HHS Public Access

Author manuscript

Adv Water Resour. Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:

Adv Water Resour. 2017 October ; 108: 367–376. doi:10.1016/j.advwatres.2016.11.013.

Climate-driven endemic cholera is modulated by human mobility in a megacity

Javier Perez-Saez[a], **Aaron A. King**[b], **Andrea Rinaldo**[a], **Mohammad Yunus**[c], **Abu S. G. Faruque**[c], and **Mercedes Pascual**[d]

[a]Laboratory of Ecohydrology, Ecole Polytechnique Fédérale de Lausanne, CH-1015, Switzerland [b]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA [c]International Centre for Diarrheal Disease Research, Dhaka 1000, Bangladesh [d]Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA

Abstract

Although a differential sensitivity of cholera dynamics to climate variability has been reported in the spatially heterogeneous megacity of Dhaka, Bangladesh, the specific patterns of spread of the resulting risk within the city remain unclear. We build on an established probabilistic spatial model to investigate the importance and role of human mobility in modulating spatial cholera transmission. Mobility fluxes were inferred using a straightforward and generalizable methodology that relies on mapping population density based on a high resolution urban footprint product, and a parameter-free human mobility model. In accordance with previous findings, we highlight the higher sensitivity to the El Niño Southern Oscillation (ENSO) in the highly populated urban center than in the more rural periphery. More significantly, our results show that cholera risk is largely transmitted from the climate-sensitive core to the periphery of the city, with implications for the planning of control efforts. In addition, including human mobility improves the outbreak prediction performance of the model with an 11 month lead. The interplay between climatic and human mobility factors in cholera transmission is discussed from the perspective of the rapid growth of megacities across the developing world.

Keywords

human mobility; endemic cholera; El Niño

1. Introduction

Seasonal outbreaks of endemic cholera are responsible of an estimated 3 million cases of acute watery diarrhea and 90’000 deaths each year [1]. The disease is associated with poor sanitary conditions given that fecal-oral transmission occurs either via the environment-to-

Correspondence to: Mercedes Pascual.

human or human-to-human routes, bringing into play climatic, environmental and socio-economic factors [2, 3]. Megacities of the developing world are exposed to cholera (and other water-borne infections) due to their high population density and all too often inadequate water and sanitation infrastructure prompting exposure to the pathogens through various mechanisms [4, 5].

Dhaka, the capital of Bangladesh with an estimated 17 million people, is located in a floodplain and experiences two cholera incidence peaks per year [2, 6]. Rapid urbanization in recent decades has been accompanied by an increase in the prevalence of severe cholera cases during flooding events [7]. When studied at the regional scale, cholera is known to be influenced by the El Niño Southern Oscillation (ENSO) [8, 9], but the spatio-temporal patterns of incidence within large cities are not necessarily synchronized (contrary to what one would expect from the Moran effect [10]). This lack of synchrony points to the non-trivial relationship between global climatic drivers and local transmission dynamics [11], in particular in urban contexts [12]. The strong degree of spatial heterogeneity in population density and sanitary infrastructure in Dhaka has been hypothesized to underlie a differential sensitivity to climatic forcing between the densely populated city core and the more rural periphery (Fig. 1) [13]. Indeed, informal settlements, which often constitute hotspots of cholera transmission [14], are widespread in the city [15], and vulnerable to the intense yearly monsoon flooding which gains particular intensity in high-ENSO years [16, 17]. Although these conditions would explain the high cholera attack rates in the core, the pathways by which cholera spreads across the megacity remain unclear.

Hydrological transport and human mobility are the two main mechanisms for the spatial spread of *Vibrio cholera*, although food networks can also play an important role in bacterial dissemination [for eg. see 18]. Both of the former have been considered in modeling studies at the national scale for epidemic cholera either together [19, 20], or separately for human mobility [21, 22] and hydrology [23]. Potential pathways of hydrological pathogen transport in Dhaka, through both the urban drainage system and the surface during flooding, have recently been explored using a high-resolution 2D advection-dispersion model [24]. This study suggested small-scale heterogeneities in pathogen exposure due to the mixing of wastewater with flood water, but the highlighted effects were mainly local and were subject to unresolved sources of uncertainty [24]. Although hydrology may indeed be an important explanatory factor of cholera attack rates at the neighborhood-scale, it remains to be shown that it also acts as a city-scale control of disease spread. Another mechanism underling spread would be human mobility which has recently been shown to be a crucial driver at the national level [22], but that to our knowledge has neither been quantified nor included in spatial models for cholera dynamics at the urban scale. This gap has also recently been stressed for the transmission of all climatically-forced infectious diseases [12]. Models of human mobility that depend only on the spatial distribution of population have been used in waterborne disease models in the absence of mobility data such as mobile phone records, but not within cities [25, 26]. The recent development of a high resolution (12m) Global Urban Footprint (GUF) extracted from remote sensing data opens the door to the application of mobility models at the megacity scale [27, 28]. We present here an approach to rely on such mobility models to improve on default near neighbor assumptions of connectivity in transmission models, when direct data on movement is not available.

Addressing the interplay between climate forcing and spreading mechanisms could prove useful in designing vaccination campaigns that take spatial structure into account, especially given the recent evidence showing the feasibility and protective effect of this kind of intervention in Dhaka [29, 30]. Indeed cholera models that account for spatial connectivity have been proposed to guide reactive vaccination interventions during urban cholera epidemics [31].

Here, we build on the existing framework of a multi-dimensional inhomogeneous Markov Chain model (MDIMC) for cholera transmission [13, 32]. We specifically incorporate inferred human mobility fluxes based on population distribution to investigate the spatio-temporal dynamics of cholera within the city. We then discuss the relevance of taking human mobility into account in water-borne disease transmission within megacities.

2. Data and Methods

2.1. Dhaka and cholera

Dhaka covers around 2'300 km^2 of a complex patchwork of land and water located North of the intersection between the Padma and the Meghna rivers. The city experiences regular flooding of particular intensity during high-ENSO years, of which 1998 and 2003 are the most memorable examples [17]. The part of the city we consider in this study is composed of 21 administrative units called *thanas*, which can be categorized into a densely populated city core and a more rural periphery (Fig 1 and [13, 33]).

Cholera, which has been reported for this region since the British colonial presence in the late 1800's [34], continues to represent a public health issue for the city of Dhaka. Outbreaks are still common today, showing different temporal patterns in different parts of the city (Fig. 1). The epidemiological data, its collection protocol and the processing required by the MDIMC framework have been described elsewhere [13]. Briefly, cholera case data consisted of monthly cholera hospitalization counts at the International Center for Diarrhoeal Diseases Research, Bangladesh from 1995 to 2008. As mentioned in [13], the data correspond to El Tor infections confirmed by systematic laboratory analysis of a 4% sample of incoming patients with cholera symptoms from in 1995, and a 2% sample from 1996 onward. The attack rate data used in this study are the same as the ones used in [13], which were corrected for the difference in sampling effort. For modeling purposes and following [13], the monthly attack rate data was categorized into 3 discrete cholera states 0,1 and 2 corresponding to no (0 cases/10'000 people/month), medium (less than 1.85 cases/10'000 people/month) and high (more than 1.11 cases/10'000 people/month) cholera incidence respectively (see Fig. 1). The average incidence in the medium and high cholera states were 1.11 cases/10'000 people/month and 4.58 cases/10'000 people/month respectively.

2.2. The model

A large palette of modeling techniques have been applied to study cholera (for an overview see [32]). Here we choose the MDIMC framework proposed by Reiner et al. [13] in which the observed monthly cholera attack rates are categorized into discrete states of cholera intensity, and the space-time dynamics of transmission modeled by a inhomogeneous

Markov Chain [13, 32]. The transition from cholera in state i at time t to state j at time $t+1$ ($X_{k,t} = i$) $\rightarrow$ ($X_{k,t+1} = j$) occurs with baseline probability $p_{i,j}$. The baseline transition probability matrix is given by:

$$P = \begin{pmatrix} p_{1,1} & (1-p_{1,1}-p_{1,3}) & p_{1,3} \\ p_{2,1} & (1-p_{2,1}-p_{2,3}) & p_{2,3} \\ p_{3,1} & (1-p_{3,1}-p_{3,3}) & p_{3,3} \end{pmatrix}.$$

The effect of covariates can be accounted for through the modification of matrix P. In the case of Dhaka, the consideration of two groups of thanas with differential baseline probabilities was significant to explain the spatio-temporal attack rate patterns [13]. Furthermore, the seasonal variation of cholera risk and the effect the ENSO were shown to play important roles depending on the group to which each thana belongs (core or periphery, which we refer to interchangeably as inner and outer groups). We take both of these factors into account, in addition to the effect of human mobility which we specifically develop. More formally, the transition probability from state i to j with $i < j$ or $i = 3$, of thana k in group g (inner or outer) is assumed to be modulated as:

$$p'_{i,j,k,t} = p_{i,j}\,(1+f^{seas}_{g,t})^{\beta_{i,sgn,g}}\,(1+f^{mob}_{k,t})^{\nu_g}\,(1+f^{ENSO}_{g,t}), \quad \beta,\nu>0 \qquad (1)$$

where $f^{seas}_{g,t}$, $f^{mob}_{k,t}$ and $f^{ENSO}_{g,t}$ correspond to the time varying seasonal, human mobility and ENSO forcings respectively, the β parameters depend on the starting cholera state i and the sign (increase or decrease) of f^{seas}, and all exponents depend on the group g of the thana. The forcing functions are defined as:

$$\begin{aligned} f^{seas}_{g,t} &= \alpha_{m,g}, & \alpha>-1, \\ f^{mob}_{k,t} &= \gamma_{in,g}\sum_{l\neq k}\mathcal{Q}_{lk}X_{l,t} + \gamma_{out,g}\sum_{l\neq k}\mathcal{Q}_{kl}X_{l,t} & \gamma>0, \\ f^{ENSO}_{g,t} &= A_g \tan\left(\tfrac{k_g}{2}\tfrac{ENSO(t-\tau)}{M_g}\right)\tan\left(\tfrac{k_g}{2}\right)^{-1}, & A, M>0; k\in(0,\pi) \end{aligned}$$

where the subscript m indicates the month, $\mathcal{Q}_{kl}$ is the human mobility matrix describing the probability for a person to move from the origin thana, k, to destination, l, and to come back during a finite time interval, the parameters γ_{in}(γ_{out}) modulate multiplicatively the effects of the incoming (outgoing) mobility fluxes, and $ENSO(t-\tau)$ is the sea surface temperature (SST) anomaly corresponding to the El Nino 3.4 index at τ lags (in months) ([13]). The methodology to obtain $\mathcal{Q}_{kl}$ relies on the spatial distribution of population density as explained below. In the original formulation, the effect of spatial connectivity was incorporated through a simple effect of the state of neighboring thanas. Here, we replace this assumption by estimating human mobility fluxes.

The functional form for f^{ENSO} is a flexible parametrization of the influence of the ENSO on cholera risk with $k_g \rightarrow 0$ indicating a linear effect, whereas $k_g \rightarrow \pi$ a nonlinear one (see [13] for an illustration of functional shapes). Here we set $\tau = 10$ months based on the results by [13] which identify this value as the most significant lag between ENSO and cholera

outbreaks, consistent with previous studies [8, 35, 9]. This results in an 11-month lead prediction capacity. For any of the forcing functions, a negative (positive) value indicates a decrease (increase) in the probability of transitioning to a higher cholera state. The modified probabilities are normalized so that they sum to 1:

$$\begin{aligned} p'_{i,1,k,t} &= 1 - p'_{i,2,k,t} - p'_{i,3,k,t} \quad \text{if } i=1; \\ p'_{i,1,k,t} &= (1 - p'_{i,3,k,t}) \frac{p_1}{p_1+p_2} \quad \text{if } i>1 \\ p'_{i,2,k,t} &= (1 - p'_{i,3,k,t}) \frac{p_2}{p_1+p_2}. \end{aligned}$$

2.3. Population density and human mobility

Human mobility models have been extensively employed to infer mobility matrices (here Q) in the framework of spatially-explicit models for the spread of waterborne diseases [19, 25, 20, 36]. Indeed, we are here interested in the spatial repartition of mobility, rather than mobility fluxes themselves (people per unit time). A common choice is to use the gravity model and its different parametrizations, which state that the probability of moving from one location to another is proportional to their respective populations and inversely linked to a function of the distance which separates them [37]. Recently, the radiation model was introduced as a generalization of the gravity model and has been validated based on commuting data [38]. The radiation model draws from physical sciences in assuming that the movement of people follows radiation emission and absorption processes, where both the absorption threshold of trips emitted from a given location and the absorbance of the surroundings are a function of population density. Though tied to an underlying assumption of relatively uniform spatial distribution of mobile population, it presents the advantage of being ”parameter free” when inferring mobility patterns, and to directly depend on the spatial distribution of population density. The radiation model predicts that the fraction of journeys originating from a given origin u that reach destination v is given by:

$$Q_{uv} = \frac{H_u H_v}{(H_u + s_{uv})(H_u + H_v + s_{uv})}, \quad (2)$$

where $H_u(H_v)$ denotes the origin(destination) population and s_{uv} is the sum of the population within a circle centered at the location of the origin, of radius equal to the distance between u and v, d_{uv}, excluding the origin and destination populations. Given the large size and population differences between thanas (Fig. 1) the radiation model was not applied to the administrative units directly, but instead to a pixel-to-pixel basis using a gridded population density raster map.

The population density map was produced by dasymetric mapping (*sensu* [39]) of census population counts available at the thana-level to the distribution of built up areas in the city. The methodology consisted of four steps: 1) computing the thana-specific population density of built-up areas, 2) aggregating population counts at a 1km^2 resolution, 3) computing pixel-to-pixel mobility fluxes, and 4) extracting thana-to-thana fluxes. The two first steps are illustrated in Fig. 2. The starting point of the methodology is a raster map of the built-up area produced and made available by the German Aerospace Center (DLR). The product

results from the analysis of the Synthetic Aperture Radar (SAR) data from the TanDEM-X (TDM) mission merged with multispectral Landsat data (for methodological details and applications see [16, 40, 28, 41]). It consist of a raster with a 12m resolution in which values correspond to the percentage of each pixel's area covered by constructions in 2013 (mapped in terms of "greenness", i.e. the fraction of the pixel covered by vegetation, in Fig. 2 A). Based on the built-up raster, the population density map was computed by multiplying the built-up surface of each pixel by the thana-specific population density in built-up areas, which was obtained by dividing the thana's population by the total amount of built-up surface in the thana (Fig. 2 B). Since the population census data was not available for 2013, an exponential projection of the most recent census was used (see [13]). The population raster was then aggregated to a 1km resolution for the computation of mobility patterns (Fig. 2 C). The pixel-to-pixel mobility matrix, denoted by Q_{pq} hereafter, was computed using the radiation model in eq. 2 applied to the aggregated population density map. The final thana-to-thana mobility matrix $\mathcal{Q}_{kl}$ was computed as:

$$\mathcal{Q}_{kl} = \frac{1}{H_k} \sum_{p \in k} \sum_{q \in l} Q_{pq} \times H_p \quad (3)$$

where $H_k = \Sigma_{p \in k} H_p$ is the sum of the pixel population densities H_p of thana k. The diagonal of $\mathcal{Q}_{kl}$ indicates the probability of intra-thana movement, and the sum over each row equals to 1. The whole procedure can be easily applied to other urban settings since it only depends on (1) total population counts by administrative area (which could also be the overall population in the city) and (2) the built-up area map which can be provided on request for arbitrary areas of interest via DLR.

2.4. Model fitting

One of the advantages of the MDIMC formulation of the model is the simplicity of evaluating the agreement of a set of parameters to the data through a likelihood-based approach. Here we use the subplex algorithm [42], an extension of the straightforward *Nelder-Mead* (NM) algorithm used in previous MDIMC studies [43, 13, 32]. The subplex algorithm consists of iteratively applying the simplex algorithm on subspaces of the investigated parameter space, thus reducing computation time and improving convergence properties [42]. To impose parameter constraints (including positivity and probabilities summing to 1), we used the barrier method as done in previous studies [13, 32]. The optimal parameter set was obtained by running the subplex algorithm 100 times from different starting points using the subplex package in R [44, 45]. The subplex algorithm does not enable the direct computation of uncertainty bounds on the optimal parameter sets due to its heuristic nature. To give a quantification of uncertainty, the likelihood surface was further explored by profiling the parameters of interest. Profiling was performed on the parameters $p_{\cdot,3}$ of the transition matrix for each thana group which capture the intrinsic difference in cholera risk between the different parts of the city. Methodological details and results on parameter profiling are given in the Supplementary Material (SM). The computations were run in parallel with the `foreach` package [46] on the CASTOR cluster of the Scientific IT and Application Support Center of the Ecole Polytechnique Fédérale de Lausanne.

2.5. Cross-validation and predictive capacity

To further assess model predictive capacity, we performed a 14-fold cross-validation following the methodology used by [13]. Briefly, the procedure consisted of sequentially removing one year of data (out of a total of 14 years), refitting the model to the remaining years with the subplex algorithm, and simulating cholera dynamics for the given year using the new best parameter set. Simulating cholera dynamics in the MDIMC framework is straightforward: since the model is stochastic, random numbers between 0 and 1 are drawn to update cholera states based on the probabilities of the state transition matrix (see [13] for details). A total of 10’000 simulations were performed for each data point of the time series for which an 11-month lead was feasible (corresponding to the ENSO forcing lag). To quantify outbreak prediction performance, the observed and simulated outbreaks were compared in a contingency matrix where an outbreak was defined as city-average cholera incidence higher than the 75th percentile of the observed aggregated monthly incidence a the city scale (2.06 cases/10’000 people). For the purpose of outbreak prediction, the simulated outbreak probability (i.e. the percentage of simulations that give an attack rate larger than the above-mentioned threshold) was used as a classification score. Following the methodology in [13], the optimal outbreak probability threshold retained for prediction was found using the Kolmogorov-Smirnov test. The threshold corresponds to the point of maximum distance between the Cumulative Mass Functions of the predicted outbreak probability in months where an outbreak did and did not occur [47]. Given the identified optimal threshold, the confusion matrix of observed and predicted outbreak was used a basis for the computation of predictive performance, including accuracy and false positive and negative rates.

To establish a baseline for comparison for the predictive ability of the MDIMC, we also fitted a seasonal autoregressive integrative moving average model (SARIMA) to the mean city-wide incidence time series. Details on model fitting and prediction results are given in the Supplementary Material.

3. Results

3.1. Predicted thana-to-thana mobility in Dhaka

The thana-to-thana mobility fluxes produced using the radiation model present strong heterogeneties in terms of both directionality and spatial connectivity. Indeed movement is predicted to occur with higher probability between the thanas of the core of the city, and to a lesser extent between the core and the periphery, except for some exceptions (for instance between Narayanganj Sadar (thana 12) and Keraniganj (thana 7), Fig. 3). In addition, outer-outer movement is less frequent probably because of the spatial configurations of the thanas that impose crossing the city core to reach each other. Outer-to-inner movement was predicted to occur with a similar probability to the inner-to-outer one (an average across thanas of around 0.4% of trips).

3.2. Cholera, ENSO and the effect of human mobility

Model performance was quantified in terms of log-likelihood, the Akaike Information Criterion (AIC) [48], and the significance of the covariates evaluated using likelihood ratio

tests applied to the nested models (for details see [13]). The model with all forcing effects (full model with 60 parameters) presented the best AIC score, and the inclusion of all effects was significant compared to all nested models (Table 1). The climatic forcing by ENSO was found to be the most significant effect after seasonality. Interestingly, the best-fitting parameter set of the full model suggests a very similar functional form for the ENSO effect in the two regions of the city ($k_{inner} \approx k_{outer} \approx 2.5$, corresponding to a nonlinear functional form more sensible to extremes, Table 2). In accordance with previous findings [13], cholera dynamics in thanas of the core of the city presented a stronger absolute sensitivity to the ENSO than the periphery (Fig. 4). Indeed, the final selected model suggests a stronger interannual variability of the probabilities to transition to high cholera in the inner thanas than in the outer ones. This is the result of the higher base probability of transitioning to high cholera, $p_{i,3}$, in the inner group and similar I^{ENSO} values (the forcing functions have a multiplicative effect on baseline probabilities). The inner vs. outer difference in cholera risk was further highlighted by the profiling results which indicate lower values of $p_{i,3}$ in the outer group (see SM Fig S1).

Human mobility was identified as a significant effect when compared to all nested models (Table 1). Furthermore, our MIDCM model of cholera dynamics presented an overall better fit than the best-performing model reported in [13], which accounted for connectivity with the simpler near-neighbor assumption, thus highlighting the relevance of a more realistic representations of human mobility. The significant role of this spreading mechanism provides an opportunity to elucidate the spatial directionality of climate-cholera transmission in Dhaka. To this end, we computed the risk contribution of each thana through human mobility to all other thanas. To account for the stochastic variation in cholera states we performed 1-month ahead predictions via simulations with the best parameter set, sequentially setting the state of each thana to 0 and recomputing the value of the mobility forcing $\tilde{f}^{mob}_{t,-k}$. The vector difference between the simulated mobility forcing, $\tilde{f}^{mob}_{t}$, and the simulated forcing without the contribution of the thana represents the increase in cholera risk due to mobility of thana k to each other thana. The simulations were run 1000 times and confidence intervals computed from the resulting difference vectors [$\tilde{f}^{mob}_{t} - \tilde{f}^{mob}_{t,-k}$]. The grouping of risk contribution by origin and destination group (inner vs. outer) reveals that cholera risk spreads mainly from the inner to the outer thanas, and among the latter (Fig. 5). Although risk contribution varies between thanas in the inner-outer and outer-outer directions (see Supplementary Video 1 and SM Fig. S4–S5), their values are consistently larger than the outer-inner risk contribution which is close to 0 throughout the study period. It is interesting to note that the directionality of cholera risk propagation appears clearly in the modeling results, although the predicted inter-group mobility fluxes are very similar. Interestingly the mean risk contribution between thanas (averaged across simulations and the study-period) was found be linearly linked to the intensity of mobility between thanas, although the slope of the relation varied significantly between groups (p-value< 0.01, see SM Table 2 and SM Fig. S6), thus confirming the fundamental differences between the role that human mobility plays depending on the region of the city. This probably stems from the larger probability of experiencing high attack rates in the core than in the periphery.

3.3. Model predictive capacity

Cross-validation confirmed the forecasting capacity of the full model both in terms of inter- and intra-annual variation of cholera incidence and large outbreaks (Fig. 6 A). The cross-validation parameter sets tend to over-estimate cholera attack rates during the period 2002–2003, years in which the lagged ENSO anomaly was high but not accompanied by increased incidence. The optimal threshold for outbreak prediction was found to be an outbreak probability of 18.7% (Fig. 6 B), yielding false negative (21.1%) and positive (24.4%) rates that slightly outperform previously obtained results of ~ 25% by [13]. The overall accuracy of outbreak prediction is of 76.4%.

Furthermore, the MDIMC framework clearly outperformed the best-fitted SARIMA model, thus justifying the additional parametrization required for the MDIMC framework. Indeed, the SARIMA model reached an overall accuracy of 68.9% and much higher false negative rate (37.5%) which is of particular interest for operational reasons (see the SM and SM Fig. S2–S3).

4. Discussion

Dhaka with its heterogeneous landscape in population density and socioeconomic conditions presents a complex yet emblematic setting for waterborne disease transmission in urban centers across megacities of the developing world. Through our MDIMC modeling framework, we have disentangled the roles of climatic forcing and human mobility in shaping disease seasonality, inter-annual variability and spread.

The human mobility fluxes in Dhaka, as predicted by the radiation model, were found to be characterized by a strong intra-group connectivity, especially in the core of the city for which the majority of fluxes were predicted to occur among these densely populated thanas. The inter-group fluxes were predicted to be fairly symmetric. Fluxes were inferred using a simple and generalizable methodology based solely on the spatial distribution of population density. This approach could thus be implemented in other urban settings to study the effect of human mobility on the spread of cholera and other infectious diseases [49, 31, 33], provided urban settings are large enough to allow for the assumptions of the radiation model to apply. We here chose to use the radiation model by [38], although other parameter-free options exist [50], and refined models with calibrated parameters have been suggested to outperform the original formulation [51]. An assumption of our method is that mobility fluxes were computed based on the 2013 distribution of urban population in the city although the cholera data refers to the period 1995–2008. We therefore assume that mobility fluxes between thanas had the same structure throughout the study period, although Dhakas land cover has evolved significantly in the past 30–40 years with built-up areas rapidly extending into the rural peripheries of the city [16]. A full extension of our methodology would therefore require the availability of a time-series for the GUF product, which is not available. We nevertheless posit that the main result of the work, i.e. the directionality of the cholera spread from the inner thanas to the outer ones, would hold. In fact, progressive urbanization occurred mainly in the peripheral thanas [52], and thus the simplification made here would appear reasonable for the purpose of modeling cholera transmission, as it implies that weaker outer-to-outer fluxes with respect to prior urban settings. Furthermore, a

sensitivity analysis (see the SM for details) showed that a variation of ±20% in the outer-to-inner mobility fluxes had only a marginal effect on the prediction outcomes at the city scale (difference in the attack rates < 0.1%, SM Fig S7), thus suggesting that our results are robust to the constant mobility matrix simplification made in this work. Despite these issues, our methodology could be easily modified to incorporate other model formulations, although the lack of access of mobility data such as mobile phone records hinders the applicability of models with free parameters. On the other hand, the use of mobile phone data has been shown to be relevant at the urban scale [53, 54], thus the increasing availability of mobile phone records in cholera-stricken countries could soon overcome the need for human mobility models altogether [21, 22, 55]. In an intermediate phase, and to inform efforts in locations where phone data are not readily available, comparisons of the fluxes based on methods such as the present one against phone data, where available, would be of interest. Indeed the built-up area raster dataset on which we based our current approach can be provided on request for arbitrary areas of interest via DLR. In summary, we have presented the application of an approach to examine whether a model of human movement provides a basis for modeling the connectivity of a spatio-temporal transmission model. In the absence of movement data, one can test whether a model of movement improves the fit of a transmission model to disease data. When this is the case, one has indirect evidence for the model of movement itself, and for the importance of further pursuing data on movement.

Transmission dynamics in the core and the periphery of the city appear to be fundamentally different, in particular with respect to the intrinsically higher probability of the inner thanas to transition to high cholera states. Poor water and sanitation infrastructure conditions, particularly in informal settlements and possibly exacerbated by rapid urbanization in the megacity, have already been pointed out as underlying factors of high cholera risk in the most densely populated areas of Dhaka and other cities [14, 7, 13]. Indeed, the thanas of the inner group presented higher average population densities and poorer sanitation conditions [13]. Our results support previous findings on the key role of global climatic forcing on cholera dynamics in Bangladesh [8, 35, 9], and its differential effect in the city of Dhaka [13]. Indeed, the inner thanas were found to be more sensitive to ENSO and presented a larger inter-annual variation of the probability of transition to high cholera states. These results are conditions on the choices both of the discretization of cholera attack rates into levels, and on the definition of the spatial grouping (here inner vs. outer). A future research direction would be to develop a more systematic way of determining the levels and groupings themselves (for instance through nonparamteric bayesian methods [56]), which would be extremely relevant to the application of the MDIMC framework to other settings and diseases.

The main finding of our work is that the inclusion of human mobility in the model reveals a clear directionality in the spread of cholera risk from the inner to the outer thanas, which had previously been proposed but not quantified. These results echo recent studies on epidemic cholera in urban settings that have revealed transmission hotspots, with implications for vaccination strategies [31]. Interestingly, these findings would suggest that even in the endemic context of Dhaka, secondary (i.e. human-mediated) transmission to (rather than from) the areas with suitable habitats for free-living *V. cholerae* is a main driver of disease outbreaks. Support for this transmission route could re-frame the emphasis placed on the

physical and chemical characteristics of the environment that are suggested to produce outbreaks in endemic regions such as Bangladesh [57, 58, 59]. Although it is undeniable that local environmental indicators are a key piece in understanding cholera outbreaks in urban settings, the results of this study are illustrative of the significance of human-related factors in the dynamics of the disease in a megacity like Dhaka.

Finally, including human mobility not only enabled a better understanding of disease transmission within the city, but also improved to some extent the prediction capacity of the model. Indeed we preserved the 11-month lead prediction horizon, but improved upon predictability, with a 21.1% false negative rate using cross-validation. The MDIMC framework also clearly outperformed the baseline SARIMA model, in particular in terms of false negative rate which is crucial for operational decision-making, thus suggesting that the added complexity in the model is worthwhile. This result is encouraging for the usefulness of MDIMC models for the prediction of cholera, which has already been shown in other countries [32], and for their applicability to other climatic-driven diseases.

5. Conclusion

The transmission of waterborne diseases in urban settings is inextricably linked to the interactions between environmental, socio-economic and climatic factors. Given the rate at which megacities in the developing world are emerging and growing, it is crucial to better understand these linkages to better support disease management in the highly complex and heterogeneous landscapes that characterize them.

Seasonal outbreaks of endemic cholera in Dhaka serve as a striking example of these interactions when accounting for human mobility. The heterogeneous physical landscape of the city, implicitly captured here via population density, could not only be an explanation of the different sensitivity to ENSO between the core and the periphery as suggested by [13], but also yields predictions of mobility fluxes that act to spread the disease. By providing insight into the spatial nature of transmission, the increased understanding these processes may help in the design of large-scale vaccination campaigns supported by promising field studies [29, 30].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to T. Esch and the team ”Urban Areas and Landmanagement” from DLR’s Earth Observation Center (EOC) for providing the built-up area raster of Dhaka used in this study. We thank R. C. Reiner for insight into the modeling framework. The cholera case data used in this paper were collected with the support of ICDDR, B and its donors who provide unrestricted support to ICDDR, B for its operation and research. Current donors providing unrestricted support include the Government of the Peoples Republic of Bangladesh, the Canadian International Development Agency, the Swedish International Development Cooperative Agency, and the Department for International Development, United Kingdom. We gratefully acknowledge these donors for their support and commitment to ICDDR, B’s research efforts. J.P.S and A.R thank the Swiss Agency for Development and Cooperation for its support [grant ”Projet 3E Afrique, Burkina Faso” - 2iE partie scientifique]. A.R. was supported by the Swiss National Science Foundation [grant # CR23I2-138101/1]. M.P. and A.A.K. are grateful for the support of the National Oceanic and Atmospheric Administration’s Program on Oceans and Health [grant #

F020704] and for earlier support by the National Science Foundation (NSF) through the program on the Ecology of Infectious Diseases [grant # 0430120]. A.A.K. was supported by the Research and Policy in Infectious Disease Dynamics program of the Science and Technology Directorate, Department of Homeland Security, the Fogarty International Center, National Institutes of Health and by MIDAS, National Institute of General Medical Sciences [grant # U54-GM111274]. This work was supported by the Ecole Polytechnique Fédérale de Lausanne through the use of the facilities of its Scientific IT and Application Support Center.

References

1. Ali M, Lopez AL, You YA, Kim YE, Sah B, Maskery B, Clemens J. The global burden of cholera. Bulletin of the World Health Organization. 2012; 90(3):209–218. http://www.scielosp.org/scielo.php?script=sci{_}arttext{&}pid=S0042-96862012000300013{&}lng=en{&}nrm=iso{&}tlng=en. DOI: 10.1590/S0042-96862012000300013 [PubMed: 22461716]
2. Miller C, Feachem R, Drasar B. Cholera epidemiology in developed and developping countries: new thoughts on transmission, seasonality, and control. The Lancet. 1985; 325(8423):261–263. http://www.sciencedirect.com/science/article/pii/S0140673685910360. DOI: 10.1016/S0140-6736(85)91036-0
3. Sarker AR, Islam Z, Khan IA, Saha A, Chowdhury F, Khan AI, Qadri F, Khan JAM. Cost of illness for cholera in a high risk urban area in Bangladesh: an analysis from household perspective. BMC Infectious Diseases. 2013; 13(1):518. http://bmcinfectdis.biomedcentral.com/articles/10.1186/1471-2334-13-518. doi: 10.1186/1471-2334-13-518 [PubMed: 24188717]
4. Bugliarello G. Megacities of the developing world, their particularities and their challenges. The Bridge. 29(4)
5. Marshall J. Environmental health: megacity, mega mess. Nature. 2005; 437(7057):312–4. http://dx.doi.org/10.1038/437312a. DOI: 10.1038/437312a [PubMed: 16163320]
6. Akanda AS, Jutla AS, Islam S. Dual peak cholera transmission in Bengal Delta: A hydroclimatological explanation. Geophysical Research Letters. 36(19)
7. Chowdhury F, Rahman MA, Begum YA, Khan AI, Faruque ASG, Saha NC, Baby NI, Malek MA, Kumar AR, Svennerholm A-M, Pietroni M, Cravioto A, Qadri F. Impact of rapid urbanization on the rates of infection by *Vibrio cholerae* O1 and enterotoxigenic *Escherichia coli* in Dhaka, Bangladesh. PLoS Neglected Tropical Diseases. 2011; 5(4):e999. http://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0000999. doi: 10.1371/journal.pntd.0000999 [PubMed: 21483709]
8. Pascual M, Rodo X, Ellner SP, Colwell R, Bouma MJ. Cholera Dynamics and El Niño-Southern Oscillation. Science. 2000; 289(5485):1766–1769. http://www.sciencemag.org/content/289/5485/1766. DOI: 10.1126/science.289.5485.1766 [PubMed: 10976073]
9. Koelle K, Rodó X, Pascual M, Yunus M, Mostafa G. Refractory periods and climate forcing in cholera dynamics. Nature. 2005; 436(7051):696–700. http://dx.doi.org/10.1038/nature03820. DOI: 10.1038/nature03820 [PubMed: 16079845]
10. Hudson PJ, Cattadori IM. The Moran effect: a cause of population synchrony. Trends in Ecology & Evolution. 1999; 14(1):1–2. [PubMed: 10234236]
11. Cash BA, Rodó X, Kinter JL. Links between Tropical Pacific SST and cholera incidence in Bangladesh: Role of the Eastern and Central Tropical Pacific. Journal of Climate. 2008; 21(18):4647–4663. http://journals.ametsoc.org/doi/abs/10.1175/2007JCLI2001.1. DOI: 10.1175/2007JCLI2001.1
12. Santos-Vega, M., Martinez, PP., Pascual, M. Climate forcing and infectious disease transmission in urban landscapes: integrating demographic and socioeconomic heterogeneity; Annals of the New York Academy of Sciences. 2016. p. 1-12.http://doi.wiley.com/10.1111/nyas.13229
13. Reiner RC, King AA, Emch M, Yunus M, Faruque ASG, Pascual M. Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109(6):2033–6. http://www.pnas.org/content/109/6/2033.short. DOI: 10.1073/pnas.1108438109 [PubMed: 22308325]
14. Penrose K, de Castro MC, Werema J, Ryan ET. Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. PLoS Neglected Tropical Diseases. 2010; 4(3):e631. http://journals.plos.org/

plosntds/article?id=10.1371/journal.pntd.0000631. doi: 10.1371/journal.pntd.0000631 [PubMed: 20300569]

15. World Bank. Dhaka: Improving conditions for the urban poor, Report 35824 -BD. Wrold Bank; 2006.
16. Dewan AM, Yamaguchi Y. Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. Applied Geography. 2009; 29(3):390–401. http://www.sciencedirect.com/science/article/pii/S0143622809000058. DOI: 10.1016/j.apgeog.2008.12.005
17. Faisal IM, Kabir MR, Nishat A. The disastrous flood of 1998 and long term mitigation strategies for Dhaka city. Natural Hazards. 2003; 28(1):85–99. http://link.springer.com/article/10.1023/A%3A1021173832234. DOI: 10.1023/A:1021173832234
18. Taylor JL, Tuttle J, Pramukul T, Brien KO, Barrett TJ, Jolbitado B, Lim Y, Vugia D, Morris JG, Tauxe RV, et al. An outbreak of cholera in Maryland associated with imported commercial frozen fresh coconut milk. Journal of Infectious Diseases. 1993; 167(6):1330–1335. [PubMed: 8501322]
19. Bertuzzo E, Casagrandi R, Gatto M, Rodriguez-Iturbe I, Rinaldo A. On spatially explicit models of cholera epidemics. Journal of the Royal Society Interface. 2010; 7(43):321–333. DOI: 10.1098/rsif.2009.0204
20. Mari L, Bertuzzo E, Righetto L, Casagrandi R, Gatto M, Rodriguez-Iturbe I, Rinaldo A. Modelling cholera epidemics: the role of waterways, human mobility and sanitation. Journal of the Royal Society Interface. 2012; 9(67):376–388. DOI: 10.1098/rsif.2011.0304
21. Bengtsson L, Gaudart J, Lu X, Moore S, Wetter E, Sallah K, Rebaudet S, Piarroux R. Using mobile phone data to predict the spatial spread of cholera. Scientific Reports. :5.
22. Finger F, Genolet T, Mari L, de Magny GC, Manga NM, Rinaldo A, Bertuzzo E. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. Proceedings of the National Academy of Sciences of the United States of America. 2016 201522305.
23. Bertuzzo E, Mari L, Righetto L, Gatto M, Casagrandi R, Rodriguez-Iturbe I, Rinaldo A. Hydroclimatology of dual-peak annual cholera incidence: Insights from a spatially explicit model. Geophysical Research Letters. 2012; 39(5) n/a n/a l05403. http://dx.doi.org/10.1029/2011GL050723. doi: 10.1029/2011GL050723
24. Mark, O., Jrgensen, C., Hammond, M., Khan, D., Tjener, R., Erichsen, A., Helwigh, B. A new methodology for modelling of health risk from urban flooding exemplified by cholera case Dhaka, Bangladesh. Journal of Flood Risk Management. 2015. n/a–n/ahttp://dx.doi.org/10.1111/jfr3.12182
25. Righetto L, Bertuzzo E, Casagrandi R, Gatto M, Rodriguez-Iturbe I, Rinaldo A. Modelling human movement in cholera spreading along fluvial systems. Ecohydrology. 2011; 4(1):49–55. DOI: 10.1002/eco.122
26. Rinaldo A, Bertuzzo E, Mari L, Righetto L, Blokesch M, Gatto M, Casagrandi R, Murray M, Vesenbeckh SM, Rodriguez-Iturbe I. Reassessment of the 2010–2011 Haiti cholera outbreak and rainfall-driven multiseason projections. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109(17):6602–6607. DOI: 10.1073/pnas.1203333109 [PubMed: 22505737]
27. Leinenkugel P, Esch T, Kuenzer C. Settlement detection and impervious surface estimation in the Mekong Delta using optical and SAR remote sensing data. Remote Sensing of Environment. 2011; 115(12):3007–3019. http://dx.doi.org/10.1016/j.rse.2011.06.004http://www.sciencedirect.com/science/article/pii/S0034425711002288.
28. Taubenböck H, Esch T, Felbier A, Wiesner M, Roth A, Dech S. Monitoring urbanization in mega cities from space. Remote Sensing of Environment. 2012; 117:162–176. http://www.sciencedirect.com/science/article/pii/S0034425711003427. DOI: 10.1016/j.rse.2011.09.015
29. Khan IA, Saha A, Chowdhury F, Khan AI, Uddin MJ, Begum YA, Riaz BK, Islam S, Ali M, Luby SP, Clemens JD, Cravioto A, Qadri F. Coverage and cost of a large oral cholera vaccination program in a high-risk cholera endemic urban population in Dhaka, Bangladesh. Vaccine. 2013; 31(51):6058–64. http://www.sciencedirect.com/science/article/pii/S0264410X13013868. DOI: 10.1016/j.vaccine.2013.10.021 [PubMed: 24161413]

30. Qadri F, Ali M, Chowdhury F, Khan AI, Saha A, Khan IA, Begum YA, Bhuiyan TR, Chowdhury MI, Uddin MJ, Khan JAM, Chowdhury AI, Rahman A, Siddique SA, Asaduzzaman M, Akter A, Khan A, Ae You Y, Siddik AU, Saha NC, Kabir A, Riaz BK, Biswas SK, Begum F, Unicomb L, Luby SP, Cravioto A, Clemens JD. Feasibility and effectiveness of oral cholera vaccine in an urban endemic setting in Bangladesh: a cluster randomised open-label trial. Lancet (London, England). 2015; 386(10001):1362–71. http://www.sciencedirect.com/science/article/pii/S0140673615611400. DOI: 10.1016/S0140-6736(15)61140-0

31. Azman AS, Luquero FJ, Rodrigues A, Palma PP, Grais RF, Banga CN, Grenfell BT, Lessler J. Urban cholera transmission hotspots and their implications for reactive vaccination: evidence from Bissau city, Guinea Bissau. PLoS Neglected Tropical Diseases. 2012; 6(11):e1901. http://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0001901. doi: 10.1371/journal.pntd.0001901 [PubMed: 23145204]

32. Finger F, Knox A, Bertuzzo E, Mari L, Bompangue D, Gatto M, Rodriguez-Iturbe I, Rinaldo A. Cholera in the Lake Kivu region (DRC): Integrating remote sensing and spatially explicit epidemiological modeling. Water Resources Research. 2014; 50(1):5624–5637. DOI: 10.1002/2014WR015521

33. Martinez PP, King AA, Yunus M, Faruque ASG, Pascual M. Differential and enhanced response to climate forcing in diarrheal disease due to rotavirus across a megacity of the developing world. Proceedings of the National Academy of Sciences of the United States of America. 2016; 113(15): 4092–4097. arXiv:http://www.pnas.org/content/113/15/4092.full.pdf. http://www.pnas.org/content/113/15/4092.abstract. DOI: 10.1073/pnas.1518977113 [PubMed: 27035949]

34. Bouma MJ, Pascual M. Seasonal and interannual cycles of endemic cholera in Bengal 1891–1940 in relation to climate and geography. Hydrobiologia. 2001; 460(1–3):147–156. http://link.springer.com/article/10.1023/A%3A1013165215074. DOI: 10.1023/A:1013165215074

35. Koelle K, Pascual M. Disentangling Extrinsic from Intrinsic Factors in Disease Dynamics: A Nonlinear Time Series Approach with an Application to Cholera. The American Naturalist. 2004; 163(6):901–913. http://www.jstor.org/stable/10.1086/420798http://www.jstor.org/stable/pdfplus/10.1086{%}2F420798.pdf?acceptTC=true. DOI: 10.1086/420798

36. Perez-Saez J, Mari L, Bertuzzo E, Casagrandi R, Sokolow SH, De Leo GA, Mande T, Ceperley N, Froehlich J-M, Sou M, Karambiri H, Yacouba H, Maiga A, Gatto M, Rinaldo A. A theoretical analysis of the geography of schistosomiasis in Burkina Faso highlights the roles of human mobility and water resources development in disease transmission. PLoS neglected tropical diseases. 2015; 9(10):e0004127. http://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0004127. doi: 10.1371/journal.pntd.0004127 [PubMed: 26513655]

37. Erlander, S., Stewart, NF. The Gravity Model in Transportation Analysis: Theory and Extensions. VSP. 1990. http://books.google.ch/books?id=tId3PU1leR8C

38. Simini F, Gonzalez MC, Maritan A, Barabasi A-L. A universal model for mobility and migration patterns. Nature. 2012; 484(7392):96–100. DOI: 10.1038/nature10856 [PubMed: 22367540]

39. Mennis J. Generating surface models of population using dasymetric mapping. The Professional Geographer. 2003; 55(1):31–42. http://www.tandfonline.com/doi/abs/10.1111/0033-0124.10042. DOI: 10.1111/0033-0124.10042

40. Esch T, Taubenböck H, Roth A, Heldens W, Felbier A, Thiel M, Schmidt M, Müller A, Dech S. TanDEM-X mission - new perspectives for the inventory and monitoring of global settlement patterns. Journal of Applied Remote Sensing. 2012; 6(1):061702–1. http://remotesensing.spiedigitallibrary.org/article.aspx?articleid=1378855. DOI: 10.1117/1.JRS.6.061702

41. Esch T, Marconcini M, Felbier A, Roth A, Heldens W, Huber M, Schwinger M, Taubenbock H, Muller A, Dech S. Urban Footprint Processo - Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. IEEE Geoscience and Remote Sensing Letters. 2013; 10(6):1617–1621. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6566061. DOI: 10.1109/LGRS.2013.2272953

42. Rowan, TH. PhD thesis. Department of Computer Sciences, University of Texas; Austin: 1990. Functional stability analysis of numerical algorithms.

43. Nelder J, Mead R. A simplex-method for function minimization. Computer Journal. 1965; 7(4): 308–313.

44. King, AA. subplex: Unconstrained optimization using the subplex algorithm. R package version 1.1–6. 2015. https://CRAN.R-project.org/package=subplex
45. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2015. https://www.R-project.org/
46. Revolution Analytics, S. Weston, foreach: Foreach looping construct for R. r package version 1.4.2 (2014). http://CRAN.R-project.org/package=foreach
47. Rounds E. A combined nonparametric approach to feature selection and binary decision tree design. Pattern Recognition. 1980; 12(5):313–317.
48. Akaike H. A new look at the statistical model identification, Automatic Control. IEEE Transactions on. 1974; 19(6):716–723.
49. Martens P, Hall L. Malaria on the move: human population movement and malaria transmission. Emerging Infectious Diseases. 2000; 6(2):103. [PubMed: 10756143]
50. Yan X-Y, Zhao C, Fan Y, Di Z, Wang W-X. Universal predictability of mobility patterns in cities. Journal of The Royal Society Interface. 11(100) arXiv:http://rsif.royalsocietypublishing.org/content/11/100/20140834.full.pdf. http://rsif.royalsocietypublishing.org/content/11/100/20140834.
51. Kang C, Liu Y, Guo D, Qin K. A generalized radiation model for human mobility: Spatial scale, searching direction and trip constraint. PloS One. 2015; 10(11):e0143500. [PubMed: 26600153]
52. Ahmed B, Ahmed R. Modeling urban land cover growth dynamics using multi-temporal satellite images: A case study of Dhaka, Bangladesh. ISPRS International Journal of Geo-Information. 2012; 1(1):3–31.
53. Sevtsuk A, Ratti C. Does urban mobility have a daily routine? learning from the aggregate data of mobile networks. Journal of Urban Technology. 2010; 17(1):41–60.
54. Calabrese F, Colonna M, Lovisolo P, Parata D, Ratti C. Real-time urban monitoring using cell phones: A case study in Rome, Intelligent Transportation Systems. IEEE Transactions on. 2011; 12(1):141–151.
55. Lu X, Wrathall DJ, Sundsøy PR, Nadiruzzaman M, Wetter E, Iqbal A, Qureshi T, Tatem A, Canright G, Engø-Monsen K, et al. Unveiling hidden migration and mobility patterns in climate stressed regions: A longitudinal study of six million anonymous mobile phone users in Bangladesh. Global Environmental Change. 2016; 38:1–7.
56. Baskerville EB, Bedford T, Reiner RC, Pascual M. Nonparametric bayesian grouping methods for spatial time-series data. arXiv preprint arXiv:1306.5202.
57. de Magny GC, Murtugudde R, Sapiano MRP, Nizam A, Brown CW, Busalacchi AJ, Yunus M, Nair GB, Gil AI, Lanata CF, Calkins J, Manna B, Rajendran K, Bhattacharya MK, Huq A, Sack RB, Colwell RR. Environmental signatures associated with cholera epidemics. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(46):17676–17681. WOS:000261225600022. DOI: 10.1073/pnas.0809654105 [PubMed: 19001267]
58. Emch M, Feldacker C, Yunus M, Streatfield PK, DinhThiem V, Canh DG, Ali M. Local environmental predictors of cholera in Bangladesh and Vietnam. American Journal of Tropical Medicine and Hygiene. 2008; 78(5):823–832. WOS:000255568100026. [PubMed: 18458320]
59. Jutla AS, Akanda AS, Griffiths JK, Colwell R, Islam S. Warming oceans, phytoplankton, and river discharge: Implications for cholera outbreaks. American Journal of Tropical Medicine and Hygiene. 2011; 85(2):303–308. WOS:000293613000021. DOI: 10.4269/ajtmh.2011.11-0181 [PubMed: 21813852]

Highlights

- A model of cholera transmission in Dhaka which includes human mobility is proposed
- Human mobility was inferred based on gridded estimates of on population density
- Climatic forcing drives transmission inter-annual variability at the megacity scale
- Cholera was found to spread from the densely populated city core to the rural periphery

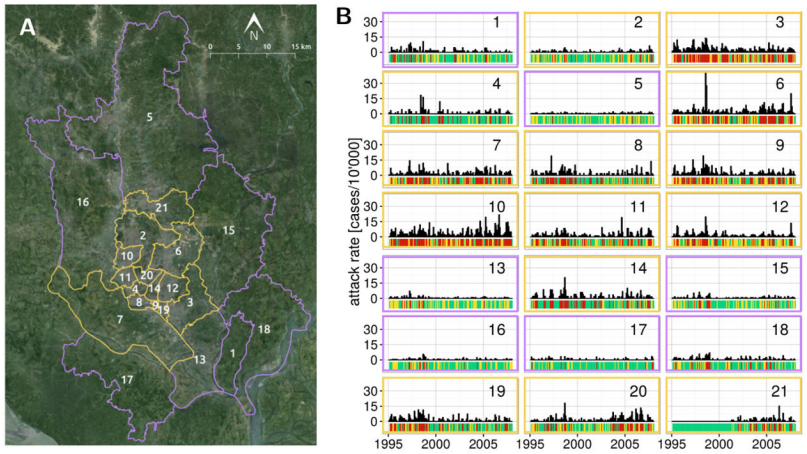

Figure 1.

Dhaka situation map and cholera incidence. (A) The 21 thanas of Dhaka are grouped in an inner core (yellow outline) and the periphery (purple outline). (B) Monthly cholera incidence in the thanas of Dhaka from 1995 to 2008, the bars beneath each timeseries indicates the discrete disease state in which the month is categorized, either no cholera (green), low cholera (yellow) or high cholera (red). The color of the frame indicates if the thana is in the core (yellow) or the periphery (purple) of the city.

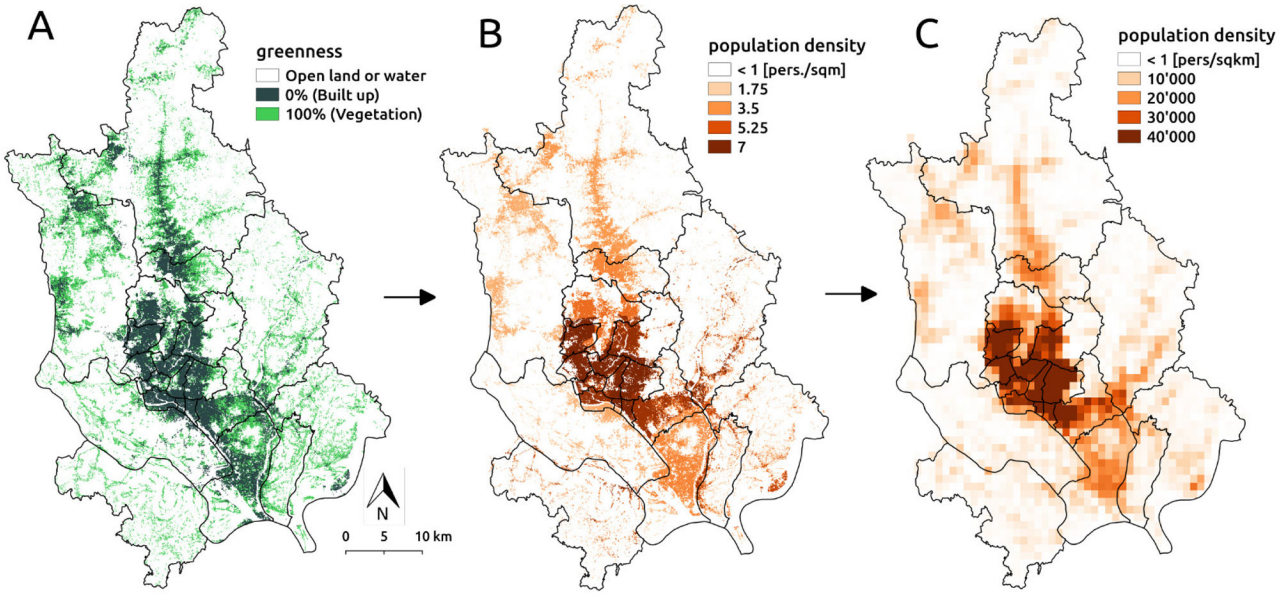

Figure 2.

Population aggregation methodology. (A) Built-up area map shown in terms of the percentage of each pixel covered by vegetation (”greenness” index). (B) Dasymetric mapping of thana-specific population density at 12m resolution. (C) Aggregated population density at 1km resolution used to compute mobility fluxes. Black lines represent thana administrative boundaries.

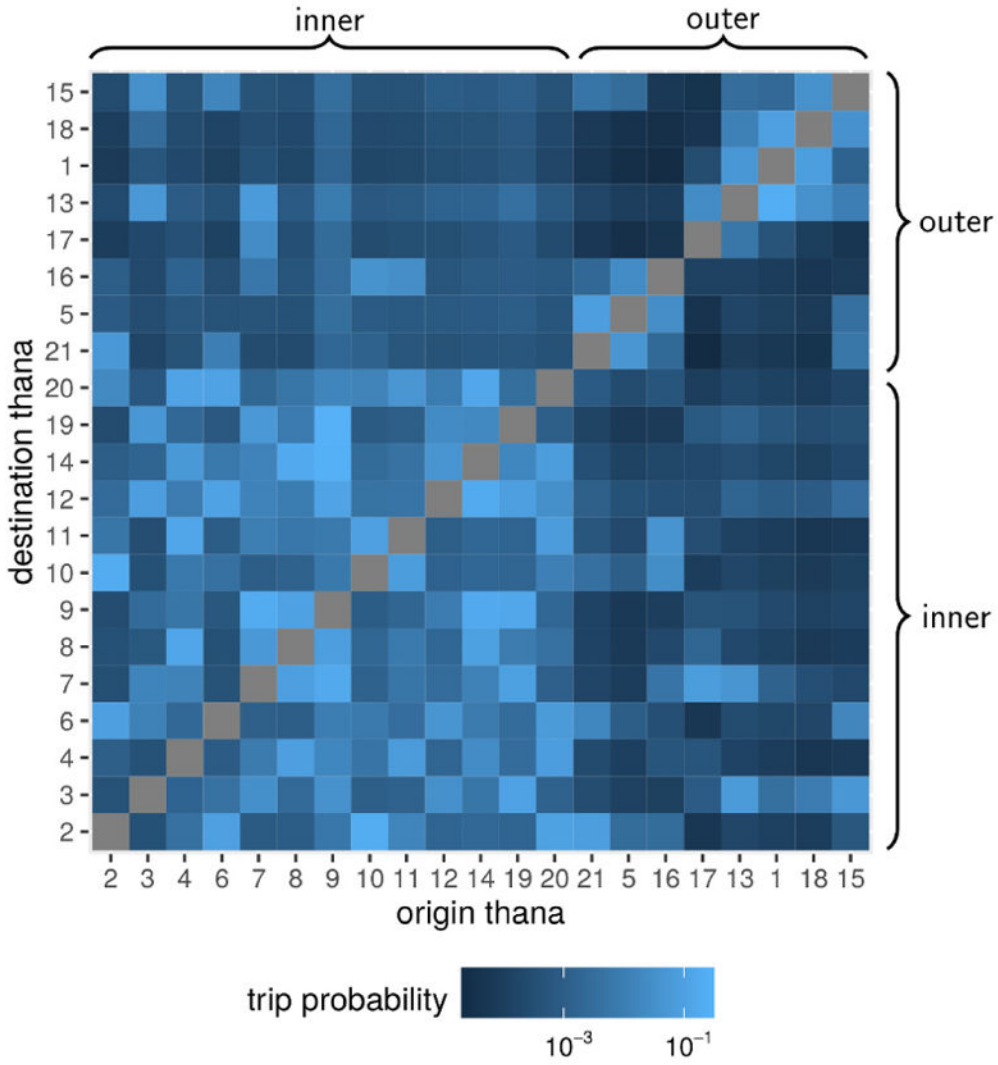

Figure 3.

Mobility fluxes between thanas in Dhaka. The mobility patterns in the city are represented as the percentage of outgoing trips that are predicted to occur between the origin(row) to the destination(column) thana in the log10-color scale of each pixel in the grid. Thana groups (inner core and outer periphery) are indicated by braces.

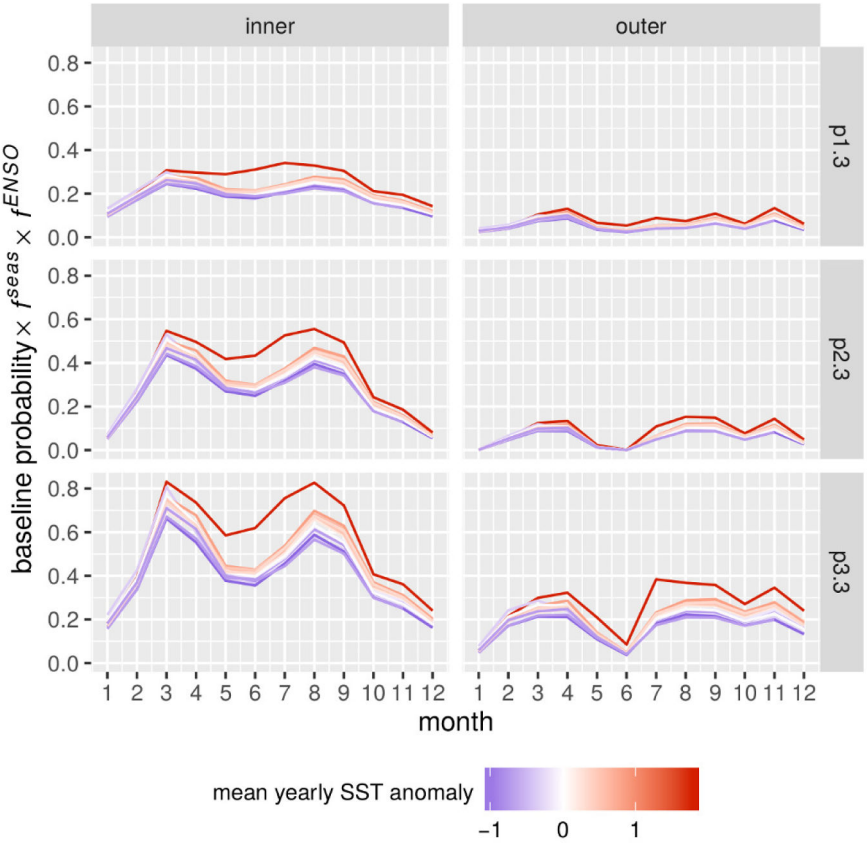

Figure 4.

Climate effect on the interannual variability of cholera risk. The probability of transitioning to a high cholera state from no ($p_{1,3}$), low ($p_{2,3}$) and high cholera ($p_{3,3}$) is shown accounting for the effect of seasonality (f^{seas}) and the ENSO (f^{ENSO}) by month, year, and thana group. The intensity of the ENSO is indicated by the mean value of the SST anomaly in each year from low (blue) to high (red).

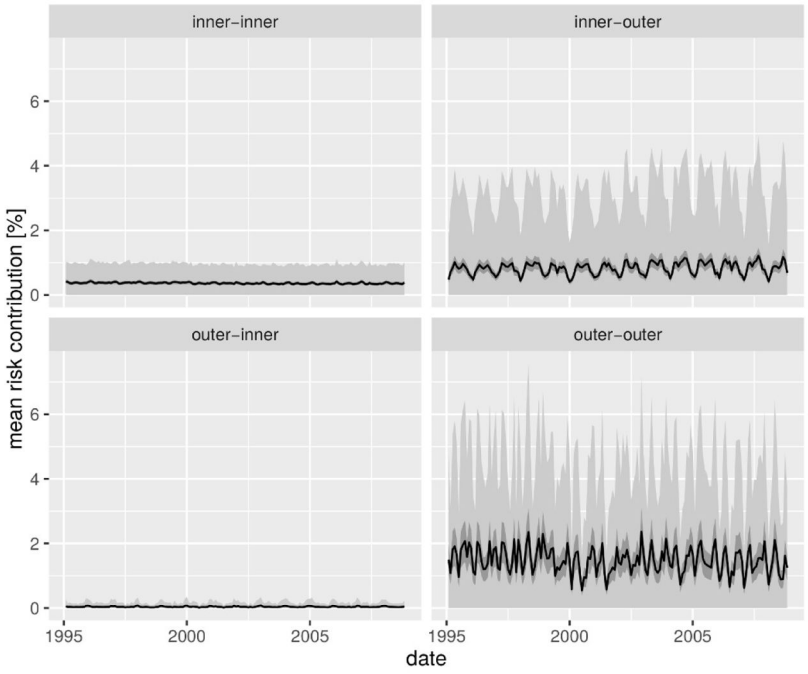

Figure 5.

Cholera risk spread in the city due to human mobility. The thana-to-thana contribution in cholera risk is illustrated in terms of the overall mean percentage increase in the baseline transition probability between the two regions of the city (black line) along with the 95% confidence intervals of the value of the mean (dark gray ribbon) over 1000 simulations of the state transitions. The variability of the difference over realizations is given in terms of the mean ± standard deviation (light gray difference).

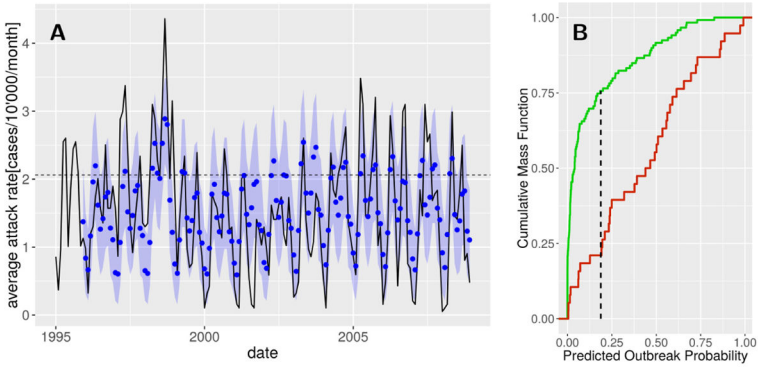

Figure 6.

Cholera outbreak prediction performance. (A) Eleven-month predictions of cholera incidence. The means (blue dots) and 95% confidence intervals (blue shadings) of the simulations are given together with the observed average city-wide cholera state (solid line). The outbreak threshold is indicated by the horizontal dashed line. (B) Kolmogorov-Smirnov test comparing the cumulative mass function of the predicted outbreak probability in months in which an outbreak did (green) and did not (red) occur. The point corresponding to the maximal distance between the two curves (vertical dashed line) is used as a threshold for outbreak prediction.

Table 1

Model fit comparison. The presence(absence) of covariate effects is indicated by the symbol +(−). Nested models are compared against the full model (first line) by sequentially removing 1 covariate (grey rows), and then testing the removal of the other two covariates (white rows). Models are ranked by the AIC score using the number of parameters in each model (df) and the maximal log-likelihood ($\mathcal{L}$). The differences in AIC scores (ΔAIC) were computed against the overall best-performing model. Covariate significance (*p-value*) was computed using the likelihood ratio test on the Δ$\mathcal{L}$s assuming a χ^2 distribution.

f^{seas}	f^{ENSO}	f^{mob}	df	$\mathcal{L}$	AIC	ΔAIC	*p-value*
+	+	+	60	−3110.36	6340.72	0.00	-
+	+	−	54	−3122.64	6353.27	12.56	1.75e-04
+	−	−	48	−3142.83	6381.66		1.74e-07
−	+	−	18	−3297.40	6630.80		2.35e-53
+	−	+	54	−3130.22	6368.45	27.73	2.33e-07
+	−	−	48	−3142.83	6381.66		1.33e-04
−	−	+	18	−3298.65	6633.29		7.11e-51
−	+	+	24	−3230.93	6509.86	169.14	1.47e-32
−	+	−	18	−3297.40	6630.80		1.50e-26
−	−	+	18	−3298.65	6633.29		4.46e-27

Table 2

Optimal parameter set of the full model. Parameters are given by thana group (inner city core and outer periphery) for the baseline probability transition matrix P, the mobility effect f^{mob} and the ENSO effect f^{ENSO}. Parameters γ_{in}(γ_{out}) modulate the effect of incoming(outgoing) human mobility fluxes on the increased risk of cholera, ν regulates the overall effect of f^{mob}, and M, k and A correspond to the scale, phase and amplitude parameters of the ENSO functional form. The parameters of the seasonality forcing f^{seas} are not shown.

	Base probability			Mobility			ENSO		
	$P_{\cdot 0}$	$P_{\cdot 1}$	$P_{\cdot 2}$	γ_{in}	γ_{out}	ν	M	k	A
Inner	$P_{0\cdot}$ 0.582	0.17	0.201	330.7	13.9	0.032	2.568	2.581	0.534
	$P_{1\cdot}$ 0.370	0.341	0.289						
	$P_{2\cdot}$ 0.305	0.292	0.403						
Outer	$P_{0\cdot}$ 0.761	0.188	0.0505	6.46	6.39	0.227	2.432	2.507	0.678
	$P_{1\cdot}$ 0.498	0.398	0.104						
	$P_{2\cdot}$ 0.475	0.275	0.251						