact, strong biannual and triannual periodicities respectively, which closely match the ones characteristic of chaotic patterns in region D of Figure 2.4 (see panels F_{PS} and D_{PS}). Along with the main periodic component, both data and model show many other excited frequencies, of intra-annual, annual and multi-annual periodicity. More specifically, Niger FFT spectrum shows important biannual and annual components, which can be also observed in panel D_{PS} of Figure 2.4. Similarly, Cameroon epidemiological pattern is characterized by a set of different periodicities, not only involving the biannual main component, but also significant frequencies of lesser – 1 year and less – and greater – 3 years and more – period. This pattern is also caught by the spectrum of the chaotic trajectories of the biannual attractor, as seen in panel F_{PS} of Fig.2.4. The Philippines, instead, are characterized by an important annual component, with a much lower peak at the 2-year periodicity, a pattern also emerging from the chaotic trajectories of region E of Figure 2.4 (see panel E_{PS}).

One might wonder whether adding random noise – e.g. a lognormal factor to water input, to simulate rainfall daily variability – to models with parameters corresponding to nonchaotic patterns would lead to more realistic spectra. The answer is negative, as it emerges from these simulations: there would be only slight variations in high frequency components, as expected. Therefore, it looks like the broad and diversified spectra of actual time series can only be obtained with parameters corresponding to chaotic patterns, and not simply by adding noise to more regular patterns.

These findings raise the question of whether the difference between endemic regions, where cholera appears every year (the Philippines, for instance), and epidemic regions (Cameroon and Niger), where outbreaks are sparsely distributed along the years (Ruiz-Moreno et al., 2007; Sack et al., 2003) may well be simply ascribed to different values of R_0, which cause outbreaks characterized by different frequency. It is also worth noting that, compared to airborne diseases, such as measles (Kuznetsov and Piccardi, 1994) or influenza (Casagrandi et al., 2006), environmental drivers of waterborne diseases such as cholera can display a seasonality degree ϵ as high as 80% or more. In fact, it is expected the values of ϵ in historically endemic regions, where monsoonal regimes prevail, to be typically high, as water input is definitely concentrated in specific periods of time during the year. According to the model outcomes, it should therefore be expected, in general, complex epidemiological patterns of the disease.

To verify the robustness of this global picture of the model behaviors the bifurcation diagram for a different value of the water retention time λ^{-1}, which is the third parameter that can widely vary in different situations, is also analyzed. In particular, the value of λ is increased to 0.1 day^{-1}, thus shortening the average water retention time to 10 days. The previous value (0.01 day^{-1}), in fact, identifies reservoirs with residence time of approximately 3 months. In Figure 2.6 the opposite extreme of much faster

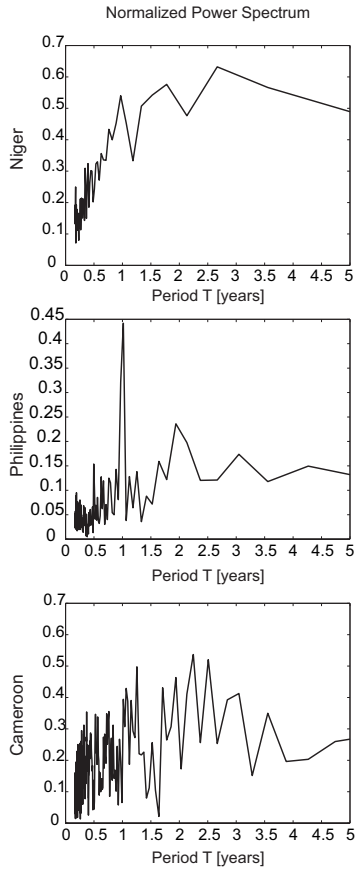

Figure 2.5: Normalized power spectra of monthly aggregated weekly epidemiological reports (WHO) for Niger, Philippines and Cameroon.

hydro-epidemiological dynamics is instead considered, in which the water reservoir in the system gets inundated or refilled by rainfall, but is rapidly drained. This new parameter setting leads to a slightly more complicated diagram, as shown in Figure 2.6. In fact, for a fixed value of R_0, the sequence of period doubling bifurcations leading to a region of chaotic patterns occurs for smaller ϵ's, compared to Fig. 2.4. For instance, chaos is possible with $\epsilon = 0.2$ and $R_0 = 2$, whereas these parameters corresponded to a 2-year cycle in the system with $\lambda = 0.01$ days^{-1}.

At the same time, the robustness of the previous analysis is confirmed by observing that the basic topological structure of the bifurcation diagram is respected: one can still find the bistability region (again, between the blue and the red solid lines) and the two chaos regions, for different values of R_0. The upper chaotic region is closed and restricted by the seasonality range $0.25 < \epsilon < 0.85$. Higher or lower values of the forcing amplitude will lead to periodic dynamics, with return time of one year. Important differences arise, however, from an epidemiological viewpoint, between period 1 cycles at low and high values of ϵ. It is found, for instance, that the peak

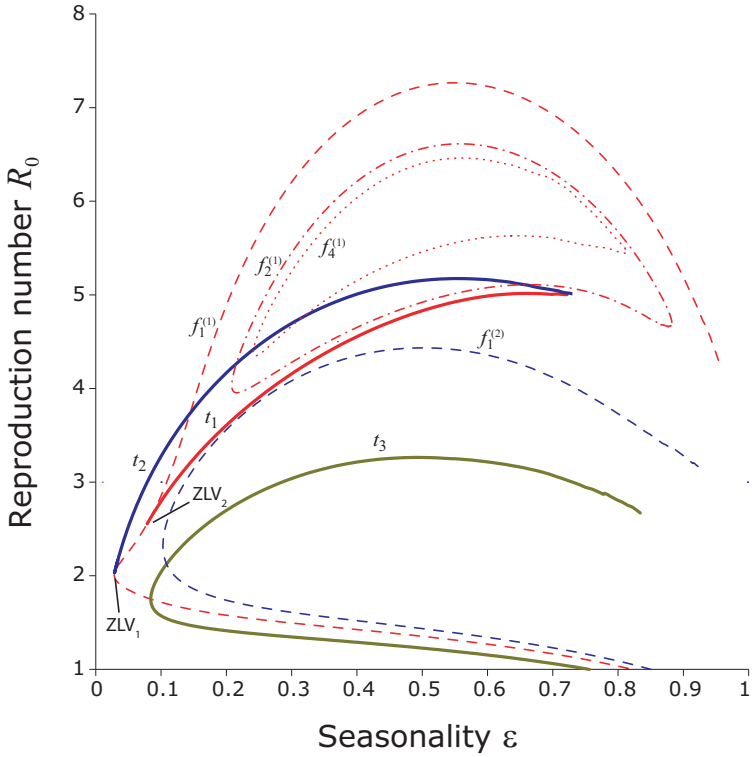

Figure 2.6: As Fig. 2.4, but for $\lambda = Q_0 = 0.1\, day^{-1}$.

prevalence for $\epsilon = 0.1$ is 0.0129, while a value of 0.065 for $\epsilon = 0.9$ is obtained at $R_0 = 6$. Moreover, the tangent bifurcation of period three cycles (t_3 curve) is again observed.

Another parameter whose value has been recently questioned is the duration of immunity after recovery from cholera infection. Even if the great majority of immunological studies and modeling exercises indicate that immunity should last between 3 and 10 years, King et al. (2008) reconstructed from cholera incidence time series a short-term immunity of only 12 weeks. To compare the behavior of the model for different values of the rate of loss of immunity, in Figure 2.7 the results pertaining six different values of immunity duration are shown, namely 12 weeks, 1 year, 3 years, 5 years, 7 years and 10 years. The analysis here shown is a so-called Feigenbaum diagram, displaying the values of cholera prevalence peaks as a function of the degree of seasonality ϵ, for a fixed basic reproductive number $R_0 = 5.5$. Results prove to be consistent with these also for $R_0 = 2.5$ (not shown). Panels (A), (B) and (C), corresponding to the shorter durations of immunity, do not show any complexity arising. Apart from an increasing amplitude of the prevalence peaks, driven by the more pronounced seasonal oscillations of water input, the annual periodic pattern is qualitatively maintained, except for a flip bifurcation emerging in diagram (B), corresponding to $1/\rho = 1$ year, which does not affect the general annual periodicity of such pattern. Chaotic patterns only appear in the diagram related to the longest immunity periods (panels

D-E-F), coherently with what was obtained in Figure 2.4. The difference is explained by the fact that the faster the loss of immunity is, the more a SIRS-based model like the one presented resembles a simple SIS model, with practically no role of the R class. This induces faster replenishment dynamics of the Susceptible compartment and a simple synchronization of epidemiological dynamics with the external forcing.

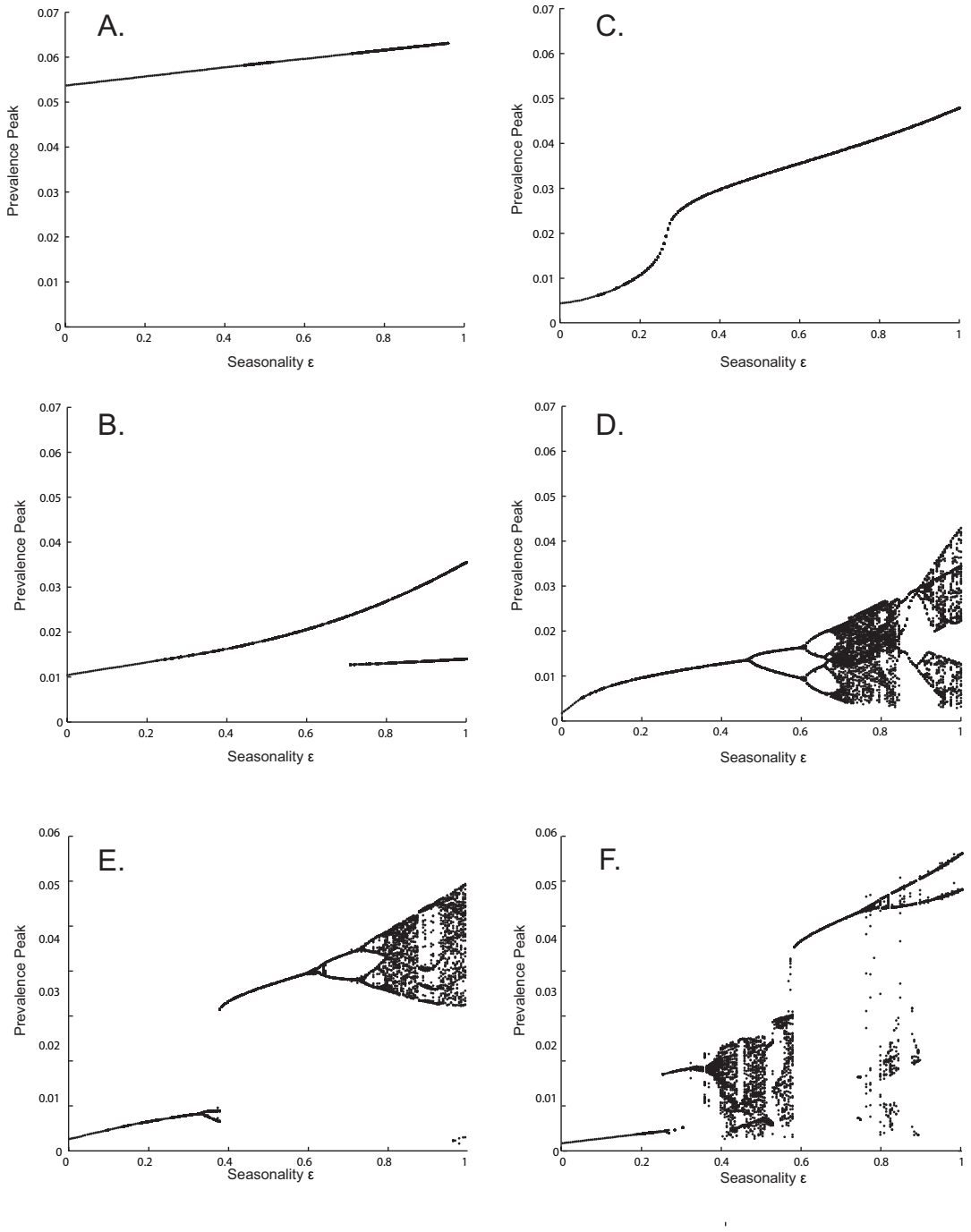

Figure 2.7: Feigenbaum diagrams for varying values of loss of immunity rate: $1/\rho = 12$ weeks (A.); 1 year (B.); 3 years (C.); 5 years (D.); 7 years (E.); 10 years (F.). $R_0 = 5.5$ and $\lambda = Q_0 = 0.01\, days^{-1}$. Unspecified parameters as in Figure 2.2.

2.5 Conclusions

A new model to study cholera dynamics has been proposed, which explicitly takes into account temporal fluctuations of the water volume hosting the pathogen *Vibrio cholerae*. To this end, two compartments have been added to the standard SIB model for cholera, originally proposed by Codeço (2001): the recovered compartment R, which is important to understand the long term dynamics of the disease, and the hydrologic fluctuations of the water volume W available to a certain human community. In order to mimic the processes of dilution that control bacterial concentrations in the system, the input (rainfall or runoff) to W, which varies over time following a simple (and regular) seasonal pattern, is periodically forced. It has been shown that, even with this simple description of hydrological phenomena, the model can reproduce the epidemiological courses that characterize the periodic insurgence of the disease in various areas of the world. In particular, the model can quantitatively describe the delay between drought seasons – during which there is a high vibrio concentration in the water – and prevalence peaks. This has been observed in countries where cholera has been present for decades (e.g. Iran). The effect of bimodal forcings, which better describe the yearly rainfall patterns of other geographical regions (e.g. India), has been also investigated. If forced with the actual values of water input that are observed for endemic regions in the province of Madras, India, the model is able to quite well reproduce the recorded patterns of seasonal cholera which, in this particular case, occurs twice per year. Not only the timing, but also the relative amplitude of peaks is in fact quite satisfactorily reproduced. Finally, the whole range of model behaviors has been analyzed in the parameter space of the degree of seasonality ϵ and of the basic reproduction number R_0, which was chosen as most significant and subject to wider variations (other parameters are assumed from the relevant literature or from available data). Bifurcation analysis revealed the presence of behaviors also found in real data, including non-periodic, chaotic dynamics. It was observed an increased complexity of model behaviors for intermediate or high values of ϵ coherently with previous results Kuznetsov and Piccardi (1994) . Also, two distinct parameter regions have been found where different chaotic attractors emerge, which however display different average epidemic frequencies (which increase with R_0). Power spectra referring to chaotic trajectories show significant similarities with real data time series in both endemic and epidemic countries. It is thus hypothesized that the difference between endemic and epidemic regions might be different intensities of the disease, as expressed by different basic reproductive numbers. Moreover, different values of the loss of immunity rate show that only a relatively long time of immunity can induce complex dynamics in the system.

Despite the ability of such a simple model to describe many actual temporal patterns of cholera, there exist some cases that cannot be explained by the dilution mechanism only. Patterns, such as those typically observed for certain Bangladesh regions

(Matlab in particular; see e.g. Emch et al. (2010)), display two outbursts per year - one between April and May (spring peak), the other around October/November (autumn peak), even if there is only one rainfall peak per year. It is speculated, though, that this peculiarity might be due to the concurrence of two mechanisms: a) the disruption of sanitation systems during floods or hyper-supply of pathogens, which counter the dilution effect, and b) spatially explicit dynamics, which is not considered here but that are accounted for in several studies (see e.g. Bouma and Pascual, 2001).

As the recent Haiti epidemics testifies, the development of analytical tools that can help not only to predict but also to understand the basic processes underlying the intertwinement between environmental variability and the insurgence of cholera is much needed. Models including both the epidemiological compartments and the environmental matrix in which the disease unfolds can greatly help in setting up intervention policies and emergency management alternatives. Making them space and time-specific can strengthen their effectiveness and reliability. Needless to say, such effort is most effective when supported by consistent data surveys, which can help in detailing the environmental and epidemiological characteristics of the area hit by the disease outbreaks.

3 Local interactions of *Vibrio cholerae* with hydro-climatology and phytoplankton

3.1 Introduction

Even though the major outbreaks of cholera of recent years have been reported in countries newly invaded by the disease (WHO, 2011), understanding the dynamics of survival of *Vibrio cholerae* O1 in endemic contexts may provide useful information on the conditions of survival of vibrios in the environment and of seasonal resurgence of the disease. Bangladesh, in this respect, represents an ideal case study, as cholera has been present in the region since its appearance in the XIX century.

Seasonality of cholera in Bangladesh and other endemic areas has been linked to various hydro-climatological signatures, at both local and regional scale. Strong hydrological forcings are characteristic of the monsoonal climate of Bangladesh and have been shown to drive the epidemiology of cholera in the region (Akanda et al., 2009; Bertuzzo et al., 2012). In fact, at the country scale there exist two peaks of cholera prevalence within a year, one coinciding with the period of more intense drought (April-May), the other (October-November) following the monsoonal season. Locally, climatological drivers such as rainfall and sea surface temperature have been shown to reproduce convincingly epidemiological patterns of cholera insurgence also at interannual scale (Pascual et al., 2000; Huq et al., 2005; de Magny et al., 2008). Indeed, such drivers are also a proxy of the seasonal fluctuations of planktonic populations, with which *V. cholerae* interacts often in complex ways (Lipp et al., 2002).

Vibrios, in fact, have been found to survive in a viable-but-non-culturable state (Colwell et al., 1985), in which they are able to cope with unfavorable environmental conditions in association with other, more complex organisms, such as phyto-plankton (Islam et al., 1994) and zoo-plankton (de Magny et al., 2011). Accordingly, the viability of *V. cholerae* populations in freshwater, even in inter-epidemic periods, has been confirmed by several studies (Colwell et al., 1977; Alam et al., 2006), including the one presented in Chapter 5 of this thesis.

Other than direct association with the planktonic fauna – for instance, due to *V. cholerae* using algal metabolites as nutrients (Islam et al., 1994) – indirect boosting of vibrio growth is assumed to be induced by phytoplankton blooms, during which O_2 and CO_2 water content is altered and the subsequent higher levels of pH seemingly favor *V. cholerae* over other bacteria (Cockburn and J.G., 1950). Also, decomposition of algal matter may constitute a further reserve for vibrio growth in the post-bloom, post-monsoonal season (Lipp et al., 2002). Moreover, phyto-plankton blooms drive also zoo-plankton growth, whose chitin is degraded by bacteria, including vibrios, for nutrition (Nalin, 1979).

In this Chapter it is argued that the hydrologic fluctuations of the water reservoir are also important in determining the patterns of concentration of vibrios and other organisms, directly and indirectly. In Bangladesh, in fact, the first cholera peak coincides with the lowest yearly river discharge (Akanda et al., 2009). Also, rainfall runoff has been shown to cause the reinsurgence of the disease due to contamination of drinking water reservoirs with open air defecation sites – for instance, in Haiti (Rinaldo et al., 2012). In the case of rural Bangladesh, which is here analyzed, the common water reservoir is represented by a pond of variable size, whose volume fluctuates in time due to hydrological forcings. Pond water is used by the population for domestic purposes and, after boiling, for drinking. It has been speculated that also in this case rainfall runoff may induce nutrient transport into the pond and, thus, further growth of the microbial community (Islam et al., 1994). In many locations, in fact, especially in inland regions (Akanda et al., 2009; Bertuzzo et al., 2012) a post-monsoonal peak of cholera prevalence is observed. The theoretical analysis carried out in the previous Chapter is thus extended, in this case, to an experimental setting.

Data collected from a site in the region of Matlab (Bangladesh), an acknowledged endemic area for cholera which has been monitored by the International Center for Diarrheal Diseases, Bangladesh, since 40 years ago, is analyzed. Both meteorological and biochemical data was collected, along with the concentrations of *V. cholerae* and phytoplankton. The aim is to unify evidences of dependence of vibrio growth on abiotic and biotic factors, taking into account both hydrological dynamics and ecological processes of interaction of *V. cholerae* with other organisms.

Phytoplankton dynamics have been a subject of intense analysis in the modeling literature (Scheffer, 1991). The effects of the interaction with fish (Doveri et al., 1993; Scheffer et al., 2000; Medvinsky et al., 2002) and zoo-plankton (Scheffer et al., 1997b; Gragnani et al., 1999) populations and of phyto-plankton interspecies competition (Scheffer et al., 1997a) have been studied widely. Here the nutrient-vibrio-phytoplankton dynamics, coupled with hydrologic fluctuations, which contribute to modify concentrations of the ecological compartments, are coupled together. The final aim is to synthesize all the processes – hydroclimatological and ecological – which have been identified to lead to vibrio growth and cholera endemic resurgence in the region, in a

unique mechanistic system.

3.2 Data

3.2.1 Data collection

A parallel collection of hydro-climatological and biochemical data has been conducted in a pond located in the rural area of Matlab (Bangladesh), located approximately 55 km south of the capital Dhaka. The pond is surrounded by elevated ground, where villages lie, on every side. A topographic survey of the area was carried out in order to reconstruct in detail the morphology of the pond. In Figure 3.1 such survey is summarized by a map of the area, including the contour lines defining the pond bottom profile and the sections of the pond on the two main directions. There one can appreciate some of the features of the area: above all that, although being close to the river, there is an embankment that protects the pond and the surrounding villages even in times of flooding, which are frequent during the monsoons. An additional characteristic of the pond, not evident here, is that infiltration of water can be neglected due to the clay surface layer.

In the beginning of May 2011 two Sensorscope® stations were installed. Sensorscope® is a system which is able to measure physical quantities with high temporal and spatial density – these data can also be accessed remotely from an online platform (Fig. 3.2), in which also the status of the stations can be monitored in real-time. The station is completely energy-independent thanks to an integrated system of solar cells plus rechargeable batteries. It requires a GPRS-activated SIM to be able to connect to the Internet. A backup save for the recorded data is provided by a memory card where data is continuously stored.

One of these was set up on the edge of the pond, the other served as "control", operating inside the ICDDR,B institute in Matlab and potentially supplying information in periods of limited connectivity or in case of other technical problems. The stations have been recording sensible data of rainfall, air and water surface temperature, air humidity and solar radiation from the sites with very high temporal resolution ($\simeq 2$ minutes). The station installed beside the pond also measures the distance between the water surface and a mechanical arm leaning on the side of the pond (Fig.3.1). Note that the station was installed by the edge of the water at the time of minimum water volume in the pond and opportunely elevated from ground to avoid submersion of sensors during monsoons, when water level rises of a few meters.

Concurrently, from 16/05/2011 water samples from the pond were analyzed on a biweekly basis, in order to evaluate presence and concentration of *V. cholerae* and phyto-plankton. Physico-chemical parameters were measured contextually. Air, wa-

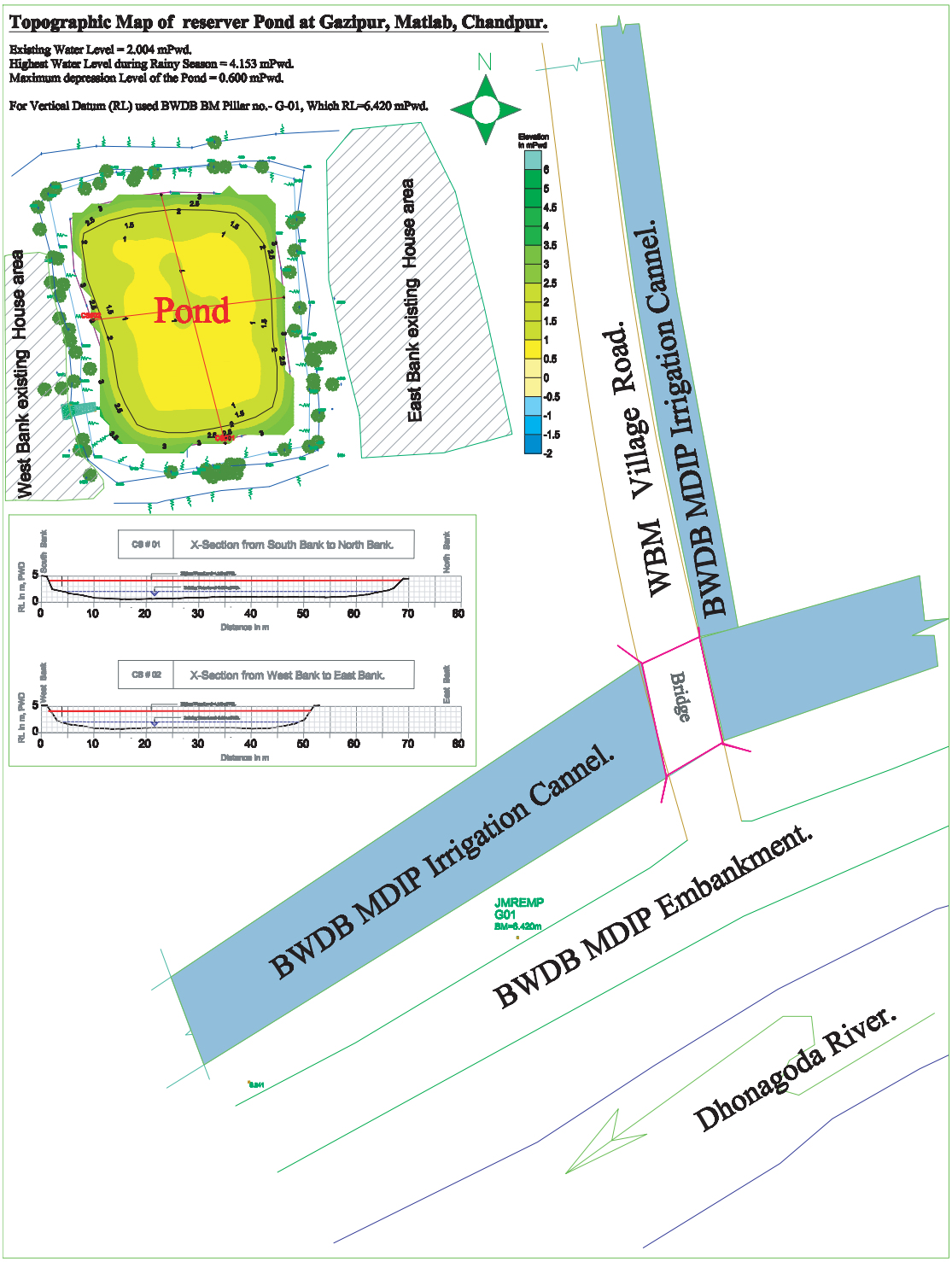

Figure 3.1: Topographic survey of the study area. The profile of the pond along the two main geographical directions is also shown.

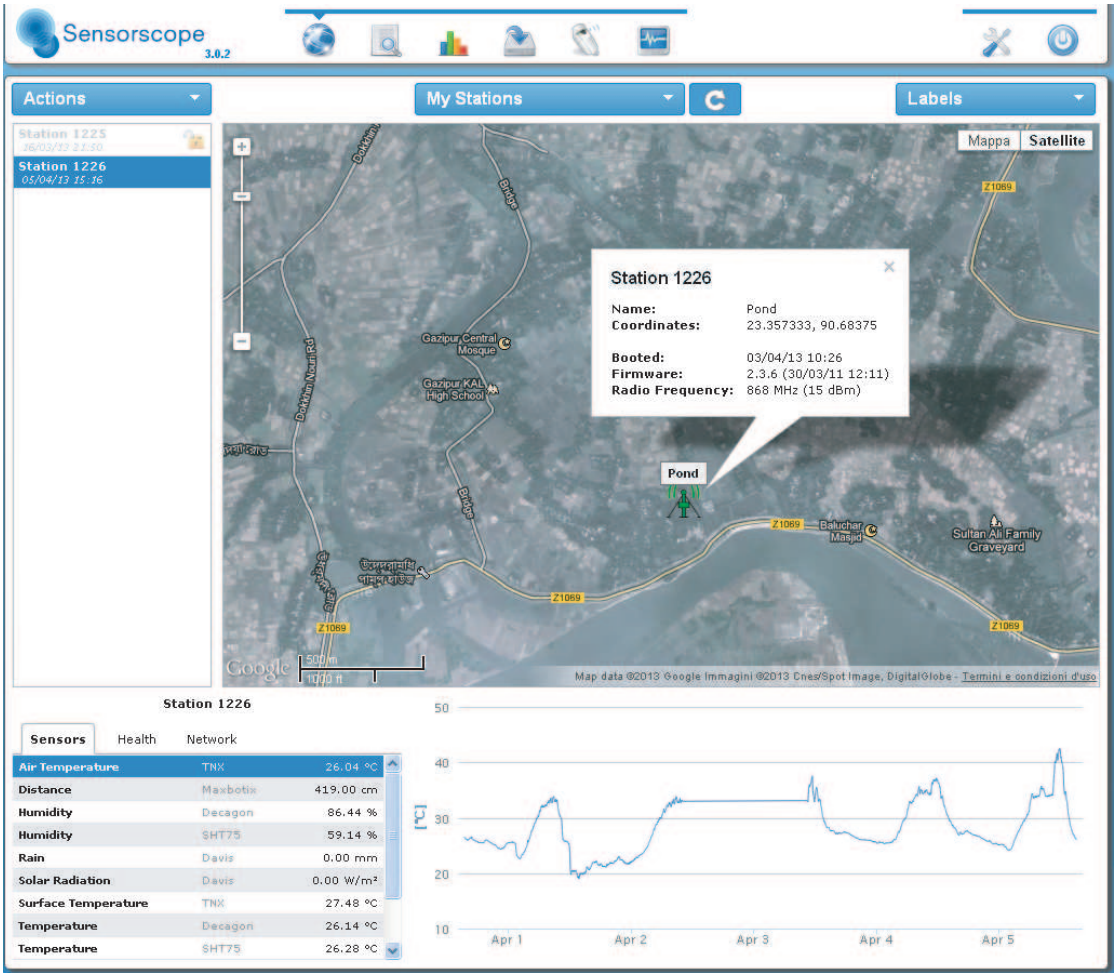

Figure 3.2: Screenshot of the online Sensorscope® interface, named Climaps®. The position of the station is shown, together with the time series of the measured variables. The connection status (green indicating that the station has connected to the Internet in the last few hours), the charge of the batteries and the space remaining on the memory card.

Figure 3.3: Panel A: installation of the Sensorscope® station in the study site; panel B: view of one of the villages surrounding the pond under study.

ter temperature and pH were measured with a portable meter (HANNA model HI991001), while turbidity, conductivity, salinity and total dissolved solid content were measured with a portable HACH Sension 5.

Two liters of pond water were collected and transported to Dhaka for analysis in sterile dark Nalgene bottles (Nalgene Nunc International, St.Louis, MO), placed in an insulated plastic box with ice packs maintaining temperature at $4-8$C. Within 24h from collection, the samples were checked for either phyto-plankton or *V. cholerae* O1 counts. For the former, Water was discarded from the upper portion by siphoning method from the bottles in which the collected water samples were kept. More than 900 ml of water was discarded from the 1 liter sample. During siphoning any jerking was avoided so that the settled phytoplankton was not disturbed or mixed up again. Then the 100 ml of concentrated phytoplankton sample was carefully transferred from the bottom of the bottle into a sterile 4-oz bottle. The volume of the concentrated plankton sample was written on the 4-oz bottle for the back-calculation of phytoplankton abundance. For long-term preservation, formalin was added to make 4% final concentration of formaldehyde. The preserved concentrated phytoplankton samples were gently shaken and properly mixed; one ml sub-sample was taken by pipette and transferred into a Sedgewick-Rafter (S-R) counting cell. The S-R cell was carefully covered with a glass-slide so that no bubble was formed inside phytoplankton sample. The S-R cell with phytoplankton sample was placed under a compound binocular microscope (Olympus). Then the phytoplankton samples were observed at a magnification of 10x for identification and enumeration.

Vibrio counts were instead performed with the direct fluorescent antibody (DFA) technique (Kogure et al., 1979). A fluorescent dye was added to a fraction of the sample, so that *V. cholerae* cells would be recognized by direct microscopy. A small quantity of

nalidixic acid (NA) and yeast extract (YE) was required for pre-incubation of samples. Cell division in gram-negative bacteria was prevented by nalidixic acid; a specific inhibitor of DNA synthesis that prevented cross-wall formation but other synthetic pathways continued to function. The result was formation of elongated, metabolically active cells in the presence of nutrient from yeast extract. Therefore, viable cells appeared elongated when stained with fluorescent antibody. A species-specific antiserum with a fluorescent dye (New Horizon Diagnostics Corporation) was used for the direct fluorescent antibody (DFA) technique.

3.2.2 Data analysis

In Figure 3.2 the patterns of a sub-sample of the dataset are shown in the whole timespan of collection (16/05/2011-31/12/2012). In Panel A of Fig.3.2 one can observe almost simultaneous patterns of phytoplankton and *V. cholerae* O1 abundance ($p - value \simeq 10^{-7}$ for linear correlation between the two quantities against the constant model null hypothesis), showing both post-monsoonal and pre-monsoonal peaks, which coincide with the observed pattern of cholera cases in the region, which show two annual peaks in the same periods, the spring one being usually higher than the other. The highest of these appears in April 2012, when water volume is at its lowest in the pond. Several mechanisms may contribute to this algal bloom, chiefly the increasing hours of sunlight and the higher concentration of nutrients. In this period, pH (panel B, Fig.3.2) increases, as a result of photosynthetic activity. In this situation, *V. cholerae* is expected to be under favorable competitive conditions, as high-pH solutions are often used to isolate vibrios from environmental samples (Lipp et al., 2002).

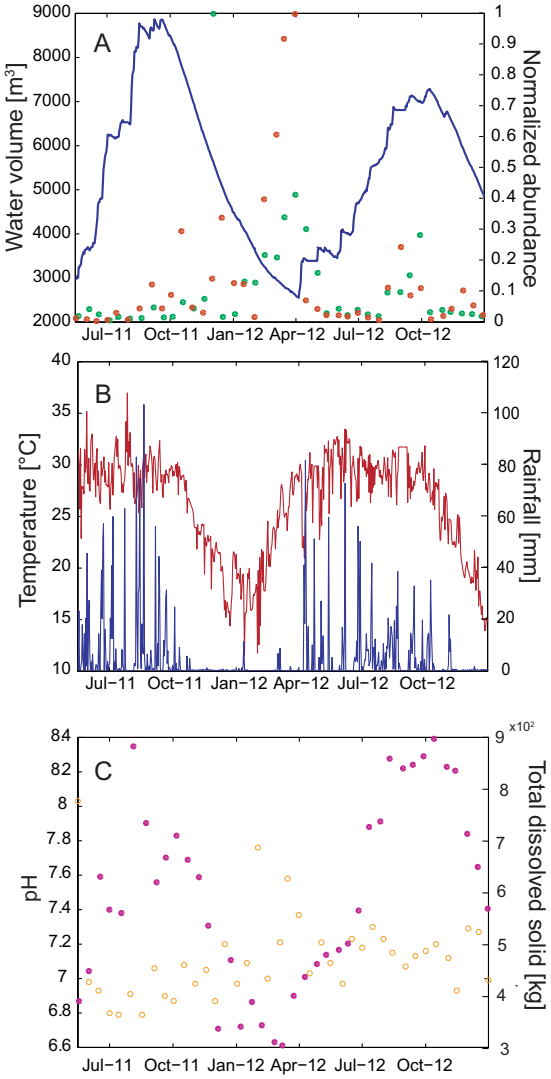

Figure 3.4: Patterns of sampled and estimated variables between 16/05/2011 and 31/12/2012. Panel A: volume of water in the study pond (blue solid line), as evaluated by water stage measurement and stage-volume curve emerging from land survey; phytoplankton/*V. cholerae* O1 normalized abundances in the pond (green/red dots); Panel B: sub-sample of the observed hydro-climatological variables: rainfall (blue solid line) and air temperature (red solid line) patterns; Panel C: sub-sample of the physicochemical parameters: pH (yellow circles) and Total Dissolved Solids (lilac dots).

Clear evidence emerges, though, of additional mechanisms intervening in determining the observed patterns. Even though it may be difficult to disentangle, at this stage, the effect of the extreme rainfall events of the summer monsoon, washing out the system, from self- and trophic-regulated dynamics of planktonic population, wet seasons display higher quantities of Total Dissolved Solids (TDS), a measure of both inorganic and organic substances contained in water, in molecular, ionized or microgranular suspended form. This may be a possible indication of rainfall runoff also transporting nutrients into the pond, causing further multiplication of microbial populations – also in conditions of low photosynthetic activity, as many phytoplankton species are also heterotroph. This mechanism of rainfall-induced contamination of local water reservoirs has been observed also in the case of Haiti epidemics, in which strong rainfall events seemed to have caused a revamping of the disease in the region after the first outbreak Rinaldo et al. (2012). The complexity of the interactions of the biota – especially the microbial fraction – with hydro-climatological and physicochemical parameters stands out also from the fact that no significant correlation emerges from ANOVA analysis between the observed quantities and abundances of either *V. cholerae* O1 or phytoplankton abundance (highest p-value being $p = 0.3$ for correlation between solar radiation and phytoplankton counts). For this reason, a process-based framework is proposed to understand and reproduce the dynamics of *V. cholerae* in aquatic habitats.

To disentangle the effect of water volume fluctuations driving the observed concentrations from the growth caused by intrinsic ecological dynamics, a simple experiment is run and shown in Fig. 3.3. There the abundance of phytoplankton and vibrios is correlated to the inverse of the water volume, i.e. the hypothesis that there is a constant population, whose concentration is merely controlled by the increase or decrease of water volume, is tested. This is equivalent, in brief, to writing:

$$C^* = \frac{\overline{P}}{W} \tag{3.1}$$

where C^* is the observed concentration of the populations under study, $\overline{P}$ is a constant value representing the abundance of a given population and W is the water volume in the pond. In Fig.3.3 $\overline{P} = 1$ is assumed, for simplicity. The correlation analysis between the two quantities reveals that this simple model, which only includes the effect of hydrology in driving the observed concentrations, is able to explain only a fraction, albeit non-negligible, of the variations that have been recorded ($R^2 = 0.272$ for phytoplankton, $R^2 = 0.377$ for *V. cholerae* O1). This suggests that a more comprehensive description of the ecological dynamics is required to correctly reproduce the observed patterns of vibrio concentration but, also, that the role of the hydrological forcings cannot be neglected. Consequently, both the hydrological and the ecological

components will concur to the definition of the modeling framework, which will be detailed in the following section.

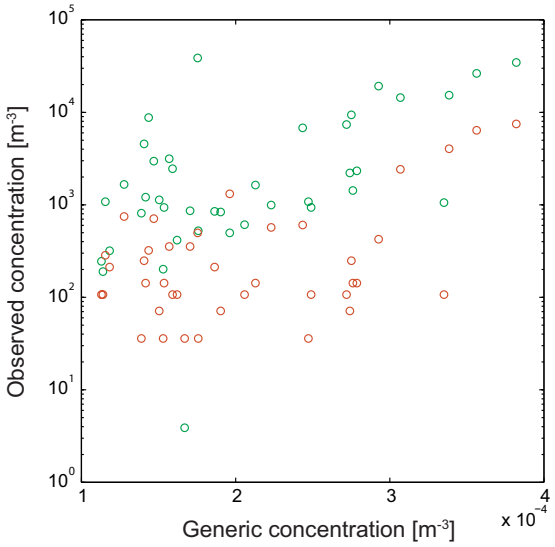

Figure 3.5: Scatter plot of the measured concentration of phytoplankton (green circles) and *V. cholerae* O1 (red circles) populations against the inverse of water volume, representing the concentration of a generic, constant population $\overline{P} = 1$.

3.3 The Models

3.3.1 Water volume dynamics

In order to verify whether the hydrological dynamics, that are governing the fluctuations of the water volume in the pond, are being understood correctly, all water inputs and outputs are first characterized. It is assumed that the variation of water volume in the pond can be described by the following differential equation:

$$\dot{W} = (R(t) - E(t))A_{pond} + Q_{corr} \tag{3.2}$$

where $R(t)$ is the amount of rainfall falling directly on the pond, $E(t)$ is the evaporation rate and Q_{corr} expresses the amount of rainfall that falls on the area surrounding the pond and is drained into it.

The evaporation rate is first estimated from the formulation indicated by Priestley and Taylor (1972), which has been widely used for characterizing open water evaporation

in particular (Wang et al., 2006; Arasteh and Tajrishy, 2008; Agam et al., 2010):

$$E = a\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda} \tag{3.3}$$

where R_n is the net incoming radiation. This can be expressed as

$$R_n = (1-\alpha)R_s - R_{nl} \tag{3.4}$$

where α is the surface albedo, which is taken here equal to 0.04, a typical value for open water surfaces. R_s is the measured (daily averaged) incoming short-wave solar radiation, while R_{nl} stands for the isothermal long-wave net radiation emitted by the water body:

$$R_{nl} = f_c f_h \sigma({T_K}^4) \tag{3.5}$$

where f_c is the cloudiness factor, f_h is the air humidity correction, σ is the Stefan-Boltzmann constant, T_K is the daily mean water surface temperature.

In Table 3.1 the calculations required to estimate the parameters contained in this formulation are shown, together with their data requirements.

Table 3.1: Parameters required for the estimation of the evaporation rate with Priestley-Taylor formulation.

Parameter	Formulation	Units	Data requirement
Slope of saturation vapor pressure Δ	$\frac{4098 \cdot 0.611 exp(17.27T/T+237.3)}{(T+237.3)^2}$	kPa/°C	T (air temperature)
Latent heat of vaporization λ	$(2.501 - 0.002361T) \cdot 10^6$	J/kg	T (air temperature)
Psychrometric constant γ	$\frac{C_p P}{0.622\lambda}$	kPa/°C	Specific heat capacity of moist air (C_p), atmospheric pressure (P)
Extraterrestrial solar radiation R_a	$\frac{1}{\pi} K_{sc}(1 + 0.033cos(\frac{2\pi}{365}J))(\omega_s sin(\vartheta)sin(\delta) + cos(\vartheta)cos(\delta)sin(\omega_s)$	J/(m^2 d)	Solar constant (K_{sc}), latitude (ϑ) and solar declination (δ) angles, Julian day (J)
Cloudiness factor f_c	$1.35(\frac{R_s}{0.75R_a} - 0.35)$	unitless	R_s (solar radiation)
Air humidity correction factor f_a	$0.34 - 0.140.611 exp(17.27T/T + 237.3)RH/100^{1/2}$	unitless	Air temperature (T), relative humidity (RH)

Only one free parameter is left, that is the correction factor a, which is calibrated from the first dry period observed (01/10/2011-05/04/2012; Fig.3.3 Panel A), in which the corrivation input is neglected. From this assumption it is possible to write the equation describing water volume dynamics as $\dot{W} = (R(t) - E(t))A_{pond}$. The result shown in Figure 3.3, panel A, is obtained for $a = 3.231$. As the validation test shows (Fig. 3.3 panel B, validation window: 07/11/2012-01/03/2013), the model can reproduce very well the behavior of the evaporating pond during the greater part of the dry season. A time-varying correction factor, which goes beyond the scope of this work, would possibly increase the performance of the model also when solar radiation begins to increase.

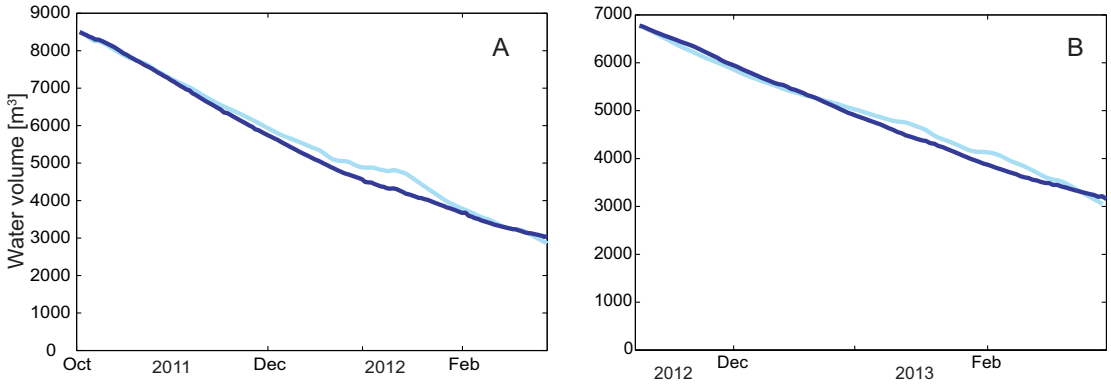

Figure 3.6: Observed/simulated water volume in the study pond (blue/light blue solid line). Panel A: calibration window (01/10/2011-05/04/2012); panel B: validation window (07/11/2012-01/03/2013).

In order to reproduce the entire water balance dynamics, the input of water, caused by rainfall runoff into the pond, still needs to be characterized. It is assumed that such inflow term is proportional to the drainage area of the pond:

$$Q_{corr} = \eta R(t) A_{drain} \tag{3.6}$$

Since detailed information on the topography of the area surrounding the pond is lacking, this term is expressed as a function of the area of the pond:

$$Q_{corr} = \phi R(t) A_{pond} \tag{3.7}$$

Model 1 is then fitted for the whole observation period (12/05/2011-01/03/2013), calibrating parameter ϕ. The best fit result is shown in Figure 3.4. The value obtained is

$\phi = 3.0612$, which means, reasonably, that $A_{drain} \sim 3A_{pond}$.

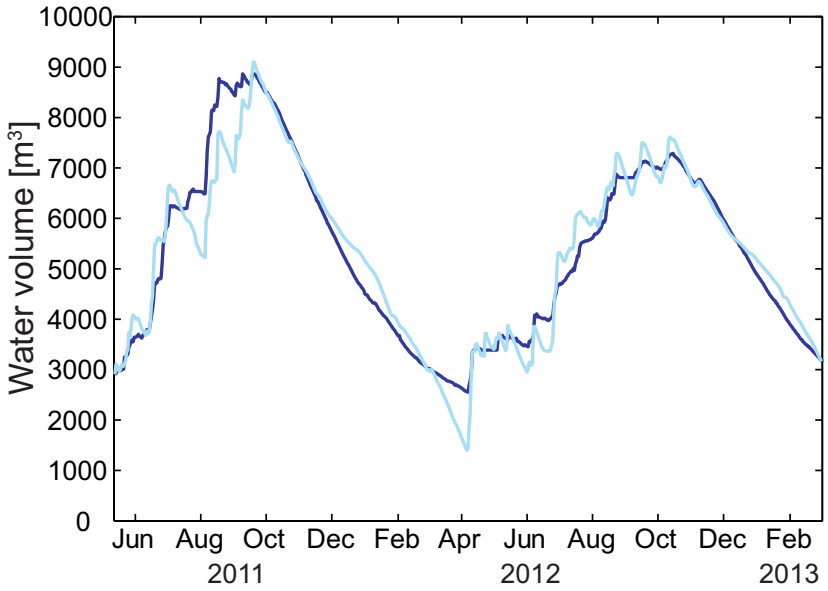

Figure 3.7: Observed/simulated water volume in the study pond (blue/light blue solid line) in the time span (12/05/2011-01/03/2013).

3.3.2 A possible approach to modeling the ecology of *V. cholerae*

Here a possible description of the combined dynamics of phytoplankton and *V. cholerae* O1 is proposed, that takes into account also the patterns of limiting factors of growth, such as nutrient concentration and temperature, and the water volume fluctuations, which are described by model 1.

The complete model would read as:

$$\dot{P} = r_P \mathcal{F}_{phs} \mathcal{G}_{nut,P} P - \mu_P P \tag{3.8}$$

$$\dot{V} = r_V (\mathcal{G}_{nut,V} + \mathcal{G}_{P,V}) V - \mu_V V \tag{3.9}$$

$$\dot{N} = \frac{Q_{corr}}{W} N_{out} - \mu_N N - g_P \mathcal{G}_{nut,P} P - g_V \mathcal{G}_{nut,V} \tag{3.10}$$

$$\dot{W} = (R(t) - E(t)) A_{pond} + Q_{corr} \tag{3.11}$$

where P is the number of phytoplankton cells in the pond, V is the abundance of *V. cholerae* O1, N is the amount of nutrients in the water and W is the water vol-

ume in the pond. The dynamics of phytoplankton are expressed as a growth limited by light availability ($\mathcal{F}_{phs}$) and by the concentration of nutrients in the pond ($\mathcal{G}_{nut,P}$), as observed in other studies (Scheffer et al., 1997a). Vibrios can instead feed on the free nutrients N ($\mathcal{G}_{nut,V}$) but also on the byproducts of phytoplankton photosynthesis (Islam et al., 1990, 1994), so that an additional growth term is regulated by phytoplankton concentration P ($\mathcal{G}_{P,V}$). The parameters r_P and r_V represent the maximum growth rates of the respective compartments, while μ_P and μ_V represent the specific mortalities.

The function $\mathcal{F}_{phs}$ can be expressed as:

$$\mathcal{F}_{phs} = \frac{h_s}{h_s + q(E_b + k_P P)} \tag{3.12}$$

where h_s depends on the number of sunshine hours (i.e. on solar radiation). This photosynthetic term also declines with increasing turbidity of the system, which is described as a sum of background (E_b) and of plankton-induced turbidity ($k_P P$). Moreover, functional responses $\mathcal{G}_{x,y}$ are usually written as Holling type II functions:

$$\mathcal{G}_{x,y} = \frac{x/W}{h_{xy} + x/W} \tag{3.13}$$

The variation of nutrient content in the water is here given by the influx of runoff water Q_{corr}, which carries along soil content, such as fertilizers, that are present in the pond drainage area in a quantity N_{out}. Nutrients are degraded naturally at rate μ_N and are uptaken by phytoplankton and vibrios according to their respective grazing rates g_P and g_V.

3.4 Conclusions

The following conclusions are worth mentioning:

- the concentrations of phytoplankton and *V. cholerae* O1 have been monitored in a pond in rural Bangladesh, one of the common water reservoirs for domestic use in the area, together with abiotic and biotic factors; the observed patterns, which are closely correlated, show both pre-monsoonal and post-monsoonal peaks in both concentration and abundance of such organisms, something that matches the observations of cholera cases in this particular region and in Bangladesh in general;

- the correlation between all the measured quantities and vibrio/phytoplankton abundances has been tested, finding no evidence of significance for any possible pairing; it is suggested that complex interactions between hydro-climatological and physicochemical parameters and the ecological dynamics of phytoplankton and vibrio populations may explain the observed patterns;
- a simple model is tested in order to describe the water volume fluctuations observed in the pond, in order to identify the relevant hydrological processes; this model has been fitted using data collected locally, with high spatial and temporal resolution, and could thus be applied, with limited corrections, to other similar cases in the area, to reproduce the water volume dynamics of open surface ponds, in which infiltration processes can be neglected; in this case, remote satellite measures of hydro-climatological variables could be used as inputs of the model;
- an ecological model is proposed to reproduce the combined dynamics of phytoplankton and *V. cholerae*, taking also into account the process of nutrients inflow and degradation, together with the water volume fluctuations which have been previously characterized; this model can be used to reproduce the observed patterns and to test parameter dependence on external inputs (e.g. the effect of temperature on vibrio growth), identifying the essential processes (for instance, by evaluating Akaike Information Criterion as the number of parameters increases).

4 Detection of *Vibrio cholerae* O1 and O139 in environmental waters of rural Bangladesh

4.1 Introduction

In light of recent cholera epidemics that have struck Haiti (see e.g. PAHO, 2010; Piarroux et al., 2011; Andrews and Basu, 2011; Bertuzzo et al., 2011; Chao et al., 2011; Rinaldo et al., 2012), South-American and in particular African countries (WHO, 2011), tools for providing quick and reliable monitoring of surface waters can be of vital importance. Such tools could enable to understand local epidemiological dynamics and could play an important role in for the control of epidemics. This is also important to pin-point specific areas of aquatic contamination in order to either restrict the access to those areas or to eradicate the bacteria from such sites by improving sanitation conditions. In this respect, portable flow cytometry could be an ideal method for monitoring *V. cholerae* concentrations in the environment due to the rapidity of sample analysis and the absolute volumetric counting that flow cytometry provides (Hammes and Egli, 2010).The flow-cytometry-based assay presented here allows measuring of both O1 and O139 serogroups of *V. cholerae*, which are the most prominent serogroups associated with large cholera epidemics and pandemics. Flow cytometry of *Vibrios* has already been applied to *in vitro* growth studies (Vital et al., 2007, 2010). However, this method has never been applied to a field-work setting, which it was aimed to demonstrate in this study. It was confirmed the growth and persistence of *V. cholerae* in freshwater even in presence of pre-existing, autochthonous bacterial communities and while subjected to variations of environmental conditions.

Here, samples from water bodies in the rural area of Matlab, Bangladesh, a well-known endemic region for cholera, were tested. Matlab is located about 55 km south-east of Dhaka, the capital of Bangladesh. In Matlab, the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) has been conducting a demographic surveillance for more than 40 years encompassing a population of 200,000 individuals. Seasonal occurrence of the disease is characteristic of this region, with two peaks of prevalence every year. The first peak usually occurs in spring (around April-May)

and the second outbreak is mostly linked to late autumn (November-December Glass et al., 1982). Even though the O139 serogroup was the cause of major outbreaks when it emerged around the Bay of Bengal in 1992 (Albert et al., 1993), reports of cholera cases associated with the O139 serogroup are meager since 1996 (except sporadic reports, see Faruque et al., 2003). The surface water system in this area is composed of ponds and irrigation canals that are diverted from the river Dhonagada, a branch of the Meghna (see Fig. 4.1 for a map of the area and its position with respect to the Bengalese territory). Pond waters in rural areas are frequently used by villagers for domestic purposes. However, prior to drinking any environmental water the people living in this area usually first boil the water. In the close proximity of villages the ponds are surrounded by elevated ground which protects the villages from flooding even during monsoon season. The volume of those ponds can vary greatly between each other and in time so that some may dry up completely in times of drought whereas others don't. Survival of *V. cholerae* in ponds has been associated with the presence of other organisms, such as blue-green algae and other members of the phytoplankton as ponds are generally rich in organic matter. The high concentration of organic matter is due to the drainage of latrines and to rainfall runoff inducing transport of fertilizers and cattle feces into the ponds (Islam et al., 1994).

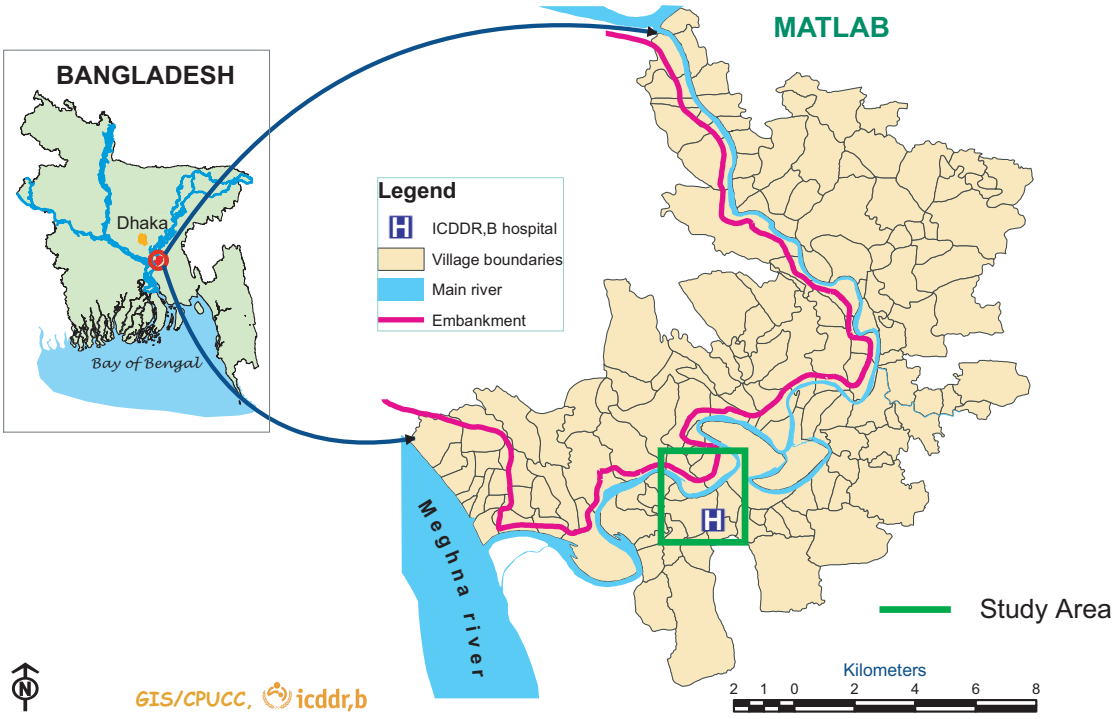

Figure 4.1: Location of the study area (green square) with respect to Bangladesh and to the Matlab area. The location of the ICDDR,B hospital is also highlighted.

The interest of this work lies, in particular, in a procedure that may aid in the field vali-

dation of mathematical models of cholera spreading. Such modeling approaches take the bacterial concentration dynamics in the water reservoirs into account thereby determining the local infection probability (Codeço, 2001). Recently, spatially explicit models of cholera have emerged as being essential in capturing the dynamics of propagation of the disease (Bertuzzo et al., 2008; Mari et al., 2012a; Chao et al., 2011; Tuite et al., 2011), but they still lack a thorough description of the local change of *V. cholerae* concentrations. In particular, it is still uncertainwhether coastal environments in general contain a viable, resident population of *V. cholerae* as it is the case around the Bay of Bengal Colwell (1996); Collins (2003); Worden et al. (2006), or whether pathogens present in those coastal regions are simply a result of the output of the hydrologic network. This uncertainty is especially true for epidemic and newly invaded areas. Thus, rapid and reliable methods such as the one presented in this work may provide an invaluable tool to refine mathematical models. Such models would be important for monitoring and predicting ongoing epidemics in time and space For the limited scope of this work, and for a proof of principle the analysis is limited to the endemic and well-studied area of Matlab.

Standard methods (and in particular culturing techniques) have often proved ineffective in detecting the presence of *V. cholerae* in environmental waters both during epidemic (Islam et al., 1994) and interepidemic (Huq et al., 1990) periods. Such an inability in detecting *V. cholerae* has been attributed to the fact that *Vibrios* can survive in the aquatic environment in a viable-but-non-culturable state (VBNC; Colwell et al., 1985). More complex analyses, based on e.g. fluorescent staining of cell surfaces or DNA are capable of detecting *V. cholerae* even in interepidemic periods (as has been shown by Alam et al., 2006). Still, these authors made use of a microscopy-based approach, which relies on manual counting (Alam et al., 2006). Instead, flow cytometry is an automated approach and allows the processing of larger samples sizes over a relatively short time. Hammes et al. (2008) showed that flow cytometry has an accuracy of less than 5% standard error and can be considered reliable also at concentrations as low as $100 - 200$ cells per millilitre (Hammes and Egli, 2010). Here, the presence was tested and the concentration was also determined of *V. cholerae* O1 and O139 at several sites in Matlab (46 ponds, 10 canal and 10 river waters) in a limited period of time (September-October 2012). This time falls in-between the two seasonal cholera peaks described for this region. For each sampling site the total concentration of cells was measured, concomitantly with basic physicochemical parameters.

4.2 Methods

Sample collection and processing Water samples were collected in the Matlab area (Fig. 4.1) between September and October 2012. Samples were taken from from 46 ponds, 10 locations in irrigation canals and 10 sampling sites within the river Dhonagada. All samples were collected aseptically in sterile dark bottles (Nalgene Nunc In-

ternational, St.Louis, MO) in the morning and placed in an insulated plastic box with ice packs maintaining the temperature at $4 - 8$C. The samples were then transported to the Environmental Microbiology Laboratory of the International Center For Diarrheal Disease Research, Bangladesh (ICDDR,B) in Dhaka for further processing on the same day. The analysis of the samples was done on the next day. From each independent bottle, 9 samples of 1 ml each (preserved in Sarstedt vials) were filtered through tilted nylon gauze filters (Celltrics Partec). The pore-size equalled 10 μm in six cases (for pathogens count) and 30 μm for the other residual 3 samples (for counting total number of cells). Thus, three parallel measurements per biological sample could be performed.

Total cell count Enumeration of total cell concentrations was done according to the method described in Vital et al. (2007). Briefly, the samples processed through the 30 μm filters were mixed with the DNA dye SYBR green (Molecular probes; 1x final concentration). The flow cytometer used for these experiments was a portable Cyflow SL Blue (Partec, Münster) equipped with an argon laser. The green fluorescent signal of the DNA-bound SYBR green dye was exited at 488 nm and the emission signal was detected at 532 nm.

Specific detection of *Vibrio cholerae* O1 and O139 via flow cytometry Specific detection of V. cholerae was done with slight modification from the method described by (Vital et al., 2007). On the collection day, all samples (1 ml each) devoted to pathogen count were subjected to UV irradiation inside a laminar air flow cabinet (Clyde-Apac HWS 120/75) equipped with a UV sterilizing chamber (Philips TUV 30W/G30T8), in order to arrest growth. Next, 10 μl of the *V.cholerae* O1 or O139 DFA reagent (New Horizon Diagnostics Corporation) were added to each sample and incubated as described in Vital et al. (2007). To verify the ability of the flow cytometer to correctly measure the concentration of *V.cholerae* O1 or O139 in water samples containing different populations of microorganisms – aside from standard positive and negative controls – environmental water was spiked with known concentrations of *V.cholerae* spanning between 10^2 and 10^5 cells per ml. The instrumental counts confirmed in all cases the order of magnitude of the cells expected in each sample. As described earlier by Hammes and Egli (2010), the accuracy of the method could not be tested below the minimum concentration of $100 - 200$ cells/ml. Thus, counts below this threshold, even though reported below for completeness, should be considered as unreliable. Presence of *V.cholerae* O1 was confirmed by DFA microscopy for the samples showing higher counts. No cross-reaction of the antibodies between *V.cholerae* O1 and O139 serogroups was observed. The mean concentration of *V.cholerae* (averaged over the three replicate measurements) sampled from independent sites were found to be significantly different from each other (ANOVA $p = 0.0003$ for O1 and $p \backsimeq 10^{-35}$ for O139).

Physicochemical parameters Air, water temperature and pH were measured with a

portable meter (HANNA model HI991001), while conductivity, salinity and total dissolved solid content were measured with a portable HACH Sension 5.

4.3 Results and discussion

To correctly enumerate the number of *V.cholerae* cells present in each sample, one has first to devise a procedure that can be applied to safely discriminate between stained cells, other organisms and background noise in the instrumental output. Previous studies indicate (Hammes and Egli, 2010), in fact, that the post-processing steps contain all the possible sources of arbitrariness and error, given that true positives need to be separated from the background and from false positives. Shown in Fig. 4.2 are the different output signals obtained by flow cytometry using the differentially treated samples, which were derived from the same environmental water source. In order to limit this source of errors the enumeration of pathogens was based on the following post-processing steps. Forward SCatter (FSC) provides a measure of the dimension of particles/cells and Sideward SCatter (SSC) is related to the irregularity of particles/cells. The processing software was set in order to only show particles that emit a green fluorescent signal. In this respect, non-stained water samples were analyzed to identify naturally fluorescing mineral particles or microorganisms, which might constitute false positives in our analysis (Fig. 4.2, panel A). Next, stained samples were analyzed . In panel B and panel C of Figure 4.2 the "cloud" that is observed in O1- and O139-positive samples is shown, generally centered around an FSC value corresponding to the one characteristic of the beads used for calibration. The calibration beads had a fixed diameter of 3 μm (red line in Fig. 4.2, panel B) and are therefore comparable in size to free-living, single cells of *V. cholerae*. Other samples showed higher FSC values (see Fig. 4.3), which may indicate the presence of small clusters of Vibrios, a feature that has been observed previously especially in interepidemic periods (Alam et al., 2006). Alternatively, such an increase in forward scatter could be caused by elongated cells. Note that bigger clusters, as the ones shown therein, would be filtered out in assay due to the small pore-size of the filters used for sample processing. The enumeration is then followed through gating of the area of interest in order to neglect background noise and residual stain (appearing at low values of both FSC and SSC).

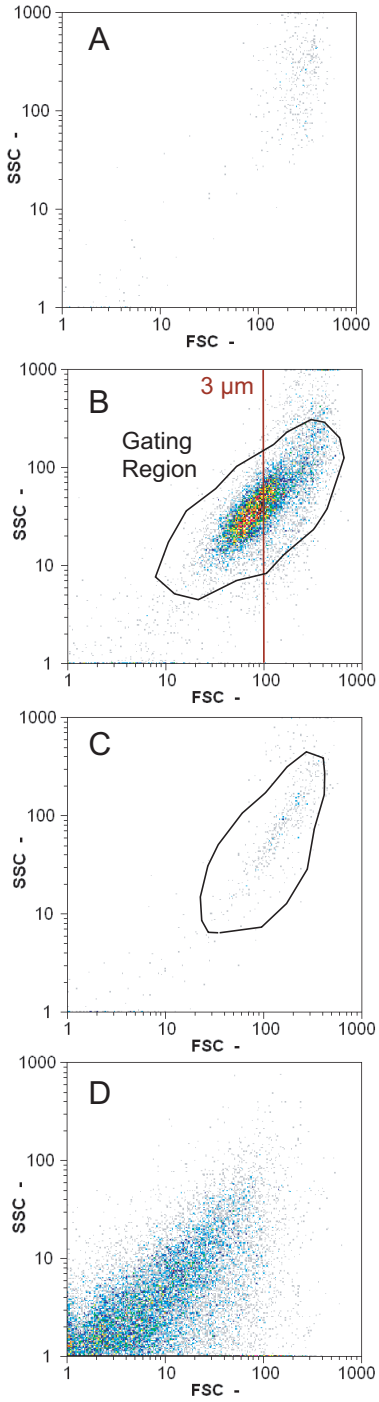

Figure 4.2: Water samples taken from pond 36 were analyzed for the concentrations of *V. cholerae* O1, O139, and total cells. The counting was done in the green fluorescent channel in order to detect the FITC-labeled antibody attached to *V. cholerae* or the SYBR green-stained DNA within cells. Colors indicate the number of counts in each pixel in relation to the total count. Panel A: Negative control to measure background fluorescence in the water sample without antibody or DNA staining. Panel B:: 10 micron-filtered water sample with O1-specific staining. The red line indicates the calibration value for spherical beads of 3μm diameter. The black solid line depicts the gated region used for final counting. Panel C: 10 micron-filtered water sample stained with O139-specific antibodies. Panel D: 30 micron-filtered water sample stained with the non-specific DNA dye SYBR-green. FSC: Forward SCatter; SSC: Sideward SCatter.

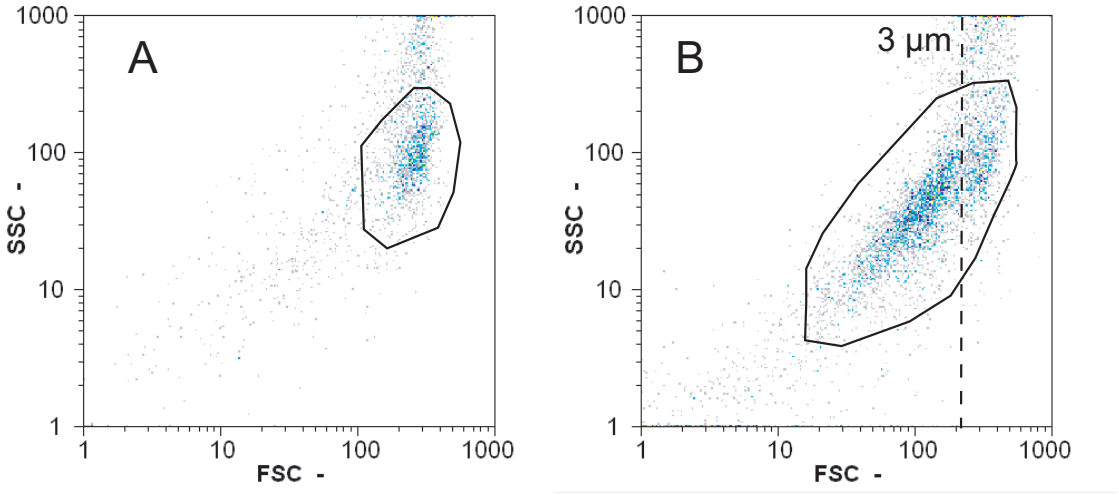

Figure 4.3: Flow cytometry-based *V. cholerae* O1 analysis of environmental water sampled from two independent ponds. The sole presence of counts in the upper SSC and FSC range shown in panel A (pond 18) is indicative of cluster formation by the *V.cholerae* cells. Panel B: Water analysis of samples derived from pond 34 suggests coexistence between free-living and clustered *V.cholerae* cells.

After setting up the technique the interest lay in the general distribution of *V. cholerae* in the different water bodies. In particular it was intriguing to see whether there is any significant difference with respect to the presence of *V. cholerae* O1, O139, and total cell counts in ponds, irrigation canals, and rivers. As depicted in Fig. 4.4 the concentration of *V. cholerae* O1 was consistently higher in ponds than the concentration of the O139 serogroup with a maximum concentration reaching approximately 70,000 bacteria per ml. The concentration range measured here complies with laboratory-based data derived from experiments of *in-vitro* competition and microcosm experiments (Worden et al., 2006; Vital et al., 2007, 2010). Note that the possible presence of bacterial clusters of bacteria – which in this analysis would count as single events – might also lead to a slight underestimation of the total concentration. The concentrations of the O139 serogroup were in the great majority of cases inferior to the one observed for the O1 serogroup, possibly indicating coexistence between the two serogroups that favors the latter in environmental conditions. The ratio of O139 to O1 amounted on average to 0.35. However, it should be noted that, even though no cross-reactions between the O1 and O139-specific antibodies towards the opposite serogroup have been previously observed and that cross-reaction against other natural aquatic bacteria were also undetectable in laboratory microcosms (Vital et al., 2007), non-specific attachment of the antibodies might still occur to a certain extent. Also, it is possible that there exist seasonal patterns of the concentrations of different Vibrio serogroups which are not synchronized, due to different adaptation to varying biotic and abiotic conditions. Interestingly, water samples taken from canals and rivers consistently showed negligible counts, with a median concentration of *V. cholerae* O1 and O139 below 100 cells per ml (Fig. 4.4, panels B and C). In contrast, the

total cell count measured in parallel from each water sample and processed through 30 micron filters before being stained for DNA did not show significant differences between the three water bodies (ponds, canals and rivers). Indeed, the average concentration was within a range of $10^6 - 10^7$ cells per ml (Fig. 4.4).

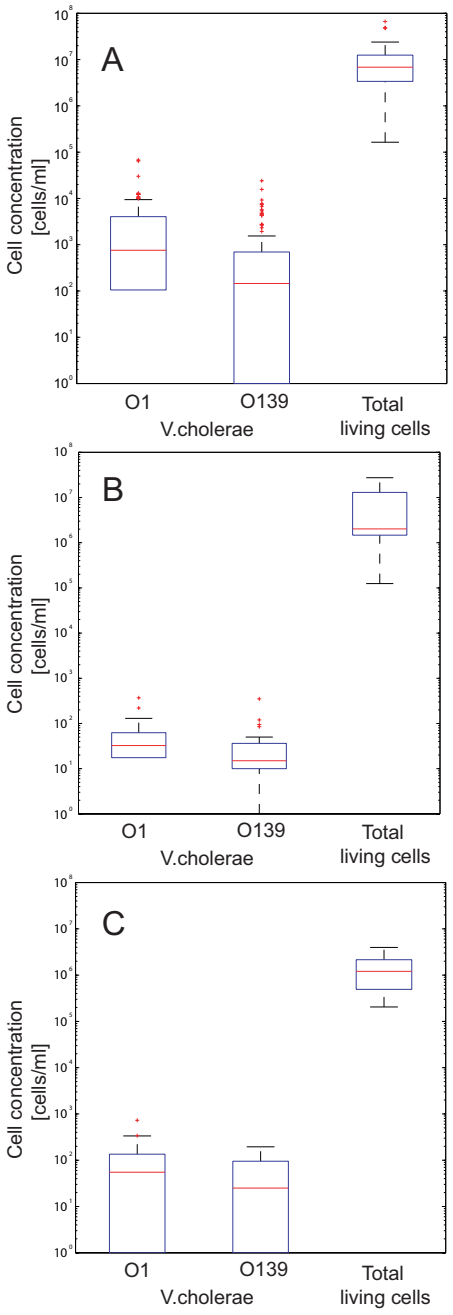

Figure 4.4: Boxplot of measured concentrations of *V.cholerae* O1 and O139 (10 micron-filtered samples) and of total living cells (30 micron-filtered samples) in water samples of the different water bodies of the study area. Panel A: ponds, panel B: canal waters and panel C: river waters). The red line indicates the median value, while the edges of the box correspond to the 25th-75th percentile values and whiskers extend to the extreme datapoints not considered outliers, which are shown with red crosses. Values from all replicative measurements were considered.

Our findings confirm that resident populations of *V. cholerae* O1 and O139 persist in freshwater environments, which is in contrast to other studies concerning the survival of these species in the Bangladesh aquatic environment. Those studies indicated saline aquatic environments as major natural habitats of *V. cholerae* (Collins, 2003; Worden et al., 2006). Indeed a clear spatial and temporal pattern of cholera spreading in Bangladesh has been reported (Akanda et al., 2009). According to this and other studies, cholera outbreaks usually start around the coastal area, in spring, to reach the inland regions in autumn. Further insight is needed to understand if transport of pathogens – as proposed by Bertuzzo et al. (2012) – may determine different patterns in different regions of Bangladesh or not. Alternatively, spatio-temporal patterns of environmental drivers, such as rainfall (as in the case of Haiti, see e.g. Rinaldo et al., 2012; Righetto et al., 2012) or spatial heterogeneity of social and environmental conditions (Mari et al., 2012a; Gatto et al., 2012) could contribute to the observed cholera seasonality in Bangladesh.

Next the fraction of pond-derived water samples falling into a given concentration range was determined for *V. cholerae* O1 and O139. As shown in Fig. 4.5 in half of the sites the concentration of *V. cholerae* O1 was above 10^3 cells per ml in half of the sites, and in few cases even above 10^4 cells per ml. In contrast, average concentration of the O139 serogroup exceeded $1,000$ cells per ml in only four of the 46 ponds analyzed (Fig.4.5).

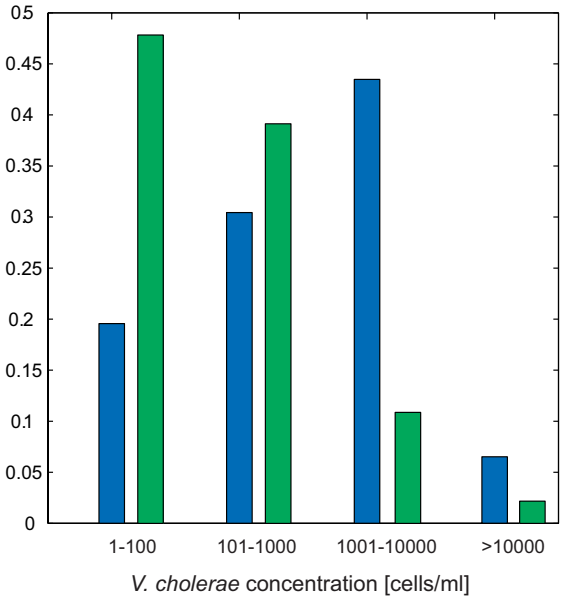

Figure 4.5: The distribution refers to the averages of the 3 replicate measurements performed for each location. Blue/green bars stand for *V.cholerae* O1/O139.

The association of *V.cholerae* O1 and O139 with other organisms has been a topic

of great importance in the study of the ecology of these bacteria, especially with respect to survival strategies in inter-epidemic periods (Tamplin et al., 1990; Islam et al., 1994; de Magny et al., 2011). In Fig. 4.6 the relationship between the concentration of *V. cholerae* O1 and O139 and the total cell count in the ponds was analyzed. Even though the linear correlation between these quantities had low R^2 values ($R^2 = 0.04$ for both *V. cholerae* O1 and O139 serogroups), p-values showed that such correlation is significant with respect to the constant model at a 20% significance level ($p \simeq 0.19$ in both cases).

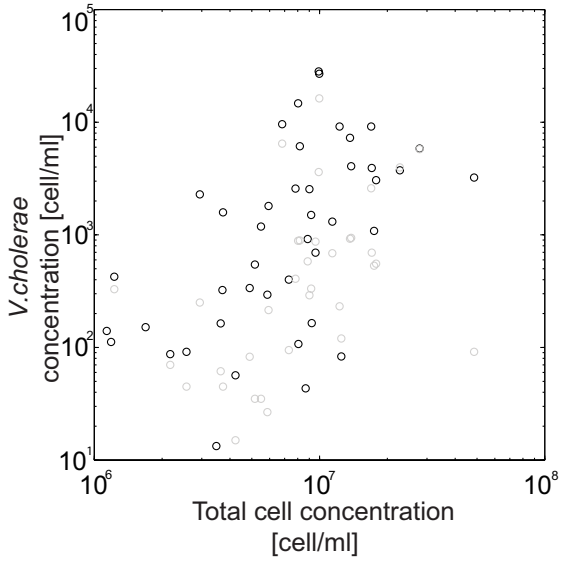

Figure 4.6: Log-log plot of *V.cholerae* serogroups O1 (black dots) and O139 (gray dots) (y-axis) compared to total cell counts (x-axis). Each dot represents the average concentration of the three replicative flow cytometry measurements from different pond-derived water samples.

Correlations between physicochemical parameters and *V. cholerae* O1 and O139 concentration were not significant for all tested parameters (Table 4.1)even though water temperatures were close to significance with respect to serogroup O1 abundance. This general lack of correlation is most likely a reflection of the limited time span of this study and future long-term investigations will shed further light on the association of *V. cholerae* with abiotic factors.

Spatial distribution of *V. cholerae* O1 and O139 concentrations showed instead a definite pattern (Fig. 4.7). The Matlab area is crossed by a branch of the river Meghna, which is embanked on one side to protect settlements close to the river. Our results indicated that ponds located in flood-protected areas host larger populations (with median concentrations of both O1 and O139 serogroups being an order of magnitude higher than in non-protected areas). A correlation between the risk of being infected with cholera and living in flood-controlled areas has previously been described for

Table 4.1: Observed physico-chemical parameters in ponds and their statistical significance with respect to ANOVA N-way analysis on correlation to *V.cholerae* O1 concentrations

Parameter	**Median value**	**Range**	**F**	**Prob>F**
Water temperature (° C)	29.3	$27.1 - 31.8$	3.42	0.072
Salinity (g/l)	0.1	$0 - 0.2$	0.38	0.539
Total Dissolved Solid (mg/l)	74.35	$23.60 - 240$	0.61	0.438
Conductivity (μ S/cm)	159.90	$50.70 - 498.00$	0.12	0.736
pH	6.88	$6.60 - 7.08$	0.04	0.839

the same study area, but no further explanations have been given so far (Emch, 1999). Our results suggest that specific and more favorable environmental conditions may support survival of *V. cholerae* in such areas thereby increasing the chance of contracting cholera.

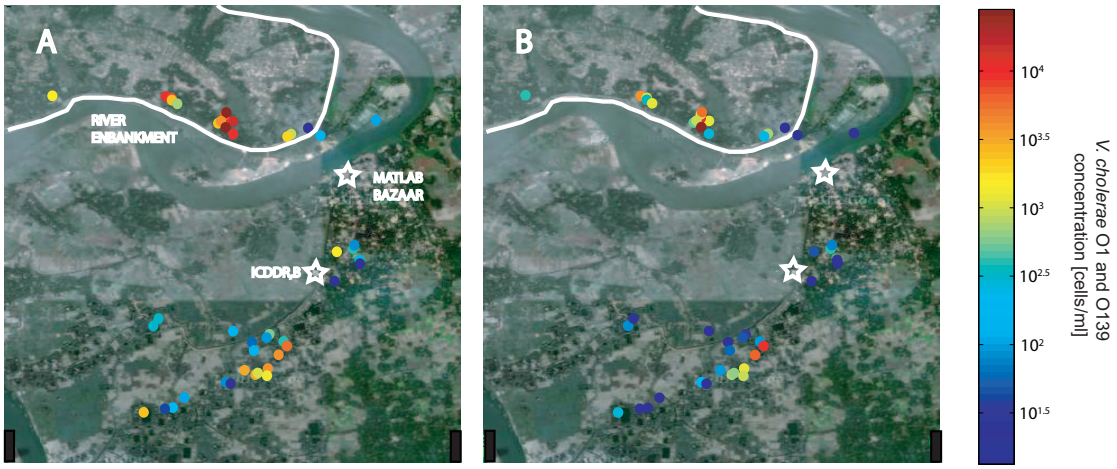

Figure 4.7: Map indicating the location of the sampled ponds and the relative concentration of *V. cholerae* serogroups O1 (panel A) and O139 (panel B). The ICDDR,B hospital, the regional market place and the outline of the river enbankment are highlighted in white. The heat map on the right indicates the concentration range.

Table 3: Comparison of the abundance of *V.cholerae* O1/O139 and the total cell count in ponds located in flood-protected and non-protected areas.

	Flood-controlled bank	**Non flood-controlled bank**
O1 conc. (cells/ml)	$3252.50(0-28420)$	$323.34(0-5837)$
O139 conc. (cells/ml)	$638.34(0-16330)$	$61.67(0-5765)$
Water temperature (° C)	$28.55(27.1-30.7)$	$29.4(28.4-31.8)$
Salinity (g/l)	$0.134(0-0.2)$	$0.063(0-0.2)$
Total Dissolved Solid (mg/l)	$111.2(25.8-240)$	$62.9(0-225)$
pH	$6.85(6.69-7.01)$	$6.98(6.60-7.08)$

Table 4.3 indicates that ponds specifically located in these regions show significantly different values of physicochemical parameters (e.g. higher salinity and increased total dissolved solid content). Flood protection might indeed prevent the washing out of the system by extreme events of flooding and thus promote the establishment of a thriving resident community of *V. cholerae*.

4.4 Conclusions

Based on the data presented here the following conclusions are worth mentioning:

- a procedure for the rapid assessment of *V. cholerae* O1 and O139 concentrations was tested under field setting conditions in diverse water bodies. This flow cytometry-based approach allowed a rapid test of several water samples within a short time period. Such data on pathogen abundance in water environments (including reservoirs) serves as a much-needed field validation of spatially explicit mathematical models of the spread of infections.
- The procedure proved effective in detecting the presence of *V. cholerae* O1 and O139 where standard techniques might fail, due to the VBNC state of *V. cholerae*.
- The method allowed measuring of *V. cholerae* O1 and O139 concentrations, rather than simply detecting the presence/absence of the pathogen, and is thus particularly suitable in epidemic management control and in long-term assessment of eradication efforts.

The applicability of this method remains limited, though, especially in large-scale studies and in emergencies, as more cost-effective detection methods do exist for *V. cholerae*, whose efficacy has been tested especially for diagnostic purposes (Dick et al., 2012). However, the method presented here in a field work situation allows measuring of the concentration, rather than simply detecting the presence of the pathogen, and may thus be more accurate to study the efficacy of eradication efforts. Also, the portability of the specific model of flow cytometer used in this work reduces greatly the costs associated with this kind of analysis, as a single instrument can be carried elsewhere and used for other field campaigns.

Although further efforts should be devoted especially to the establishment of reliable and fast procedures for the evaluation of the fraction of *V. cholerae* O1 and O139 harboring toxigenic genes, which are ultimately responsible for cholera outbreaks, it is concluded that the procedure tested here is a fundamental tool for the field validation of spatially explicit mathematical models of epidemic cholera. Such models require calibration of parameters describing the dynamics of Vibrio concentration over time and space. Thus, the importance of this work is to demonstrate that modern epidemiology of water-borne disease should be based on a balanced use of field work and mathematics.

5 Modeling human movement in cholera spreading along fluvial systems

5.1 Introduction

Despite the medical and socio-economic importance of the problem, modeling Cholera dynamics is a relatively recent endeavour, as described in the Introduction of this Thesis. A first mathematical approach dates back to thirty years ago, with a seminal work on the 1973 Mediterranean epidemic (Capasso and Paveri-Fontana, 1979). Notably, Codeço (2001) developed a model which considered three ordinary differential equations where, in addition to the compartments of susceptibles (S) and infectives (I) that characterize traditional epidemiological models, one equation accounts for the population dynamics of bacteria *Vibrio cholerae* (B) in the water reservoir. A key aspect concerns, in fact, the role of the water reservoir, because it has been shown that *V. cholerae* is a member of the surface water ecosystem where it may thrive and disperse in association with zooplankton and phytoplankton (Colwell et al., 1977; Islam et al., 1990, ; see also Chapter 2 of this Thesis). The Codeco model constitutes the spatially implicit basis of the reaction-drift-diffusion model that will be presented in the following.

Bertuzzo et al. (2008) were the first who developed a spatially explicit approach based on Codeco's model to account for epidemic spreading through a river network. In particular, the model proved reliable in describing both temporal and spatial evolutions of the disease for a significant case study concerning a South Africa outbreak occurred in the Kwazulu-Natal province in 2000. This epidemic lasted two years, causing more than $140,000$ confirmed Cholera cases. The model demonstrated that there does exist notable correlation between hydrologic controls and epidemiological dynamics.

This encouraging outcome suggested to Bertuzzo et al. (2010) to study the model behavior in more detail and in broader context. Therein, the speed of disease propagation along theoretical rivers, described by one- and two-dimensional lattices or more complex river networks, was evaluated both analytically and numerically via a

methodology which implied the formulation of a reaction-diffusion model brought to a spatially continuous form. The study postulates that bacterial diffusion is the spatial transport mechanism underlying the spreading of cholera in a given system. Of particular interest was the finding that the speed of a cholera travelling wave c in a substrate characterized by a uniform spatial distribution of reproductive numbers R_0 (defined as the ratio of the local population size to the threshold for epidemic outbreak), is given by the relationship $c \propto \sqrt{D(R_0 - 1)}$ where D is the bacterial diffusion coefficient. The comparison of basic timescales (that of epidemic propagation throughout the system and that intrinsic to the disease dynamics) then dictates whether or not spatially explicit schemes are needed to describe an outbreak (Bertuzzo et al., 2010).

To obtain the results shown in Bertuzzo et al. (2008, 2010), however, models were formulated where the spreading of the disease is only caused by the movement of the bacterial component. This implies that the equation of bacteria diffusion implicitly blends many relevant processes, among which hydrodynamic dispersion, symbiotic transport (e.g. via zooplankton) and human movements have to be cited. Here, such an assumption is relaxed by explicitly specifying different mechanisms of bacterial diffusion by proposing a more general formulation of the same model which explicitly prescribes different transport mechanisms. In particular, human movement can play a prominent role in the disease spreading, because most of the infected individuals are asymptomatic (King et al., 2008) and thus can move freely in the environment, transporting the vibrios far from where they were infected by excreting large numbers of *Vibrio cholerae* cells as part of their stool. To include this important process, the new model does explicitly account for the movement of human hosts by assuming that the equations of susceptibles and infectives include a diffusive component. In this way, the fact that not only free-living bacteria can move through the water network but also the bacteria transported by their hosts will be considered.

Needless to say, even the present formulation entails strong approximations for the description of movement of both the bacterial and the human component. In fact, describing human movement via a diffusive term implies that humans are supposed to spread homogeneously without any gravitational mechanism that would return individuals, eventually, to their original sites or other individual mobility patterns (Gonzalez et al., 2008). However, even this relatively simple way of accounting for human movement in the model leads to interesting and novel results. In particular, if the system is initialized with a few infected individuals in a given location of the river, two different wave fronts originate, one progressive and one regressive (see Figure 5.1 inset). If the speed of the regressive front is negative, the propagation is strictly upstream. When significant advection occurs, the upstream front can propagate more rapidly and the downstream front more slowly, with consequences described in detail herein.

5.2 The Model

The dynamics of Cholera are considered along the simplest fluvial ecological corridor. More precisely, the attention is limited here to a one-dimensional linear stream, because the (slower) propagation of travelling waves along branching backbones of complex networks can be recovered *a posteriori* (Campos and Mendez, 2005; Campos et al., 2006; Bertuzzo et al., 2008, 2010). Bertuzzo et al. (2010) formulated the following model:

$$\frac{\partial S}{\partial t} = \mu(1-S) - \frac{\beta B}{1+B}S \tag{5.1}$$

$$\frac{\partial I}{\partial t} = \frac{\beta B}{1+B}S - \eta I \tag{5.2}$$

$$\frac{\partial B}{\partial t} = -\mu_B B + \theta I - v\frac{\partial B}{\partial x} + D_B\frac{\partial^2 B}{\partial x^2} \tag{5.3}$$

where $S = S(x,t)$ and $I = I(x,t)$ are, respectively, the fractions of susceptible and infected individuals in the human population at location x and time t, and $B = B(x,t)$ is the associated concentration of *Vibrio cholerae* in the water reservoir. Note that the epidemiological and/or hydrological parameters are described in Table 5.1. Here the model is generalized, adding diffusive terms in the equations of susceptibles and infectives. It thus obtained:

$$\frac{\partial S}{\partial t} = \mu(1-S) - \frac{\beta B}{1+B}S + D_S\frac{\partial^2 S}{\partial x^2} \tag{5.4}$$

$$\frac{\partial I}{\partial t} = \frac{\beta B}{1+B}S - \eta I + D_I\frac{\partial^2 I}{\partial x^2} \tag{5.5}$$

$$\frac{\partial B}{\partial t} = -\mu_B B + \theta I - v\frac{\partial B}{\partial x} + D_B\frac{\partial^2 B}{\partial x^2} \tag{5.6}$$

In addition to the reaction process described via the Codeco's equations, in the (4-6) model the spatial displacement of human populations and bacteria along and within the river corridor is explicitly accounted for. In fact, both advection and diffusion processes are specified, along with their respective transport parameters v and D. Diffusion is assumed to occur in all compartments, while advection affects the bacterial compartment only. In fact, it is hypothesized that human movement is practically

Symbol	Description
μ	population natality and mortality rate (day^{-1})
β	rate of exposure to contaminated water (day^{-1})
K	concentration of *V. cholerae* in water that yields 50% chance of being infected with Cholera ($cells/m^3$)
p	per capita, per cubic meter excretion rate of humans ($cells\,infected^{-1}m^{-3}$)
H	abundance of local human population
W	volume of water reservoir (m^3)
θ	$\theta = \frac{pH}{KW}$; rate of production by infected individuals of *V. cholerae* that reach the water body ($day^{-1}\cdot$)
η	rate of Infected abduction; it represents the sum of natural mortality, disease induced mortality (virulence) and recovery
μ_B	mortality of *V.cholerae* in the aquatic environment (day^{-1})
v	advection or "drift" velocity ($m\,day^{-1}$)
$D_{S/I}$	diffusion coefficient of humans ($m^2 day^{-1}$)
D_B	diffusion coefficient of bacteria ($m^2 day^{-1}$)
c	Cholera spreading velocity (travelling wave solution of models (1)-(3) and (4)-(6)) ($m\,day^{-1}$)

Table 5.1: Description of the symbols used in the text

unaffected by the direction of river flow because humans can also move on land. This stands as a substantial change with respect to model (1-3), where human movement term is not considered explicitly and all transport mechanisms are condensed into the single equation of bacterial mass balance. As pointed out in the introduction, the diffusion coefficient associated to the bacterial movement in model (1-3) implicitly accounted for all transport mechanisms, from hydrodynamic turbulence to human movement. Under typical conditions, it is observed that the values to be attributed to human diffusion coefficients are orders of magnitude larger than that of bacterial transport. While other modes of human displacement may be of interest e.g. via some observational understanding of human mobility patterns (Gonzalez et al., 2008), it is now shown that interesting patterns emerge even in this simplified scheme.

5.3 Cholera Travelling Wave Fronts

The velocities (both progressive and regressive) with which the disease spreads from where the infection first occurs are pursued. To this end, traveling wave solutions to the equations (4-6) are to be found. Using a standard procedure (Murray, 2002), the

variable $u = x - ct$ is introduced, with c being the speed of the disease travelling front. The original PDE model (4-6) can thus be transformed to the following 6-th order ODE system:

$$\frac{\partial S}{\partial u} = Z \tag{5.7}$$
$$\frac{\partial Z}{\partial u} = \frac{1}{D_S}\left[-cZ - \mu(1-S) + \beta\frac{B}{1+B}S\right] \tag{5.8}$$
$$\frac{\partial I}{\partial u} = W \tag{5.9}$$
$$\frac{\partial W}{\partial u} = \frac{1}{D_I}\left[-cW - \beta\frac{B}{1+B}S + \eta I\right] \tag{5.10}$$
$$\frac{\partial B}{\partial u} = Y \tag{5.11}$$
$$\frac{\partial Y}{\partial u} = \frac{1}{D_B}\left[-cY + \mu_B B - \theta I + vY\right] \tag{5.12}$$

Model (7)-(12) has two equilibria, namely $\overline{X}_0 = [1\,0\,0\,0\,0\,0]^T$ and $\overline{X}_+ = [\overline{S}_+\,0\,\overline{I}_+\,0\,\overline{B}_+\,0]^T$. These equilibria represent two space homogeneous steady state solutions of the original PDE. In particular, $\overline{X}_0$ corresponds to the disease-free situation in which only susceptible individuals are present in the human population. $\overline{X}_+$ is the endemic equilibrium, where a constant fraction $\overline{I}_+$ of individuals is infected by Cholera. The latter equilibrium is non negative, thus epidemiologically feasible, as long as $\beta\,\theta > \eta\mu_B$, i.e. if $R_0 = \beta\theta/\eta\mu_B > 1$ (where R_0 is the reproduction number of the disease (Anderson and May, 1992)). A progressive [regressive] wave front of the original model (4-6) can be seen as a heteroclinic orbit in model (7)-(12) that connects $\overline{X}_+$ to $\overline{X}_0$ [$\overline{X}_0$ to $\overline{X}_+$]. Since $\overline{X}_0$ is placed on one of the state-space axes, any orbits that would involve oscillations around it would break the constraint of $S(u)$, $I(u)$ and $B(u)$ being non negative. It is therefore crucial for a traveling wave to exist that the linearized system (7)-(12) around $\overline{X}_0$ has the the eigenvalue(s) closest to the imaginary axis with no imaginary part. If $\overline{X}_0$ is a hyperbolic equilibrium, in fact, the heteroclinic connection under study will be tangent to the eigenvectors associated to those eigenvalues while converging to (progressive wave) or departing from (regressive wave) $\overline{X}_0$. In particular, the values of c which are being sought are then those that lead to the transition of such eigenvalues from complex to real form. It is found that there is a range of values $c_{reg} < c < c_{prog}$ for which the eigenvalues are complex, thus propagating fronts with these speeds for the disease cannot exist. It has been shown (Aronson (1978); in the epidemiological context see Lin et al. (2003)) that, among all infinite possible waves propagating at $c \geq c_{prog}$, the selected velocity for large t is exactly c_{prog}. In the same way, among all the potential speeds $c \leq c_{reg}$, the one that establishes is actually

$c = c_{reg}$. The Jacobian matrix of model (7)-(12) at $\overline{X} = [\overline{S}\,\overline{Z}\,\overline{I}\,\overline{W}\,\overline{B}\,\overline{Y}]^T$ is given by:

$$J\left(\bar{X}\right) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{\mu}{D_S} + \frac{\beta}{D_S}\frac{\overline{B}}{1+\overline{B}} & -\frac{c}{D_S} & 0 & 0 & \frac{\beta}{D_S}\frac{\overline{S}}{(1+\overline{B})^2} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{\beta}{D_I}\frac{\overline{B}}{1+\overline{B}} & 0 & \frac{\eta}{D_I} & -\frac{c}{D_I} & -\frac{\beta}{D_I}\frac{\overline{S}}{(1+\overline{B})^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{\theta}{D_B} & 0 & \frac{\mu_B}{D_B} & \frac{v-c}{D_B} \end{bmatrix} \tag{5.13}$$

At $\overline{X}_0$ the Jacobian is a block upper triangular matrix , whose diagonal blocks are:

$$J_1\left(\bar{X}_0\right) = \begin{bmatrix} 0 & 1 \\ \frac{\mu}{D_S} & -\frac{c}{D_S} \end{bmatrix} \tag{5.14}$$

$$J_2\left(\bar{X}_0\right) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{\eta}{D_I} & -\frac{c}{D_I} & -\frac{\beta}{D_I} & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{\theta}{D_B} & 0 & \frac{\mu_B}{D_B} & \frac{v-c}{D_B} \end{bmatrix} \tag{5.15}$$

The determinant of $J_1\left(\bar{X}_0\right)$ is negative (i.e., $\overline{X}_0$ is a saddle) and the two eigenvalues are real for all c's. Therefore, no conditions on c emerge from block $J_1\left(\bar{X}_0\right)$. This result implies that the diffusion of Susceptibles is not relevant to the propagation speed of Cholera, because the parameter D_S appears only in block $J_1\left(\bar{X}_0\right)$.

The characteristic polynomial of $J_2\left(\bar{X}_0\right)$ is:

$$\lambda^4 + b_1\lambda^3 + b_2\lambda^2 + b_3\lambda + b_4, \tag{5.16}$$

where $b_1 = \frac{c}{D_I} + \frac{c-v}{D_B}$, $b_2 = \frac{c(c-v)}{D_I D_B} - \frac{\eta}{D_I} - \frac{\mu_B}{D_B}$, $b_3 = \frac{\eta(v-c)-c\mu_B}{D_I D_B}$ and $b_4 = \frac{\eta\mu_B - \beta\theta}{D_I D_B}$. It is to be noticed that b_4 is independent of c and always negative if $R_0 > 1$. This implies that the quartic equation has always at least one positive, real root. In order to find the possible values of c for which the characteristic polynomial has two real coinciding solutions, the method proposed in Jury and Mansour (1981) is followed. They found out that the characteristic polynomial (5.16) has two real coinciding solutions whenever

$$\Delta = 4M^3 - L^2 = 0 \tag{5.17}$$

$$
\begin{aligned}
M &= b_2^2 + 12b_4 - 3b_1b_3 \\
L &= 72b_2b_4 + 9b_1b_2b_3 - 2b_2^3 - 27b_3^2 - 27b_4b_1^2
\end{aligned}
$$

In the present case, relation (5.17) turns out to be an algebraic equation of the 12-th order in c. As shown in the following section, for every fixed set of parameters (but c) within epidemiologically and hydrologically meaningful ranges, two values of c are found, say c_{reg} and c_{prog}, such that at $c = c_{prog}$ [$c = c_{reg}$] the largest negative [smallest positive] solutions of (5.16) are double roots.

5.4 Results

As stated above, the main goal is to determine the effects of the explicit accounting for human movement in a framework of Cholera propagation along river corridors. To this end, all parameters are fixed - except those related to space - to the values used in Bertuzzo et al. (2008) for the Kwa Zulu - Natal epidemics. To compare the results of the present model with those of Bertuzzo et al. (2010), a diffusion coefficient for bacteria that is a fraction of the Infected one, is considered, i.e.:

$$D_B = \alpha D_I, \tag{5.18}$$

To focus on the effects of human movement, it shall be initially assumed $\alpha = 0.001$, that is $D_I \gg D_B$. This assumption reflects the case where bacterial movement is dominated by human transport. The analytical considerations outlined in the previous section are used to determine how the speed of Cholera propagation is influenced by D_I, i.e. $c = c_I(D_I)$. To facilitate comparisons with Bertuzzo et al. (2010), the model (1-3) is used and the methods therein with their $D_B = D_I$, in order to determine a speed of Cholera propagation that is named $c_B(D)$. The case of a purely diffusive process (i.e. $v = 0$ [km d^{-1}]) shall be considered first. For each value of D, it is obviously found in both the present model and in model (1-3) two c's, having the same absolute value but opposite sign. As previously stated, this stands for the existence of two fronts, progressing in opposite directions, with a speed having the same absolute value but opposite sign. As previously stated, this stands for the existence of two fronts, progressing in opposite directions, with speeds ($c_{prog} > 0$ and $c_{reg} = -c_{prog} < 0$; see Figure 5.1 inset). Figure 5.1 shows that, consistently with previous findings (Bertuzzo et al., 2010), in the absence of advection, a scaling of the form $c \propto \sqrt{D}$ is recovered. It is important to notice that only a slight difference between the two models emerges, in absence of advection (Figure 5.2). Actually, it turns out that

just a proportionality constant allows to map one model into the other, i.e. $c_B = p\, c_I$, with $p \sim 1$, provided the value of D_B in model (1-3) is set equal to that of D_I in model (4-6). Thus all results obtained in Bertuzzo et al. (2010) hold also for model (4-6) in the case of negligible advection, namely when the river flows slowly.

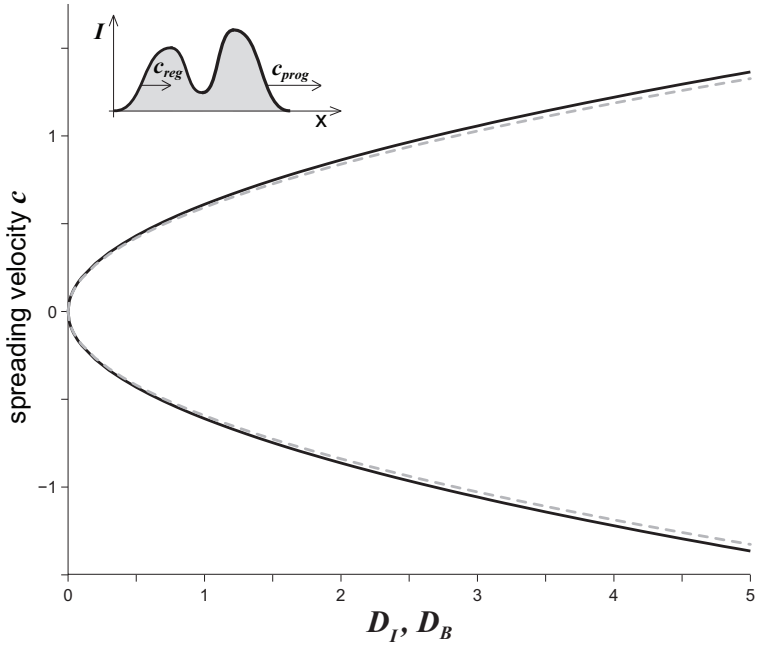

Figure 5.1: Comparison of disease spreading velocity c as a function of the diffusion coefficient D obtained with model (4)-(6) ($D_I = D = 10^3 D_B$; solid line) and model (1)-(3) (dashed line) in absence of advection ($v = 0$ [km d^{-1}]). All other parameters as in Bertuzzo et al. (2010).

Interestingly, however, when advection is not negligible (e.g. assuming $v = 2$ [km d^{-1}]; see Figure 5.2) the results obtained via model (1-3) can differ significantly from those of (4-6). As it appears from Figure 5.2, in fact, for every fixed value of the diffusion coefficient, the speed of propagation of the disease is systematically higher in model (1-3) than in the present model. This result might seem paradoxical at a first glance, but it is not. In fact, while model (1-3) really focuses on a diffusive-advective process, in the sense that diffusion acts on the top of advection within the river domain, here diffusion in a compartment (I) that is not subject to the advection is mainly considered, which in turn characterizes the main transport process for bacteria in water.

An important consequence of these results concerns the regressive wave. In both the model presented (solid curve) and model (1-3) (dashed curve) threshold values are found for the diffusion coefficient D^*_{reg} above which the regressive wave exhibits negative speeds. Such threshold values are significant because they mark the condition for the disease to propagate upstream, against the river flow. In fact, *Vibrio*

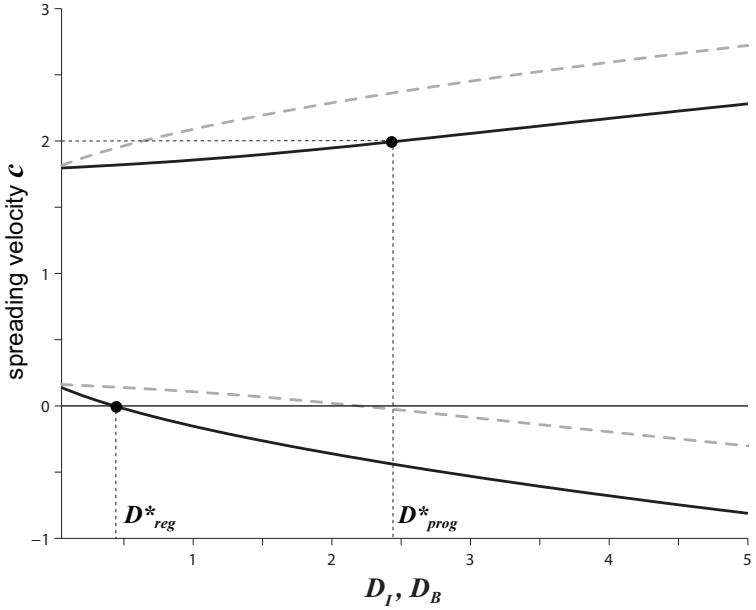

Figure 5.2: Comparison of disease spreading velocity c as a function of the diffusion coefficient D obtained with the present model ($D_I = D = 10^3 D_B$ and model (1)-(3) in presence of advection ($v = 2$ [km d^{-1}]). All other parameters as in Bertuzzo et al. (2010).

cholerae being an autochtonous inhabitant of coastal environments, it is often the case that infections start from coastal environments and progressively reach inland regions. The value of D^*_{reg} turns out to be much lower for the present model than for model (1-3). More quantitatively, for $v = 2$ [km d^{-1}] (Figure 5.2), it is found that $D^*_{reg} \sim 2.2$ [km^{2}d^{-1}], with model (1-3), while it amounts to $D^*_{reg} = 0.431$ [km^{2}d^{-1}] with the present model. This may have profound consequences on epidemiological control policies. In fact, in a diffusion process, the average distance travelled by a particle at time t scales as $\sqrt{Dt}$. As a consequence, if the human movement is assumed to occur within the river as in equations (1-3), then bacteria should move approximately 1.48 km during one day, in order to originate a regressive front spreading upstream. In contrast, infected individuals must travel in the average a distance of 0.66 [km day^{-1}] upstream to let the epidemics spread against the river flow. This second spatial scale seems much more realistic. It demonstrates that the movement of infected individuals can greatly influence the spreading of the disease. This result might sustain control policies that aim at using quarantine during a cholera epidemic in order to avoid its spreading towards inland regions. However, it should not be forgotten that a large portion of the infections is asymptomatic and thus it may be difficult to intervene drastically to limit the movement of infected individuals.

Hydrological control of water-borne diseases also calls for a better understanding of

the relationship between the river velocity and the disease propagation speed. If it is assumed that the drift speed v is representative of the river velocity, it is possible to estimate, for example, the value of D that implies $c > v$. Defining it as D^*_{prog} (see again Figure 5.2), this threshold for the diffusion coefficient is much higher in the present model than it is in model (1-3). Consequently, the downstream propagation of the disease is also much slower. Given their potential interest for the new point of disease control, it is important to understand how the quantities D^*_{prog} and D^*_{reg} (altogether named D^*) scale with hydrological and epidemiological parameters. Noticeably, it is found out that the square root of both thresholds scales linearly with the advection velocity v.

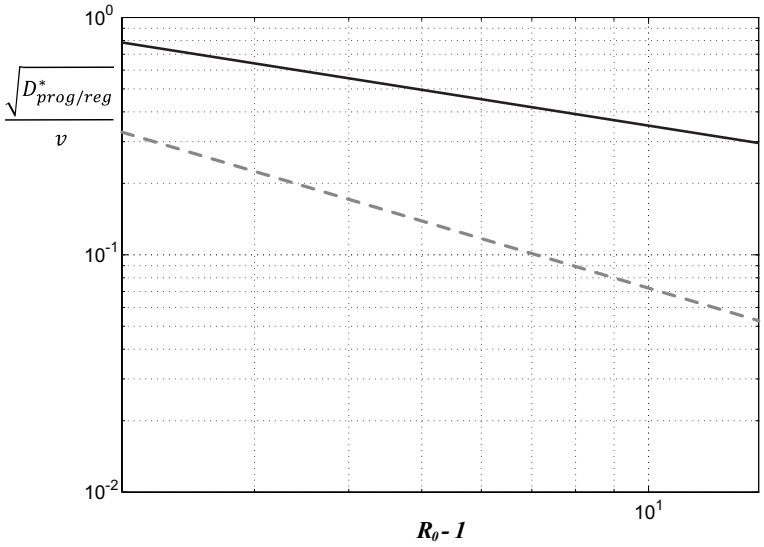

Figure 5.3: Log-log plot showing the relationship between $\frac{\sqrt{D^*}}{v}$ and $R_0 - 1$ (loglog scale). The solid black line represents the progressive wave threshold (D^*_{prog}), while the dashed gray line portrays the regressive wave (D^*_{reg}). All unspecified parameters as in Fig. 5.1

This is non-trivial because, as shown in Figure 5.2, the relationship between c and $\sqrt{D}$ is nonlinear for $v \neq 0$. It is thus interesting to investigate how the $\frac{\sqrt{D^*}}{v}$ ratios depend upon the reproduction number of the disease or, more precisely, $R_0 - 1$. It turns out that the ratios scale with $R_0 - 1$ (Figure 5.3) following different power laws. In fact it is obtained:

$$\frac{\sqrt{D^*_{prog}}}{v} \propto (R_0 - 1)^{-0.5}$$
$$\frac{\sqrt{D^*_{reg}}}{v} \propto (R_0 - 1)^{-0.9}$$

The results obtained up to now might be influenced by the value of bacterial diffusion

hypothesized for model (4-6). To investigate how bacterial diffusion within the river may affect the spreading of the disease, the analysis has been repeated by changing the ratio between D_I and D_B (see eq.5.18). For example, the outcome with $\alpha = 0.1$ is shown in Figure 5.4. In this case, it is found that D_B has a pronounced effect on the exponent of the power law of the progressive wave speed, lowering it to approximately -0.63, while it has almost no effect on the regressive wave speed. Therefore it is suggested that bacterial diffusion due to water turbulence and symbiotic transport has very little impact on the upstream spreading of the disease, at least in the framework of the present model.

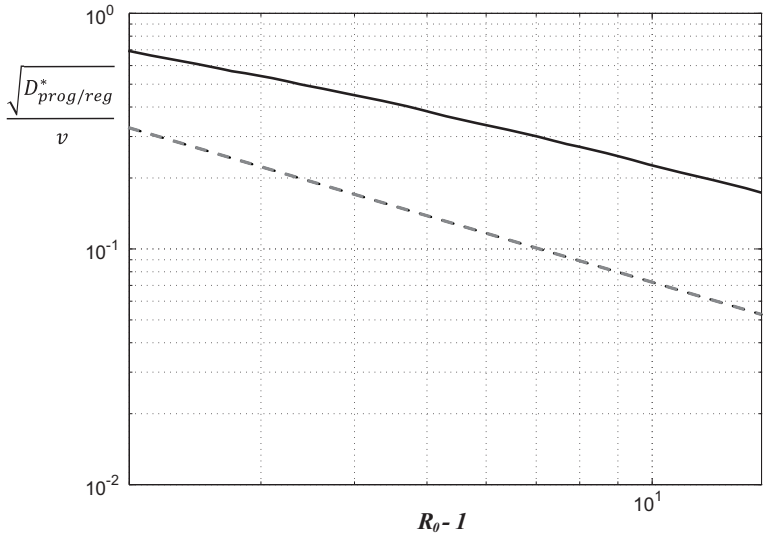

Figure 5.4: As Figure 5.3 but with $D_B = 0.1 D_I$.

5.5 Conclusions

The following conclusions are worth mentioning:

- A new spatially explicit model of Cholera spreading along river corridors has been proposed and analyzed to determine disease wave fronts. The model extends previous approaches in that it adds a diffusion term in both the susceptible and the infected components. It was shown that the displacement of susceptibles does not have any influence on the main features of disease spreading. The model presented in this Chapter suggests instead that the diffusion of Infected individuals may have quite an important epidemiological impact;
- A comparison with previous spatially explicit models (where humans are immobile) has been carried out in detail in the case of a purely diffusive process (no advection or $v = 0$) and of biased transport ($v > 0$) for bacteria. In the first case, the two models predict very similar spreading velocities for the disease. Instead, when the drift velocity v is positive, the progressive wave exhibits lower celerity but a faster regressive wave is observed. Thus if infected humans are allowed to move along the river, the predicted epidemics runs faster upstream but slower downstream;
- It is also found that there exist scaling laws between some of the key model parameters and both the advection drift v and the disease basic reproduction number R_0. In particular, the ratio between the minimum value of infected diffusion coefficient required to obtain a regressive wave (D^*_{reg}) and v scales approximately as the inverse of $R_0 - 1$. For plausible values of the basic reproductive rate for cholera, which is tipically found to vary between 3 and 15, it is found that $\frac{\sqrt{D^*_{reg}}}{v}$ ranges approximately between 0.3 and 0.05. If the advection speed v is identified with river velocity, this implies that the minimum distance to be traveled by infected people in one time unit to propagate cholera upstream is in the range 5-30% of the distance traveled downstream by river water in one time unit.

6 Rainfall mediations in the spreading of epidemic cholera

6.1 Introduction

The recent, still ongoing cholera outbreak that has struck Haiti has brought to broad public attention the loss of human lives and the social and economic disruption caused, even to date, by epidemics of the disease. The global relevance of the problem and the need for preventive assessment and control of cholera spreading is manifest also in view of other emerging outbreaks registered in 2011-2012 in the Congo river basin, Cuba, Sierra Leon and the Sahel region (Luque Fernandez et al., 2009; Kelvin, 2011; Nkoko et al., 2011; Al-Tawfiq and Memish, 2012).

While the role of weather conditions, and rainfall in particular, on patterns of waterborne infections have long been studied especially in empirical frameworks (Altizer et al., 2006; de Magny et al., 2008; Emch et al., 2008; Koelle, 2009; Lipp et al., 2002; Pascual et al., 2000), hydrologically-driven, spatially explicit mathematical models of cholera epidemics have only recently been developed (Bertuzzo et al., 2008, 2010). They have been applied to study the course of the Haitian epidemic, starting from the very first months of its insurgence in late 2010 (Bertuzzo et al., 2011; Tuite et al., 2011; Chao et al., 2011), and following disease resurgence occurred in May 2011 in connection with unusually intense tropical rains (Rinaldo et al., 2012). Even though concerns for correct surveillance, monitoring and intervention planning have been on the rise in international institutions debate, regarding cholera in particular (e.g. WHO, 2011), few of these models have been utilized to date to test their effectiveness as predictive and control tools. Such models could be in principle applied, for instance, to deploy medical staff and life-saving supplies through projections of the patterns of cholera infections, and to implement pro-active rather than reactive policies as commonplace in epidemiological control strategies. The increasing number of spatially explicit mathematical models suggest that their impact on public health practice is gaining momentum.

In this work all the information gathered in the studies illustrated in previous Chapters has been used to build up a detailed framework to simulate the course of the epidemic in Haiti, taking into account several mechanisms that control the spreading of the disease in the country (i.e. human mobility fluxes coupled with hydrological transport) and that have caused the reappearance of the disease after its first onset (waning of the immunity and increase of water contamination due to rainfall runoff).

The Haitian epidemic represents more than just another test case. In fact, cholera had never been reported in Haiti before 2010 and therefore it is likely that the population had no significant prior exposure or acquired immunity to the disease, suggesting that the entire population was susceptible to infection. Moreover, once a cholera epidemic starts, infected patients excrete huge numbers of *Vibrio cholerae* bacteria which spread either through water pathways (via active and passive dispersal; Bertuzzo et al., 2008, 2010; Chao et al., 2011; Righetto et al., 2011; Mari et al., 2012a) or through human mobility networks involving both susceptibles and infected individuals (Tuite et al., 2011; Chao et al., 2011; Mari et al., 2012b). Thus the poor sanitation conditions experienced especially after the disastrous 2010 earthquake facilitated both types of spread and fostered the abundance of microorganisms in the water system, thus rendering the Haitian outbreak exemplary (Rinaldo et al., 2012).

The first assessment of the disease made by the group (Bertuzzo et al., 2011) described the spreading of the disease by a simple diffusion of vibrios among neighboring nodes and was used to simulate the first, more virulent months of the disease. Panel B of Fig. 6.1 compares the projected course of the epidemic as published in January 2011 with the actual reported case counts reported at the end of September 2011 (PAHO, 2010). Highlighted (dark gray) in Fig. 6.1 (panel A) is the dataset used for calibration, which is limited to the end of December 2010 before the first decline of the incidence of the disease. The original prediction (Bertuzzo et al., 2011) ran up to the end of May 2011 (solid line in Fig. 6.1A). The 5-mo forecast, judged in retrospect, was quite robust and could have been used to make practical decisions and act in time. To facilitate a further assessment of our model reliability, the original prediction is here extended to the end of September 2011 (Fig. 6.1A, dashed line). Whereas the order of magnitude of total cumulated infections is captured up to September 2011, important features are clearly missed, such as the June–July revamping of weekly incidence (which is likely correlated to seasonal rainfall, Fig. 6.1, panel A, *Top*).

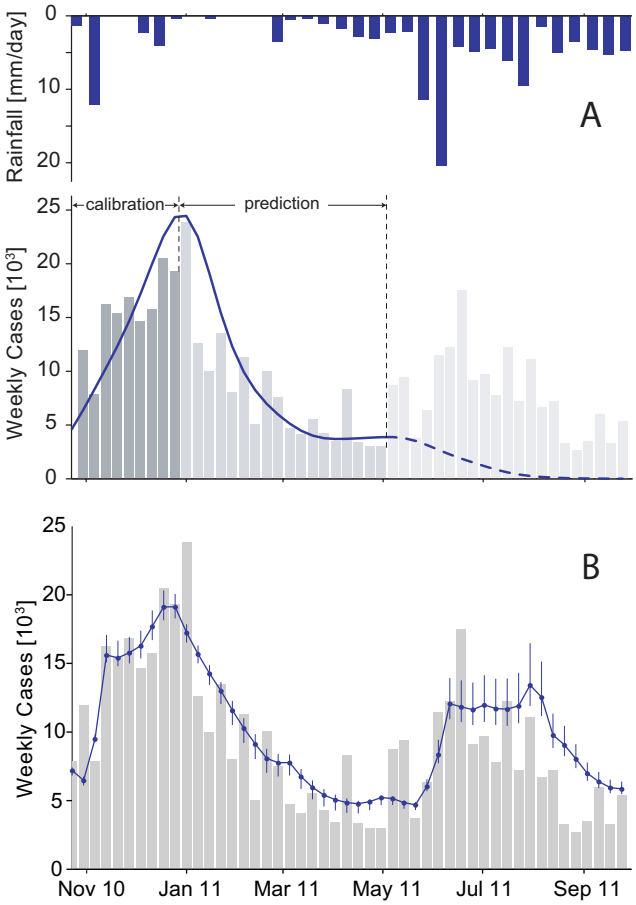

Figure 6.1: Panel A: (*Top*) Daily decadal rainfall intensity, averaged over the entire Haiti region. (*Bottom*) Weekly reported cases (gray bars Piarroux et al., 2011) compared with the simulated incidence pattern (solid line) computed by the model in Bertuzzo et al. (2011). Data from each department were collected until September 30, 2011. The calibration dataset (dark gray) was limited to the total reported cases available until December 2010. The solid line shows the published early prediction (Bertuzzo et al., 2011) that was run until the end of May 2011. To facilitate the assessment, the original prediction is extended to the end of September 2011 (dashed line). Panel B: Evolution of reported new weekly cases (gray bars) along with the simulated incidence pattern of the revised model (solid line Rinaldo et al., 2012). Error bars highlight the range of uncertainty due to parameter estimation.

In agreement with this result, the group, in an effort which has been part of the work of this Thesis, has developed a new family of models to reproduce the observed resurgence of cholera in Haiti, in concomitance with the rainfall season, calibrating them from the onset of the disease (October 2010) until September 2011. Panel B of fig. 6.1 shows the results of the best-ranking model and its estimation uncertainty. Because the analysis was then extended to a 1-y time span, the loss of acquired immunity (i.e., a flux from the pool of recovered cases back to the pool of susceptible individuals) was accounted for by all models compared. There, the metacommunity model of transmission among human settlements exposed to the infection was also revised by including specific mechanisms of hydrological transport, as suggested by empirical observation of the downstream spreading of early infections along the Artibonite River (Piarroux et al., 2011). Pathogen dispersal along waterways is there described (i) by a careful extraction of the river networks (Fig. 6.2) from digital terrain maps (DTMs), through suitable geomorphologic criteria, and (ii) as a biased random walk process on an oriented graph. Pathogen redistribution is also enhanced by contamination of the water reservoir driven by heavy seasonal rains. Moreover, human mobility patterns are explicitly modeled. It is assumed that susceptible and infected individuals engage in short-term trips from the communities where they live toward other settlements. While traveling or commuting, susceptible individuals can be exposed to pathogens and return as infected carriers to their settlement (10, 20). Similarly, infected hosts can disseminate the disease away from their home community—in many cases infected individuals are asymptomatic and thus are not barred from their usual activities. Fluxes of human mobility are described by gravity-like models where the flux between two communities owing to human mobility is proportional to the product of the respective populations and decays with the distance separating them (Erlander and Stewart, 1990).

The Haiti epidemic also provided direct and compelling evidence relating cholera resurgence to environmental drivers, specifically to rainfall. Little insight could be gained, in fact, from past empirical studies correlating rainfall to cholera cases because most, if not all, previous studies were carried out in contexts where cholera is endemic (see e.g. Lipp et al., 2002; Emch et al., 2008). In fact, reported correlations between rainfall events and resurgences – both in their sign and time lag – have been rather disparate (Luque Fernandez et al., 2009; Hashizume et al., 2010; Ruiz-Moreno et al., 2007; Akanda et al., 2009) reflecting the range of potential mechanisms through which rainfall may affect increased exposure to risk of infections (e.g. crowding effects due to flooding, raw sewage contamination of water sources, increased availability of compounds boosting *V. cholerae* survival or toxins, increased contamination due to over-exploitation of the water reservoirs, to name a few). Rinaldo et al. (2012) have shown how such correlation could be implemented in epidemiological models by forcing the contamination of the local water reservoir through rainfall-runoff transfer of *V. cholerae* from waste- to drinking-water. The model presented in Rinaldo

et al. (2012) – or, rather, family of models – works in a spatially explicit framework, in which Haiti is depicted as a network of human communities (the nodes) connected by both hydrology and human mobility (the edges). Each community is represented by an ODE system, in which the gravitating population is divided into Susceptible (S), Infected (I) and Recovered (R) individuals. In each model, the evolution of the concentration of *V. cholerae* in the environmental water reservoir is considered. Here such approach is further extended, generating scenarios of precipitation which are utilized to perform epidemiological predictions and to evaluate *a priori* the impact of intervention policies.

Unlike cholera, rainfall predictions are an established endeavour (Rodriguez-Iturbe et al., 1986) and rainfall stochastic generators have recently been widely considered for studying precipitation patterns (Laio et al., 2001), also in the light of the inclusion of a description of interannual variability (Porporato et al., 2006). Here a Poisson generator is used, such that it takes into account both the inter-annual and the spatial variabilities of rainfall intensity in order to preserve space/time correlations while generating rainfall at local scale. The identification of statistically equivalent spatio-temporal aggregates is carried out using suitable clustering techniques. Such an approach allows to generate a large number of rainfall scenarios, naturally preserving the statistical properties of the hydrologic data.

These synthetic rainfall fields are used to force the epidemiological models and to obtain, as a result, estimates of the strength of the disease resurgence. Here, two types of analysis are performed: i) an *ex post* evaluation, in which calibration, validation and prediction all belong to the past course of the outbreak. Under such premises, this approach simulates known recrudescence of the disease; and ii) multi-seasonal projections, from the current state of the epidemic to the next few years in which cholera is speculated to become endemic in the region. The first analysis simulates real-time conditions in which short-term (a few month) scenarios of cholera resurgence are used to evaluate the performance of the model as a predictive and control tool during the very course of an epidemic. Then, the effect of different, alternative scenarios of intervention (sanitation and vaccination, possibly differing in timing and in spatial distribution) on the evolution of the outbreak is analyzed, to mimic model-guided intervention policies. It is studied whether the inference of the most effective policy – say, that aiming at the maximum reduction of the total number of reported cases in a given time frame – may still hold in the face of the actual development of disease resurgence. In the long-term case, the study of correlations of cyclic resurgence of the disease to the seasonal rainfall cycle matters, as the particular initial conditions that have favored the appearance of cholera in Haiti — i.e. a high number of susceptibles -– will no longer apply in the future. The disease, in fact, can be expected to reappear in particular conditions of stress with an intensity that depends on the rate at which recovered individuals lose their temporary acquired immunity to the disease. This kind of analysis allows also to estimate the amount of sanitation or the extent of a

vaccination campaign aimed at eradicating the disease from the region, and is deeply rooted in hydrologic sciences.

6.2 Spatially explicit epidemiological models for the Haitian epidemic

Here some of the models presented in Rinaldo et al. (2012) are used, who have constructed a spatially explicit framework for the description of the Haitian epidemic and whose approach evolved from the first Haiti application by Bertuzzo et al. (2011). In particular, this analysis is restricted to the two models which emerged as best performing under absorbing or diffusive boundary conditions (Rinaldo et al., 2012, see Figure $S8$ therein). They consider n communities ($i = 1, n$) spatially distributed within a given domain that embeds the hydrologic and the human mobility networks (see Figure 6.1, panel A for details on the spatial databases considered and panel B for a graph depicting the local model and the mobility networks). Let $S_i(t)$, $I_i(t)$ and $R_i(t)$ be the local abundances of susceptible, infected and recovered individuals in each node i of the river network at time t, and let $B_i(t)$ be the concentration of *V. cholerae* in the water reservoir at site i.

6.2.1 Basic model and dynamics

Epidemiological dynamics, pathogen transport and human mobility can be described by the following set of coupled differential equations, which includes most of the mechanisms common to the models and represents the simplest of the two models here considered:

$$\begin{aligned}
\frac{dS_i}{dt} &= \mu(H_i - S_i) - \mathcal{F}_i(t)S_i + \rho R_i \\
\frac{dI_i}{dt} &= \mathcal{F}_i(t)S_i - (\gamma + \mu + \delta)I_i \\
\frac{dR_i}{dt} &= \gamma I_i - (\rho + \mu)R_i \\
\frac{dB_i}{dt} &= -\mu_B B_i - l\left(B_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}B_j\right) + \\
&+ \frac{p_i}{W_i}\left[1 + \phi J_i(t)\right]\mathcal{G}_i(t)\,.
\end{aligned} \tag{6.1}$$

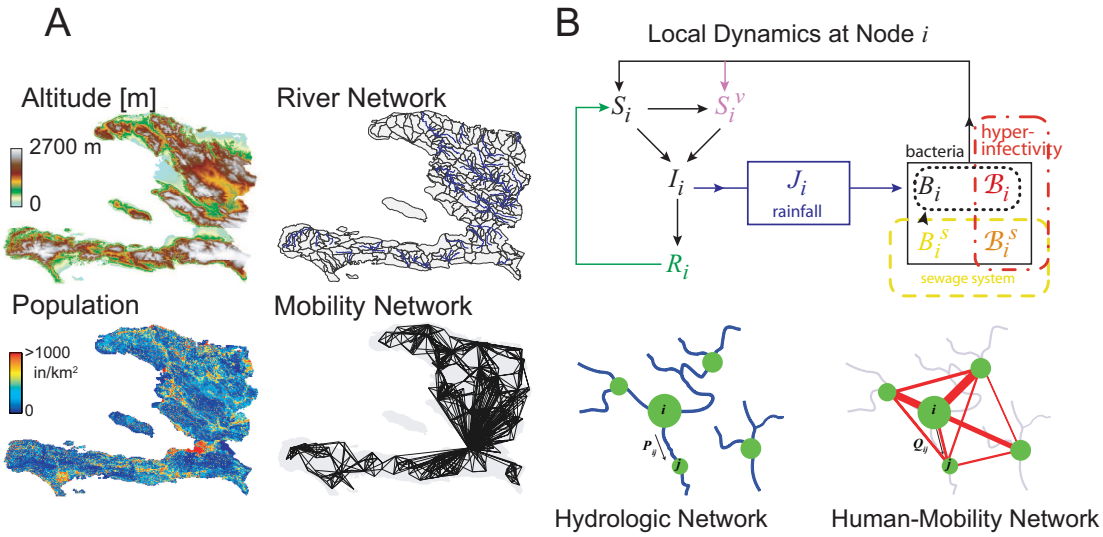

Figure 6.2: Panel A: Spatial databases. In clock-wise order: color-coded Digital Terrain elevation Map (DTM) of Haiti; the subdivision of Haitian territory in hydrological units (sub-basins) extracted from the DTM; a relevant subset of the network of human mobility, here portrayed synthetically by the four largest outbound connections for each node; spatial distribution of population density obtained by LandScan remote sensing (Rinaldo et al., 2012), which is translated into a georeferenced spatial distribution of nodes i endowed with population H_i. Panel B: Elements for a revised model. Above: block diagram of the i-th site epidemiological model (model 6.1); below: schematic representations of hydrologic (left) and human mobility (right) networks. Human communities of different sizes are assumed to be concentrated in nodes (green dots). Connections between nodes can be due to either hydrological pathways (i.e. river branches; left), which determines active and passive fluvial transport of pathogens, or human displacement (right). Figures from Rinaldo et al. (2012).

The host population is assumed to be at a demographic equilibrium, where μ is the human mortality rate, H_i is the population size of the local community and μH_i a constant recruitment. Loss of acquired immunity is also accounted for, so that recovered individuals $R_i(t)$ are returned to the compartment of susceptibles at rate ρ. Because of human mobility, infection of susceptible hosts in a given node regards all susceptibles that do not move from their node (a fraction $1-m$ of the susceptible population; note that the parameter m represents the fraction of individuals moving from their node) plus the number of susceptible individuals from their resident node i to any other node j contracting there the infection. Following this consideration, one can define the function $\mathcal{F}_i(t)$, which describes infection via contact with contaminated water, as:

$$\mathcal{F}_i(t) = \beta \left[(1-m)\frac{B_i}{K+B_i} + m\sum_{j=1}^{n} Q_{ij}\frac{B_j}{K+B_j} \right] .$$

The parameter β represents the rate of exposure to contaminated water, and $B_i/(K+B_i)$ is the probability of becoming infected due to the exposure to a concentration B_i of vibrios, K being the half-saturation constant (Capasso and Paveri-Fontana, 1979; Codeço, 2001). The matrix Q_{ij} defines the probability that an individual (a susceptible one, in this case) living in node i would reach j as a destination. Choosing a gravity-like pattern to describe human mobility (Erlander and Stewart, 1990), one can define connection probability as

$$Q_{ij} = \frac{H_j e^{-d_{ij}/D}}{\sum_{k\neq i}^{n} H_k e^{-d_{ik}/D}} ,$$

where the attractivity factor of node j depends on its population, while the deterrence factor is assumed to be dependent on distance d_{ij} and represented by an exponential kernel (with shape factor D). Infected individuals recover at a rate γ, or die for natural or cholera-induced mortality at a rate μ or δ, respectively. Similarly to the mechanism already presented, one can suppose that bacterial shedding in node i – quantified by the rate p_i/W_i, with p_i and W_i being the contamination rate caused by bacterial shedding of infectives and the volume of the local water reservoir respectively – is regulated by infected individuals there residing (again, the fraction $1-m$ which does not move from its node), and others who have travelled from their home to node i, so that $\mathcal{G}_i(t)$ can be defined as:

$$\mathcal{G}_i(t) = (1-m)I_i + m\sum_{j=1}^{n} Q_{ji}I_j ,$$

In order to express the worsening of sanitation conditions caused by rainfall-induced runoff, which causes additional loads of pathogens to be cast into the water reservoir, the contamination rate p_i is increased by rainfall $J_i(t)$ via a runoff coefficient ϕ. Bacteria are also supposed to have an environmental mortality μ_B and to be transported on the hydrologic network at rate l. It is assumed that pathogens can travel from node i to j with probability P_{ij}. In this case it is assumed $P_{ij} = 1$ if j is a downstream nearest neighbor of node i and zero otherwise. One modification in introduced here with respect to Rinaldo et al. (2012), in which absorbing boundaries were imposed at the outlets (therein such possibility was only briefly addressed in the Supplementary Material, see e.g. their Figure S8). Here instead a partial reflection of particles flowing downstream from an outlet node, in order to simulate specific coastal hydrologic conditions, such as upstream transport or local environmental conditions (i.e.

a higher salinity) is accounted for, causing higher viability of bacterial populations in coastal regions. The fraction of reflected particles is fixed at 0.5, which corresponds to diffusive conditions. As in previous applications (Bertuzzo et al., 2011; Rinaldo et al., 2012), it is assumed that all the population is susceptible, i.e. $S_i(0) = H_i$ because of the lack of any pre-existing immunity (Enserink, 2010; Walton and Ivers, 2011; Sack, 2011; Piarroux et al., 2011). The model is also initialized with some infected individuals being placed in the locations of the first reported cases (see again Piarroux et al., 2011).

6.2.2 Hyperinfectivity

The second model tested here accounts for a hyperinfective state of *V. cholerae*, caused by passage through human intestine (Merrell and al., 2002; Alam and al., 2005), which has already been applied to modeling exercises (Hartley et al., 2006), also in the context of the Haiti epidemic (Andrews and Basu, 2011; Chao et al., 2011). Thus, an equation describing the dynamics of the hyperinfective stage of vibrios has to be added to model 6.1:

$$\begin{aligned}
\frac{d\mathcal{B}_i}{dt} &= -\xi\mathcal{B}_i - l\left(\mathcal{B}_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}\mathcal{B}_j\right) + \frac{p_i}{W_i}\mathcal{G}_i(t) \\
\frac{dB_i}{dt} &= \xi\mathcal{B}_i - \mu_B B_i - l\left(B_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}B_j\right),
\end{aligned}$$

where $\mathcal{B}_i$ is the concentration of hyperinfective pathogens in the water reservoir and ξ is the rate at which vibrios lose hyperinfectivity and convert to normal state (here $1/\xi = 1$ [day] is imposed). The total contact rate $F_i(t)$ has to take into account the joint effect of hyperinfective and normal vibrios:

$$\begin{aligned}
\mathcal{F}_i^{HI}(t) &= \beta\left[(1-m)\left(\frac{B_i}{K+B_i} + \frac{\mathcal{B}_i}{K_{HI}+\mathcal{B}_i}\right) + \right. \\
&+ \left. m\sum_{j=1}^{n} Q_{ij}\left(\frac{B_j}{K+B_j} + \frac{\mathcal{B}_j}{K_{HI}+\mathcal{B}_j}\right)\right],
\end{aligned}$$

where K_{HI} is the half-saturation constant for hyperinfective bacteria ($K/K_{HI} \approx 50$; Hartley et al., 2006).

This model can thus be written as follows:

$$
\begin{aligned}
\frac{dS_i}{dt} &= \mu(H_i - S_i) - \mathcal{F}_i^{HI}(t)S_i + \rho R_i \\
\frac{dI_i}{dt} &= \mathcal{F}_i^{HI}(t)S_i - (\gamma + \mu + d)I_i \\
\frac{dR_i}{dt} &= \gamma I_i - (\rho + \mu)R_i \\
\frac{d\mathcal{B}_i}{dt} &= -\xi\mathcal{B}_i - l\left(\mathcal{B}_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}\mathcal{B}_j\right) + \frac{p_i}{W_i}\mathcal{G}_i(t) \\
\frac{dB_i}{dt} &= \xi\mathcal{B}_i - \mu_B B_i - l\left(B_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}B_j\right) .
\end{aligned} \tag{6.2}
$$

6.2.3 Parameter calibration and model selection

While several parameters are estimated from the literature (see Table 6.1 for the numerical values and for the relevant references), others are obtained through calibration. The models are calibrated using a Markov Chain Monte Carlo sampling algorithm (ter Braak and Vrugt, 2008, see the Appendix for details). The goodness of each single simulation is computed as the residual sum of squares (RSS) between weekly reported new cholera cases in each Haitian department as recorded in the epidemiological dataset and simulated by the model being tested. Note that the reported cases of the dataset actually represent a fraction σ of the weekly total cases simulated by the models, because of the high fraction of asymptomatic cases that characterizes the spreading of cholera (here $\sigma = 0.2$, PAHO, 2010). The numerical values of the best-fit parameters of both models are reported in Table 6.2. Model selection is then carried out using the Akaike Information Criterion (AIC, Akaike, 1974, see the Appendix).

Here a first test calibration run is performed, from the start of the epidemic (23/10/2011) until 28/05/2011, for both the presented models. The set of tuning parameters is R_0, l, m, D, ϕ for both candidate models (see the definition for the basic reproduction number R_0 in Rinaldo et al., 2012).

The results of this run, expressed with the RSS and AIC values, are shown in Table 6.3. This fitting is used to perform an *ex-post* evaluation of the performance of model predictions, in particular regarding the resurgence of the disease of the summer/autumn of 2011 (so until 31/12/2011). This calibration run will be referred to as run "♣".

In this first run, none of the models emerges as performing significantly better (as the Akaike difference must be > 4 for significance, Akaike, 1974; Burnham and Anderson,

Table 6.1: Estimated parameter values, relevant units, and cited literature for both calibration runs and both models.

Parameter	Units	Value	References
μ	day^{-1}	$1/(61*365)$	CIA (2009)
β	day^{-1}	1.0	Codeço (2001)
ρ	day^{-1}	$1/(3*365)$	Koelle et al. (2005)
γ	day^{-1}	0.20	Codeço (2001)
d	day^{-1}	$4.0 \cdot 10^{-3}$	PAHO (2010)
μ_B	day^{-1}	0.2	Codeço (2001)
σ	-	0.2	WHO (2011)

Table 6.2: Fitted parameter values and relevant units for both calibration runs and both models.

		♣		♢
Parameter	**Units**	*Model 1*	*Model 2*	*Model 2*
R_0		2.71	2.02	2.77
l	day^{-1}	0.70	1.37	2.54
m	-	0.05	0.06	0.05
D	km	95.47	50.82	150.33
ϕ	day mm^{-1}	$6.23 \cdot 10^{-2}$	$4.77 \cdot 10^{-2}$	$2.43 \cdot 10^{-2}$
ρ	day^{-1}	-	-	$2.6 \cdot 10^{-3}$(1.04 years)
σ	-	-	-	0.21

Table 6.3: Number of calibrated parameters and AIC scores for the two tested models (see text for technical details on the candidate models) in the calibration runs 23/10/2010-28/05/2011 (named "♣") and 23/10/2010-14/01/2012 ("♢"). Columns indexed with Δ_{AIC} show Akaike differences (with respect to the best-ranked model), which must be > 4 for significance (Akaike, 1974; Burnham and Anderson, 1983; Corani and Gatto, 2007).

		♣			♢
Model	Θ	RSS	AIC	Δ_{AIC}	RSS
model 6.1	5	$1.905 \cdot 10^8$	4267	1	–
model 6.2	5	$1.902 \cdot 10^8$	4266	–	$3.34 \cdot 10^8$

1983; Corani and Gatto, 2007). Model 2 is chosen as best ranked in the first run and as the model including the higher level of detail and realism and a second calibration run is performed for this model only, until 14/01/2012 (which is called "♢"). In this case, the parameters which play a major role in the dynamics of endemic periodic resurgence are also tuned: the loss of immunity rate (which expresses also the rate of replenishment of the susceptibles compartment), ρ, and the fraction of symptomatic infected individuals (which dictates the depletion of susceptibles in the calibration period), σ. These two parameters control the rate at which the susceptible pool is emptied (σ) and replenished (ρ). At the beginning of an epidemic the dynamics is not controlled by the pool of susceptible which is, particularly in the case of Haiti, very large. Therefore these two parameters have been safely assumed from literature values in the calibration run ♣. In Figure 6.2 the temporal sequence of calibration runs and the subsequent validation/prediction windows for each of them are depicted.

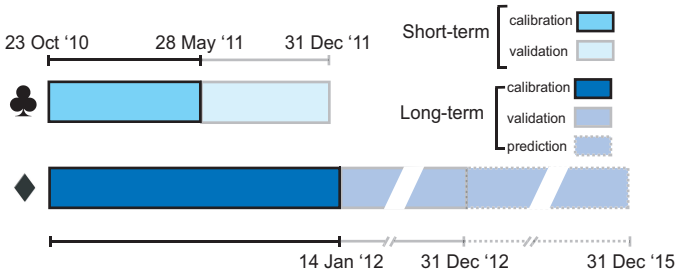

Figure 6.3: Gantt diagram of the calibration runs and validation/prediction windows.

6.2.4 Modeling of interventions

One of the possible uses of epidemiological, large-scale predictions is the evaluation of the effectiveness of different intervention strategies. In the case of a waterborne disease such as cholera, two major categories of intervention can be taken into account: sanitation of the water supply system and vaccination. In the former case, several active (e.g. chlorination tablets, direct clean water supply) or pro-active (education campaigns) actions can be taken to increase the sanitation level in a country. The application of such measures decreases the probability of ingesting contaminated water and/or the probability that critical concentrations of pathogens may reach the water reservoir. They can be expressed as a decrease, respectively, of the contact rate β and/or of the contamination rate p. Instead, vaccinations is modeled as an outgoing flux from the susceptible compartment of individuals being immunized from the disease. It has to be recalled that two doses, administered separately with a 10-14 days delay between the two, are required for current standard cholera vaccination, in which immunization builds up 7 days after the second dose (Jertborn et al., 1993). Also, it has to be considered that standard cholera vaccines only grant temporary immunization, so that immunity loss has to be taken into account in the evaluation of long-term intervention scenarios. Field trials (Clemens, 1990; Girard et al., 2006) suggest parameter values of immunity loss for the only vaccine acknowledged by WHO. In particular, here the immunization is modeled as in the observations by Clemens (1990) in Bangladesh. The vaccine grants 85% immunization in the first 6 months, 60% in the following 18 months and 20% in the third and last year of (partial) immunization (see the new classes $V_{i,I/II/III}$ in the model below). This means that vaccinated people still have a finite probability of contracting the disease, and they may become fully susceptible again once their immunity is lost, after three years. The implementation of interventions in model 2 is detailed here:

$$
\begin{aligned}
\frac{dS_i}{dt} &= \mu(H_i - S_i) - \mathcal{F}_i^{HI}(t)S_i + \rho R_i - \nu_i(t)\frac{S_i}{H_i - C_i - \int_t^{t_v}\nu_i(t)dt} + \rho_{V,III}V_{i,III} \\
\frac{dS_i^v}{dt} &= -\mu S_i^v - \mathcal{F}_i^{HI}(t)S_i^v + \nu_i(t)\frac{S_i}{H_i - C_i - \int_t^{t_v}\nu_i(t)dt} - \omega S_i^v \\
\frac{dV_{i,I}}{dt} &= \omega S_i^v - 0.15\mathcal{F}_i^{HI}(t)V_{i,I} - \rho_{V,I}V_{i,I} \\
\frac{dV_{i,II}}{dt} &= \rho_{V,I}V_{i,I} - 0.4\mathcal{F}_i^{HI}(t)V_{i,II} - \rho_{V,II}V_{i,II} \\
\frac{dV_{i,III}}{dt} &= \rho_{V,II}V_{i,II} - 0.8\mathcal{F}_i^{HI}(t)V_{i,III} - \rho_{V,III}V_{i,III} \\
\frac{dI_i}{dt} &= \mathcal{F}_i^{HI}(t)(S_i + S_i^v + 0.15V_{i,I} + 0.4V_{i,II} + 0.8V_{i,III}) - (\gamma + \mu + \delta)I_i \\
\frac{dR_i}{dt} &= \gamma I_i - (\rho + \mu)R_i \\
\frac{d\mathcal{B}_i}{dt} &= -\xi\mathcal{B}_i - l\left(\mathcal{B}_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}\mathcal{B}_j\right) + \frac{p}{W_i}\mathcal{G}_i(t) \\
\frac{dB_i}{dt} &= \xi\mathcal{B}_i - \mu_B B_i - l\left(B_i - \sum_{j=1}^{n} P_{ji}\frac{W_j}{W_i}B_j\right)
\end{aligned}
\tag{6.3}
$$

where $\nu_i(t)$ is the vaccination rate (number of doses administered per day, possibly varying in time/space) and ω is the rate at which vaccinated individuals gain immunity to the disease (here $\omega = 0.5$ days^{-1}). The denominator $H_i - C_i - \int_t^{t_v}\nu_i(t)dt$, where $C_i = \sigma\mathcal{F}_i^{HI}(t)(S_i + S_i^v)$ is the pool of potential candidates for vaccination. Here all local population is included, except for reported cases (C_i) and people who have already been vaccinated ($\int_t^{t_v}\nu_i(t)dt$). The variables $V_{i,I,II,III}$ represent the abundance of individual in node i which has been vaccinated less than 6, between 6 and 24, and more than 24 months ago, respectively. Loss of immunity of vaccinated individuals progresses at rates $1/\rho_{V,I} = 1/2$ year, $1/\rho_{V,II} = 1.5$ years and $1/\rho_{V,III} = 1$ year. Other symbols as in model 2.

6.3 Rainfall generation patterns

In the hydrological literature, stochastic rainfall generation is often modeled as a marked Poisson process, where rainfall events are treated as a series of point events in continuous time where the associated mark represents the rainfall depth of the event (see e.g. Rodriguez-Iturbe and Porporato, 2004; Laio et al., 2001). This implies that no temporal evolution of a single event is taken into account, such that the amount of rainfall falling at a given time scale – which is usually assumed as daily – is modeled by a point

process. Arrivals of rainfall events is modeled as a Poisson process with rate λ and therefore the inter-arrival time τ between rainfall events is exponentially distributed with mean $\sim 1/\lambda$ i.e.:

$$f_T(\tau) = \lambda e^{-\lambda\tau} \text{ for } \tau \geq 0. \tag{6.4}$$

The depth of each rainfall event is then sampled, again, from an exponential distribution, described by the following probability density function, in which α is the mean rainfall depth of all events:

$$f_H(h) = \frac{1}{\alpha} e^{-h/\alpha} \text{ for } h \geq 0. \tag{6.5}$$

This model has been shown to perform well in describing daily rainfall statistics (see e.g. Benjamin and Cornell, 1970) but it is better suited to applications to confined climatic regions, showing similar precipitation regimes. In Figure 6.3 one can appreciate the peculiarity of the Haitian rainfall patterns in both space (panel A; mean daily rainfall for the period 1998-2012) and time (panel B: average yearly pattern for whole Haiti). Daily rainfall estimates have been obtained from data collected by the NASA-JAXA's Tropical Rainfall Measuring Mission (TRMM_3B42 precipitation estimates, see `http://trmm.gsfc.nasa.gov/` for details) through the IRI/LDEO data portal set up by Columbia University (`http://iridl.ldeo.columbia.edu/`). Rainfall data are spatially distributed with the resolution of 0.25 degrees of latitude and longitude and are then downscaled at the node level with nearest neighbor interpolation.

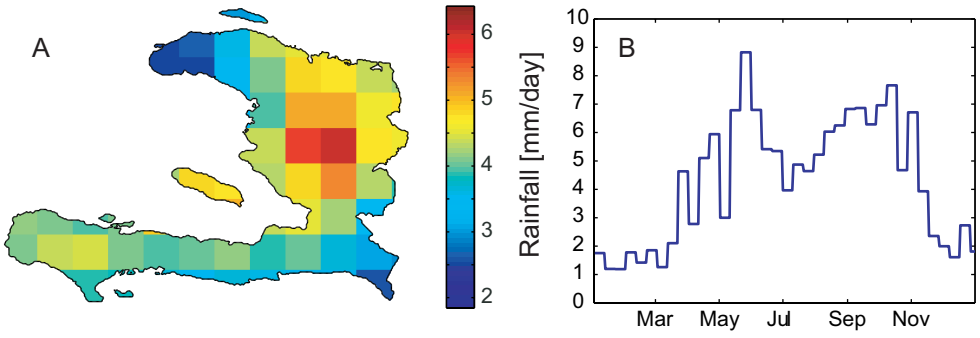

Figure 6.4: Panel A: Raster map of Haiti in which each pixel shows the average depth of rainfall events in the period 1998-2012 (source: Columbia University (see text); resolution: 0.25 degrees); Panel B: yearly time series of rainfall intensity averaged over the whole territory of Haiti and over the whole period of observation; resolution: 10 days.

In order to derive climatically homogeneous regions (and seasons) from data, the mean depth of observed rainfall occurrences (which expresses α) and their mean interarrival time ($1/\lambda$, in days) are calculated at the node scale, for each month. Then a cluster analysis on those parameters is performed, taking also into account the coordinates of each node to preserve spatial continuity of each cluster. Using the k-means clustering technique (Xu and Wunsch, 2005), 20 clusters are identified – over 12 months times 301 spatial nodes – of statistically similar monthly rainfall regimes that are contiguous in space. k-means clustering uses an iterative procedure which is initialized by fixing k means in the space of data points and by progressively changing their position until the distance among such centroids and each point of the cluster is minimized for all clusters (so that the contiguity of the points of each cluster is highest). As the procedure depends on the initial position of the centroids, it has been repeated $10,000$ times choosing the partition which gives the minimum distance among points belonging to the same cluster. In the case at hand, each data point of the 4-dimensional clustering space is characterized by the normalized values of: a) the monthly depth and interarrival time of rainfall events, averaged over the whole period of observation (1998-2012), in each node; b) the spatial coordinates of the node. In order to limit the number of clusters in the spatial subset of the data space, coordinate values are weighted less than rainfall statistics (in this case, the weight is equal to 0.5). Figure 6.4 shows the performance of the clusterization with respect to rainfall statistics (panel A) and to the spatial distribution of clusters (panel B).

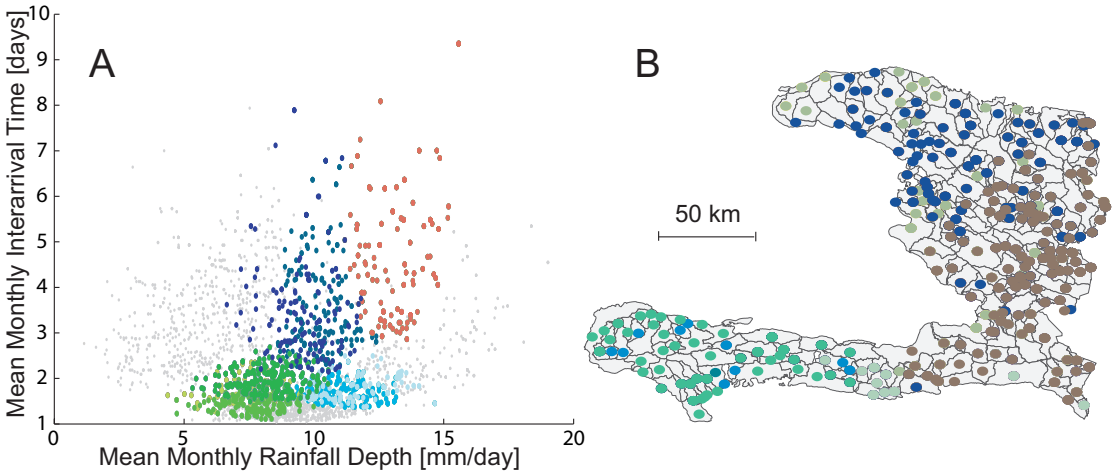

Figure 6.5: Panel A: Scatter plot in which each point represents the average, in the period 1998-2012, of mean monthly depth/interarrival time of rainfall events at the node level; colored dots belong to the same cluster (here the aggregates that emerge for the month of June are shown). Panel B: Distribution of clusters in space, in the month of June.

To generate the maximum range of plausible scenarios of rainfall events in the Haiti territory, a multi-variate Poisson generator is also utilized, such as the one described

in Porporato et al. (2006), in which not only intra-annual, but also inter-annual variability is taken into account. It was observed that different yearly patterns may not be merely described by different realizations of the same stochastic process, but by explicit changes in the statistical properties of the process – the parameters α and λ. This procedure is carried out by assuming that the parameters of the exponential distributions of rainfall depths and inter-arrival times change from year to year and are gamma-distributed random variables (Porporato et al., 2006):

$$g_x(x) = \frac{(b_x)^{(a_x)}}{\Gamma(a_x)} x^{(a_x-1)} e^{-b_x x} \tag{6.6}$$

where x is alternatively α or λ. The parameters a_x and b_x of the distributions of each cluster have been evaluated fitting the empirical distribution extracted from the rainfall dataset using the moments method. The resulting distributions of depths and inter-arrival times can be expressed as:

$$p(h) = \frac{(b_\alpha)^{(a_\alpha+1)/2}}{\Gamma(a_\alpha)} h^{(a_\alpha+1)} / 2K_{1-a_\alpha}[2\sqrt{b_\alpha h}] \tag{6.7}$$

$$p(\tau) = \frac{a_\gamma b_\gamma^{a_\gamma}}{(\tau + b_\gamma)^{a_\gamma+1}} \tag{6.8}$$

where K_{1-a_α} is the Bessel function of order $1 - a_\alpha$ (Abramowitz and Stegun, 1965). To generate daily rainfall events according to the distributions (7) and (8) the following procedure is used:

- for each year of generation and each clusterα's and λ's are extracted from their respective Gamma distribution;

- for each cluster, rainfall events are generated using the corresponding α and λ, assigning the generated daily depth to each node belonging to the respective cluster.

The results are illustrated in Figure 6.5, where the probability density function of the whole ensemble of inter-arrival times and of rainfall depths for the observed rainfall and for the multivariate Poisson generator is shown, for both the global and the local scale. As in Porporato et al. (2006), the so-called "super-statistics" generator shows a good agreement with data.

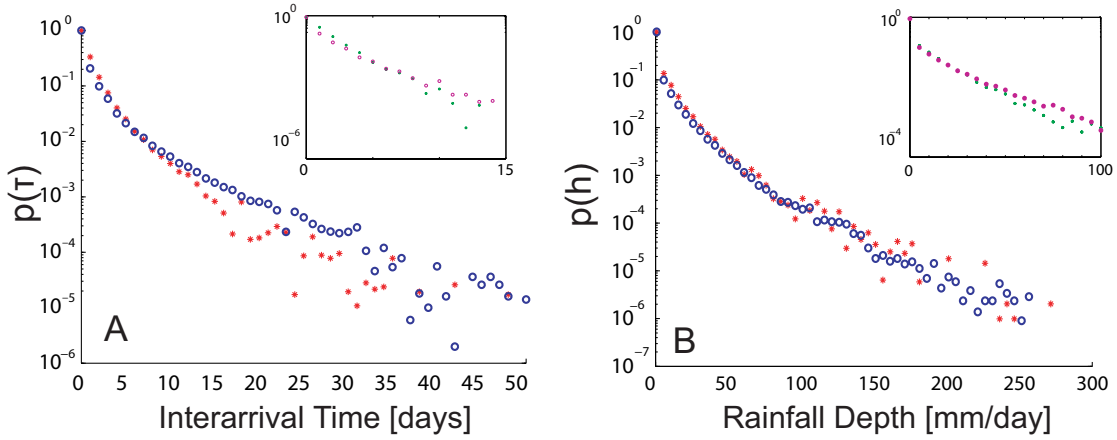

Figure 6.6: Distribution of the inter-arrival times (A) and precipitation depths (B) of rainfall events at the Haiti scale, derived from the multi-variate (red stars) Poisson process or from the data (blue circles). In the inset the same distributions are shown at the cluster scale (one the clusters containing the node representing Port-au-Prince is taken as example), as outcomes of generated rainfall (green dots) or data values (purple dots).

6.4 Results and Discussion

As specified in the Introduction, it is tested whether the models presented previously can be used as an effective tool for the prediction and the control of the course of an ongoing epidemic – in this case, the Haitian one. To this aim, the short-term evolution of the disease is specifically focused on, in order to simulate conditions of epidemiological emergency. For this analysis, the interval between the beginning of the epidemic (October 2010) and the end of May, 2011, just before the June 2011 resurgence driven by the Haitian rainy season, is used as calibration time horizon. It has to be emphasized that a limited dataset is specifically used to calibrate these test models to represent a worst-case scenario, i.e. a situation in which decision makers face limited information missing key drivers. Notice in particular that the calibration window ends right before the spring rainy season and therefore it contains few indications on the effect of rainfall on the epidemic dynamics. This limited information refers to the autumn rainfalls of 2010 and is possibly clouded by the initial boost of the epidemic. All these elements make this prediction exercise particularly challenging. Rainfall scenarios generated by the multi-variate Poisson model presented in the previous section are used to force epidemiological model 2 – which shows a slightly better AIC value – and simulate the course of the epidemic in the following months, until the end of 2011. In all panels the trajectory of the models when forced by the rainfall pattern that was actually observed in Haiti and the range of possible epidemiological scenarios emerging from the generated rainfall patterns are shown.

The different timing of the epidemiological peak – observed in June 2011 from the data and at the end of October in the model simulations – can be explained by the peculiar rainfall pattern that was observed in Haiti in 2011, with a highly concentrated event of very rare intensity at the beginning of June followed by an abnormally wet summer. It is deemed that the difficulty in generating such particular rainfall pattern lies in the limitedness of the dataset (15 years) available to fit the rainfall stochastic model. Simulations performed with the actually observed rainfall pattern display, in fact, a far better synchronization with reported cases (Fig.6.6).

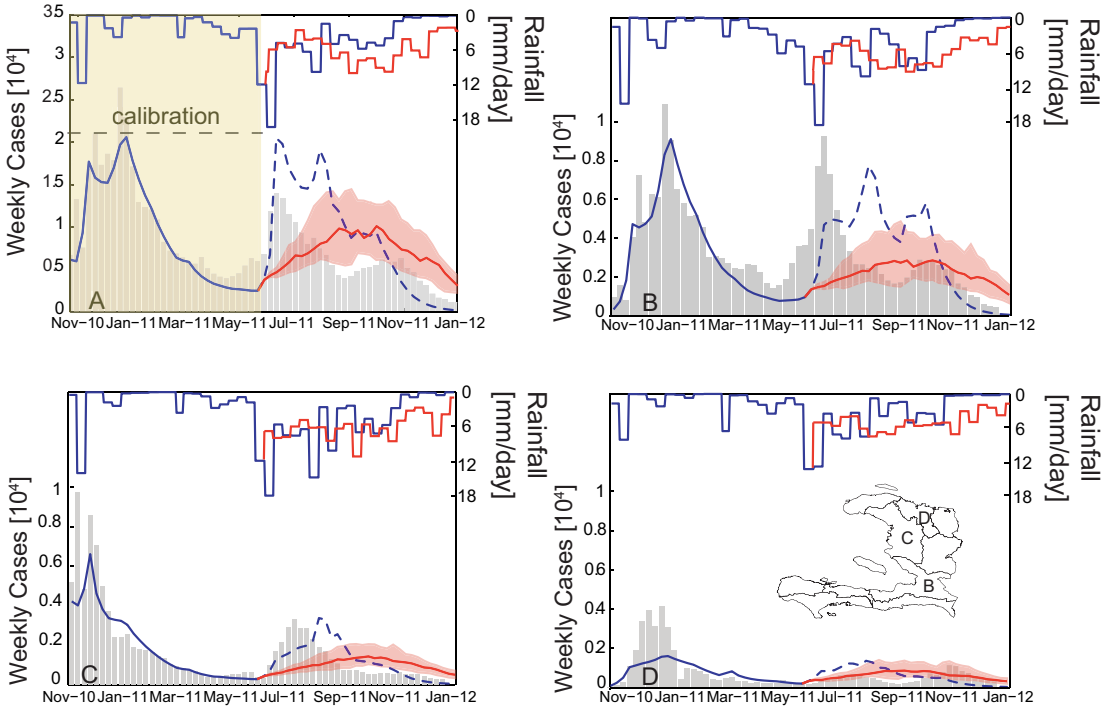

Figure 6.7: Rainfall and epidemiological patterns for the period 23/10/2010-31/12/2011; blue solid lines refer to observed rainfall patterns (decadal and averaged over the whole Haiti, upper part) and to the corresponding model outputs as new weekly cholera cases (blue dashed lines in the validation period); red solid lines show instead one realization in 200 generated rainfall patterns between 28/05/2011 and 31/12/2011 and the median of the corresponding model outputs (the shaded range also shows the 25th-75th percentile span); grey bars depict the weekly reported cases over the simulation horizon. Panel A shows results at the country scale, while patterns for the most populated departments are shown in panels B-D (Ouest, Artibonite, Nord respectively; see inset in panel D). Parameters are taken from the best fit in calibration run ♣.

An important result of this prediction effort is, on the other side, that the model is able to grasp quite nicely the order of magnitude of the new outbreak. The cumulated reported cases between May-December 2011 amount to approximately $188,000$, while

the model forced with generated rainfall events predicts in the median $230,000$ cases. One can appreciate this more clearly when looking at the course of cumulated cases, which is shown in Figure 6.7. This result is somehow reassuring if one aims at using mathematical models as prediction and control tools. Knowing the order of magnitude of the upcoming spreading of the disease in space and time, public health organization can set the scale of the reaction needed to act in time and possibly mitigate the resurgence of the epidemic.

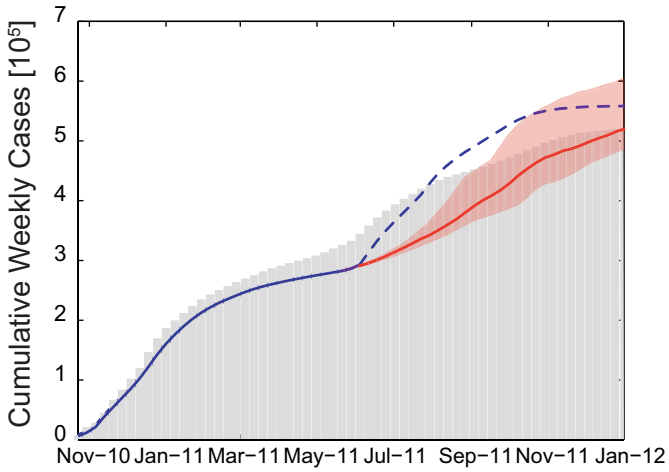

Figure 6.8: Cumulative weekly cases, reported (gray bars) and simulated by model 2 with generated rainfall. Blue solid/dashed line: simulation with the observed rainfall pattern; red solid line (median)/shaded range (25th-75th percentile).

To this regard, mathematical models like the ones presented here offer a invaluable tool to evaluate *a priori* the effectiveness of policy scenarios on the course of the epidemic. An assessment of different intervention scenarios in the same time span (23/10/2010-31/12/2011) is presented (Figure 6.8), using the best fit parameters of the best performing model (model 2, Table 6.3). This time interval, for which it is known that the model is able to reproduce at least the expected number of cases, allows us to make quantitative considerations on the effects of possible control actions. Panels A and C of Figure 6.8 show the effect of a reduction in the contact rate β with contaminated water, which is supposed to decrease linearly, and uniformly in the territory of Haiti, in the span of one month since the start of the campaign. Given such prediction of resurgence peaking in late October and starting in September, it is set either at 01/07/2011 or a month later, at 01/08/2011, depending on the rapidity of decision and intervention deployment. In panels B and D the effect of a spatially uniform vaccination of a fraction of the Haitian population is displayed, under the same assumptions at the start of the campaign (i.e. July or August 2011) and in the same time span of intervention (1 month). A constant vaccination effort is assumed. For details of the mathematical implementation of interventions in the model, see Section 2.4.

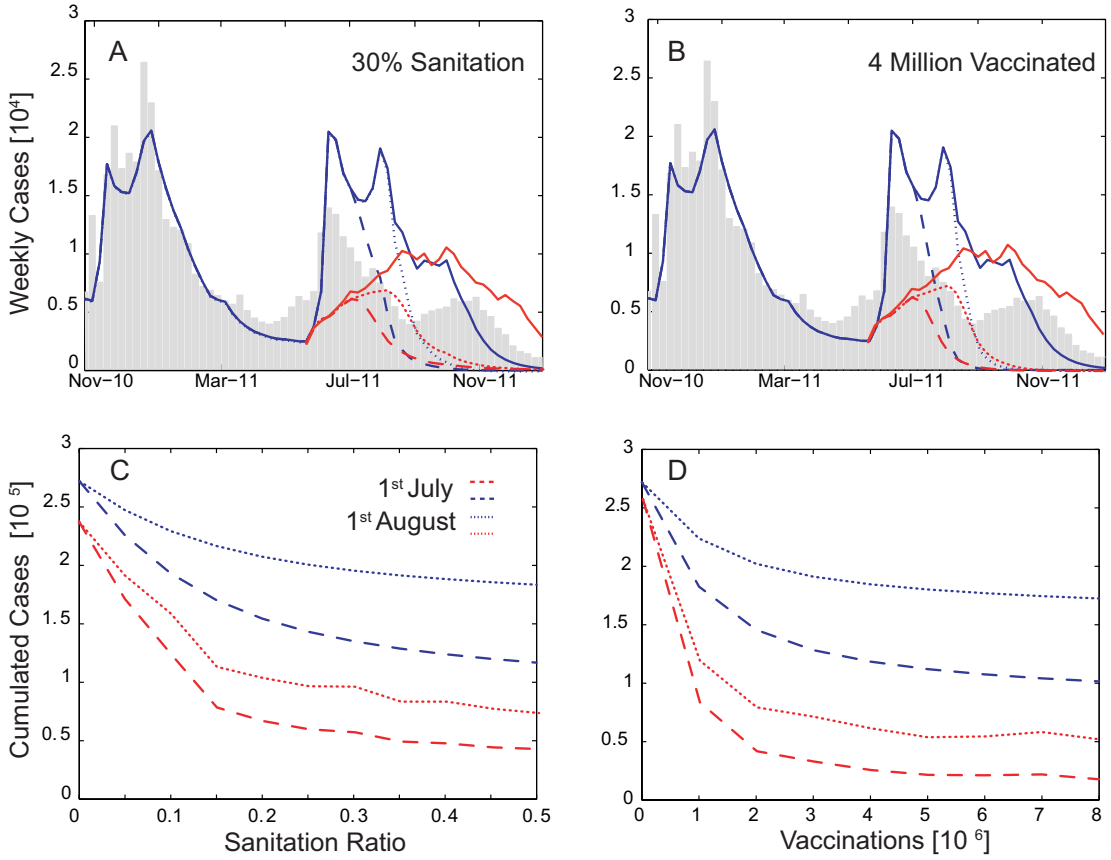

Figure 6.9: Effect of intervention policies on the predicted course of the epidemic between 28/05 and 31/12/2011 (blue lines: simulations with the observed rainfall pattern; red lines/shaded range: median/25th-75th percentile range of simulations with generated rainfall). A) Effect of a reduction of the 30% of the contact rate β, applied in one month starting from the 1st of July (dashed blue/red lines) or the 1st of August (dotted blue/red lines) 2011. B) Effect of a vaccination of 4 million individuals, implemented in one month starting from the 1st of July (dashed blue/red lines) or the 1st of August (dotted blue/red lines) 2011. C) Number of new cases in the period 28/05-31/12/2011 as a function of the reduction of the contact rate β. D) Number of new cases in the period 28/05-31/12/2011 as a function of the number of vaccinated individuals.

While the application of the two types of interventions (which could also be applied simultaneously, of course) presents different kinds of difficulties (therefore the effort needed to implement a given policy can be different and not easily estimated), it is still possible to compare the effectiveness of both sanitation and vaccination at reducing the impact of cholera resurgence on the Haiti population. In panels C and D of Figure 6.7 the total number of new cholera cases between 28/05 and 31/12/2011 as a function of the reduction of the contact rate (C) and of the number of vaccines deployed (D) is shown. Here a maximum effort of 8 million vaccinations in the latter case (corresponding to the entire Haitian population), and of 50% sanitation rate in the former are chosen. Note that such a rate implies that half of the population of each node would never come into contact with contaminated water.

Timely intervention represents an essential feature of any public health policy, be it focused on sanitation or on vaccination. The number of "avoided infections" (the difference between the sum of cases simulated from 28/05/2011 to 31/12/2011 in absence of interventions and the average number of cases when measures are taken) in fact amounts to around $105,000$ (vaccination) and $79,000$ (sanitation) when policies start being applied in July 2011, and to $42,000$ (vaccination) and $25,500$ (sanitation) for actions starting in August 2011. This means a rough average of 40% of total cases avoided should interventions have started in July, while figures drop to 15% had campaigns begun in August.

Another use of the tools implemented deals with multi-seasonal projections of epidemiological predictions, which allows to estimate possible future resurgences of the disease in the case should it become endemic (Tappero and Tauxe, 2011). Figure 6.9 shows a projection of future outbreaks spanning until 31/12/2015. Several (500) simulation runs are performed, generating rainfall events and sampling from the *a posteriori* distribution of the fitting parameter set, which now includes also the rate of immunity loss ρ and the asymptomatic ratio σ. These predictions show a first resurgence of the disease in 2012 and then a settling on an endemic, seasonal pattern of constant annual incidence (Figure 6.9). The comparison with the actual observed cases shows that, although within the percentile range, the model generally overestimates the intensity of the new prevalence peaks. In this respect, an increased awareness regarding common sanitizing practices and the risk connected to the ingestion of contaminated water could have led to lower contact/contamination rates and, thus, to a lower intensity of cholera resurgence. Another factor that can concur to this slight overestimation is the uncertainty related to the calibration of the duration of the immunity of Recovered individuals $1/\rho$ and the ratio of asymptomatic infections σ. In fact, the calibration run $\diamondsuit$ converges to a duration of the immunity of $1-2$ years $(5-95th$ percentile range of the parameter posterior distribution), which causes a fast replenishment of the susceptible pool. The asymptomatic ratio σ ranges between 0.2 and 0.25 in the $5-95th$ percentile range of the parameter posterior distribution. One can observe – see panel B – how the intensity of new outbreaks is set by the ini-

tial size of the susceptible pool and that a threshold size is needed for the outbreak to start (which reminds the concept of generalized reproduction numbers for spatially explicit models; see Gatto et al., 2012). Overall, in an epidemic management context, the results of this long-term prediction are deemed particularly valuable.

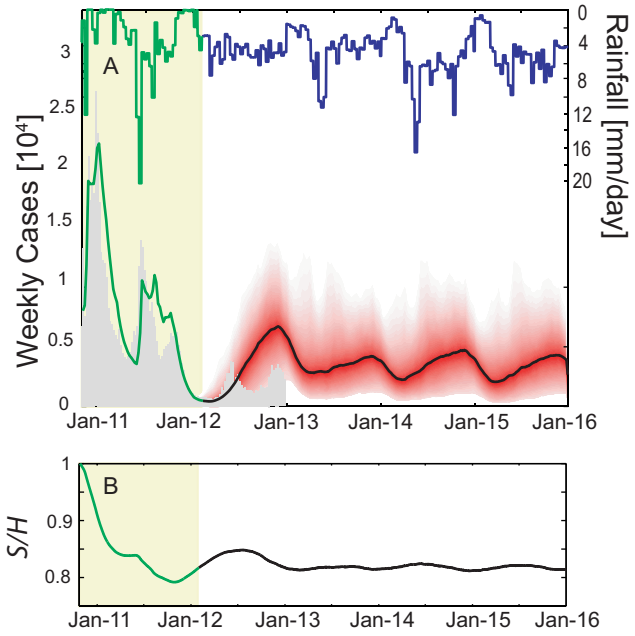

Figure 6.10: Panel A: multi-seasonal projection of the course of the epidemics in the period 15/01/2012-31/12/2015. Calibration is performed between 23/10/2010 and 14/01/2012 (shaded yellow). New weekly cases in Haiti in the calibration (green solid line) and in the prediction (black solid line/shaded range for the median/5th-95th percentile range) periods. Gray bars show the reported cases used for calibration and the cases registered until 31/12/2012. One rainfall realization is also shown in the upper part (green: observed; blue: generated pattern. Panel B: temporal evolution of the median ratio between susceptibles and total population.

The possibility of cholera becoming endemic in Haiti and the possible strategies to eradicate it have been the subject of an intense debate (see e.g. Mukandavire et al., 2013). If one considers vaccination more easily deployed in the field in times of emergency especially when living conditions are precarious and the sanitation system is nearly absent, effects of different intervention policies on the long-term evolution of the disease must be accounted for. Figure 6.10 illustrates the effects of different intervention scenarios on the predicted course of the epidemic in the period 14/01/2012-31/12/2015 (policy implementation starting on 01/08/2012 and lasting 1 year in the case of sanitation and 90 days in the case of vaccination, to reflect the different effort required to implement these measures). In panel A the effect of vaccinating 1 to 4 million susceptibles is displayed.

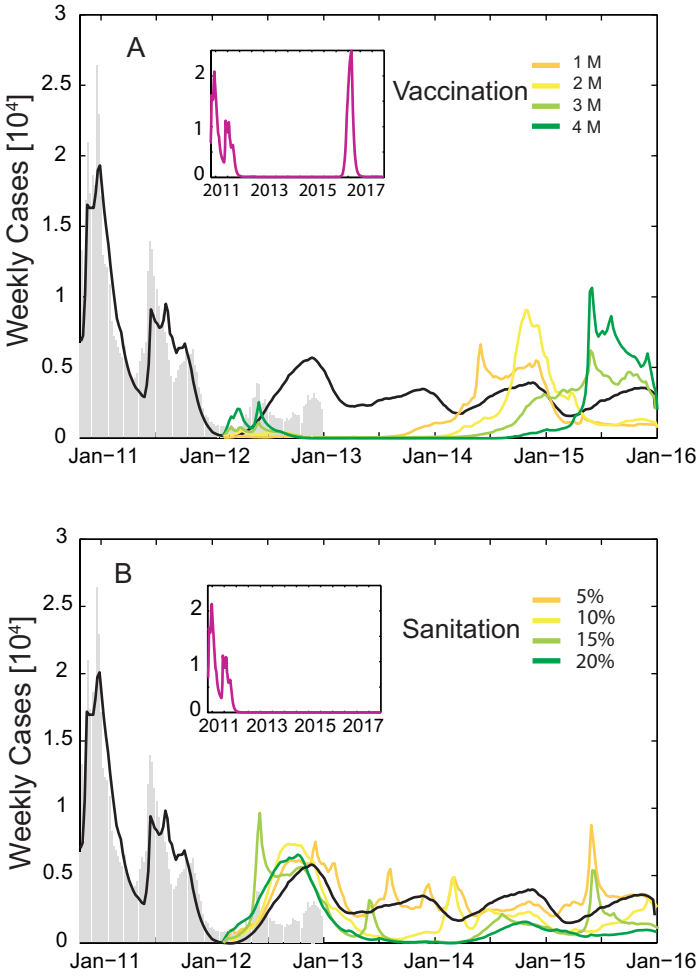

Figure 6.11: Effect of intervention policies on the predicted course of the epidemic between 08/2012 and 12/2015. The black solid line shows the median value of the prediction with no planned intervention. Other solid lines (from yellow to green) show the average predicted pattern of new weekly cases with increasing effort of vaccination (A) or sanitation (B). In the inset of both panels the lilac solid line shows the same pattern until 12/2017, for 8 million vaccinations (A) or 30% sanitation (B). Notice that small differences before the implementation of interventions may arise from different realizations of the stochastic rainfall scenarios.

One can already appreciate the effect of a mass vaccinations campaign when carried out only once: increasing the number of vaccinations has the sole effect of delaying – and of, maybe, exacerbating – the resurgence of the disease, even in the extreme event of vaccinating the entire population of Haiti(see inset of panel A, portraying the predicted evolution of new weekly cases until 31/12/2017). Note, though, that this is a worst case scenario as no concurrent improvement of sanitation conditions is accounted for. In the other case (panel B), it is shown that a sanitation campaign, besides other positive effects brought by an increased drinking water quality, can effectively eradicate the disease from the country, provided that the final effort exceeds a certain threshold. In this analysis, the threshold lies somewhere between a $20 - 30\%$ reduction of the contact rate in all nodes of the Haitian network (see the inset of panel B, showing the effect of 30% sanitation on the course of the epidemic until 31/12/2017). Focusing only on vaccination as intervention policy, the eradication of the disease could have been obtained only if periodic vaccination campaigns were set up via effective vaccines. These periodical campaigns, however, may well imply an effort comparable to that needed to minimize exposure probabilities, say by sanitizing the water supply system on a permanent basis. No complete eradication of the disease would be reached, however, without improving concurrently the sanitation conditions of water supply in the country. The limited stock of vaccines that is currently available has to be considered as well (which counts less than $400,000$ doses currently; Waldor et al., 2010), which would not allow, of course, vast vaccinations campaigns such as the ones which have been simulated. It has to be reminded that mass vaccinations, however, remain one of the potentially fastest interventions to be deployed in an emergency.

6.5 Conclusions

The following conclusions are worth mentioning:

- Mathematical, spatially explicit models provide a tool to predict and control the course of ongoing cholera epidemics. The relevance of this new class of models relates to the fact that inappropriate responses can be avoided by providing adequate and timely information to policy-makers, decision-makers, the media and the public. The theoretical framework proposed could be applied, with appropriate modifications, to other cholera epidemics on the premises that infection mechanisms can be tackled reasonably. While several issues remain open, like the field validation of parameters defined node-by-node, major public health policy challenges like those involving limitations of human mobility, structural measures (construction of hospitals, placement of field hospitals, construction of water sanitation infrastructures) and other interventions (vaccination and/or sanitation campaigns, antibiotics administration) can be thor-

oughly addressed by the proposed class of models;

- rainfall patterns have been generated and used to force the infection model. A Poisson generator integrated with interannual variability and differentiated at the spatio-temporal scale has been used. Poisson generators constitute a simple, synthetic way to generate rainfall patterns preserving the basic statistics of observed precipitation events (depth and interarrival times), but they are indeed flawed when it comes to reproducing widespread events, in both space and time, such as monsoons and tropical storms in general. An altered intensity and frequency of these particular rainfall events may influence dramatically the conditions of resurgence of the disease;

- multi-season projection of the disease patterns under diverse assumptions are provided. It is suggested that vaccination alone, still considered in many studies as the key form of outbreak control, may be effective in the short term but would avoid resurgence of the disease only if sanitation conditions improve concurrently. Relative merits of the various interventions can therefore be weighed on a quantitative basis;

- applications of spatially explicit models prove powerful monitoring tools. The epidemiological framework was in fact coupled with projections of rainfall scenarios that have been used to force the infection patterns. Short- and long-term assessments of the possible evolution of the disease have been produced and discussed.

7 Conclusions and Recommendations

The main effort characterizing the activities presented in this Thesis regards the deepening of the knowledge on the processes regulating cholera occurrence in both epidemic and endemic contexts. All steps taken in this work are instrumental to the creation of modeling tools that may encompass the mechanisms that drive cholera invasion of new countries as well as the processes that regulate its reappearance and, ultimately, its possible endemic residence in such countries. Such tools may aid public health organizations in monitoring and predicting ongoing epidemics and in the setting up of intervention policies.

All the work presented here lies within a common framework, i.e. a mathematical representation of a given system taking into account both local dynamics and their interaction with spatial, regional scale processes. These spatially explicit models all rely on the description of a system composed of a network of local human communities, which are connected by hydrologic transport of *V. cholerae* and human mobility. The latter is also important for local contamination, as susceptible individuals may contract the disease outside their resident node and then bring there the infection, but also infected asymptomatic individuals – usually the larger share of infections – may travel to other communities, shedding bacteria where cholera was not previously present.

In Chapter 2, the impact of seasonally varying hydrologic conditions at local scale has been investigated. As open water bodies are often used for domestic purposes in developing countries (sometimes even for drinking), periodic drought periods can cause progressive increases of the concentration of *V. cholerae*, leading to sudden reoccurrences of the disease. Observed seasonal patterns in various areas of the world can be reconstructed by simply accounting for regular seasonal fluctuations of the volume of the local water reservoir. The analysis of all possible epidemiological pat-

terns for varying amplitude of the forcing and average intensity of the disease reveals that also chaotic dynamics are possible, in such a way that spectral characteristics observed in real case studies are reproduced by regular seasonal forcings, without any interannual variations.

The proposed mechanism is then tested in a field setting, in the experiment described in Chapter 3. The field work has been carried out in the area of Matlab in rural Bangladesh, a well-known area of endemicity of cholera. There, the common water reservoir is represented by a pond, which is surrounded by elevated ground, where human settlements lie. The water level in the pond has been monitored for one year and a half, together with relevant hydro-climatological variables, such as rainfall and solar radiation, which are responsible for such variations of water volume. Concurrently, the concentrations of phytoplankton and *V. cholerae* in the water were periodically measured. These displayed peaks during the drought period, when the water volume was at its lowest, but also in post-monsoonal conditions. This observation enlightens the need for the consideration of a second mechanism of infection, one that takes into account the revamping of bacterial and microbial populations caused by nutrient transport in local water reservoirs. In the area of Matlab, peaks of cholera prevalence are observed, in fact, in both pre- and post-monsoonal periods. A simple water balance model has been described, to identify the major inputs and outputs regulating the water volume in the pond, and an ecological model of phytoplankton-vibrios interaction has been proposed, to further describe the processes that control the survival and the growth of bacteria in aquatic environments. In fact, the patterns of the two populations turned out to be closely correlated, showing peaks of abundance in both pre-monsoonal – corresponding with the lowest water level – and post-monsoonal – when the quantity of nutrients in the water is supposedly higher – periods.

As the complexity of the modeling framework increases, it is vital that model assumptions, together with calibrated parameters and non-observed state variables, can be all verified on the field. A procedure to investigate rapidly for presence of *V. cholerae* in water samples has been tested in Chapter 4. There the results of a survey of the ponds and of other water bodies surrounding the area of the Matlab market have been shown. In this survey, water samples collected between September and October 2012 were analyzed with a flow cytometer, an instrument that allows to perform screenings specifically for *V. cholerae* O1 and O139 serogroups, the ones responsible for cholera epidemics. Using this approach it is shown that both serogroups are present in the environmental waters with a consistent dominance of *V. cholerae* O1. These results extend earlier studies where *V. cholerae* O1 and O139 were mostly detected in times of cholera epidemics using standard culturing techniques. Furthermore, these results confirm that specific ponds host a population of vibrios, which is able to self-sustain even when cholera cases are scarce. Such contaminated ponds may constitute a natu-

ral reservoir for endemicity in the Matlab region. The procedure is replicable in other contexts, as the instrument itself is portable and only basic (or portable) laboratory equipment is required for processing of the samples.

A preliminary study of the influence of human mobility on cholera spreading on a theoretical 1-D lattice is carried out in Chapter 5. On such a structure, in fact, the celerity of the propagation wave can be derived analytically. Human mobility has been shown to be important, in particular, to propel the upstream spreading of the disease. When human movement is taken into account, the minimum value of the diffusion coefficient needed for the formation of a regressive wave is significantly lower than in absence of the human contribution. This work has been part of the definition of a model of human mobility describing the daily fluxes of humans between settlements, which is responsible, also, for the seeding of the disease outside the hydrologic boundaries.

The theoretical and experimental evidence which has been collected in these first four chapters all contributed to the development of a more detailed modeling framework, which was applied to the case of Haiti (Bertuzzo et al., 2011; Rinaldo et al., 2012; Gatto et al., 2012). Several of the investigated processes were needed, in fact, to characterize the resurgence of the disease in the region:

- human mobility, because of the observations of fast intercatchment transmission of the infection that would not be explained by water-mediated pathogen dispersal in times of cholera;
- immunity loss, as it influences the intensity of the disease resurgence, because it also stands for the rate of replenishment of the susceptibles compartment;
- hydrological forcings, as they were found to be responsible of increased water contamination through rainfall-runoff transfer of *V. cholerae* from waste- to drinking-water and necessary to reproduce the disease resurgence.

A final chapter deals with the integration of the previously developed spatially explicit model for the Haitian epidemic with generated rainfall scenarios, to test the model as a prediction and control tool. The rainfall model included spatially and temporally varying parameters of intensity and inter-arrival time, following a correlation structure that was derived from data. The epidemiological model required, in fact, rainfall as an input, as it has been shown that the resurgence of the disease in Haiti has been driven by major rainfall events, that have caused contamination of water reservoirs with open surface defecation sites. Generated rainfall events were used to project the model simulations onwards and to evaluate the predictive power of the model with a hindcast analysis. The model is able to capture quite well the expected number of

cases in the hindcast interval. This type of information can be of vital importance for setting up intervention strategies, as their impact can be safely evaluated *a priori*, as it is shown in this Chapter. Moreover, rainfall scenarios are used to produce multi-seasonal projections, to investigate for future, possible reoccurrences of the disease in the region. Finally, intervention strategies to eradicate the disease from the country are devised.

All the topics addressed in this Thesis encompass possible research developments. The ecological model illustrated in Chapter 3 could shed some light on the local dynamics of *V. cholerae*, highlighting the crucial processes that govern bacteria survival and growth in the environment. As cholera is expected to become – or already is – endemic in other areas of the world, mostly in Africa, where the population is still vulnerable to the disease, such knowledge would be useful to identify the environmental conditions that regulate the seasonal occurrence of the disease (Chapter 3). In this respect, the procedure which has been tested in the field for the detection of *V. cholerae* O1 and O139 (Chapter 4) can be replicated in other contexts and would be possibly more relevant, as less evidence on the environmental reservoirs of the bacteria exists outside Bangladesh.

Human mobility patterns (Chapters 5 and 6) also deserve further work. Until now, the movements of susceptible and individuals among settlements have been estimated on the assumption of gravitational fluxes – i.e. fluxes being proportional to the product of the populations of the origin and of the destination of travel. All parameters (say, the fraction of resident population in a given community travelling daily towards another destination, or the average travelled distance) have been left free in the calibration process of the epidemiological model, so far, but new insight could be gained from mobile phone data. These data can provide information on the position of a user during the day, i.e. its resident node and its preferred travelling destinations. This would allow not only to infer some of the parameters of the human mobility model, but also to devise and evaluate the effect of intervention policies based on mobility limitations.

In the case of Haiti, for instance, more insight is required on the effect on cholera infection of storm events which, as cloud masses, sweep across the land. These highly correlated, extreme rainfall events have been associated with all the reoccurrences of cholera in the region. Predicting in advance at least these storm events, which possibly act as a trigger mechanism for the resurgence of the epidemic, would be an important help in setting up a sort of “alert system” for cholera in Haiti.

The present Thesis aims, in general, at contributing to the research on the processes relevant to the epidemiology of cholera,in both epidemic and endemic settings. This is key to developing usable real-time tools for the assessment and control of epidemics. Quantitative evaluation of the effects of interventions, be them "soft" (i.e. vaccination, chlorination, distribution of antibiotics) or "hard" (i.e. construction of new hospitals or of sanitation infrastructures), can in fact be crucial for policy-makers – especially in this case, in which delicate decisions on public health emergencies have to be taken. Still, no real decision support system has been developed up to now to evaluate the effectiveness of combined strategies in the short and long term.

The framework this Thesis has contributed to develop can be used to assess the effectiveness of policies varying their application in space and time. Many of the possible interventions can – or have to be – distributed in time (for scarcity of vaccination doses, for instance) and space (e.g. ring vaccination). This poses technical and computational challenges, as no application of optimal control theory has been attempted, so far, for spatially explicit models of cholera. In fact, it should be possible to devise policies adapting constantly to the newly available data. In this respect and more generally, the application of modeling tools to the management of epidemics has to rely on a timely monitoring of new cholera cases, with the highest possible spatial and temporal resolution. In the case of Haiti, data on cases and mortality were made available on a weekly basis, at the Department level. A multi-criteria approach would be a necessary choice, as optimal strategies for, say, the minimization of cases or victims of the disease in a region may still not be the most acceptable at social level (Belton and Stewart, 2003). In this framework, specific techniques of decision analysis (decision trees, Pareto optimality, utility and risk functions) must be applied.

Finally, a user-friendly interface for policy-makers to evaluate the effect of intervention strategies would be an additional instrument to bridge the gap between mathematical epidemiology and public health management, in both emergency and standard conditions. Such endeavour, in general, would be a necessary step for giving quantitative bases to decisions in terms of public health management.

Bibliography

Abramowitz, M., Stegun, I., 1965. Handbook of Mathematical Functions. Orlando, FL 32887-0405 (USA): Harcourt Brace Janovich.

Agam, N., Kustas, W., Anderson, M., Norman, J., Colaizzi, P., Howell, T., Prueger, J., Meyers, T., Wilson, T., 2010. Application of the priestley-taylor approach in a two-source surface energy balance model. Journal of Hydrometeorology 11, 185–198.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716–723.

Akanda, A. S., Jutla, S., Islam, S., 2009. Dual peak cholera transmission in Bengal Delta: A hydroclimatological explanation. Geophysical Research Letters 36, L19401.

Al-Tawfiq, J. A., Memish, Z. A., 2012. The Hajj: updated health hazards and current recommendations for 2012. Eurosurveillance 17 (41), 6–10.

Alam, A., al., 2005. Hyperinfectivity of human-passaged V. *cholerae* can be modeled by growth in the infant mouse. Infection and Immunity 73, 6674–6679.

Alam, M., Sultana, M., Nair, G., Sack, R., Sack, D., Siddique, A., Ali, A., Huq, A., Colwell, R., 2006. Toxigenic Vibrio cholerae in the aquatic environment of Mathbaria, Bangladesh. Applied and Environmental Microbiology 72 (4), 2849–2855.

Albert, M., Ansaruzzaman, M., Bardhan, P., Faruque, A., Faruque, S., Islam, M., Mahalanabis, D., Sack, R., Salam, M., Siddique, A., Yunus, M., Zaman, K., 1993. Large epidemic of cholera-like disease in Bangladesh caused by *Vibrio cholerae* O139 synonym Bengal. Lancet 342 (8868), 387–390.

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., Rohani, P., 2006. Seasonality and the dynamics of infectious diseases. Ecology Letters 9 (4), 467–484.

Anderson, R., May, R., 1978. Regulation and stability of host-parasite population interactions: I. regulatory processes. Journal of Animal Ecology 47 (1), 219–247.

Anderson, R., May, R., 1979. Population biology of infectious diseases: Part I. Nature 280 (5721), 361–367.

Anderson, R., May, R., 1992. Infectious Diseases of Humans. Oxford, UK: Oxford University Press.

Andrews, J., Basu, S., 2011. Transmission dynamics and control of cholera in Haiti: an epidemic model. Lancet 377, 1248–1252.

Arasteh, P. D., Tajrishy, M., 2008. Calibrating Priestley-Taylor model to estimate open water evaporation under regional advection using volume balance method-case study: Chahnimeh reservoir, Iran. Journal of Applied Sciences 8, 4097–4104.

Aronson, D., 1978. A comparison method for stability analisys of nonlinear parabolic problems. SIAM Review 20, 245–264.

Becker, N., 1979. The uses of epidemic models. Biometrics 35 (1), 295–305.

Belton, V., Stewart, R., 2003. Multiple Criteria Decision Analysis: an Integrated Approach. Kluver Academic Publishers.

Benjamin, J., Cornell, C., 1970. Probability, Statistics and Decision for Civil Engineers. New York, McGraw-Hill.

Bertuzzo, E., Azaele, S., Maritan, A., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., 2008. On the space-time evolution of a cholera epidemic. Water Resources Research 44, W01424.

Bertuzzo, E., Casagrandi, R., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., 2010. On spatially explicit models of cholera epidemics. Journal of The Royal Society of Interface 7, 321–333.

Bertuzzo, E., Mari, L., Righetto, L., Gatto, M., Casagrandi, R., Blokesch, M., Rodriguez-Iturbe, I., Rinaldo, A., 2011. Prediction of the spatial evolution and effects of control measures for the unfolding Haiti cholera outbreak. Geophysical Research Letters 38, L06403.

Bertuzzo, E., Mari, L., Righetto, L., Gatto, M., Casagrandi, R., Rodriguez-Iturbe, I., Rinaldo, A., 2012. Hydroclimatology of dual-peak annual cholera incidence: Insights from a spatially explicit model. Geophysical Research Letters 39, L05403.

Bertuzzo, E., Maritan, A., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., 2007. River networks and ecological corridors: reactive transport on fractals, migration fronts, hydrochory. Water Resources Research 43, W04419.

Bouma, M., Pascual, M., 2001. Seasonal and interannual cycles of endemic cholera in Bengal 1891-1940 in relation to climate and geography. Hydrobiologia 460, 147–156.

Brutsaert, W., 2005. Hydrology: An Introduction. Cambridge University Press.

Burnham, K., Anderson, D., 1983. Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach. Springer-Verlag, New York, N.Y., USA.

Campos, D., Fort, J., Mendez, V., 2006. Transport on fractal river networks: Application to migration fronts. Theoretical Population Biology 69 (1), 88–93.

Campos, D., Mendez, V., 2005. Reaction-diffusion wave fronts on comblike structures. Physical Review E 71 (5, Part 1).

Capasso, V., Paveri-Fontana, S., 1979. A mathematical model for the 1973 cholera epidemic in the European Mediterranean region. Révue Epidemiologique et Santé Publique 27, 121–132.

Casagrandi, R., Bolzoni, L., Levin, S., Andreasen, V., 2006. The SIRC model and influenza A. Mathematical Biosciences 200 (2), 152–169.

Chao, D. L., Halloran, M. E., Longini, Jr., I. M., 2011. Vaccination strategies for epidemic cholera in Haiti with implications for the developing world. Proceedings of the National Academy of Sciences of the United States of America 108 (17), 7081–7085.

CIA, 2009. CIA's World Factbook, available online at https://www.cia.gov/library/publications/the-world-factbook/index.html.

Clemens, J. e., 1990. Field trial of oral cholera vaccines in Bangladesh: results from three-year follow-up. Lancet 335, 270–273.

Cockburn, T., J.G., C., 1950. Epidemiology of endemic cholera. Public Health Reports 75, 791–803.

Codeço, C., 2001. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. BMC Infectious Diseases 1 (1).

Collins, A., 2003. Vulnerability to coastal cholera ecology. Social Science & Medicine 57 (8), 1397–1407.

Colwell, R., 1996. Global climate and infectious disease: the cholera paradigm. Science 274, 2025–2031.

Colwell, R., Brayton, P., Grimes, D., Roszak, D., Huq, S., Palmer, L., 1985. Viable but non-culturable *Vibrio-cholerae* and related pathogens in the environment - Implications for release of genetically engineered microorganisms.

Colwell, R., Kaper, J., S.W., J., 1977. *Vibrio cholerae*, *Vibrio parahaemolyticus*, and other vibrios: occurrence and distribution in chesapeake bay. Science 198, 394–396.

Corani, G., Gatto, M., 2007. Structural risk minimization: a robust method for density-dependence detection and model selection. Ecography 30, 400–416.

de Magny, G. C., Mozumder, P. K., Grim, C. J., Hasan, N. A., Naser, M. N., Alam, M., Sack, R. B., Huq, A., Colwell, R. R., 2011. Role of zooplankton diversity in *Vibrio cholerae* population dynamics and in the incidence of cholera in the Bangladesh sundarbans. Applied And Environmental Microbiology 77 (17), 6125–6132.

de Magny, G. C., Murtugudde, R., Sapiano, M. R. P., Nizam, A., Brown, C. W., Busalacchi, A. J., Yunus, M., Nair, G. B., Gil, A. I., Lanata, C. F., Calkins, J., Manna, B., Rajendran, K., Bhattacharya, M. K., Huq, A., Sack, R. B., Colwell, R. R., 2008. Environmental signatures associated with cholera epidemics. PNAS 105 (46), 17676–17681.

Dhooge, A., Govaerts, W., Kuznetsov, Y., 2003. MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. ACM Transactions on Mathematical Software 29 (2), 141–164.

Dick, M. H., Guillerm, M., Moussy, F., Chaignat, C.-L., 2012. Review of two decades of cholera diagnostics - how far have we really come? Plos Neglected Tropical Diseases 6 (10).

Doveri, F., Scheffer, M., Rinaldi, S., Muratori, S., Kuznetsov, Y., 1993. Seasonality and Chaos in a Plankton Fish Model.

Earn, D., Rohani, P., Bolker, B., Grenfell, B., 2000. A simple model for complex dynamical transitions in epidemics. Science 287 (5453), 667–670.

Emch, M., 1999. Diarrheal disease risk in Matlab, Bangladesh. Social Science & Medicine 49 (4), 519–530.

Emch, M., Feldacker, C., Islam, M. S., Ali, M., 2008. Seasonality of cholera from 1974 to 2005: A review of global patterns. International Journal of Health Geographics 7, 31.

Emch, M., Yunus, M., Escamilla, V., Feldacker, C., Ali, M., 2010. Local population and regional environmental drivers of cholera in Bangladesh. Environmental Health 9.

Enserink, M., 2010. Infectious diseases: Haiti's outbreak is latest in cholera's new global assault. Science 330, 738–739.

Erlander, S., Stewart, N. F., 1990. The Gravity Model in Transportation Analysis – Theory and Extensions. VSP Books, Zeist, The Netherlands.

Faruque, S., Bin Naser, I., Islam, M., Faruque, A., Ghosh, A., Nair, G., Sack, D., Mekalanos, J., 2005. Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 102 (5), 1702–1707.

Faruque, S., Chowdhury, N., Kamruzzaman, M., Ahmad, Q., Faruque, A., Salam, M., Ramamurthy, T., Nair, G., Weintraub, A., Sack, D., 2003. Reemergence of epidemic Vibrio cholerae O139, Bangladesh. Emerging Infectious Diseases 9 (9), 1116–1122.

Gatto, M., Mari, L., Casagrandi, R., Bertuzzo, E., Righetto, L., Rodriguez-Iturbe, I., Rinaldo, A., 2012. Generalized reproduction numbers and the prediction of patterns in waterborne disease. Proceedings of the National Academy of Sciences of The United States of America 109 (48), 19703–19708.

Girard, M., Steele, D., Chaignat, M.L. Kieny, M., 2006. A review of vaccine research and development: human enteric infections. Vaccine 24, 2732–2750.

Glass, R., Becker, S., Huq, M., Stoll, B., Khan, M., Merson, M., Lee, J., Black, R., 1982. Endemic Cholera in Rural Bangladesh, 1966-1980. American Journal of Epidemiology 116 (6), 959–970.

Gonzalez, M., Hidalgo, C., Barabasi, A., 2008. Understanding individual human mobility patterns. Nature 453, 7196, 779–782.

Gragnani, A., Scheffer, M., Rinaldi, S., 1999. Top-down control of cyanobacteria: A theoretical analysis. AMERICAN NATURALIST 153 (1), 59–72.

Hamer, W., 1906. Epidemic disease in England: the evidence of variability and the persistence of type. Lancet 167, 733–739.

Hammes, F., Berney, M., Wang, Y., Vital, M., Koester, O., Egli, T., 2008. Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. Water Research 42 (1-2), 269–277.

Hammes, F., Egli, T., 2010. Cytometric methods for measuring bacteria in water: advantages, pitfalls and applications. Analytical And Bioanalytical Chemistry 397 (3), 1083–1095.

Hartley, D., Morris, J., Smith, D. L., 2006. Hyperinfectivity: A critical element in the ability of *V. cholerae* to cause epidemics? PLoS Medicine 3, 63–69.

Hashizume, M., Faruque, A. S. G., Wagatsuma, Y., Hayashi, T., Armstrong, B., 2010. Cholera in Bangladesh: climatic components of seasonal variation. Epidemiology 21 (5), 706–710.

Huq, A., Colwell, R., Rahman, R., Ali, A., Chowdhury, M., Parveen, S., Sack, D., Russekcohen, E., 1990. Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and culture methods. Applied And Environmental Microbiology 56 (8), 2370–2373.

Huq, A., Sack, R., Nizam, A., Longini, I., Nair, G., Ali, A., Morris, J., Khan, M., Siddique, A., Yunus, M., Albert, M., Sack, D., Colwell, R., 2005. Critical factors influencing

the occurrence of *Vibrio cholerae* in the environment of Bangladesh. Applied and Environmental Microbiology 71 (8), 4645–4654.

Iran Meterological Service, 2011. Iran rainfall data are available online at http://www.irimo.ir/english/monthly&annual/map/index.asp.

Islam, M., Drasar, B., Bradley, D., 1990. Long-term persistence of toxigenic *Vibrio cholerae* 01 in the mucilaginous sheath of a blue-green alga, *Anabaena variabilis*. J Trop Med Hyg 93, 133–139.

Islam, M., Drasar, B., Sack, D., 1994. Probable role of blue-green algae in maintaining endemicity and seasonality of cholera in bangladesh: A hypothesis. Journal of Diarrhoeal Diseases Research 12, 245–256.

Islam, M., Hasan, M., Miah, M., Yunus, M., Zaman, K., Albert, M., 1994. Isolation of *Vibrio cholerae* O139 synonym Bengal from the aquatic environment in Bangladesh - implications for disease transmission. Applied And Environmental Microbiology 60 (5), 1684–1686.

Islam, M. S., Sharker, M. A. Y., Rheman, S., Hossain, S., Mahmud, Z. H., Islam, M. S., Uddin, A. M. K., Yunus, M., Osman, M. S., Ernst, R., Rector, I., Larson, C. P., Luby, S. P., Endtz, H. P., Cravioto, A., 2009. Effects of local climate variability on transmission dynamics of cholera in Matlab, Bangladesh. Transactions of the Royal Society of Tropical Medicine and Hygiene 103 (11), 1165–1170.

Jertborn, M., Svennerborn, A., Holmgren, J., 1993. Evaluation of different immunization schedules for oral cholera B subunit-whole cell vaccine in Swedish volunteers. Vaccine 11, 1007–1012.

Johnson, A., Hatfield, C., Milne, B., 1995. Simulated diffusion dynamics in river networks. Ecological modelling 83, 311–325.

Jury, E., Mansour, M., 1981. Positivity and nonnegativity conditions of a quartic equation and related problems. IEEE Transactions On Automatic Control 26, 444–451.

Kaper, J., Morris, J., Levine, M., 1995. Cholera. Clin. Microbio. Rev. 8, 48–86.

Kelvin, A. A., 2011. Cholera outbreak in the Republic of Congo, the Democratic Republic of Congo, and cholera worldwide. Journal Of Infection In Developing Countries 5 (10), 688–691.

Kermack, W., Kendrick, A., 1927. Contributions to the mathematical theory of epidemics (Part I). Proceedings of the Royal Society A 115, 700–721.

King, A., Ionides, E., Pascual, M., Bouma, M., 2008. Inapparent infections and cholera dynamics. Nature 454, 877–880.

Koelle, K., 2009. The impact of climate on the disease dynamics of cholera. Clinical Microbiology And Infection 15, 29–31.

Koelle, K., Rodo, X., Pascual, M., Yunus, M., Mustafa, G., 2005. Refractory periods and climate forcing in cholera dynamics. Nature 436, 696–700.

Kogure, K., Simidu, U., Taga, N., 1979. A tentative direct microscopic method for counting living marine bacteria. Canadian Journal of Microbiology 25, 415–420.

Kuznetsov, Y., 1995. Elements of Applied Bifurcation Theory. New York, US: Springer-Verlag.

Kuznetsov, Y., Piccardi, C., 1994. Bifurcation-analysis of periodic SEIR and SIR epidemic models. Journal of Mathematical Biology 32 (2), 109–121.

Laio, F., Porporato, A., Ridolfi, R., Rodriguez-Iturbe, I., 2001. Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress. probabilistic soil moisture dynamics. Advances in Water Resources 24, 707–723.

Lin, J., Andreasen, V., Casagrandi, R., Levin, S., 2003. Traveling waves in a model of influenza A drift. Journal of theoretical Biology 222, 437–445.

Lipp, E., Huq, A., Colwell, R., 2002. Effects of global climate on infectious disease: The cholera model. Clinical Microbiology Reviews 15 (4), 757+.

Luque Fernandez, M. A., Bauernfeind, A., Diaz Jimenez, J., Linares Gil, C., El Omeiri, N., Herrera Guibert, D., 2009. Influence of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia, 2003-2006: analysis of a time series. Transactions Of The Royal Society Of Tropical Medicine And Hygiene 103 (2), 137–143.

Mari, L., Bertuzzo, E., Righetto, L., Gatto, M., Casagrandi, R., Rodriguez-Iturbe, I., Rinaldo, A., 2012a. Modelling cholera epidemics: the role of waterways, human mobility and sanitation. Journal of the Royal Society Interface 9, 376–388.

Mari, L., Bertuzzo, E., Righetto, L., Gatto, M., Casagrandi, R., Rodriguez-Iturbe, I., Rinaldo, A., 2012b. On the role of human mobility in the spread of cholera epidemics: towards an epidemiological movement ecology. Ecohydrology 5, 531–540.

Medvinsky, A., Petrovskii, S., Tikhonova, I., Malchow, H., Li, B., 2002. Spatiotemporal complexity of plankton and fish dynamics. SIAM Review 44 (3), 311–370.

Merrell, D. S., al., 2002. Host-induced epidemic spread of the cholera bacterium. Nature 417, 642–645.

Mukandavire, Z., Smith, D. L., Morris, Jr., J. G., 2013. Cholera in Haiti: Reproductive numbers and vaccination coverage estimates. Scientific Reports 3.

Murray, J., 2002. Mathematical Biology. New York, US: Springer-Verlag.

Nalin, D., 1979. Cholera, copepods and chitinase. Lancet 30, 958.

Neilan, R. L. M., Schaefer, E., Gaff, H., Fister, K. R., Lenhart, S., 2010. Modeling optimal intervention strategies for cholera. Bulletin of Mathematical Biology 72 (8), 2004–2018.

Nkoko, D. B., Giraudoux, P., Plisnier, P.-D., Tinda, A. M., Piarroux, M., Sudre, B., Horion, S., Tamfum, J.-J. M., Ilunga, B. K., Piarroux, R., 2011. Dynamics of Cholera Outbreaks in Great Lakes Region of Africa, 1978-2008. Emerging Infectious Diseases 17 (11), 2026–2034.

PAHO, 2010. Eoc situation report #17. Tech. rep., Pan-American Health Organization, Regional Office of the World Health Organization, online at http://new.paho.org/hq/images/Atlas_IHR/CholeraHispaniola/atlas.html.

Pascual, M., Bouma, M. J., Dobson, A. P., 2002. Cholera and climate: revisiting the quantitative evidence. Microbes and Infection 4, 237–245.

Pascual, M., Chaves, L. F., Cash, B., Rodó, X., Yunus, Md., 2008. Predicting endemic cholera: The role of climate variability and disease dynamics. Climate Research 36, 131–140.

Pascual, M., Rodo, X., Ellner, S., Colwell, R., Bouma, M., 2000. Cholera dynamics and El Nino-Southern oscillation. Science 289, 1766–1769.

Piarroux, R., Barrais, R., Faucher, B., Haus, R., Piarroux, M., Gaudart, J., Magloire, R., Raoult, D., 2011. Understanding the cholera epidemic, Haiti. Emerging Infectious Diseases 17, 1161–1168.

Porporato, A., Vico, G., Fay, P., 2006. Superstatistics of hydro-climatic fluctuations and interannual ecosystem productivity. Geophysical Research Letters 33, L15402.

Priestley, C. H. B., Taylor, R. J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly Weather Review 100 (2), 82–92.

Righetto, L., Bertuzzo, E., Casagrandi, R., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., 2011. Modelling Human Movement in Cholera Spreading along Fluvial Systems. Ecohydrology 4 (1), 49–55.

Righetto, L., Casagrandi, R., Bertuzzo, E., Mari, L., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., 2012. The role of aquatic reservoir fluctuations in long-term cholera patterns. Epidemics 4, 33–42.

Rinaldo, A., Bertuzzo, E., Mari, L., Righetto, L., Blokesch, M., Gatto, M., Casagrandi, R., Murray, M., Vesenbeckh, S., Rodriguez-Iturbe, I., 2012. Early predictions of the 2010-2011 Haiti cholera outbreak: Assessments and perspectives. Proceedings of the National Academy of Science of the United States of America 109 (17), 6602–6607.

Rinaldo, A., Beven, K., Nicotina, L., Davies, J., Fiori, A., Russo, D., Botter, G., 2011. Catchment travel time distributions and water flow in soils. Water Resources Research 47, W07537.

Rinaldo, A., Iturbe, I. R., 1996. Geomorphological Theory of the Hydrological Response. Hydrological Processes 10 (6), 803–829.

Rodo, X., Pascual, M., Fuchs, G., Faruque, A., 2002. ENSO and Cholera: A Nonstationary Link Related to Climate Change? Proceedings Of The National Academy Of Sciences Of The United States Of America 99 (20), 12901–12906.

Rodriguez-Iturbe, I., Cox, D., Eagleson, P., 1986. Spatial Modeling Of Total Storm Rainfall. Proceedings of the Royal Society of London Series A-Mathematical and Physical Sciences 403 (1824), 27–50.

Rodriguez-Iturbe, I., Porporato, A., 2004. Ecohydrology of Water-Controlled Ecosystems: Soil Moisture and Plant Dynamics. Cambridge University Press, Cambridge, UK.

Rodriguez-Iturbe, I., Rinaldo, A., 1997. Fractal River Basins. Chance and Self-Organization. Cambridge University Press, New York, US.

Ruiz-Moreno, D., Pascual, M., Bouma, M., Dobson, A., Cash, B., 2007. Cholera seasonality in Madras (1901-1940): Dual role for rainfall in endemic and epidemic regions. Ecohealth 4 (1), 52–62.

Sack, D., 2011. How many cholera deaths can be averted in Haiti? Lancet 377, 1214–1216.

Sack, R., Siddique, A., Longini, I., Nizam, A., Yunus, M., Islam, M., Morris, J., Ali, A., Huq, A., Nair, G., Qadri, F., Faruque, S., Sack, D., Colwell, R., 2003. A 4-year study of the epidemiology of *Vibrio cholerae* in four rural areas of Bangladesh. Journal of Infectious Diseases 187 (1), 96–101.

Scheffer, M., 1991. Fish and nutrients interplay determines algal biomass: a minimal model. Oikos 62, 271–282.

Scheffer, M., Rinaldi, S., Gragnani, A., Mur, L. R., Vannes, E. H., 1997a. On the Dominance of Filamentous Cyanobacteria in Shallow, Turbid Lakes. Ecology 78 (1).

Scheffer, M., Rinaldi, S., Kuznetsov, Y., 2000. Effects of fish on plankton dynamics: a theoretical analysis. Canadian Journal of Fish and Aquatic Sciences 57, 1208–1219.

Scheffer, M., Rinaldi, S., Kuznetsov, Y., Vannes, E., 1997b. Seasonal dynamics of Daphnia and algae explained as a periodically forced predator-prey system. OIKOS 80 (3), 519–532.

Tamplin, M., Gauzens, A., Huq, A., Sack, D., Colwell, R., 1990. Attachment of *Vibrio cholerae* serogroup-O1 to zooplankton and phytoplankton of Bangladesh waters. Applied And Environmental Microbiology 56 (6), 1977–1980.

Tappero, Tauxe, 2011. Lessons learned during public health response to cholera epidemic in haiti and the dominican republic. Tech. rep., United States Center for Disease Control.

ter Braak, C. J. F., Vrugt, J. A., 2008. Differential Evolution Markov Chain with snooker updater and fewer chains. Statistics and Computing 18, 435–446.

Tien, J. H., Earn, D. J. D., 2010. Multiple Transmission Pathways and Disease Dynamics in a Waterborne Pathogen Model. Bulletin Of Mathematical Biology 72 (6), 1506–1533.

Tuite, A., Tien, J., Eisenberg, M., Earn, D., Ma, J., Fisman, D., 2011. Cholera epidemic in Haiti, 2010: using a transmission model to explain spatial spread of disease and identify optimal control interventions. Annals of Internal Medicine 154, 593–601.

Vital, M., Fuechslin, H. P., Hammes, F., Egli, T., 2007. Growth of *Vibrio cholerae* O1 Ogawa Eltor in freshwater. Microbiology 153 (Part 7), 1993–2001.

Vital, M., Stucki, D., Egli, T., Hammes, F., 2010. Evaluating the growth potential of pathogenic bacteria in water. Applied And Environmental Microbiology 76 (19), 6477–6484.

Waldor, M. K., Hotez, P. J., Clemens, J. D., 2010. A National Cholera Vaccine Stockpile - A New Humanitarian and Diplomatic Resource. New England Journal Of Medicine 363 (24), 2279–2282.

Walton, D., Ivers, L., 2011. Responding to cholera in post-earthquake Haiti. The New England Journal of Medicine 364, 3–5.

Wang, K., Li, Z., Cribb, M., 2006. Estimation of evaporative fraction from a combination of day and night land surface temperatures and NDVI: A new method to determine the Priestley-Taylor parameter. REMOTE SENSING OF ENVIRONMENT 102 (3-4), 293–305.

WHO, 2011. Annual report on cholera, available online at http://www.who.int/cholera/statistics/en/. Tech. rep., World Health Organization.

Worden, A., Seidel, M., Smriga, S., Wick, A., Malfatti, F., Bartlett, D., Azam, F., 2006. Trophic regulation of *Vibrio cholerae* in coastal marine waters. Environmental Microbiology 8 (1), 21–29.

Xu, R., Wunsch, D., 2005. Survey of clustering algorithms. IEEE Transactions On Neural Networks 16 (3), 645–678.

Personal information

Name / Surname	Lorenzo Righetto
Address	Laboratory of Ecohydrology, IIE/ENAC Ecole Polytechnique Fédérale de Lausanne Station 2 - EPFL, 1015 Lausanne, Switzerland
Telephone	+41.021.693.6348
Fax	+41.021.693.3739
Email	lorenzo.righetto@epfl.ch
Web site	http://people.epfl.ch/lorenzo.righetto
Nationality	Italian
Date of birth	April 22, 1984
Current position	*Doctoral Student* Laboratory of Ecohydrology, IIE/ENAC, Ecole Polytechnique Fédérale de Lausanne

Research statement

Research interests

I graduated in environmental engineering (with specific background in information technology applied to the management of natural resources) with a thesis on a metapopulation model based on allometry relationships between animal body mass and model parameters. In my current work in the Laboratory of Ecohydrology at EPFL, I integrated my knowledge in system theory and ecology with hydrology and network theory, applying them to epidemiological models of waterborne diseases (cholera in particular). My research there has a high degree of interdisciplinarity, as it lies at the boundary between the ecology of populations and the study of transport dynamics of particles (pathogens, in this case) in complex networks, describing hydrologic and/or human mobility.

Research activities

During my PhD I did preliminary studies of the role of human mobility on the spreading of waterborne diseases and of hydroclimatological fluctuations on purely theoretical frameworks ([a1,a2]). I then participated in extracting the Haiti hydrologic network and in the calibration of models for the large cholera outbreak that struck the country in recent years ([a3]), also developing a model for the prediction of rainfall and the consequent resurging of the disease caused by worsening of sanitation conditions ([a4]). Part of my PhD work was also spent conducting a field campaign in Bangladesh, where we applied new microbiological techniques (i.e. flow cytometry) to the study of the ecology of the cholera bacterium in an endemic setting ([a5]) and we studied the relation of bacterial concentration with hydroclimatological variables during the span of an year ([a6]).

In addition to scientific relevance, some of these topics have clear social and/or economic implications. This is the case, for instance, of building models for cholera epidemics (like the one that stroke Haiti in 2010 and has been responsible of nearly 650,000 reported cases and more than 8,000 deaths to date [a16]). Mathematical models of human epidemics or biological invasions are key tools to understand drivers and controls of their spread, and to design effective control measures.

Current research and perspectives

I recently supervised a Master thesis on the study of hydroclimatological forcings on the cholera patterns in Bangladesh, based on the Bengalese hydrologic network. This study could lead to a better understanding of the influence of hydrology and climate on cholera outbreaks in a spatially explicit framework. I plan also to continue working on metapopulations and apply the model ([a7]) to real cases and upgrade it to a metacommunity - rather than metapopulation - framework.

Employment

Feb 2009 – Doctoral Student, Laboratoire d'Ecohydrologie (ECHO), École Polytechnique Fédérale de Lausanne (EPFL)

Education

Cursus studiorum

Oct 2006 – Dec 2008 MSc, Environmental and Land Engineering, I Facoltà di Ingegneria, PoliMi

Thesis title: "Effetti della taglia e del movimento attivo sulla persistenza in metapopolazioni animali" (*The effects of size and active dispersal on persistence in animal metapopulations*). Advisor: Prof. R. Casagrandi; co-advisor: Prof. M. Gatto. Mark: 110/110 *cum laude*.

Sept 2003 – Sep 2006 BSc, Environmental and Land Engineering, I Facoltà di Ingegneria, PoliMi

Thesis title: "Impatti del cambiamento climatico globale sull'ecologia del gallo forcello nell'area del parco nazionale del Veglia-Devero" (*Impacts of global climate change on the ecology of the black grouse in the area of Veglia-Devero National park*). Advisor: Prof. M.Gattto (DEI, PoliMi). Mark: 104/110

Awards

Nov 2010 LATSIS Poster Prize, EPFL

Scientific memberships

2010– Member of the American Geophysical Society (AGU) and European Geosciences Society (EGU)

Involvement in research programs

2012 – Project "Dynamics and controls of large-scale cholera outbreaks" (FNS project 138104, principal investigator: Prof. A. Rinaldo)

2009 – Project "River networks as ecological corridors for biodiversity, populations and waterborne disease" (ERC project 227612, principal investigator: Prof. A. Rinaldo)

2009 – Project "Hydrologic controls on ecological processes: river networks as corridors for biodiversity, populations and pathogens of water-borne diseases" (FNS project 124930, principal investigator: Prof. A. Rinaldo)

Academic services

2009 – Acting as reviewer for *Ecological Indicators*, *Weather, Climate and Society*, *Journal of Insect Science*

Teaching activities

A.Y. 2012–2013 Teaching assistant (20 hours) for the course "Hydrology" (Prof. A. Rinaldo, ECHO, EPFL), Bachelor Degree in Environmental Sciences and Engineering, EPFL

A.Y. 2011–2012 Lecturer (4 hours) for the course "Spatial dynamics in Biology" (Prof. M. Gatto), PhD Program in Information Technology, DEI, PoliMi

2010–2012 Co-advisor of BSc and MSc theses, Corso di Laurea (Magistrale) in Ingegneria per l'Ambiente e il Territorio, PoliMi

120

Publications

Journal articles

[a09] M. Gatto, L. Mari, E. Bertuzzo, R. Casagrandi, **L. Righetto**, I. Rodriguez-Iturbe, A. Rinaldo (2012)
Generalized reproduction numbers and the prediction of patterns in waterborne disease
Proceedings of the National Academy of Sciences of the USA, 109:19703–19708. doi:10.1073/pnas.1217567109

[a08] L. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo (2012)
On the role of human mobility in the spread of cholera epidemics: towards an epidemiological movement ecology
Ecohydrology, 5:531–540. doi:10.1002/eco.262

[a07] A. Rinaldo, E. Bertuzzo, L. Mari, **L. Righetto**, M. Blokesch, M. Gatto, R. Casagrandi, M. Murray, S. Vesenbeckh, I. Rodriguez-Iturbe (2012)
Reassessment of the 2010–2011 Haiti cholera outbreak and rainfall-driven multi-season projections
Proceedings of the National Academy of Sciences of the USA, 109:6602–6607. doi:10.1073/pnas.1203333109

[a06] E. Bertuzzo, L. Mari, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo (2012)
Hydroclimatology of dual-peak annual cholera incidence: insights from a spatially explicit model
Geophysical Research Letters, 39:L05403. doi:10.1029/2011GL050723

[a05] **L. Righetto**, R. Casagrandi, E. Bertuzzo, L. Mari, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo (2012)
The role of aquatic reservoir fluctuations in long-term cholera patterns
Epidemics, 4:33–42. doi:10.1016/j.epidem.2011.11.002

[a04] L. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo (2012)
Modeling cholera epidemics: the role of waterways, human mobility and sanitation
Journal of the Royal Society Interface, 9:376–388. doi:10.1098/rsif.2011.0304

[a03] A. Rinaldo, M. Blokesch, E. Bertuzzo, L. Mari, **L. Righetto**, M. Murray, M. Gatto, R. Casagrandi, I. Rodriguez-Iturbe (2011)
A transmission model of the 2010 cholera epidemic in Haiti
Annals of Internal Medicine, 155:403–404. doi:10.1059/0003-4819-155-6-201109200-00018

[a02] E. Bertuzzo, L. Mari, **L. Righetto**, M. Gatto, R. Casagrandi, M. Blokesch, I. Rodriguez-Iturbe, A. Rinaldo (2011)
Prediction of the spatial evolution and effects of control measures for the unfolding Haiti cholera outbreak
Geophysical Research Letters, 38:L06403. doi:10.1029/2011GL046823

[a01] **L. Righetto**, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo (2011)
Modelling human movement in cholera spreading along fluvial systems
Ecohydrology, 4:49-55. doi:10.1002/eco.122

Submitted

L. Righetto, E. Schild, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo (2012)
Rainfall mediations in the spreading of epidemic cholera

121

M. Gatto, L. Mari, E. Bertuzzo, R. Casagrandi, **L. Righetto**, I. Rodriguez-Iturbe, A. Rinaldo (2012)
Spatially explicit conditions for waterborne pathogen invasion

International conferences

[i19] A. Rinaldo, M. Gatto, L. Mari, R. Casagrandi, **L. Righetto**, E. Bertuzzo, I. Rodriguez-Iturbe
Spatially explicit models, generalized reproduction numbers and the prediction of patterns of waterborne disease
AGU Fall Meeting, San Francisco (CA), USA, December 3–7 2012

[i18] **L. Righetto**, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Rainfall-driven epidemic cholera: hydrologic controls on water-borne disease and multi-season projections
AGU Fall Meeting, San Francisco (CA), USA, December 3–7 2012

[i17] L. Mari, M. Gatto, E. Bertuzzo, R. Casagrandi, **L. Righetto**, I. Rodriguez-Iturbe, A. Rinaldo
A novel spatially-explicit condition for the onset of waterborne diseases in complex environments
AGU Fall Meeting, San Francisco (CA), USA, December 3–7 2012

[i16] E. Bertuzzo, L. Mari, **L. Righetto**, M. Gatto, R. Casagrandi, I. Rodriguez-Iturbe, A. Rinaldo
Hydroclimatology of dual-peak annual cholera incidence: insights from a spatially explicit model
AGU Fall Meeting, San Francisco (CA), USA, December 3–7 2012

[i15] **L**. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, I. Rodriguez-Iturbe, A. Rinaldo, M. Gatto
An ecohydrological model of cholera dynamics
SIDISA 2012, Milan, Italy, June 26–29 2012

[i14] **L. Righetto**, S. Islam, Z.H. Mahmud, E. Bertuzzo, L. Mari, R. Casagrandi M. Gatto, I. Rodriguez-Iturbe, M. Blokesch, A. Rinaldo
Presence and viability of *V. cholerae* in the waters of rural Bangladesh (Matlab area)
EGU General Assembly, Vienna, Austria, April 22–27 2012

[i14] E. Bertuzzo, L. Mari, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
An epidemic model for the future progression of the current Haiti cholera epidemic
EGU General Assembly, Vienna, Austria, April 22–27 2012

[i13] E. Bertuzzo, L. Mari, **L. Righetto**, M. Gatto, R. Casagrandi, I. Rodriguez-Iturbe, A. Rinaldo
A spatially explicit model for the future progression of the current Haiti cholera epidemic
AGU Fall Meeting, San Francisco (CA), USA, December 5–9 2011

[i12] **L. Righetto**, R. Casagrandi, E. Bertuzzo, L. Mari, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Modeling environmental drivers of cholera seasonality in Bengal endemic areas
European Conference on Ecological Modeling, Riva del Garda, Italy, May 30–June 2 2011

[i11] A. Rinaldo, E. Bertuzzo, L. Mari, **L. Righetto**, M. Gatto, R. Casagrandi, I. Rodriguez-Iturbe
Reactive transport on multiscale networks: controls and drivers of large-scale cholera outbreaks
EGU General Assembly, Vienna, Austria, April 3–8 2011

[i10] L. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Hydrological transport, human mobility and cholera epidemics: a spatially explicit modeling approach
EGU General Assembly, Vienna, Austria, April 3–8 2011

[i09] E. Bertuzzo, L. Mari, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
A spatially distributed model for the future evolution of the current Haiti cholera outbreak
EGU General Assembly, Vienna, Austria, April 3–8 2011

[i08] A. Rinaldo, E. Bertuzzo, L. Mari, **L. Righetto**, M. Gatto, R. Casagrandi, I. Rodriguez-Iturbe
On spatially explicit models of cholera epidemics: hydrologic controls, environmental drivers, human-mediated transmissions
AGU Fall Meeting, San Francisco (CA), USA, December 13–17 2010

[i07] **L. Righetto**, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
The role of the aquatic reservoir in long-term cholera dynamics
AGU Fall Meeting, San Francisco (CA), USA, December 13–17 2010

[i06] L. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Human mobility patterns and cholera epidemics: a spatially explicit modeling approach
AGU Fall Meeting, San Francisco (CA), USA, December 13–17 2010

[i05] E. Bertuzzo, L. Mari, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Hydroclimatology of dual peak cholera incidence in Bengal region: inferences from a spatial explicit model
AGU Fall Meeting, San Francisco (CA), USA, December 13–17 2010

[i04] **L. Righetto**, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Ecohydrological drivers of cholera spreading along fluvial systems
Water and health: where science meets policy, Chapel Hill (NC), USA, October 25–26 2010

[i03] **L. Righetto**, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Ecohydrologic drivers and controls for cholera epidemics
LATSIS Symposium 2010, Lausanne, Switzerland, October 17–20 2010

[i02] E. Bertuzzo, L. Mari, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo.
Spreading of cholera through surface water
LATSIS Symposium 2010, Lausanne, Switzerland, October 17–20 2010

[i01] **L. Righetto**, E. Bertuzzo, L. Mari, R. Casagrandi, M. Gatto, A. Rinaldo
Hydroclimatological and anthropogenic drivers for cholera spreading
European Geosciences Union General Assembly 2010, Wien, Austria, May 2–7 2010

National conferences

[n02] L. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, I. Rodriguez-Iturbe, A. Rinaldo, M. Gatto
A novel condition for the onset of waterborne diseases in complex environments
XXII Congresso della SItE, Alessandria, September 10–13 2012

[n01] L. Mari, E. Bertuzzo, **L. Righetto**, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, A. Rinaldo
Hydrologic transport, human mobility and the spread of cholera epidemics: insights from a spatially explicit model
XXI Congresso della SItE, Palermo, October 3–6 2011

Other selected publications

[o02] **L. Righetto** (2008)
Effetti della taglia e del movimento attivo sulla persistenza in metapopolazioni animali
MSc thesis in Environmental Engineering, PoliMi, A.Y. 2007/2008
Advisor: Prof. R. Casagrandi; co-advisor: Prof. M. Gatto

[o01] **L. Righetto** (2006)
Impatti del cambiamento climatico globale sull'ecologia del gallo forcello nell'area del parco nazionale del Veglia-Devero
BSc thesis in Environmental Engineering, PoliMi, A.Y. 2005/2006
Advisor: Prof. M. Gatto

Lausanne, April 10, 2013