

Generalized reproduction numbers and the prediction of patterns in waterborne disease

Marino Gatto^{a,b}, Lorenzo Mari^{a,b}, Enrico Bertuzzo^b, Renato Casagrandi^a, Lorenzo Righetto^b, Ignacio Rodriguez-Iturbe^c, and Andrea Rinaldo^{b,d,1}

^aDipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milan, Italy; ^bLaboratory of Ecohydrology, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; ^cDepartment of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544; and ^dDipartimento di Ingegneria Civile, Edile ed Ambientale, Università di Padova, 35131 Padua, Italy

Contributed by Andrea Rinaldo, October 10, 2012 (sent for review July 14, 2012)

Understanding, predicting, and controlling outbreaks of waterborne diseases are crucial goals of public health policies, but pose challenging problems because infection patterns are influenced by spatial structure and temporal asynchrony. Although explicit spatial modeling is made possible by widespread data mapping of hydrology, transportation infrastructure, population distribution, and sanitation, the precise condition under which a waterborne disease epidemic can start in a spatially explicit setting is still lacking. Here we show that the requirement that all the local reproduction numbers R_0 be larger than unity is neither necessary nor sufficient for outbreaks to occur when local settlements are connected by networks of primary and secondary infection mechanisms. To determine onset conditions, we derive general analytical expressions for a reproduction matrix G_0 , explicitly accounting for spatial distributions of human settlements and pathogen transmission via hydrological and human mobility networks. At disease onset, a generalized reproduction number Λ_0 (the dominant eigenvalue of G_0) must be larger than unity. We also show that geographical outbreak patterns in complex environments are linked to the dominant eigenvector and to spectral properties of G_0 . Tests against data and computations for the 2010 Haiti and 2000 KwaZulu-Natal cholera outbreaks, as well as against computations for metapopulation networks, demonstrate that eigenvectors of G_0 provide a synthetic and effective tool for predicting the disease course in space and time. Networked connectivity models, describing the interplay between hydrology, epidemiology, and social behavior sustaining human mobility, thus prove to be key tools for emergency management of waterborne infections.

ecohydrology | aquatic ecosystems | invasion | bifurcations

Waterborne diseases, mainly due to protozoa or bacteria and often resulting in profuse diarrhea (cholera is a prominent example), are one of the leading causes of death and especially strike infants and children in low-income countries (1). Therefore, it is fundamental to develop realistic dynamical models that can provide insights into the course of past and ongoing epidemics, assist in emergency management, and allocate health-care resources via an assessment of alternative intervention strategies (2–7). These models must properly account for the relevant disease and transport timescales (8–10). Whereas some diarrheal infections, like Rotavirus, have basically a fecal–oral transmission and their spread can thus be modeled as susceptible–infected–recovered (SIR) systems (11), traditional waterborne disease models (12, 13) include, in addition to susceptibles (S) and infectives (I) (14), the population dynamics of bacteria (B) in water reservoirs. More recent modeling has considered hyperinfectivity of newly excreted vibrios (15), prey–predator interactions with phages (16), and seasonal and climate forcings (17–22). All these models, however, have not considered explicitly the spatial spread that, in waterborne diseases, occurs primarily along hydrological pathways, from coastal to inland regions or vice versa and from inland epidemic sites to neighboring areas. Integrating hydrology into epidemiological models is clearly of paramount importance and represents a recent accomplishment (5, 6, 9, 10). Moreover, infected individuals are often asymptomatic—as much as 75% for cholera (23) and 80% for amoebiasis caused by *Entamoeba*

histolytica (24)—and thus can move and spread the disease to communities other than those where they were infected. Therefore, including human mobility—via either diffusion-based (25) or gravity-like models (2–6, 26)—is mandatory. The widespread availability of georeferenced data regarding hydrological and transportation networks, human demography, water sanitation, and treatment centers distribution has facilitated the introduction of explicitly spatial models (3, 5). However, a systematic analysis of the conditions under which a waterborne disease epidemic can start within a specific territory is not yet available. Here, we determine the onset conditions and link them to explicit geographic descriptions of demographic, epidemiological, climatic, and socio-economic characteristics. From these conditions, spatiotemporal patterns of the infection are derived in a predictive manner.

Model and the Onset Conditions

The basis of our analysis is a spatially explicit nonlinear differential model (*Materials and Methods*) that accounts for both the hydrological and the human mobility network (3, 6). The hydrological networks can be those of river basins, extracted from landscape topography (27), or those of human-made water distribution and sewage systems (or both). Depending on spatial resolution, network nodes can be cities, towns, or villages. In the i th community of size H_i , with $i = 1, n$ (n is the number of nodes) the state variables at time t are the local abundances of susceptibles, $S_i(t)$, infected/infective individuals, $I_i(t)$, and the concentration of pathogens $B_i(t)$ in water reservoirs of volume W_i . Infectives release pathogens at a site-dependent rate p_i . Susceptibles are exposed to contaminated water at a rate β_i and become infected according to a saturating function of B_i ($K =$ half-saturation constant). Connections between communities are described by two matrices (3, 6), $\mathbf{P} = [P_{ij}]$ (hydrologic network) and $\mathbf{Q} = [Q_{ij}]$ (human mobility) with $i, j = 1, n$. Pathogens die at a rate μ_B and move in water from node i to node j with a probability P_{ij} at a rate l depending on downstream advection and other water-mediated transport pathways (e.g., attachment to plankton). They are also spread by human mobility: individuals leave their home node i with probability m_S for susceptibles and m_I for infectives (usually $m_S \geq m_I$), reach their target j with probability Q_{ij} , and then come back to i . Consistency requires $\sum_{j=1}^n P_{ij} = 1$ and $\sum_{j=1}^n Q_{ij} = 1$ for any i . We assume that the union of \mathbf{P} - and \mathbf{Q} -associated graphs is strongly connected; namely the infection can spread to any community along either network.

Disease onset is determined by the instability of the disease-free equilibrium \mathbf{X}_0 ($S_i = H_i$, $I_i = 0$, $B_i = 0$ for all i). If the communities were isolated (no hydrological or mobility connections),

Author contributions: M.G., L.M., E.B., R.C., I.R.-I., and A.R. designed research; M.G., L.M., E.B., R.C., L.R., and A.R. performed research; M.G., L.M., E.B., R.C., L.R., and A.R. analyzed data; and M.G., L.M., I.R.-I., and A.R. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: andrea.rinaldo@epfl.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1217567109/-DCSupplemental.

the onset condition in each community (13) would require that the local basic reproduction number

$$R_{0i} = \frac{p_i H_i \beta_i}{W_i K \mu_B (\gamma + \mu + \alpha)} > 1,$$

where γ is the recovery rate, and μ and α are, respectively, human baseline and disease-induced mortality rates. The connection with more traditional derivations of local values of R_0 for classic SIR models is reported in *SI Text*. Note that various versions of SIR-like local models would command different expressions of R_{0i} , accounting for parameters of the relevant mechanisms (e.g., ref. 28). For the spatially connected system, instead, onset in the metacommunity does not correspond to one or more local reproduction numbers > 1 .

Technically, our approach is similar to that of the next-generation matrix (NGM) (29–31), which has been used for compartmental models rather than spatially explicit ones. Both methods are based on dominant eigenvalue analysis (32). More specifically, however, we use a bifurcation analysis (33) to determine how the eigenvalues of the Jacobian at X_0 vary with the model parameters (*Materials and Methods*). As the system is positive and X_0 is characterized by null values of I_i and B_i , the bifurcation can occur only via an exchange of stability; i.e., the disease-free equilibrium switches from stable node to saddle through a transcritical bifurcation (the Jacobian has one zero eigenvalue).

Define $\phi = \gamma + \mu + \alpha$ and introduce the diagonal matrices β , H , p , W , and R_0 (whose diagonals consist of parameters β_i , H_i , p_i , W_i , and R_{0i} with $i = 1, n$, respectively). W^{-1} is also diagonal with elements equal to $1/W_i$; thus, $R_0 = p H \beta W^{-1} / (K \mu_B \phi)$. The bifurcation corresponds to

$$\det \left(U_n - \frac{l}{\mu_B + l} P^T - \frac{\mu_B}{\mu_B + l} T_0 \right) = 0, \quad [1]$$

where U_n is the (n -dimensional) identity matrix and

$$T_0 = (1 - m_I)(1 - m_S)R_0 + m_I(1 - m_S)R_0^I + (1 - m_I)m_S R_0^S + m_S m_I R_0^{IS}$$

is a transmission matrix accounting for different probabilities of movement. R_0^I , R_0^S , and R_0^{IS} are matrices that correspond, respectively, to metacommunities with (i) infectives only, (ii) susceptibles only, or (iii) both infectives and susceptibles being mobile. T_0 thus depends on mobility Q through the terms HQ (movement to a community), $Q^T H$ (movement from a community), and $Q^T H Q$ (movement to and from) (*Materials and Methods*).

Equivalent to Eq. 1 is that the dominant eigenvalue $\lambda_{\max}(G_0)$ of the generalized reproduction matrix

$$G_0 = \frac{l}{\mu_B + l} P^T + \frac{\mu_B}{\mu_B + l} T_0 \quad [2]$$

equals unity. Our main result is therefore that the onset of the disease can be triggered whenever $\lambda_{\max}(G_0)$ switches from being less than to being larger than 1; namely

$$\Lambda_0 = \lambda_{\max}(G_0) > 1. \quad [3]$$

G_0 is the sum of two matrices, one depending (linearly) on the hydrological matrix P and the other (nonlinearly) on the human mobility matrix Q . Therefore, the two networks interplay in a complex manner to determine disease insurgence and spread.

The geographical distribution of disease insurgence can be linked to the dominant eigenvector of G_0 (*SI Text*). When Λ_0 is slightly larger than unity and no other eigenvalue has modulus > 1 , X_0 is a saddle and the dominant eigenvector corresponds to the unique unstable direction in the state space along which the system orbit will diverge from the equilibrium. Once the eigenvector is suitably projected onto the subspace of infectives and normalized,

its n components are the relative proportions of the infectives in the n communities (*SI Text*). Whenever the dominant eigenvalue is sufficiently larger than one, there may be other eigenvalues of G_0 with modulus > 1 and more than one unstable direction of X_0 . However, after a short-term transient due to initial conditions, the disease will mainly propagate to the nodes corresponding to the largest components of the dominant eigenvector. These communities are those where the number of infectives will be highest during the onset and will thus act as the main foci of the disease.

Case Studies

Our approach is applied to the cholera epidemic that struck Haiti in 2010 and is still ongoing (details in *SI Text*). We use Haiti epidemiological data from November 2010 to May 2011 (Fig. 1 *A* and *B*) and a model (based on a network of 366 hydrologic/demographic entities, *SI Text*) derived from the reassessment (6) of the disease unfolding. Attractivity of a certain destination site is supposed to be given by a constrained gravity model (34),

$$Q_{ij} = \frac{H_j \exp(-d_{ij}/D)}{\sum_{k \neq i} H_k \exp(-d_{ik}/D)},$$

where d_{ij} is the distance between node i and node j and D is the average travel distance. Results of the analysis are reported in Fig. 1 *C–F*. In the Haitian case $\Lambda_0 = 1.08$ is practically insensitive to changes in pathogen movement rate l and average mobility distance D and to increases of the human mobility rate m (Fig. 1*D*), whereas it is quite sensitive to variations of the exposure and contamination rates β and p , of the pathogen mortality rate μ_B , and of the recovery rate (which is the largest component of parameter ϕ). Therefore, an effective way to prevent the onset of cholera would have been to implement sanitation measures to decrease the exposure (or contamination) rate by more than 40%. The alternative would have been to increase the rate of recovery by more than 50% via, e.g., the use of antibiotics, a measure that is not recommended, however, on a massive scale (35). The dominant eigenvector is a good indicator of the spatial distribution of recorded cases at both the coarse administrative level (10 departments, Fig. 1*C*) and the fine-grained scale (1-km² population pixels, Fig. 1*E*), as demonstrated by the corresponding coefficients of determination R^2 (Fig. 1 legend). A sensitivity analysis has been run to determine how the predictive ability of the dominant eigenvector and the value of Λ_0 change with parameter variations. The coefficients of determination (Fig. 1*F*) exceed 75% even for variations as high as 50%, thus indicating robustness in the prediction of the spatial pattern.

The Haitian study was purely retrospective. We attempt a real-time predictive use for another monitored case of cholera epidemic (Thukela river basin in KwaZulu-Natal, South Africa, which was struck by cholera in 2000) (Fig. 2 *A* and *B*). The structure of the water network (287 nodes) is borrowed from a previous work (9). Model parameters have been estimated via Markov chain Monte Carlo sampling (36), using only the first weeks of epidemiological records as calibration data (Fig. 2*A* and *SI Text*). The results are reported in Fig. 2 *C–F*. The dominant eigenvector (associated to $\Lambda_0 = 1.47$) is a good indicator of the spatial patterns of disease spread (Fig. 2*C*), with coefficients of determination $R_0^2 = 0.86$ for disease onset (gray area in Fig. 2*A*) and $R_7^2 = 0.89$ for the epidemic course in 2000–2001. On the contrary, the local reproduction number in excess of one is neither necessary nor sufficient to pinpoint the ensuing epidemic in a spatially connected system. Fig. 2*D*, in fact, shows a map of the local reproduction numbers R_{0i} at the onset of the epidemic that can be contrasted with the spatial distribution of infection cases in excess of 10 (Fig. 2*E*). Strikingly, the comparison shows that several regions where $R_{0i} \leq 1$ still have disease because they are connected to other infected areas. Furthermore, a correlation analysis (details in *SI Text*) between the number of cases and the local reproduction numbers reveals that the distribution of R_{0i} has a much lower explanatory power than the dominant

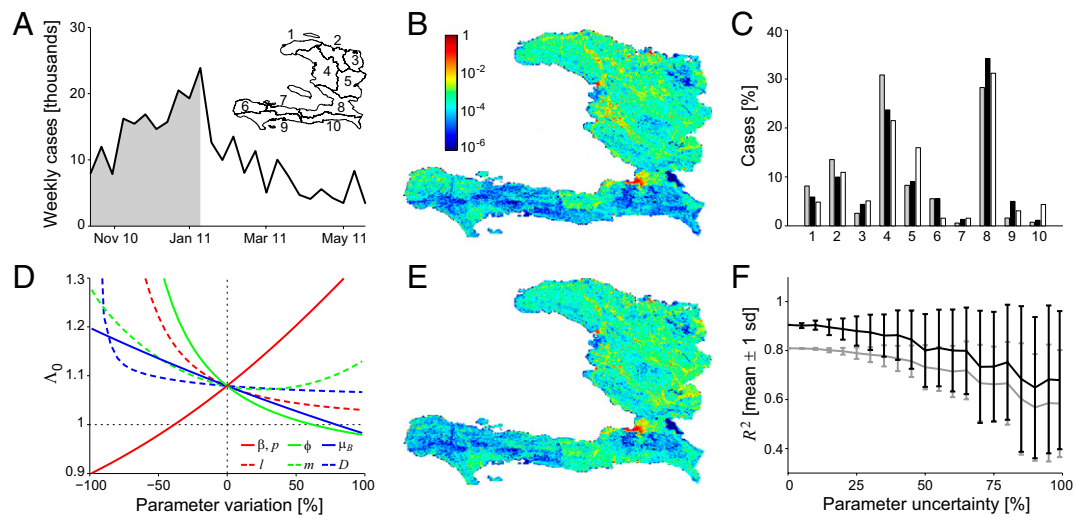


Fig. 1. Data and model predictions of Haiti epidemic. (A) Total incidence data (weekly cases) from November 2010 to May 2011. (B) Fine-grained spatial distribution of recorded cases cumulated during the onset phase of the epidemic (defined as period from beginning to peak, gray area in A). (C) Comparison of coarse-grained data (department level, labels as in A, *Inset*) on spatial distributions of cumulated cases (gray bars, onset phase; black bars, whole period) vs. predictions based on the dominant eigenvector (white bars). $R_O^2 =$ coefficient of determination for the onset phase = 0.81, $R_T^2 =$ coefficient of determination for the whole period = 0.90. (D) Sensitivity to parameter variations of the dominant eigenvalue of G_0 . The dotted horizontal line indicates the value below which the disease cannot start. (E) Fine-grained spatial distribution as predicted by dominant eigenvector (*SI Text*). $R_O^2 = 0.92$, $R_T^2 = 0.95$. (F) Sensitivity to parameter variations of correlations between spatial distribution as predicted by dominant eigenvector and actual spatial distribution of cumulated cases (gray, onset phase; black, whole period). Parameter values and details are in *SI Text*.

eigenvector of G_0 . Similar results hold for the correlation with the local population sizes H_i and the fraction of households without piped water or toilet facilities, as well as for the related model parameters (exposure probabilities β_i and contamination rates p_i ; *SI Text*). Their coefficients of determination cannot compete with those resulting from the eigenvector analysis, thus suggesting that not only local living conditions but also pathogen transport (along waterways and induced by human mobility) must be properly taken into account to understand the formation of the spatial patterns of disease spread.

A sensitivity analysis of the Thukela model (Fig. 2F) shows that the dominant eigenvector is quite a robust spatial indicator of disease incidence. Performances are lower than for the Haiti case study, in which, however, model parameters were calibrated using the whole epidemiological dataset. The dominant eigenvector of G_0 , as calculated at the end of initial calibration, is not only a good indicator, but also a satisfactory predictor of the disease spatial distribution in the months following the calibration's end. In fact, R^2 of the eigenvector components vs. the fractions of cumulative cases from the end of the calibration window to the epidemic peak is 0.91, whereas it is 0.84 if calculated against the cases from the end of calibration to epidemic fading.

Metapopulation Networks

More insight derives from exploring complex theoretical landscapes with different characteristics of river basin network, population distribution, and human mobility. To this end, we construct abstract hydrological networks endowed with a realistic topology, the Peano networks (37–39), which are deterministic fractals (40) whose scaling topology and metrics have been solved analytically (41) (*SI Text*). We study both the spatially homogeneous case (all the site-dependent parameters are identical and so are all the R_{0i}) and the more realistic one in which the population H_i of different communities is distributed according to Zipf's law (42) (*SI Text*), mobility is gravity-like, and water volumes W_i are proportional to H_i . Fig. 3A and B shows the parameter combinations corresponding to disease onset for the homogeneous case. The higher the values are of either l ($1/l$ is the average residence time of pathogens in each node) or the downstream transport bias b (proportional to stream velocity; *SI Text*), the higher the local

reproduction number must be for an epidemic to be possibly triggered (or triggered with high probability in the case of inhomogeneous population distributions, Fig. 3C). Conversely, for low values of l and b the disease can start even if all the R_{0i} are smaller than unity. We term this phenomenon a locally subthreshold epidemic as was done for compartmental models (43). In fact, when $l, m_S, m_I \neq 0$, the local onset conditions ($R_{0i} > 1$) are neither necessary nor sufficient for disease outbreak. Human mobility can remarkably favor disease onset, especially if large population shares move (high m) over either intermediate (homogeneous population, Fig. 3B) or short (Zipf-like distribution, Fig. 3D) distances D . The analysis can be extended to different mobility models based on small-world and scale-free graphs (*SI Text*). The predictive ability of the dominant eigenvector against the spatial distribution of cases for both homogeneous and Zipf-like population distributions is shown in *SI Text*. It is very good at disease emergence and during the onset phase, whereas it usually decreases as the epidemic outbreak unfolds (homogeneous population distribution, $R_E^2 = 0.93$ for disease emergence; $R_O^2 = 0.83$ for the onset phase; $R_T^2 = 0.68$ for the whole epidemic course; and Zipf-like population distribution, $R_E^2 = 0.96$, $R_O^2 = 0.92$, $R_T^2 = 0.88$). This negative trend is especially apparent for the homogeneous case, whereas it is much less evident for Zipf-like population distributions, suggesting that population heterogeneity may play an important role in the definition of long-term spatial patterns of epidemic spread.

Conclusion

We have derived an epidemiological matrix G_0 accounting for both hydrological transport and human mobility. Its dominant eigenvalue Λ_0 is the generalized reproduction number controlling the onset of waterborne disease whereas the dominant eigenvector well characterizes the geography of disease insurgence. Despite the simplicity of the model, the method not only is successful from a theoretical viewpoint, but also proves valid as a tool for analyzing real epidemics (Haiti and KwaZulu-Natal) and inspiring appropriate control measures and sanitation actions, which are the subject of ongoing debate (44, 45). In particular, the model is a good compromise between traditional space-homogeneous approaches (13) and individual-based simulation (5), thus coupling realism with a sufficiently simple structure that allows statistically significant

dynamics of infectives (second equation) are a balance between newly infected individuals and losses due to recovery or natural/pathogen-induced mortality, with γ and α being recovery and disease-induced mortality rates, respectively. The dynamics of recovered individuals are neglected, because waterborne diseases confer at least temporary immunity and its loss neither determines conditions for disease onset nor affects its evolution in the immediate development following the epidemic peak. The dynamics of free-living pathogens concentration (third equation) assume that bacteria or protozoa are released in water (e.g., excreted) by infective individuals and immediately diluted in a well-mixed local water reservoir of volume W_i at a site-dependent rate p_i . Free-living pathogens are also assumed to die at a constant, site-independent rate μ_B . Regarding the hydrological transport, the spread of pathogens over the river network is described as a biased random-walk process on an oriented graph (9). Here, we assume that pathogens can move at a rate l from node i to node j of the hydrological network with a probability P_{ij} . The rate depends on both downstream advection and other possible pathogen transport pathways along the hydrological network, e.g., short-range distribution of water for consumption or irrigation or bacterial attachment to phyto- and zooplankton. Possible topological structures for the hydrological network range from simple one-dimensional lattices to realistic mathematical characterizations of existing river networks. The nodes of the human mobility network are assumed to correspond to those of the hydrological layer, whereas edges are defined by connections among communities. We also assume that susceptible and infective individuals can undertake 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 the settlement where they usually live. Similarly, infected hosts can disseminate the disease away from their home community. In many cases infected individuals are asymptomatic and thus are not barred, or are only partially 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 (respectively m_S for susceptibles and m_I for infectives, 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 networks used in epidemiology can be based on suitable measures of node-to-node distance like in gravity models (34); on the actual transportation network; or on models based on conceptually different interactions, such as Erdős-Rényi random graphs (51), scale-free networks (52), small-world-like graphs (53), and radiation models (54).

Derivation of Onset Conditions. To analyze stability, we consider the Jacobian of the linearized system evaluated at the disease-free equilibrium X_0 ($S_i = H_i$, $I_i = 0$, $B_i = 0$ for all i), which is given by

$$J = \begin{bmatrix} j_{11} & 0 & j_{13} \\ 0 & j_{22} & j_{23} \\ 0 & j_{32} & j_{33} \end{bmatrix},$$

where

$$\begin{aligned} j_{11} &= -\mu U_n \\ j_{13} &= -m_S H Q \beta - (1 - m_S) H \beta \\ j_{22} &= -\phi U_n \\ j_{23} &= m_S H Q \beta + (1 - m_S) H \beta \\ j_{32} &= \frac{m_I}{K} p W^{-1} Q^T + \frac{1 - m_I}{K} p W^{-1} \\ j_{33} &= -(\mu_B + l) U_n + l W^{-1} P^T W. \end{aligned}$$

Note that the variables for pathogen have been scaled as $B_i^* = B_i/K$. 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

$$J^* = \begin{bmatrix} j_{22} & j_{23} \\ j_{32} & j_{33} \end{bmatrix}.$$

J^* is a proper Metzler matrix (55); namely its off-diagonal entries are all nonnegative and at least one diagonal entry is negative. Thus its eigenvalue with maximal real part (dominant eigenvalue) is real. If the union of the graphs associated with matrices P and Q is strongly connected, then the graph associated with J^* is also strongly connected. Therefore one can apply the Perron-Frobenius theorem (56) for irreducible matrices and

state that the dominant eigenvalue is a simple real root of the characteristic polynomial. The condition for the transcritical bifurcation of the disease-free equilibrium is that the dominant eigenvalue crosses the imaginary axis at zero; namely, the determinant of J^* is zero (33). Actually, when the disease-free equilibrium is stable, all the eigenvalues have negative real parts and $\det(J^*)$ is positive because J^* is a matrix of order $2n$. So the disease-free equilibrium becomes unstable when $\det(J^*)$ switches from positive to negative or equivalently the dominant eigenvalue becomes zero.

As U_n obviously commutes with any matrix, we have (ref. 57 and *SI Text*)

$$\begin{aligned} \det(J^*) &= \det \left[\phi (\mu_B + l) U_n - \phi l W^{-1} P^T W + \right. \\ &\quad - \frac{m_S m_I}{K} p W^{-1} Q^T H Q \beta - \frac{m_I (1 - m_S)}{K} p W^{-1} Q^T H \beta + \\ &\quad \left. - \frac{(1 - m_I) m_S}{K} p W^{-1} H Q \beta - \frac{(1 - m_I) (1 - m_S)}{K} p W^{-1} H \beta \right]. \end{aligned}$$

Because H , β , and W^{-1} are diagonal, thus commuting, matrices, we can state that $R_0 = \frac{p}{K \mu_B \phi} H \beta W^{-1} = \frac{p}{K \mu_B \phi} W^{-1} H \beta$ and rework the determinant of J^* (*SI Text*). The condition $\det(J^*) = 0$ is thus given by

$$\begin{aligned} \det \left\{ U_n - \frac{l}{\mu_B + l} P^T - \frac{\mu_B}{\mu_B + l} \left[(1 - m_I) (1 - m_S) R_0 + \frac{m_S m_I}{K \mu_B \phi} p Q^T H Q \beta W^{-1} + \right. \right. \\ \left. \left. + \frac{m_I (1 - m_S)}{K \mu_B \phi} p Q^T H \beta W^{-1} + \frac{(1 - m_I) m_S}{K \mu_B \phi} p H Q \beta W^{-1} \right] \right\} = 0. \end{aligned}$$

In addition to the matrix $R_0 = \frac{p}{K \mu_B \phi} H \beta W^{-1}$ we can now introduce three other matrices of reproduction numbers; namely,

$$R_0^I = \frac{p Q^T H \beta W^{-1}}{K \mu_B \phi}, \quad R_0^S = \frac{p H Q \beta W^{-1}}{K \mu_B \phi}, \quad \text{and} \quad R_0^{IS} = \frac{p Q^T H Q \beta W^{-1}}{K \mu_B \phi},$$

corresponding to metacommunities with infectives only being mobile, susceptibles only being mobile, and both infectives and susceptibles being mobile, respectively. If we account for the different probabilities of movement in the metacommunity, we can define a transmission matrix averaged over nonmobile individuals, mobile infectives, and mobile susceptibles as

$$T_0 = (1 - m_I) (1 - m_S) R_0 + m_S m_I R_0^{IS} + m_I (1 - m_S) R_0^I + (1 - m_I) m_S R_0^S.$$

Therefore, the bifurcation of the disease-free equilibrium corresponds to the condition

$$\det \left(U_n - \frac{l}{\mu_B + l} P^T - \frac{\mu_B}{\mu_B + l} T_0 \right) = 0.$$

Equivalently, the dominant eigenvalue Λ_0 of the matrix

$$G_0 = \frac{l}{\mu_B + l} P^T + \frac{\mu_B}{\mu_B + l} T_0,$$

which is a convex combination of P^T and T_0 , must equal unity. Actually, the disease-free equilibrium switches from being stable to being a saddle, thus triggering the start of the disease, whenever the dominant eigenvalue of J^* switches from positive to negative and hence whenever Λ_0 switches from being less than 1 to being larger than 1.

ACKNOWLEDGMENTS. The authors thank Paolo Gatto for technical suggestions on numerical simulations. L.M., E.B., L.R., and A.R. acknowledge the support provided by the European Research Council (ERC) advanced grant program through the project "River networks as ecological corridors for biodiversity, populations and waterborne disease" (RINEC-227612) and by the Swiss National Science Foundation (SNF/FNS) projects 200021_124930/1 and CR2312_138104/1. M.G. and A.R. acknowledge the support from the SNF/FNS project IZK022_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).

1. World Health Organization (2008) *The Global Burden of Disease: 2004 Update* (WHO, Geneva).
2. Bertuzzo E, et al. (2011) Prediction of the spatial evolution and effects of control measures for the unfolding Haiti cholera outbreak. *Geophys Res Lett* 38: L06403.
3. Mari L, et al. (2012) Modelling cholera epidemics: The role of waterways, human mobility and sanitation. *J R Soc Interface* 9(67):376–388.
4. Tuite AR, et al. (2011) Cholera epidemic in Haiti, 2010: Using a transmission model to explain spatial spread of disease and identify optimal control interventions. *Ann Intern Med* 154(9):593–601.
5. Chao DL, Halloran ME, Longini IM, Jr. (2011) Vaccination strategies for epidemic cholera in Haiti with implications for the developing world. *Proc Natl Acad Sci USA* 108(17):7081–7085.
6. Rinaldo A, et al. (2012) Reassessment of the 2010–2011 Haiti cholera outbreak and multi-season projections via inclusion of rainfall and waning immunity. *Proc Natl Acad Sci USA* 109:6602–6607.
7. Mukandavire Z, et al. (2011) Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe. *Proc Natl Acad Sci USA* 108(21):8767–8772.
8. Rohani P, Earn DJ, Grenfell BT (1999) Opposite patterns of synchrony in sympatric disease metapopulations. *Science* 286(5441):968–971.
9. Bertuzzo E, et al. (2008) On the space-time evolution of a cholera epidemic. *Water Resour Res* 44:W01424.
10. Bertuzzo E, Casagrandi R, Gatto M, Rodriguez-Iturbe I, Rinaldo A (2010) On spatially explicit models of cholera epidemics. *J R Soc Interface* 7(43):321–333.
11. Pitzer VE, et al. (2009) Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* 325(5938):290–294.
12. Capasso V, Paveri-Fontana SL (1979) A mathematical model for the 1973 cholera epidemic in the European Mediterranean region. *Rev Epidemiol Sante Publique* 27(2): 121–132.
13. Codeço CT (2001) Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir. *BMC Infect Dis* 1:1.
14. Anderson R, May R (1991) *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ Press, Oxford).
15. Hartley DM, Morris JG, Jr., Smith DL (2006) Hyperinfectivity: A critical element in the ability of *V. cholerae* to cause epidemics? *PLoS Med* 3(1):e7.
16. Jensen MA, Faruque SM, Mekalanos JJ, Levin BR (2006) Modeling the role of bacteriophage in the control of cholera outbreaks. *Proc Natl Acad Sci USA* 103(12): 4652–4657.
17. Colwell RR (1996) Global climate and infectious disease: The cholera paradigm. *Science* 274(5295):2025–2031.
18. Pascual M, Bouma MJ, Dobson AP (2002) Cholera and climate: Revisiting the quantitative evidence. *Microbes Infect* 4(2):237–245.
19. Lipp EK, Huq A, Colwell RR (2002) Effects of global climate on infectious disease: The cholera model. *Clin Microbiol Rev* 15(4):757–770.
20. Koelle K, Rodó X, Pascual M, Yunus M, Mostafa G (2005) Refractory periods and climate forcing in cholera dynamics. *Nature* 436(7051):696–700.
21. Reiner RC, Jr., et al. (2012) Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. *Proc Natl Acad Sci USA* 109(6):2033–2036.
22. Righetto L, et al. (2012) The role of aquatic reservoir fluctuations in long-term cholera patterns. *Epidemics* 4(1):33–42.
23. World Health Organization (2010) *Prevention and Control of Cholera Outbreaks: WHO Policy and Recommendations*, Technical Report (WHO Global Task Force on Cholera Control, Geneva, Switzerland).
24. Guerrant R (1986) The global problem of amebiasis: Current status, research needs, and opportunities for progress. *Rev Infect Dis* 8:218–227.
25. Righetto L, et al. (2011) Modelling human movement in cholera spreading along fluvial systems. *Ecohydrology* 4:49–55.
26. Mari L, et al. (2012) On the role of human mobility in the spread of cholera epidemics: Towards an epidemiological movement ecology. *Ecohydrology* 5(5):531–540.
27. Rodriguez-Iturbe I, Rinaldo A (1997) *Fractal River Basins: Chance and Self-Organization* (Cambridge Univ Press, New York).
28. Grad YH, Miller JC, Lipsitch M (2012) Cholera modeling: Challenges to quantitative analysis and predicting the impact of interventions. *Epidemiology* 23(4):523–530.
29. Diekmann O, Heesterbeek JAP, Roberts MG (2010) The construction of next-generation matrices for compartmental epidemic models. *J R Soc Interface* 7(47):873–885.
30. Roberts MG, Heesterbeek JAP (2003) A new method for estimating the effort required to control an infectious disease. *Proc Biol Sci* 270(1522):1359–1364.
31. Lopez LF, Coutinho FAB, Burattini MN, Massad E (2002) Threshold conditions for infection persistence in complex host-vectors interactions. *C R Biol* 325(11):1073–1084.
32. Diekmann O, Heesterbeek JA, Metz JA (1990) On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J Math Biol* 28(4):365–382.
33. Kuznetsov Y (1995) *Elements of Applied Bifurcation Theory* (Springer, New York).
34. Erlander S, Stewart NF (1990) *The Gravity Model in Transportation Analysis – Theory and Extensions* (VSP Books, Zeist, The Netherlands).
35. Nelson EJ, Nelson DS, Salam MA, Sack DA (2011) Antibiotics for both moderate and severe cholera. *N Engl J Med* 364(1):5–7.
36. Gilks WR, Richardson S, Spiegelhalter DJ (1995) *Markov Chain Monte Carlo in Practice* (Chapman & Hall, New York).
37. Peano G (1890) Sur une courbe qui remplit toute une air plane [On an area-filling curve]. *Math. Ann* 36:157–160.
38. Campos D, Fort J, Méndez V (2006) Transport on fractal river networks: Application to migration fronts. *Theor Popul Biol* 69(1):88–93.
39. 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 Resour Res* 43:W04419.
40. Mandelbrot B (1983) *The Fractal Geometry of Nature* (Freeman, New York).
41. Marani A, Rigon R, Rinaldo A (1991) A note on fractal channel networks. *Water Resour Res* 27:3041–3049.
42. Zipf GK (1949) *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology* (Addison-Wesley, Cambridge, MA).
43. van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci* 180:29–48.
44. Farmer P, et al. (2011) Meeting cholera's challenge to Haiti and the world: A joint statement on cholera prevention and care. *PLoS Negl Trop Dis* 5(5):e1145.
45. Walton DA, Ivers LC (2011) Responding to cholera in post-earthquake Haiti. *N Engl J Med* 364(1):3–5.
46. Bertuzzo E, et al. (2012) Hydroclimatology of dual-peak cholera epidemics: Inferences from a spatially explicit model. *Geophys Res Lett* 39:L05403.
47. Siddique AK, et al. (2011) Epidemiology of rotavirus and cholera in children aged less than five years in rural Bangladesh. *J Health Popul Nutr* 29(1):1–8.
48. Vezzulli L, Pruzzo C, Huq A, Colwell R (2010) Environmental reservoirs of *Vibrio cholerae* and their role in cholera. *Environ Microbiol Rep* 2:27–33.
49. Lara R, et al. (2011) *Vibrio cholerae* in waters of the Sunderban mangrove: Relationship with biogeochemical parameters and chitin in seston size fractions. *Wetlands Ecol Manage* 19:109–119.
50. Faruque SM, et al. (2003) Pathogenic potential of environmental *Vibrio cholerae* strains carrying genetic variants of the toxin-coregulated pilus pathogenicity island. *Infect Immun* 71(2):1020–1025.
51. Erdős P, Rényi A (1959) On random graphs. *Publicationes Mathematicae* 6:290–297.
52. Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74:4797.
53. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393(6684):440–442.
54. Simini F, González MC, Maritan A, Barabási A-L (2012) A universal model for mobility and migration patterns. *Nature* 484(7392):96–100.
55. Farina L, Rinaldi S (2000) *Positive Linear Systems: Theory and Applications* (Wiley Interscience, New York).
56. Gantmacher F (1959) *Theory of Matrices* (AMS Chelsea Publishing, Providence, RI).
57. Silvester J (2000) Determinants of block matrices. *Math Gazette* 84:460–467.