



UNIVERSITY OF LEEDS

This is a repository copy of *On the space-time evolution of a cholera epidemic*.

White Rose Research Online URL for this paper:

<http://eprints.whiterose.ac.uk/78955/>

Version: Published Version

Article:

Bertuzzo, E, Azaele, S, Maritan, A et al. (3 more authors) (2008) On the space-time evolution of a cholera epidemic. *Water Resources Research*, 44 (1). W01424. ISSN 0043-1397

<https://doi.org/10.1029/2007WR006211>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

On the space-time evolution of a cholera epidemic

E. Bertuzzo,^{1,4,5} S. Azaele,^{2,4} A. Maritan,^{2,6} M. Gatto,³ I. Rodriguez-Iturbe^{1,4}
and A. Rinaldo^{1,5}

Received 30 May 2007; revised 22 September 2007; accepted 5 October 2007; published 19 January 2008.

[1] We study how river networks, acting as environmental corridors for pathogens, affect the spreading of cholera epidemics. Specifically, we compare epidemiological data from the real world with the space-time evolution of infected individuals predicted by a theoretical scheme based on reactive transport of infective agents through a biased network portraying actual river pathways. The data pertain to a cholera outbreak in South Africa which started in 2000 and affected in particular the KwaZulu-Natal province. The epidemic lasted for 2 years and involved about 140,000 confirmed cholera cases. Hydrological and demographic data have also been carefully considered. The theoretical tools relate to recent advances in hydrochory, migration fronts, and infection spreading and are novel in that nodal reactions describe the dynamics of cholera. Transport through network links provides the coupling of the nodal dynamics of infected people, who are assumed to reside at the nodes. This proves a realistic scheme. We argue that the theoretical scheme is remarkably capable of predicting actual outbreaks and, indeed, that network structures play a controlling role in the actual, rather anisotropic propagation of infections, in analogy to spreading of species or to migration processes that also use rivers as ecological corridors.

Citation: Bertuzzo, E., S. Azaele, A. Maritan, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo (2008), On the space-time evolution of a cholera epidemic, *Water Resour. Res.*, 44, W01424, doi:10.1029/2007WR006211.

1. Introduction

[2] Cholera is an intestinal disease caused by the bacterium *Vibrio cholerae*, which colonizes the human intestine. The dynamics of cholera epidemics have been studied since the 1800s, when John Snow established the link between cholera cases and exposure to contaminated water of a well in London. *V. cholerae* is also a natural member of the aquatic microbial community [Colwell, 1996; Lipp *et al.*, 2002]. Thus the spatial and temporal patterns of cholera epidemics are strongly related to the ecology of the bacterium in the environment which is itself driven by meteorological and climatic variability. Time series analyses of cholera cases in endemic regions, such as Bangladesh, show a time variability with subannual, annual, and interannual components. The low-frequency variability has been established to be related to long-term climatic oscillations [Pascual *et al.*, 2000; Koelle *et al.*, 2005]. In nonendemic regions the annual component is much more important. Also, the spatial distribution of the disease plays a fundamental role, which is,

however, usually neglected in the existing cholera models. The aim of this paper is to understand the spatio-temporal evolution of cholera by explicitly accounting for the environmental matrix within which the disease can spread.

[3] The model will be tested against a very well documented case. In 2000, after several years without cholera outbreaks a new epidemic spread in South Africa, affecting in particular the KwaZulu-Natal province. The epidemic lasted for 2 years with only a few cases recorded during the third year, and ultimately involved about 140,000 confirmed cholera cases. The epidemic was caused by the 01 el Tor strain [Mugero and Hoque, 2001], which is more easily transmitted through contamination of aquatic environments than the classical biotype of *Vibrio cholerae*.

[4] Models of cholera dynamics are relatively recent. Capasso and Paveri-Fontana [1979] proposed a mathematical model to describe the 1973 cholera epidemic in Bari (Italy) with two equations describing the dynamics of the infected population and the free-living pathogens. Codeço [2001] extended Capasso and Paveri-Fontana's model, adding an equation for the dynamics of the susceptible population, and studied the role of the aquatic reservoir in the endemic-epidemic dynamics of cholera. In the model proposed by Pascual *et al.* [2002], another equation is added to describe the temporal evolution of the volume of water hosting the free-living bacteria. Recent laboratory findings suggest that passage of the bacterium through the gastrointestinal tract results in a short-lived hyperinfectious state that can enhance the human-to-human versus environmental-to-human transmission of cholera. Hartley *et al.* [2006] incorporate the hyperinfectious state into Codeço's model to achieve a better explanation of explosive cholera outbreaks.

¹Dipartimento di Ingegneria Idraulica, Marittima Ambientale e Neotenia, Università di Padova, Padua, Italy.

²Dipartimento di Fisica "G. Galilei," Università di Padova, Padua, Italy.

³Dipartimento di Elettronica e Informazione, Politecnico di Milano, Milan, Italy.

⁴Now at Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

⁵International Center for Hydrology "Dino Tonini," Università di Padova, Padua, Italy.

⁶Now at Consorzio Interuniversitario per le Scienze Fisiche della Materia, Istituto Nazionale Fisica della Materia, Padua, Italy.

Table 1. Description of the Symbols Used in the Text

Symbol	Description	Value	Note
S_i	number of susceptibles at node i		equation (1)
I_i	number of infected at node i		equation (1)
B_i	concentration of <i>V. cholerae</i> in aquatic environment ^a		equation (1)
C_i	number of cumulated cases of node i		equation (7)
H_i	total human population size of node i at the disease-free equilibrium		input data
n	population natality and mortality rate ^b	5×10^{-5}	estimated
a	rate of exposure to contaminated water ^b	1	estimated
K	concentration of <i>V. cholerae</i> in water that yields 50% chance of being infected with cholera ^a		
r	rate at which people recover from cholera ^b	0.2	estimated
m	mortality rate due to cholera ^b	4×10^{-4}	estimated
n_B	net growth rate (usually negative) of <i>V. cholerae</i> in the aquatic environment ^b	-0.228	calibrated
p	rate of production by one person infected of <i>V. cholerae</i> that reach the water body ^c		
W_i	volume of water ^d at node I		
S_{C_i}	critical threshold of node i		equation (3)
b	transport bias ($P_{out} - P_{in}$)	0.08	calibrated
l	<i>V. cholera</i> mobility ^b	3.5	calibrated
V_i	vulnerability of node i : H_i/S_{C_i}		equation (6)
c	per capita water volume ^e		
H_T	threshold on the node population	29,000	calibrated
$p/(Kc)$	combined parameters ratio ^b	4.76×10^{-6}	calibrated

^aMeasured in cells per m^{-3} .

^bMeasured per day.

^cMeasured in cells per day per person.

^dMeasured in m^3 .

^eMeasured in m^3 per person.

[5] All the above models do not consider space explicitly. They assume a unique community of people who interact and share the same resources. We believe, however, that the spatial distribution of the communities and how they interact is crucial to understanding the spatial spreading of the epidemic in a disease-free region, particularly if travel times of pathogens are comparable with the characteristic time of the virulent infection. The spatial distribution of different communities, along with the distribution of their population size and how they are interconnected, could indeed affect the dynamics of the process, especially in the case of a nonendemic region.

[6] *V. cholerae* can survive in the aquatic environment in associations with chitinous zooplankton like copepods and shellfish and also with the aquatic vegetation [Colwell, 1996]. Therefore *V. cholerae* (and the disease) can spread from the coastal region, where it is autochthonous, to the inland area through waterways and river networks. In the same manner the infection can spread from inland regions with epidemic outbursts into the surrounding areas.

[7] We propose a model that explicitly accounts for the role of the river network in transporting and redistributing *V. cholerae* between several human communities, and we proceed to apply the model to the real conditions of the 2000 cholera epidemic in the KwaZulu-Natal province. The model explicitly recognizes a role for network structures acting as support for the infection, in analogy to recent studies on migrating fronts constrained by landscape heterogeneities or spreading of species along riparian ecological corridors [Campos et al., 2006; Bertuzzo et al., 2007; Muneeppeerakul et al., 2007].

[8] The paper is organized as follows. Section 2 describes the theoretical approach and the model used in detail. The complete data set and the case study are presented in a specific chapter (section 3). The main results and the

discussion are collected in section 4. A set of conclusions closes the paper.

2. Theoretical Approach

[9] Spreading of epidemics in networks is addressed by viewing the environmental matrix as an oriented graph (i.e., a directed graph having no symmetric pair of directed edges). Nodes represent human communities (cities, towns, and villages) in which the disease can be diffused and grow. The edges represent links between the communities, typically hydrological links. Edge direction is chosen accordingly to the flow direction. The model is assembled by coupling two models: (1) a local epidemic model at nodes of the graph and (2) a transport model for the spreading of the disease vector through the edges of the support. Details of the two models follow.

[10] As for the local dynamics, we use a continuous model of the susceptible, infected, and recovered class with a reservoir of free-living infective propagules. It is obtained by a slight modification of the cholera epidemic model introduced by Codeço [2001]. The model has three state variables: the number of susceptibles (S), the number of infected (I), and the concentration of *V. cholerae* in the aquatic environment (B), whose respective temporal dynamics are described by the following system of first-order differential equations:

$$\begin{aligned}
 \frac{dS}{dt} &= n(H - S) - a \frac{B}{K + B} S \\
 \frac{dI}{dt} &= a \frac{B}{K + B} S - (r + m + n)I \\
 \frac{dB}{dt} &= n_B B + \frac{p}{W} I
 \end{aligned} \tag{1}$$

[11] The meaning of the parameters is explained in Table 1. The first equation describes the dynamics of susceptibles in

a community of size H . Susceptible individuals are born and die on average at rate n . Newborn individuals are considered susceptible. Susceptible people become infected at a rate $aB/(K+B)$, where a is the rate of contact with contaminated water and $B/(K+B)$ is a logistic dose response curve that links the probability of becoming infected to the concentration of vibrios B in water. Infected people (whose dynamics are described by the second equation) die at a rate which is the sum of natural mortality n and disease-caused mortality m and recover with rate r . The third equation describes the dynamics of the free-living infective propagules in the reservoir. Infected people contribute to the concentration of vibrios at a rate p/W , where p is the rate at which bacteria are produced by one infected person and W is the volume of the contaminated water body. The growth rate n_B of the free-living bacteria in the water body is usually negative because bacteria mortality in natural environments exceeds reproduction. If n_B is positive, the model would predict an exponential growth of the vibrios concentration, and all the susceptible population would be affected by the disease. The hidden equation for the recovered is $(dR/dt) = rI - nR$. People recovered from cholera are considered immune. The model does not take into account any loss of immunity (i.e., a flux from recovered to susceptibles) because immunity usually lasts for a period longer than the 2 years of the epidemic we consider here [Koelle et al., 2005]. Nevertheless, the immunity loss could play an important role in the dynamics of cholera in regions where it is endemic [Koelle et al., 2005].

[12] As long as we are not interested in the numerical value of the concentration B of *V. cholerae*, we can introduce the dimensionless concentration $B^* = B/K$, thus obtaining from equation (1) the system of equations

$$\begin{aligned} \frac{dS}{dt} &= n(H - S) - a \frac{B^*}{1 + B^*} S \\ \frac{dI}{dt} &= a \frac{B^*}{1 + B^*} S - (r + m + n)I \\ \frac{dB^*}{dt} &= n_B B^* + \frac{p}{KW} I. \end{aligned} \quad (2)$$

Notice that equation (2) has the advantage of merging the parameters K , p , and W (which can hardly be directly estimated) into a unique ratio. This ratio will be the control parameter of the process jointly with the *V. cholerae* growth rate n_B . On the contrary, the mortality rates, both natural (n) and due to cholera (m); the recovery rate (r); and the exposure rate (a) can reasonably be estimated from demographic and epidemiological studies, as we show below.

[13] A linear stability analysis shows that, given an initial condition of the type $S(0) = H$; $I(0) > 0$; $B^*(0) = 0$, the model predicts an epidemic outbreak only if the population size is greater than a certain critical threshold S_C given by [Codeço, 2001]

$$H > S_C = \frac{-(r + n + m)Kn_B W}{ap}; \quad (3)$$

otherwise the infected population decreases to zero. It is important to remark on the dilution effect: The larger the volume of the water body is, the higher the critical threshold will be.

[14] We model the spreading of *V. cholerae* through the network with a biased random walk process on oriented graph [Bertuzzo et al., 2007]. For a detailed discussion of the process, see also Johnson et al. [1995]. An infectious propagule can move with some probability from a node to one of the adjacent nodes, which are all the nodes that are connected to it through an inward or outward edge. We assign to each edge of the graph an orientation according to the flow direction. Consider first a particular case of the network in which every node has only one inward and one outward edge (i.e., a one-dimensional lattice). We define as P_{out} (P_{in}) the probability that a propagule leaving a node moves to another node along an outward (inward) edge. We have then $P_{\text{out}} + P_{\text{in}} = 1$.

[15] We now turn to the analysis of a random walk process on a generic-oriented graph in which every node can have an arbitrary number of inward and outward edges. We assume that a propagule can move following an outward or inward edge with a probability proportional to P_{out} and P_{in} , respectively. In this case, the probability P_{ij} for a propagule to be transported from node i to node j can be expressed as follows:

$$P_{ij} = \begin{cases} \frac{P_{\text{out}}}{d_{\text{out}}(i)P_{\text{out}} + d_{\text{in}}(i)P_{\text{in}}} & \text{if } i \rightarrow j \\ \frac{P_{\text{in}}}{d_{\text{out}}(i)P_{\text{out}} + d_{\text{in}}(i)P_{\text{in}}} & \text{if } i \leftarrow j \\ 0 & \text{if } i \not\leftrightarrow j, \end{cases} \quad (4)$$

where $d_{\text{out}}(i)$ and $d_{\text{in}}(i)$ are the outdegree and indegree of node i , respectively (i.e., the number of outward and inward edges of node i , respectively). Since $P_{\text{out}} + P_{\text{in}} = 1$, one has $\sum_{j=1}^N P_{ij} = 1$, where N is the total number of nodes. We term $b = P_{\text{out}} - P_{\text{in}} = 2P_{\text{out}} - 1$ the bias of the transport.

[16] When we apply the local epidemic model at each node of the network, we have $3N$ state variables S_i , I_i , and B_i^* , where the subscript i identifies the nodes. We assume that vibrios are removed at every node with a certain rate l (d^{-1}) and transported through the network following the transition probabilities from equation (4). Then the equations that describe the coupled process are

$$\begin{aligned} \frac{dS_i}{dt} &= n(H_i - S_i) - a \frac{B_i^*}{1 + B_i^*} S_i \\ \frac{dI_i}{dt} &= a \frac{B_i^*}{1 + B_i^*} S_i - (r + m + n)I_i \\ \frac{dB_i^*}{dt} &= n_B B_i^* + \frac{p}{KW_i} I_i - l B_i^* + \sum_{j=1}^N l P_{ji} B_j^* \frac{W_j}{W_i}, \end{aligned} \quad (5)$$

for $i = 1, 2, \dots, N$. Note that all the parameters are node-independent except for the population size H_i and the water volume W_i . The latter represents the whole set of water supplies available for that community, not only the one provided by the river. The network acts as a link through which different sets of water supplies of different communities can be connected and contaminated. In order to further minimize the number of parameters, we assume that the water volume is a nondecreasing function of the population size: $W_i = f(H_i)$. Different choices of the function f can lead to different scenarios of the epidemic. In fact,

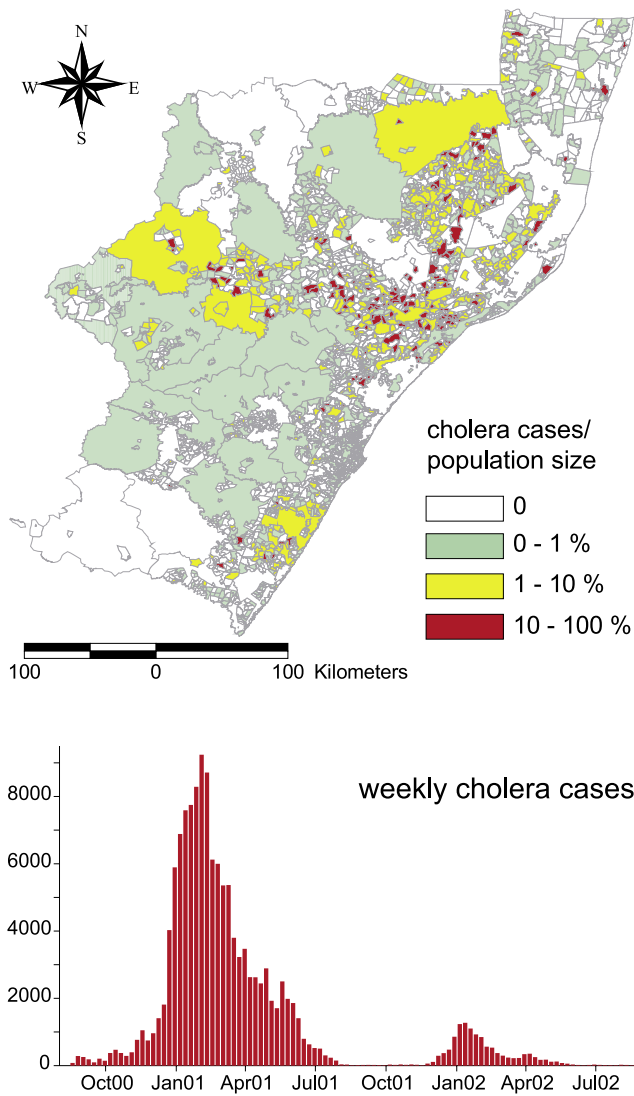


Figure 1. (a) Spatial representation of the health districts of the KwaZulu-Natal province, South Africa. Colors represent the percentage incidence of cholera cases over the population size of each district. (b) Also shown is temporal evolution of the new weekly cholera cases for the whole province.

consider the ratio V_i between the population size and the critical threshold equation (3).

$$V_i = \frac{H_i}{S_{C_i}} = \frac{H_i a p}{-(r+n+m) K n_{Rf}(H_i)} \propto \frac{H_i}{f(H_i)}. \quad (6)$$

[17] This is an index of the node vulnerability to an epidemic. Let us first analyze the case in which the water availability is constant for all the nodes ($W_i = \text{constant}$). This corresponds to assuming that the communities utilize water resources that are quite uniformly distributed in space. In this case, the vulnerability of the nodes increases linearly with the population size, namely, $V_i \propto H_i$. In this scenario then, if an epidemic occurs, the most affected communities would be the most populated. Another different scenario derives from assuming that larger communities manage to increase their own water supply so that the per capita

available water is kept constant. This is equivalent to assuming that the water volume of a node is proportional to the population size: $W_i \propto H_i$, and then $V_i = \text{constant}$. Under such a scenario an epidemic would affect, even if with different dynamics, all the communities regardless of their size. This assumption seems reasonable for large and developed communities, but it is unsatisfactory for the nodes with small population size because it would imply a small water body associated with the node regardless of the natural presence of water. A more general and realistic hypothesis could derive from the combination of the two assumptions described above. In particular, we assume that the nodes with population size H_i smaller than a certain threshold H_T have a constant water volume associated with them, whereas the nodes with $H_i > H_T$ have a constant per capita water availability. Summarizing, we have $W_i = \max(cH_T, cH_i) = c \times \max(H_T, H_i)$, where c is the per capita volume of water resources.

[18] Substituting the last relationship into the second term of the right-hand side of the third equation of system (5), we get $pI_i/(K c \times \max(H_T, H_i))$. Thus the parameters to be separately estimated are the ratio $p/(K c)$ and the population threshold H_T . These are parameters that depend on social and environmental factors, like hygiene, health conditions, eating habits, and lifestyle, and on how these variables vary with population density. Nevertheless, the particular choice of parameters that link vulnerability to the actual distribution of population has to be evaluated for each case.

3. Case Study

[19] We apply the model to a well-documented case of cholera epidemics that occurred in South Africa. The data were provided by the KwaZulu-Natal Health Department and consist of a record of each single cholera case specified by the date and health subdistrict where it occurred. The spatial representation of the districts is shown in Figure 1a. The record starts from August 2000 and runs continuously until present time. Our analysis focuses on the two largest epidemic outbreaks which occurred during the 2000–2001 and 2001–2002 summers and involved 135,000 cholera cases. The data set also provides the population size of each district. The total population of the province is about 8.5 million inhabitants. The temporal evolution of the weekly cholera cases is reported in Figure 1b. The data exhibit a clear seasonality with the outbreaks occurring during the warmest months of the austral summer. This is probably due to the increased growth rate of vibrios at warm temperatures in association with plankton blooms. Evidence for this phenomenon comes from the higher rates of isolation of the bacterium in the environment during warm periods [Lipp *et al.*, 2002].

[20] In order to apply the model, we need to build the network along which infection is transported. First of all, we have derived the mathematical model of the river networks from the hydrological geographic information systems data provided by the South Africa Department of Water Affairs and Forestry shown in Figure 2. All the channels of perennial rivers are considered edges, and all the endpoints of these channels are considered as nodes. Second, we had to transfer the information from districts to network nodes. This is done by assigning the population and the cholera cases of a subdistrict to the nearest network node, with the

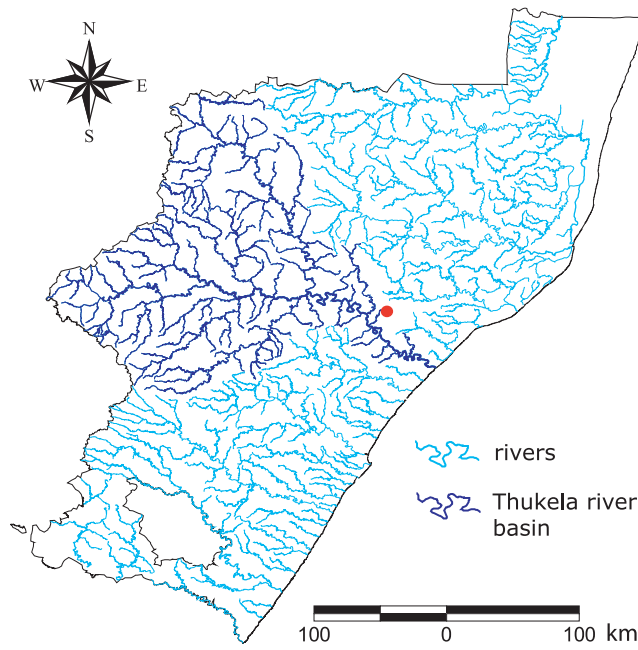


Figure 2. Hydrographic map of KwaZulu-Natal province with the Thukela river basin evidenced. The dot reports the location of the first epidemic outbreak in the basin studied.

distances being computed from the centroid of the subdistrict to the node. The results of this interpolation for the population and the total cumulated cholera cases are shown in Figure 3, where the color coding is obtained by spatial linear interpolation of the node values. Comparing the spatial distribution of population sizes and total cases in the nodes, one can note that the high population density areas recorded few cases of cholera as did the low-density ones. The most affected nodes are those with intermediate population size. This clearly appears by plotting the cholera average incidence (i.e., the total number of cases divided by the population size) as a function of the population size (see Figure 4). The highest incidence was recorded for population sizes between 2000 and 30,000. This is probably due to the fact that the highest population density regions correspond, in this particular case, with the most developed ones. These cities can then rely on wastewater treatment and treated water supply that help to reduce cholera transmission.

[21] The framework described in this paper addresses the spread of an epidemic in a single river basin, and for the time being we avoid any modeling of the flux of bacteria across different catchments. For this reason, in order to test the validity of the model, we have applied it to the basin of river Thukela, the largest of the region (see Figure 2). The 29,000 km² area drained by this river is populated by 1.5 million people, and cholera cases recorded there amounted to 29,000 (21% of the total cases of the whole province) during the two epidemic outbreaks considered.

[22] We estimated the birth and mortality rate of the population as the inverse of the average lifetime for this region (about 60 years), so $n \simeq 5 \times 10^{-5} \text{ d}^{-1}$. Because the average duration of the cholera disease in an infected person is approximately 5 d [Codeço, 2001; Hartley et al., 2006], we set the recovery rate at $r = 0.2 \text{ d}^{-1}$. The deaths due to cholera for the epidemic analyzed were 0.2% of the cholera

cases. Thus we can estimate the cholera mortality rate m by assuming that after the duration of the disease, 99.8% of the infected population survive, that is, $\exp(-m/r) = 0.998$, and then $m = 4 \times 10^{-4} \text{ d}^{-1}$. From this simple analysis we conclude that, given the order of magnitude of the parameters involved, we can simplify the model setting $r + m + n \simeq r$. Following Codeço [2001], we assume that people ingest contaminated water or food once a day ($a = 1 \text{ d}^{-1}$).

[23] To model the seasonality of the bacterium ecology as discussed in section data, we let the net growth rate of *V. cholerae* in the aquatic environment vary periodically in

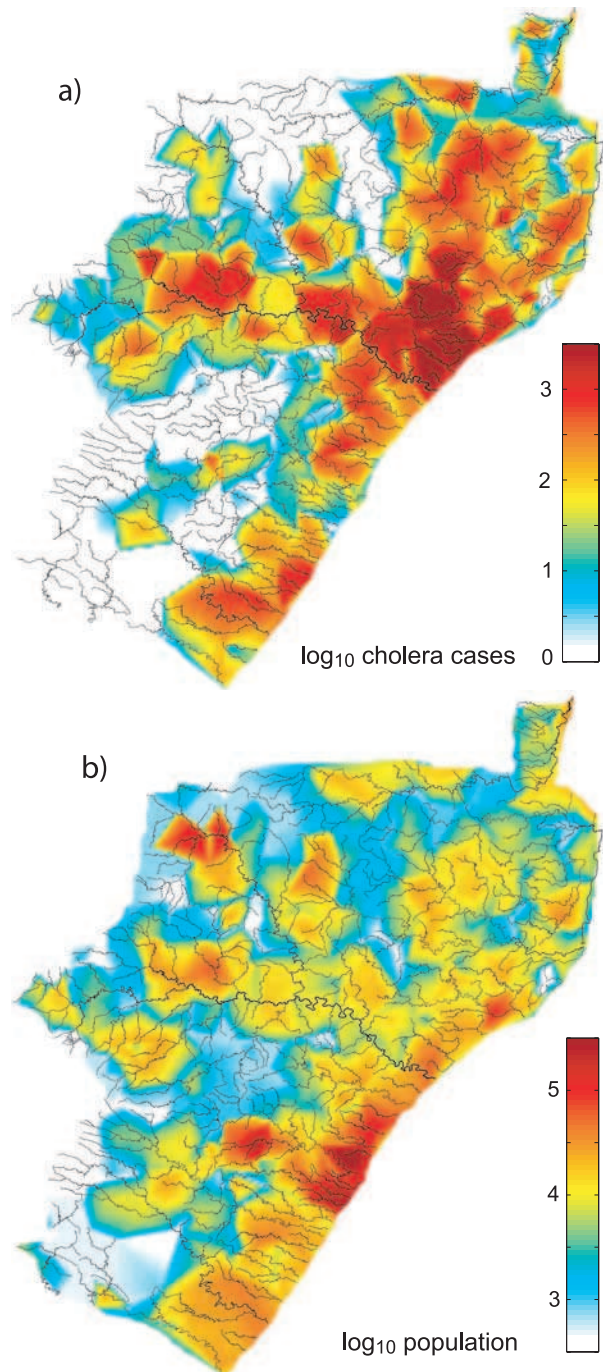


Figure 3. Spatial linear interpolation of network nodes value of (a) cholera cases and (b) population size.

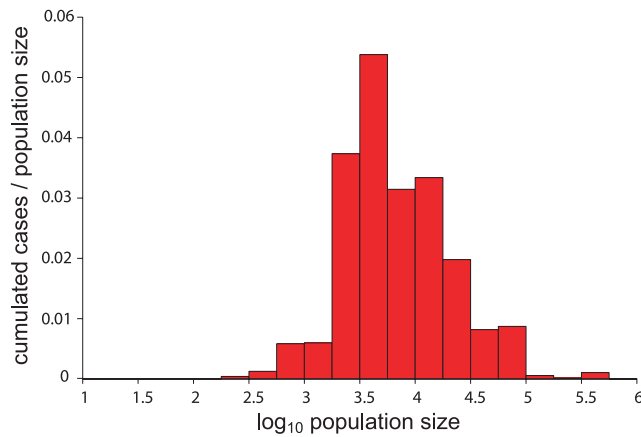


Figure 4. Average cholera incidence (i.e., total cholera cases over population size). Nodes have been grouped in a logarithmic bin on the basis of their population size.

time according to $n_B(t) = n_B(1 + \sin(2\pi t/365))$ (with t in days and $t = 0$ corresponding to 1 October) following the water temperature cycle [Jury, 1998]. We implicitly assume that the free-living bacteria are in demographic equilibrium ($n_B(t) = 0$) during the warmest season. After the above considerations, the model parameters to be calibrated are the mean net *V. cholerae* growth rate in the aquatic environment (n_B), the rate at which *V. cholerae* are removed and transported (l), the ratio ($p/(Kc)$), the threshold (H_T), and the bias of the transport (b).

4. Results and Discussion

[24] In order to calibrate the model, we subdivided the set of parameters into two groups: The first contains the parameters related to the transport and spatial distribution of cases (l , b , and H_T), while the second set groups the parameters closely related to the local model and the temporal dynamics (n_B and $p/(Kc)$). For every randomly chosen combination of the three parameters of the first group, we calibrated the two parameters of the second group by minimizing the mean square error between the data of the temporal evolution of cumulated cases in the whole basin and the simulations. Note that the temporal evolution of the cumulated cases for each node can be obtained via the equation

$$\frac{dC_i}{dt} = a \frac{B_i^*}{1 + B_i^*} S_i, \quad (7)$$

which takes into account only the flux of susceptibles that actually become infected. By adding the cumulated cases at time t throughout all the nodes, we get the cumulative evolution for the entire catchment. Every simulation starts with the initial conditions of one infected person in a single node where, according to data, the first case of cholera was recorded and runs continuously for 2 years (the location of this node is reported on the map in Figure 2). Then, for each combination of the five parameters (the triplet l , b , H_T and the corresponding calibrated pair n_B , $p/(Kc)$) we computed the mean square error between the total cumulated cases of cholera at each node after 2 years obtained from the

simulations and from the data. Note that the error from the first calibration measures the likelihood of the temporal patterns, while the second calibration compares simulated and recorded spatial patterns of the epidemic. Finally, we chose the combination of parameters that minimizes the weighted sum of the two errors. The weight for each error is computed as the inverse of the minimum value of the error itself found among all the realizations. The values of the calibrated parameters are listed in Table 1.

[25] Figure 5 shows the comparison between data (dots) and model simulation (solid line) of the temporal dynamics of the weekly (Figure 5a) and cumulated (Figure 5b) cholera cases in the whole Thukela river basin. Figure 6 compares the spatial distribution of data with that obtained via simulation. It shows the distribution of the cumulated cholera cases after the first and second epidemic outbreak. As in Figure 3, the colors are graded via spatial linear interpolation of nodal values.

[26] The model does well in reproducing the distribution of the cholera cases during the two outbreaks as well as their spatial spreading. It is interesting to note that cases from the second outbreak are mainly located in new regions with respect to the first one. This is related to the spread of cholera from the regions involved in the first epidemic outbreak into disease-free ones. This supports our hypothesis that the dynamics of cholera epidemics in nonendemic regions depend on the spatially anisotropic spreading along an environmental matrix defined by river corridors as well as on inner local dynamics. Similar results in a different context were obtained by Campos *et al.* [2006], Bertuzzo *et al.* [2007], and Muneeppeerakul *et al.* [2007].

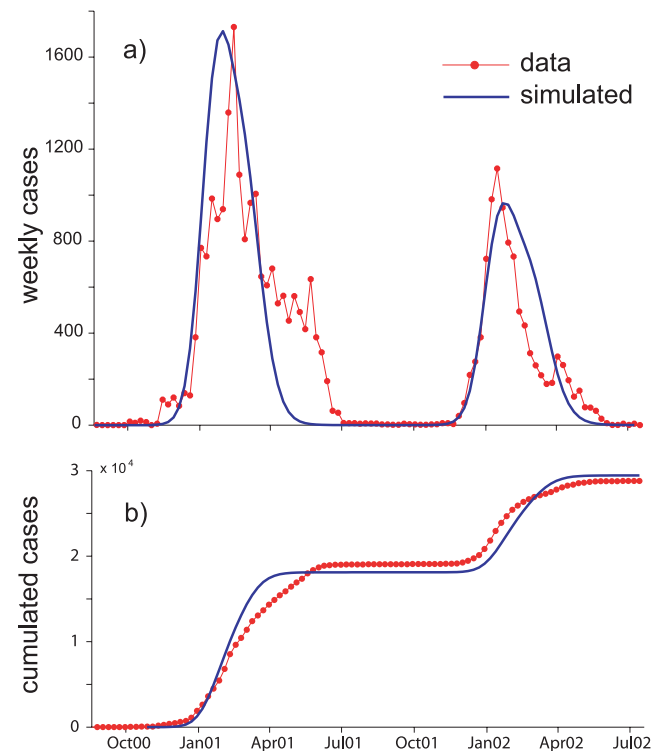


Figure 5. Comparison between the data (dots) and simulated (solid line) temporal evolution of (a) weekly cholera cases and (b) cumulated cases for the Thukela river basin of the KwaZulu-Natal province.

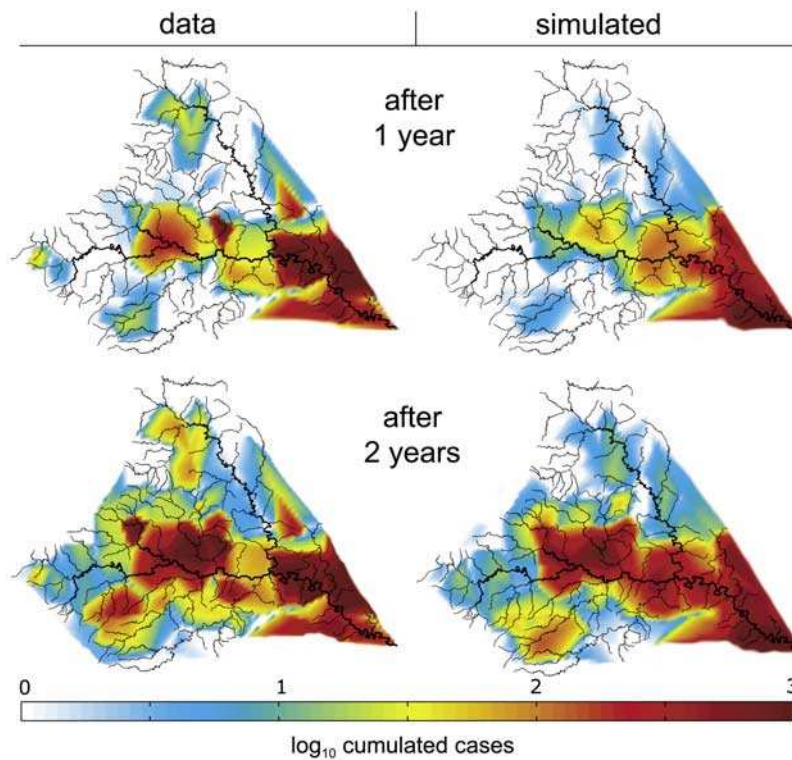


Figure 6. Comparison between data and simulated spatial distribution of the cumulated cholera cases after the first and second epidemic outbreak for the Thukela river basin. Colors are obtained via spatial linear interpolation of the node values.

[27] Finally, we have checked whether our model is able to reproduce the relation discussed in section 3 between cholera incidence and population size of each node. The comparison for the Thukela river basin is reported in Figure 7. It demonstrates that the model results regarding the incidence distribution agree quite well with the data.

[32] In conclusion, we suggest that this approach represents a first step toward understanding how hydrology and population distribution along the water network control the spreading of water-borne diseases.

5. Conclusions

[28] The following main conclusions can be drawn from our results:

[29] 1. Our model of cholera dynamics explicitly accounts for the spatial distribution of the communities and their interconnections, proving capable of reproducing the main spatial and temporal patterns of the spreading for a well-documented epidemic into a disease-free region in the KwaZulu-Natal province of South Africa.

[30] 2. A significant role emerges for the ecological corridors defined by waterways and river networks. Such hydrologic control derives from the transportation and redistribution of the free-living infective propagules. In particular, because *vibrios* can spread along stream both upstream and downstream with a slightly biased propagation downstream (which is to be expected, of course), the infection patterns are anisotropic.

[31] 3. Despite its satisfactory performance, the model is not completely reliable in reproducing secondary peaks of infections in the tail of both annual outbreaks. We speculate that they might be the combined results of seasonality in the epidemiological parameters and the presence of short-lived hyperinfectious bacteria. This hypothesis remains to be tested and will be the object of future work.

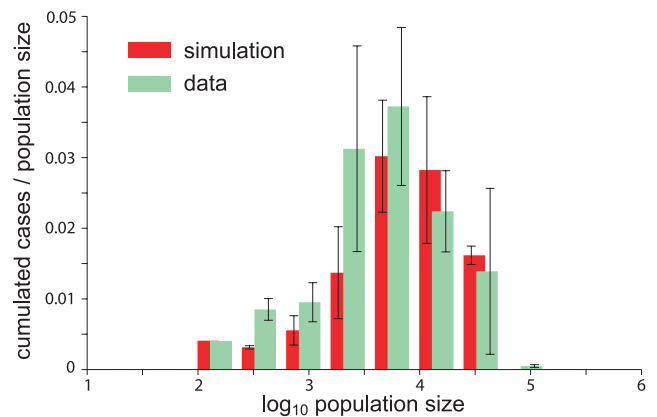


Figure 7. Comparison between data and simulated cholera incidence (i.e., the ratio of total cholera cases to population size) for the Thukela river basin. Nodes have been grouped in a logarithmic bin on the basis of their population size. Bars represent the mean incidence for each bin, while the error bars represent the standard error of the mean. The Kolmogorov-Smirnov goodness of fit test value is 0.0781, and that leads us to accept the hypothesis that the model fits the data with a significance level of 0.05.

[33] **Acknowledgments.** Special thanks go to the KwaZulu-Natal Department of Health for providing the data set that made this work possible. Research funding was provided by AQUATERRA EU project GOCE-505428, PRIN 2006 “Fenomeni di trasporto sul ciclo idrologico,” Politecnico di Milano, and the University of Padua. I.R.-I. gratefully acknowledges the support of the James S. McDonnell Foundation under the grant “Studying Complex System” (220020138).

References

- 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 Resour. Res.*, *43*, W04419, doi:10.1029/2006WR005533.
- Campos, D., J. Fort, and V. Méndez (2006), Transport on fractal river network: Application to migration fronts, *Theor. Popul. Biol.*, *69*, 88–93, doi:10.1016/j.tpb.2005.09.001.
- Capasso, V., and S. L. Paveri-Fontana (1979), A mathematical model for the 1973 cholera epidemic in the European Mediterranean region, *Rev. Epidemiol. Sante Publique*, *27*, 121–132.
- Codeço, C. T. (2001), Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir, *BMC Infect. Dis.*, *1*, 1, doi:10.1186/1471-2334-1-1.
- Colwell, R. R. (1996), Global climate and infectious disease: The cholera paradigm, *Science*, *274*(5295), 2025–2031, doi:10.1126/science.274.5295.2025.
- Hartley, D. M., J. G. Morris, Jr., and D. L. Smith (2006), Hyperinfectivity: A critical element in the ability of *V. cholerae* to cause epidemics?, *PLoS Med.*, *3*(1), 63–69, doi:10.1371/journal.pmed.0030007.
- Johnson, A. R., C. A. Hatfield, and B. T. Milne (1995), Simulated diffusion dynamics in river networks, *Ecol. Modell.*, *83*, 311–325, doi:10.1016/0304-3800(94)00107-9.
- Jury, M. R. (1998), Statistical analysis and prediction of KwaZulu-Natal climate, *Theor. Appl. Climatol.*, *60*, 1–10, doi:10.1007/s007040050029.
- Koelle, K., X. Rod, M. Pascual, M. Yunus, and G. Mostafa (2005), Refractory periods and climate forcing in cholera dynamics, *Nature*, *436*(7051), 696–700, doi:10.1038/nature03820.
- Lipp, E. K., A. Huq, and R. R. Colwell (2002), Effects of global climate on infectious disease: The cholera model, *Rinsho to Biseibutsu*, *15*(4), 757–770, doi:10.1128/CMR.15.4.757-770.2002.
- Mugero, C. and A. Hoque (2001), Review of cholera epidemic in South Africa with focus on KwaZulu-Natal Province, technical report, KwaZulu-Natal Dept. of Health, Pietermaritzburg, South Africa.
- Muneepeerakul, R., J. Weitz, S. Levin, A. Rinaldo, and I. Rodriguez-Iturbe (2007), A neutral metapopulation model of riparian biodiversity, *J. Theor. Biol.*, *245*(2), 351–363, doi:10.1016/j.jtbi.2006.10.005.
- Pascual, M., X. Rod, S. P. Ellner, R. Colwell, and M. J. Bouma (2000), Cholera dynamics and El Niño–Southern Oscillation, *Science*, *289*(5485), 1766–1769, doi:10.1126/science.289.5485.1766.
- Pascual, M., M. J. Bouma, and A. P. Dobson (2002), Cholera and climate: Revisiting the quantitative evidence, *Microbes Infect.*, *4*(2), 237–245, doi:10.1016/S1286-4579(01)01533-7.

S. Azaele and I. Rodriguez-Iturbe, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08540, USA.

E. Bertuzzo and A. Rinaldo, Dipartimento di Ingegneria Idraulica, Marittima Ambientale e Geotecnica, Università di Padova, via Loredan 20, I-35131 Padova, Italy. (bertuzzo@idra.unipd.it)

M. Gatto, Dipartimento di Elettronica e Informazione, Politecnico di Milano, via Ponzio 34, I-20133 Milano, Italy.

A. Maritan, Consorzio Interuniversitario per le Scienze Fisiche della Materia, Istituto Nazionale Fisica della Materia, via Marzolo 8, I-35131 Padova, Italy.