

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/255734900>

Spatially Explicit Conditions for Waterborne Pathogen Invasion

ARTICLE in THE AMERICAN NATURALIST · SEPTEMBER 2013

Impact Factor: 3.83 · DOI: 10.1086/671258 · Source: PubMed

CITATIONS

10

READS

109

7 AUTHORS, INCLUDING:



Marino Gatto

Politecnico di Milano

153 PUBLICATIONS 1,689 CITATIONS

SEE PROFILE



Renato Casagrandi

Politecnico di Milano

69 PUBLICATIONS 1,273 CITATIONS

SEE PROFILE



Lorenzo Righetto

Politecnico di Milano

24 PUBLICATIONS 212 CITATIONS

SEE PROFILE



Andrea Rinaldo

École Polytechnique Fédérale de Lausanne

165 PUBLICATIONS 3,884 CITATIONS

SEE PROFILE

Spatially Explicit Conditions for Waterborne Pathogen Invasion

Marino Gatto,^{1,2,*} Lorenzo Mari,^{1,2} Enrico Bertuzzo,² Renato Casagrandi,¹ Lorenzo Righetto,² Ignacio Rodriguez-Iturbe,³ and Andrea Rinaldo^{2,4}

1. Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milan, Italy; 2. Laboratory of Ecohydrology, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; 3. Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey 08544; 4. Dipartimento di Ingegneria Civile, Edile ed Ambientale, Università di Padova, 35131 Padua, Italy

Submitted June 28, 2012; Accepted April 11, 2013; Electronically published July 18, 2013

Online enhancements: appendixes, supplementary tables and figures, video.

ABSTRACT: Waterborne pathogens cause many possibly lethal human diseases. We derive the condition for pathogen invasion and subsequent disease outbreak in a territory with specific, space-inhomogeneous characteristics (hydrological, ecological, demographic, and epidemiological). The criterion relies on a spatially explicit model accounting for the density of susceptible and infected individuals and the pathogen concentration in a network of communities linked by human mobility and the water system. Pathogen invasion requires that a dimensionless parameter, the dominant eigenvalue of a generalized reproductive matrix J_0 , be larger than unity. Conditions for invasion are studied while crucial parameters (population density distribution, contact and water contamination rates, pathogen growth rates) and the characteristics of the networks (connectivity, directional transport, water retention times, mobility patterns) are varied. We analyze both simple, prototypical test cases and realistic landscapes, in which optimal channel networks mimic the water systems and gravitational models describe human mobility. Also, we show that the dominant eigenvector of J_0 effectively portrays the geography of epidemic outbreaks, that is, the areas of the studied territory that will be initially affected by an epidemic. This is important for planning an efficient spatial allocation of interventions (e.g., improving sanitation and providing emergency aid and medicines).

Keywords: spatially explicit models, hydrological transport, human mobility, stability analysis, incidence map, dominant eigenvalue.

Introduction

Diarrheal waterborne diseases are caused by pathogenic microorganisms that are transmitted when contaminated water (or food contaminated by water) is consumed and are thus directly or indirectly hydrologically controlled. The pathogens include protozoa (e.g., *Entamoeba histolytica*), bacteria (e.g., *Vibrio cholerae* and *Shigella dysenteriae*), and viruses (e.g., *Rotavirus gastroenteritis*). These diseases are among the major causes of death in the world

(World Health Organization 2008). Cholera is possibly the best-known lethal diarrheal disease, although other infections, such as rotavirus, cause more casualties (about 500,000 in 2004 alone, according to the World Health Organization).

To understand, control, and predict epidemics, development of appropriate mechanistic models is fundamental, as they can play an important role in devising appropriate intervention measures (Bertuzzo et al. 2011; Tuite et al. 2011; Mari et al. 2012a; Rinaldo et al. 2012). After the seminal model proposed by Capasso and Paveri-Fontana (1979), Codeço (2001) developed a system of three ordinary differential equations where, in addition to the compartments of susceptible (S) and infected (I) that characterize traditional microparasitological models, one equation accounts for the population dynamics of bacteria (B) in the water reservoir. Subsequently, refinements have been introduced to account for, for example, hyperinfectivity (Hartley et al. 2006) or the prey-predator interaction between bacteria and their specific phages (Jensen et al. 2006). Also, seasonal and climate forcings have been analyzed (Colwell 1996; Lipp et al. 2002; Pascual et al. 2002; Koelle et al. 2005; Akanda et al. 2011; Reiner et al. 2012) to include the influence of different hydrometeorological regimes on the time course of the disease. All these models, however, have relied on mean-field approximations, without considering that very often epidemics display spatially inhomogeneous patterns. The spatial component must be explicitly accounted for to develop realistic policies, which have to be specifically linked to the territory where public-health measures are implemented.

While some diarrheal infections are basically linked to fecal-oral direct transmission (e.g., rotavirus) and their dynamics can thus be modeled as SIR (susceptible-infected-recovered) systems (Pitzer et al. 2009), the spatial spread of waterborne diseases is mediated primarily by the hydrological network (Bertuzzo et al. 2008, 2010, 2012; Akanda et al. 2009). However, in many cases, human mo-

* Corresponding author; e-mail: gatto@elet.polimi.it.

bility plays an additional important role. In fact, infected individuals are often asymptomatic—for example, about 75% for cholera (World Health Organization 2010) and 80% for amoebiasis caused by *E. histolytica* (Guerrant 1986)—and thus can inadvertently spread the pathogens by moving to another community. Similarly, susceptible individuals can be infected by pathogens far from their home sites when traveling or commuting. While metapopulation dynamics was incorporated into traditional SI or SIR models many years ago (Bolker and Grenfell 1995; Swinton et al. 1998; Gog et al. 2002; Arino et al. 2005), the addition of human mobility to spatial models of waterborne diseases is more recent. This has been accomplished via either diffusion-based (Righetto et al. 2011) or gravity-like models (Bertuzzo et al. 2011; Chao et al. 2011; Tuite et al. 2011; Gatto et al. 2012; Mari et al. 2012a, 2012b; Rinaldo et al. 2012). The wide availability, via geographic information systems, of spatial data on hydrology, road networks, population distribution, and sanitation makes these models applicable to specific situations in each country and for each disease to be studied.

Here we extend, generalize, and make more explicit the approach outlined by Gatto et al. (2012). Specifically, in this work we explicitly account for spatial heterogeneities of both hydrological conditions and pathogen ecology, which were not considered in previous approaches. The latter extension, in particular, is of paramount importance in the study of pathogen invasion conditions at regional spatial scales characterized by inhomogeneous climatic and ecological conditions (temperature, pH, salinity, nutrient availability). In addition, most of the results described here refer to a spatial setting much more realistic than those reported elsewhere (Gatto et al. 2012). To portray hydrological connectivity, in fact, here we make use of so-called optimal channel networks, which are mathematical objects whose scaling forms are de facto indistinguishable from those of real river networks (Rinaldo et al. 1992; Rodriguez-Iturbe et al. 1992). We carry out a systematic analysis to determine the conditions under which pathogen invasion and ensuing disease outbreak are possible within a specific territory with certain demographic, epidemiological, climatic, and socioeconomic characteristics. The main aim of this work is to find waterborne pathogen invasion conditions while accounting for (1) the spatial inhomogeneity of population density, sewage water disposal and treatment, and the availability of clean, safe water, as well as that of the hydrological system and pathogen ecology and (2) the spread of pathogens via different pathways (typically, the hydrological network and the human-mobility network). Conditions for invasion are studied under varying model parameters (population density distribution, contact and water contamination rates, pathogen growth rates) and network characteristics (connectivity, directional transport,

water retention times, mobility patterns). The second aim of the study is analysis of the geography of epidemic outbreaks, that is, determining the areas of the studied territory that will be initially affected by the epidemic, if invasion conditions are met. This is especially important because knowledge of an outbreak epicenter can assist in an efficient allocation of interventions (e.g., improving sanitation, providing emergency aid and medicines, enforcing quarantine), which can minimize the extent of a disease outbreak.

The Model

Our starting point is a generalization of a spatially explicit metacommunity model that includes both the hydrological and the human-mobility network (Gatto et al. 2012; Mari et al. 2012a). Although parameterizable to particular cases (see below), the model of the network of human communities (metacommunity) possibly subject to infection by waterborne pathogens is general and can be summarized as follows. Let $S_i(t)$ and $I_i(t)$ be the local abundances of susceptible and infected individuals, respectively, in each node i of the hydrological network at time t , and let $B_i(t)$ be the concentration of the pathogens (i.e., bacteria, viruses, or protozoa) in water. The hydrological network can be a river basin, human-made water distribution and/or sewage systems, or both, and network nodes can be cities, towns, or villages. Epidemiological dynamics and pathogen transport over the hydrological and human-mobility networks are described by the following set of ordinary differential equations:

$$\begin{aligned} \frac{dS_i}{dt} &= \mu(H_i - S_i) - \left[(1 - m_s)\beta_i f(B_i) + m_s \sum_{j=1}^n Q_{ij}\beta_j f(B_j) \right] S_i, \\ \frac{dI_i}{dt} &= \left[(1 - m_s)\beta_i f(B_i) + m_s \sum_{j=1}^n Q_{ij}\beta_j f(B_j) \right] S_i \\ &\quad - (\gamma + \mu + \alpha)I_i, \\ \frac{dB_i}{dt} &= -(\mu_{B_i} + l_i)B_i + \sum_{j=1}^n l_j p_{ji} \frac{W_j}{W_i} B_j \\ &\quad + \frac{p_i}{W_i} \left[(1 - m_i)I_i + \sum_{j=1}^n m_i Q_{ji} I_j \right]. \end{aligned} \quad (1)$$

The evolution of the susceptible compartment (first equation of model [1]) is a balance between population demography and infections due to contact with the pathogen. The host population, if uninfected, is assumed to be at demographic equilibrium H_i (the size of the i th local community), with μ being the human mortality rate. The parameter β_i represents the site-dependent rate of exposure to contaminated water, and $f(B_i)$ is the probability of becoming infected because of the exposure to a concentra-

tion B_i of pathogens. The dose-response function $f(B_i)$ (Zmirou-Navier et al. 2006) usually saturates for high pathogen concentrations. Beta-Poisson, exponential, and hyperbolic are the most-used models. In accordance with much of the literature on cholera initiated by Codeço (2001), hereafter we use the hyperbolic function $f(B_i) = B_i/(K + B_i)$ (with K being the half-saturation constant). This does not represent a loss of generality, because disease outbreak is linked to the linear approximation of $f(B_i)$ at low pathogen concentrations, as shown below. Although originally devised for cholera, this transmission mechanism could be extended to other waterborne diseases. Also, if the network is sufficiently fine grained, the model comprises the so-called human-to-human transmission pathway as well (Weil et al. 2009; Tuite et al. 2011). In this kind of transmission, in fact, infection rarely occurs through direct contact, as in airborne diseases, but rather occurs via contact by susceptibles with water (or food) contaminated by infected individuals living in close proximity and thus utilizing the same water reservoir.

The dynamics of the infected compartment (second equation of model [1]) is a balance between newly infected individuals and losses due to recovery or natural/pathogen-induced mortality, with γ and α being the rates of recovery and mortality due to the disease, respectively. Waterborne diseases generally confer at least temporary immunity (in cholera, about 2–5 years; Clemens et al. 1990; Koelle et al. 2005). Therefore, we do not consider the dynamics of recovered individuals, because our goal is to assess pathogen dynamics during the initial conditions of an outbreak. The evolution of the local concentration of pathogens that live free in the aquatic environment (third equation of model [1]) assumes that bacteria, viruses, or protozoa are released in water (e.g., excreted) by infected individuals and immediately diluted in a well-mixed local water reservoir of volume W_i at a site-specific rate p_i , depending on local sanitation conditions. Free-living pathogens are also assumed to die at a constant, site-dependent rate $\mu_{B,i}$, which can widely vary according to inhomogeneous climatic and ecological conditions.

As regards the hydrological transport, the spread of pathogens over the river network is described as a biased random-walk process on an oriented graph (Bertuzzo et al. 2007). Specifically, we assume that pathogens can move, at a rate l_{ij} , from node i to node j of the hydrological network with a probability P_{ij} . The rate l_{ij} depends on both downstream advection and other possible pathogen transport pathways along the hydrological network, for example, short-range distribution of water for consumption or irrigation or pathogen attachment to phyto- and zooplankton. This rate can be, therefore, quite variable from node to node. The transport process is assumed to be conservative, that is, $\sum_{j=1}^n P_{ij} = 1$. Possible topological

structures for the hydrological network range from simple one-dimensional lattices to more realistic mathematical characterizations, such as Peano basins (as in Gatto et al. 2012), optimal channel networks (Rinaldo et al. 1992; Rodriguez-Iturbe et al. 1992; see below for details), or real river systems (e.g., Bertuzzo et al. 2008; Mari et al. 2012a; Rinaldo et al. 2012). As for the human-mobility network, we assume that the nodes of this second layer correspond to those of the hydrological layer, whereas edges are defined by connections among communities. We also assume that susceptible and infected individuals can undertake short-term trips from the communities where they live toward other nodes. While traveling or commuting, susceptible individuals can be exposed to pathogens and return as infected carriers to the community where they usually live. Similarly, infected hosts can disseminate pathogens away from their home community. It should be remembered that in many cases infected individuals are asymptomatic and thus are not barred from their usual activities by the presence of the pathogen in their intestine.

Human mobility patterns are defined according to a connection matrix in which individuals leave their original node (say i) with an infection-dependent probability (m_s for susceptibles and m_i for infecteds, usually with $m_s \geq m_i$), reach their target location (say j) with a probability Q_{ij} , and then come back to node i . Topological and transition probability structures for human mobility patterns used in epidemiology are quite varied (Eubank et al. 2004; Riley 2007). They can be based on suitable measures of node-to-node distance, as in gravity models (Erlander and Stewart 1990), on the actual transportation network (roads/airports/waterways; Colizza et al. 2006), on proxies for human movement (e.g., mobile phone data; Bengtsson et al. 2011), or on paradigmatic models based on conceptually different interactions, such as Erdős-Rényi random graphs (Erdős and Rényi 1959), scale-free networks (Albert and Barabási 2002), and small-world-like graphs (Watts and Strogatz 1998).

Pathogen Invasion Conditions

The derivation of the conditions for pathogen invasion in the network system can be outlined as follows. The nonlinear model of the waterborne disease at the global scale is always characterized by the disease-free equilibrium \mathbf{X}_0 , that is, a state where $S_i = H_i$, $I_i = 0$, and $B_i = 0$ for all $i = 1, \dots, n$. If \mathbf{X}_0 is unstable, the epidemic can start and the system evolves toward another equilibrium (or more generally, an attractor) characterized by positive values of pathogen concentrations and densities of infected people. On the contrary, if \mathbf{X}_0 is stable, the pathogen cannot invade the metacommunity and disease outbreak is not possible. If the communities are isolated (no connections, either

through the hydrological network or through human mobility), successful pathogen invasion in each community is determined by whether the basic reproductive number,

$$R_{0_i} = \frac{p_i H_i \beta_i}{W_i K \mu_{B_i} (\gamma + \mu + \alpha)},$$

(Codeço 2001) is smaller or larger than unity.

Here, we analyze system (1), taking into account the network connections, and find out which parameter combinations correspond to a transition from \mathbf{X}_0 being stable to its being unstable. Technically, this is called a “bifurcation” (Kuznetsov 1995). Finding bifurcation conditions requires analysis of how the eigenvalues associated with the disease-free equilibrium vary with the system parameters. In practice, to analyze stability, we consider the linearized system, which is given by

$$\begin{aligned} \frac{d\Delta\mathbf{S}}{dt} &= -\mu\Delta\mathbf{S} - (1 - m_s)\mathbf{H}\beta\mathbf{B}^* - m_s\mathbf{H}\mathbf{Q}\beta\mathbf{B}^*, \\ \frac{d\mathbf{I}}{dt} &= (1 - m_s)\mathbf{H}\beta\mathbf{B}^* + m_s\mathbf{H}\mathbf{Q}\beta\mathbf{B}^* - \phi\mathbf{I}, \\ \frac{d\mathbf{B}^*}{dt} &= -(\mathbf{M}_B + \mathbf{L})\mathbf{B}^* + \mathbf{W}^{-1}\mathbf{P}^T\mathbf{W}\mathbf{L}\mathbf{B}^* \\ &\quad + \frac{1 - m_1}{K}\mathbf{p}\mathbf{W}^{-1}\mathbf{I} + \frac{m_1}{K}\mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T\mathbf{I}, \end{aligned}$$

where the superscript T indicates matrix transposition; $\Delta\mathbf{S} = [S_1 - H_1, \dots, S_n - H_n]^T$; $\mathbf{I} = [I_1, \dots, I_n]^T$; $\mathbf{B}^* = [B_1/K, \dots, B_n/K]^T$; $\phi = \gamma + \mu + \alpha$; β , \mathbf{p} , \mathbf{M}_B , \mathbf{L} , \mathbf{H} , and \mathbf{W} are diagonal matrices whose nonzero elements are made up by the parameters β_p , p_p , μ_{B_p} , l_p , H_p , and W_p with $i = 1, 2, \dots, n$, respectively (\mathbf{W}^{-1} is also obviously diagonal, with elements equal to $1/W_i$); and $\mathbf{P} = [P_{ij}]$ and $\mathbf{Q} = [Q_{ij}]$ are, respectively, the hydrological and human-mobility connection matrices, assumed to be Markov matrices, that is, $\sum_{j=1}^n P_{ij} = 1$ and $\sum_{j=1}^n Q_{ij} = 1$ for any i . As for \mathbf{P} and \mathbf{Q} , we further assume that the union of the graphs associated with them is strongly connected, that is, that there always exists a path between any two nodes along either the hydrological or the human-mobility network. In some cases, we use the stronger assumption that the graphs associated with the two networks are strongly connected (so that matrices \mathbf{P} and \mathbf{Q} are irreducible). In this case, according to the Perron-Frobenius theorem for nonnegative Markov matrices (Gantmacher 1959), we can state that both \mathbf{P} and \mathbf{Q} have unique dominant eigenvalues equal to 1. As the eigenvalues of a matrix and of its transpose coincide, the dominant eigenvalues of \mathbf{P}^T and \mathbf{Q}^T are also 1.

Because the model is a positive system (namely, the variables S_p , I_p , and B_i can never become negative if the

system is initialized at nonnegative conditions) and the disease-free equilibrium is characterized by null values of infected abundances and pathogen concentrations, the bifurcation from stable to unstable can occur only via an exchange of stability. This implies that the disease-free equilibrium switches from being a stable node to being a saddle through a transcritical bifurcation. Thus, the condition for the bifurcation to occur (epidemiologically equivalent to the condition for successful pathogen invasion) is that the Jacobian evaluated at the disease-free equilibrium has one zero eigenvalue (Kuznetsov 1995). Appendix A reports the detailed calculations necessary to derive the pathogen invasion condition, which we state below. An alternative derivation of the onset condition, based on a next-generation matrix approach (Diekmann and Heesterbeek 2000; Diekmann et al. 2010), is described in appendix B; appendixes B–D available online.

Define the matrix

$$\mathbf{R}_0 = \begin{bmatrix} R_{0_1} & 0 & \dots & 0 \\ 0 & R_{0_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{0_n} \end{bmatrix} = \frac{\mathbf{p}}{K\phi} \mathbf{M}_B^{-1} \mathbf{H} \beta \mathbf{W}^{-1},$$

which is made up of the basic reproductive numbers of each community when isolated from the others (R_{0_i}). In addition to matrix \mathbf{R}_0 , three other matrices of reproductive numbers can be introduced, namely,

$$\begin{aligned} \mathbf{R}_0^I &= \frac{\mathbf{p} \mathbf{M}_B^{-1} \mathbf{Q}^T \mathbf{H} \beta \mathbf{W}^{-1}}{K\phi}, \\ \mathbf{R}_0^S &= \frac{\mathbf{p} \mathbf{M}_B^{-1} \mathbf{H} \mathbf{Q} \beta \mathbf{W}^{-1}}{K\phi}, \\ \mathbf{R}_0^{IS} &= \frac{\mathbf{p} \mathbf{M}_B^{-1} \mathbf{Q}^T \mathbf{H} \mathbf{Q} \beta \mathbf{W}^{-1}}{K\phi}, \end{aligned}$$

corresponding to metacommunities where only infecteds are mobile, only susceptibles are mobile, or both are mobile, respectively. If we account for the different probabilities of movement in the metacommunity, we can compute the average reproductive matrix as

$$\begin{aligned} \mathbf{R}_0^{\text{ave}} &= (1 - m_1)(1 - m_s)\mathbf{R}_0 + m_1(1 - m_s)\mathbf{R}_0^I \\ &\quad + (1 - m_1)m_s\mathbf{R}_0^S + m_s m_1 \mathbf{R}_0^{IS}. \end{aligned}$$

In appendix A, we show that the bifurcation of the disease-free equilibrium corresponds to the condition

$$\det(\mathbf{U}_n - \mathbf{R}_0^{\text{ave}} - \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L}) = 0, \quad (2)$$

where \mathbf{U}_n is the identity matrix of dimension n . Equivalently, if we define a generalized reproductive matrix

$$\mathbf{J}_0 = \mathbf{R}_0^{\text{ave}} + \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L}, \quad (3)$$

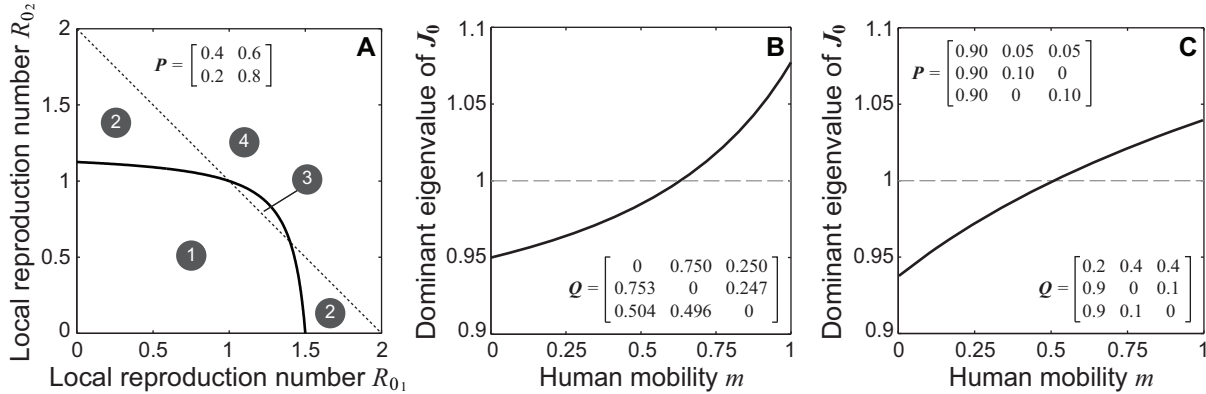


Figure 1: Special and prototypical cases. A, The case $n = 2$ for model (1) without human mobility. The solid line shows the bifurcation curve given by equation (2), while the dotted line shows where the average of the two basic reproductive numbers is equal to 1. Below the solid curve, the pathogen cannot invade the metacommunity. In region 1, the average reproductive number is smaller than 1 and the pathogen cannot invade; in region 2, the average reproductive number is smaller than 1, but the pathogen can invade; in region 3, the average reproductive number is larger than 1, yet the pathogen cannot invade; in region 4, the average reproductive number is larger than 1 and the pathogen can invade. Parameter values: $l_1 = l_2 = \mu_{B_1} = \mu_{B_2} = 0.23$. B, Dominant eigenvalue of $J_0 = \mathbf{R}_0^{\text{ave}}$ (solid line) as a function of human mobility: $m_s = m_i = m$ for a simple topology with three nodes having the same community size H and connected via a gravitational model (no hydrological transport). Pairwise distances between nodes are $d_{12} = 1$, $d_{13} = 5$, $d_{23} = (d_{12}^2 + d_{13}^2)^{1/2}$. The connection matrix \mathbf{Q} is computed under the assumption that $g(d) = 1/(1 + d)$. The values of the local reproductive numbers are $R_{0i} = 0.95$ ($i = 1, 2, 3$). C, Dominant eigenvalue of J_0 as a function of mobility: $m_s = m_i = m$ for a simple topology with three hydrologically and ecologically networked nodes, connected by both hydrology and human mobility. Parameter values: $l_1 = l_2 = l_3 = 0.7$, $\mu_{B_1} = \mu_{B_2} = \mu_{B_3} = 0.23$, $R_{0i} = 0.75$ ($i = 1, 2, 3$). Connection matrices \mathbf{P} and \mathbf{Q} are reported in the figure. See text and appendix D, available online, for details.

and indicate by $\lambda_{\max}(\mathbf{A})$ the dominant eigenvalue of a matrix \mathbf{A} , the condition for waterborne pathogen invasion to be successful is

$$\xi_0 = \lambda_{\max}(\mathbf{J}_0) > 1. \quad (4)$$

Epidemics can be triggered by positive perturbations of pathogen concentrations or densities of infected individuals if and only if ξ_0 , the dominant eigenvalue of \mathbf{J}_0 , is larger than 1. Condition (4), based on matrix \mathbf{J}_0 , constitutes a nontrivial advancement over invasion criteria proposed elsewhere (Gatto et al. 2012), specifically because it allows for the study of systems where the parameters relevant to pathogen ecology (μ_B) and transport (l) vary in space, which is most likely the case in large-scale epidemics.

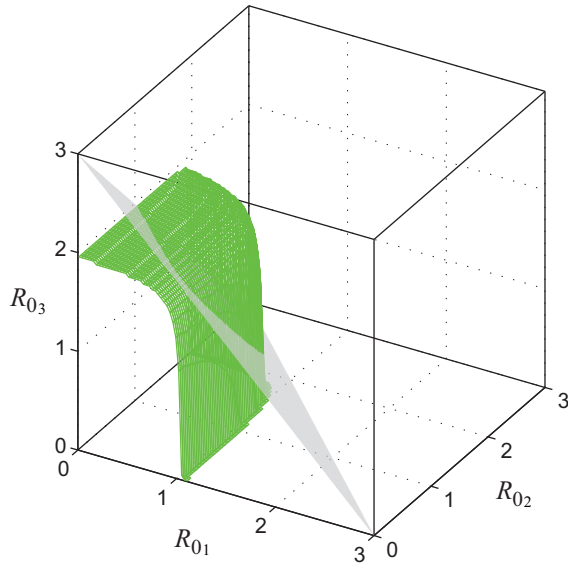
Criterion (4), coupled with equation (3), synthesizes the intertwined effects of epidemiological, hydrological, and human-mobility dynamics on the outbreak of waterborne-disease epidemics. As a matter of fact, \mathbf{J}_0 results from the sum of two matrices, one depending linearly on the hydrological network \mathbf{P} and the other nonlinearly on the human-mobility network \mathbf{Q} . Pathogen invasion condition (4) is based on the dominant eigenvalue of \mathbf{J}_0 , which is not simply deducible from the eigenvalues of the two addenda. Therefore, the hydrological and human-mobility networks interplay in a complex manner to determine disease outbreak and spread. If the pathogen transport rates l_i and the human mobilities m_s and m_i are equal to

0 (completely isolated communities), then equation (2) becomes

$$\det(\mathbf{U}_n - \mathbf{R}_0) = 0.$$

This condition is obviously satisfied when the largest R_{0i} is equal to 1—which corresponds to the classical criterion obtained in a spatially implicit context. However, the introduction of hydrological connections and human mobility changes the picture completely and makes condition (4) absolutely nontrivial, as demonstrated in the next sections.

The above criterion for successful pathogen invasion is also a key to understanding the geography of pathogen persistence and ensuing epidemic outbreaks (see app. C). In fact, when condition (4) is verified, the disease-free equilibrium becomes a saddle, and the dominant eigenvector (corresponding to the dominant eigenvalue ξ_0) pinpoints the direction in the state space along which the system orbit, after a transient period due to initial conditions, will diverge from the equilibrium. The eigenvector lies in the subspace $\Delta S_i = 0$, $i = 1, 2, \dots, n$ and is characterized by strictly positive components (according to Perron-Frobenius theorem for nonnegative matrices; see Gantmacher 1959). The components of the eigenvector thus correspond to the values of the infected individuals' abundances and the pathogen concentrations in the different human communities. Whenever the dominant ei-



Video 1: Still image from video 1 (available online). The case $n = 3$ for model (1) without human mobility, with $l_i = 0.23$ for all i , $\mu_{B_i} = 0.23$ for all i , $P_{11} = 0.9$, $P_{12} = P_{13} = 0.05$, $P_{21} = P_{31} = 0.99$, $P_{22} = P_{33} = 0.01$, $P_{32} = P_{23} = 0$. The green surface shows the bifurcation in the parameter space (R_{01}, R_{02}, R_{03}) , while the gray plane shows where the average of the three local basic reproductive numbers is equal to 1; below the green surface, the pathogen cannot invade the metacommunity and epidemics cannot be triggered.

genvalue is sufficiently larger than 1, there may be other eigenvalues larger than unity. Thus, in the very short term the system orbit might point in the direction of the subdominant eigenvector before converging to the dominant. Therefore, we can outline the geography of the epidemic as follows. At the very beginning (transient period related to initial conditions), the spread will obviously occur in the locations most connected to those initially infected, but then the epidemic will mainly propagate along the locations that correspond to the largest components of the dominant eigenvector (or, possibly, first along the components of the subdominant and then along those of the dominant eigenvector). These communities are where the number of infected individuals and the pathogen concentrations will be highest during disease outbreak, thus acting as the main foci of the epidemic.

Special and Prototypical Cases

In this section, we examine three special cases that, despite their simplicity, demonstrate how invasion criterion (4) can be applied and lead to interesting results. The situations we consider are those of no human mobility (hydrological network only), no connection among water reservoirs (human-mobility network only), and coupling of

the two dissemination pathways in a spatially homogeneous setting. The numerical examples (fig. 1) refer to prototypical networks consisting of only a few nodes, but the results (technical details in app. D) can be extended to more complex and realistic networks.

Hydrological Transport Only

If the human mobilities are equal to 0 ($m_s = m_i = 0$), then equation (2) becomes

$$\det(\mathbf{U}_n - \mathbf{R}_0 - \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L}) = 0.$$

The previous equation yields a surface in the space $(R_{01}, R_{02}, \dots, R_{0n})$. The simple case with two hydrologically and ecologically homogeneous nodes ($n = 2$, $l_1 = l_2$, $\mu_{B_1} = \mu_{B_2}$) already shows some distinctive features, which are summarized in figure 1A. In fact, the average reproductive number, $(R_{01} + R_{02})/2$, is not a good predictor of disease outbreak, because there are cases in which pathogen invasion is possible even if the average reproductive number is smaller than 1 (regions labeled 2 in fig. 1A). In contrast, there are also cases in which the pathogen cannot invade even if the average reproductive number is larger than 1 (region 3 in fig. 1A). Letting P_{11} vary, we obtain figure S1 (figures S1–S6 are available online), which shows how the size of the various parametric regions changes; for instance, if there is little transport of water between node 1 and node 2 ($P_{11} = P_{22} = 0.8$), a large size is obtained for region 2, where the average reproductive number is smaller than 1 but the pathogen can still invade. Similar results extend to the case $n = 3$. When node 1 is quite close to being absorbing—that is, when P_{11} is close to 1—the shape of the bifurcation surface is such that successful pathogen invasion is possible even if R_{02} and R_{03} are small and R_{01} is slightly larger than 1 (see video 1, available online).

Human Mobility Only

If $\mathbf{L} = 0$, then equation (2) becomes

$$\det(\mathbf{U}_n - \mathbf{R}_0^{\text{ave}}) = 0,$$

or, otherwise stated, the dominant eigenvalue of $\mathbf{R}_0^{\text{ave}}$ must be 1. As stated above, there are several human-mobility models that can be used in the epidemiological context. The class of gravitational models (Erlander and Stewart 1990) is sufficiently simple, yet flexible and amenable to theoretical analysis. Here, we assume that the connection matrix \mathbf{Q} is determined by a gravity model in which

$$Q_{ij} = \frac{H_j g(d_{ij})}{\sum_{k=i}^n H_k g(d_{ik})} \quad (5)$$

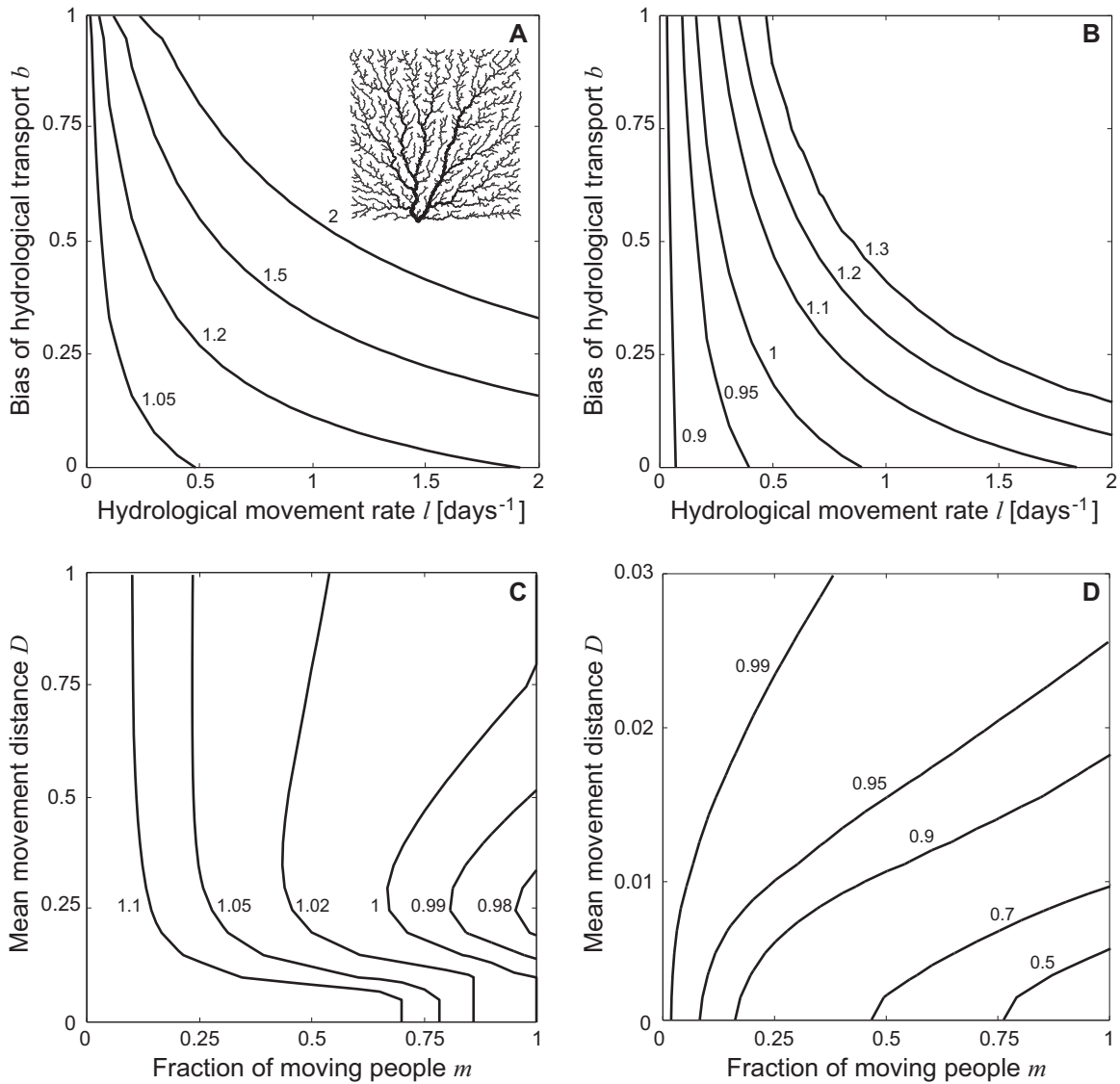


Figure 2: Effects of hydrological transport parameters, human mobility, and local reproductive number on waterborne pathogen invasion in an optimal channel network (OCN; inset of *A*) with a spatially homogeneous population distribution and a water reservoir system. *A* is obtained in the absence of human mobility, *B* and *C* account for both hydrological transport and gravity-like human mobility, and *D* is obtained in the absence of hydrological transport. The solid lines show the bifurcation curves given by equation (2) for different values of ρ_0 ($R_{0i} = \rho_0$ for all i , as labeled on curves). The pathogen can invade for the parameter combinations lying to the left (right) of the curves in *A* and *B* (*C* and *D*). The OCN is embedded in a square of side 1. Parameter values: *A*, $m_s = m_1 = 0$; *B*, $m_s = m_1 = 0.5$, $D = 0.01$; *C*, $l_i = 0.7$ for all i , $b = 0.7$, $m_s = m_1 = m$; *D*, $l = 0$, $m_s = m_1 = m$. Other parameters: $\beta = 1$, $\phi = 0.2$, $\mu_{B_i} = 0.23$ for all i . All rates are per day.

(where $i \neq j$), $Q_{ii} = 0$, and $g(d)$ is a suitable decreasing function of distance d . If we assume that water availability is proportional to population size (Bertuzzo et al. 2008; Mari et al. 2012a), then it is quite easy to prove (see app. D for technical details) that pathogen invasion conditions strongly depend on the topology of the human-mobility

network and that there might exist network topologies for which an epidemic can be triggered even if all the R_{0i} are less than 1 (subthreshold epidemic; see van den Driessche and Watmough 2002). Figure 1*B* shows a simple prototypical case that has the required property. Even if the network consists of three nodes characterized by local re-

productive numbers smaller than 1, for increasing values of human mobility, the dominant eigenvalue of \mathbf{J}_0 increases and finally exceeds unity.

Interplay between Hydrological Connections and Human Mobility in Spatially Homogeneous Networks

Let us now assume that both hydrological transport and human mobility are fully operating but that all p_i , β_i , $\mu_{B,i}$, l_i , H_i , and W_i are spatially homogeneous, that is, independent of i . Then all $R_{0,i}$ are equal to a constant ρ_0 . It can be shown (see app. D for details) that if \mathbf{Q} is nonsymmetric and $m_s > 0$, a pathogen can invade even if $\rho_0 < 1$. An example is shown in figure 1C. For low mobility, the dominant eigenvalue is smaller than 1; this is so because for $m_s = 0$, the dominant eigenvalue is controlled by ρ_0 even if \mathbf{Q} is nonsymmetric. However, for high mobility, the dominant eigenvalue becomes larger than 1. This demonstrates that even if the epidemic could never start in isolated, homogeneous communities (because $\rho_0 < 1$), nonsymmetric mobility can give rise to pathogen invasion (another case of subthreshold epidemic; see van den Driessche and Watmough 2002). The condition that \mathbf{Q} is nonsymmetric is quite common, because even symmetric gravity models (where all H_i are equal) can yield nonsymmetric \mathbf{Q} matrices, as shown, for example, in figure 1B.

Epidemic Spread in Complex Landscapes

The analytical framework proposed above can be used to derive pathogen invasion conditions for waterborne disease epidemics spreading on more complex, realistic landscapes. Gatto et al. (2012) analyzed epidemic outbreaks in Peano basins with spatially homogeneous rates of pathogen growth and hydrological transport. Peano basins are deterministic, space-filling fractals with a tree-like structure whose topological measures are similar to those of real rivers but fail to satisfy the statistics of aggregation and upstream/downstream distances (Rodriguez-Iturbe and Rinaldo 1997; Rinaldo et al. 1999). Furthermore, the degree of the inner nodes of Peano networks is four, while the inner nodes of real river networks have no more than three connected neighbors (Rodriguez-Iturbe and Rinaldo 1997). Here, we want to introduce a more realistic description of hydrological connectivity. To this end, we apply model (1) to an optimal channel network (OCN; inset of fig. 2A), that is, a mathematical structure characterized by optimal scaling forms that closely conform to the observed geomorphological features of real river networks. OCNs hold fractal characteristics that are obtained through a specific selection process based on the principle of minimum energy expenditure in the network as a whole (Rinaldo et al. 1992; Rodriguez-Iturbe et al. 1992). The scaling

behavior of the probability distributions of contributing area, channeled lengths, and elongation in OCNs matches that observed in the field (Rodriguez-Iturbe and Rinaldo 1997), to the extent that OCNs pass all tests—however stringent—in reproducing distinctive river network statistics, both topological and metric (Rinaldo et al. 1999).

The OCN is described as an oriented graph whose nodes and edges represent, respectively, human communities and the hydrological interconnections among them (Bertuzzo et al. 2007). The inner nodes of the network have two inward edges and one outward edge, while the leaves have just one outward link. These interconnections are subsumed into matrix \mathbf{P} . Specifically, the fraction P_{ij} of pathogens that move between two nodes of the hydrological network (say from i to j) is given by

$$P_{ij} = \begin{cases} \frac{P_{\text{out}}}{d_{\text{out}}(i)P_{\text{out}} + d_{\text{in}}(i)P_{\text{in}}} & \text{if } i \rightarrow j, \\ \frac{P_{\text{in}}}{d_{\text{out}}(i)P_{\text{out}} + d_{\text{in}}(i)P_{\text{in}}} & \text{if } i \leftarrow j, \\ 0 & \text{otherwise,} \end{cases}$$

where P_{out} (P_{in}) is the fraction of pathogens moving along an outward (inward) edge and d_{out} (d_{in}) is the outdegree (indegree) of node i , that is, the number of outward (inward) edges. Note that the quantity P_{ij} can be derived from the discretization of the standard advection-dispersion equation for water flow or, equivalently, from a biased random-walk process on an oriented graph (Bertuzzo et al. 2010). The transport process is assumed to be conservative (see above); that is, $\sum_{j \in N_i} P_{ij} = 1$, where N_i is the set of neighbors connected to node i (of cardinality $d(i) = d_{\text{out}}(i) + d_{\text{in}}(i)$). The bias of hydrological transport along the river (which is related to downstream velocity) can thus be defined as $b = P_{\text{out}} - P_{\text{in}} = 2P_{\text{out}} - 1$. Note also that matrix \mathbf{P} must account for proper boundary conditions (BCs) for the upstream nodes of the network (leaves of the graph) and the outlet of the OCN. “Absorbing” BCs at a certain node describe the situation when the pathogens reaching that node will never leave it. “Reflecting” BCs, on the contrary, imply that at least some of the pathogens reaching that node can be transported to nearby nodes. At the outlet (labeled “node 1”), in particular, absorbing BCs can be used to characterize river basins with coastal regions where pathogens are not found in interepidemic periods, while reflecting BCs are better suited for, for example, areas where cholera is endemic, brackish and estuarine water represents a reservoir for pathogens, and seawater can move landward. To properly define BCs, we introduce a fictitious node 0 downstream of the network outlet (i.e., node 0 is connected with node 1 only), so that $\sum_{j \in N_1} P_{1j} = 1$, with N_1 including node 0. Absorbing BCs thus correspond to setting $P_{01} = 0$ and $P_{00} = 1$, while

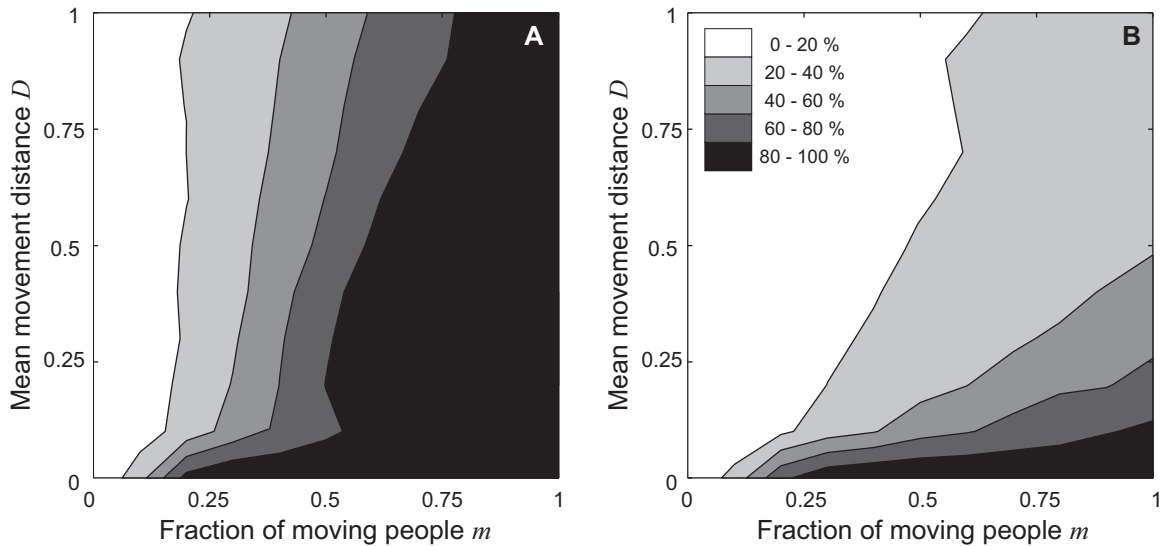


Figure 3: Effects of the human-mobility parameters $m_s = m_i = m$ and D on pathogen invasion in an optimal channel network with a Zipf-like population distribution. Results are shown for $\rho_0 = 1$ (A) and $\rho_0 = 0.9$ (B), with $R_{0_i} = \rho_0$ for all i . For each parameter combination, 1,000 simulations with different stochastic realizations of population distribution were performed. The grayscale indicates the percentage of realizations in which invasion condition (4) is met. Parameter values: $l_i = 1$ for all i , $b = 0.5$, $\beta = 1$, $\phi = 0.2$, and $\mu_{B_i} = 0.23$ for all i . All rates are per day.

(purely) reflecting BCs can be obtained by imposing $P_{01} = P_{10}$ and $P_{00} = 1 - P_{10}$. In the same way, we can define fictitious nodes upstream of the network leaves. All the numerical examples described in the remainder of the article are obtained with reflecting BCs for the leaves and absorbing BCs for the outlet of the river network, so as to mimic nonendemic coastal settings.

Analysis of the Invasion Conditions

Figure 2A shows how hydrological transport alone influences pathogen invasion/epidemic outbreak conditions in an OCN with spatially homogeneous population ($H_i = h$), water reservoir system ($l_i = l$, $W_i = cH_i$) and pathogen mortality rate ($\mu_{B_i} = \mu_B$ for all i). Specifically, the higher the values of either the hydrological transport rate l ($1/l$ is the average residence time of pathogens in each node) or the hydrological downstream bias b , the higher the local reproductive number must be for an epidemic to be triggered. As a matter of fact, in the presence of absorbing BCs at the outlet of the river system, high values of the product lb imply that a large number of pathogens are washed out through the terminal node of the network (Mari et al. 2012b), de facto increasing overall pathogen mortality—which could be compensated for by higher contamination and/or exposure rates at the local scale (i.e., higher R_{0_i}).

As in the prototypical examples proposed in the pre-

vious section, local communities can also be interlinked through human mobility, which is described by matrix \mathbf{Q} . Here we assume that the entries of \mathbf{Q} are given by gravity model (5), in which $g(d) = \exp(-d/D)$, where D is the distance parameter of the exponential kernel (Bertuzzo et al. 2011; Mari et al. 2012b). Although this mobility model is not expected to fully capture the complexity of real human movement patterns, gravity-like models have been widely applied in the epidemiological literature to describe the impact of human mobility on the course of a suite of human diseases, including influenza (Viboud et al. 2006; Eggo et al. 2010), HIV infection (Thomas 1999), measles (Xia et al. 2004; Bharti et al. 2008), and, very recently, cholera (Bertuzzo et al. 2011; Chao et al. 2011; Tuite et al. 2011; Mari et al. 2012a, 2012b; Rinaldo et al. 2012). Figure 2B shows that human mobility can significantly favor waterborne disease epidemics. In fact, it is apparent from comparing figures 2A and 2B that epidemic outbreaks related to human movement may occur for lower values of the local basic reproductive number ρ_0 ($R_{0_i} = \rho_0$ for all i) than in the corresponding case in which human mobility is negligible ($m_s = m_i = 0$). Properly taking into account host movement is thus essential to correctly estimate the likelihood of waterborne disease epidemics. Even in this realistic spatial setting, human mobility can allow for the occurrence of subthreshold epidemics (van den Driessche and Watmough 2002), that is, for the spread of epidemics characterized by $\rho_0 < 1$. This requires that a

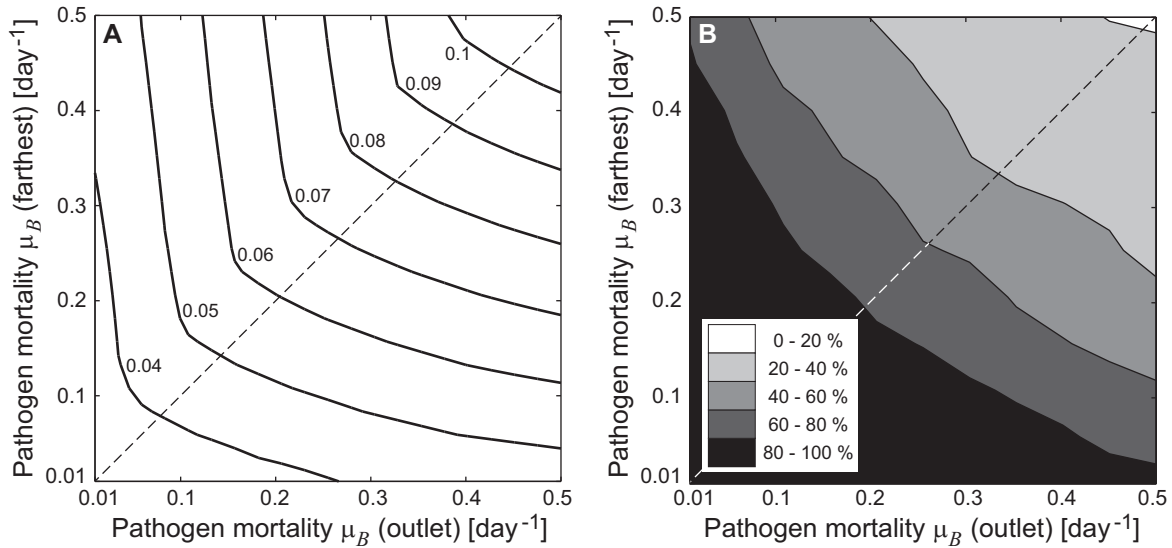


Figure 4: Effects of the spatial distribution of pathogen mortality rates and of the local contamination rate on waterborne pathogen invasion in an optimal channel network with spatially homogeneous (A) or Zipf-like (B) population distribution. The parameters $\mu_B(\text{outlet})$ and $\mu_B(\text{farthest})$ indicate the values of the pathogen mortality rate at the network outlet and in the farthest node from it, respectively. Mortality rates in the remaining nodes have been computed by linear interpolation according to distance from the outlet. The solid lines in A show the bifurcation curves given by equation (2) for different values of the local contamination rate ($p_i = p$ for all i , as labeled on curves). The pathogen can invade for mortality combinations lying below the curves. In B ($p = 0.03$), for each parameter combination, 1,000 simulations with different stochastic realizations of population distribution were performed; the grayscale indicates the percentage of realizations in which invasion condition (4) is met. The dashed lines indicate spatially homogeneous distributions of pathogen mortality rates. Other parameter values are as in figure 3.

large fraction of individuals move over intermediate distances (fig. 2C). In the absence of hydrological transport (or for low values of the product lb), epidemics can be triggered even for $\rho_0 \ll 1$ (fig. 2D).

All the results shown in figure 2 refer to the idealized case of spatially homogeneous population distribution ($H_i = h$ for all i), which obviously represents a somewhat crude simplification of the observed spatial arrangement of human communities. Empirical observations show, in fact, that a much more realistic model for the size of human settlements is given by the so-called Zipf's law (Zipf 1949; Newman 2005), according to which the size distribution of human communities can be well represented by a power-law distribution, namely, $\Pr(H_i) \propto H_i^{-2}$. We have thus repeated the analysis of model (1), sampling the size of each local community from a power-law distribution. To allow comparison with the results of the homogeneous case, we imposed the normalizing constraint $\sum_{i=1}^n H_i = nh$. For each parameter setting analyzed, we extracted 1,000 independent realizations of sample size n from the population distribution and distributed the n population sizes randomly in the landscape. For each replicate, human-mobility fluxes were computed with the exponential form of gravity model (5). Again we find that increased movement distance favors epidemic outbreaks, especially for high values of the

fraction of moving individuals (fig. 3). Quite remarkably, coupling a spatially heterogeneous population distribution with a gravity-like mobility pattern allows the conditions for successful pathogen invasion to be met with a high probability, even for low values of the local reproductive number (e.g., $R_{0i} = \rho_0 = 0.9$ for all i in fig. 3B)—which thus turns out to be a rather poor indicator of the likelihood of pathogen invasion in realistic networked landscapes.

Pathogen invasion condition (4) displays all its power when the further effects of heterogeneities in the spatial distributions of pathogen mortality and/or hydrological transport rates are studied. Figure 4 shows the results for a case in which pathogen mortality rates follow a geographical pattern (for instance, because of heterogeneous water temperature, pH values, or salinity) in which the network outlet and the node farthest from it are characterized by the extreme values of the mortality range (either minimum or maximum), while mortality rates in the remaining nodes assume intermediate values. In particular, the pathogen mortality rate can often be assumed to be minimal at the network outlet and maximal at the upstream headwaters. This condition is representative, for instance, of the ecology of *Vibrio cholerae*. The causative agent of cholera is known to be a regular guest of coastal

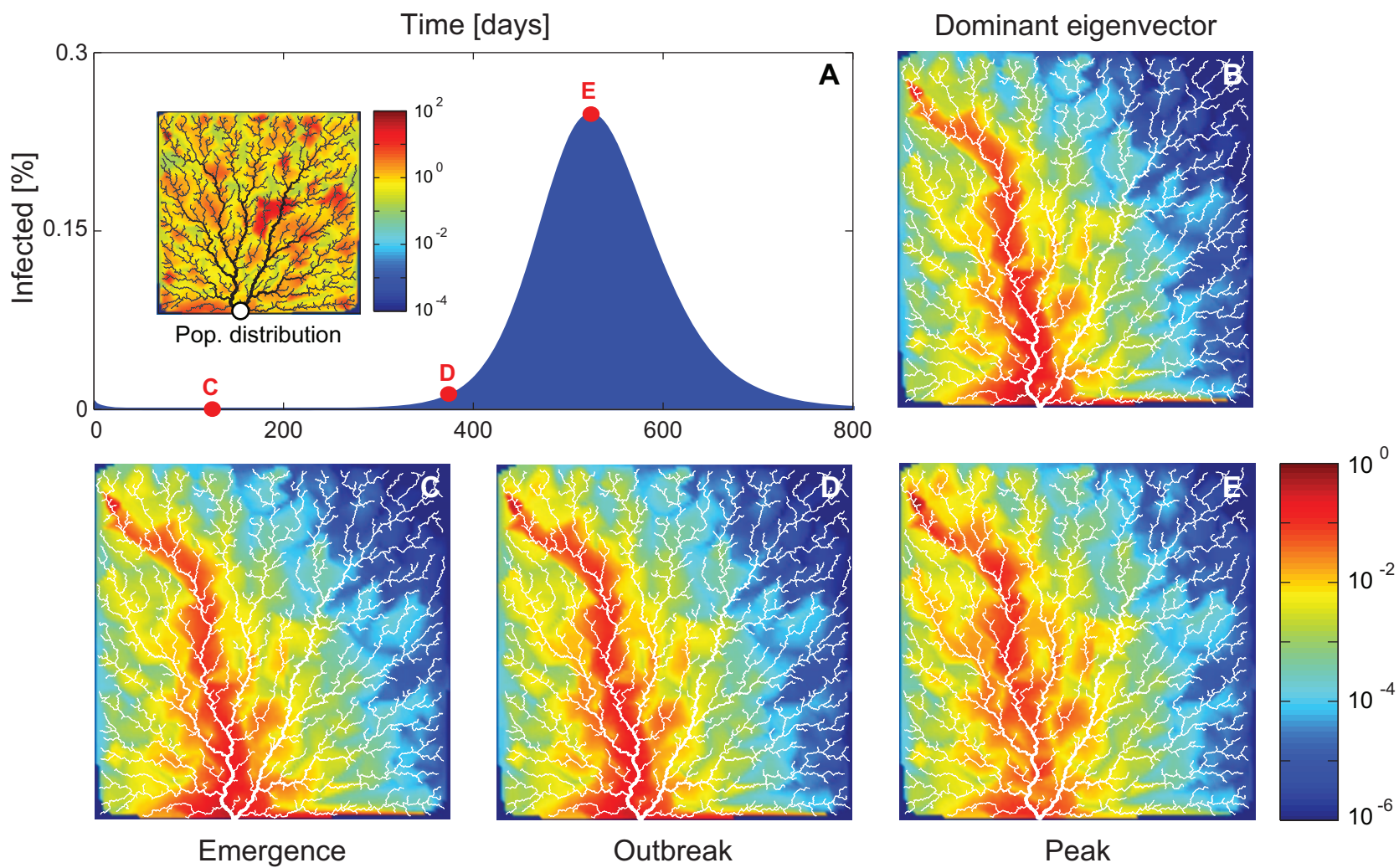


Figure 5: Simulation of a waterborne disease epidemic close to the bifurcation in an optimal channel network with hydrological transport and human mobility. Populations of different human communities are distributed according to Zipf's law (inset of A; see text). A, Time series of total infected individuals (normalized by total population); the red dots indicate the emergence of the epidemic (C), the full outbreak (D, identified numerically as the week after which daily incidence and its first and second time derivatives exceed 5% of the respective maximum values recorded in the simulation), and the peak (E). B, Dominant eigenvector of matrix J_0 , rescaled so as to portray the infected persons' components (see app. C, available online). C, Geography of disease emergence, evaluated from the numerical simulation as the number of cases in the week following point C in A. D, Geography of epidemic outbreak,

brackish water ecosystems, where it finds suitable conditions for long-term persistence (Colwell 1996), while its survival in inland freshwaters is quite low. Figure 4 shows that as the pathogen mortality rate increases, so does the local contamination rate necessary for a pathogen outbreak to occur. However, as the bifurcation curves are not symmetric with respect to the bisector of the parameter plane, different geographical patterns of pathogen mortality can determine different pathogen invasion conditions. Spatial heterogeneities in pathogen mortality (and transport) may thus play an important role in determining epidemic outbreaks, especially when regional spatial scales are considered.

The Geography of Epidemic Outbreaks

The analytical framework outlined above can be used to characterize the geography of epidemic outbreaks in complex landscapes. In fact, close to the transcritical bifurcation through which the disease-free equilibrium loses stability, the components of the dominant eigenvector of matrix \mathbf{J}_0 (corresponding to ξ_0) indicate where the epidemic is expected to hit hardest at the beginning of the epidemic (see app. C). Numerical simulations of model (1) can be used to assess the predictive power of matrix \mathbf{J}_0 and its eigenvalues and eigenvectors. Figures 5 and 6 show simulations of the model in the most complex setting considered so far (realistic OCN landscape, hydrological and human-mediated connections, heterogeneous population distribution). The outbreak is assumed to start close to the outlet of the river network, as is often observed in real-world epidemics (Bertuzzo et al. 2008; Mari et al. 2012a), with 0.1% of the local population representing the initially infected pool.

Figure 5 illustrates the case of spatially homogeneous mortality rates of pathogens. Close to the bifurcation, the dominant eigenvector of \mathbf{J}_0 (fig. 5B) is actually a very good indicator of incidence at disease emergence (defined as the time when the derivative of the time series of total infected people turns from negative to positive after the fading of initial conditions; see fig. 5A, 5C), at the full development of the outbreak (i.e., when a sizable number of cases begin to be reported; see fig. 5D legend for details), and even at the epidemic peak (corresponding to the global maximum

of the time series for total infected individuals; fig. 5E). The agreement between the quantity depicted in figure 5B (\log_{10} of the dominant eigenvector components) and those reported in figures 5C–5E (\log_{10} of simulated weekly incidence in the network nodes) is demonstrated by the relevant coefficients of determination: $R_E^2 = 0.99$ for disease emergence (fig. 5C), $R_O^2 = 0.99$ for epidemic outbreak (fig. 5D), and $R_P^2 = 0.98$ for epidemic peak (fig. 5E). Moreover, the dominant eigenvector of \mathbf{J}_0 is also a good indicator of the whole course of the epidemic evaluated as cumulative cases in the two years following disease emergence ($R_T^2 = 0.98$; not shown in fig. 5).

The dominant eigenvector of \mathbf{J}_0 is a satisfactory indicator not only of disease incidence but also of prevalence in each node, evaluated as the number of cases recorded in a network node (in a given time span) divided by the size of the local community. The relevant coefficients of determination range, respectively, from $R_E^2 = R_O^2 = 0.99$ for emergence and outbreak to $R_P^2 = 0.97$ and $R_T^2 = 0.98$ for epidemic peak and the whole epidemic course. The disease dynamics is also related to the location of the first epidemic hotbed (namely, being more or less close to the dominant eigenvector). Figure 5 was obtained with the first infections occurring at the outlet of the OCN. Different choices of initial conditions have been used in figures S2 and S3. These show that, after a transient period due to varied initial conditions, the dominant eigenvector of \mathbf{J}_0 still provides a good description of the geography of disease spread.

Similar results are found for spatially heterogeneous μ_B s. Figure 6 shows a simulation in which pathogen mortality is at a minimum at the outlet and at a maximum at the headwaters of the river network. It is evident from contrasting the spatial patterns reported in figures 5 and 6 that the high values of the pathogen mortality rate far from the network outlet limit the upstream propagation of the epidemic in its initial phases. As the outbreak unfolds, however, the effect of infected individuals shedding pathogens in the river system becomes more important than the geographical distribution of pathogen mortality rates. Nevertheless, the dominant eigenvector of \mathbf{J}_0 remains a satisfactory indicator of disease incidence in the long run as well ($R_E^2 = R_O^2 = 0.95$, $R_P^2 = 0.91$, $R_T^2 = 0.92$).

It is also possible to properly characterize the geography

evaluated as the number of cases in the week following point D in A. E, Geography of epidemic peak, evaluated as the number of cases in the week following point E in A. In B–E, large (small) values of the depicted quantities (\log_{10} of either eigenvector or incidence components, normalized so as to have the norm equal to 1) are colored in red (blue; see color scale in E); note also that linear interpolation in space has been performed in these four panels. All the parameters are spatially homogeneous. The outlet of the river network (white dot in the inset of A) is chosen as the initial condition for the simulation. Parameter values: $R_{0_i} = \rho_0 = 2$ for all i , $l_i = 3.1$ for all i , $b = 0.8$, $m_s = m_t = 0.2$, $D = 0.05$, $\beta = 1$, $\phi = 0.2$, $\mu_{B_i} = 0.23$ for all i . All rates are per day. For this parameter setting, the dominant eigenvalue of matrix \mathbf{J}_0 is slightly larger than 1 ($\xi_0 = 1.07$).

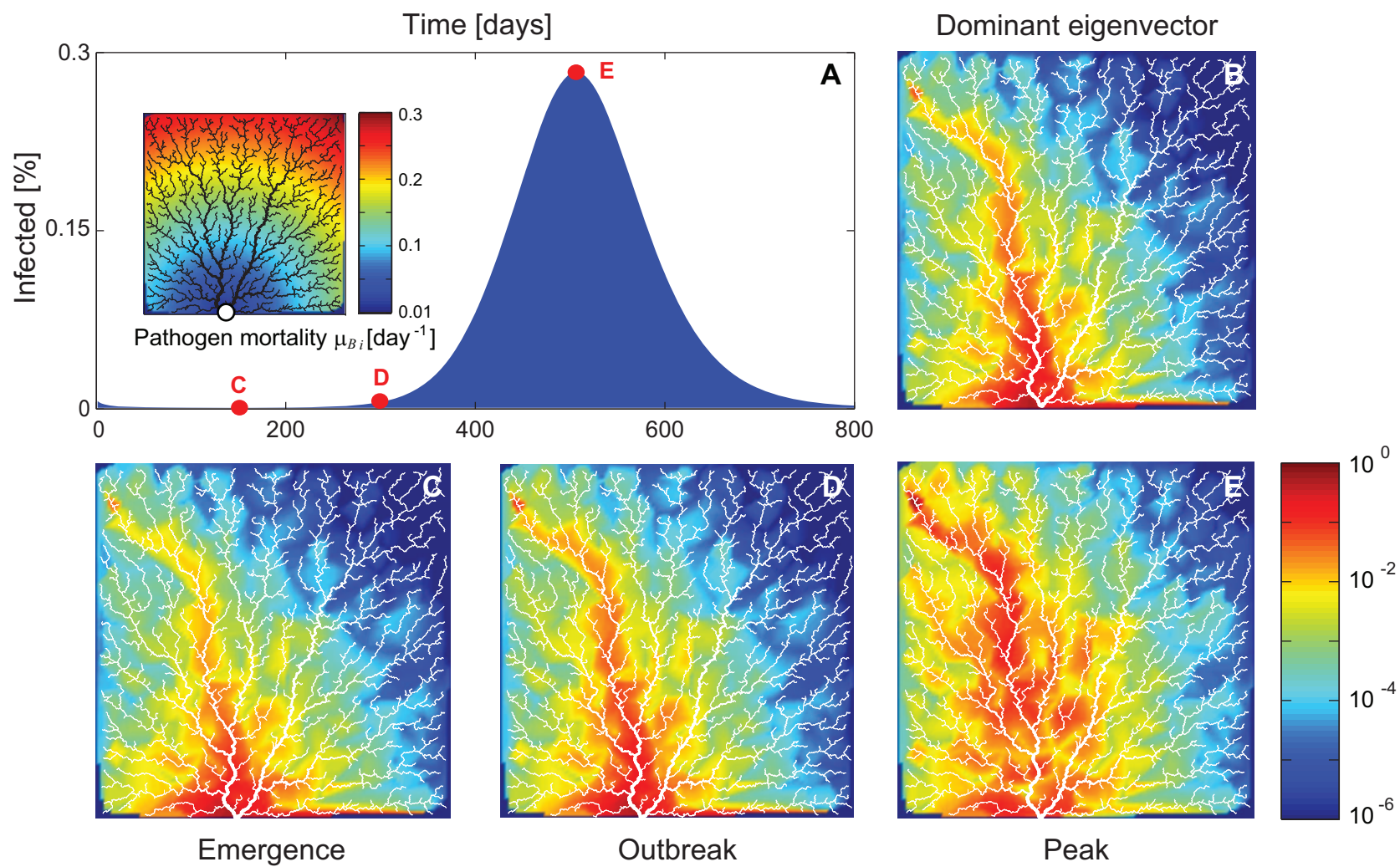


Figure 6: Simulation of a waterborne disease epidemic close to the bifurcation in an optimal channel network with spatially heterogeneous pathogen mortality rates. Parameter values: $p_i = 0.092$ for all i , μ_{B_i} as in the inset of A. Other parameters and simulation details are as in figure 5. For this parameter setting, the dominant eigenvalue of \mathbf{J}_0 is 1.06.

of pathogen invasion and disease spread far from the bifurcation, when the dominant eigenvalue of \mathbf{J}_0 is much larger than unity (and \mathbf{J}_0 may possibly have several unstable eigenvalues), by computing the eigenvalues (and the eigenvectors) of matrix \mathbf{J}^* (instead of \mathbf{J}_0 ; see app. C for the definition of \mathbf{J}^* and its link to \mathbf{J}_0). Numerical simulations show that far from the bifurcation, the localization of the first cases is well characterized by the components of the subdominant eigenvector of matrix \mathbf{J}^* , while the subsequent phases of the epidemic are better captured by the dominant eigenvector (see figs. S4 and S5). The quantitative agreement between the eigenvector components and the geography of disease incidence is very good for disease emergence and outbreak; however, at the epidemic peak the accordance is weaker (table S1, available online).

Discussion and Conclusions

Diekmann et al. (1990) and van den Driessche and Watmough (2002) showed how to properly define a generalized reproductive number for epidemiological models in which the host population can be divided into groups whose members are distinguishable by age, behavior, spatial position, and/or stage of disease. In this work, we have followed a similar approach to derive a novel expression that defines space-dependent conditions for waterborne pathogen invasion and disease outbreak. To this end, we have introduced a generalized reproductive matrix \mathbf{J}_0 that accounts not only for local epidemiological dynamics but also for different mechanisms of pathogen spread along hydrological and human-mobility networks. The dominant eigenvalue ξ_0 of \mathbf{J}_0 is the quantity controlling the epidemic. The condition for pathogen invasion is $\xi_0 > 1$ and does not require that the local R_0 be larger than 1. Subthreshold epidemics (van den Driessche and Watmough 2002) are therefore possible. This invasion condition generalizes previous results obtained under the assumption of spatially homogeneous pathogen ecology and transport (Gatto et al. 2012). Therefore, it is particularly suited to addressing large-scale epidemic outbreaks, that is, epidemics involving spatial scales at which climatic and ecological conditions cannot reasonably be considered spatially homogeneous.

With respect to traditional mean-field modeling of waterborne diseases, our approach is more realistic because there are many cases in which spatial homogeneity of the host population and of the pathogen distribution cannot be reasonably assumed. Actually, the hypothesis of homogeneity would be tenable if there were fast spatial mixing, that is, if the spatial spread of pathogens were much faster than the local increase of pathogen concentration and disease incidence. However, this is rarely the case. For instance, the timescales of spatial spread and local epi-

demiology in cholera dynamics are often comparable; therefore, realistic models for cholera epidemics must account for the spatial distribution of the host population and for the relevant pathogen relocation mechanisms (Bertuzzo et al. 2010). The importance of replacing R_0 with \mathbf{J}_0 and its dominant eigenvalue ξ_0 is also demonstrated by the ability of our approach to describe the geography of disease outbreak, which is well characterized by the dominant eigenvector of matrix \mathbf{J}_0 (or, more generally, by the dominant and subdominant eigenvectors of matrix \mathbf{J}^*). This result can be easily applied to realistic landscapes described by networks of any given complexity, ranging from a few to thousands of nodes. It is, in fact, based on simple matrix computations that are routinely performed with standard scientific software.

Network structure can also influence pathogen propagation along river systems (Bertuzzo et al. 2010) and, in turn, invasion conditions. As an example, the different topological and metric properties of Peano basins and OCNs make pathogen invasion more likely in the former than in the latter, all else being equal (fig. S6). In other terms, approximating a river network with a Peano construct (as done in Gatto et al. 2012), rather than with a more realistic OCN, can lead to an overestimation of the actual likelihood of waterborne pathogen invasion (and epidemic outbreak).

Our method can also be used to determine conditions for real-world outbreaks (as already attempted by Gatto et al. [2012]) and to predict which communities will be hit hardest during the epidemic. Such outcomes would greatly help in the planning of prognostic sanitation actions and/or the management of emergency interventions in the aftermath of disease outbreak. Both are crucial goals of public-health policies, as shown by the discussion on the recent cholera epidemic in Haiti (Cyranoski 2011; Walton and Ivers 2011). The application of the framework to real case studies would clearly require some spatially distributed knowledge of local conditions. However, georeferenced data on population distribution, hydrology, and transportation networks are becoming increasingly available worldwide, including data for developing countries. This has recently allowed spatially explicit models of epidemic outbreaks to be set up in the very course of an epidemic (e.g., Andrews and Basu 2011; Bertuzzo et al. 2011; Chao et al. 2011; Tuite et al. 2011; Rinaldo et al. 2012). Data on sanitation conditions (which influence contact and contamination rates and, in turn, determine the local value of the basic reproductive number of the disease) are usually less common. When available, these data can be readily incorporated into spatially explicit analyses (e.g., Sasaki et al. 2008; Oguntoket et al. 2009; Osei et al. 2010; Mari et al. 2012a). When sanitation data are not available, the framework outlined here can still be

useful for evaluating possible intervention scenarios in a more qualitative way.

Our approach, however, is not exempt from limitations. First of all, it is useful to analyze the space/time patterns of the disease outbreak in a broad sense at the beginning of the epidemic or during the following midterm evolution, but in general our approach cannot be reliably used to derive long-term patterns into the distant future. This task would in fact require determination of the attractor of the model and analysis of its characteristics. Also, the model should be modified to include waning immunity, because individuals recovered from waterborne diseases usually become susceptible again after a certain time. This could be done by adding a compartment of “recovered” in each node (Rinaldo et al. 2012), thus obtaining a space-explicit SIRB-like model (i.e., an SIR model with a component for bacterial dynamics). The theoretical analysis, however, would not be simple, because the attractor might not be an equilibrium. Second, our approach might be made more realistic by including seasonality in the model parameters (as in the space-implicit model of Righetto et al. [2012] or the hydroclimatological model of Bertuzzo et al. [2012]) to account for temperature and precipitation variations that are known to influence pathogen growth rate, water availability, and exposure and contamination rates. The stability analysis of the disease-free equilibrium in this space-explicit, time-varying model would be considerably more complex. It would involve determination not of the eigenvalues but of the Lyapunov exponents associated with the equilibrium (Ferrière and Gatto 1995) and is thus left to future work.

Despite these caveats, we think that the mathematical approach used here to derive waterborne pathogen invasion conditions in network models is quite general. It can be usefully applied to other infectious diseases for which an SI-like compartmental model can be devised, specifically in the presence of spatially inhomogeneous host distribution. Possible examples include directly transmitted human diseases (i.e., pathologies that are transmitted

through airborne or fecal-oral contact), in which the intertwining of mobility and social networks may play a key role in spreading the epidemic, and vector-borne/zoonotic diseases, in which spatial interactions of both host and vector/carrier populations might be relevant. In general, the mathematical framework outlined in this work can be used to approach other (apparently unrelated) topics in spatial and conservation ecology, such as the persistence of populations living in fragmented landscapes (e.g., Casagrandi and Gatto 1999, 2002, 2006), advective environments (e.g., Speirs and Gurney 2001; Pachepsky et al. 2005), dendritic networks (Campbell Grant et al. 2007, 2010), and webs of marine protected areas (e.g., White et al. 2010; Aiken and Navarrete 2011; Watson et al. 2011). In all these cases, in fact, population persistence can be established by properly accounting for the relevant spatial interactions and studying the conditions under which the extinction equilibrium becomes stable or unstable. Eigenvector analysis can assist in designing conservation efforts that should be specifically calibrated in both time and space.

Acknowledgments

We are grateful to J. L. Bronstein, G. Dwyer, and two reviewers for providing very valuable suggestions on how to improve the article, and to L. Lloyd-Jones for reading and commenting on the manuscript. L.M., E.B., L.R., and A.R. acknowledge the support provided by the European Research Council advanced grant program, through project RINEC-227612, and by the Swiss National Science Foundation (SNSF) projects 200021_124930/1 and CR2312_138104/1. M.G. and A.R. acknowledge the support from SNSF project IZK0Z2_139537/1 for international cooperation. I.R.-I. acknowledges the support of the James S. McDonnell Foundation through a grant for Studying Complex Systems (220020138).

APPENDIX A

Derivation of Pathogen Invasion Conditions

To analyze stability, we consider the Jacobian of the linearized system evaluated at the disease-free equilibrium \mathbf{X}_0 (see main text), which is given by

$$\mathbf{J} = \begin{bmatrix} -\mu\mathbf{U}_n & 0 & -m_s\mathbf{H}\mathbf{Q}\boldsymbol{\beta} - (1 - m_s)\mathbf{H}\boldsymbol{\beta} \\ 0 & -\phi\mathbf{U}_n & m_s\mathbf{H}\mathbf{Q}\boldsymbol{\beta} + (1 - m_s)\mathbf{H}\boldsymbol{\beta} \\ 0 & \frac{m_l}{K}\mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T + \frac{1 - m_l}{K}\mathbf{p}\mathbf{W}^{-1} & -\mathbf{M}_B - (\mathbf{U}_n - \mathbf{W}^{-1}\mathbf{P}^T\mathbf{W})\mathbf{L} \end{bmatrix}.$$

Because of its block-triangular structure, the Jacobian has obviously n eigenvalues equal to $-\mu$; therefore, instability is determined by the eigenvalues of the block matrix

$$\mathbf{J}^* = \begin{bmatrix} -\phi\mathbf{U}_n & m_s\mathbf{H}\mathbf{Q}\boldsymbol{\beta} + (1 - m_s)\mathbf{H}\boldsymbol{\beta} \\ \frac{m_1}{K}\mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T + \frac{1 - m_1}{K}\mathbf{p}\mathbf{W}^{-1} & -\mathbf{M}_B - (\mathbf{U}_n - \mathbf{W}^{-1}\mathbf{P}^T\mathbf{W})\mathbf{L} \end{bmatrix}.$$

The off-diagonal entries of \mathbf{J}^* are all nonnegative, and at least one diagonal entry is negative; thus, \mathbf{J}^* is a proper Metzler matrix (Farina and Rinaldi 2000), and its eigenvalue with the maximal real part (dominant eigenvalue) is real. If the union of graphs associated with \mathbf{P} and \mathbf{Q} is strongly connected, then the graph associated with \mathbf{J}^* is also strongly connected. Therefore, the Perron-Frobenius theorem for irreducible matrices (Gantmacher 1959) states that the dominant eigenvalue is a simple real root of the characteristic polynomial. Transcritical bifurcation of the disease-free equilibrium occurs when one eigenvalue is 0, that is, when the determinant of \mathbf{J}^* is 0 (Kuznetsov 1995). As long as the disease-free equilibrium is stable, all the eigenvalues have negative real parts and $\det(\mathbf{J}^*)$ is positive, because \mathbf{J}^* is a matrix of even order $2n$. The disease-free equilibrium becomes unstable when $\det(\mathbf{J}^*)$ switches from positive to negative or, equivalently, when the dominant eigenvalue becomes 0. For matrices of the kind

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

with all blocks being square and matrix \mathbf{A} commuting with matrix \mathbf{C} , the following equality holds:

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det(\mathbf{A}\mathbf{D} - \mathbf{C}\mathbf{B})$$

(see Silvester 2000). As \mathbf{U}_n obviously commutes with any matrix, we have

$$\begin{aligned} \det(\mathbf{J}^*) &= \det \left[\phi(\mathbf{M}_B + (\mathbf{U}_n - \mathbf{W}^{-1}\mathbf{P}^T\mathbf{W})\mathbf{L}) - \frac{m_s m_1}{K} \mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T\mathbf{H}\mathbf{Q}\boldsymbol{\beta} - \frac{m_1(1 - m_s)}{K} \mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T\mathbf{H}\boldsymbol{\beta} \right. \\ &\quad \left. - \frac{(1 - m_1)m_s}{K} \mathbf{p}\mathbf{W}^{-1}\mathbf{H}\mathbf{Q}\boldsymbol{\beta} - \frac{(1 - m_1)(1 - m_s)}{K} \mathbf{p}\mathbf{W}^{-1}\mathbf{H}\boldsymbol{\beta} \right]. \end{aligned}$$

The matrices \mathbf{L} and \mathbf{W}^{-1} are diagonal, and thus commuting, matrices; hence, we can rework the determinant of \mathbf{J}^* as

$$\begin{aligned} \det(\mathbf{J}^*) &= \phi^n \det \left[\mathbf{M}_B\mathbf{W}^{-1}\mathbf{W} - \mathbf{M}_B\mathbf{M}_B^{-1}\mathbf{W}^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L}\mathbf{W} - \frac{m_s m_1}{K\phi} \mathbf{M}_B\mathbf{M}_B^{-1}\mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T\mathbf{H}\mathbf{Q}\boldsymbol{\beta} - \frac{m_1(1 - m_s)}{K\phi} \mathbf{M}_B\mathbf{M}_B^{-1}\mathbf{p}\mathbf{W}^{-1}\mathbf{Q}^T\mathbf{H}\boldsymbol{\beta} \right. \\ &\quad \left. - \frac{(1 - m_1)m_s}{K\phi} \mathbf{M}_B\mathbf{M}_B^{-1}\mathbf{p}\mathbf{W}^{-1}\mathbf{H}\mathbf{Q}\boldsymbol{\beta} - \frac{(1 - m_1)(1 - m_s)}{K\phi} \mathbf{M}_B\mathbf{M}_B^{-1}\mathbf{p}\mathbf{W}^{-1}\mathbf{H}\boldsymbol{\beta} \right]. \end{aligned}$$

Introducing the basic reproductive numbers of each community, when isolated from the others, that is,

$$R_{0_i} = \frac{p_i H_i \beta_i}{W_i K \mu_{B_i} \phi},$$

we can define the matrix

$$\mathbf{R}_0 = \begin{bmatrix} R_{0_1} & 0 & \dots & 0 \\ 0 & R_{0_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{0_n} \end{bmatrix} = \frac{\mathbf{P}}{K\phi} \mathbf{M}_B^{-1} \mathbf{H} \boldsymbol{\beta} \mathbf{W}^{-1}.$$

Since \mathbf{p} , \mathbf{H} , $\boldsymbol{\beta}$, \mathbf{M}_B and \mathbf{W}^{-1} are diagonal, and thus commuting, matrices, we can write

$$\det(\mathbf{J}^*) = \phi^n \det(\mathbf{M}_B) \det\left(\mathbf{W}^{-1}\left\{\mathbf{U}_n - \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L} - \left[(1 - m_1)(1 - m_s)\mathbf{R}_0 + \frac{m_s m_1}{K\phi} \mathbf{pM}_B^{-1}\mathbf{Q}^T\mathbf{H}\mathbf{Q}\beta\mathbf{W}^{-1} + \frac{m_1(1 - m_s)}{K\phi} \mathbf{pM}_B^{-1}\mathbf{Q}^T\mathbf{H}\beta\mathbf{W}^{-1} + \frac{(1 - m_1)m_s}{K\phi} \mathbf{pM}_B^{-1}\mathbf{H}\mathbf{Q}\beta\mathbf{W}^{-1}\right]\right\}\mathbf{W}\right).$$

The determinant of a product of square matrices is the product of the determinants, the determinant of an inverse matrix is the inverse of the determinant, and $\det(\mathbf{M}_B)$ and $\det(\mathbf{W})$ are positive. Therefore, the condition of transcritical bifurcation (i.e., $\det(\mathbf{J}^*)$ switching from positive to negative) is finally given by

$$\det\left\{\mathbf{U}_n - \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L} - \left[(1 - m_1)(1 - m_s)\mathbf{R}_0 + \frac{m_s m_1}{K\phi} \mathbf{pM}_B^{-1}\mathbf{Q}^T\mathbf{H}\mathbf{Q}\beta\mathbf{W}^{-1} + \frac{m_1(1 - m_s)}{K\phi} \mathbf{pM}_B^{-1}\mathbf{Q}^T\mathbf{H}\beta\mathbf{W}^{-1} + \frac{(1 - m_1)m_s}{K\phi} \mathbf{pM}_B^{-1}\mathbf{H}\mathbf{Q}\beta\mathbf{W}^{-1}\right]\right\} = 0.$$

In addition to the matrix $\mathbf{R}_0 = (\mathbf{p}/K\phi)\mathbf{M}_B^{-1}\mathbf{H}\beta\mathbf{W}^{-1}$, we introduce three other matrices of reproductive numbers:

$$\mathbf{R}_0^{\text{IS}} = \frac{\mathbf{pM}_B^{-1}\mathbf{Q}^T\mathbf{H}\mathbf{Q}\beta\mathbf{W}^{-1}}{K\phi},$$

$$\mathbf{R}_0^{\text{S}} = \frac{\mathbf{pM}_B^{-1}\mathbf{H}\mathbf{Q}\beta\mathbf{W}^{-1}}{K\phi},$$

$$\mathbf{R}_0^{\text{I}} = \frac{\mathbf{pM}_B^{-1}\mathbf{Q}^T\mathbf{H}\beta\mathbf{W}^{-1}}{K\phi},$$

corresponding, respectively, to metacommunities in which both infected and susceptible individuals are mobile, only infecteds are mobile, and only susceptibles are mobile. If we account for the different probabilities of movement in the metacommunity, we can define the average reproductive matrix as

$$\mathbf{R}_0^{\text{ave}} = (1 - m_1)(1 - m_s)\mathbf{R}_0 + m_s m_1 \mathbf{R}_0^{\text{IS}} + m_1(1 - m_s)\mathbf{R}_0^{\text{I}} + (1 - m_1)m_s \mathbf{R}_0^{\text{S}}.$$

Therefore, in more compact form the condition for pathogen invasion is

$$\det(\mathbf{U}_n - \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L} - \mathbf{R}_0^{\text{ave}}) < 0.$$

Equivalently, the bifurcation condition is that the dominant eigenvalue of the matrix

$$\mathbf{J}_0 = \mathbf{R}_0^{\text{ave}} + \mathbf{M}_B^{-1}(\mathbf{P}^T - \mathbf{U}_n)\mathbf{L}$$

must equal unity. Actually, the disease-free equilibrium switches from being stable to being a saddle, thus triggering pathogen invasion, whenever the dominant eigenvalue ξ_0 of this matrix switches from being less than 1 to being larger than 1.

Literature Cited

- Aiken, C., and S. Navarrete. 2011. Environmental fluctuations and asymmetrical dispersal: generalized stability theory for studying metapopulation persistence and marine protected areas. *Marine Ecology Progress Series* 428:77–88.
- Akanda, A. S., A. S. Jutla, M. Alam, G. C. de Magny, A. K. Siddique, R. B. Sack, A. Huq, R. R. Colwell, and S. Islam. 2011. Hydroclimatic influences on seasonal and spatial cholera transmission cycles: implications for public health intervention in the Bengal Delta. *Water Resources Research* 47:W00H07.
- Akanda, A., S. Jutla, and S. Islam. 2009. Dual peak cholera transmission in Bengal Delta: a hydroclimatological explanation. *Geophysical Research Letters* 36:L19401.
- Albert, R., and A.-L. Barabási. 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics* 74:47–97.
- Andrews, J., and S. Basu. 2011. Transmission dynamics and control of cholera in Haiti: an epidemic model. *Lancet* 377:1248–1252.
- Arino, J., J. R. Davis, D. Hartley, R. Jordan, J. Miller, and P. van den Driessche. 2005. A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology* 22:129–142.
- Bengtsson, L., X. Lu, A. Torson, R. Garfield, and J. von Schreeb. 2011. Improved response to disasters and outbreaks by tracking popu-

- lation movements with mobile phone network data: a post-earthquake geospatial study in Haiti. *PLoS Medicine* 8:e1001083.
- Bertuzzo, E., R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. 2010. On spatially explicit models of cholera epidemics. *Journal of the Royal Society Interface* 7:321–333.
- Bertuzzo, E., M. Gatto, A. Maritan, S. Azaele, I. Rodriguez-Iturbe, and A. Rinaldo. 2008. On the space-time evolution of a cholera epidemic. *Water Resources Research* 44:W01424.
- Bertuzzo, E., L. Mari, L. Righetto, M. Gatto, R. Casagrandi, M. Blokesch, I. Rodriguez-Iturbe, and A. Rinaldo. 2011. Prediction of the spatial evolution and effects of control measures for the unfolding Haiti cholera outbreak. *Geophysical Research Letters* 38: L06403.
- Bertuzzo, E., L. Mari, L. Righetto, M. Gatto, R. Casagrandi, I. Rodriguez-Iturbe, and A. Rinaldo. 2012. Hydroclimatology of dual-peak cholera epidemics: inferences from a spatially explicit model. *Geophysical Research Letters* 39:L05403.
- Bertuzzo, E., A. Maritan, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. 2007. River networks and ecological corridors: reactive transport on fractals, migration fronts, hydrochory. *Water Resources Research* 43:W04419.
- Bharti, N., Y. Xia, O. N. Bjørnstad, and B. T. Grenfell. 2008. Measles on the edge: coastal heterogeneities and infection dynamics. *PLoS ONE* 3(4):e1941.
- Bolker, B., and B. Grenfell. 1995. Space, persistence and dynamics of measles epidemics. *Philosophical Transactions of the Royal Society B: Biological Sciences* 348:309–320.
- Campbell Grant, E., W. Lowe, and W. Fagan. 2007. Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters* 10:165–175.
- Campbell Grant, E., J. Nichols, W. Lowe, and W. Fagan. 2010. Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. *Proceedings of the National Academy of Sciences of the USA* 107:6936–6940.
- Capasso, V., and S. L. Paveri-Fontana. 1979. A mathematical model for the 1973 cholera epidemic in the European Mediterranean region. *Revue d'Epidemiologie et de la Santé Publique* 27:121–132.
- Casagrandi, R., and M. Gatto. 1999. A mesoscale approach to extinction risk in fragmented habitats. *Nature* 400:560–562.
- . 2002. A persistence criterion for metapopulations. *Theoretical Population Biology* 61:115–125.
- . 2006. The intermediate dispersal principle in spatially explicit metapopulations. *Journal of Theoretical Biology* 239:22–32.
- Chao, D., M. Halloran, and I. Longini. 2011. Vaccination strategies for epidemic cholera in Haiti with implications for the developing world. *Proceedings of the National Academy of Sciences of the USA* 108:7081–7085.
- Clemens, J. D., D. A. Sack, J. R. Harris, F. Van Loon, J. Chakraborty, F. Ahmed, M. R. Rao, et al. 1990. Field trial of oral cholera vaccines in Bangladesh: results from three-year follow-up. *Lancet* 335:270–273.
- Codeço, C. T. 2001. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. *BMC Infectious Diseases* 1:1.
- Colizza, V., A. Barrat, M. Barthélemy, and A. Vespignani. 2006. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the USA* 103:2015–2020.
- Colwell, R. R. 1996. Global climate and infectious disease: the cholera paradigm. *Science* 274:2025–2031.
- Cyranoski, D. 2011. Cholera vaccine plan splits experts. *Nature* 469: 273–274.
- Diekmann, O., and J. A. P. Heesterbeek. 2000. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Wiley, Chichester.
- Diekmann, O., J. A. P. Heesterbeek, and J. A. J. Metz. 1990. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* 28:365–382.
- Diekmann, O., J. A. P. Heesterbeek, and M. G. Roberts. 2010. The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface* 7:873–885.
- Eggo, R. M., S. Cauchemez, and N. M. Ferguson. 2010. Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. *Journal of the Royal Society Interface* 55:233–243.
- Erdős, P., and A. Rényi. 1959. On random graphs. *Publicationes Mathematicae* 6:290–297.
- Erlander, S., and N. F. Stewart. 1990. *The gravity model in transportation analysis: theory and extensions*. VSP, Zeist, Netherlands.
- Eubank, S., H. Guclu, V. S. Anil Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429:180–184.
- Farina, L., and S. Rinaldi. 2000. *Positive linear systems: theory and applications*. Wiley Interscience, New York.
- Ferrière, R., and M. Gatto. 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theoretical Population Biology* 48:126–171.
- Gantmacher, F. 1959. *Theory of matrices*. Chelsea, New York.
- Gatto, M., L. Mari, E. Bertuzzo, R. Casagrandi, L. Righetto, I. Rodriguez-Iturbe, and 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.
- Gog, J., R. Woodroffe, and J. Swinton. 2002. Disease in endangered metapopulations: the importance of alternative hosts. *Proceedings of the Royal Society B: Biological Sciences* 269:671–676.
- Guerrant, R. 1986. The global problem of amebiasis: current status, research needs, and opportunities for progress. *Reviews of Infectious Diseases* 8:218–227.
- Hartley, D., J. Morris, and D. Smith. 2006. Hyperinfectivity: a critical element in the ability of *V. cholerae* to cause epidemics? *PLoS Medicine* 3:63–69.
- Jensen, M., S. Faruque, J. Mekalanos, and B. Levin. 2006. Modeling the role of bacteriophage in the control of cholera outbreaks. *Proceedings of the National Academy of Sciences of the USA* 103: 4652–4657.
- Koelle, K., X. Rodó, M. Pascual, M. Yunus, and G. Mostafa. 2005. Refractory periods and climate forcing in cholera dynamics. *Nature* 436:696–700.
- Kuznetsov, Y. 1995. *Elements of applied bifurcation theory*. Springer, New York.
- Lipp, E., A. Huq, and R. R. Colwell. 2002. Effects of global climate on infectious disease: the cholera model. *Clinical Microbiology Reviews* 15:757–770.
- Mari, L., E. Bertuzzo, L. Righetto, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. 2012a. Modelling cholera epidemics: the role of waterways, human mobility and sanitation. *Journal of the Royal Society Interface* 9:376–388.
- . 2012b. 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.
- Newman, M. E. J. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46:323–351.
- Oguntoke, O., O. Aboderin, and A. Bankole. 2009. Association of water-borne diseases morbidity pattern and water quality in parts of Ibadan City, Nigeria. *Tanzania Journal of Health Research* 11: 189–195.
- Osei, F., A. Duker, E. Augustijn, and A. Stein. 2010. Spatial dependency of cholera prevalence on potential cholera reservoirs in an urban area, Kumasi, Ghana. *International Journal of Applied Earth Observation and Geoinformation* 12:331–339.
- Pachepeksy, E., F. Lutscher, R. Nisbet, and M. Lewis. 2005. Persistence, spread and the drift paradox. *Theoretical Population Biology* 67: 61–73.
- Pascual, M., M. J. Bouma, and A. P. Dobson. 2002. Cholera and climate: revisiting the quantitative evidence. *Microbes and Infection* 4:237–245.
- Pitzer, V., C. Viboud, L. Simonsen, C. Steiner, C. Panozzo, W. Alonso, M. Miller, et al. 2009. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* 325: 290–294.
- Reiner, R. C., Jr., A. A. King, M. Emch, M. Yunus, A. S. G. Faruque, and M. Pascual. 2012. Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. *Proceedings of the National Academy of Sciences of the USA* 109:2033–2036.
- Righetto, L., E. Bertuzzo, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. 2011. Modelling human movement in cholera spreading along fluvial systems. *Ecohydrology* 4:49–55.
- Righetto, L., R. Casagrandi, E. Bertuzzo, L. Mari, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. 2012. The role of aquatic reservoir fluctuations in long-term cholera patterns. *Epidemics* 4:33–42.
- Riley, S. 2007. Large-scale spatial-transmission models of infectious disease. *Science* 316:1298–1301.
- Rinaldo, A., E. Bertuzzo, L. Mari, L. Righetto, M. Blokesch, M. Gatto, R. Casagrandi, M. Murray, S. Vesenbeckh, and I. Rodriguez-Iturbe. 2012. Reassessment of the 2010–2011 Haiti cholera outbreak and rainfall-driven multiseason projections. *Proceedings of the National Academy of Sciences of the USA* 109:6602–6607.
- Rinaldo, A., R. Rigon, and I. Rodriguez-Iturbe. 1999. Channel networks. *Annual Review of Earth and Planetary Sciences* 26:289–306.
- Rinaldo, A., I. Rodriguez-Iturbe, R. Rigon, R. Bras, E. Ijjasz-Vasquez, and A. Marani. 1992. Minimum energy and fractal structures of drainage networks. *Water Resources Research* 28:2183–2195.
- Rodriguez-Iturbe, I., and A. Rinaldo. 1997. *Fractal river basins: chance and self-organization*. Cambridge University Press, New York, NY.
- Rodriguez-Iturbe, I., A. Rinaldo, R. Rigon, R. Bras, E. Ijjasz-Vasquez, and A. Marani. 1992. Fractal structures as least energy patterns: the case of river networks. *Geophysical Research Letters* 19:889–892.
- Sasaki, S., H. Suzuki, K. Igarashi, B. Tambatamba, and P. Mulenga. 2008. Spatial analysis of risk factor of cholera outbreak for 2003–2004 in a peri-urban area of Lusaka, Zambia. *American Journal of Tropical Medicine and Hygiene* 79:414–421.
- Silvester, J. 2000. Determinants of block matrices. *Mathematical Gazette* 84:460–467.
- Speirs, D., and W. Gurney. 2001. Population persistence in rivers and estuaries. *Ecology* 82:1219–1237.
- Swinton, J., J. Harwood, B. T. Grenfell, and C. A. Gilligan. 1998. Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *Journal of Animal Ecology* 67:54–68.
- Thomas, R. 1999. Reproduction rates in multiregion modeling systems for HIV/AIDS. *Journal of Regional Science* 39:359–385.
- Tuite, A., J. Tien, M. Eisenberg, D. Earns, J. Ma, and D. Fisman. 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.
- van den Driessche, P., and J. Watmough. 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* 180:29–48.
- Viboud, C., O. N. Bjørnstad, D. L. Smith, L. Simonsen, M. A. Miller, and B. T. Grenfell. 2006. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* 312:447–451.
- Walton, D., and L. Ivers. 2011. Responding to cholera in post-earthquake Haiti. *New England Journal of Medicine* 364:3–5.
- Watson, J., D. Siegel, B. Kendall, S. Mitarai, A. Rassweiler, and S. Gaines. 2011. Identifying critical regions in small-world marine metapopulations. *Proceedings of the National Academy of Sciences of the USA* 108:E907–E913.
- Watts, D., and S. Strogatz. 1998. Collective dynamics of small-world networks. *Nature* 393:440–442.
- Weil, A., A. Khan, F. Chowdhury, R. Larocque, A. Faruque, E. Ryan, S. Calderwood, F. Qadri, and J. Harris. 2009. Clinical outcomes in household contacts of patients with cholera in Bangladesh. *Clinical Infectious Diseases* 49:1473–1479.
- White, J., L. Botsford, A. Hastings, and J. Largier. 2010. Population persistence in marine reserve networks: incorporating spatial heterogeneities in larval dispersal. *Marine Ecology Progress Series* 398: 49–67.
- World Health Organization. 2008. *The global burden of disease: 2004 update*. WHO Press, Geneva.
- . 2010. *Prevention and control of cholera outbreaks: WHO policy and recommendations*. Technical report. World Health Organization, Regional Office for the Eastern Mediterranean, Cairo.
- Xia, Y., O. N. Bjørnstad, and B. T. Grenfell. 2004. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *American Naturalist* 164:267–281.
- Zipf, G. K. 1949. *Human behavior and the principle of least effort: an introduction to human ecology*. Addison-Wesley, Cambridge, MA.
- Zmirou-Navier, D., L. Gofti-Laroche, and P. Hartemann. 2006. Waterborne microbial risk assessment: a population-based dose-response function for *Giardia* spp. (E.MI.R.A study). *BMC Public Health* 6:122.