



Cite this article: Kraemer MUG, Perkins TA, Cummings DAT, Zakar R, Hay SI, Smith DL, Reiner Jr RC. 2015 Big city, small world: density, contact rates, and transmission of dengue across Pakistan. *J. R. Soc. Interface* **12**: 20150468.
<http://dx.doi.org/10.1098/rsif.2015.0468>

Received: 25 May 2015
 Accepted: 23 September 2015

Subject Areas:

computational biology

Keywords:

dengue, epidemiology, heterogeneity, mixing, spatial dynamics, spatial accessibility

Author for correspondence:

M. U. G. Kraemer
 e-mail: moritz.kraemer@zoo.ox.ac.uk

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsif.2015.0468> or via <http://rsif.royalsocietypublishing.org>.

Big city, small world: density, contact rates, and transmission of dengue across Pakistan

M. U. G. Kraemer¹, T. A. Perkins^{2,3}, D. A. T. Cummings⁴, R. Zakar⁵, S. I. Hay^{3,6,7}, D. L. Smith^{1,3,8} and R. C. Reiner Jr^{3,9}

- ¹Department of Zoology, University of Oxford, Oxford OX1 3PS, UK
²Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556, USA
³Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA
⁴Department of Epidemiology, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD 21205, USA
⁵Department of Public Health, University of Punjab, Lahore 54590, Pakistan
⁶Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK
⁷Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA 98121, USA
⁸Sanaria Institute for Global Health and Tropical Medicine, Rockville, MD 20850, USA
⁹Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN 47405, USA

Macroscopic descriptions of populations commonly assume that encounters between individuals are well mixed; i.e. each individual has an equal chance of coming into contact with any other individual. Relaxing this assumption can be challenging though, due to the difficulty of acquiring detailed knowledge about the non-random nature of encounters. Here, we fitted a mathematical model of dengue virus transmission to spatial time-series data from Pakistan and compared maximum-likelihood estimates of ‘mixing parameters’ when disaggregating data across an urban–rural gradient. We show that dynamics across this gradient are subject not only to differing transmission intensities but also to differing strengths of nonlinearity due to differences in mixing. Accounting for differences in mobility by incorporating two fine-scale, density-dependent covariate layers eliminates differences in mixing but results in a doubling of the estimated transmission potential of the large urban district of Lahore. We furthermore show that neglecting spatial variation in mixing can lead to substantial underestimates of the level of effort needed to control a pathogen with vaccines or other interventions. We complement this analysis with estimates of the relationships between dengue transmission intensity and other putative environmental drivers thereof.

1. Introduction

The transmission dynamics of all infectious diseases depend on a few basic but key determinants: the availability of susceptible and infectious hosts, contacts between them and the potential for transmission upon contact. Susceptibility is shaped primarily by historical patterns of transmission, the natural history of the pathogen, the host’s immune response and host demography [1]. What constitutes an epidemiologically significant contact depends on the pathogen’s mode of transmission [2], and structure in contact patterns can be influenced by transportation networks and the spatial scale of transmission [3], by host heterogeneities such as age [4], and dynamically in response to the pathogen’s influence on host behaviour [5]. Whether transmission actually occurs during contact between susceptible and infectious hosts often depends heavily on environmental conditions [6–8]. Disentangling the relative roles of these factors in driving patterns of disease incidence and prevalence is a difficult but central

pursuit in infectious disease epidemiology, and mathematical models that capture the biology of how these mechanisms interact represent an indispensable tool in this pursuit [9].

The time-series susceptible–infected–recovered (TSIR) model was developed by Finkenstädt & Grenfell [10] to offer an accurate and straightforward way to statistically connect mechanistic models of infectious disease transmission with time-series data. Among other features, TSIR models readily account for inhomogeneous mixing in a phenomenological way by allowing for rates of contact between susceptible and infectious hosts to depend on their densities nonlinearly. Although this is a simple feature that can be incorporated into any model based on mass-action assumptions—indeed, earlier applications pertained to inhomogeneous mixing in predator–prey systems [11]—the ‘mixing parameters’ that determine the extent of this nonlinearity have primarily been fitted to empirical data in applications of the TSIR model to measles, cholera, rubella and dengue [12–15]. Applied to discrete-time models such as the TSIR, mixing parameters also have an interpretation as corrections for approximating a truly continuous-time process with a discrete-time model [16,17]. In no application of the TSIR model to date has the potential for variation in these parameters been assessed, leaving the extent to which inhomogeneity of mixing varies across space and time as an open question in the study of infectious disease dynamics.

There are a number of reasons why mixing might vary in time or space. Seasonal variation in mixing might arise because of travel associated with labour [18], religious events [19] or vacation [20], or because of the timing of school openings in the case of influenza [21]. Spatial variation in mixing could arise because of cultural differences at geographical scales [3,22,23], because of variation in the density and quality of roads [24], or because of variation in human densities and myriad-associated factors [13,25]. For vector-borne diseases, variation in mixing is amplified even further by variation in vector densities [26], which effectively mediate contact between susceptible and infectious hosts.

Dengue is a mosquito-borne viral disease with a strong potential for spatial variation in mixing [27,28]. The dominant vectors of dengue viruses (*Aedes* spp.) thrive in areas where they are able to associate with humans, as humans provide not only a preferred source of blood but also water containers that the mosquitoes use for egg laying and for larval and pupal development [29]. Two additional aspects of *Aedes* ecology—limited dispersal distance of the mosquito [30] and daytime biting [31]—imply that human movement should be the primary means by which the viruses spread spatially [2]. Indeed, analyses of dengue transmission dynamics at a variety of scales have strongly supported this hypothesis [32–35]. To the extent that human movement in dense urban environments is more well mixed than elsewhere, there are likely to be differences in the extent of inhomogeneous mixing in peri-urban and rural areas. This is also presumably the case for directly transmitted pathogens, but with a potentially even stronger discrepancy for dengue due to the urban–rural gradient in mosquito densities.

To assess the potential for spatial variation in the inhomogeneity of mixing as it pertains dengue transmission, we performed an analysis of district-level time series of dengue transmission in the Punjab province of Pakistan using a TSIR model with separate mixing parameters for urban and rural districts. We likewise made estimates of the relationships between density-independent transmission potential

Table 1. Reported cases by year in Lahore and all other districts.

	2011	2012	2013	2014	total
Lahore	18 020	4013	11 516	1799	35 348
other	3162	500	2356	2790	8808

and putative drivers thereof, such as temperature, to allow for the relative roles of extrinsic and intrinsic factors to be teased apart. Finally, we performed mathematical analyses of the fitted model to assess the significance of spatial variation in mixing inhomogeneity for how time-series data are interpreted and used to guide control efforts.

2. Material and methods

We obtained daily dengue case data aggregated at hospital level from Punjab province provided by the Health Department Punjab, Pakistan, between 2011 and 2014. In total, 47 156 suspected and confirmed dengue cases were reported in 109 hospitals. All hospitals were subsequently geo-located using ‘Google maps’ (<http://www.maps.google.com>) similar to methods described here [36]. Hospitals that could not be identified were removed from the database. The hospitals were then assigned to a district within Punjab, Pakistan by their spatial location. A total of 21 182 cases alone were reported from the year 2011, which affected almost the entire province. Many more cases occurred in Lahore (35 348) compared to all other districts (8808) (table 1). A breakdown per year and each province is provided in electronic supplementary material, table S1, and additional information about collection can be found in the electronic supplementary material, appendix. No information on dengue serotypes were available. However, the predominant serotype circulating in Punjab province, Pakistan is that of DENV-2 [37].

2.1. Covariate selection and processing

Environmental conditions are instrumental in defining the risk of transmission of dengue [28]. Transmission is limited by the availability of a competent disease vector. Due to a lack of resources and political instability, no comprehensive nationwide entomological surveys have been performed in Pakistan. Therefore, we use a probabilistic model to infer the probability of occurrence of *Aedes aegypti* and *Aedes albopictus* in Pakistan derived from a globally comprehensive dataset containing more than 20 000 records for each species (figure 1*a,b*). In short, a boosted regression tree model was applied that predicts a continuous spatial surface of mosquito occurrence from a fitted relationship between the distribution of these mosquitoes and their environmental drivers. A detailed description of the occurrence database and modelling outputs are available here [36,38,39]. Such model outputs have proved useful in identifying areas of risk of transmission of dengue as well as malaria [28,40,41]. Other important environmental conditions defining the risk of transmission of dengue are temperature, water availability and vegetation cover [42]. To account for such variation, raster layers of daytime land surface temperature were processed from the MOD11A2 satellite, gap-filled to remove missing values, and then averaged to a monthly temporal resolution for all 4 years [43]. The density of vegetation coverage has been shown to be associated with vector abundance [44]. Moreover, vegetation indices are useful proxies for precipitation and may be used to derive the presence of standing water buckets that are habitats for the *Aedes* mosquitoes [45]. The same method was again applied to derive the enhanced vegetation index (EVI) from the MOD11A2 satellite to produce 16-day and monthly pixel-based estimates for 2011–2014

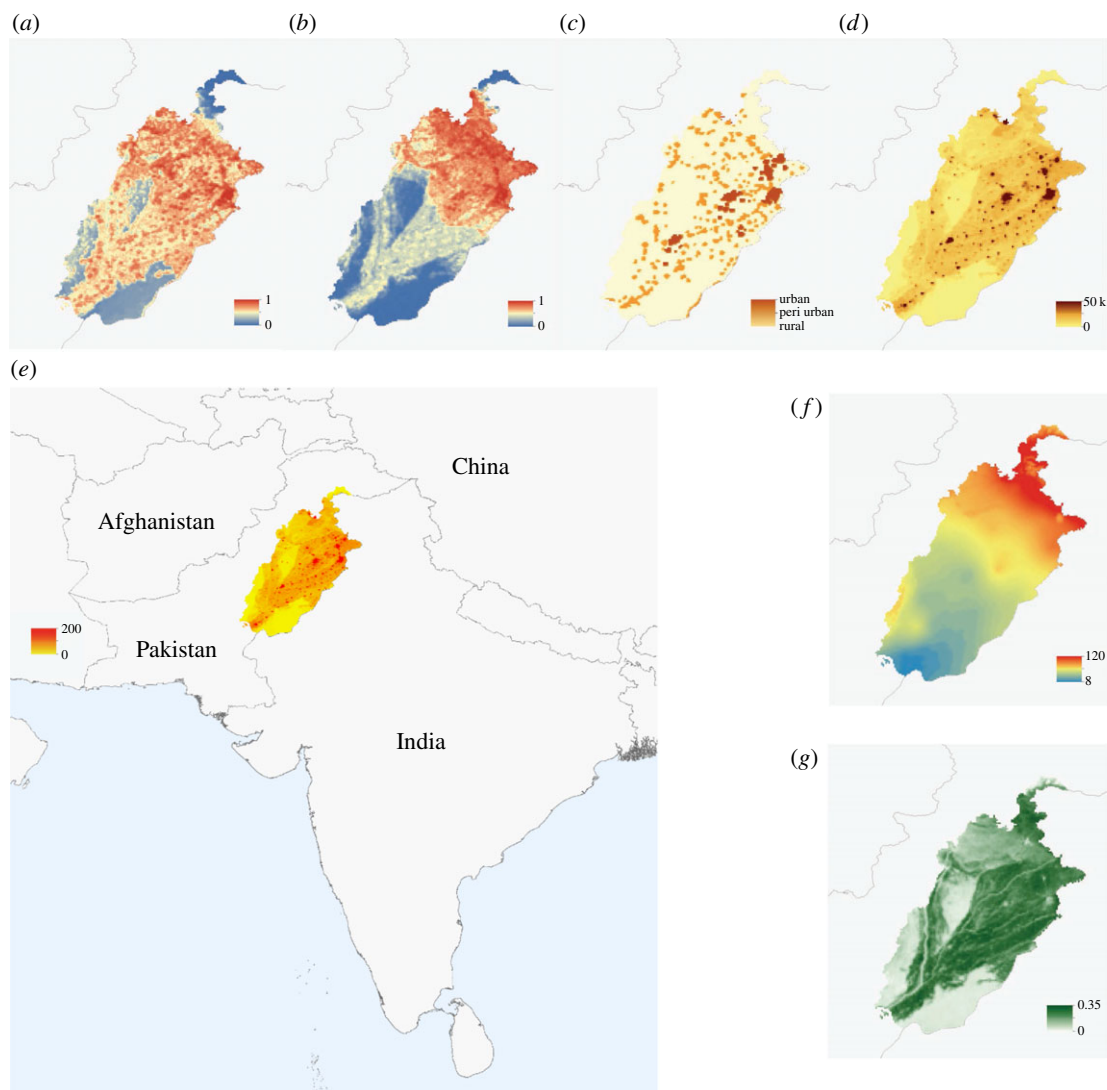


Figure 1. Covariates used in this study to derive environmental drivers of transmission. *Aedes aegypti* probability of occurrence (a); *A. albopictus* probability of occurrence (b); urbanicity (c); weighted urban accessibility (d); population density and study area (e); precipitation (f); EVI mean (g).

(figure 1g) [46]. Due to the inherent delay between rainfall and daily temperature influencing mosquito population dynamics and those mosquitoes contributing to an increase in DENV transmission, we consider both, the influence of the current temperature, vegetation indices and precipitation, data on current transmission as well as the values of those covariates the time step before (figure 1f).

We used population count estimates on a 100 m resolution that were subsequently aggregated up to match all other raster layers to a 5×5 km resolution for the year 2015 (<http://www.worldpop.org>) (figure 1e). In a follow-up analysis to our primary investigation into the climatological drivers of dengue transmission, we included several density-based covariates. We derived a weighted accessibility metric that includes both, population density and urban accessibility, a metric commonly used to derive relative movement patterns [24,47]. This map shows a friction surface, i.e. the time needed to travel through a specific pixel (figure 1d). We also used an urban, peri-urban and rural classification scheme to quantify patterns of urbanicity based on a globally available grid [48] (figure 1c). All covariates and case data were aggregated and averaged (where appropriate) to a district level.

2.2. Model

Following Finkenstädt & Grenfell [10], we assume a general transmission model of

$$I_{t,i} = \beta_{t,i} \cdot \frac{I_{t-1,i}^{\alpha_i}}{N_i} \cdot S_{t-1,i} \cdot \epsilon_{t,i}, \quad (2.1)$$

where $I_{t,i}$ is the number of infected and infectious individuals and $S_{t,i}$ the number of susceptible individuals, at time t in district i , N_i is the population of district i , and $\beta_{t,i}$ is the covariate driven contact rate. We assume each individual to be susceptible as the 2011 epidemic is the first large dengue outbreak. The mixing parameter for the i th district is given by α_i ; when α_i is equal to 1, the population has homogeneous mixing where values less than one can either indicate inhomogeneous mixing or a need to correct for the discretization of the continuous-time process. $\beta_{t,i}$ was fitted using covariates shown in figure 1. Finally, the error terms $\epsilon_{t,i}$ are assumed to be independent and identically log-normally distributed random variables.

2.3. Model selection

The term $\beta_{t,i}$ is fit using generalized additive model regression [49–51]. The time-varying climatological covariates are all fit as a smooth spline, while all other covariates enter $\beta_{t,i}$ linearly. For example, if covariate X_1 and X_2 are time varying and X_3 and X_4 are temporally constant, then we fit β as

$$E[\beta(t)] = s[X_1(t)] + s[X_2(t)] + X_3 + X_4, \quad (2.2)$$

where s are smooths.

Additionally, unexplained seasonal variation is accounted for using a 12-month periodic smooth spline.

Model selection was performed using backwards selection. Two base models were investigated. First, a climate-only model

was created using only the climatological and environmental suitability covariates. Second, a 'full' model using the density-dependent covariates as well as the climatological covariates were combined into a single model which was then subjected to backwards selection. For both models, the mixing coefficient was initially set equal for each district and once a final model was arrived upon, the mixing coefficient for Lahore was allowed to vary separately from the other coefficients. All model fitting was conducted using R [52] and the 'mgcv' package [51]. Models are fit by maximizing the restricted maximum likelihood [53] to reduce bias and over-fitting of the smooth splines. The model source code and processing of covariates will be made available in line with previous projects [54].

2.4. Model analysis

To explore the potential significance of spatial variation in mixing parameters, we conducted an analysis to probe the inherent mathematical trade-off between the mixing parameter α and the density-independent transmission coefficient β . Specifically, to answer the question, what difference in local transmission would be necessary to account for a difference in mixing while achieving identical transmission dynamics. To explore this, we used equation (2.1) to establish:

$$\beta_1 \frac{I_{1,t}^{\alpha_1}}{N_1} = \beta_2 \frac{I_{2,t}^{\alpha_2}}{N_2}, \quad (2.3)$$

from which we obtained

$$\frac{\beta_1}{\beta_2} = I^{\alpha_2 - \alpha_1}. \quad (2.4)$$

We then examined how variation in I and $\alpha_2 - \alpha_1$ affected the left-hand side of equation (2.4) and likewise the critical proportion of the population to control in order to effect pathogen elimination, which, under our model, is $p_c = 1 - (1/\beta)$.

3. Results

3.1. Description of case distribution

The majority of cases were clustered in Lahore, the capital of Punjab province. Ongoing transmission appeared to be focal in three (Vehari, Rawalpindi and Lahore) districts and to have spread through infective 'sparks' to smaller more rural provinces.

3.2. Model selection

To disentangle the different aspects of dengue dynamics and their drivers, we used a model containing only the climatological covariates and performed backwards model selection until each covariate in the model was significant at a 5% level. This resulted in a model that explained 76.9% of the deviance and that had an adjusted R^2 of 0.746. Among the yearly averaged covariates, EVI and precipitation remained in the model, as well as the derived *A. albopictus* range map ($p = 8.21 \times 10^{-4}$, 0.01, and 3.9×10^{-5} , respectively). Interestingly, when we substituted the derived *A. aegypti* map for the *A. albopictus* map, the deviance explained changed very little to 76.8%. For climatological covariates that were fit as smooth splines, temperature, lagged temperature and EVI remained in the model (figure 2, p -value of 0.010, 0.030 and 0.030 with effective degrees of freedom 7.55, 5.47 and 1.83, respectively). There was a significant amount of periodic variation unexplained by the climatological covariates alone, as the 'seasonality' covariate was retained by the model selection

algorithm (figure 2, $p = 0.0034$). The estimated median values for R_0 per district were clustered around 2 (mean = 2.1), and their geographical distribution indicated a clear trend towards districts with larger populations (figure 3). Finally, the mixing coefficient was significantly lower than 1 ($\alpha = 0.69$, 95% CI = (0.614, 0.771), $p = 1.6 \times 10^{-14}$).

To understand these differences, the final model was then compared to a nested model in which the coefficient for Lahore was allowed to vary independently of all other districts. Deviance explained increased to 77% and adjusted R^2 increased to 0.753. Further, the mixing coefficient for Lahore ($\alpha = 0.74$) was significantly larger than the mixing coefficient for the other districts ($\alpha = 0.59$, $p = 0.0068$) (electronic supplementary material, figure S1). The median R_0 for Lahore was estimated at 3.28, the highest among all districts.

To assess the extent to which the variation in mixing coefficients could be explained by other covariates, we considered the possibility that movement accounted for the differences in the mixing coefficients between Lahore and the other districts. The density-dependent covariates (described earlier) were then added to the full model and backwards selection was performed again. The resulting model explained 78.6% of the deviance, had an adjusted R^2 of 0.763 and was superior to the final climatological model based on AIC (699.23 versus 714.83). Yearly averaged EVI, normalized difference vegetation index and precipitation were all significant ($p = 8.7 \times 10^{-5}$, 0.00024 and 0.00028, respectively). Again, the derived *A. albopictus* map was significant ($p = 0.00816$). For climatological covariates fit as smooth splines, only temperature and lagged temperature were found to be significant (figure 2b, $p = 4.0 \times 10^{-5}$ and 0.0013, and effective degrees of freedom 7.61 and 4.81, respectively), and there was still a significant 'seasonality' effect (figure 2b, $p = 4.0 \times 10^{-7}$, effective degrees of freedom 4.48). The best-fit mixing coefficient was $\alpha = 0.58$, barely lower than the mixing coefficient for non-Lahore districts in the climatological model. The estimated median R_0 again clustered around 2 (mean = 1.8), and again the R_0 for Lahore was largest, but in this model it was considerably larger than in the climatological model (Lahore $R_0 = 7.82$, figure 2b). Full details of the best-fit model parameters are shown in electronic supplementary material, table S2–S5.

Two of the density-dependent covariates remained in the model: the urban map ($p = 0.01$) and the weighted access map ($p = 3 \times 10^{-5}$). When the nested model that allowed Lahore's mixing coefficient to vary was fitted, there was no significant difference between the two mixing coefficients ($p \approx 1$).

3.3. Model analysis

Given a difference in estimates of the mixing parameters between Lahore and elsewhere of 0.15, we analysed equation (2.3) to assess the bias in estimates of the transmission coefficient that would result from ignoring this extent of variation in the inhomogeneity of mixing displayed between two areas. For the purpose of *ceteris paribus* comparisons, we assumed equal force of infection but varied it across several orders of magnitude. Depending on the order of magnitude, estimates of transmission coefficients made if overestimating the mixing parameter by 0.15 could easily result in a two- to threefold underestimate in the transmission coefficient (figure 4). For realistic ranges of the transmission coefficient for dengue, and equivalently for the basic reproductive

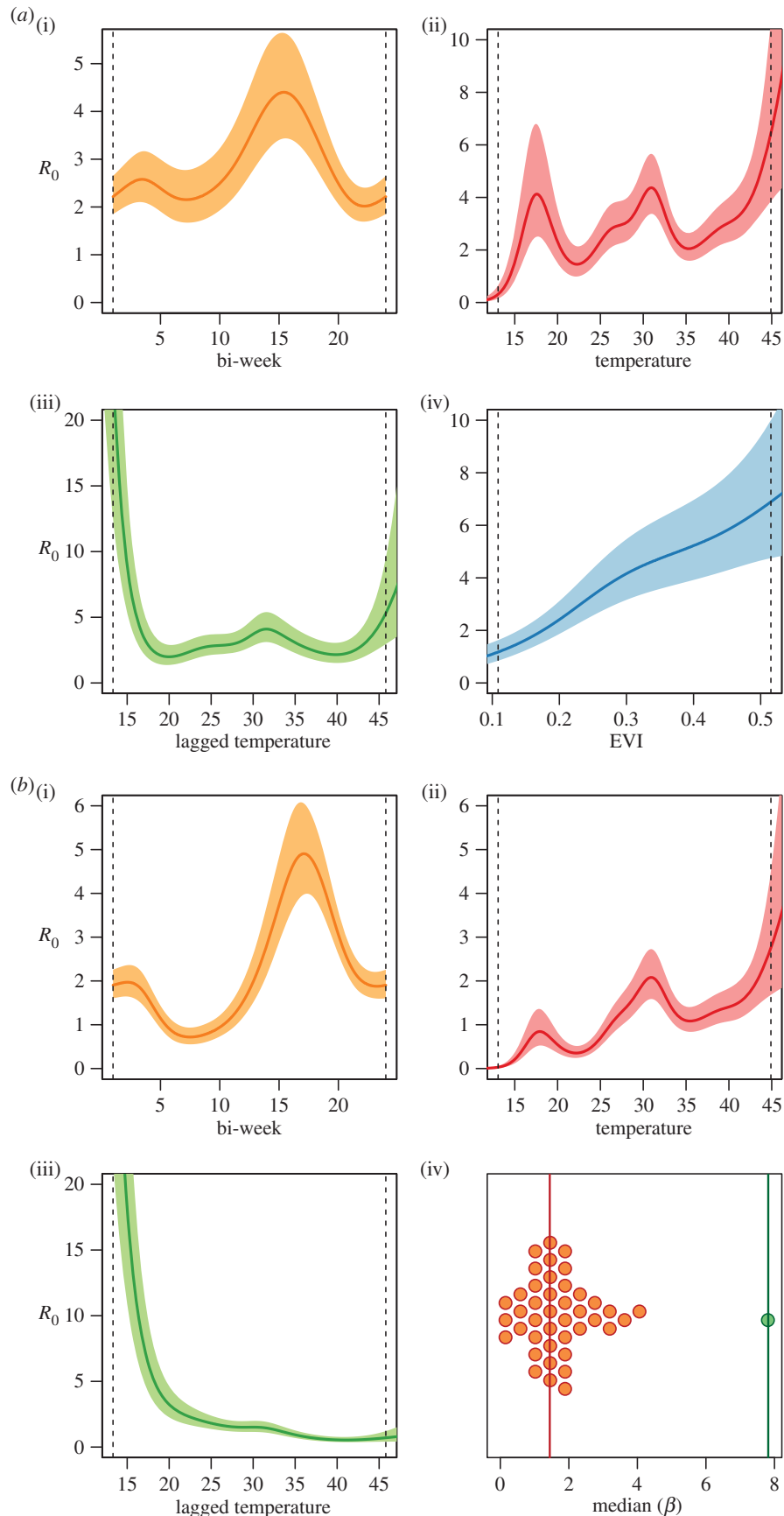


Figure 2. Model outputs using a backwards model selection procedure in the model using climatological variables (a, i–iv), and including the density-dependent variables (b, i–iv). Every subplot shows the predictions of the model for the indicated parameter carrying across the indicated range and every other parameter set to their mean. Figure (b, iv) shows the differences in the transmission coefficient from Lahore (green) and all other districts (red).

number R_0 under the present model, this extent of underestimation of R_0 could lead to underestimating the critical proportion of the population to which vaccines or other interventions must be applied by 20–30% (figure 5).

4. Discussion

Our results point to considerable spatial heterogeneity in the inhomogeneity of mixing and the strength of an associated

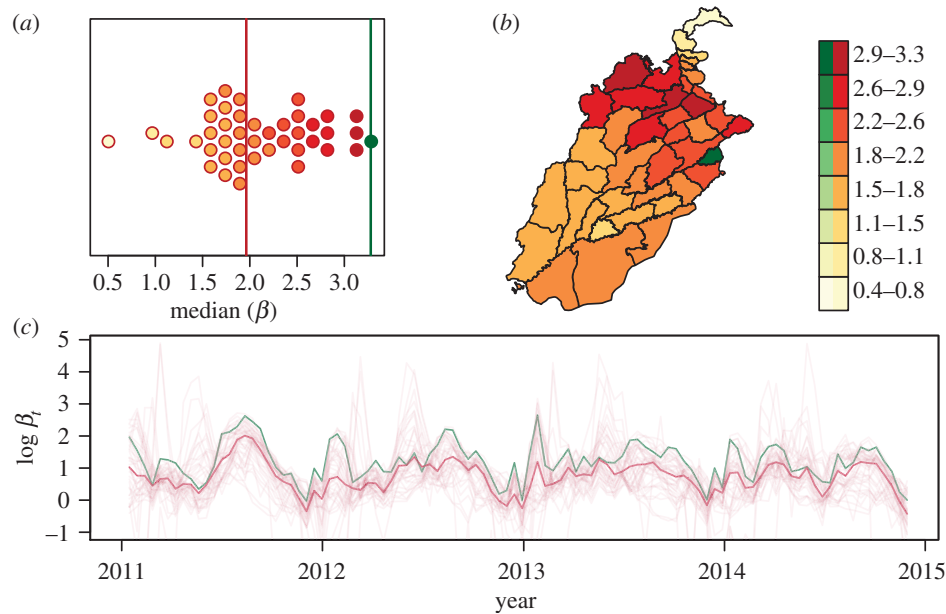


Figure 3. Average distribution of R_0 with green representing Lahore versus red all other districts (a), their geographical distribution (b) and over time in which the green line again is representing Lahore versus the red line representing all other districts (c). Outputs shown here correspond to the model without including density-related variables.

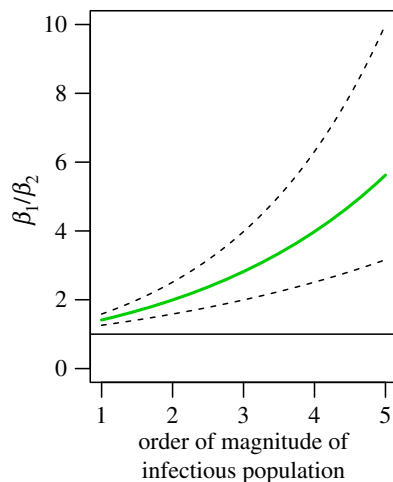


Figure 4. Ratio of betas (R_0) assuming equal force of infection and a difference in $\alpha_2 - \alpha_1$ of 0.2, 0.15 (green), and 0.1, from top to bottom. The straight line indicates a ratio of 1. Order of magnitude of the infectious population refers to number of infectious people in powers of 10; i.e. the range is 10–100 000.

nonlinearity in transmission along an urban–rural gradient. This regional variability in mixing has direct implications for estimates of the basic reproductive number of dengue in our study region and elsewhere. Although the potential for such bias in estimates of the basic reproductive number has been shown in a theoretical context [26,55], we provide quantitative estimates of the extent of this problem by interfacing models with a rich spatio-temporal dataset. Our results have implications for estimates of population-level parameters not only for dengue but also for other infectious diseases [13,56–60] and possibly even more broadly in ecology [11].

Our analysis revealed significant differences in the inhomogeneity of mixing between urban and rural settings and found that a population-weighted urban accessibility metric was able to account for differences in mixing between these settings. Mixing is presumably influenced directly by human behaviour and has been shown to be highly

unpredictable, largely dependent on the local context and the spatial and temporal scale [61]. In this study, however, we could show that the density-dependent covariate we considered was able to capture the influence of these behavioural effects on a district level. Once differences in the inhomogeneity of mixing were accounted for, estimated R_0 values indicated considerably larger differences between transmission potential in Lahore and all other districts. Synchronizing more accurate geo-referenced data would allow for the assessment of the extent to which the relationship between ‘mixing parameters’ and urban accessibility is dependent on the spatial scale at which data are aggregated [26,62]. In the case of dengue, this has been limited specifically by the availability of high-resolution data [63]. Complementing such an analysis with measurement of social contact patterns could be important for exploring this relationship in even more detail [22,64,65] and could be informed by mathematical models that explored this relationship previously for other diseases [13,66]. Another encouraging result from our analysis was the finding that large-scale mosquito suitability surfaces helped capture the environmental determinants of dengue transmission [28].

Intervention strategies are contingent on both understanding key environmental drivers of transmission and the dynamics of ongoing human-to-human transmission, particularly in outbreak situations [67]. Environmental drivers such as seasonal fluctuations in rainfall, temperature, vegetation coverage or mosquito abundance will help guide surveillance and control efforts targeted mostly towards the mosquito vector and its ecology [68]. Once infection occurs, an important and unresolved question for dengue is how to best optimize the delivery of intervention strategies to reduce disease incidence, which is largely determined by R_0 . Our analysis shows that the interaction between mixing parameters and force of infection has potentially large implications for optimizing targeted intervention, particularly in countries where transmission is high and resources are scarce [69]. In fact, this may be even more important in areas of low transmission where incidence appears to be

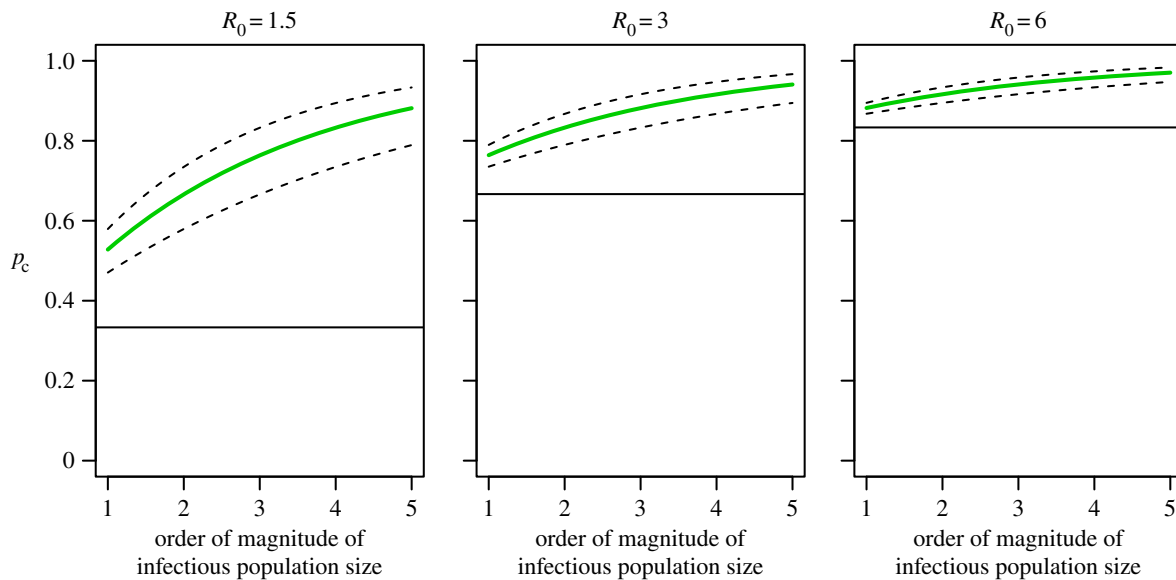


Figure 5. Critical proportion of the population to control in population 2 as a function of R_0 in population 1, the order of magnitude of the infectious numbers in each population, and a difference in $\alpha_2 - \alpha_1$ of 0.1, 0.15 (green) and 0.2. The straight line indicates the critical proportion assuming the α in each population are equal. Order of magnitude of the infectious population refers to number of infectious people in powers of 10; i.e. the range is 10–100 000.

more focal [70]. Again, however, more attention is needed to determine the spatial and temporal resolution of appropriate intervention strategies and the effects of key covariates and model parameters [62]. Empirical understanding of the spatial scale that is most appropriate for carrying out large-scale interventions remains unknown.

Once transmission has occurred in one place, understanding not only spatial heterogeneity in transmission dynamics but also their subsequent spread in mechanistic stochastic models would help to empirically determine the propagation of the disease [71]. Interest in spatial spread dynamics has risen with increasing importation of dengue into heretofore non-endemic areas due to travel and trade continentally and internationally [72]. Exploration of the case data in Pakistan that we analysed here suggests that the virus spreads along major transport routes from Lahore to Karachi and north to Rawalpindi. Using results presented here on mixing coefficients and environmental drivers will help pinpoint areas of major risk of importation more accurately, especially in the case of recurring epidemics. We explored the consequences of a spatially differentiated mixing coefficient in the context of transmission potential within this analysis. Using the fitted relationships of the environmental

drivers of transmission and R_0 will enable future analyses and comparisons between diseases and geographical regions. In this context, it will be instrumental to integrate a variety of movement and social network models with the evidence presented here to infer more accurately how the geographical spread of dengue is determined.

Authors' contributions. M.U.G.K., R.C.R., T.A.P. and D.L.S. conceived and designed the experiment. R.C.R., T.A.P. and M.U.G.K. performed the experiments. R.C.R. and M.U.G.K. analysed the data. D.L.S., S.I.H., D.A.T.C. and Z.R. edited the manuscript. M.U.G.K., R.C.R. and T.A.P. wrote the paper.

Competing interests. We declare we have no competing interests.

Funding. M.U.G.K. acknowledges funding from the German Academic Exchange Service (DAAD). T.A.P., S.I.H. and R.C.R. acknowledge the support from the RAPIDD programme of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health. S.I.H. is funded by a Senior Research Fellowship from the Wellcome Trust (no. 095066) and a grant from the Bill & Melinda Gates Foundation (no. OPP1093011). Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements. The authors thank the WHO Punjab office, Health Department Punjab and PID for providing the epidemiological data, and all participants that helped collect the data.

References

- Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, Holmes EC. 2004 Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332. (doi:10.1126/science.1090727)
- Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz Soldan V, Kochel TJ, Kitron U, Elder JP, Scott TW. 2009 The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl. Trop. Dis.* **3**, e481. (doi:10.1371/journal.pntd.0000481)
- Brockmann D, Helbing D. 2013 The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342. (doi:10.1126/science.1245200)
- Kilpatrick AM, Daszak P, Jones MJ, Marra PP, Kramer LD. 2006 Host heterogeneity dominates West Nile virus transmission. *Proc. R. Soc. B* **273**, 2327–2333. (doi:10.1098/rspb.2006.3575)
- Fenichel EP *et al.* 2011 Adaptive human behavior in epidemiological models. *Proc. Natl Acad. Sci. USA* **108**, 6306–6311. (doi:10.1073/pnas.1011250108)
- Weiss DJ, Bhatt S, Mappin B, Van Boeckel TP, Smith DL, Hay SI, Gething PW. 2014 Air temperature suitability for *Plasmodium falciparum* malaria transmission in Africa 2000–2012: a high-resolution spatiotemporal prediction. *Malar. J.* **13**, 171. (doi:10.1186/1475-2875-13-171)
- Gilbert M *et al.* 2014 Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. *Nat. Commun.* **5**, 4116. (doi:10.1038/ncomms5116)
- Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. 2010 Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol.* **8**, e1000316. (doi:10.1371/journal.pbio.1000316)

9. Anderson RM, May RM. 1991 *Infectious diseases of humans: dynamics and control*. Oxford, UK: Oxford University Press.
10. Finkenstädt BF, Grenfell BT. 2000 Time series modelling of childhood diseases: a dynamical systems approach. *Appl. Stat.* **49**, 187–205. (doi:10.1111/1467-9876.00187)
11. Pascual M, Levin SA. 1999 From individuals to population densities: searching for the scale of intermediate determinism. *Ecology* **80**, 2225–2236. (doi:10.1890/0012-9658(1999)080[2225:FITPDS]2.0.CO;2)
12. Reich NG *et al.* 2013 Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. *J. R. Soc. Interface* **10**, 20130414. (doi:10.1098/rsif.2013.0414)
13. Bjornstad ON, Finkenstädt BF, Grenfell BT. 2002 Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecol. Monogr.* **72**, 169–184. (doi:10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2)
14. Koelle K, Pascual M. 2004 Disentangling extrinsic from intrinsic factors in disease dynamics: a nonlinear time series approach with an application to cholera. *Am. Nat.* **163**, 901–913. (doi:10.1086/420798)
15. Metcalf CJE, Bjornstad ON, Ferrari MJ, Klepac P, Bharti N, Lopez-Gatell H, Grenfell BT. 2011 The epidemiology of rubella in Mexico: seasonality, stochasticity and regional variation. *Epidemiol. Infect.* **139**, 1029–1038. (doi:10.1017/S0950268810002165.The)
16. Liu WM, Hethcote HW, Levin SA. 1987 Dynamical behavior of epidemiological models with nonlinear incidence rates. *J. Math. Biol.* **25**, 359–380. (doi:10.1007/BF00277162)
17. Glass K, Xia Y, Grenfell BT. 2003 Interpreting time-series analyses for continuous-time biological models: measles as a case study. *J. Theor. Biol.* **223**, 19–25. (doi:10.1016/S0022-5193(03)00031-6)
18. Bharti N, Tatem AJ, Ferrari MJ, Grais RF, Djibo A, Grenfell BT. 2011 Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science* **334**, 1424–1427. (doi:10.1126/science.1210554)
19. Lessler J *et al.* 2014 Estimating potential incidence of MERS-CoV associated with Hajj pilgrims to Saudi Arabia, 24 November 2014, 1st edn (referring to: <http://currents.plos.org/outbreaks/article/estimating-potential-incidence-of-mers-cov-associated-with-hajj-pilgrims-to-saudi-arabia-2014/>). (doi:10.1371/currents.outbreaks.c5c9c9a9bd636164a9b6fd4dbda974369)
20. Deville P, Linard C, Martin S, Gilbert M, Stevens FR, Gaughan AE, Blondel VD, Tatem AJ. 2014 Dynamic population mapping using mobile phone data. *Proc. Natl Acad. Sci. USA* **111**, 15 888–15 893. (doi:10.1073/pnas.1408439111)
21. Salathé M, Kazandjieva M, Woo J, Levis P, Feldman MW, Jones JH. 2010 A high-resolution human contact network for infectious disease transmission. *Proc. Natl Acad. Sci. USA* **107**, 22 020–22 025. (doi:10.1073/pnas.1009094108)
22. Vazquez-Prokopec GM *et al.* 2013 Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment. *PLoS ONE* **8**, e58802. (doi:10.1371/journal.pone.0058802)
23. Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, Guan Y, Jiang CQ, Cummings DAT. 2014 Social mixing patterns in rural and urban areas of southern China. *Proc. R. Soc. B* **281**, 20140268. (doi:10.1098/rspb.2014.0268)
24. Background Paper for World Development Report 2009: Reshaping Economic Geography. Washington, DC: World Bank.
25. Simini F, González MC, Maritan A, Barabási A-L. 2012 A universal model for mobility and migration patterns. *Nature* **484**, 96–100. (doi:10.1038/nature10856)
26. Perkins TA, Scott TW, Le Menach A, Smith DL. 2013 Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. *PLoS Comput. Biol.* **9**, e1003327. (doi:10.1371/journal.pcbi.1003327)
27. Brady OJ *et al.* 2012 Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl. Trop. Dis.* **6**, e1760. (doi:10.1371/journal.pntd.0001760)
28. Bhatt S *et al.* 2013 The global distribution and burden of dengue. *Nature* **496**, 504–507. (doi:10.1038/nature12060)
29. Morrison AC, Zielinski-Gutierrez E, Scott TW, Rosenberg R. 2008 Defining challenges and proposing solutions for control of the virus vector *Aedes aegypti*. *PLoS Med.* **5**, e68. (doi:10.1371/journal.pmed.0050068)
30. Harrington LC *et al.* 2005 Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. *Am. J. Trop. Med. Hyg.* **72**, 209–220.
31. Akram W, Hafeez F, Ullah UN, Kim YK, Hussain A, Lee JJ. 2009 Seasonal distribution and species composition of daytime biting mosquitoes. *Entomol. Res.* **39**, 107–113. (doi:10.1111/j.1748-5967.2009.00204.x)
32. Allicock OM *et al.* 2012 Phylogeography and population dynamics of dengue viruses in the Americas. *Mol. Biol. Evol.* **29**, 1533–1543. (doi:10.1093/molbev/msr320)
33. Reiner RC, Stoddard ST, Scott TW. 2014 Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. *Epidemics* **6**, 30–36. (doi:10.1016/j.epidem.2013.12.003)
34. Stoddard ST *et al.* 2013 House-to-house human movement drives dengue virus transmission. *Proc. Natl Acad. Sci. USA* **110**, 994–999. (doi:10.1073/pnas.1213349110)
35. Bhooniboonchoo P *et al.* 2014 The spatial dynamics of dengue virus in Kamphaeng Phet, Thailand. *PLoS Negl. Trop. Dis.* **8**, e3138. (doi:10.1371/journal.pntd.0003138)
36. Kraemer MUG *et al.* 2015 The global compendium of *Aedes aegypti* and *Ae. albopictus* occurrence. *Sci. Data* **2**, 150035. (doi:10.1038/sdata.2015.35)
37. Khan MA *et al.* 2013 Emergence and diversification of dengue 2 cosmopolitan genotype in Pakistan, 2011. *PLoS ONE* **8**, e56391. (doi:10.1371/journal.pone.0056391)
38. Kraemer MUG *et al.* 2015 The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *Elife* **4**, e08347. (doi:10.7554/eLife.08347)
39. Kraemer MUG *et al.* 2015 Data from the global compendium of *Aedes aegypti* and *Ae. albopictus* occurrence. *Dryad Digit. Repository*. (doi:10.5061/dryad.47v3c)
40. Sinka ME *et al.* 2012 A global map of dominant malaria vectors. *Parasit. Vectors* **5**, 69. (doi:10.1186/1756-3305-5-69)
41. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IRF, Johnston GL, Tatem AJ, Hay SI. 2011 A new world malaria map: *Plasmodium falciparum* endemicity in 2010. *Malar. J.* **10**, 378. (doi:10.1186/1475-2875-10-378)
42. Messina JP, Brady OJ, Pigott DM, Golding N, Kraemer MUG, Scott TW, Wint GRW, Smith DL, Hay SI. 2015 The many projected futures of dengue. *Nat. Rev. Microbiol.* **13**, 230–239. (doi:10.1038/nrmicro3430)
43. Weiss DJ, Mappin B, Dalrymple U, Bhatt S, Cameron E, Hay SI, Gething PW. 2015 Re-examining environmental correlates of *Plasmodium falciparum* malaria endemicity: a data-intensive variable selection approach. *Malar. J.* **14**, 1–18. (doi:10.1186/s12936-015-0574-x)
44. Eisen L, Lozano-Fuentes S. 2009 Use of mapping and spatial and space-time modeling approaches in operational control of *Aedes aegypti* and dengue. *PLoS Negl. Trop. Dis.* **3**, e411. (doi:10.1371/journal.pntd.0000411)
45. Maciel-de-Freitas R, Marques WA, Peres RC, Cunha SP, De Oliveira RL. 2007 Variation in *Aedes aegypti* (Diptera: Culicidae) container productivity in a slum and a suburban district of Rio de Janeiro during dry and wet seasons. *Mem. Inst. Oswaldo Cruz* **102**, 489–496. (doi:10.1590/S0074-02762007005 000056)
46. Weiss DJ, Atkinson PM, Bhatt S, Mappin B, Hay SI, Gething PW. 2014 An effective approach for gap-filling continental scale remotely sensed time-series. *ISPRS J. Photogramm. Remote Sens.* **98**, 106–118. (doi:10.1016/j.isprsjprs.2014.10.001)
47. Tatem AJ, Hemelaar J, Gray RR, Salemi M. 2012 Spatial accessibility and the spread of HIV-1 subtypes and recombinants. *AIDS* **26**, 2351–2360. (doi:10.1097/QAD.0b013e328359a904)
48. CIESIN/FPRI/WB/CIAT. Global Rural Urban Mapping Project (GRUMP): Gridded Population of the World, version 3. See <http://sedac.ciesin.columbia.edu/gpww>.
49. Hastie TJ, Tibshirani RJ. 1990 *Generalized additive models*, vol. 43. Chapman and Hall/CRC, CRC Press.
50. Dominici F, McDermott A, Zeger SL, Samet JM. 2002 On the use of generalized additive models in time-series studies of air pollution and health. *Am. J. Epidemiol.* **156**, 193–203. (doi:10.1093/aje/kwf062)
51. Wood SN. 2011 Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat.*

- Soc. Ser. B Stat. Methodol.* **73**, 3–36. (doi:10.1111/j.1467-9868.2010.00749.x)
52. R Core Team. 2014 *R: a language and environment for computing*. Vienna, Austria: the R Foundation for Statistical Computing.
53. Patterson HD, Thompson R. 1971 Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554. (doi:10.1093/biomet/58.3.545)
54. Pigott DM, Kraemer MU. 2014 Enhancing infectious disease mapping with open access resources. *Eurosurveillance* **19**, pii: 20989.
55. Hu H, Nigmatulina K, Eckhoff P. 2013 The scaling of contact rates with population density for the infectious disease models. *Math. Biosci.* **244**, 125–134. (doi:10.1016/j.mbs.2013.04.013)
56. Bartlett MS. 1960 The critical community size for measles in the U.S. *J. R. Stat. Soc. A* **123**, 37–44. (doi:10.2307/2343186)
57. Keeling MJ, Eames KTD. 2005 Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307. (doi:10.1098/rsif.2005.0051)
58. Kiss IZ, Green DM, Kao RR. 2008 The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *J. R. Soc. Interface* **5**, 791–799. (doi:10.1098/rsif.2007.1272)
59. Smith DL, Lucey B, Waller LA, Childs JE, Real LA. 2002 Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proc. Natl Acad. Sci. USA* **99**, 3668–3672. (doi:10.1073/pnas.042400799)
60. Keeling MJ, Grenfell BT. 1997 Disease extinction and community size: modeling the persistence of measles. *Science* **275**, 65–67. (doi:10.1126/science.275.5296.65)
61. Yang Y, Herrera C, Eagle N, González MC. 2014 Limits of predictability in commuting flows in the absence of data for calibration. *Sci. Rep.* **4**, 5662. (doi:10.1038/srep05662)
62. Mills HL, Riley S. 2014 The spatial resolution of epidemic peaks. *PLoS Comput. Biol.* **10**, e1003561. (doi:10.1371/journal.pcbi.1003561)
63. Ruberto I, Marques E, Burke DS, Van Panhuis WG. 2015 The availability and consistency of dengue surveillance data provided online by the World Health Organization. *PLoS Negl. Trop. Dis.* **9**, e0003511. (doi:10.1371/journal.pntd.0003511)
64. Heesterbeek H *et al.* 2015 Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, paaa4339. (doi:10.1126/science.aaa4339)
65. Bauch CT, Galvani AP. 2013 Epidemiology. Social factors in epidemiology. *Science* **342**, 47–49. (doi:10.1126/science.1244492)
66. Reiner RC, King AA, Emch M, Yunus M, Faruque ASG, Pascual M. 2012 Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. *Proc. Natl Acad. Sci. USA* **109**, 2033–2036. (doi:10.1073/pnas.1108438109)
67. Perkins TA, Metcalf CJE, Grenfell BT, Tatem AJ. 2015 Estimating drivers of autochthonous transmission of chikungunya virus in its invasion of the Americas. *PLoS Currents Outbreaks*, 10 February 2015, 1st edn. (doi:10.1371/currents.outbreaks.a4c7b6ac10e0420b1788c9767946d1fc)
68. Johansson MA. 2015 Chikungunya on the move. *Trends Parasitol.* **31**, 43–45. (doi:10.1016/j.pt.2014.12.008)
69. Cesare MDI, Bhatti Z, Soofi SB, Fortunato L, Ezzati M, Bhutta ZA. 2015 Geographical and socioeconomic inequalities in women and children's nutritional status in Pakistan in 2011: an analysis of data from a nationally representative survey. *Lancet Glob. Health* **3**, e229–e239. (doi:10.1016/S2214-109X(15)70001-X)
70. Salje H *et al.* 2012 Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. *Proc. Natl Acad. Sci. USA* **109**, 9535–9538. (doi:10.1073/pnas.1120621109)
71. Grenfell BT, Björnstad ON, Kappey J. 2001 Travelling waves and spatial hierarchies in measles epidemics. *Nature* **414**, 716–723. (doi:10.1038/414716a)
72. Schaffner F, Mathis A. 2014 Dengue and dengue vectors in the WHO European region: past, present, and scenarios for the future. *Lancet. Infect. Dis.* **14**, 1271–1280. (doi:10.1016/S1473-3099(14)70834-5)