THE UNIVERSITY OF CHICAGO

DISENTANGLING THE ROLE OF SOCIAL, ECONOMIC AND ENVIRONMENTAL HETEROGENEITIES IN URBAN MALARIA TRANSMISSION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECOLOGY AND EVOLUTION

BY
OSCAR MAURICIO SANTOS VEGA

CHICAGO, ILLINOIS
AUGUST 2018

A mi madre y hermanos Ruby, Diego, Juan y Yanet

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am tremendously grateful to Mercedes, my advisor and mentor, for the continuous support, motivation, and source of immense knowledge during my Ph.D studies and related research. Her continuous guidance helped me achieve my research goals and ultimately, the writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study; Mercedes has been a source of inspiration for me, not only for her academic capabilities but also for the involvement and time she dedicated to every member of the lab.

I would also like to thank the members of my committee, Stefano Allesina, Sarah Cobey, Luc Anselin, and Greg Dwyer, for their feedback and insightful comments on this dissertation. I would also like to thank Menno Bouma, who taught me how to connect ideas across entomology, disease and climate, and provided me with endless and insightful conversations during our trips to India. Thanks to my graduate student collaborators, especially Andres Baeza and Amir Siraj whom were always open to discuss new ideas and to develop exciting side projects with me.

I would like to more generally thank the Pascual lab. Thanks to Ed, Andres, Shai, Xiangjun, Qixin, Victoria, Ruby, Rahul and many others who have been members of the lab for the last years, for their support and thoughtful discussions. Especially thanks to Pamela, who has been always there as a collaborator, lab mate and friend to me. I also would like to thank the people in the administrative office, Audrey, Bonnie, Connie, and Mary, for always being willing to help. I am very grateful to have met wonderful people in both Ann Arbor and Chicago. Thanks to Dea, Clarisse, Leslie, Roberto, Valen, Kery, Carlos, Dani, Shane, Darli, Dan. Bea, Jime, Clarisse, Roberto, Valen and Lucia, for being friends that turned into family. Thank you Alejandro, my confident, constant supporter, and the epitome of what a friend is.

Finally, I woud like to thank my family, and in particular my mom and siblings to whom this dissertation is dedicated. They have always been supportive, in bad and good times. I feel proud to have them as my family.

ABSTRACT

The growth of urban populations and a changing climate are two defining environmental challenges of the 21st century, especially for the developing world and for human health. An increase in urbanization poses new challenges for the spread and control of communicable diseases. Several major infectious diseases, including malaria, exhibit significant variation in the size of seasonal outbreaks across years in urban environments. To date, questions regarding inter-annual variation have focused on the temporal variability of environmental factors (such as climate ones) and on the nonlinearity of disease dynamics (laneri et al. 2010, Roy et al. 2013). However, the intersection of environmental forcing and socioeconomic heterogeneity in the context of relevant spatial scales for transmission remains unexplored in the field of infectious disease dynamics, especially for urban landscapes (Reiner et al. 2014).

Heterogeneity in socio-economic and associated demographic conditions can interact with environmental variation in space and time. These factors create differential conditions across cities, which can modulate the effect of climate factors at local spatial scales in large urban environments of the developing world (Santos-Vega et al. 2016). It is my general working hypothesis that the understanding, control, and prediction of the population dynamics of climate-sensitive infectious diseases within cities require consideration of the pronounced spatial heterogeneity of urban environments. In concert, the environmental and socio-economic dimensions could define the relevant spatial scales at which to address malaria transmission dynamics. Thus, understanding and controlling transmission dynamics within cities will necessitate consideration of their pronounced spatial heterogeneity and of the relevant spatial scales at which to address temporal variation in incidence.

Despite an increasing interest in the role of spatial heterogeneity, the population dynamics of vector-transmitted diseases has typically been addressed with temporal surveillance records aggregated at the level of whole towns, cities, or regions, including the development

of recent statistical inference methods for confronting process-based models to time series data. Such coarse resolutions have been the norm in part because climate variability is thought to operate at relatively large, regional scales, synchronizing dynamics in space (the Moran effect). These (mean-field) models assume for the most part that host populations are well mixed, so that each individual is equally susceptible. In particular, they do not take into account how spatial variation in environmental, climatic and socio-economic conditions affect vector habitat, biting rates, vector control, contact rates and host susceptibility.

Here, we used urban malaria to understand the respective roles of host density, economic level, and environmental variation in urban malaria prevalence, and to determine the critical spatial scales of these drivers. In chapter 2, I studied the determinants of within-city spatiotemporal heterogeneity in the incidence patterns of malaria by combining statistical analyses and a phenomenological transmission model applied to an extensive spatio-temporal dataset on cases of malaria in the city of Ahmedabad, India. These results showed that climate forcing and socio-economic heterogeneity act synergistically at local scales on the population dynamics of urban malaria in this city. With these models, in chapter 3 I described and explained patterns of spatial variation at the resolution of units (wards), as well as both higher or lower resolutions, and identified the spatial scales of aggregation at which the spatiotemporal variation in reported cases is best captured. Finally, the fourth chapter addresses the role of humidity in influencing the seasonality and inter-annual variability of urban malaria. Our analysis shows a strong and significant effect of humidity in the inter-annual variability of the disease. Simulations of the transmission model for twenty years capture remarkably well the observed epidemic patterns when humidity is incorporated, and also show that forecasting malaria risk based on a dynamical transmission model driven by this climate factor is possible.

CHAPTER 1

CLIMATE FORCING AND INFECTIOUS DISEASE TRANSMISSION IN URBAN LANDSCAPES: INTEGRATING DEMOGRAPHIC AND SOCIOECONOMIC HETEROGENEITY

Urbanization and environmental changes are the dominant forces altering the social and ecological landscape of the planet today (Harpham 2009, Dye 2008). These two processes encompass variation in multiple concurrent factors that affect ecological systems at different scales (Grim et al. 2008). The remarkable expansion of cities in the last 60 years has positioned urban environments as the new dominant ecosystem around the globe (UN 2014). Importantly, the fast growth of cities has co-occurred in time with climate change, including the 1-3 C rise in global temperatures and shifts in rainfall patterns in several regions (Coumou et al. 2012). In cities, climate variation can directly affect the emergence and transmission of environmentally driven infectious diseases, such as cholera, malaria, and dengue (Reiner et al. 2012, Keiser et al. 2004, Lowe et al. 2013). Climate variation, together with demographic and economic factors, can also modulate the transmission of infectious diseases that have not been traditionally considered climate sensitive, such as rotavirus, the major cause of diarrheal morbidity and mortality in children worldwide (Martinez et al. 2016). Consideration of the effects of spatial heterogeneity in different socio-economic components is required to better understand and control the population dynamics of infectious diseases within cities. In the last decade, some studies have shown that differences in socioeconomic factors (e.g., income, poverty, and healthcare access), demographic factors (population density), and morphological features (e.g., growth patterns and slum density) of cities can affect disease transmission independently and in unison (Paul et al. 2012, Perkins et al. 2013, Vazquez-Prokopec 2016).

Despite the growing interest in disease dynamics within cities, an understanding of the

interplay of climate forcing with the socioeconomic factors that modulate transmission is still lacking. What is needed is the development of an integrated approach to the population dynamics of infectious diseases within urban environments that takes into account the rapid development in the monitoring and modeling of cities. In this paper, we argue that progress in this field requires a better integration of empirical and theoretical advances, and we review some major questions related to such integration, including: (1) How does the complex environment of cities influence disease transmission?, (2) How do climate factors in particular affect disease transmission in urban environments?, (3) How do climate and demography interact in space and time within an urban setting? And what are the consequences of this interaction in disease incidence?, and finally, (4) What are the open research questions and directions in this field?.

1.1 Spatial heterogeneity and complexity of urban environments

More than 50 percent of the worlds population lives in cities. Urban areas have become the largest land cover type, and the number of large cities has increased almost 10-fold since 1950 (UN 2012). Such rapid urbanization has raised concerns about the potential consequences for biological systems and human well-being,2 and understanding of demographic, economic, and environmental patterns within cities is growing. The myriad processes underlying the development and expansion of cities appear to create remarkable scaling patterns, which are reflected in power-law relationships between population size and a number of key quantities, including road length, gross domestic product (GDP), income, and poverty (Batty 2008, Bettencourt 2013). Parallels have been drawn between cities and other complex systems, from river networks (Rodriguez-Iturbe and Rinaldo 1997) and physiology (West et al. 1999), to insect colonies (Waters et al. 2010) and biodiversity patterns (Dendrinos and Mullaly 1985). Characteristic properties that cities share with complex systems include emergent patterns in space and time, dynamics far from equilibrium, considerable energy consump-

tion required for their maintenance, and large heterogeneity/inequality due to clustering of economic activities (i.e., economic agglomeration) and competition for space (Batty 2008).

In particular, the complexity of cities relevant to human health results from the multiple interactions between social, economic, and ecological forces acting at different spatial scales. Below, we highlight three main aspects of cities that are important for the study of infectious diseases.

Population Density

One relevant aspect of cities is variation in population density, which can differ from that in population size-a distinction that is rarely considered explicitly in the literature. Population density can affect contact rates but also susceptibility to disease, depending on its association with socioeconomic conditions. Without reference to the explicit denominator of area, the effect of a given population size on contact rates and susceptibility to disease cannot be fully specified. These epidemiological parameters are influenced by how crowded a location is and also in relation to associated socioeconomic conditions. Although population density in cities of developed countries has been shown to be, in general, positively correlated with the magnitude of socioeconomic activities and, therefore, with wealth, (Bettencourt 2013, Segal 1976, Bettencourt et al. 2007, Bettencourt and West 2010) the opposite can be the case for developing countries where more crowded areas can be poorer, and higher overall population size can translate into higher spatial socioeconomic inequality. Thus, consequences for local disease spread can differ for similar total population sizes and vary in complex ways across cities. Also, the relationship between several urban indicators, such as income and GDP and population size, has been shown to follow power-law patterns, reflecting underlying nonlinear interactions (Bettencourt et al. 2007, Bettencourt and West 2010).

Mobility Patterns

With regard to mobility patterns **(Fig 1.1 A-C)**, which have been associated with spatial variation in economic variables, urban planning, and infectious disease transmission (Bharti et al. 2011, Riley 2007, Cosner et al. 2009), several models have been proposed to infer mobility fluxes between urban environments and at different scales (Jung et al. 2008, Tarasov et al. 2013, Simini et al. 2012, Lu et al. 2013). To date, the most frequently used model is the gravity model in its different parametrizations, where the probability of moving between two locations is directly proportional to their respective populations and inversely proportional to their distance (Erlander 1990). More recently, the radiation model has been introduced as a generalization of the gravity model and has been validated on the basis of commuting data (Simini et al. 2011).

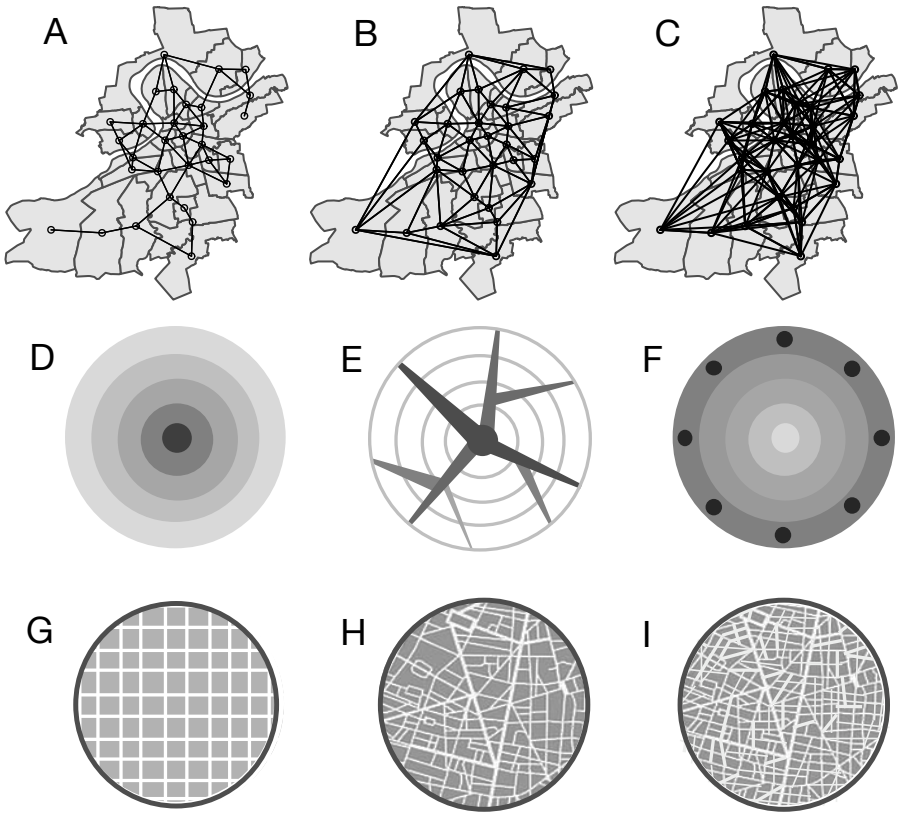

Figure 1.1: **Features of urban environments.** (A-C) Movement patterns influenced by geographical distance, social relatedness, and economic activities, among others. These patterns can ultimately have an effect on disease risk. (D-F) Patterns of city growth, including (D) compact cities, where most of the growth and development are located in the core, (E) corridor cities, with clear belts of development and urbanization, and (F) fringe cities, where the urbanization and expansion are concentrated in the periphery. Dark gray denotes high development and high population density (from Newton 1997). (G-I) Different street structures, from (D) regular patterns, characteristic of developed and wealthier areas within the cities, to (H, I) irregular patterns, characteristic of slums and economically deprived urban communities.

Urban morphology

Variation in urban morphology, which encompasses urban structures, land use, street patterns, types of buildings, and open spaces (Conzen 1988, Gauthier and Gilliard 2006), can be considered along two major axes: the first one concerns different growth patterns, ranging from those of compact cities, with most of the growth located in the core, to those of fringe cities, whose expansion is located in the periphery (Newton 1997, **Fig. 1 D-F**). The second class encompasses variation in characteristic urban features, including the regularity of the houses and streets (ranging from completely regular to irregular), which can be used to classify socioeconomic heterogeneity into typologies **(Fig 1.1 G-I)**.

Consideration of these three aspects (population density, mobility patterns, and urban morphology) is of relevance to address how spatial heterogeneity within urban environments influences disease risk. For instance, it is known that crowded areas tend to increase the contact rate among hosts, but disease transmission can also be affected by population density in a way that is not necessarily monotonic. The effect of population density, separately from that of population size, remains poorly understood in general for many infectious diseases and in particular for vector-transmission diseases, in which vector behavior may respond to host density (Antonovics et al. 1995). Likewise, movement can also determine host and/or vector mixing and contact rates, with unclear consequences for disease prevalence. Transmission models have primarily incorporated movement between cities (Xia et al. 2004, Viboud et al. 2006, Mari et al. 2012) Weselowski et al. 2015, Buckee et al. 2013), and only a few studies have considered movement within cities. For the latter, geographical distance, social proximity, and host movement to areas of high economic activity have been found to be good predictors of dengue risk (Reiner et al. 2014, Perkins et al. 2014, Honorio et al. 2009, Perez-Saez et al. 2016).

These results suggest that hosts can also move vector-borne diseases among neighbor-

hoods (Honorio et al. 2009). In the case of waterborne diseases, an extension of a transmission model of cholera that downscaled population size to map population density at high spatial resolution, and then inferred movement fluxes using a radiation model, fits the spatiotemporal burden of the disease significantly better than its predecessor with only near-neighbor effects (Perez-Saez et al. 2016).

Last, variation in urban typologies can exhibit association with disease transmission as a surrogate of socioeconomic and demographic factors. For example, urban development can trigger higher migration to areas that are perceived as changing and becoming wealthier, providing opportunities for work, which would increase the number of susceptible individuals and the likelihood of arrival of a new disease or strain. In addition, environmental and socioeconomic conditions may change rapidly in the expanding fringe of cities, affecting the suitability for vectors and the contact rates.

Although it is clear that a better understanding of disease dynamics and urbanization should encompass the physical, social, and economic environments (WHO 2008), the relationship between disease transmission and factors such as population density, poverty, or movement can be nonlinear and multidimensional, in ways that challenge inference of causality. Double causality is likely to be common, with variable x influencing variable y, at the same time that y influences x. For instance, socioeconomic inequality can influence disease incidence, and, in turn, spatial variation in infection levels can promote differences in wealth. Moreover, despite a growing literature on the association between urbanization and infectious diseases (Keiser et al. 2004, Lowe et al. 2013, Bharti et al. 2011, Stoddard et al. 2013, Delgado et al. 2011), transmission models typically consider urban systems at coarse scales, with no explicit consideration of different levels of spatial resolution and patterns of connectivity (Reiner et al. 2014, Perkins et al. 2014,Honorio et al. 2009, Perez- Saez et al. 2016, WHO 2008, Delgado et al. 2011, Barton and Grant 2006, Rydin et al. 2012)

(but see Thrall and Burdon 1999, Read and Keeling 2003 and Watts et al. 2005). Thus, the joint effects of environmental and socioeconomic factors on transmission processes and their dependency on spatial scale remain poorly understood. To a large extent, we still lack a quantitative framework that incorporates the interaction of climate forcing with demography and poverty in space and time into both dynamical models of transmission and statistical analyses at multiple scales.

1.2 Climate variation within cities

The remarkable expansion of cities creates extremely heterogeneous physical environments. For instance, urban zones in the city core may exhibit values of temperature and precipitation that are distinct from those observed in the periphery, with important consequences for pathogens and their vectors (Oke 1982, Arnfield 2003, Shepperd et al. 2010,Niyogi et al. 2011). Similarly, the replacement of natural soil and vegetation with built surfaces can elevate temperatures several degrees above those of the surrounding region (Oke 1982). This environmental variation at a higher resolution than that of the whole city can influence the spatial distribution of diseases and their spread to new locations (Jetten and Focks 1997). A well-known effect is that of urban heat islands, caused by daytime heat being retained by the fabric of buildings and by a reduction in cooling vegetation (Santamouris 2001). At temperate latitudes, this effect generates an increase in temperature of 15 ^{o}C during the night, whereas in the tropics, the temperature can increase to 10 ^{o}C by nights end, especially during the dry season (Santamouris 2001). Besides direct effects of such local changes in temperature on mortality risk (Honda et al. 1014), there are documented indirect consequences for the transmission of several infectious diseases (LaDeau et al. 2015).

Another important consideration in relation to climate is that a large proportion of the global population lives near coastal areas, and many of the worlds large cities are located either on the coast or near river mouths and estuarine wetlands on the coast (Mcgranahan et

al. 2007). The risk of flooding in cities near rivers that have been dammed for hydropower or water supply can be intensified by changes in sediment deposition (Syvitski and Higgings 2012). Likewise, changes in rainfall regimes may favor the maintenance of disease vectors, such as mosquito breeding sites, or may alter cholera risk via flood events that contaminate water sources with raw sewage (Jetten and Focks 1997, Ruiz-Moreno et al. 2007). This proximity to water bodies, combined with more extreme precipitation events and rising sea levels due to climate change, has the potential to increase the risk of flooding (IPCC 2007). Urbanization also directly creates flood-prone conditions by covering the ground with pavement and buildings and by the addition of urban drains, which cause runoff water to move more rapidly into rivers than under natural conditions (Douglas et al. 2008). In flooded areas, ditches, latrines, and septic tanks are key reservoirs that perpetuate malaria, dengue, and yellow fever (Unger and Riley 2007). Flood events also facilitate the exposure to severe waterborne diseases, such as typhoid, hepatitis A, cholera, and rotavirus (Lipp et al. 2002, Patz et al. 2000). Excessive or heavy rainfall can move pathogens into the environment, into rivers, coastal waters, and wells through the increased runoff of water from fields (Cann et al. 2013). Increased raw water turbidity can intensify gastrointestinal illness (Tinker et al. 2010). Thus, extreme climatic events can affect disease risk through an increase in host exposure and susceptibility (Kovats 1999).

The relationship between rainfall and waterborne or vector-borne diseases can be nonlinear, with effects in opposite directions at different extremes. For example, positive effects on larval development of mosquito vectors are known for adult Anophelines, the vectors of malaria, which reproduce in small natural ponds of clean water (Gage et al. 2009). But excessive precipitation may also have negative impacts on mosquito populations by washing away breeding sites (Kuhn et al. 2005). Drought in wet regions may decrease flow velocity in brooks and provide mosquitoes with more pools of stagnant water as breeding sites (Chretien et al. 2007). These considerations do not distinguish rural from urban environments.

In cities, variation and uncertainty in water supply, especially in arid regions with extreme seasonal changes in rainfall patterns, introduce the additional complexity of how humans manage water. This is relevant for vector-borne diseases, since many vectors rely on water stored by humans.

In addition to temperature and rainfall, humidity can also affect the temporal and spatial variation of disease prevalence within cities, although, of all climate variables, it is the least studied. Local humidity is influenced by proximity to rivers and other water bodies and can alter disease risk locally, as was shown for malaria, through an effect on parasite development within the Anopheline mosquito (Patz et al. 2003). Experimental studies have also shown that temperatures in the approximate range of 21-32 °C and a relative humidity of at least 60 percent are most conducive to the maintenance of malaria transmission (Metha 1934, Paaijmans et al. 2010, Cator et al. 2013, Mordecai et al. 20013).

Because the vectors need to live for at least 8 days in order for the parasite to complete its development and become infectious to humans, higher humidity can increase transmission rates by lengthening mosquito survival and enhancing activity rates (Snodgrass 1939, Gaaboub 1971). For dengue, Thu et al. 1998 found that both temperature and humidity favor the spread of the virus in mosquitoes during the rainy season, contributing to outbreaks of dengue haemorrhagic fever. It is not yet clear, however, whether humidity acts as an independent determinant of longevity or as an inversely correlated parameter with temperature.

Up to this point, we have considered how separate variation in population size/density, socioeconomic levels, movement patterns, or climate factors influence disease risk in urban environments. Next, we outline what is known about the interaction of these drivers in the context of heterogeneous cities and how those interactions ultimately define disease risk.

1.3 Interplay of climate forcing and urban heterogeneity: consequences for disease risk

Climate variability can operate at regional scales, synchronizing the population dynamics of infectious diseases across space (i.e., the Moran effect (Moran 1953)). However, this synchrony can break down within cities, owing to the spatial heterogeneity in the host population and environmental factors, as is the case for malaria (Santos-Vega et al. 2016) **(Fig 1.2 A and C)**. In particular, human density and water storage practices are strongly associated with poverty, urbanization, and population growth. Inequalities in access to water can exacerbate water storage practices that increase the risk of waterborne and vectorborne infection locally (Kovats et al. 2001).

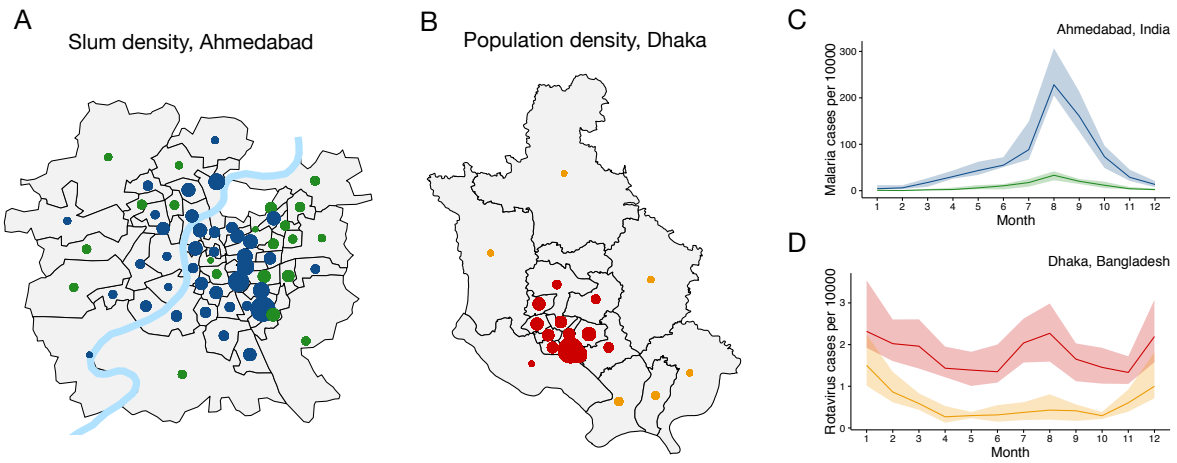

Figure 1.2: **Heterogeneity of urban landscapes and disease risk** (A) Slum density in the city of Ahmedabad, India (slums are usually illegal structures characterized by inadequate access of safe water, sanitation, and infrastructure, with poor structural quality of housing and overcrowding). The size of the dots is proportional to the number of slums per km2 of each administrative subdivision, with values that vary from 1 to 96 slums/km2. Green dots refer to administrative subdivisions of low malaria risk and blue dots to those with high malaria risk (Santos-Vega et al. 2016). (B) Population density of Dhaka, Bangladesh. The size of the dots is proportional to the number of individuals per km2 of each administrative subdivision, and the range of the values is 2500-465000 ind/km2. Red dots refer to the core part of the city, while orange dots illustrate the periphery, according to the grouping proposed by Reiner et al. 2013 and Martinez et al. 2016 (C) Malaria cases per 10,000, by month per region. The lines correspond to the median and shaded areas, to the firstthird quartile, with the colors following the same scheme as that in A. High risk is related to the proximity to the river (high humidity) and inversely correlated with economic level (Moran 1953). (D) Rotavirus cases per 10,000, by month per region. Medians and firstthird quantiles are also shown here with the corresponding color from B. High incidence is associated with the older, more urban, and dense part of the city and is depicted in red (Martinez et al. 2016).

The propagation of vectorborne diseases, such as malaria and dengue, which are responsible for significant morbidity and health risk in cities of developing countries,48 can be directly affected by the spatial distribution of the host population (Lowe et al. 2013,Perkins et al. 2013, Bousema et al. 2012, Trape et al. 1992, Wang et al. 2006, Qi et al. 2012, Texeira et al. 2002, Eisen et al. 2007, Banu et al. 2014). The behavior of the vectors is also influenced by the degree of urbanization: African malaria vectors, such as *Anopheles gambiae* and *Anopheles arabiensis*, exploit open water ponds for laying eggs and as their larval environment; they are therefore found mostly in peri-urban areas (Keiser et al. 2004).

By contrast, *Anopheles stephensi*, a malaria vector found in the Indian subcontinent, as well as the global dengue vector *Aedes aegypti*, are adapted to breed in water storage containers, making disease transmission a truly urban phenomenon (Cator et al. 2013, Hammond et al. 2007).

At coarse scales, transmission heterogeneity has been described in terms of spatial hotspots (Kovats et al 2001) and mostly related to vector biting patterns (Cator et al. 2014). Studies have found that there is important variation among households, owing to factors such as proximity to aquatic habitats containing immature mosquitoes, crowded urban slums, open city parks, suburban fringe residences (Bousema et al. 2012), type of housing (Woolhouse et al 1997), and the prevailing direction of the wind, among other factors associated with mosquito movement (Lindsey et al. 1988, Midega et al 2012). For malaria in a city of arid Northwest India, a recent study showed the existence of a largely stationary spatial pattern in disease incidence that was statistically associated with socioeconomic indicators and with temperature and humidity (Santos-Vega et al. 2016). A model accounting for climate, socioeconomic, and population factors has the highest predictive skill, suggesting that climate forcing and socioeconomic heterogeneity act synergistically at local scales on the population dynamics of urban malaria (Santos-Vega et al. 2016).

Furthermore, recent studies in the megacity of Dhaka, Bangladesh, have shown significant spatial heterogeneity in the interannual response of diarrheal diseases, such as cholera and rotavirus, to the El Nio Southern Oscillation (ENSO) and to flooding (Reiner et al. 2012, Martinez et al. 2016). The distinction between the core and periphery of the city seems important to capture the spatiotemporal dynamics of the disease at the level of administrative units, known as thanas, and to predict the response of cholera to ENSO at the level of the whole city. For rotavirus, the spatial variation of host population density is consistent with differences in the force of infection between the core and periphery (Martinez et al. 2016,

Fig. 1.2B and D). The higher populated core exhibits higher overall rotavirus transmission.

Population density and urban poverty may not only affect the response of diseases to climate forcing by exacerbating contact and susceptibility, but also by the related enhancement of asymptomatic reservoirs, which can, in turn, play an important role in the endemism of urban diseases. More specifically, in places with a higher force of infection, a larger fraction of the population may harbor the pathogen and be infectious to others without displaying clinical symptoms. The existence of such reservoirs may facilitate the response to environmental factors in diseases that are not considered particularly climate sensitive to begin with, as shown for the monsoons and rotavirus in Dhaka.9 The degree of endemism is reflected in the depth of the temporal troughs in incidence in between main seasons (**Fig 1.2 D)**. In the core of the city, disease persists at higher levels in between the two transmission seasons, and this persistence may enhance the ability of the transmission system to respond to flooding and generate a second season, typically absent in the periphery. Interestingly, the abovementioned study describes an enhanced seasonal response to the monsoons in the core of the city but higher interannual variability related to flooding in the periphery (Martinez et al. 2016). The response of disease dynamics to climate forcing at seasonal rather than interannual time scales may be a characteristic of more endemic systems and was also found for rural malaria in Senegal (Carter et al. 2000, Laneri et al. 2015).

Beyond these incipient studies in waterborne diseases and urban malaria, the role of socioeconomic and demographic factors in modulating responses to climate forcing is poorly understood in vectorborne diseases in general. Of particular relevance, the transmission of dengue is influenced by temperature and precipitation, through their effects on mosquito biology and ecology (Pascual 2015, Campbell et al. 2015, Chowell et al 2008, Descloux et al. 2012, Stewart- Ibarra et al. 2013). How this response to climate varies within urban settings and how the mosquito population dynamics can vary from strong to weak seasonality within

a heterogeneous city, also in the context of emerging diseases such as Zika and Chikungunya, remain open areas of investigation (Pascual 2015).

1.4 What is still missing?

An understanding of disease transmission in urban environments necessitates a quantitative conceptual framework that incorporates the diverse drivers reviewed above, and considers how the temporal variation in climate factors, at seasonal, interannual, and decadal scales, interacts with the spatial variation in susceptibility and contact patterns **(Fig 1.3)**. Such a framework would greatly benefit from exploiting ongoing rapid developments on the remote sampling of urban environments at high spatial and temporal resolutions, and from developing ways to take advantage of this information in the formulation of transmission models. Below, we outline some open areas that would contribute basic components of such a framework.

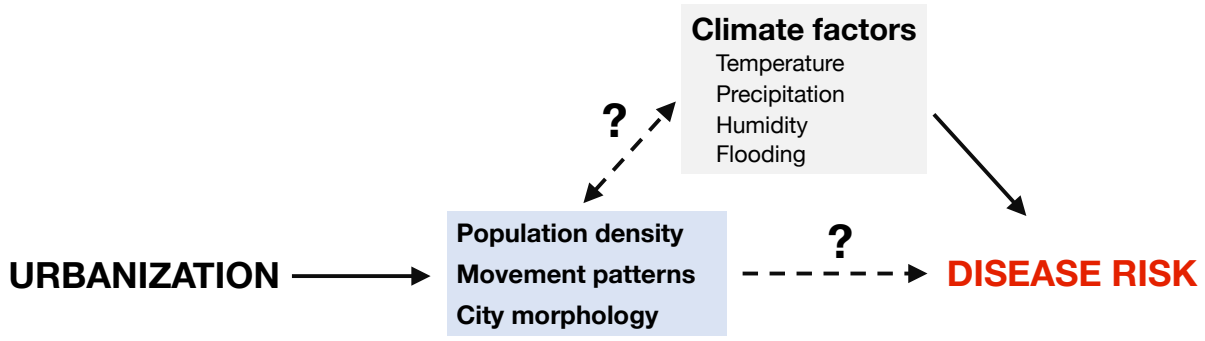

Figure 1.3: **Flow diagram** summarizing interactions between urbanization and disease risk. Dashed lines indicate poorly understood effects.

1.5 Conclusions

The growth of cities makes it imperative that we better understand how transmission is affected by the heterogeneity of urban environments and how this heterogeneity interacts with climate forcing. A few recent studies have begun to illustrate how the spatial heterogeneity of

socioeconomic and demographic conditions modulates the effects of temporal climate forcing on disease dynamics. Incorporating a finescale spatial component into transmission models represents an important technical and conceptual challenge that should be addressed in the next couple of years. Practical insights on the population dynamics of infectious diseases will be gained from incorporating urban features and their spatial structures. Transmission models that account for these complex dynamics will offer insights for control strategies that differ from those currently under-way, by achieving better informed localized interventions and, thus, more costeffective results. They will also contribute to spatiotemporal prediction based on climate variability in ways that account for different responses in different parts of the city and to the development of longterm scenarios of urbanization under climate change.

CHAPTER 2

POPULATION DENSITY, CLIMATE VARIABLES AND POVERTY SYNERGISTICALLY STRUCTURE SPATIAL RISK IN URBAN MALARIA IN INDIA

2.1 Introduction

Addressing health problems associated with urban growth will be one of the major challenges of the 21st century, especially for the developing world (Harpham 2009). City life is associated with significant variation in socioeconomic and environmental conditions of potential relevance to vector-borne diseases (Hay et al. 2005,Omumbo et al. 2005, Dye 2008). In particular, the pronounced and on-going increase in urban population (Keiser et al 2005, UN 2014), combined with climate change and economic disparities could act synergistically on the transmission dynamics of malaria (Robert et al. 2003, Donelly et al. 2003, Reiner et al. 2014). Although, Indian cities harbor both malaria parasites, *Plasmodium falciparum* and *Plasmodium vivax*, there is an increasing appreciation of the latter as a threat to global health (Roy et al. 2013), in particular in urban environments where Plasmodium vivax has become the most prevalent parasite. *Plasmodium vivax* has re-emerged in areas previously cleared of malaria, and has done so with higher mortality and morbidity than previously documented (Baird 2007,Sattabongkot et al. 2004, Price et al. 2007). In Indian cities with seasonal transmission, the incidence of vivax malaria starts to rise earlier than that of falciparum malaria, and also earlier than transmission via the vector would allow. This earlier part of the vivax season is dominated and enabled by relapses. The parasite has the ability to delay the development of a fraction of the infectious load of sporozoites in the liver (Gross et al. 1989), which results in the relapse of the disease after the primary infection is cleared from the bloodstream (Klinkenberg et al. 2008). The later part of the season is shared with *P. falciparum*, reflecting largely vector-borne transmission following the monsoon rains

(Baird 2007). A better understanding of the spatial heterogeneity of malaria vivax risk (defined based on the wards incidence) within cities remains an open area of research in the population dynamics of the disease, of relevance to both prediction and intervention.

Historically, urbanization has led to economic and social transformations associated with profound improvements in sanitation and hygiene (Qi et al. 2012, Trape et al. 1992). For malaria, the process of urbanization is generally thought to reduce transmission, primarily because urban environments are largely unsuitable for malaria vectors due to a lack of breeding sites and the pollution of potential larval habitats (Afrane et al. 2004). Other explanations for reduced malaria risk include better access to health care services and an increased ratio of humans to mosquitoes (Hay et al. 2005). There is concern however that areas with rapid unplanned urbanization and poor sanitation may not experience this marked decrease in malaria transmission (Keiser et al 2005).

The most common hypotheses for the persistence of malaria in cities include spatial variation in: 1) environmental conditions (relative humidity, temperature, precipitation), land use, and stored water, which create a favorable environment for Anopheles breeding in cities (Robert et al. 2003, Noor et al. 2003, Hay et al. 2009, Hay et al. 2007); 2) socioeconomic factors (income, human movement, population density and the failure of local malaria intervention among others), which hamper the effectiveness of case management and the promotion of intermittent antimalarials (Donelly et al. 2005, Stoddard et al. 2009, De Silva and Marshall 2012, Hasibeder and Dye 1988, Cosner et al 2009); and 3) at more local scales, variation in mosquito behavior and ecology which can influence transmission intensity (De Silva 2012, Hasibeder and Dye 1988). Despite increased interest in the role of spatial heterogeneities in the population dynamics of vector-borne diseases (Reiner et al. 2013, Perkins et al. 2013, Cator et al. 2013, Reiner et al. 2013), malaria models typically assume spatially homogeneous transmission and tend to aggregate temporal dynamics over space. In

particular, they do not take into account how spatial variation in environmental, climatic and socio-economic conditions affect vector habitat, contact rates, host susceptibility and the effectiveness of control (Cator et al. 2013, Reiner et al. 2013, Lawless 1987).

Importantly, these considerations are focused on Africa where endemic malaria remains a predominantly rural problem, because the main mosquito vectors are themselves rural, and in cities, largely peri-urban (Hay et al. 2005, Omumbo et al. 2005, keiser et al. 2004, Robert et al. 2005, Donelly et al 2005). By contrast, the Indian subcontinent harbors a truly urban vector, *Anopheles stephensi*, particularly thrives in man-made environments, and breeds in various artificial containers within homes and in water collected in construction sites (catch basins, seepage canals, wells) (Finger et al. 2014), whereas its sister (sub) species (myorensis) is associated with rural areas. The existence of this particular vector within cities poses a unique challenge to the elimination of malaria in India. Cities can act as a reservoir for the persistence of the disease beyond their administrative limits, and prevent elimination despite considerable gains in the fight against rural transmission. But even beyond the Indian subcontinent, malaria can no longer be considered only a rural problem, given the increasing proportion of the world population living in cities and transmission of the disease in urban environments sometimes surpassing that of rural ones (Hay et al. 2005, Omumbo et al. 2005).

Here, we describe the spatial pattern of urban *Plasmodium vivax* risk within a large city of India, and investigate its association with socio-economic and environmental heterogeneity. By combining statistical analyses and adapting a probabilistic modeling framework previously proposed for cholera (Metha 1934), we show that spatial heterogeneity in population size, environmental and economic factors, modulates malaria risk, and that the temporal effect of climate variables on malaria risk interacts with this spatial heterogeneity. These findings emphasize the importance of considering the interplay of climate forcing and socio-economic heterogeneity in the population dynamics of urban malaria in India. They also

provide a basis for more targeted intervention, such as vector control, based on identifying transmission hotspots. Comparison to findings for *Plasmodium falciparum* reinforce the evidence for a role of spatial heterogeneity in the transmission of urban malaria in this region.

2.2 Methods

2.2.1 Data

We took advantage of a highly disaggregated dataset of monthly malaria cases collected by the Municipal Corporation of the city of Ahmedabad, the capital of state of Gujarat in Northwest India **(Fig 2.1 A and 2.1 B)**. In this largely semi-arid state where malaria is seasonally epidemic, Ahmedabad reports more than 1000 cases every year. The city presents ideal conditions to investigate malaria transmission dynamics in an urban environment, since it has experienced rapid development, unplanned urbanization and large population growth. It is also located on the banks of the Sabarmati River which creates environmental variation in this otherwise arid setting. The malaria data consist of monthly cases for the dominant parasite, *Plasmodium vivax*, over the last decade, confirmed through microscopy examination of blood slides from clinical (febrile) individuals self-presenting at the health facilities **(Fig 2.1 B and 2.1C)**. The resulting time series span a total of 12 years (from 2002 to 2014). Because some administrative units (known as wards) were subdivided into smaller wards in 2007, we aggregated the 64 units of the city into 59 wards, in order to maintain the same geographical subdivision through time. For comparison purposes, a similar data set for *Plasmodium falciparum* is also considered.

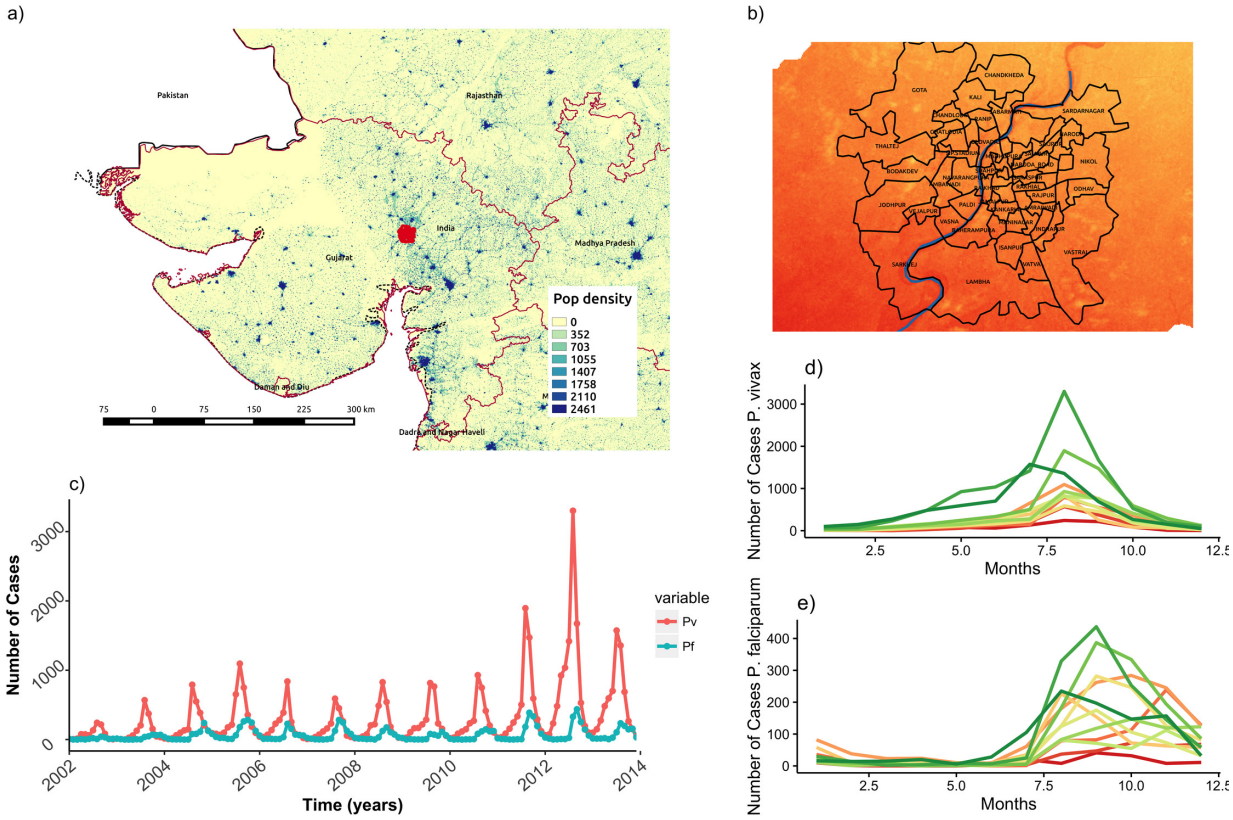

Figure 2.1: **Study Area**. Location of study area (A), and temporal patterns of incidence of *P. falciparum* (red solid) and *P. vivax* (blue dashed) (B, C). Boxplots are shown in B to illustrate seasonality, and time series are shown in C to illustrate interannual variation.

Apart from decadal census data, annual population data were provided by the Ahmedabad Municipal Corporation to approximate the population of each ward. Socio-economic data were obtained from the District Census Handbook of the concerned district for the year 2001 from the Directorate of Census Operations, Gujarat. Monthly time series (from 2002 to 2014) for mean temperature, mean rainfall and relative humidity at 8 am are those from the meteorological station of the city of Ahmedabad (and were provided by the Indian Institute of Technology in Gandhinagar).

2.2.2 Spatial regularities in malaria transmission

In order to investigate the existence of a spatial pattern in malaria incidence within the city of Ahmedabad, we performed a series of statistical analyses to address whether malaria risk varied within the city and what factors explained this variation. First, we analyzed the spatial and temporal variation of malaria incidence and identified regions of high and low risk based on incidence. Second, we performed a series of statistical analyses on the role of socioeconomic and environmental factors in the spatial, and spatio-temporal patterns of malaria incidence. These analyses ranged from a simple t-test comparing socio-economic factors between the two regions of differential malaria risk, to time series models incorporating the autocorrelation in the data and the external drivers (including climatic ones), to a full spatio-temporal general linear mixed model with random effects. Third, based on these results a probabilistic dynamical model was formulated for malaria transmission at the ward level, and predictions of this model were evaluated at the city level.

To consider a measure of vivax malaria risk independent from interannual variation, we normalized malaria incidence for each ward in a given year by the total number of cases throughout the city. We then ranked these normalized values across wards to determine if high risk locations were consistently so over time. To examine the robustness of the patterns, we complemented the estimation of the intensity of infection with the Slide Positivity Rate (SPR) (Smith et al. 2004) **S1 Fig 2.5**. SPR, defined as the number of laboratory-confirmed malaria cases per 100 suspected cases examined, provides a rapid and inexpensive means of assessing the burden of malaria in the population that relies on health care facilities. SPR provides a better estimate of the true intensity of malaria than malaria incidence (reported cases divided by population), as it is more robust to spatial sampling errors/biases resulting from variation in the population size of the wards (Smith et al. 2004). SPR is independent from incidence, although it can be prone to different biases (as the result for example of other febrile diseases which would increase the number of blood tests taken). Confirming

that the patterns are consistent across these two measures of malaria intensity is important because the population denominator that is relevant for the reporting may vary in ways that are difficult to determine. Also, the local population size might be under- or overestimated by a census carried out only once every 10 years in India.

To further characterize spatial variation in risk we applied a 2k-means cluster algorithm to the incidence data at the wards level, and examined the existence of at least two regions differing in malaria risk. (We pre-determined two groups of wards to consider the hypothesis of different transmission intensity in the core and periphery of the city). Also, two groups allow us to consider the subdivision of wards into high and low risk regions. We hypothesized that differences in malaria risk in these two main areas are largely explained by demographic and socio-economic factors. To test if the two regions differed significantly in those variables, we first extracted socioeconomic indicators, including slum density (number of slums/ward area), unemployment (number of unemployed people), number of marginal workers, literacy, population below 6 years, total population, number of households, vulnerable and economically deprived communities, from the 2001 Ahmedabad census and calculated the density of slums per ward based on cartographic information on the slums distribution within the city **S1 Fig 2.6**. A t-test was applied to evaluate if these two regions differed significantly in these covariates.

We then addressed whether the temporal variation in malaria incidence responded differentially to climate variables (rainfall, temperature and humidity) across the two regions. To determine which predictors explain the temporal variation in vivax cases, we considered models with autoregressive terms to account for serial correlation in the data. The correlation structure in the malaria time series was assessed by inspecting the autocorrelation function ACF. Then we applied a generalized linear model (GLM) framework. Because observed count data, such as reported cases in infectious diseases, often exhibit significant

over-dispersion (Bansal et al. 2007), a negative binomial distribution of cases was used. Backwards model selection was based on the minimum Akaike Information Criterion (AIC) **(S1 Table 2.4)**. Variables with coefficient significantly different from zero were selected. Spatial correlation was incorporated by assuming a conditional autoregressive (CAR) process in the random effect ν_{ij} :

$$\nu_i \mid \nu_j = N\left(\frac{\sum_j a_{ij}\nu_{ij}}{\sum_j a_{ij}}, \frac{\sigma^2}{\sum_j a_{ij}}\right) \tag{2.1}$$

where σ^2 controls the strength of the local dependence, and a_{ij} are neighborhood weights for each ward based on distance to the river. We additionally compared the best model to a model with a different distribution (Zero Inflated Poisson) but this model was not significantly better **(S1 Table 2.5)**. Finally, we considered the full spatio-temporal variation at the level of wards, by fitting a generalized linear mixed model including as covariates the effects of temperature and humidity, and a spatially-structured random effect weighted by the distance to the river (SSRE). Parameters and their distributions were estimated with Bayesian Markov Chain Monte Carlo (MCMC) parameter sampling implementation in OPENBUGS.

2.2.3 Probabilistic model for malaria dynamics

To model malaria risk within the city of Ahmedabad an inhomogeneous Markov chain model was used, following the theoretical framework developed for cholera by Reiner et al. 2012. In this approach the monthly malaria cases are categorized into discrete states of malaria incidence, which we chose as low malaria, mild malaria, and high malaria. The three discrete states partitioned the distribution of monthly incidence based on the 25th, lower than 75th and above 75th quantiles. Then, the model assigns baseline probabilities $P_{i,j}$ to the tran-

sitions between these states in a defined time step as described by the following transition probability matrix P

$$P = \begin{bmatrix} P_{11} & (1 - P_{11} - P_{13}) & P_{13} \\ P_{21} & (1 - P_{21} - P_{23}) & P_{23} \\ P_{31} & (1 - P_{31} - P_{33}) & P_{33} \end{bmatrix} \tag{2.2}$$

Where the baseline probabilities depend on the state of the system only in the previous time step. Modification of this basic matrix can introduce the effect of covariates, including the consideration of different regions such as the two groups of wards, identified in previous analyses. Thus, the resulting Markov chain model can be made inhomogeneous by allowing transition probabilities to depend on temporal and spatial environmental drivers; here seasonality, the state of the neighbouring wards, temperature, Relative humidity and socio-economic heterogeneity (summarized by the two groups of wards):

$$P_{i,j,k,t} = P_{i,j,d} \cdot Seas_{i,j,t,d} \cdot Neigh_{i,j,v,d} \cdot temp_{t,d} \cdot RH_{t,d} \tag{2.3}$$

Where $P_{i,j,k,t}$ is the probability that ward k goes from state i to j from time t to time $t+1$. This probability is dependent on: (1) $P_{i,j,d}$, the baseline transition probabilities of moving from state i to state j for a ward in risk region d; (2) a seasonal factor $Seas_{i,j,t,d} = (1 + \beta_{i,j,d})^{Se_{t,d}}$, where the seasonality exponent $Se(t,d)$ is periodic over the 12 months of the year and each group d has its own seasonality; (3) a neighbourhood effect , where v corresponds to the malaria state (0,1,2) of the neighbouring ward with the highest value. This function reduces to 1 when none of the neighbouring wards have malaria reported. The parameters of these functions, $\beta_{i,j,d}$ and $\alpha_{i,j,d}$ are estimated. Finally, the effects of temperature and humidity are included as sigmoidal functions, similar to the formulation for ENSO and its effect on cholera in Reiner et al 2012: where the temperature

or humidity with temporal lag , in this case 2 months for the humidity and 1 month for the temperature, A is the maximal amplitude, M controls the scale (normalization by the maximum values allowed for the climate factors), and h is shape factor varying between 0 and π to go from a linear to a nonlinear effect **(S1 Fig 2.7)**.

We considered different models obtained by including or neglecting the effect of a subset of the following factors: temperature, relative humidity, the state of the neighboring wards and the two different risk regions. We compared each of the models to a null model) employing a likelihood ratio test. The most complex model has 78 parameters **(S1 Table 2.6)**. Under Markovian assumptions, the transitions for the different time steps (months) are independent from each other (Metha 1934, Bui et al. 2011), which allows us to explicitly write a likelihood. The constraint that each transition probability must be between 0 and 1 was imposed by a barrier method (i.e. by setting the probability to 0 whenever its estimated value falls outside these limits) (Metha 1934). Each model was fitted by maximizing the likelihood with a Nelder-Meade simplex algorithm, which allows for the incorporation of such constraints.

Finally, to assess prediction performance, a cross-validation approach was implemented by sequentially removing the epidemic months (August-November) that follow the monsoon in a given year, refitting the model to the remaining data, and simulating it four months ahead starting from August to predict the course of the seasonal outbreak for the omitted period. Forecasting accuracy was estimated by computing the likelihood of the observed state. To that end, we inferred the probability distribution of the predicted state by performing 5000 independent simulations. This procedure is then sequentially repeated removing, one at a time, all the epidemic seasons available. To quantify the accuracy of our predictions, we calculate the percentage mean absolute error in our predictions, as well as a second quantity more practical and possibly relevant to public health, based on the definition of

a large outbreak. We defined such as event as one where the peak of the epidemic at the whole city level exceeds the 75 quantile of the distribution of this quantity. We quantify the fraction of times the model correctly predicts the observed malaria incidence state (above the 75th quantile). Then, to examine and illustrate the importance of the climate covariates to the predictions, we simulated the model using different combinations of humidity and temperature data. In particular, predictions for the epidemic months in the anomalous, low incidence years, were obtained using: 1) monthly observed temperature and average humidity 2) monthly average temperature and observed humidity 3) monthly average humidity and average temperature. Monthly averages were computed based on the mean of all previous years for a given month. We also examined the performance of the model by obtaining a one-step ahead prediction, where we removed 1 month of data at a time for all the wards **(S1 Fig 2.8)**.

2.3 Results

We initially addressed whether the spatial distribution of normalized *Plasmodium vivax* risk is heterogeneous throughout the city, and how the ranking of risk varies in time. The top left panel of Fig 2 shows that several locations systematically rank high or low based on their *P. vivax* incidence through time. The top right panel in this figure reveals a stable spatial regularity of the locations with the highest malaria burden through time. This remarkable temporal stability of the spatial pattern suggests the existence of strong underlying determinants that are largely stationary at the temporal scales of malaria transmission within the city.

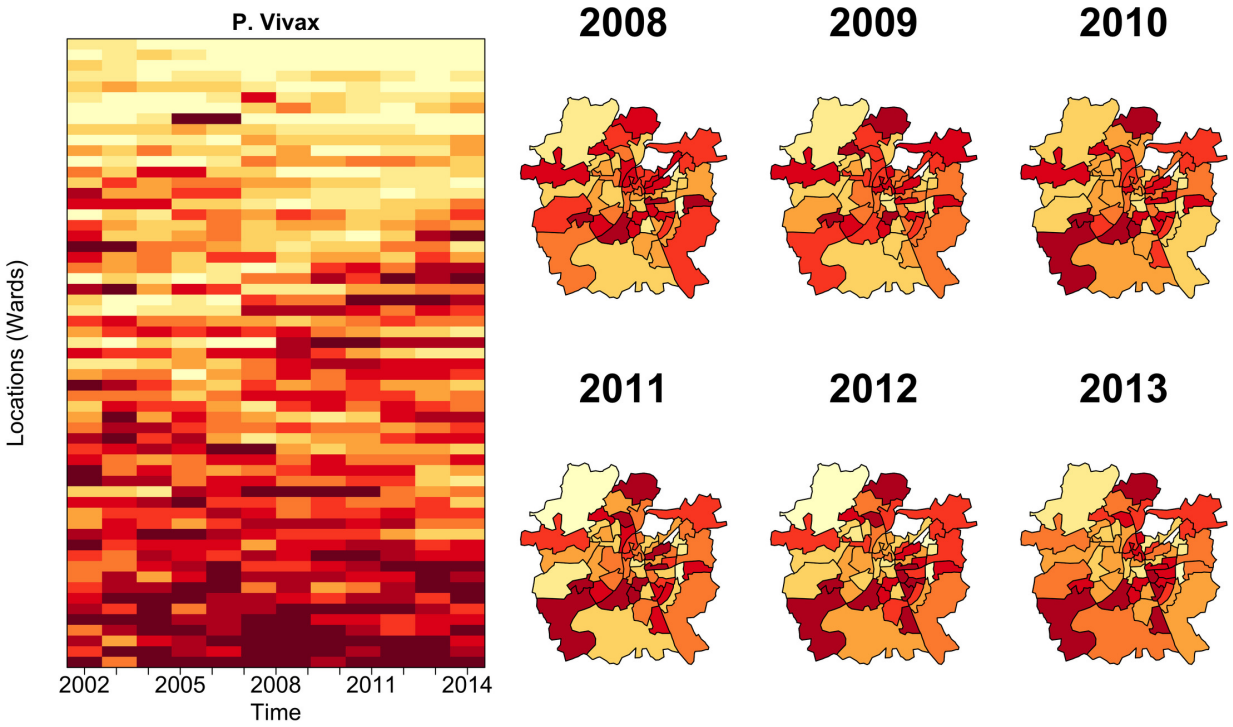

Figure 2.2: **Analysis of spatio-temporal patterns of malaria vivax incidence in the city of Ahmedabad.** The panels show the distribution of the cases normalized by population, with the intensity of the color (from low yellows to high reds) corresponding to the ranking of incidence. There is striking consistency from one year to the next in the places exhibiting the highest burden of the disease. Some of this regularity also extends to the two parasites.

The spatial pattern observed in **Fig 2.2** also indicates the existence of two distinct malaria risk regions within Ahmedabad: one comprised of the wards close to the Sabarmati River and the core of the city (region 1), and the other, of those in the newer urban periphery (region 2).

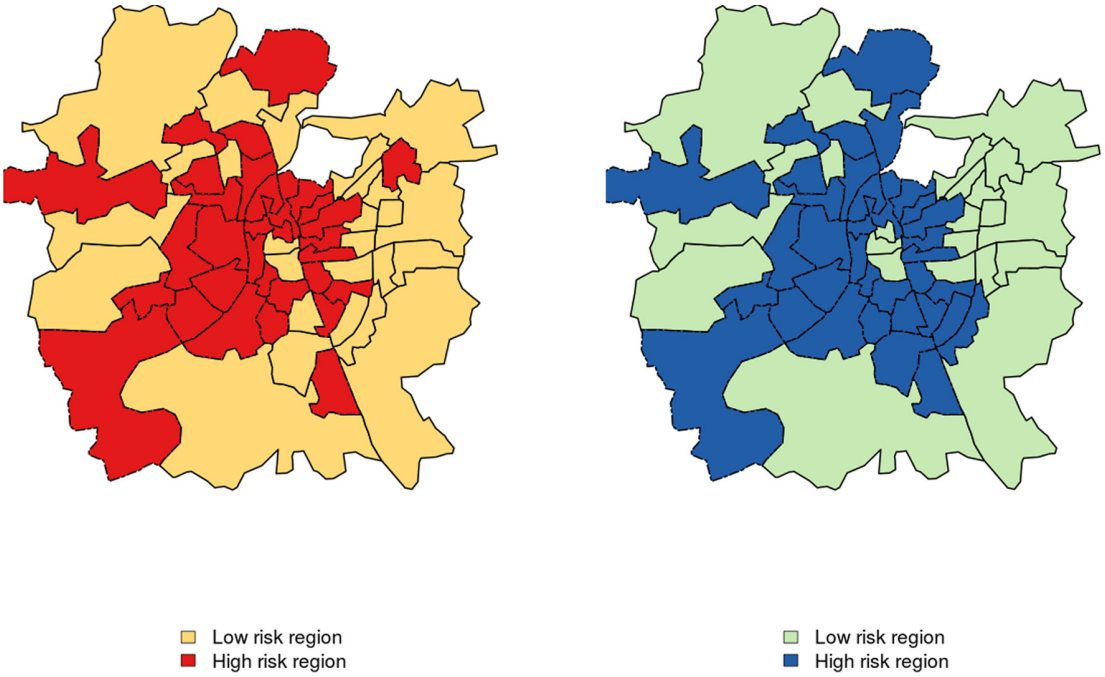

Figure 2.3: **Map depicting the two groups of wards (administrative units).** Map depicting the two groups of wards (administrative units), with high and low malaria risk respectively, *P. vivax* (left) and P. falciparum (right). There are significant differences in the annual malaria incidence between the two regions ($p<0.001$), for both parasites.

Fig 2.3 shows the results of the cluster analysis using the rankings data, demonstrating that the group of wards that are close to the river and in the inner part of the city (referred to as high risk region hereafter) have a different malaria risk than those in the in the periphery (low risk region). *Plasmodium vivax* cases in the region defined as high risk are significantly ($p<0.001$) higher than those in the region defined as low risk, whereas the seasonal pattern remains the same for both areas. This results hold for both parasites **S1 Fig 2.9**.

We then asked whether the differences in malaria risk between the low and high malaria risk regions are associated with differences in population density (number of slums, population size, number of households) or economic level (income, unemployment, literacy). Table 1 summarizes the results from statistical comparisons (t-test) between the high and low risk regions (for *P. vivax*) based on socioeconomic indicators from the 2001 census. We found that the high risk region has significantly higher unemployment, slum density, total popu-

lation and number of households. In addition, we did not find significant differences in the mean area of the wards. The differences in slum density, unemployment, literacy, economically disadvantaged people and vulnerable communities (both Scheduled Castes, SC, and Scheduled Tribes, ST, the official designations given to two different groups of historically disadvantaged people) are pronounced, with the high risk region encompassing for example a density of slums that is at least 1.2 higher than that of the low risk region **(Table 2.1)**.

Table 2.1: **Statistical analysis of differences between the two regions** for *P. vivax*, based on the socioeconomic information of the 2001 census.

variable	**statistic.t**	**parameter.df**	**log(mean of High risk)**	**log(mean of low risk)**
Slum density	1.834*	54.99	3.261	2.747
Unemployment	3.059***	29.21	10.82	8.828
Marginal workers	3.197***	30.56	6.967	5.592
Literacy	3.217***	29.18	10.884	8.566
pop_below_6years	3.275***	29.43	9.095	7.120
Total population	3.237***	29.18	11.20	8.808
Area	0.336	45.51	1.379	1.438
Number of Households	3.178**	29.29	9.565	7.544
Vulnerable communities	3.210**	40.75	6.207	4.825
Economically deprive communities	3.934**	36.14	8.841	6.682

Moreover, the temporal variation in malaria incidence between the two regions could be influenced differentially by the environmental covariates. **Table 2.2** show the results of the stepwise multiple linear regression between *Plasmodium vivax* cases and climate covariates (temperature, humidity and rainfall) and **S1 Fig 2.10** show the residuals of our best model for each region. The temporal variation in the low risk region responds predominantly to changes in temperature and humidity, whereas that of the high risk region is explained predominantly by humidity, although both parameters are seasonally associated with rainfall. Humidity would act in both the temporal and spatial dimensions, as the water table and associated ground moisture should be higher in the high risk region closer to the river, and

this parameter is known to affect survival of the adult mosquitoes (Sachs and Malaney 2002). Thus, increased humidity (through changes in the water table) can lead to an earlier onset of the *Plasmodium vivax* season and to higher incidence **S1 Fig 2.11** Finally, results for the general linear mixed model show that the best model includes the effects of temperature and humidity, and the random spatial variation in the data weighted by the distance to the river **(S1 Table 2.7 and S1 Table 2.8)**.

Table 2.2: **Results of the best model** for monthly cases as a function of environmental covariates for Plasmodium vivax.

Low risk region						
Variable	Estimate	Std. Error	z value	Pr(>—z—)	2.50%	97.50%
ar1	0.63891	0.064	9.917	0	0.448	0.718
intercept	-1.257	2.126	-0.591	0.554	-13.702	9.159
temp	0.140	0.069	2.006	0.044	-0.035	0.725
RH	0.034	0.025	1.329	0.0183	0.027	0.258
High risk region						
Variable	Estimate	Std. Error	z value	Pr(>—z—)	2.50%	97.50%
ar1	0.5835	0.0689	8.4637	0.0000	0.4483	0.7186
intercept	-2.2716	5.8323	-0.3895	0.6969	-13.7027	9.1594
RH	0.1775	0.0235	3.2948	0.0010	0.1332	0.3952

Interestingly, *Plasmodium falciparum* also exhibits stable regularities in risk levels within the city **S1 Fig 2.12**. We find the existence of two different regions with contrasting incidence, largely consistent with those for *Plasmodium vivax* **Fig 2.3**, and a significant difference in socioeconomic level and environmental conditions for these two regions **(S1 Table 2.9 and S1 Table 2.10)**.

For a more dynamical perspective, we used an inhomogeneous Markov chain model that incorporates the effect of spatial and temporal variation on malaria risk. Results are also consistent for both parasites. Specifically, the comparison of the different models analysed **(S1 Table 2.7 and S1 Table 2.8)**, shows that the best model is the one that accounts for

the effects of seasonality, neighbours, temperature, humidity, and includes the two different risk regions identified above **(model 5 Table 3)**. To test the significance of the individual components of each model (alternative hypothesis) against a model including seasonality only (model number 1, null hypothesis) we employed a likelihood ratio test. Improvements in likelihood for models 3 to 8 are significant, and so are the effects of seasonality, temperature, neighbours and the two regions ($p<0.05$). Improvements in likelihood for model 2 (seasonal effect + neighbourhood effect) are not significant at $p = 0.05$. Moreover, models that account for the effect of spatial heterogeneity, represented by including the two risk regions, tend to perform better than those that do not incorporate this effect **(S1 Table 2.7 and S1 Table 2.8)**.

Table 2.3: Likelihood comparison of the models showing the covariates included in each model.

models	seasonality	Neighbors	Temperature	RH	**Regions**	**Log lik**	**DF**	**delta AIC**	**LRT**
Model 1	+					-6656.621	27		–
Model 2	+	+				-6648.966	33	3.31	
Model 3	+	+	+			-6573.846	36	144.24	*
Model 8	+				+	-5932.35	54	1246.99	*
Model 4	+	+	+	+		-5894.961	39	104.78	*
Model 7	+	+			+	-5670.935	66	394.05	*
Model 6	+	+	+		+	-4903.554	72	1522.763	*
Model 5	+	+	+	+	+	-4821.559	78	151.988	*

Finally, **Fig 2.4** shows the predictions for the epidemic months for *Plasmodium vivax*, generated with the cross-validation procedure described above and the best model, which includes seasonality, temperature, the state of the neighbouring wards, and spatial heterogeneity. The predicted rate of cases for the peak of the season is coherent with the observed rate. Further explorations of the model show the importance of temperature and humidity in the prediction of malaria risk **Fig 2.4 B**. For example, the pronounced dip in the cases for years 2009 and 2010 is captured by the dependencies of the model on temperature and humidity, as those two years are dryer, warmer and less humid **Fig 2.4 C and D**. Although our best model exhibits a tendency to under-predict the size of the peaks which we discuss below, it is able to capture the seasonality and to a reasonable extent, the interannual vari-

ation in the data. The anomalous decrease in the number of cases in 2009 and 2010 seasons can be explained by the variation in the climate covariates.

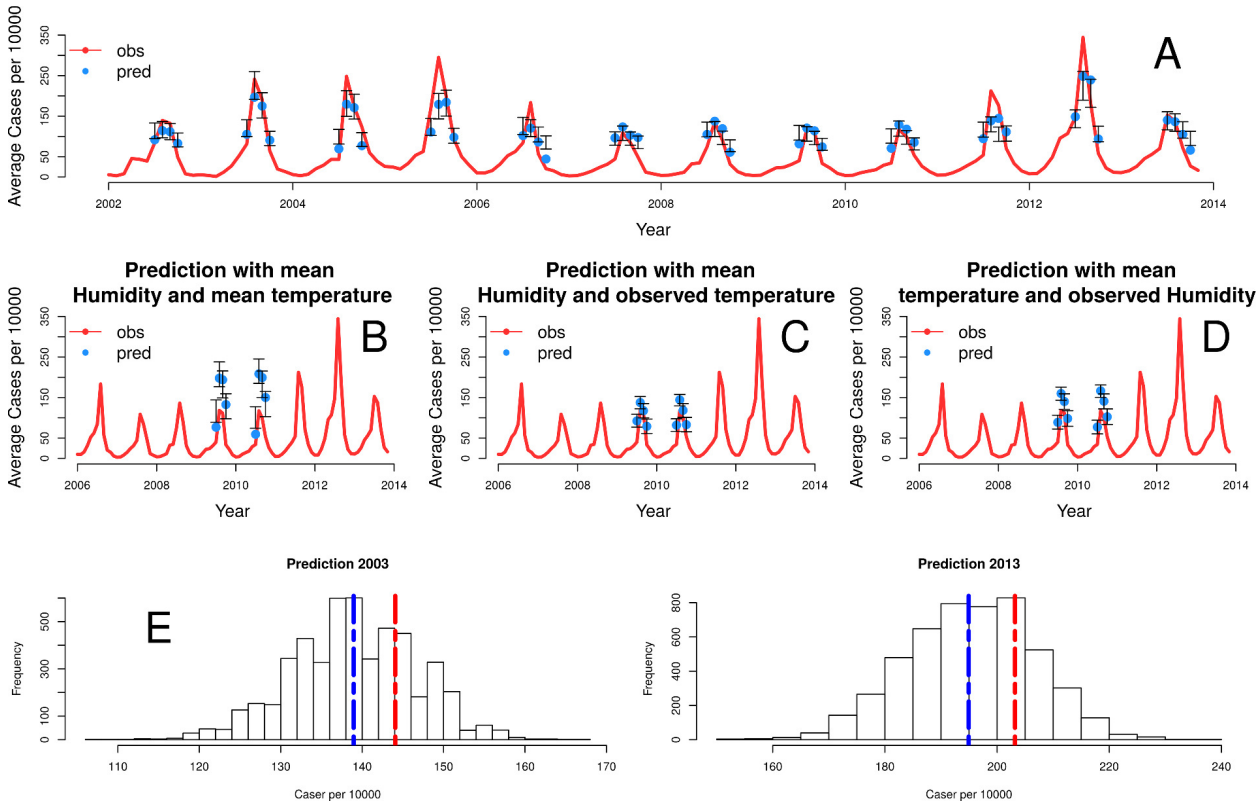

Figure 2.4: **Comparison of observed and predicted cases with the best model.** In (A), the red line corresponds to the average number of cases per 1000 for the 59 wards. The blue dots correspond to predictions given by the median of 5000 simulations, and the gray bars correspond to the 5th and 95th percentiles. In (B-D), simulations of the model predict the seasonal epidemics of 2009 and 2010 starting from the end of the monsoons (August) under modifications of the observed climate covariates. The different panels show the effect of fixing temperature and/or humidity at their mean monthly values, to remove their effect on the interannual variation of these anomalous years. When the interannual effect of both is removed (B), the model clearly over-estimates the cases. Individual effects are less pronounced (C, D) although predictions are also higher than observations. Our best model has a mean absolute error of 68% for predicting the peak of the epidemic in a year with a high number of cases (2013). **Fig 2.4 (E, left),** shows the distribution of model forecasts from 5000 runs for October 2003 based on October 2002 data. Although the mean prediction differs from the observation, almost all (84%) model simulations resulted in large events for 2003. The figure on the right repeats this hindcast analysis for October 2013 (using data from 2012). Here, we find a reduced but still large (87%) probability of a large outbreak.

Because the model is stochastic and it considers discrete states (no malaria, low and high), we simulated repeatedly, and from these ensemble of simulations computed the mean number of cases in a given month for a given ward. Our simulations generate realizations of the stochastic process and therefore, configurations of the discrete states. To convert the discrete states to cases, we used the mean number of cases for each class. The red line corresponds to the wards mean observed cases of the city. The blue dots show the median of the simulated values and the blue shaded regions correspond to the 5th 95th percentile range over 5000 simulations.

2.4 Discussion

Most transmission models of vector-borne diseases tend to aggregate the data at large scales and treat transmission homogeneously in space (Cator et al. 2013, Reiner et al. 2012, Lawless 1987). However, at local scales spatial heterogeneity can significantly influence the risk of infection. In particular, urban environments can exhibit pronounced heterogeneity from rapid and unplanned urbanization. Spatial heterogeneity in the environmental conditions such as temperature and humidity or in socioeconomic level can affect mosquito ecology, such as habitat distribution, vector longevity, biting rate or host finding ability (Levin 2007), and in factors related to human exposure and susceptibility respectively (Donelly et al. 2005, Qi et al. 2012,Worrall et al. 2005).

For *Plasmodium vivax* in Ahmedabad, we found defined heterogeneity in malaria risk that is slow-changing and therefore largely stationary in time, relative to the characteristic temporal scales of the population dynamics of the disease. The presence of such stable pattern suggests strong and spatially-structured determinants of malaria risk. In particular, the existence of two main regions with different risk was shown to be associated with socioeconomic level. Higher risk is largely concentrated in the inner part of the city where socio-economic indicators reveal higher poverty on average **(Fig 2.3 and Table 1.1)**. These

results are consistent with the previously described negative associations between malaria and socioeconomic status (Ajayi et al. 2013, Snodgrass 1939, Gaaboub et al. 1971, Kessler and Guerin 2008). Disease persistence decreases with increasing employment, literacy and income **S1 Fig 2.14**, with poor people more vulnerable to ineffective diagnosis and treatment for financial and cultural reasons, and less able to access antimalarial and anti-mosquito protection (Mayne 1930). Consistent results for *Plasmodium falciparum* emphasize the driven nature of the patterns.

The two different risk regions within the city were also shown to exhibit differential temporal responses to climate forcing. This finding underscores the importance of humidity to malaria transmission, with a higher water table in the high risk region possibly increasing relative humidity and affecting vector ecology. The spread of malaria requires favourable conditions for the survival of both the mosquito and the parasite. Temperatures in the approximate range of 21-32C and a relative humidity of at least 60% are most conducive to transmission (Ajayi et al. 2013). Malaria vectors need to live at least 8 days in order to transmit malaria, and higher humidity increases both survival rate and activity rate (Snodgrass 1939, Gaaboub et al. 1971). These relationships explain why *Anopheles stephensi* is more active and prefers feeding during the night when relative humidity is higher. The low risk region should have lower humidity (ground moisture) given its distance from the Sabarmati river and other water bodies. It has been established that if the average monthly relative humidity (measured at 8 am) is below 60%, then the lifespan of the mosquito is too short leading to very low or no malaria transmission (Dye 2008, Kessler and Guerin 2008, Mayne 1930). The effect of humidity is also evident in the estimated regression coefficients **(Table 1.2)**, where the value corresponding to humidity is one order of magnitude higher for the high than the low risk region. This means that a typical annual change of 10% in humidity with the rest of the covariates kept the same for both regions, will result in an increase of 403 to 1030 additional malaria cases (or 15% to 39% relative to the mean) in

the high-risk region, and 70 to 411 additional cases (or 7% to 33% relative to the mean) in the low risk region. Rainfall itself, which is closely related to both humidity and temperature, is not retained in the regression as a significant explanatory variable probably because of its collinear effects with the other climatic factors. Our results from the mixed model indicate however that while climate factors (temperature and humidity) play an important role in disease transmission, it is their combined effect with spatial heterogeneity in the risk weighted by the distance to the Sabarmati river that better explains the spatio-temporal variation in the malaria incidence. More mechanistic models informed by time series data on disease incidence and vector abundance, will provide further insight on the role and interplay of climate covariates through their effects on parasite mortality, as well as on vector breeding and longevity.

Social and economic elements such as the quality of housing can also favor the biological development of mosquitoes (Kirby 2008). It is common in urban areas of India for water to be supplied irregularly; this leads to water storage within houses which creates multiple breeding sites for the mosquito in overhead tanks, cisterns and cement tanks (Sharma 1996). Higher population density would result in higher water storage concentrations in close proximity to people. At the local scale spatial heterogeneity in urban malaria risk would follow from high density, especially where coexisting with poverty, as well as environmental microclimates from the proximity of water bodies.

Our dynamical and stochastic model captures the seasonal pattern and the main trends in the interannual variation of the malaria cases. Interestingly, most of the models that incorporate spatial structure, namely the two regions, perform better than the models that do not. This conclusion is consistent with the results for diarrheal diseases in Dhaka, where consideration of different parts of the city also improved model performance [31, 48]. Our best model identifies significant spatial effects at two different scales: (1) that of neighboring

wards ($p<0.01$), where the probability of transitioning to a higher risk level depends of the level of the surrounding wards, and (2) that of the two regions ($p<0.001$) influenced socioeconomic and demographic level. Although our model under-predicts the size of outbreaks, this tendency is expected from the discretization of the cases into a small number of levels, which preserves the rankings of the cases over time but tends to reduce the magnitude of the peaks. Predictions in that scale can still be useful when evaluating the risk of an outbreak larger than a given selected threshold [31]. Although additional classes could be incorporated in the formulation of the model, this would rapidly increase the number of parameters.

The model is able to capture the interannual trends and in particular, the lower outbreaks of years like 2009 and 2010, based on the effect of climate covariates. These two years exhibit anomalous high temperature and low humidity (associated with low monsoon rainfall), only comparable to values in 2002, another year with low incidence **S1 Fig 2.14**. The predictions of **Fig 2.4B**, for temperature and humidity kept at monthly averages, show that the model over-predicts the number of cases **Fig 2.4**. Thus, the anomalous incidence is explained by the lower temperature and higher humidity. **Fig 2.4C and 2.4D** isolate the effect of each climate covariate on the prediction, showing a stronger effect of temperature on the temporal reduction of cases of those two years.

Besides prediction, the phenomenological modeling framework applied here is also useful to address the spatial scale at which to aggregate the data to consider process-based epidemiological models, in a way that balances reducing the noise with representing dominant spatial heterogeneity. Questions on the spatial scale of aggregation are specifically relevant to addressing climate forcing in the context of socio-economic heterogeneity. Given that pronounced changes in urbanization will co-occur with those in climate, these are fundamental questions for infectious disease dynamics within cities.

Although our approach is able to capture the interannual variation in the data and predict the peak of the epidemic, it could be improved in several directions. For example, one could incorporate in the model: (1) mobility fluxes derived from the spatial distribution of the population with movement models, to replace the near-neighbor effects on transition probabilities; (2) the explicit effect of population density on group-dependent parameters explicitly; (3) further analysis of the local effect of environmental heterogeneities such as river discharge and soil moisture on malaria incidence at higher resolution by increasing the number of groups in the model. Moreover, temporal changes of the city itself would be of interest, including changes in the local speed of urbanization, and their implications for mobility, population distribution and economic level.

The region has experienced strong malaria interventions in the last three decades reflected in the pronounced negative trends in the number of reported cases from the 1980s and 19990s to the 2000s. From 2000 onwards, malaria prevalence in the city of Ahmedabad has remained however fairly stationary. Although the spatio-temporal variation in intervention efforts could influence the results of our models. This is unlikely given that the interventions within the city are largely homogeneous in space **(S1 Fig 2.15)**, with small differences between the high and low risk regions.

Our probabilistic model tends to underpredict the size of outbreaks. This bias results from the transformation of the incidence data into discrete malaria levels, which smooths out extreme events. The effect of this bias can be assessed and corrected by lowering the threshold probability (the proportion of simulations with large outbreaks) below 50% using ROC (Receiver-Operating-Curves) (Reiner et al., 2012). The number of discrete classes describing the malaria levels could also be increased or estimated to balance complexity and accuracy. At the limit, one could move to stochastic models that do not require such discretization, although their parametrization would present challenges related to model complexity.

A better understanding of urbanization and malaria is needed, since urban environments can contribute to the persistence of the disease and frustrate elimination efforts more broadly at a regional level, by creating a reservoir for the disease in cities that contributes to transmission in rural areas. In India, the earlier National Eradication Programs focused on rural areas, with urban malaria contributing to the resurgence of disease in the 1970s [48]. Urbanization and growing populations also exacerbate inequalities in access to water, and in so doing introduce variation in another fundamental but poorly understood environmental factor for vector-transmitted infections. In concert these two dimensions, environmental and socio-economic, define the relevant spatial scales at which to address transmission dynamics. They also define the existence of spatial 'hotspots' of high disease risk in urban environments, especially those of the developing world where economic inequality and variation in disease vulnerability can be pronounced. Identifying these hotspots for targeted intervention in urban environments, can contribute to the control of vector-transmitted diseases, and eventually to the elimination of malaria in India.

2.5 Supplementary information

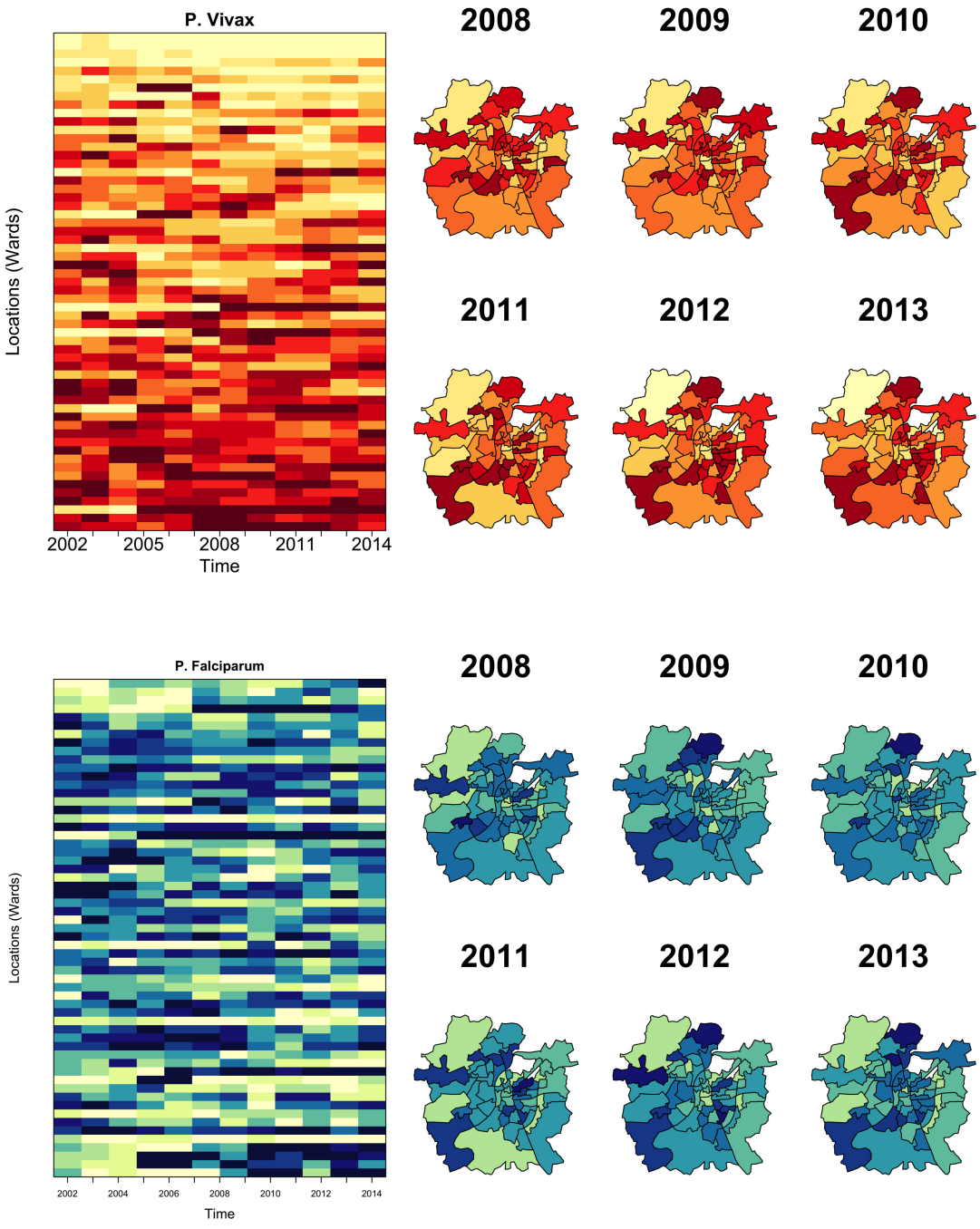

Figure 2.5: **Analysis of spatio-temporal patterns of malaria slide positivity rate in the city of Ahmedabad.** The panels show the distribution of SPR with the intensity of the color (from low yellows to high reds for P. vivax and from light yellow to blue for *P. falciparum*) corresponding to the ranking based on the intensity of the transmission

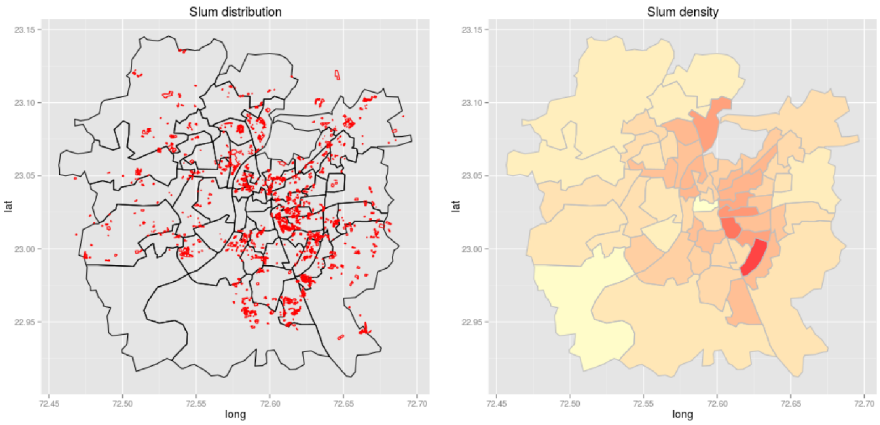

Figure 2.6: **Slum distribution in the city (left) and slum density calculation (right).** The latter was generated by overlaying the slum distribution map with the wards map provided by the municipal corporation, and calculating the number of slums per ward divided by the ward area.

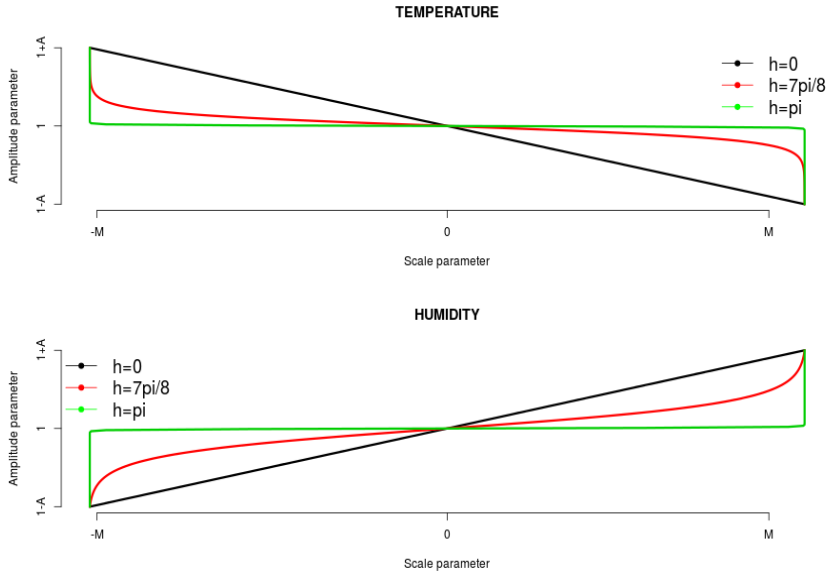

Figure 2.7: **Functional forms of the humidity (bottom) and temperature (top) effects illustrating the flexibility of the formulation, with different values of parameter h leading to different shapes.** (Here we show these functional forms for an arbitrary amplitude A and scale M, and different shape values: h = 0 (black line), h = 7/8pi (red line) and h = pi (green line).

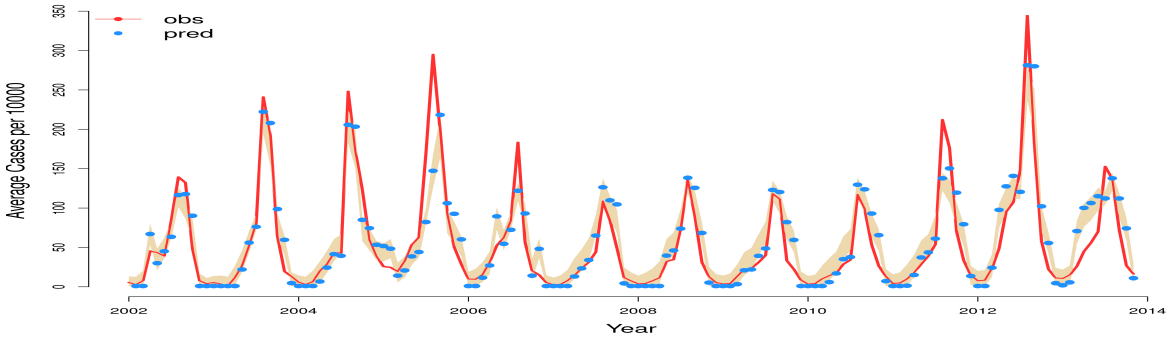

Figure 2.8: The red line corresponds to the average monthly cases per 10000 for the 59 wards. The blue dotted line corresponds to one-month ahead predictions for the median of the 5000 simulations values, and the light brown shaded region corresponds to the interval between the 5th and 95th percentiles for these simulations.

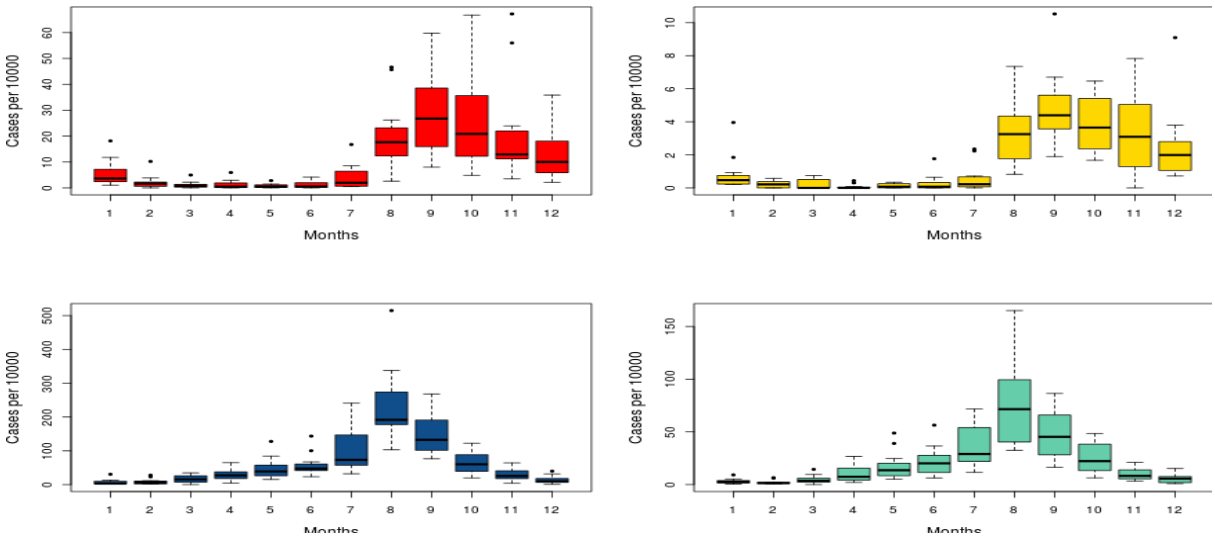

Figure 2.9: The top panels represent the seasonal pattern for *Plasmodium falciparum*, for the high risk region in the left and the low risk region in the right. The bottom panel shows the corresponding patterns for *Plasmodium vivax*.

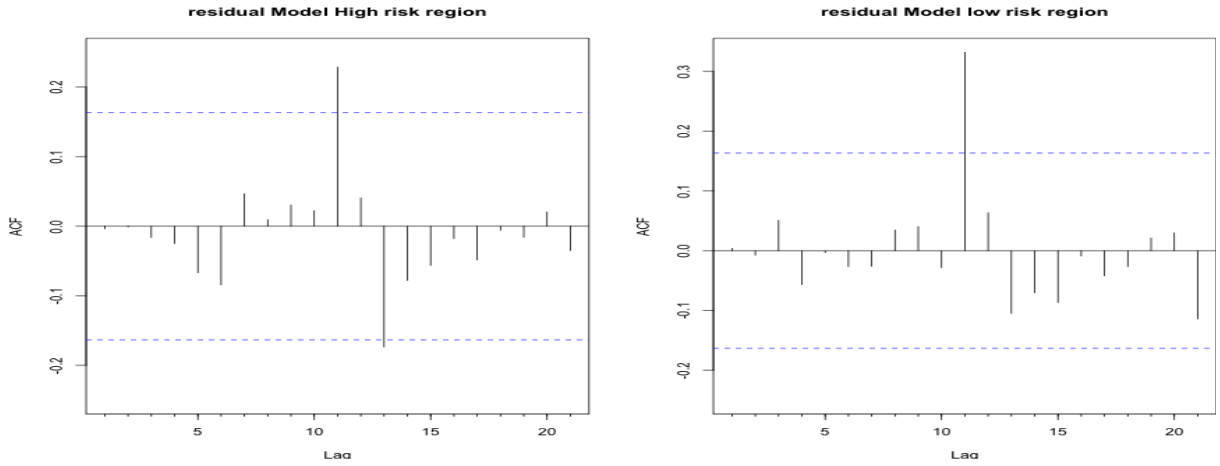

Figure 2.10: Although most of the autocorrelations fall within the confidence intervals, there is a small autocorrelation at lags of 11 and 12 months (seen in the significant spike of the ACF plot). This suggests that the model can be slightly improved by capturing the remaining seasonal variation.

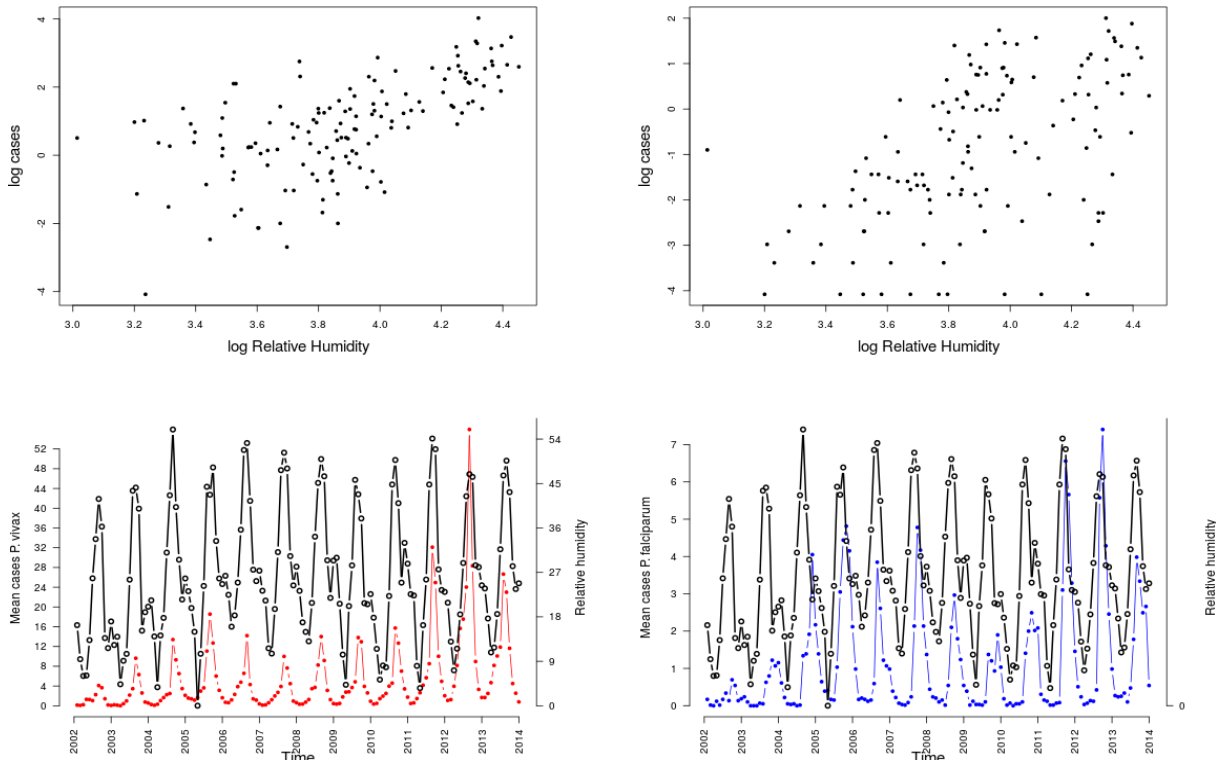

Figure 2.11: The top panels show scatterplots of humidity vs cases for P. vivax (left panels) and P. falciparum (right panels), and the bottom ones, the time series for the cases (in red for vivax and in blue for falciparum) together with those for humidity (in black).

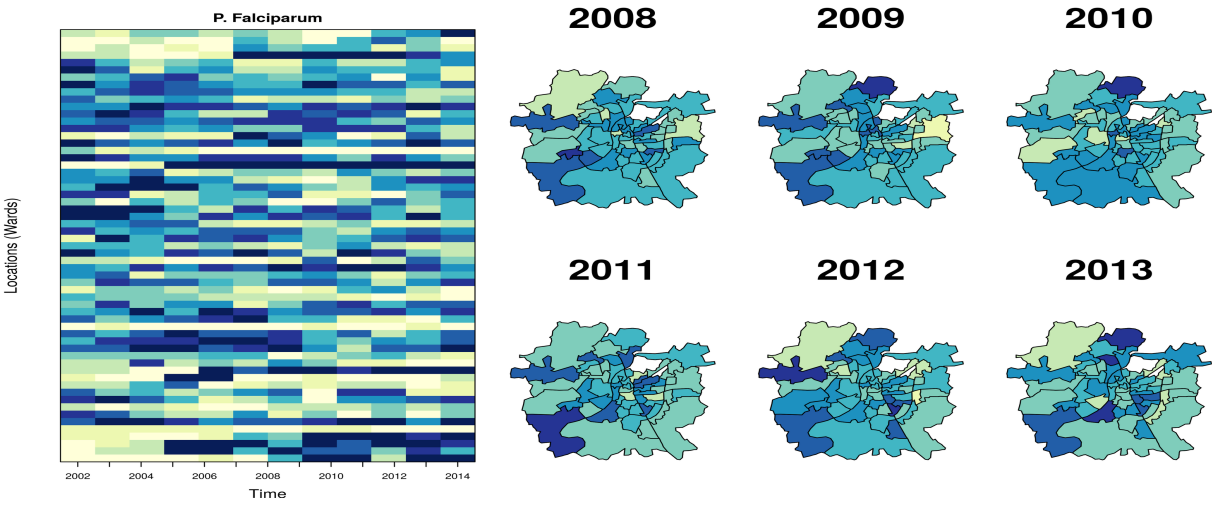

Figure 2.12: **Analysis of spatio-temporal patterns of malaria incidence in the city of Ahmedabad (from low yellow to high blue)**. The panels show the distribution of the cases normalized by population, with the intensity of the color corresponding to the ranking of incidence.

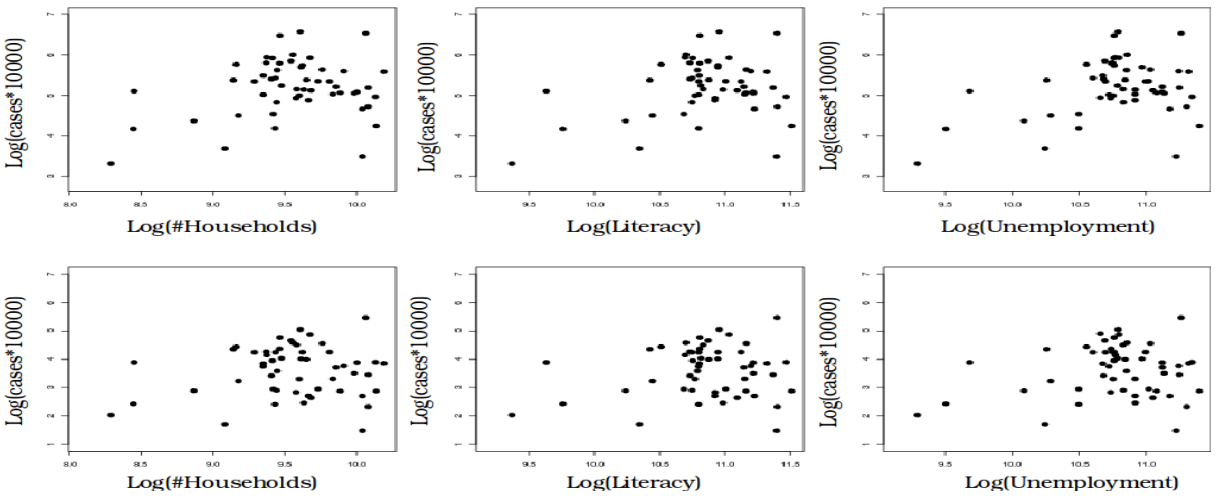

Figure 2.13: Scatterplot of socioeconomic indicators (number of households, literacy and unemployment) and the number of cases of *P. falciparum* and *P. vivax* per ward.

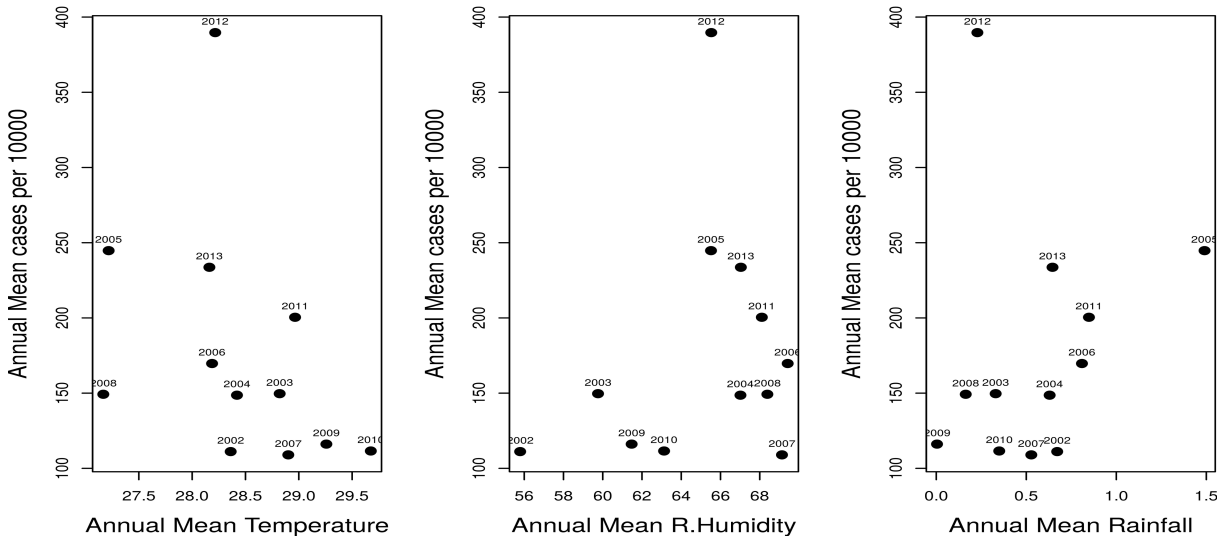

Figure 2.14: From left to right, scatterplots of the annual mean temperature, humidity and rainfall vs annual *Plasmodium vivax* incidence.

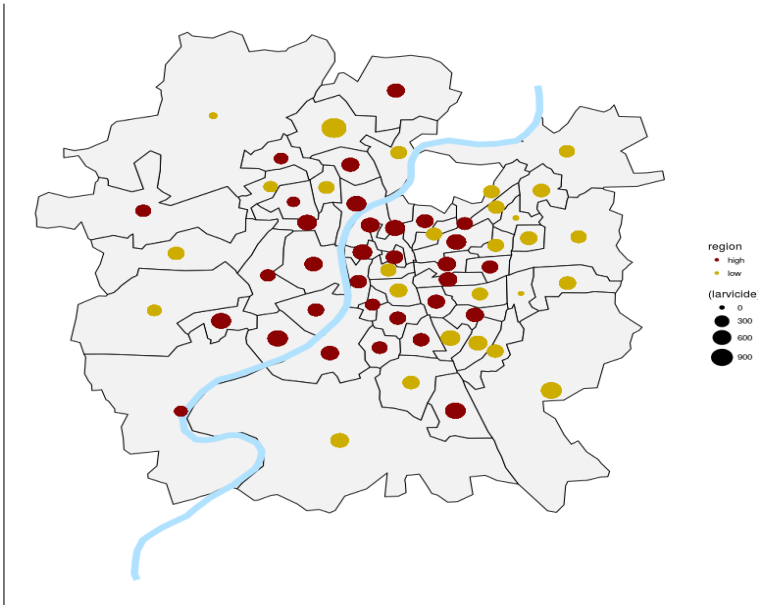

Figure 2.15: **The map depicts the mean number of containers treated with larvicide throughout the city.** Red dots represent wards in the high risk regions, and yellow dots, those in the low risk region.

Table 2.4: **Comparisons of nested models based on the likelihood ratio test** (where asterisks indicate significance relative to the previous model). *** significant at 95% ** significant at 99%

Model	HRR (High risk region)	LRR(low risk region)
Full (autoregressive)	–	–
M1 (autoregressive + temperature) vs Full	***	***
M2 (Autoregressive + temperature + rain) vs M1	–	–
M3 (Autoregressive + temperature + RH) vs M1	**	–
M4 (Autoregressive + temperature + RH + NDVI) vs M3	–	–
M5 (Autoregresive + temperature + RH + cases in the previous year) vs M3	–	–

Table 2.5: Show comparisons of the zero inflated Poisson and negative binomial models fitted to the data.

test	**Zero inflated**	**Negative binomial**
T-test (likelihood ratio test)		***
AIC	749.1740	399.2133
BIC	785.5930	439.2742
Log lik	-364.5870	-188.6066

Table 2.6: Different parameterizations of the probabilistic model. The table shows the effects included in the model formulation and the corresponding total number of parameters.

Symbol	**effect**	**Different effect for the two groups**	**Number of parameters (# groups)**
P	Base	Yes	18
β	Neighbors	Yes	6(2)
α	Seasonality	Yes	6(2)
Se	Seasonality	Yes	12(1)
Temp	Climate variability	Yes	6 (2)
RH	Climate variability	Yes	6 (2)

Table 2.7: Model comparisons highlight the best model which incorporates the random effects, as well as the effect of temperature and humidity.

model	**DIC**
Fixed effects model	
Temperature	5645
Temperature + Relative humidity	5631
Temperature + Relative humidity + rainfall	5646
Fixed effect + Random effect	
Temperature	5628
Temperature + Relative humidity	5622
Temperature + rainfall + humidity	5629

Table 2.8: **Estimated parameters for the best model** which includes the effects of temperature and relative humidity, and the random effects, whose values are significantly different from zero. Credible intervals (CI) from posterior distributions (from two chains that are well mixed and have converged).

covariate	**median**	**95% CI**
Temperature	0.1400	[0.117, 0.280]
Relative humidity	0.0532	[0.0245,0.0712]
Spatial structured parameter	0.2485	[0.1125,0.3578]
Overdispersion parameter	1.5195	[1.025,2.036]

Table 2.9: Statistical analysis of differences between the two regions identified using the socioeconomic information of the 2001 census for *P. falciparum*.

variable	**statistic.t**	**parameter.df**	**p.value**	**log(mean of High risk)**	**log(mean of low risk)**
Slum density	2.4701	47.6618	0.0171	3.3136	2.6065
Unemployment	3.0032	24.2349	0.0061	10.8310	8.5201
Marginal workers	2.8050	25.3346	0.0095	6.8985	5.4690
Literacy	2.3591	32.5098	0.0245	10.5786	8.5977
pop below 6years	2.4111	32.5914	0.0217	8.8374	7.1430
Total population	2.3596	32.4772	0.0245	10.8835	8.8465
Area	1.5635	49.3197	0.1243	1.5391	1.2545
Number of Households	2.3569	32.5793	0.0246	9.3066	7.5628
Vulnerable communities	4.4633	29.4517	0.0001	6.3893	4.3684
Economically deprive communities	4.3176	26.7627	0.0002	8.8945	6.2612

Table 2.10: Results of the best negative binomial model selected by AIC for *Plasmodium falciparum* with climate covariates. This model includes an autoregressive term that accounts for serial autocorrelation, the effect of relative humidity in both regions, and the effect of temperature in the low risk region.

			Low risk region			
	Estimate	**Std. Error**	**z value**	**Pr(>—z—)**	**2.50%**	**97.50%**
ar1	0.6295	0.0701	8.985	0	0.4658	0.7437
intercept	-1.5647	1.438	-1.0881	0.2766	-1.6098	0.2194
temp	-0.0244	0.0446	-0.5461	0.585	-0.0283	0.0271
RH	0.0297	0.015	1.9855	0.0471	-0.0023	0.0179
			High risk region			
	Estimate	**Std. Error**	**z value**	**Pr(>—z—)**	**2.50%**	**97.50%**
intercept	-0.695	0.467	-1.49	0.136	-1.61	0.219
ar1	0.605	0.071	8.533	0	0.466	0.744
RH	0.037	0.016	2.354	0.019	0.006	0.068

Table 2.11: Likelihood comparison of the models showing the covariates included in each model. The best likelihood is obtained for the model that incorporates seasonality, temperature, two regions and the effect of neighbors. The last column shows the likelihood ratio test between the null model (model 1) and each of the other models.

models	**seasonality**	**Neighbors**	**Temperature**	**RH**	**Regions**	**Log lik**	**DF**	δ **AIC**	**LRT**
Model 1	+					-5094.041	27		–
Model 2	+	+				-5073.091	33	29.9	
Model 3	+	+	+			-5003.043	36	134.09	*
Model 4	+	+	+	+		-4986.389	39	27.31	*
Model 8	+				+	-4615.592	54	711.596	*
Model 7	+	+			+	-4066.56	66	1074.063	*
Model 6	+	+	+		+	-3272.393	72	1576.334	*
Model 5	+	+	+	+	+	-3078.985	78	374.816	*

CHAPTER 3

SPATIO-TEMPORAL VARIABILITY OF URBAN MALARIA ACROSS SCALES: ROLE OF CLIMATE AND SOCIOECONOMIC CONDITIONS.

3.1 Introduction

Urbanization and environmental change are pivotal forces affecting the ecological landscape of our planet (Dye 2008). These processes encompass changes in multiple concurrent factors that affect the transmission of communicable diseases at different scales (Grimm et al. 2008). Urban areas have become the new dominant ecosystem around the world and are characterized by large spatial heterogeneity (Harpham 2009). Urbanization is a complex phenomenon that has been associated with rapid and strong environmental variation, such as flooding events caused by extreme rainfall or variation in the increase of sensible heat, which in turn can increase temperature by 2-10 C in highly urbanized areas (Vittal et al. 2016, Shepperd et al. 2002). The remarkable expansion of cities has also increased heterogeneity in terms of population density, economy, and infrastructure, creating inequalities in the provision of urban services (Jacobi et al. 2010, Ahern 2013, Pickett et al. 2017, Zhou et al. 2017). These pronounced socio-economic inequalities are evident in the unprecedented scale and growth of vast informal settlements (slums) in low and middle-income cities (Bolay 2006, Mitlin and Satterthwaite 2010). These factors have important consequences for the spatio-temporal population dynamics of vector borne diseases, such as malaria. There is limited understanding of the multiple drivers of these dynamics in heterogeneous city environments, and how climatic and demographic/socio-economic conditions interact across a range of spatial scales (Zhao et al. 2014, Mishra et al. 2015).

Urbanization can alter important ecological and physiological parameters of the mosquito,

which can promote the transmission, spread or emergence of mosquito borne diseases. However, urbanization can also improve infrastructure and environmental health, leading to better health care provision (Alirol et al 2011). Our understanding of the complex relationship between urbanization and the threats posed by vector borne diseases thus remains incomplete (Dye 2008). In particular, studies of urban malaria have focused on Africa where the disease remains a predominantly rural problem and mosquito vectors are themselves adapted to breed in rural environments (Donelly et al. 2005, Hay et al. 2005, Keiser et al 2004, Robert et al. 2003,). Given the lack of suitable breeding sites for the vectors in African cities (Qi et al. 2012), the better access to health care services and an increased ratio of humans to mosquitoes (Hay et al. 2005), a common assumption is that urbanization reduces malaria transmission. Nevertheless, malaria transmission continues to persist in cities, in some cases at even higher levels than in surrounding (rural) areas (De Silva and Marshall 2012), with several factors thought to explain this unexpected outcome, including wealth, population density, and the climate (temperature and humidity).

In contrast to Africa, cities of the Indian subcontinent harbor a truly urban vector, *Anopheles stephensi*, which breeds in various artificial containers within homes and construction sites (Murdock et al. 2014, Cator et al. 2013), causing malaria to become an urban phenomenon. Given the rapid urbanization rate in this region, there is growing interest in understanding how the spatiotemporal structure of urban malaria reflects socio-economic and environmental heterogeneities of large cities (Ahern et al. 2005, Santos-Vega et al. 2016). Although a few studies have described fine-grained spatial patterns in malaria risk in relation to spatial heterogeneity in population size, environmental and economic factors (Santos-Vega et al. 2016), several key drivers remain poorly understood, especially across different spatial resolutions. They include humidity and population density (as distinct from population counts), and temperature at the high end of the spectrum, where a negative effect is expected from a number of experimental physiological dependencies (Cator et al. 2013).

Moreover, despite the increased interest in the role of spatial heterogeneity in the population dynamics of vector borne infectious diseases, mathematical models typically assume homogeneous transmission. Retrospective spatio-temporal data with sufficient resolution and extent to parametrize these models remain rare (Rezza et al. 2007, Smith et al 2004, Acevedo et al. 2007, Cosner et al. 2009).

Climate variability and climate change are expected to impact urban areas in particular ways, and as such are major determinants of global health (Watts et al. 2015). Some authors have argued that at coarser spatial resolutions (e.g. ¿10 km), the effect of climate on urban areas could be negligible (Trusilova et al. 2013). Other authors have argued that even at coarse aggregated scales, the effects of climate change should be evident, not only in urban areas, but also in larger neighboring areas (Georgescu et al. 2014). Finally, some authors claim that in the case of vector borne diseases, mosquitoes would experience environmental variation at fine scales (Potter et al. 2013, Pincebourde et al. 2016) and this variation would interact with local differences in features such as density of housing, housing material, vegetation cover, and distance to water (Afrane et al. 2004, Murdock et al. 2016). Since climatic quantities at coarse scales are averages over large regions, land types and populations, such as those of metropolitan areas, they tend to hide extremes and can lead to spurious correlations with disease risk (Baker 2008, Crane and Daniere 1996). Thus, it remains unclear the optimum level of aggregation at which infectious disease models should be formulated to capture important effects of spatial heterogeneity in climatic factors.

Here, we investigate the spatial distribution of urban malaria risk and the influence of climatic, demographic (population density), and socio-economic drivers in a large city in India. Based on extensive surveillance data and using a Bayesian hierarchical modeling framework we document a largely stationary spatial pattern of malaria. We identify locations within the city that can be aggregated into larger groups and analyze the factors that explain pat-

terns of spatial variation at a medium resolution. The existence of socio-economic household surveys for the city allow the characterization of socioeconomic variation into three major dimensions, reflecting wealth and water access, occupation and movement, and population density. This variation is shown to interact with local temperature and global humidity. We then investigate if changes in spatial scale would modify the association with climate and economic variables and the prediction accuracy of the model, by zooming-in and out to higher and lower spatial resolution. Finally, we provide an evaluation of prediction accuracy for our best statistical model at these different resolutions over a 10-year time period (Lowe et al. 2013) . We then discuss the consequences for control efforts of the documented spatial structure of the disease and its drivers.

3.2 Methods

3.2.1 Study Site and data description

Surat is located on the banks of the Tapi River in the western part of India in the state of Gujarat **(Fig 3.1)**. It is one of the fastest growing Indian cities due to immigration from various parts of Gujarat and other Indian states. In this largely semi-arid state where malaria is seasonally epidemic, Surat reports more than 1300 *Plasmodium falciparum* cases every year, and an even larger number of *Plasmodium vivax* cases. We concentrate our analyses on the former since it is largely the target of control efforts. The current burden of malaria for both parasites has been greatly reduced in the city after the higher levels of the 1990s. Extensive and dedicated control efforts (indoor residual spraying, breeding sites detection or insecticide-impregnated bed nets to prevent transmission) by the SMC keep malaria outbreaks to relatively low levels, but do not eliminate the problem. Malaria exhibits seasonal outbreaks of varying size. The city presents ideal characteristics to investigate urban malaria transmission and their association with climatic and socioeconomic conditions at different

spatial scales, given pronounced environmental and socioeconomic disparities, as well as an established surveillance program.

Surat is located on the banks of the Tapi River in the western part of India in the state of Gujarat **(Fig 3.1)**. It is one of Indias cities with the fastest growth rate due to immigration from various parts of Gujarat and other Indian states. In this largely semi-arid state where malaria is seasonally epidemic, Surat reports more than 1300 *Plasmodium falciparum* cases every year, and an even larger number caused by *Plasmodium vivax*. Although, vivax malaria We concentrate our analyses on the former since it is largely the target of control efforts. The current burden of malaria for both parasites has been greatly reduced in the city after the higher levels of the 1990s. Extensive and dedicated control efforts by the SMC keep malaria outbreaks to relatively low levels, but do not eliminate the problem. Malaria exhibits seasonal outbreaks of varying size. The city presents ideal characteristics to investigate urban malaria transmission and their association with climatic and socioeconomic conditions at different spatial scales, given pronounced environmental and socioeconomic disparities, as well as an established surveillance program.

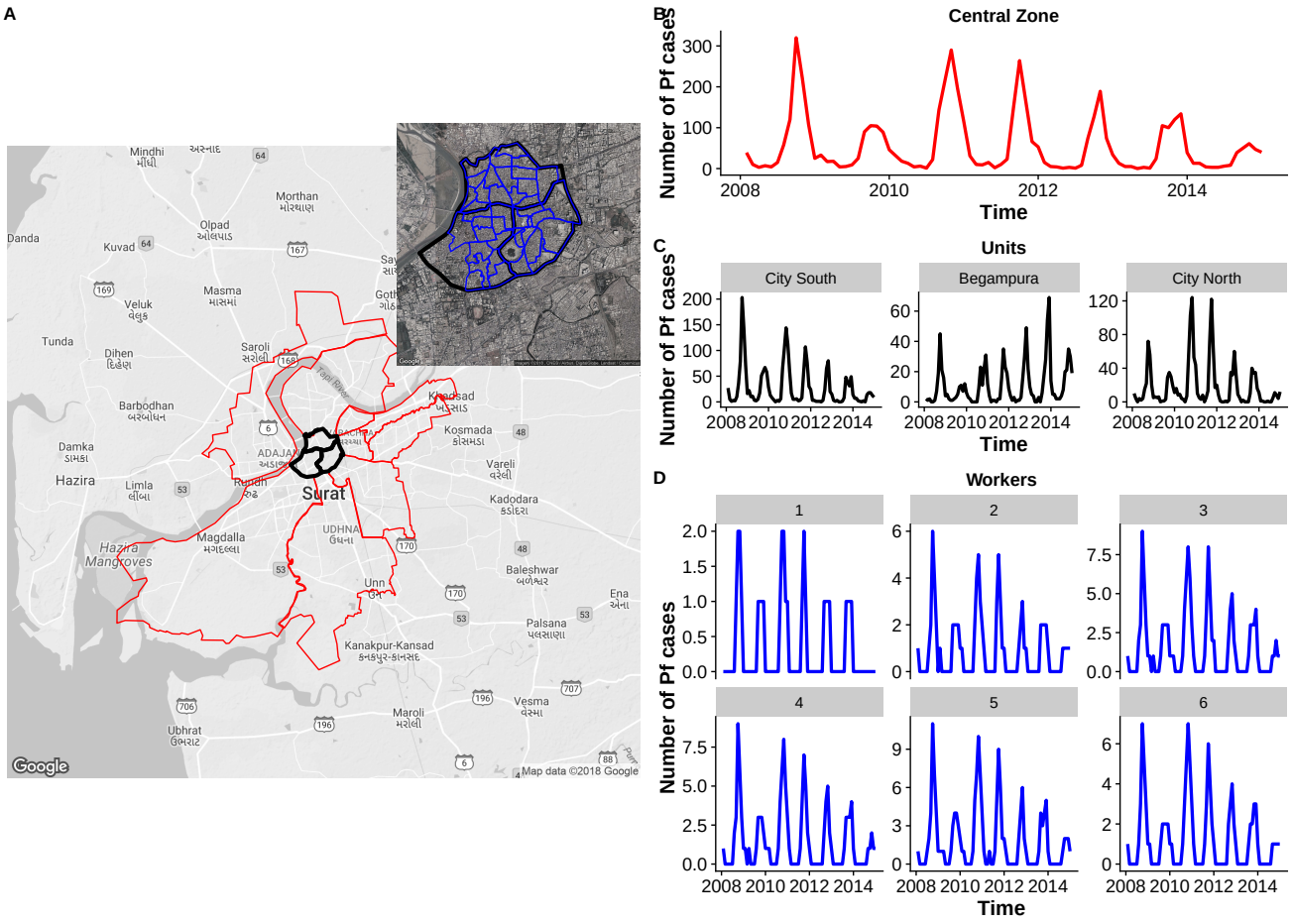

Figure 3.1: Location of the study area (A). The 7 administrative zones of the city are depicted in red, and for comparison some of the units (of a total of 32) are also shown in black for the central zone. Embedded within these, the subplot in the upper right illustrates the smallest reporting scale, that of the worker units (B, C, D). These plots show in corresponding colors the time series for the cases for *P. falciparum* for each of these levels of aggregation.

We collated a multi-sourced spatio-temporal (climatic and socioeconomic) dataset using the statistical computing software R. We reconciled data of different types and aggregation levels (i.e. economic, demographic) to the gridded climate data. The interpolation approaches and description of the generated datasets are explained below.

3.2.2 Climate and socioeconomic data

A database of candidate drivers of urban malaria risk in Surat was generated at the level of every reporting unit (zones, units, and workers) for 2010 using a universal kriging method

to generate an estimated interpolation surface for each covariate **(S1 text 3.5.1)**.For this, we used census data from the District Census Handbook at the district level for the year 2010 from the Directorate of Census Operations, Gujarat. In addition, variables such population density and slum density were calculated using annual population estimates from the Surat Municipal Corporation (SMC). Data on urbanization, housing, water provision, water storage, sanitation, and water scarcity were obtained from water management surveys conducted by Taru Leading Edge, a leading development advisory company in India (http://taru.co.in/). This household survey included 80 locations and 400 households across the city.

3.2.3 Climate data

We obtained gridded (1km x 1km latitude longitude grid) monthly mean surface air temperature data from the MODIS (Moderate Resolution Imaging Spectroradiometer). To estimate surface relative humidity (RH) within the city, we used MODIS data and meteorological parameters obtained from ground-based measurements extracted from GSOD (Global Summary of the Day from NOAA) and 10 meteorological stations located throughout the city that have measured daily temperature, humidity and dew point in the city of Surat since 2014 **(S1 text 3.5.2, Peng et al. 2006)**.

3.2.4 Data analyses

To examine spatial patterns of malaria risk independently from the inter-annual variation of incidence, we accumulated reported cases for a given year and normalized this sum by the total yearly cases for the whole city. We then conducted univariate statistical analyses to evaluate the spatial dependency of the malaria cases **(Fig 3.2 B and C)**. First, a univariate Moran Index (Moran's I) was computed through time (Anselin 1988, Anselin 1996) **(Fig 3.2 C)**. Moran's I identifies the global degree of spatial association (or how much the magnitude of an indicator in one location is influenced by the magnitude of the indicator

in an area close to it) (Anselin 1996, Anselin 2001). Then the Local Moran statistic was obtained to identify local clusters and spatial outliers (Anselin 1995) **(Fig 3.2 B)**. This indicator identifies units with spatial association in malaria incidence. Depending on the sign of the indicator (positive or negative), these local associations can express positive-positive positive-negative, negative-positive or negative-negative associations. Only positive-positive associations are taken into account.

In order to reduce the dimensionality of the socioeconomic variables **(S1 Table 3.5)** and to address the existence of a spatial pattern in these indicators, we used principal component analysis (PCA) to find the best low-dimensional representation of the variation in this multivariate data set. Graphical visualization reveals organization of the variables in a low-dimensional space as well as the spatial organization of the variables **(Fig 3.3 A-B)**. We explored the spatial distribution of the components that account together for more than the 80% of the variance by plotting the spatial distribution of the factor loadings of each component. Then we evaluated the association between the loading values of each PC and the mean malaria rankings **(Fig S1 3.6 B)**. We examined the environmental data for the existence of temporal associations between the interannual variation in Relative Humidity, temperature and malaria rankings **(S1 Fig 3.6 A)**. To explore spatial relationships between the malaria annual cases and the climate factors, we conducted a bivariate Moran scatter plot. This index shows the correlation between one variable (malaria incidence) at a location, and a different variable (temperature and humidity) at neighboring locations **(S1 Table 3.5)**.

3.2.5 Statistical models

A hierarchical model framework was applied to assess the relative contribution of climatic and socioeconomic factors to spatiotemporal urban malaria cases. Generalized linear mixed models (GLMM) were formulated, including random effects to account for unobserved confounding factors, such as additional spatial heterogeneity, quality of health care services, and

local health interventions. A negative binomial model was used to allow for overdispersion found in the urban malaria count data. We also incorporated: (i) a spatially unstructured random effect to introduce an extra source of variability/overdispersion in space (a latent effect), and (ii) a spatially structured random effect to explicitly account for spatial autocorrelations and weight relative risk in a region according to the relative risks in its neighborhood. This is consistent with the effect of increased infectious disease risk from neighboring regions of high transmission introduced in both mathematical (Longini et al. 1988, Viboud et al. 2006, Bertuzzo et al. 2011) and statistical models (Lowe et al. 2013, Lowe et al. 2016).

The general form of the GLMM is as follows:

$$Y_{it} \sim NegBin(\mu_{it}, \sigma) \tag{3.1}$$

$$log(\mu_{it}) = log(p_{it}) + \alpha + \sum_{j}^{2} \beta_j X_{ijt} + \sum_{j}^{3} \gamma_j Y_{ij} + \omega Z_{t-1} + v_i + \phi_i \tag{3.2}$$

Where μ_{it} are the malaria counts for each ward $i = 1$ to 32 and time t= 1 to 132, where the population p_{it} is treated as an offset. The variables X_{it} represent the selected climate influences: temperature one month earlier (j=1) and humidity two months earlier (j=2) and β_j their coefficient. The variables Y_{ji} are the components PC1 (j=1), PC2 (j=2) identified by the PCA analysis, described above, and population density (j=3). Z_{t1} represents a first order autoregressive term to account for temporal autocorrelation in the data. An independent diffuse Gaussian exchangeable prior was assumed for the unstructured random effect (ϕ_{it}), as well as v_i, a normal conditional autoregressive (CAR) prior distribution for the spatially structured random effects, v_i:

$$\nu_i \mid \nu_j = N\left(\frac{\sum_j a_{ij}\nu_{ij}}{\sum_j a_{ij}}, \frac{\sigma^2}{\sum_j a_{ij}}\right) \tag{3.3}$$

where σ^2 controls the strength of local spatial dependence, and a_{ij} are neighborhood weights for each unit as defined above, with simple binary values of 1 when unit i is a neighbor of unit j, and 0 otherwise. Since the CAR distribution is improper, we applied a 'sum to zero' constraint on each v_j.We fitted the Bayesian model via MCMC sampling implemented in R in conjunction with the OpenBUGS software. We generated two parallel MCMC chains, each of length 25,000 with a burn-in of 20,000 and a thinning of 10, to obtain 1000 samples from the joint posterior distribution. Convergence was assessed by inspecting plots of traces of simulations for individual parameters and monitoring the Gelman-Rubin diagnostic (Geman and Rubin 1992). We standardized the fixed explanatory variables for humidity, temperature, and population density to zero mean and unit variance, helping MCMC convergence. The goodness-of-fit of all models was assessed using different quantities, the deviance information criterion (DIC) (Spiegelhalter et al. 2002), the Watanabe-Akaike information criterion (Watanabe, 2010) and an R^2_{LR} statistic for mixed effects models, based on a likelihood ratio (LR) test between the candidate model and an intercept only (null) model (Kramer 2005; Magee 1990). Smaller values of DIC indicate a better-fitting model, whereas R^2_{LR} (for R squared based on Likelihood Ratio) is between 0 and1 the latest corresponding to a perfect fit for any reasonable model specification (Lowe et al. 2015).

3.2.6 Scale dependency

Once the best model was identified, we tested the extent to which the association and significance of the coefficients in the model varied between different levels of spatial resolution. To this end, we fitted the previously identified best model at both a higher spatial resolution (478 worker units) and a lower resolution (7 zones). Then we checked the convergence of the individual parameter estimates and calculated the potential scale reduction (see (Gelman et al. 2004) and note that values below 1.1 are considered to be acceptable in most cases). We also evaluated changes in the 95% credible interval (CI) to see if the interval contained

zero. If the CI does not contain zero, the covariates contribute significantly to the model fit **(Table 3.3)**.

3.2.7 Model comparisons and prediction

To assess the predictive ability of the best model, posterior predictive distributions of malaria incidence were obtained for each unit and month. New pseudo-observations were simulated by drawing random values from a negative binomial distribution with mean and scale parameter estimated using 10.000 samples from the posterior distribution of the parameters in the model and computing the median cases from the simulations. To summarize this information, the observed and posterior predictive mean malaria risk estimates were aggregated across space, and predictions for each unit in high and low incidence years were generated **(Fig 3.4 A-B)**. For the best model, we compared the temporal evolution of the fitted posterior median cases with the observed cases for Surat as a whole **(S1 Fig 3.7)**.

We also compared the proportion of places in which the model accurately predicts malaria incidence. For this, we classified the cases on quantiles generated by considering all zero cases in one class and by subdividing all remaining non-zero cases into four equally sized intervals. The resulting categories correspond to no cases, very low, low, high and very high cases. We then mapped the quantiles of the spatial predictions and observed cases for 2008 **(Fig 4 C-D)** and 2011 (Fig 4 E-F), quantifying the number of times the values correctly matched. We also compared predictive accuracy across spatial scales (zones, units, workers) by calculating the root mean squared error (RMSE), a measure of the difference between modeled and observed values, over the 10-year time period and for each reporting unit (Lowe et al. 2013). Smaller values of RMSE indicate a better fitting model. Regions with positive values (RMSE null - RMS alternate ¿ 0) indicate that modeling the system at the new scale locally improves the estimation of malaria relative risk. In other words, the aggregation or disaggregation of

the system results in a smaller difference between predictions and observations than when using the model at the level of the units.

3.3 Results

The pattern that results from ranking malaria risk based on incidence is largely stationary in time, as shown in **Fig 3.2** and **S1 Fig 3.8**, with locations of high or low risk persisting over time independently from the inter-annual variation in total malaria cases. This temporal regularity of the spatial pattern suggests the existence of strong underlying determinants that are themselves largely stationary over the temporal scales of malaria variation considered here **(Fig 3.2A and S1 Fig 3.8)**. This stationary pattern is supported by a significant spatial autocorrelation through time **(Fig 3.2C)**. Results from the local indicators of the spatial association test (LISA) show that the units spatially associated with high malaria risk differ significantly between the center of the city and its periphery **(Fig 3.2 B)**. Specifically, units in the center and north part of the city exhibit (positive/positive) associations and the periphery of the city shows some units with a negative association in malaria risk. These clusters identified with spatial associations by these statistics also show a striking regularity from season to season **(Fig 3.2 B)**.

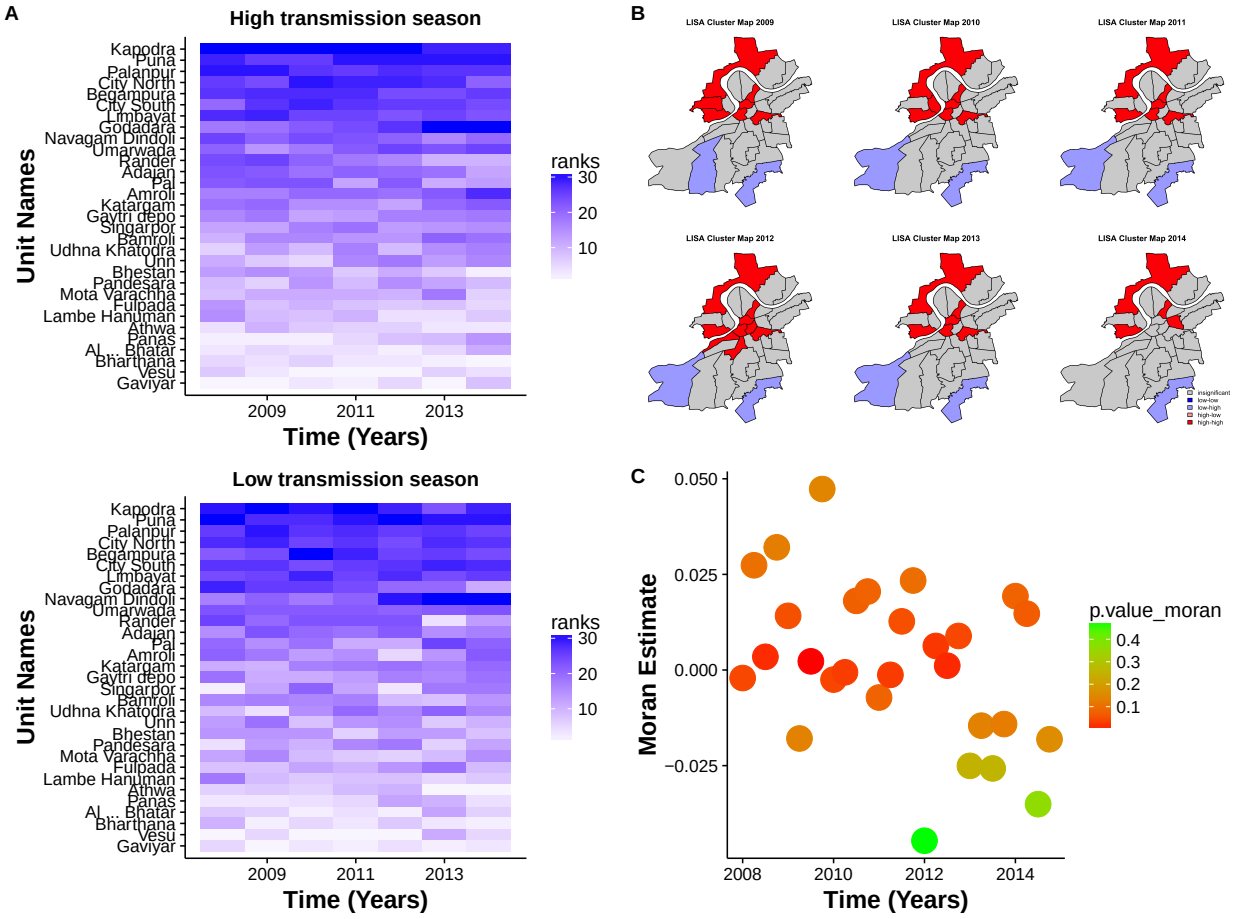

Figure 3.2: **Spatial patterns of disease.**(A) Distribution of cases normalized by population, with color intensity (from low [white] to high [blue]) corresponding to the ranking of incidence. (B) Clusters in malaria incidence for the period 2008-14, identified using the local indicators of spatial associations (LISA) test/analysis. LISA cluster maps are shown, depicting the locations of significant Local Morans I statistics, classified by type of spatial association, with red corresponding to positive/positive associations (high incidence surrounded by high incidence places). Clusters are significant at p = 0.05 (after carrying out 9999 permutations). (C) Moran's I autocorrelation index is calculated bimonthly from 2008-2014. Red dots represent significant values ($p < 0.05$) of the autocorrelation.

With the socioeconomic data from census and surveys in the city **(S1 table 3.5)**, we first addressed whether we could simplify the variation in a low dimensional space. The results from PCA showed that just the first three dimensions explained 80.9% of the variability among units. Dimensions 1, 2 and 3 respectively explained 45.1%, 27% and 3, 8.84% of the variance **(Fig. 3.3 A)**. These three dimensions were therefore retained for further analysis. Fig 2 A shows the distribution of units along the first three PCA axes (PC1 vs PC2 in the upper panel and PC2 vs PC3 in the lower panel), with greater dispersion along the horizontal than vertical axes. In the PCA bi-plot, the origin represents the average unit and the dispersion around it indicates how units differ in relation to this average. The contributions of each of the variables to the three components in the analysis are given in **(S1 Fig 3.9)**. These contributions show that PC1 largely represents economic level and is also correlated with the amount of water stored. PC2 is associated with labor and employment, likely representing the effect of movement and exposure in particular environments within the city, and PC3 exhibits a strong contribution from population density. Fig 2 B shows the spatial distribution for PC1 and PC3. The spatial pattern of the economic level and population density summarized by the PC closely matches that of disease risk **(Fig 3.2 B and S1 Fig 3.8)**.

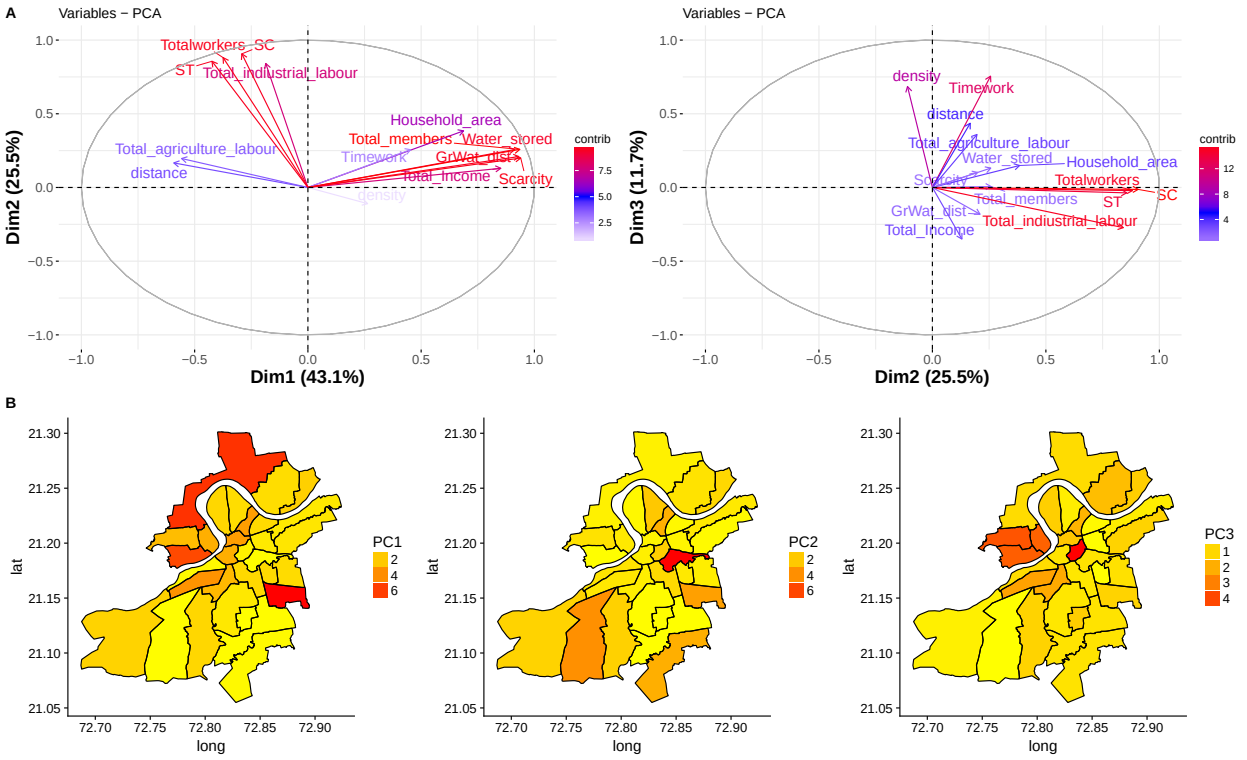

Figure 3.3: **Socioeconomic and climate covariates analyses.** A-B) Distribution of the units in the first-second and second-third dimensions of the PCA. The colors from light blue to red show the contributions of the variables to the principal components; (B) Maps show the PC1-PC3 loadings spatial distribution

We then tested if the spatially stationary pattern in malaria risk was associated with variation in environmental factors (RH and temperature) and/or in socioeconomic and demographic indicators, including population density, income level, water scarcity, amount of water stored, household size, and literacy, as summarized in the three main components of the PCA axes **(Fig 3.3)**. **S1 Fig 3.6 A** illustrates the temporal associations between ranked malaria risk and mean temperature and humidity, respectively. A statistically significant linear correlation was found only between malaria risk and humidity (p ¡ 0.01). **S1 Fig 3.6 B** shows spatial correlations between the three main principal components and the mean ranked cases. The results also detect a significant positive correlation (p¡0.05) of economy/water management and population density.

Goodness-of-fit metrics for models of increasing complexity are shown in **(Table 3.1)**. Comparisons between different models that account for, or neglect, the effect of climate, economy, demography, and neighborhood structure show that the model that best accounts for spatiotemporal variation in malaria risk/incidence based on WAIC (Watanabe-Akaike information criterion) is the one including the combined effects of temperature, humidity, population density, economy, and spatial random variations. Specifically, the effects of population density and economic level/water storage practices (represented by PC1) are significant when considered together with structured and unstructured random effects, explaining 56% of the variance based on the R^2 based likelihood test **(Table 3.1)**. Of the 56% explained variance, 28% is explained by temperature and a monthly autoregressive term of order one, and an additional 10% when considering relative humidity. To test if the climate variables act at a regional or local level, we fitted models with either spatially resolved humidity and temperature data, or temperature and humidity averaged over the spatial units. The model that relies on aggregated humidity and spatially explicit temperature is the one with the best fit. These findings suggest that humidity mostly contributes to explain the temporal variation, whereas temperature strongly influences spatial variability. The inclusion of population density and economic factors added 5% to the explained variance, while random effects added an extra 12%.

Table 3.1: **Comparison of goodness of fit** for the different models tested based on the DIC, WAIC and R^2 based on a likelihood ratio test. Global humidity or temperature consisted in a covariate averaged over the units and spatial humidity/temperature is a value of the covariate varying among units (Smaller DIC and WAIC values signify better fit.)

Model	DIC	R^2_{LR}	WAIC
fixed effect only climate factors			
(A) Autoregressive + Spatial temperature	4937	0.28	821.63
(B) Autoregressive + Global temperature + global humidity	4934	0.32	818.32
(C) Autoregressive + Spatial temperature + global humidity	4927	0.38	815.85
(D) Autoregressive + Temperature + spatial humidity	4932	0.34	816.43
fixed effect climate and socioeconomic			
(A) Autoregressive + climate factors + PC1	4919	0.41	811.22
(B) Autoregressive + climate factors + PC1 + PC2	4920	0.40	812.41
(C) Autoregressive + climate factors + PC1 + population density	4917	0.43	808.52
fixed effect, random effect			
(A) Autoregressive + climate + PC1	4902	0.51	805.17
(B) Autoregressive + climate + PC1+ population density	4896	0.56	802.76

Table 3.2 summarizes the posterior mean parameter estimates for the best model. All parameters in the best model are significantly different from zero, with posterior distributions from the two chains well mixed and converged based on the Gelman-Rubin diagnostic **(Table 3.2)**. Note that the overdispersion parameter of the negative binomial (i.e. the reciprocal of the scale parameter) has a posterior mean value of 2.5190 with a 95% credible interval (CI) of [1.456,3.243]. Thus, the estimated overdispersion parameter (κ) is significantly different from zero (the value corresponding to the Poisson special case of the negative binomial). Additionally, population density had a positive and statistically significant association with malaria incidence. Importantly, the effect of temperature is negative, consistent with the expected non-monotonic effect of temperature on the reproductive number R_0 for malaria at the high end of the spectrum (Mordecai et al. 2013, Pharman et al. 2010). Increases in temperature above a certain threshold affects mosquito and parasite physiological and demographic parameters and ultimately influence transmission intensity (Mordecai et al. 2013, Pharman et al. 2010). Also, relative humidity shows a significant and important enhancing effect on malaria risk within the city.

Table 3.2: **Parameter estimates for the coefficients** associated to variables included in the selected model (i.e. explaining more variation in urban malaria cases). The CI correspond to the 5% and 95% quantiles of the marginal posterior distribution. Each value of $\hat{R}$ correspond to the Gelman-Rubin diagnostic for chains convergence.

covariate	**mean**	**95% CI**	$\hat{\mathbf{R}}$
Temperature	-0.2001	[-0.121, -0.267]	1.005
Relative Humidity	0.6299	[0.445, 0.823]	1.002
population density	0.1055	[0.034,0.226]	1.001
Socioeconomic (PC1)	0.546	[0.369,0.723]	1.002
spatial unstructured hyperparameter	0.3257	[0.134,0.456]	1.001
spatial structured hyperparameter	0.0026	[0.0012,0.016]	1.001
lagged malaria relative risk	1.4270	[0.9674,2.465]	1.001
overdispersion parameter	2.5190	[1.4567,3.243]	1.001

A comparison of modeled and observed malaria cases is illustrated in **Fig 3.4**. In general, the predicted patterns reflect the observations for individual units **(Fig 3.4 C-F)** and for averages over the units **(Fig 3.4 A-B and S1 Fig 3.8)**. However, the best model tends to over predict the number of cases. The maps show comparisons for observed and predicted cases at the unit level for 2008 (a high incidence year) **(Fig 3.4 C and D)**, and for 2011 (a low incidence year) (C-F). We observe correct quantile predictions for 73% and 66% of the units respectively. Although the GLMM has a tendency to over predict malaria in certain areas, the model is better able to capture instances of very high malaria risk across the southwest part of the city.

A

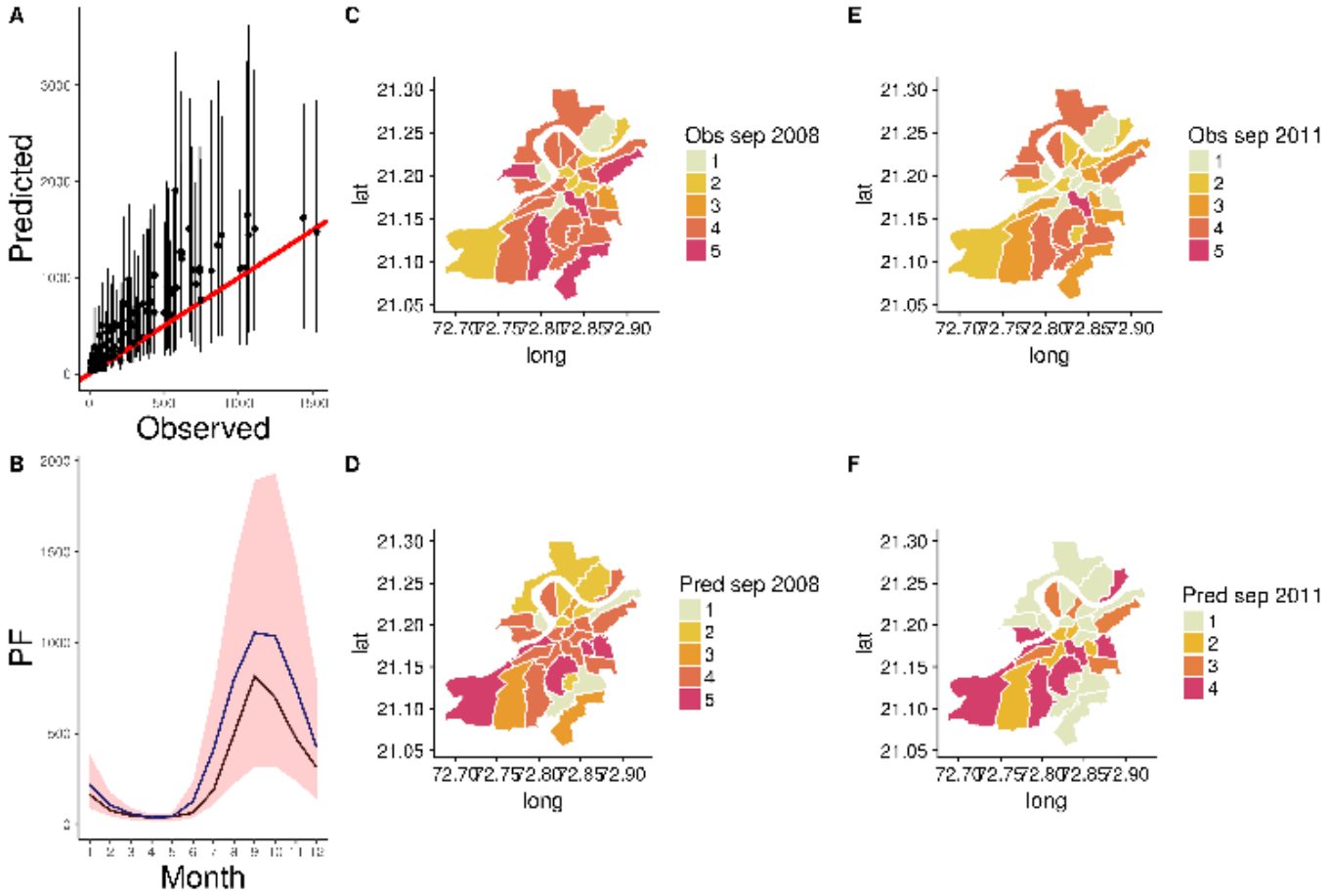

Figure 3.4: **Observed versus predicted *Plasmodium falciparum* cases**. A) Total malaria cases. The identity line (in red) is shown for comparison. B) Seasonal pattern for the observed cases (red) and the median of 1000 model simulations (blue). The 5% and 95% percentiles of the simulated data are shaded in light red. Comparison of observed and predicted cases for 2008 (C-D) and 2011 (E-F). The colors in the maps progress from yellow to red based on quantiles generated by considering all zero cases in one class, and by subdividing all remaining non-zero cases into four intervals. The resulting 5 categories correspond to no cases, very low, low, high and very high cases. Comparisons of the quantiles in the maps reveal that 73% and 66% of their values correctly match for years 2008 and 2011, respectively.

To compare the best model coefficients at different spatial levels (zones, units, workers), we refitted the unit model for 2008 to 2014 (the period of time for which we also have high resolution data), and then fitted the model at both the aggregated zone level (7 zones) and the worker level (478 worker units). Table 3 shows changes in the significance of the coefficients of the best model **Table 3.3** at the different spatial scales. Note, parameter estimates are considered to be statistically significant if their 95% credible interval did not contain zero.. Temperature and humidity show variation in their contribution through the spatial levels: when we disaggregate the system, the effect of temperature becomes more significant, whereas the trend in humidity loses significance. Also, the effect of the economy and population density become more significant at the highest resolution.

To compare the best model coefficients at different spatial levels (zones, units, workers), we refitted the unit model for 2008 2014 (the period of time for which we also have high resolution data), and then fitted the model at both the aggregated level (zone level) and the worker level (478 units). Results of changes in the significance of the coefficients are depicted in **Table 3.3**. Temperature and humidity show variation in their contribution through the spatial levels: when we disaggregate the system, the effect of temperature becomes more significant, as opposed to the trend in humidity, which loses significance. Also, the effect of the economy and population density become more significant at the highest resolution.

Table 3.3: **Comparison across scales of coefficients significance** fitted at three levels of aggregation (Zones, Units and Workers) (significance level *** 0.01, ** 0.05, * 0.1, - -not significant) the colors represents whether the credible interval include zero (red) or not (blue) and the intensity of the blue (from light to dark) represents the significance (*** 0.01, ** 0.05, * 0.1)

Level of aggregation	Temperature	Humidity	Socioeconomic (PC1)	Population Density	Spatial unstructured hyperparameter	Spatial structured hyperparameter	Lagged malaria relative risk
Zones	–	***	–	*	–	–	*
Units	**	*	*	**	*	*	*
Workers	***	–	**	**	**	***	*

Fig 3.5 shows the added value of modeling the system at different levels of aggregation,

in particular the way the significance of climatic and socioeconomic covariates changes. The model fitted at the unit level (32 units) appears to better capture the spatio-temporal dynamics of urban malaria. A model fitted at the zone level only performs better or adds significant information in 22% of the locations (Fig 5 A), and the model fitted at the workers level provides added value to 33% of the units. When the system is modeled at the zone level, malaria risk is close to observed values for only 29% of the locations. By contrast, when the system is disaggregated at the worker level, this proportion increases to 47%, and it is no larger than 61% when the system is viewed at the unit level.

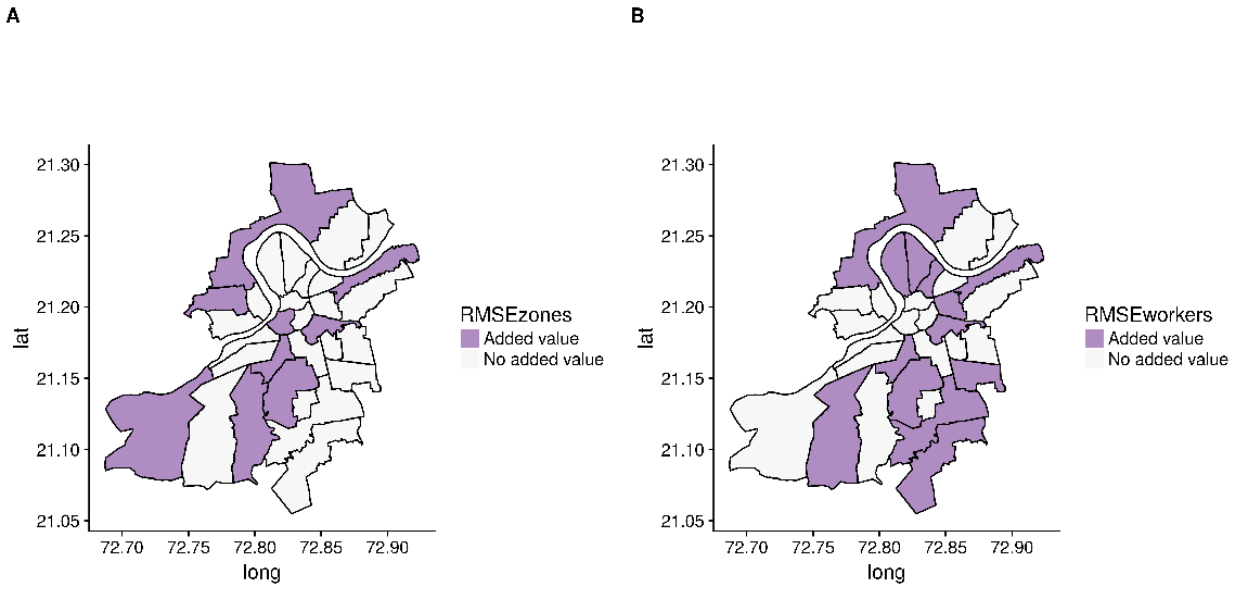

Figure 3.5: **Comparisons across scales**. The lower the RMSE (root-mean-square error) , the better the model fit. Differences between RMSE for the models fitted at the workers and zone level vs. the unit level (A, B). Districts with positive values of (RMSE null - RMSE alternate) (in purple) suggest that disaggregating/aggregating the data improves the model in these areas. Districts with positive values of this quantity (in white) suggest that the new scale does not add information to the model in that area.

3.4 Discussion

At local scales, spatial heterogeneity can significantly influence the risk of infection. In particular, urban environments exhibit pronounced heterogeneity from rapid and unplanned

urbanization (Santos-Vega et al. 2016, Reiner et al. 2012, Perkins et al. 2013). Our results strengthen and expand the results presented by Santos et al. (2016) for a different, inland city of Northwest India, by providing an explicit evaluation of the role of socioeconomic and demographic factors on the stationarity of the spatial distribution of urban malaria at different scales. The model employed in this study estimates the relative contribution of climate, economic covariates, and population density to urban malaria cases, while accounting for unobserved heterogeneity and unknown confounding factors via the inclusion of unstructured and structured random effects. This allows for explicit spatial dependence between units in the city, arising through similar malaria risk patterns, such as in neighboring densely populated urban areas as opposed to sparsely populated units. Conditional intrinsic Gaussian autoregressive (CAR) models are employed in this way to account for spatial dependency in disease mapping (Besag 1991).

Interestingly, our analyses show the presence of three distinct explanatory components in the social and economic covariates. The first component of the PCA is related to economic level and water management, such as income, the quality and size of housing, and the amount of water stored, which can all influence the recruitment of mosquitoes (Sharma 1996). Access to water is an important determinant for malaria in India, given that water is supplied irregularly, leading to water storage within houses, which in turn creates multiple breeding sites for the mosquito in overhead tanks, cisterns and cement tanks (Salje et al. 2016). The second component corresponds to labor and employment, and as such is likely related to human mobility and where people spend time and acquire the disease (Romeo-Aznar et al. 2018). Vector movement is expected to act at very local scales since experiments have demonstrated that mosquitoes do not travel very far and often stay within the same residence for days Pharman 2010).

Our study shows a statistically significant negative relationship between malaria cases

and temperature, which opposes the more commonly documented positive relationship consistent with the effects of physiological and epidemiological parameters of the mosquito, and the malaria parasite within the mosquito, at near optimal conditions (Mordecai et al. 2013). This finding at high temperatures, beyond this optimum, emphasizes the need to better understand the high end of the temperature spectrum on vector-borne infections, the part of the curve least studied (Cator et al. 2013). In addition, population density was identified as an important covariate with a positive effect on malaria risk. Higher population density could result in higher water storage concentrations in close proximity to people. At the local scale, spatial heterogeneity in urban malaria risk could result from varying population density and poverty, as well as environmental microclimates due to proximity of water bodies (Siraj et al. 2015, Arnfield 2003). Since humans are potential generators of vector breeding sites, human behavior influences the relationship between the reproductive number of the disease R0 and host density. The positive effect of population density on malaria incidence found in our model agrees with (Siraj et al. 2015). The authors found that linking vector abundance through mosquito recruitment to human population density in a transmission model can generate an increasing trend in the force of infection. This pattern is opposite to that obtained in standard models with typical assumptions of either no link or a constant ratio between total mosquitoes and human numbers. Interestingly, in our statistical model, the effect of population density is separate from, and does not simply map onto that of, other socio-economic variables.

Our model captures the seasonal pattern and the main trends in interannual variation in malaria cases, by including climatic and socioeconomic covariates, and allowing for spatial dependency between areas. Overall, a model including only autoregressive effects (to account for seasonality) and climate covariates accounted for 29% of the variance in urban malaria risk. Economic covariates and population density explained an additional 15%, and finally spatial dependencies added the biggest increase in the explained variance, underscoring the

importance of taking into account local dependencies.

Humidity and temperature show contrasting effects in our model. A model that incorporates aggregated humidity (averaged for the whole city) and spatially explicit temperature performs better than one with both spatially resolved temperature and humidity. This difference could be explained by the known and strong dependence of humidity on the winds, which can alter evaporation by changing water vapor in the air. By contrast temperature can exhibit large variation within a city at the local level, given the pronounced heterogeneity of impervious surfaces, with differing radiative, thermal, aerodynamic, and moisture properties (Siraj et al. 2015, Arnfield 2003). The two climate covariates further contribute differentially to explain the variation in the data: humidity helps in capturing inter-annual variability and peak timing of outbreaks; temperature explains some of the spatial variation synergistically with economic and demographic covariates. This effect is confirmed when we fit the models at different spatial scales and find an increase in the contribution of temperature at the highest resolution, with humidity losing explanatory power at this level.

Although our spatio-temporal statistical model is able to capture the inter-annual variation in the cases and to predict the peak of epidemics, it could be improved in several directions. In particular, the model could include: (1) mobility fluxes derived from the spatial distribution of the population with movement models, to replace the near-neighbor effects on transition probabilities; (2) the explicit effect of population density on vector abundance ; (3) further analysis of the local effect of other environmental heterogeneities such as river discharge and soil moisture. Moreover, data on temporal changes of the city structure itself would be of interest, including changes in the local speed of urbanization, and their implications for mobility, population distribution, and economic level.

The city has experienced strong malaria interventions in the last three decades reflected

in the known pronounced negative trends in the number of reported cases from the 1980s and 1990s to the 2000s. However, from 2000 onwards, the seasonal prevalence of malaria in Surat has remained fairly stationary. The stationary pattern described here, together with major drivers of this variation, indicate that targeted control could help reduce transmission even further, and that control measure could be implemented ahead of the season based on known spatial heterogeneity. In particular, ongoing efforts to provide better access to water may contribute to reduce malaria transmission, and possibly other vector-borne infections. Although we were unable to separate here the correlated effects of poverty and water access/storage, this is an important area for further study. Ultimately, at longer time scales, a reduction in poverty concomitant with better access to water is fundamental to reduce and eliminate malaria not only within cities but at a more regional level. A deeper understanding of the relationship between humidity or temperature and malaria transmission is key for determining the effect of climate trends on the distribution of vector-borne diseases. The Indian subcontinent humidity and temperature is expected to increase under future climate projections (Edmonson et al. 2016), specifically the northwest part of India will experience a rise in humidity. Under this scenario, an understanding of the relationship between humidity and malaria transmission in urban environments is pivotal to achieving India's target of malaria elimination by 2030.

3.5 Supplementary information

3.5.1 Data description and processing

RH reflects the amount of moisture in the atmosphere. To estimate surface RH in Surat using MODIS data, some meteorological parameters were obtained from ground-based measurements recorded at 10 automatic meteorological stations. MODIS Level-1 data were processed to calculate precipitable water vapor (PW), and MOD07 of MODIS Level-2 atmospheric profile products were used to extract surface air temperature in the clear sky. Data

were downloaded from NASA official website. RH is the ratio of vapor pressure (e) and saturation vapor pressure (es). Vapor pressure (e) depends on air pressure (Pa) and specify humidity (Q), while saturation vapor pressure (es) depends on air temperature (Ta) (Wang 1987). Their relationship can be described by equations (3.4), (3.5) and (3.6):

$$RH = \frac{e}{e_s} \tag{3.4}$$

$$RH = 6.11e^{(17.27\frac{T_a}{237.3+T_a})} \tag{3.5}$$

$$RH = 6.11e^{(17.27\frac{T_a}{237.3+T_a})} \tag{3.6}$$

Since precipitable water vapor (PW) can be retrieved successfully by MODIS data, it is possible to estimate RH using MODIS data. Then specific humidity was calculated with the water vapor following the equation:

$$Q = 0.001(-0.0762PW^2 + 1.753PW + 12.45) \tag{3.7}$$

Then the air pressure was calculated with the following equation:

$$P_a = 1012.30.1038(Alt) \tag{3.8}$$

Once all the parameters such as specific humidity (Q), air pressure (Pa), air temperature (Ta) were obtained, the relative humidity (RH) can be calculated using equations (3.4)(3.6).

3.5.2 Kriging description

Kriging is a technique that generates an estimated interpolation surface from a set of data points (Oliver and Webster 1990). This technique incorporates inference from the spatial

structure of data points to derive estimations at unmeasured locations. Here we used a universal Kriging that incorporates spatial trends into the interpolation process. In mathematical terms, the random function Z(x), representing the variable under study, is split into a deterministic drift m(x) and a random function e(x) with zero mean

$$Z(x_0) = m(x) + e'(x) \tag{3.9}$$

where m(x) is a structural component, associated with a constant mean value or a constant trend and $e'(x)$ a second-order stationary random function spatially correlated component, known as the variation of the regionalized variable. In Universal Kriging, the mean is a function of the site coordinates. Then, m(x) follows the equation:

$$m(x_0) = \sum_{i}^{n} \alpha_i p_i \tag{3.10}$$

where α_i are the local trend or drift coefficients and p_i are functions of the site coordinates. We assume that the drift is a smooth function of the coordinate vector x and that it can be represented by a linear function m(x)=$\alpha_0 + \alpha_i x + \alpha_2 y$ where x and y are the coordinates of the i th control point and α are the drift coefficients.

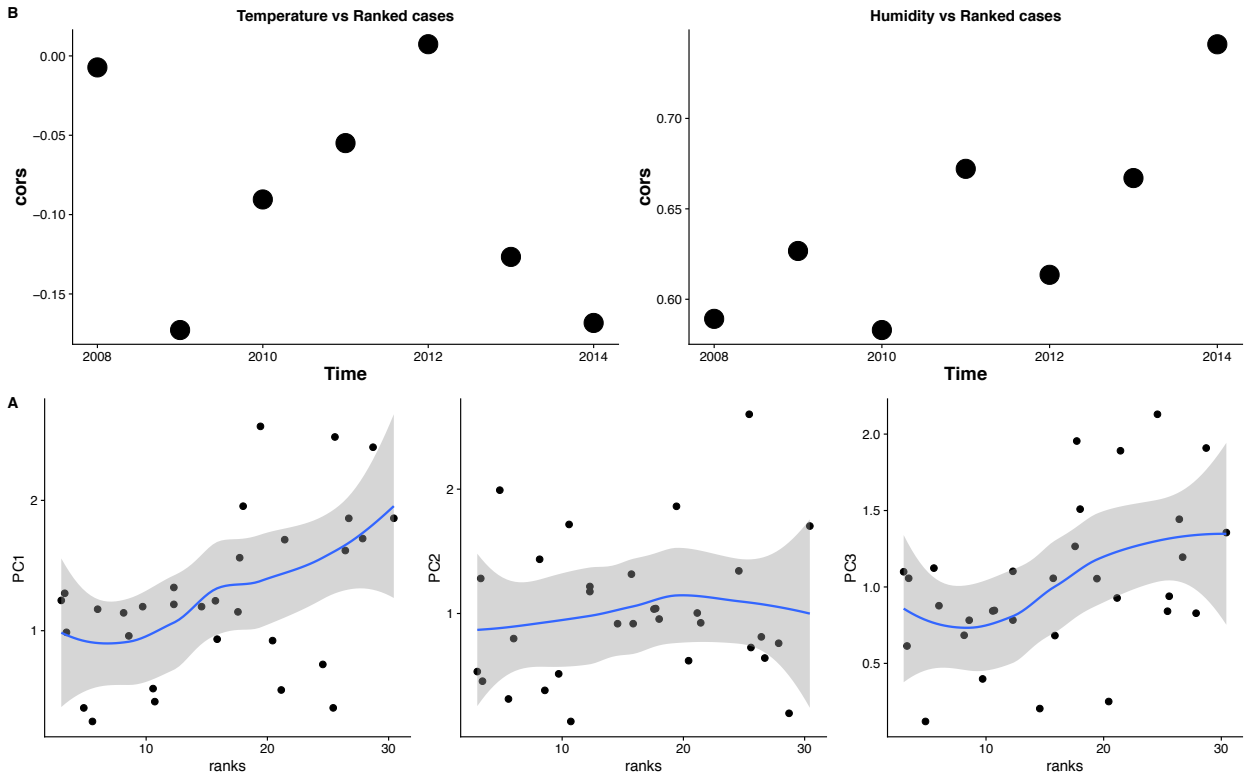

Figure 3.6: **Spatiotemporal correlations with drivers**. A) Top panels show temporal scatter plots of cases and temperature and humidity, respectively. B) Bottom panels show spatial correlations between the cases and the 3 dimensions identified with the PCA in figure 2. A bivariate Moran scatter plot. This index shows the correlation between one variable (malaria incidence) at a location, and a different variable (temperature and humidity) at the neighboring locations (S1 Table 3.6).

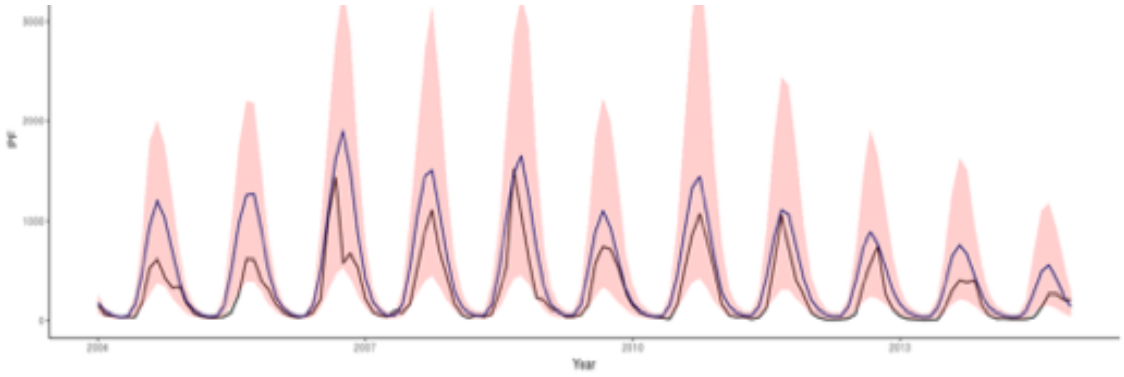

Figure 3.7: **Predictions of cases aggregated for all units through time**. From the 10 000 parameter combinations, the medians of the aggregated cases for a given season are shown together with the 95% credible intervals (CI) (the 5% and 95% quantiles of the simulated cases).

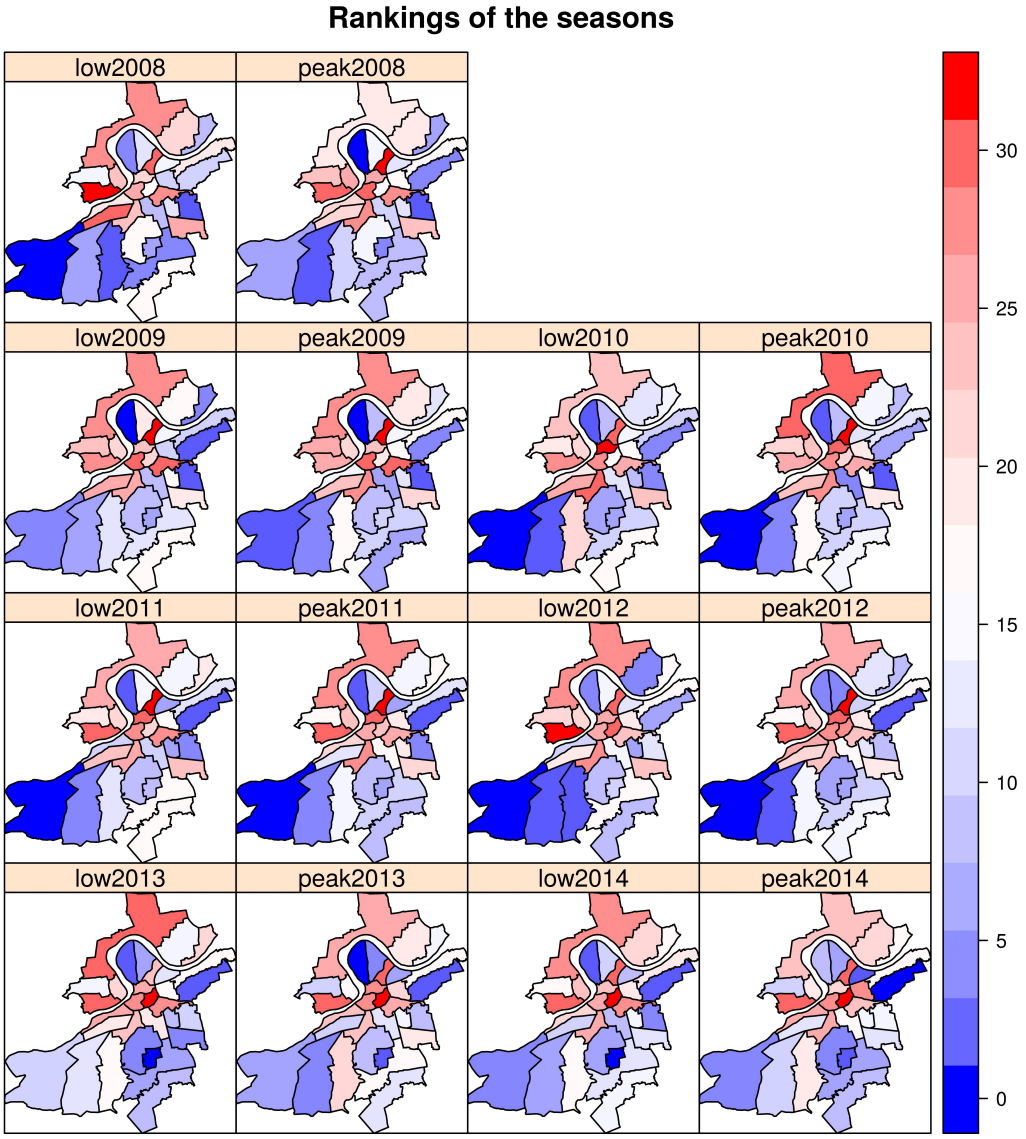

Figure 3.8: The panels show the distribution of malaria rankings based on the incidence (from low for blue to high red for *P. falciparum*).

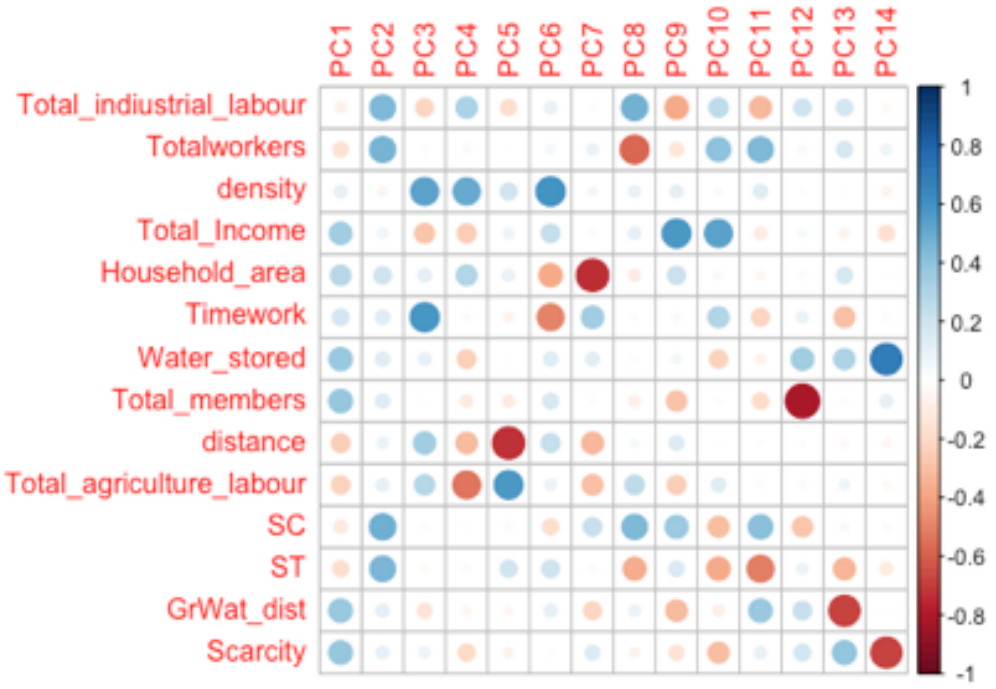

Figure 3.9: **PCA factor loadings** , i.e. the correlation coefficients between the variables (rows) and factors (columns).

Table 3.4: Values of the correlations obtained from the bivariate Moran scatterplot.

Year	**Humidity**	**Temperature**
2008	0.1826	-0.3021
2009	0.0709	-0.2766
2010	0.0964	-0.3337
2011	0.0542	-0.3436
2012	0.0503	-0.4916
2013	0.1407	-0.4059
2014	0.0656	-0.3481

Table 3.5: Socioeconomic covariates included in the analysis

Variable	**units**
Total_indiustrial_labour	persons
Totalworkers	persons
density	persons/area
Total_Income,	Rupees
Household_area	Sq meters
Water_stored	Cubic meters
Total_members	persons
Distance to river	km
Total_agriculture_labour	persons
SC	persons
ST	persons
timework	Hr/day
Distance to water soucers	meters
Scarcity	–

Table 3.6: PCA factor loadings. Which are the correlation coefficients between the variables (rows) and factors (columns)

variable	**PC1**	**PC2**	**PC3**
Total_indiustrial_labour	0.574	19.862	4.528
Totalworkers	2.323	21.777	0.028
density	1.153	0.338	28.574
Total_Income	12.022	0.474	7.470
Household_area	7.831	4.141	1.357
Water_stored	13.817	1.857	1.055
Total_members	14.507	1.873	0.002
distance	5.813	0.779	11.582
Total_agriculture_labour	5.153	1.082	7.885
SC	1.410	23.064	0.011
ST	2.972	20.560	0.081
Timework	3.406	1.842	24.783
GrWat_dist	14.386	1.233	2.004
Scarcity	14.634	1.117	0.638

CHAPTER 4

RELATIVE HUMIDITY EXPLAINS INTERANNUAL VARIABILITY OF MALARIA IN INDIAN CITIES

4.1 Introduction

Many climate sensitive infectious diseases have shown significant interannual variability in the size of seasonal epidemics (Pascual et al. 200, Zhou et al. 2004). Identification of climatic factors shaping interannual and seasonal variability is fundamental to an understanding of the transmission biology of vector-borne diseases. It also provides a basis to formulate early-warning systems for environmentally driven diseases, such as malaria, and to examine their response to both climate variability and climate change (Cazelles et al. 2005, Alonso et al. 2011). Despite many studies of vector-borne infections and climate variability, the role of humidity has been neglected in comparison to variables such as temperature, rainfall and global drivers such as the El Nino Southern Oscillation (ENSO). However, entomological studies on the effects of humidity on vectors (Nagao et al. 2003, Cator et al. 2013, Paaijmans et al. 2010) suggest that this aspect of climate could exert an important influence on temporal patterns of disease outbreaks.

Since climate variability is most important at the edge of malaria's geographical distribution, where climate variables themselves are strong limiting factors (Pascual 2015, Siraj et al. 2014, Laneri et al 2010), large arid/semi-arid regions are of primary relevance to ask about humidity from a dynamical perspective. In addition, urban areas within these regions are potentially of special relevance given their high human density, rapid growth, and low transmission rates which generate epidemic patterns and substantial interannual variation. These areas can act as important reservoirs for transmission in surrounding rural areas and for the persistence of disease regionally. In the Indian subcontinent, a truly urban vector, *Anopheles stephensi*, is present, thriving in the built human environment (Cator et al.

2013), making malaria an urban phenomenon. In e urban environments of semi-arid regions, changes in incidence should be largely determined by seasonal variation in the mosquito population density (Laneri et al. 2010, Roy et al. 2013, Baeza et al. 2014). Thus, identifying how climate influences the population dynamics of the disease in these areas of low or unstable transmission is pivotal to achieving India's plan for elimination by 2030 (Roy et al. 2013).

The presence of an urban vector that does not rely directly on rainfall for recruitment (Metha 1934) raises the central question of whether climate variability drives the interannual variation of the disease in arid regions, as the rainfall-driven models of rural areas do not apply (Laneri et al. 2010, Roy et al. 2013, Baeza et al. 2014). On the basis of experimental results, the population dynamics of *Anopheles stephensi* should be affected by changes in temperature and humidity (Murdock et al. 2014, Bayoh 2001, Mordecai et al. 2013). Specifically, experiments have shown that temperatures in the approximate range of 21-32°C and a relative humidity of at least 60% are the most conducive conditions for maintenance of transmission (Metha 1934). The vectors need to live at least eight days in order to transmit malaria and higher humidity increases both survival rate and activity rate (Snodgrass 1939, Gaaboub et al. 1971).This relationship explains why mosquitoes are most active and prefer feeding during the night when the relative humidity of the environment is higher. When average monthly relative humidity is below 60%, the lifespan of the mosquito would be too short for effective malaria transmission (Kessler et al. 2008, Mayne 1930). Most process-based mathematical models of malaria transmission that explicitly account for the effect of temperature continue to rely on the same few, valuable but limited, early studies of temperature dependence of fundamental demographic parameters for *A. gambiae* (and for the development of *P. falciparum* within this vector), such as rates of parasite sporogony, vector survival and biting rate (Detinova 1962, Bayoh 2003). There is currently a revival of entomological studies to deepen and extend the empirical basis for demographic parameters. It remains unclear however whether these effects manifest themselves at the population level

in the incidence patterns of the disease, especially for humidity and for the high end of the temperature spectrum important for questions on global warming (Murdock et al. 2014), and for humidity in general. The population effects of humidity are also of relevance more generally to other vector-transmitted diseases in urban settings that are emerging today.

The extensive surveillance records for two cities of Northwest India provide an opportunity to address the role of humidity with a combination of mathematical models and statistical inference methods for time series. Our results underscore the importance of climate forcing in the inter-annual variability of urban malaria, primarily through changes in humidity. Disease risk is already considerably higher in the more humid city of Surat than Ahmedabad. We discuss these findings in the context of current and projected future changes in this climate parameter.

4.2 Material and Methods

4.2.1 Data

In Indian cities, cases of falciparum malaria rise after the monsoon rains and peak in October-November. Interestingly, the Indian subcontinent harbors a truly urban mosquito, *Anopheles stephensi*, which thrives in the built-in human environment, with breeding sites in various artificial containers within homes and in water that collects in construction sites [6]. To address the question of whether humidity influences the seasonality and inter-annual variability of urban malaria we have chosen 2 cities (with over 3 million people each) in the state of Gujarat, India (Ahmedabad and Surat). These cities exhibit a growing population where sustained, extensive, and consistent surveillance programs have been conducted for over two decades. Despite their close proximity, these cities exhibit distinct environments. While Ahmedabad is semi-arid, Surat is a coastal city with a maritime influence on its climate is prone to flooding from the Tapi river.

The malaria data consist of monthly cases collected from 1997 to 2014 by the respective Municipal Corporations of the cities of Ahmedabad and Surat **(Figure 4.1 A-B)**. These cities are both located in the semi-arid state of Gujarat in Northwest India. The epidemiological data result from two kinds of surveillance: (a) the collection of blood slides from fever patients by house-to-house visits by a health worker and examination of these slides for positive malaria parasites at the Primary/Community Health Center (active surveillance); (b) examination of blood slides from febrile patients reporting directly to the Primary/Community Health Center (passive surveillance). Both types of data are pooled into a temporal record for each city. Monthly relative humidity and temperature for the same 18 years were obtained from local weather stations (**Figure 4.1)**. Time series for population size in each city were obtained through estimates by the municipal corporations. Monthly humidity data were recorded at a local weather station within each city, and supplied by the Indian Meteorological Department in Pune (India).

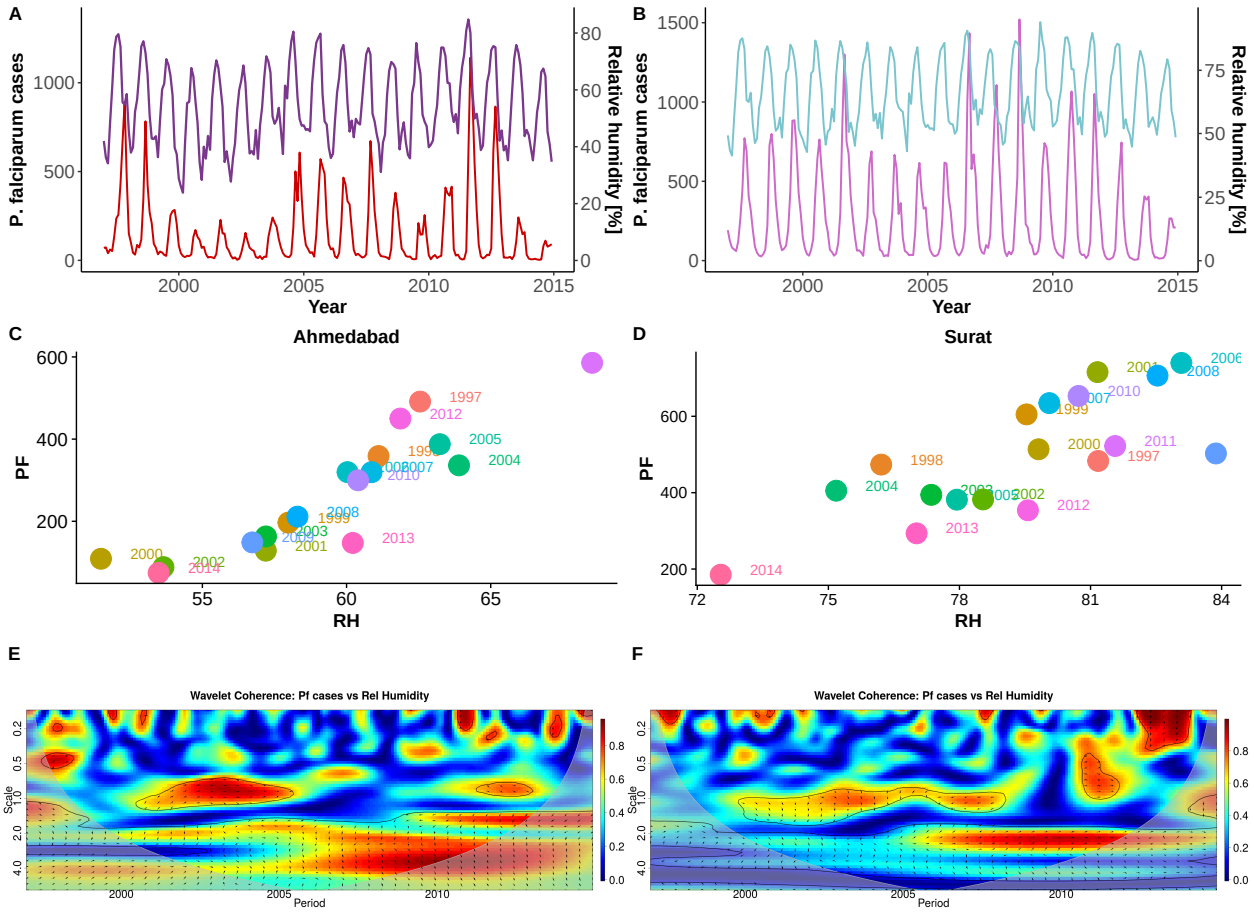

Figure 4.1: **Data description and exploratory analyses**(A-B) Monthly case data for *P. falciparum* (in red) are shown for Ahmedabad (A) and Surat (B) respectively, along with the corresponding relative humidity time series from 1997 to 2014. (C-D) The total cases during the transmission season from August to November are shown as a function of average RH in a critical period from March to July for Surat (D) and May to July for Ahmedabad (C). (E-F)The cross-coherence between the humidity in Ahmedabad (E) and Surat (F) and the monthly malaria cases. Cross-coherence varies between 0 and 1, shown here in a color scale from blue to red. The lines correspond to the 5 and 10% significance levels. Only regions within these lines exhibit significant cross-coherence at those levels. The shaded region corresponds to periods and times that are affected by the boundaries and are therefore considered outside the curve indicating the so-called cone of significance. (The RH time series has been filtered with a low-pass filter to remove seasonality and therefore periods below 1 year; similar patterns are obtained without the filtering with the additional expected coherence at the 1-year scale, results not shown)

4.2.2 Data Analyses

The temporal and possibly transient association of variability at different periods between the times series for malaria and humidity were first examined using cross-wavelet coherence analysis (Cazellez et al. 2007, Chavez et al. 2006). In contrast to Fourier cross-spectrum analysis, wavelet analysis is well suited for the study of signals whose frequency spectra changes in time. This time-frequency decomposition provides information on the variance at different periods along the temporal dimension (Cazellez et al. 2007). Specifically, using wavelet coherency analysis, we can determine whether the presence of a particular frequency at a given time in the number of cases corresponds to the presence of that same frequency at the same time in humidity. As an additional exploratory analysis, we plotted the mean number of cases aggregated over the epidemic months (Aug-Nov) against the relative humidity in a critical period preceding the epidemic. This window was selected based on maximum lagged correlations between this period and the peak case months, which correspond to March to July for Surat and May to July in Ahmedabad.

4.2.3 Transmission model

Using a stochastic transmission model (Fig 4.2) we tested the hypothesis that humidity it is important in driving the temporal dynamics of malaria in cities. We rely on a model that represents the vector dynamics implicitly in terms of its force of infection, λ, and a Gamma-distributed lag time with mean τ to account for the developmental delay of *P. falciparum* parasites within surviving mosquitoes. Since we are specifically interested in the effect of humidity on interannual variation, we simplify the formulation of the model by considering seasonality and humidity-driven interannual variation in the force of infection [10,11]. We couple a chain of classes implementing the distributed delay, to a model for transmission in the human population.

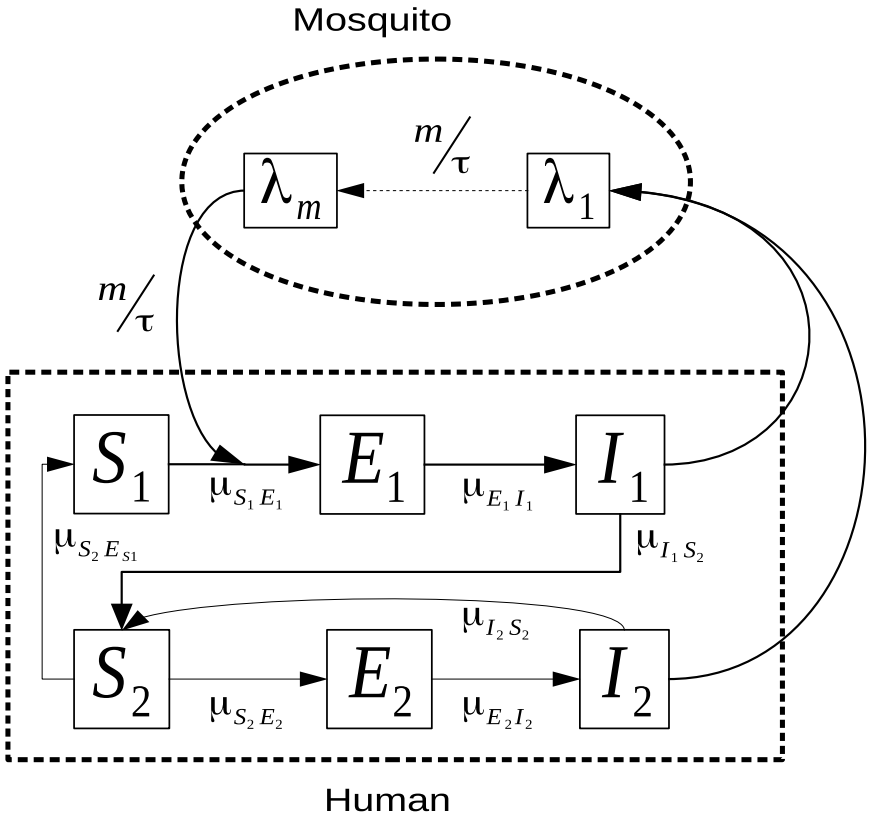

Figure 4.2: **Model description.** In our model of urban malaria, we subdivide the ,total population into two classes of infectious and susceptible individuals to allow for heterogeneity in the degree of clinical symptoms and represent a degree of protection conferred by previous infection: S1, susceptible to infection; E, exposed; I1, infected symptomatic and infectious; I2, infected asymptomatic and infectious; and S2, recovered but with subpatent load of parasites (not completely cleared). Reinfection, which here refers to only recrudescent infections that were formerly subpatent excluding new infections, is included as a transition from S2 to I2. Mosquito parasite classes are λ_1 (force of infection at previous time t- s) with arrows indicating the direction of transition between classes. The possibility of transition between classes X and Y is denoted by a solid arrow, with the corresponding rate written as μ_{XY}

The system of stochastic differential equations is given by :

$$\text{human} \rightarrow \begin{cases} \mathrm{P} = S_1 + E + I_1 + S_2 + I_2 \\ \mathrm{d}S_1/\mathrm{d}t = \left(\delta P + \frac{\mathrm{d}P}{\mathrm{d}t}\right) + \mu_{S2S_1} S_2 - \lambda_H(t) S_1 - \delta S_1, \\ \mathrm{d}E/\mathrm{d}t = \lambda_H(t) S_1 - \mu_{EI_1} E - \delta E, \\ \mathrm{d}I_1/\mathrm{d}t = \mu_{EI_1} E - \mu_{I_1 S_2} I_1 - \delta I_1, \\ \mathrm{d}S_2/\mathrm{d}t = \mu_{I_1 S_2} I_1 + \mu_{I_2 S_2} I_2 - \mu_{S_2 S_2} S_2 - \lambda_H(t) S_2 - \delta S_2, \\ \mathrm{d}I_2/\mathrm{d}t = \lambda_H(t) S_2 - \mu_{I_2 S_2} I_2 - \delta I_2, \end{cases} \tag{4.1}$$

where P represents population size. The flow of newborns combined with the death rate of each class results in population numbers equal to those observed for the overall demographic growth of the city. Furthermore, the force of infection (or rate of transmission per susceptible individual) λ is given by the following expression:

$$\lambda = \left(\frac{I_1(t) + I_2(t)}{P(t)}\right) \beta \tag{4.2}$$

where the transmission rate β includes the effects of seasonality, climate variability and environmental noise as follows,

$$\beta(t) = exp\left[\sum_{k=1}^{6} b_k S_k + b_{RH} S_4 F\right]\left[\frac{d\Gamma}{dt}\right] \tag{4.3}$$

The explicit effect of climate forcing by either relative humidity or temperature is represented in this expression on the interannual time scale. Specifically, seasonality is introduced nonparametrically through the coefficients (b_k) for a basis of periodic b-splines with $S_k(t)$, k= 1..., 6 constructed using n_s evenly spaced knots. We superimpose variability in the transmission rate across years through explicit consideration of interannual variation in temperature or humidity. For this purpose, the effect of these variables was localized in the spline S_4 which corresponds to the time of the year during the monsoons (S1 Fig 4.7). Parameter

C quantifies the strength of the effect. Finally, environmental noise was modeled with a Gamma distribution Γ to represent additional fluctuations absent in the climate covariate [details are provided in Badra et al 2011(6)].

We consider that cases are under-reported and measured with error. The measurement model is implemented with a negative binomial distribution so that: cases $\sim$ NegBin(ρC_i, k_i) with mean ρC_i and overdispersion k_i, where ρ represents the reporting rate and C_i is the symptomatic infected individuals for population i obtained as the sum of the individuals entering classes I_1 and I_2 during an interval of time t (here, monthly) **(Fig 4.2)**.

4.2.4 parameter estimation

We rely on a sequential Monte Carlo method for likelihood maximization to fit this nonlinear stochastic model to data from a partially-observed system. The estimation of parameters and initial conditions for all state variables was carried out with the iterated filtering algorithm known as MIF, for maximum likelihood iterated filtering, implemented in the R package pomp (partially observed Markov processes, (Ionides et al. 2006, king et al. 2015).This methodology has a plug-and-play property (Breto et al. 2009, He et al. 2010), meaning that its flexible application is based on numerical simulation of the model. This enables consideration of a wide class of models in terms of structure.

This algorithm further allows for consideration of both measurement and process noise, in addition to hidden variables, which are typical limitation of surveillance records providing a single observed variable for the incidence. The initial search in parameter space was performed with a grid of 10,000 random parameter combinations, and the output of this search was used as the initial conditions of a more local search. The algorithm itself consists of two loops, with the external loop essentially iterating an internal, filtering loop and in so doing generating a new, improved estimate of the parameter values at each iteration.

The filtering loop implements a selection process for a large number of particles over time. For each time step, a particle can be seen as a simulation characterized by its own set of parameter values. Particles can survive or die as the result of a resampling process, with probabilities determined by their likelihood given the data. From this selection process over the whole extent of the data, a new estimate of the parameters is generated, and from this estimate, a cloud of new particles is reinitialized using a given noise intensity adjusted by a cooling factor. This noise, as well as the stochasticity of the dynamics of the system itself, provides the variability for the selection process of the particles to act upon.

4.2.5 Prediction

To examine the ability of a process-based (mechanistic) model to predict malaria incidence in an urban environment, we compare the total number of malaria cases observed for each city to those predicted by model simulations. Since simulations of the model produce temporal trajectories predicted when one starts from estimated initial conditions, we considered prediction of out-of-fit data not used to fit the model. Thus, monthly cases from January 1997 to December 2008 were used as a training set for parameter estimation , and data from January 2009 to December 2014 was kept as the out-of-fit set to evaluate the predictive ability of the model. We specifically test whether the model can be used as an early-warning system to forecast malaria outbreaks following the monsoon season by using. Observed humidity values in 1997 to 2008 to, predict the number of cases from Jan 2009 to Dec 2014. MIF allows us to estimate the state of the system (in all classes) at the initial month for the prediction.

4.3 Results

Motivated by the need of a deeper understanding of the association between relative humidity and urban malaria transmission, here we took advantage of the extensive surveillance

records for two cities of Northwest India **(Figure 4.1 A-B)** to address the role of humidity with a combination of mathematical models and statistical inference methods for time series data. We begin by characterizing the temporal scales present in the interannual variability of malaria in relation to those of humidity. **Figure 4.1 C-D** shows a strong and significant cross-coherence between malaria cases in both cities (Ahmedabad and Surat) and relative humidity (RH) for the circa-biennial scale. In addition, the coherence between these malaria cases and humidity changes over time, with lower values in the late 1990s and early 2000s **(Figure 4.1 E-F)**. In addition, average RH in the selected critical window of the year (Figure S1) is significantly correlated to total cases of the malaria season (Aug-Nov), with a correlation coefficient of 0.72 for Ahmedabad (p=0.0002) and a coefficient of 0.69 for Surat (p=0.004) **(Figure 4.1 C-D)**.

Then with an stochastic malaria transmission model **(Figure 4.6)**, we asked next whether variation in RH is important to explain the interannual variation in urban malaria cases. Table 1 shows comparisons between the different models that include or not the effect of RH or temperature on malaria transmission. Based on a likelihood ratio test, the model incorporating RH as a covariate performs significantly better than one without this interannual effect or a model with only temperature as a driver, ($p < 0.001$ Table 1).

Table 4.1: Model Comparisons based on likelihood

model	**log likelihood**	**SE**	**# parameters**	**Delta AIC**	**LRT**
		Surat			
With Rh	-1166.9011	0.1254	25	–	<0.001
With Temp	-1176.2280	0.2174	25	-18.6539	
Without RH	-1181.6492	0.1885	24	-27.4962	
		Ahmedabad			
With Rh	-1111.3123	0.3664	25	–	<0.001
With Temp	-1120.9340	0.3210	25	-19.2434	
Without RH	-1123.9829	0.2657	24	-23.3411	

Numerical simulations of the best model including RH capture the observed variation in

the size of seasonal outbreaks for both cities **(Figure 4.3 A-B)**. In particular, the pattern of large outbreaks followed by smaller ones is well captured (Figure 4.3 A-B), as well as the main seasonal pattern of the reported cases in both cities **(Figure 4.3 C-D)**. The similarity is particularly striking given that these simulations are not next-month predictions, but the complete temporal trajectories for the whole seventeen years starting only from estimated initial conditions in 1997. Interestingly, Surat experienced large epidemics in 2001 and 2006, coincident with extremes in RH coincident with flooding events within the city (Figure S3).

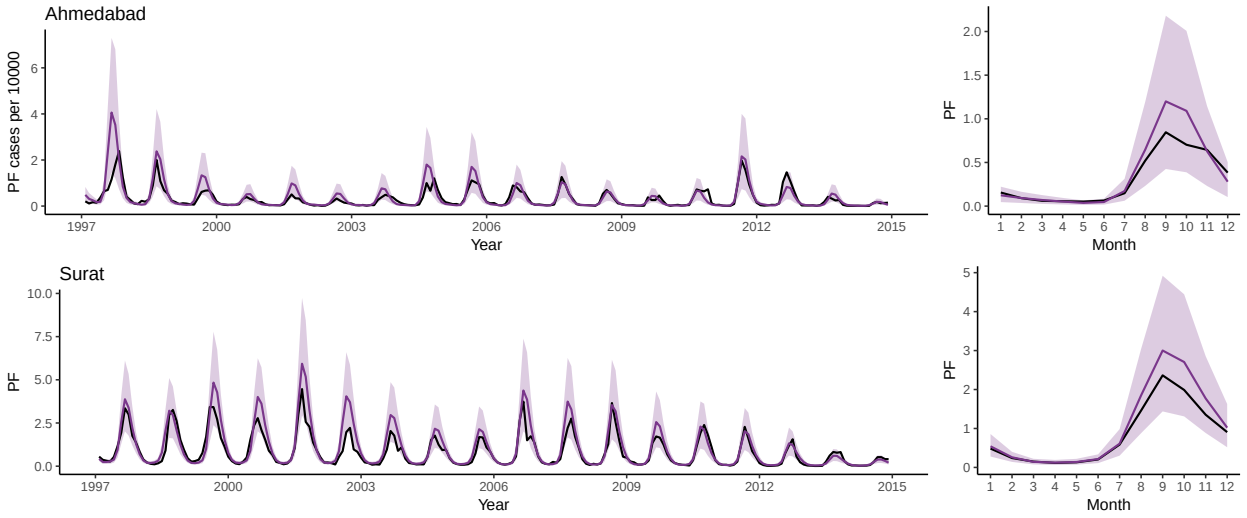

Figure 4.3: **Best model simulation.** Comparison of simulated and reported cases for Ahmedabad and Surat respectively. Time series (A) and seasonality (B) for the observed cases (obs; black) and the mean of 1,000 model simulations (sim; purple). The 10% and 90% percentiles of the simulated data are shaded in light red. The model simulations are not next step predictions but numerical simulations of the model forward for the whole time period of the study starting with estimated initial conditions

Moreover, this model reveals a difference in the force of infection (the instantaneous infection risk to each susceptible individual) between both cities, with larger values in the Surat which on average is more humid than Ahmedabad **(Figure 4.3 A and B)**. In Ahmedabad (a dryer city) the transmission rate during the monsoon season has a much lower average, yet a more pronounced response to humidity. This difference suggests a higher intensity of transmission more humid environments **(Figure 4.4 C)**. Likewise, the differential effect

of humidity is also expressed in the values of the humidity coefficient in transmission b_{RH} which is higher for the dryer city of Ahmedabad **(Figure S1 4.8)**.

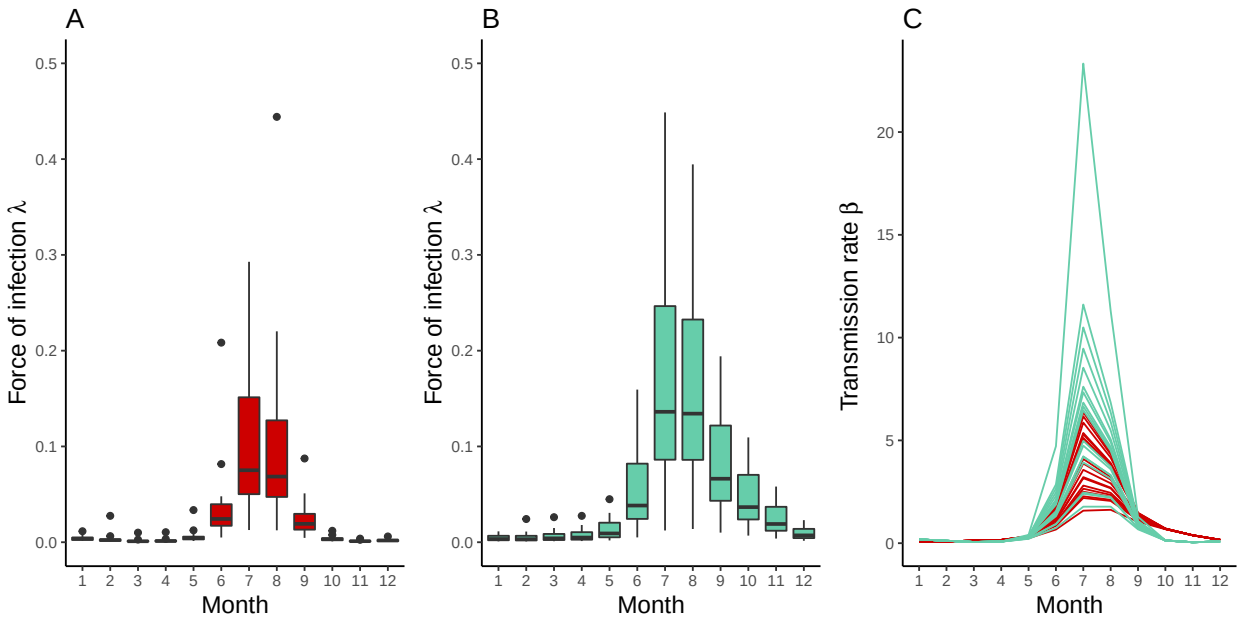

Figure 4.4: **Force of infection and transmission rate** from one simulation for the period 1997-2014. (A and B) Box plot of the force of infection by month per city. The force of infection is defined as the per capita rate at which susceptible individuals become infected. (C) Transmission rate by month per city (Ahmedabad in red; Surat in green)

The maximum likelihood estimates of the parameters reveal additional features of the dynamics (**Table 4.2**). **Table 4.2** summarizes the mean parameter estimates and their confidence intervals. The duration of immunity is very short for both cities and immunity is shorter in the more humid environment. Another parameter of interest is the delay between the latent and current force of infection caused by vector transmission. A delay was estimated of approximately 9 days for Ahmedabad compared to 16 days for Surat in the models with RH **(Table 4.2)**.

Table 4.2: Parameter estimates and confidence intervals

Description	Unit	parameter	Ahmedabad	CI	Surat	CI
Time from exposed to infected	days	$1/\mu_{EI_1}$	24	[21-27]	28	[24-33]
Mean recovery time	days	$1/\mu_{I_1S_2}$	32	[29-37]	38	[32-43]
Mean loss of immunity	days	$1/\mu_{S_2S_1}$	48.36	[46.32-52.25]	38.26	[36.53-40.21]
Recovery time from asymptomatic infection	days	$1/\mu_{I_2S_2}$	23.27	[20.66-26.23]	18.62	[15.36-22.46]
Case reporting fraction		ρ	0.014	[0.09-0.017]	0.015	[0.012-0.019]
Mean lag for mosquitoes	days	τ	9.25	[8.36-11.15]	16.24	[14.57-19.02]
Coef of humidity covariate		b_H	1.744	[1.53-2.39]	3.156	[2.45-4]

The performance of our best model suggests the possibility of forecasting malaria cases as a function of RH in the preceding monsoon season. Model predictability was evaluated for the years between 2009 and 2014 with the out-of-fit data mimicking the use of the model to forecast upcoming seasons. Because the model incorporates process noise, we generated both the median value of malaria cases and the 10-90% quantiles from 1000 predictions for each month. For the most part, the median predictions are close to the observations and reported cases fall within the uncertainty interval.

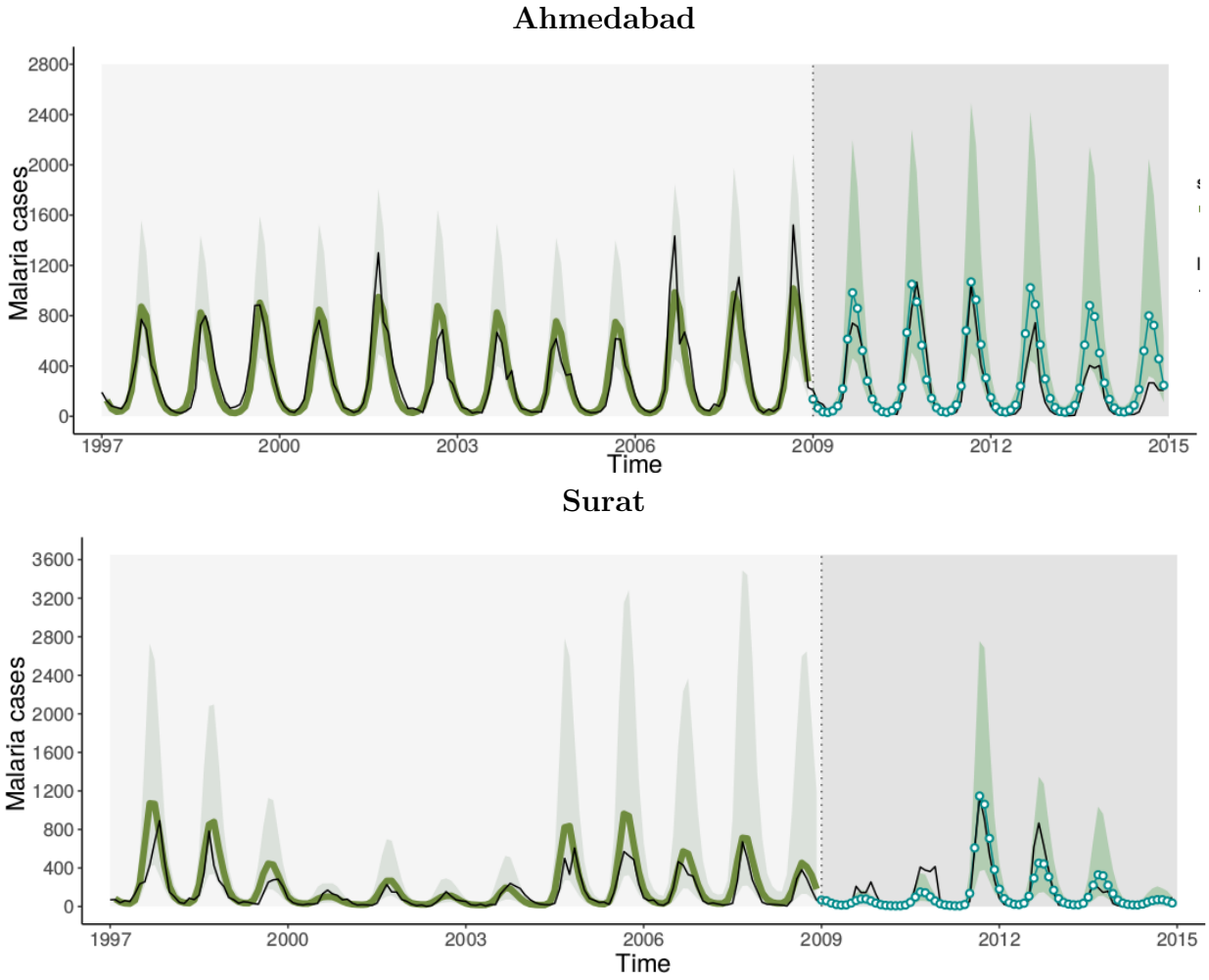

Figure 4.5: Malaria prediction. The median predicted cases are shown for Ahmedabad (A) and Surat (B) together with the observed cases (in black). The median of 1000 simulations is shown in blue, with the 10-90% quantiles of the predicted distribution of cases for each month in the shaded dark gray color. Predictions for the out-of-fit data start in 2009 in, with the median for each month in open circles, and their respective uncertainty envelope in a shaded light color.

4.4 Discussion

Consideration of relative humidity appears essential to explain the interannual variability of malaria in densely populated urban environments. Here we have applied an approach that confronts dynamical process-based models with time series data to provide evidence for a strong role of humidity in the interannual cycles of epidemic malaria in semiarid cities of India. Our results are consistent with experimental findings from (Gaaboub et al. 1971, Kessler and Guerin 2008), and they help explain epidemiological differences between the two cities. Ahmedabad, the dryest and hottest city of the two, exhibits a lower transmission intensity than Surat, a more humid environment **(Figure 4.4)**. Similarly, these results are in agreement with survival curves fitted by (Pharman and Michael 2009, Lunde et al. 2013, Yamana and Eltahir 2013), who found that low levels of humidity decrease transmission by affecting mosquito lifespan. A significant effect of humidity on vector ecology is expected to be a more general pattern applying to other vectors and associated diseases, since all insects have a limited range of tolerable humidity. Given their high surface area to volume ratio, mosquitoes are especially sensitive to desiccation at low humidity levels (Wigglesworth 2011,Liu et al. 2011). Likewise, studies have shown that extremely low levels of relative humidity are fatal to mosquitoes, ultimately determining mosquito abundance in arid regions (Liu et al. 2011). Here we have shown that these demographic and physiological effects have a clear signature in the population dynamics of a vector borne disease of a semi-arid region.

Our findings show that the effect of humidity on the intensity of transmission varies between the two cities. The variation in the effect of humidity between cities is specifically expressed in the higher value of the humidity parameter b_{RH} in the model for Ahmedabad (3.15) compared to Surat (1.74). These differences underscore the effect of humidity as a limiting factor for transmission. Humidities below 40% and temperatures above 40 degrees in Ahmedabad have a greater effect on mosquito survival and activity than in the more humid city of Surat (Bayoh 2001, Gaaboub et al. 1971)[15,18]. This higher transmission

in more humid cities is also explained by Surat's longer window of time when humidities surpasses 60%, an experimental threshold for values conducive to maintaining transmission (Metha 1934).

Our model does not represent the vector dynamics explicitly but instead implements the effect of the vector as a distributed lag on transmission. The values estimated for the delay parameter τ (Table 4.2) in the model are consistent with empirical values of the parasite's developmental time within the vector given the observed temperatures of these regions (Laneri et al. 2010, Roy et al. 2013). Here the delay represents the combined effect of multiple processes that establish a lag in transmission, relative not only to previous levels of infection but also to humidity, through the longevity of adult mosquitoes. The average longer delay in Surat (15 days) compared to Ahmedabad (9 days) is an indication of humidity affecting differentially parasites development in surviving mosquitoes. Differences in humidity levels and periods of time with adequate humidity to maintain transmission affect important parameters of the mosquito and parasite. Our model creates a lower transition rate through the chain of compartments accounting for mosquito survival and parasite development within the vector as a function of seasonality and humidity

Our comparison of different models (Table 4.1) shows that humidity is a more important driver of urban malaria dynamics than temperature in these cities. Humidity exhibits a positive effect on transmission counterbalanced by the negative effect of temperature (opposite to the typical positive relationship described in the literature near and below optimal conditions (Santos-Vega et al. 2016, Mordecai et al. 2013). The higher explanatory power of humidity compared to temperature confirms the results from Santos et al. 2018, where temperature mainly helped explain the spatial distribution of malaria at local scales. Temperature is known to exhibit large spatial variation within cities due to the pronounced heterogeneity of impervious surfaces, whose differing radiative, thermal, aerodynamic, and

moisture properties generate areas of elevated heat (IPCC 2007, Zhao et al. 2014). It is not clear whether humidity acts as a determinant of longevity independent from temperature, or simply operates as an inversely correlated parameter. We need to develop an experimental understanding of the influence of this neglected climate variable in the same way than the field is now attacking temperature parameterizations. Based on such understanding, the model presented here can be extended to encompass a process-based representation of the mosquito subcomponent, which would include explicit parameter dependencies on both temperature and humidity. Specific parameters include biting rate and adult longevity, as well as the developmental rate of the parasite within the vector. This extended model will allow direct investigation of whether long-lived mosquitoes whose survival late in the season depends on humidity, mediate transmission.

In summary, a deeper understanding of the relationship between humidity and urban malaria transmission is key for anticipating the effect of climate trends on the incidence and distribution of vector-borne diseases. For the Indian subcontinent the presence of *Anopheles stephensi*, a vector that does not rely on rainfall for larval recruitment (Metha 1934), accentuates the need to understand the role of humidity in malaria transmission. Changes in surface humidity are associated with anthropogenic warming, which is expected to increase under future climate projections (Singh et al . 2014). In particular, Northwest India is expected to experience a rise in humidity (Dai 2006,Singh et al . 2014, Mukherjee et al. 2018) as well as an increased frequency of precipitation extremes in the mid and end of the 21st century (Vittal et al. 2016). Under this scenario, an understanding of the relationship between humidity and malaria transmission in urban environments is pivotal to achieving India's target of malaria elimination by 2030. Cities can act as a reservoir for the persistence of the disease beyond their administrative limits, preventing elimination despite considerable gains in the fight against rural transmission. Finally, our results indicate the feasibility of forecasting malaria epidemics in these cities of India based on climate variability. Our epidemiological model including humidity exhibited good predictive skill for seasonal cases

and accurate predictions for most malaria years after 2009. This is a promising development towards establishing a climate-based early warning and forecasting system. The applied approach to statistical inference naturally allows for the continuous assimilation of new data, an imperative requirement given the unavoidably non-stationary behavior of malaria incidence.

4.5 Supporting information

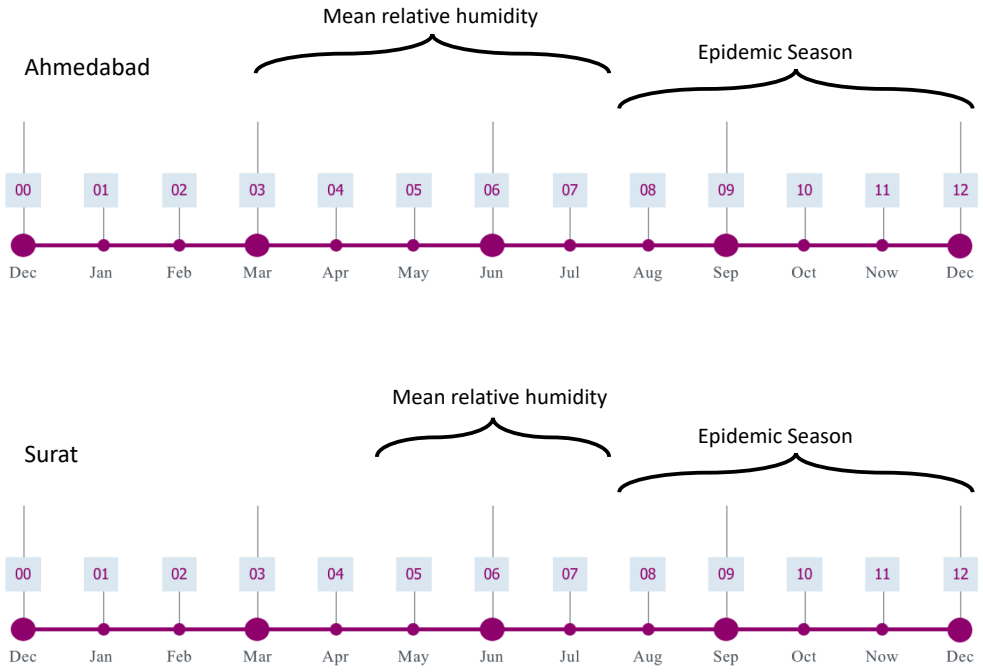

Figure 4.6: **Schematic to show the main malaria seasons and the selected** windows of time for mean temperatures for Ahmedabad (left) and Surat (right). For the purpose of our analyses, we used average relative humidity for the four months and 2 months preceding the epidemic seasons in the two Cities, i.e. Mar to Jul in Ahmedabad, and May to Jul in Surat.

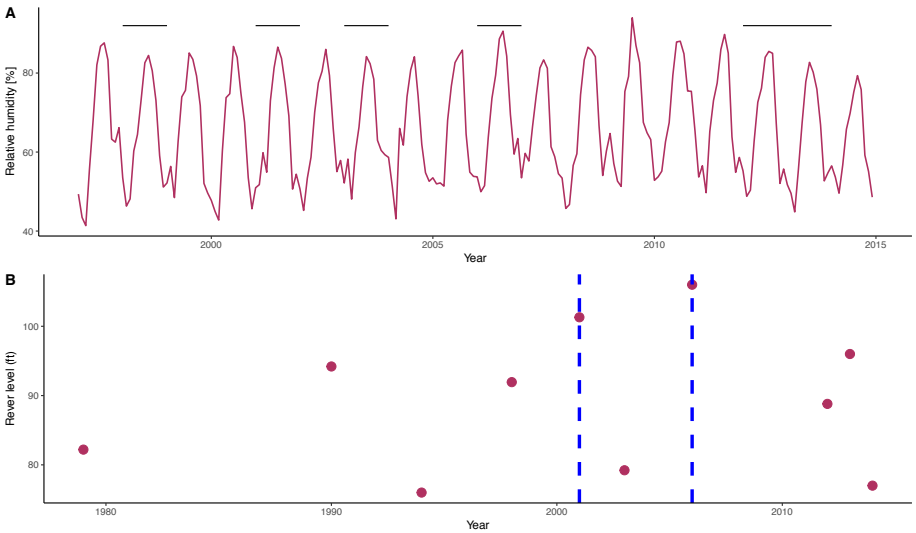

Figure 4.7: A) shows monthly relative humidity for the city of Surat from 1997 to 2014 and flooding years are marked in black. B shows the river level at the Tapi river in Surat for flooding years, 2001 and 2006 are marked in blue as the two years with river levels above 100 ft.

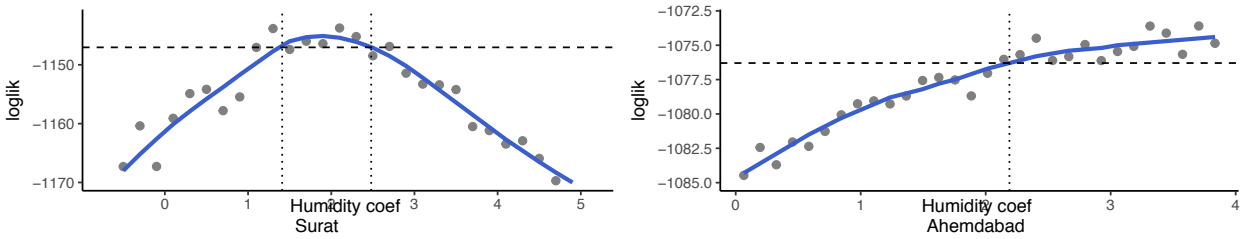

Figure 4.8: Coefficient b_{RH} representing the interannual effect of relative humidity for the core and periphery. The likelihood profile curves are shown, with the intersection between the solid and dashed lines indicating the 95% CI for each region.

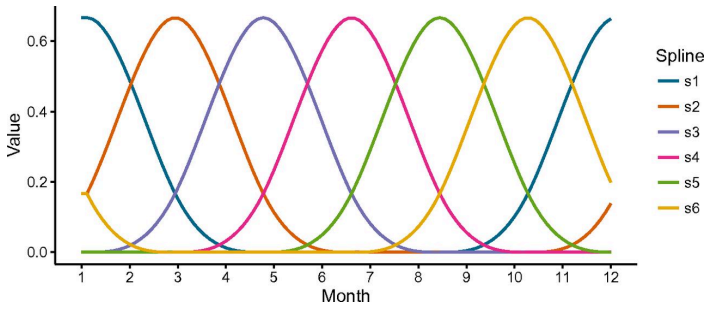

Figure 4.9: **Periodic splines**. The six beta splines (s1 ...s6) used in the expression and estimation of the transmission rate (Eq. 3) are shown. Note the localization of the fourth spline (turquoise) during the monsoon months

CHAPTER 5

CONCLUSION

5.1 Concluding remarks

The growth of cities makes it imperative that we better understand how transmission is affected by the heterogeneity of urban environments and how this heterogeneity interacts with climate forcing. A few recent studies have begun to illustrate how the spatial heterogeneity of socioeconomic and demographic conditions modulates the effects of temporal climate forcing on disease dynamics. Incorporating a fine-scale spatial component into transmission models represents an important technical and conceptual challenge that my dissertation started to addressed. Practical insights on the population dynamics of infectious diseases will be gained from incorporating urban features and their spatial structures. Transmission models that account for these complex dynamics will offer insights for control strategies that differ from those currently underway, by achieving better informed localized interventions and, thus, more cost-effective results. They will also contribute to spatiotemporal prediction based on climate variability in ways that account for different responses in different parts of the city, and to the development of longterm scenarios of urbanization under climate change.

We still have a poor understanding of proximal mechanisms that mediate the effect of environmental and social drivers on malaria incidence in urban environments. My dissertation developed a general framework for the study of transmission dynamics in malaria within urban environments. My findings shed light on the interaction between urban malaria, seasonal climate variability, socio-economic conditions, and population density. Parameter estimation for a mechanistic transmission model will also contribute to epidemiological knowledge about the disease in India, relevant to the planning of future epidemic control. Knowledge on malaria hot-spots can inform the planning of targeted interventions which can achieve more cost-effective results than current reactive policies. More generally, my

work adds a new perspective to the population ecology of infectious diseases by considering climate factors in the context of urban environments and with dynamical approaches.

In chapter 2, I show how climate forcing and socio-economic heterogeneity act synergistically at local scales on the population dynamics of urban malaria in the city of Ahmedabad. Malaria risk is then shown to be associated with socioeconomic indicators and environmental parameters, temperature and humidity. In a more dynamical perspective, an Inhomogeneous Markov Chain Model is used to predict vivax malaria risk. Models that account for climate factors, socioeconomic level, and population size show the highest predictive skill. These results provide a basis for more targeted intervention, such as vector control, based on transmission hotspots.

We continue this exploration of the drivers of urban malaria and their spatial distribution in chapter 3. Here, I found that three main axes of variation in socioeconomic factors corresponding respectively to poverty and access to water, work environment and human mobility, and population density interact with climate factors at different spatial scales. I report previously undescribed positive effects of population density and socio-economic factors, together with a negative effect of temperature, and an enhancing role of humidity. This chapter contributes to the understanding of the multiple drivers of malaria dynamics, especially across a range of spatial scales and in regard to the way in which climatic and demographic/socio-economic conditions interact. It also elucidates the level of aggregation at which one should model the infectious disease system to better capture important effects of spatial heterogeneity in environmental and socioeconomic factors.

In chapter 4, I consider the effect of relative humidity on urban malaria transmission. This climate covariate appears essential to understand the inter-annual variation of malaria in densely populated urban environments. I have applied an approach that confronts dynamical process-based models to time series data to provide evidence of a strong role of humidity

on the inter-annual cycles of epidemic malaria in semiarid cities of India. These results underscore the importance of a deeper understanding of the relationship between humidity and malaria, and provide promising results for developing climate-based early warning and forecasting systems.

5.2 Future directions

Below, I outline some open areas that would contribute would improve this work .

1. **A better empirical-based understanding of how climate factors influence transmission processess:** A basic understanding of the full range of environmental variation is typically lacking for waterborne and vector-borne infections. In particular, parameterizations of the effect of temperature on key demographic and physiological rates (vector longevity, biting rate, larval recruitment) are based on a few well-studied vectors from the early literature, and, even for those parameterizations, the effect of the high end of the temperature spectrum remains poorly documented. Studies of another important climate variable, humidity, are rare or nonexistent. These two aspects are particularly relevant in the context of global warming.

2. **The characterization of the effects of demographic factors:** Characterizing the effects of demographic factors, especially human population density on transmission intensity and disease persistence, is essential to formulate better spatiotemporal models of transmission. Most temporal models for the population dynamics of vector-transmitted diseases are extensions of the wellknown RossMacdonald equations, which assume frequencydependent transmission. Because the force of infection depends on the ratio between the number of vectors and that of the hosts, host population size dilutes transmission intensity, with no explicit distinction between the effect of density and that of population size. However, the few studies of vector-transmitted diseases that have considered human population density in the development of theory and sta-

tistical analysis of incidence patterns suggest an increasing and nonmonotonic effect of density on vector abundance. More in-depth knowledge on appropriate functional forms for the effect of density on transmission rates is needed.

3. **The understanding of how connectivity and movement fluxes within a city affect transmission:** Here too, the leveraging of novel sources of information would be valuable. Telephone data, which are already being applied in epidemiology over larger geographical regions, provide an example. Withincity telephone data would be valuable, together with studies on the accuracy of movement models that are based on easily accessible variables, such as population density. This would permit the computation of movement fluxes when and where data directly related to movement are not available.

4. **The fit of mechanistic models in space and time:** the development of statistical inference methods for stochastic dynamical systems in space and time, applicable to high spatial resolution, is still lacking. Extensions of existing methods based on likelihood and particle filters are good candidates to address this challenge. Future work to decrease the prohibitive computational cost and improve mathematical optimization would allow the inclusion of multiple spatial locations to fit models, and to have an impact on future outbreak control and public health outcomes at relevant epidemiological scales.

REFERENCES

1. Harpham T. Urban health in developing countries: What do we know and where do we go? Heal Place. 2009;15: 107-116.

2. Dye C. Health and urban living. Science. 2008;319: 766-769.

3. Grimm NB, Grimm NB, Faeth SH, Golubiewski NE, Redman CL, Wu J, et al. Global change and the ecology of cities. Science. 2008;319: 756-760. doi:10.1126/science.1150195

4. Nations U. World Urbanization Prospects: The 2014 Revision, Highlights. New York, United. 2014. doi:10.4054/DemRes.2005.12.9

5. Coumou D, Rahmstorf S. A decade of weather extremes. Nat Clim Chang. 2012;2: 1-6. doi:10.1038/nclimate1452

6. Reiner Jr. RC, King AA, Emch M, Mohammad Y, Faruque ASG, Pascual M. Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. Proc Natl Acad Sci U S A. 2012;109: 2033-2036. doi:10.1073/pnas.1108438109

7. Keiser J, Utzinger J, Caldas De Castro M, Smith T a., Tanner M, Singer BH. Urbanization in sub-Saharan Africa and implication for malaria control. American Journal of Tropical Medicine and Hygiene. 2004. pp. 118-127.

8. Lowe R, Bailey TC, Stephenson DB, Jupp TE, Graham RJ, Barcellos C, et al. The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil. Stat Med. 2013;32: 864-883. doi:10.1002/sim.5549

9. 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. Proc Natl Acad Sci . 2016; doi:10.1073/pnas.1518977113

10. Paull SH, Song S, McClure KM, Sackett LC, Kilpatrick AM, Johnson PTJ. From super-spreaders to disease hotspots: Linking transmission across hosts and space. Frontiers in Ecology and the Environment. 2012. pp. 75-82. doi:10.1890/110111

11. Perkins TA, Scott TW, Le Menach A, Smith DL. Heterogeneity, Mixing, and the Spatial Scales of Mosquito-Borne Pathogen Transmission. PLoS Comput Biol. 2013;9. doi:10.1371/journal.pcbi.1003327

12. Vazquez-Prokopec, G.M., T.A. Perkins, L.A. Waller, et al. 2016. Coupled heterogeneities and their impact on parasite transmission and control. Trends Parasitol. 32: 356-367.

13. UN-Habitat. Prosperity of Cities: State of the Worlds Cities 2012/2013 [Internet]. State of the Worlds Cities. 2012. doi:10.1080/07293682.2013.861498

14. Batty M. The size, scale, and shape of cities. Science. 2008;319: 769-771.

15. Bettencourt LM a. The origins of scaling in cities. Science (80-). 2013;340: 1438-1441. doi:10.1126/science.1235823

16. Rodriguez-Iturbe I, Rinaldo A. Fractal river basins: chance and self-organization. Power. 1997.

17. West GB, Enquist BJ, Brown JH. The fourth dimension of life: fractal geometry and allometric scaling of organisms. Science (80-). 1999;284: 1677-9.

18. Waters JS, Holbrook CT, Fewell JH, Harrison JF. Allometric scaling of metabolism, growth, and activity in whole colonies of the seed-harvester ant Pogonomyrmex californicus. Am Nat. 2010;176: 501-10. doi:10.1086/656266

19. Dendrinos D, Mullally H. Urban Evolution: Studies in the Mathematical Ecology of Cities. Oxford: Oxford University Press; 1985.

20. Segal D. ARE THERE RETURNS TO SCALE IN CITY SIZE. Rev Econ Stat. 1976;58: 339. doi:10.2307/1924956

21. Bettencourt LM a, Lobo J, Helbing D, Khnert C, West GB. Growth, innovation, scaling, and the pace of life in cities. Proc Natl Acad Sci U S A. 2007;104: 7301-7306. doi:10.1073/pnas.0610172104

22. Bettencourt L, West G. A unified theory of urban living. Nature. 2010;467: 912-3. doi:10.1038/467912a

23. Bharti N, Tatem AJ, Ferrari MJ, Grais RF, Djibo A, Grenfell BT. Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. Science. 2011;334: 1424-7. doi:10.1126/science.1210554

24. Riley S. Large-Scale Spatial-Transmission Models of Infectious Disease. Science (80-). 2007;316: 1298-1301. doi:10.1126/science.1134695

25. Cosner C, Beier JC, Cantrell RS, Impoinvil D, Kapitanski L, Potts MD, et al. The effects of human movement on the persistence of vector-borne diseases. J Theor Biol. 2009;258: 550-560. doi:10.1016/j.jtbi.2009.02.016

26. Jung, W.S., F. Wang and H.E. Stanley. 2008. Gravity model in the Korean highway.

27. Tarasov, A., F. Kling, A. Pozdnoukhov. 2013. Prediction of user location using the radiation model and social checkins. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, article No. 8, ACM, New York.

28. Simini F, Gonzlez MC, Maritan A, Barabsi A-L. A universal model for mobility and migration patterns - Supplementary Information. Nature. 2012;484: 96-100. doi:10.1038/nature10856

29. Lu X, Wetter E, Bharti N, Tatem AJ, Bengtsson L. Approaching the limit of predictability in human mobility. Sci Rep. 2013;3: 2923. doi:10.1038/srep02923

30. Erlander, S., N.F. Stewart. 1990. The Gravity Model in Transportation Analysis: Theory and Extensions. Utrecht: VSP.

31. Conzen MRG. Morphogenesis, morphological regions and secular human agency in the historic townscape, as exemplified by Ludlow. Urban Hist Geogr. 1988; 253-272.

32. Gauthier P, Gilliland J. Mapping urban morphology: A classification scheme for interpreting contributions to the study of urban form. Urban Morphol. 2006;10: 41-50.

33. Newton PW. Reshaping cities for a more sustainable future: exploring the link between urban form, air quality, energy and greenhouse gas emissions. Melbourne: Research monograph (Australian Housing and Urban Research Institute); 1997.

34. Antonovics J, Iwasa Y, Hassell MP. A Generalized Model of Parasitoid, Venereal, and Vector-Based Transmission Processes. Am Nat. 1995;145: 661. doi:10.1086/285761

35. Xia Y, Bjornstad ON, Grenfell BT. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. Am Nat. 2004;164: 267-281. doi:10.1086/422341

36. Viboud C, Bjrnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science. 2006;312: 447-51. doi:10.1126/science.1125237

37. Mari, L., E. Bertuzzo, L. Righetto, et al. 2012. Modelling cholera epidemics: the role of waterways, human mobility and sanitation. J. R. Soc. Interface 9: 376-388.

38. Wesolowski, A., A.J. Tatem, J. Lessler, et al. 2015. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. Proc. Natl. Acad. Sci. U.S.A. 112: 11114-11119.

39. Buckee, C.O., A. Wesolowski, N.N. Eagle, et al. 2013. Mobile phones and malaria: modeling human and parasite travel. Travel Med. Infect. Dis. 11: 15-22.

40. Stoddard, S.T., B.M. Forshey, A.C. Morrison, et al. 2013. Housetohouse human movement drives dengue virus transmission. Proc. Natl. Acad. Sci. U.S.A. 110: 994-999.

41. Reiner, R.C., S.T. Stoddard and T.W. Scott. 2014. Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. Epidemics 6: 30-36.

42. Perkins, T.A., A.J. Garcia, V.A. PazSoldn, et al. 2014. Theory and data for simulating finescale human movement in an urban environment. J. R. Soc. Interface 11: 20140642.

43. Honorio, N.A., R.M.R. Nogueira, C.T. Codeco, et al. 2009. Spatial evaluation and modeling of dengue seroprevalence and vector density in Rio de Janeiro, Brazil. PLoS Negl. Trop. Dis. 3: e545.

44. PerezSaez, J., A.A. King, A. Rinaldo, et al. 2016. Climatedriven endemic cholera is modulated by human mobility in a megacity. Adv. Water Resour. In press.

45. World Health Organization (WHO). 2008. Our cities, our health, our future. World Health Organization, Geneva.

46. Delgado, S., R.C. Neyra, V.R.Q. Machaca, et al. 2011. A history of Chagas disease transmission, control, and reemergence in perirural La Joya, Peru. PLoS Negl. Trop. Dis. 5: e970.

47. Barton, H. and M. Grant. 2006. A health map for the local human habitat. J. R. Soc. Promot. Health 126: 252-253.

48. Rydin, Y., A. Bleahu, M. Davies, et al. 2012. Shaping cities for health: complexity and the planning of urban environments in the 21st century. Lancet 379: 2079-2108.

49. Thrall, P.H. and J.J. Burdon. 1999. The spatial scale of pathogen dispersal: consequences for disease dynamics and persistence. Evol. Ecol. Res. 1: 681-701.

50. Read, J.M. and M.J. Keeling. 2003. Disease evolution on networks: the role of contact structure. Proc. Biol. Sci. 270: 699-708.

51. Watts, D.J., R. Muhamad, D.C. Medina and P.S. Dodds. 2005. Multiscale, resurgent epidemics in a hierarchical metapopulation model. Proc. Natl. Acad. Sci. U.S.A. 102: 11157-11162.

52. Oke, T.R. 1982. The energetic basis of the urban heat island. Q. J. R. Meteorol. Soc. 108: 1-24.

53. Arnfield, A.J. 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int. J. Climatol. 23: 1-26.

54. Marshall Shepherd, J., M. Carter, M. Manyin, et al. 2010. The impact of urbanization on current and future coastal precipitation: a case study for Houston. Environ. Plann. B Plann. Des. 37: 284-304.

55. Niyogi, D., P. Pyle, M. Lei, et al. 2011. Urban modification of thunderstorms: an observational storm climatology and model case study for the Indianapolis urban region. J. Appl. Meteorol. Climatol. 50: 1129-1144.

56. Jetten, T.H. and D.A. Focks. 1997. Potential changes in the distribution of dengue transmission under climate warming. Am. J. Trop. Med. Hyg. 57: 285-297.

57. Santamouris, M. 2001. Energy and Climate in the Urban Built Environment. London: James and James.

58. Honda, Y., M. Kondo, G. McGregor, et al. 2014. Heatrelated mortality risk model for climate change impact projection. Environ. Health Prev. Med. 19: 56-63.

59. LaDeau, S.L., B.F. Allan, P.T. Leisnham and M.Z. Levy. 2015. The ecological foundations of transmission potential and vectorborne disease in urban landscapes. Funct. Ecol. 29: 889-901.

60. McGranahan, G., D. Balk and B. Anderson. 2007. The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. Environ. Urban. 19: 17-37.

61. Syvitski, J. and S. Higgins. 2012. Going under: the world's sinking deltas. New Sci. 216: 40-43.

62. RuizMoreno, D., M. Pascual, M. Bouma, et al. 2007. Cholera seasonality in Madras (1901-1940): dual role for rainfall in endemic and epidemic regions. Ecohealth 4: 52-62.

63. IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment of the Intergovernmental Panel on Climate Change. 446: 727-728. Cambridge and New York: Cambridge University Press.

64. Douglas, I., K. Alam, M. Maghenda, et al. 2008. Unjust waters: climate change, flooding and the urban poor in Africa. Environ. Urban. 20: 187-205.

65. Unger, A. and L.W. Riley. 2007. Slum health: from understanding to action. PLoS Med. 4: 1561-1566.

66. Lipp, E.K., A. Huq and R.R. Colwell. 2002. Effects of global climate on infectious disease: the cholera model effects of global climate on infectious disease: the cholera model. Clin. Microbiol. Rev. 15: 757-770.

67. Patz, J.A., T.K. Graczyk, N. Geller and A.Y. Vittor. 2000. Effects of environmental change on emerging parasitic diseases. Int. J. Parasitol. 30: 1395-1405.

68. Cann, K.F., D.R. Thomas, R.L. Salmon, et al. 2013. Extreme waterrelated weather events and waterborne disease. Epidemiol. Infect. 141: 671-686.

69. Tinker, S.C., C.L. Moe, M. Klein, et al. 2010. Drinking water turbidity and emergency department visits for gastrointestinal illness in Atlanta, 1993-2004. J. Expo. Sci. Environ. Epidemiol. 20: 19-28.

70. Kovats, R. 1999. El Nio and health. World Health Organization, Geneva.

71. Gage, K.L., T.R. Burkot, R.J. Eisen and E.B. Hayes. 2008. Climate and vectorborne diseases. Am. J. Prevent. Med. 35: 436-450.

72. Kuhn, K., D. Campbelllendrum, A. Haines and J. Cox. 2005. Using climate to predict infectious disease epidemics. World Health Organization, Geneva.

73. Chretien, J.P., A. Anyamba, S.A. Bedno, et al. 2007. Droughtassociated chikungunya emergence along coastal East Africa. Am. J. Trop. Med. Hyg. 76: 405-407.

74. Patz, J.A., A.K. Githeko, J.P. McCarty, et al. 2003. Climate Change and Human Health: Risks and Responses. World Health Organization, Geneva.

75. Metha, D. 1934. Studies on the longevity of some Indian anophelines. Part I. Survival of Anopheles subpictus Grassi under controlled conditions of temperature and humidity. Rev. App. Ent. B 23: 18-19.

76. Paaijmans, K.P., S. Blanford, A.S. Bell, et al. 2010. Influence of climate on malaria transmission depends on daily temperature variation. Proc. Natl. Acad. Sci. U.S.A. 107: 15135-15139.

77. Cator, L.J., S. Thomas, K.P. Paaijmans, et al. 2013. Characterizing microclimate in urban malaria transmission settings: a case study from Chennai, India. Malar. J. 12: 84.

78. Mordecai, E.A., K.P. Paaijmans, L.R. Johnson, et al. 2013. Optimal temperature for malaria transmission is dramatically lower than previously predicted. Ecol. Lett. 16: 22-30.

79. Snodgrass, R.E. 1939. The principles of insect physiology. Science 90: 159.

80. Gaaboub, I.A., S.K. ElSawaf and M.A. ElLatif. 1971. Effect of different relative humidities and temperatures on eggproduction and longevity of adults of Anopheles (Myzomyia) pharoensis Theob. J. Appl. Entomol. 67: 88-94.

81. Thu, H.M., K.M. Aye and S. Thein. 1998. The effect of temperature and humidity on dengue virus propagation in Aedes aegypti mosquitos. Southeast Asian J. Trop. Med. Public Health 29: 280-284.

82. Moran, P. 1953. The statistical analysis of the Canadian lynx cycle. Aust. J. Zool. 1: 291.

83. Santos-Vega M, Bouma MJ, Kohli V, Pascual M. Population Density, Climate Variables and Poverty Synergistically Structure Spatial Risk in Urban Malaria in India. PLoS Negl Trop Dis. 2016;10. doi:10.1371/journal.pntd.0005155

84. Kovats, R.S., D.H. CampbellLendrum, A.J. McMichael, et al. 2001. Early effects of climate change: do they include changes in vectorborne disease? Philos. Trans. R. Soc. Lond. B Biol. Sci. 356: 1057-1068.

85. Bousema, T., J.T. Griffin, R.W. Sauerwein, et al. 2012. Hitting hotspots: spatial targeting of malaria for control and elimination. PLoS Med. 9: e1001165.

86. Trape, J.F., E. LefebvreZante, F. Legros, et al. 1992. Vector density gradients and the epidemiology of urban malaria in Dakar, Senegal. Am. J. Trop. Med. Hyg. 47: 181-189.

87. Wang, T.P., J. Shrivastava, M.V. Johansen, et al. 2006. Does multiple hosts mean multiple parasites? Population genetic structure of Schistosoma japonicum between definitive host species. Int. J. Parasitol. 36: 1317-1325.

88. Qi, Q., C.A. Guerra, C.L. Moyes, et al. 2012. The effects of urbanization on global Plasmodium vivax malaria transmission. Malar. J. 11: 403.

89. Teixeira, M.D.G., M.L. Barreto, M.D.C.N. Costa, et al. 2002. Dynamics of dengue virus circulation: a silent epidemic in a complex urban area. Trop. Med. Int. Heal. 7: 757-762.

90. Eisen, R.J., R.E. Enscore, B.J. Biggerstaff, et al. 2007. Human plague in the Southwestern United States, 1957-2004: spatial models of elevated risk of human exposure to Yersinia pestis. J. Med. Entomol. 44: 530-537.

91. Banu, S., W. Hu, Y. Guo, et al. 2014. Projecting the impact of climate change on dengue transmission in Dhaka, Bangladesh. Environ. Int. 63: 137-142.

92. Hammond, S.N., A.L. Gordon, E.D.C. Lugo, et al. 2007. Characterization of Aedes aegypti (Diptera: Culcidae) production sites in urban Nicaragua. J. Med. Entomol. 44: 851-860.

93. Cator, L.J., P.A. Lynch, M.B. Thomas and A.F. Read. 2014. Alterations in mosquito behaviour by malaria parasites: potential impact on force of infection. Malar. J. 13: 164.

94. Woolhouse, M.E., C. Dye, J.F. Etard, et al. 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc. Natl. Acad. Sci. U.S.A. 94: 338-342.

95. Lindsay, S.W. and R.W. Snow. 1988. The trouble with eaves; house entry by vectors of malaria. Trans. R. Soc. Trop. Med. Hyg. 82: 645-646.

96. Midega, J.T., D.L. Smith, A. Olotu, et al. 2012. Wind direction and proximity to larval sites determines malaria risk in Kilifi district in Kenya. Nat. Commun. 3: 674.

97. Carter, R., K.N. Mendis and D. Roberts. 2000. Spatial targeting of interventions against malaria. Bull. World Health Organ. 78: 1401-1411.

98. Laneri, K., R.E. Paul, A. Tall, et al. 2015. Dynamical malaria models reveal how immunity buffers effect of climate variability. Proc. Natl. Acad. Sci. U.S.A. 112: 8786-8791.

99. Pascual, M. 2015. Climate and population immunity in malaria dynamics: harnessing information from endemicity gradients. Trends Parasitol. 31: 532-534.

100. Campbell, L.P., C. Luther, D. MooLlanes, et al. 2015. Climate change influences on global distributions of dengue and chikungunya virus vectors. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370: 20140135.

101. Chowell, G., C.A. Torre, C. MunaycoEscate, et al. 2008. Spatial and temporal dynamics of dengue fever in Peru: 1994-2006. Epidemiol. Infect. 136: 1667-1677.

102. Descloux, E., M. Mangeas, C.E. Menkes, et al. 2012. Climatebased models for understanding and forecasting dengue epidemics. PLoS Negl. Trop. Dis. 6: e1470.

103. StewartIbarra, A.M. and R. Lowe. 2013. Climate and nonclimate drivers of dengue epidemics in southern coastal Ecuador. Am. J. Trop. Med. Hyg. 88: 971-981.

104. Hay SI, Guerra C a, Tatem AJ, Atkinson PM, Snow RW. Urbanization, malaria transmission and disease burden in Africa. Nat Rev Microbiol. 2005;3: 81-90.

105. Omumbo J a., Hay SI, Snow RW, Tatem a. J, Rogers DJ. Modelling malaria risk in East Africa at high-spatial resolution. Trop Med Int Heal. 2005;10: 557-566. doi:10.1111/j.1365-3156.2005.01424.x

106. Robert V, Macintyre K, Keating J, Trape JF, Duchemin JB, Warren M, et al. Malaria transmission in urban sub-Saharan Africa. American Journal of Tropical Medicine and Hygiene. 2003. pp. 169-176.

107. Donnelly MJ, McCall PJ, Lengeler C, Bates I, DAlessandro U, Barnish G, et al.

Malaria and urbanization in sub-Saharan Africa. Malar J. 2005;4: 12. doi:10.1186/1475-2875-4-12

108. Reiner RC, Smith DL, Gething PW. Climate change, urbanization and disease: Summer in the city... Trans R Soc Trop Med Hyg. 2014;109: 171-172.

109. Roy M, Bouma MJ, Ionides EL, Dhiman RC, Pascual M. The Potential Elimination of Plasmodium vivax Malaria by Relapse Treatment: Insights from a Transmission Model and Surveillance Data from NW India. PLoS Negl Trop Dis. 2013;7. doi:10.1371/journal.pntd.0001979

110. Baird JK. Neglect of Plasmodium vivax malaria. Trends in Parasitology. 2007. pp. 533-539. doi:10.1016/j.pt.2007.08.011

111. Sattabongkot J, Tsuboi T, Zollner GE, Sirichaisinthop J, Cui L. Plasmodium vivax transmission: Chances for control? Trends in Parasitology. 2004. pp. 192-198. doi:10.1016/j.pt.2004.02.001

112. Price RN, Tjitra E, Guerra CA, Yeung S, White NJ, Anstey NM. Vivax malaria: Neglected and not benign. Am J Trop Med Hyg. 2007;77: 79-87. doi:77/6 Suppl/79

113. Gross R, Schell B, Molina MC, Leo MA, Strack U. The impact of improvement of water supply and sanitation facilities on diarrhea and intestinal parasites: a Brazilian experience with children in two low-income urban communities. Rev Saude Publica. 1989;23: 214-220. doi:10.1590/S0034-89101989000300006

114. Klinkenberg E, McCall P, Wilson MD, Amerasinghe FP, Donnelly MJ. Impact of urban agriculture on malaria vectors in Accra, Ghana. Malar J. 2008;7: 151. doi:10.1186/1475-2875-7-151

115. Afrane YA, Klinkenberg E, Drechsel P, Owusu-Daaku K, Garms R, Kruppa T. Does irrigated urban agriculture influence the transmission of malaria in the city of Kumasi,

Ghana? Acta Trop. 2004;89: 125-134. doi:10.1016/j.actatropica.2003.06.001

116. Noor AM, Zurovac D, Hay SI, Ochola SA, Snow RW. Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. Trop Med Int Heal. 2003;8: 917-926. doi:10.1046/j.1365-3156.2003.01112.x

117. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al. A world malaria map: plasmodium falciparum endemicity in 2007. PLoS Med. 2009;6: 0286-0302. doi:10.1371/journal.pmed.1000048

118. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Soldan VP, Kochel TJ, Kitron U, et al. The role of human movement in the transmission of vector-borne pathogens. PLoS Negl Trop Dis. 2009;3. doi:10.1371/journal.pntd.0000481

119. De Silva PM, Marshall JM. Factors contributing to urban malaria transmission in sub-saharan Africa: A systematic review. Journal of Tropical Medicine. 2012.

120. Hasibeder G, Dye C. Population dynamics of mosquito-borne disease: Persistence in a completely heterogeneous environment. Theor Popul Biol. 1988;33: 31-53.

121. Cosner C, Beier JC, Cantrell RS, Impoinvil D, Kapitanski L, Potts MD, et al. The effects of human movement on the persistence of vector-borne diseases. J Theor Biol. 2009;258: 550-560. doi:10.1016/j.jtbi.2009.02.016

122. Reiner RC, Perkins TA, Barker CM, Niu T, Chaves LF, Ellis AM, et al. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010. J R Soc Interface. 2013;10: 20120921. doi:10.1098/rsif.2012.0921

123. Lawless JF. Negative binomial and mixed Poisson regression. Can J Stat. 1987;15: 209-225. doi:10.2307/3314912

124. Finger F, Knox A, Bertuzzo E, Mari L, Bompangue D, Gatto M, et al. Cholera in the Lake Kivu region (DRC): Integrating remote sensing and spatially explicit epidemiological modeling. Water Resour Res. 2014;50: 5624-5637. doi:10.1002/2014WR015521

125. Smith DL, Dushoff J, McKenzie FE. The risk of a mosquito-borne infection in a heterogeneous environment. PLoS Biol. 2004;2: e368. doi:10.1371/journal.pbio.0020368

126. Bansal S, Grenfell BT, Meyers LA. When individual behaviour matters: homogeneous and network models in epidemiology. J R Soc Interface. 2007;4: 879-891. doi:10.1098/rsif.2007.1100

127. Bui HM, Clements ACA, Nguyen QT, Nguyen MH, Le XH, Hay SI, et al. Social and environmental determinants of malaria in space and time in Viet Nam. Int J Parasitol. 2011;41: 109-16. doi:10.1016/j.ijpara.2010.08.005

128. Sachs J, Malaney P. The economic and social burden of malaria. Nature. 2002;415: 680-685. doi:10.1038/415680a

129. Levin S a. Introduction: Infectious diseases. Environ Dev Econ. 2007;12: 625-626. doi:10.1017/S1355770X07003798

130. Worrall E, Basu S, Hanson K. Is malaria a disease of poverty? A review of the literature. Tropical Medicine and International Health. 2005. pp. 1047-1059. doi:10.1111/j.1365-3156.2005.01476.x

131. Ajayi IO, Jegede AS, Falade CO, Sommerfeld J. Assessing resources for implementing a community directed intervention (CDI) strategy in delivering multiple health interventions in urban poor communities in Southwestern Nigeria: a qualitative study. Infect Dis Poverty. 2013;2: 25. doi:10.1186/2049-9957-2-25

132. Kessler S, Guerin PM. Responses of Anopheles gambiae, Anopheles stephensi, Aedes aegypti, and Culex pipiens mosquitoes (Diptera : Culicidae) to cool and humid refugium

conditions. J Vector Ecol. 2008;33: 145-149.

133. Mayne B. A study of the influence of relative humidity on the life and infectibility of the mosquito. Indian J Med Res. 1930;17: 1119-1137.

134. Kirby MJ, Green C, Milligan PM, Sismanidis C, Jasseh M, Conway DJ, et al. Risk factors for house-entry by malaria vectors in a rural town and satellite villages in The Gambia. Malar J. 2008;7: 1-9. doi:10.1186/1475-2875-7-2

135. Sharma VP. Re-emergence of malaria in India. Indian Journal of Medical Research. 1996. pp. 26-45.

136. Grimm NB, Grimm NB, Faeth SH, Golubiewski NE, Redman CL, Wu J, et al. Global change and the ecology of cities. Science. 2008;319: 756-760. doi:10.1126/science.1150195

137. Harpham T. Urban health in developing countries: What do we know and where do we go? Heal Place. 2009;15: 107-116. doi:10.1016/j.healthplace.2008.03.004

138. Vittal H, Ghosh S, Karmakar S, Pathak A, Murtugudde R. Lack of Dependence of Indian Summer Monsoon Rainfall Extremes on Temperature: An Observational Evidence. Sci Rep. 2016;6. doi:10.1038/srep31039

139. Shepherd JM, Pierce H, Negri AJ. Rainfall Modification by Major Urban Areas: Observations from Spaceborne Rain Radar on the TRMM Satellite. J Appl Meteorol. 2002;41: 689-701. doi:10.1175/1520-0450(2002)041¡0689:RMBMUA¿2.0.CO;2

140. Jacobi, P., Kjellen, M., McGranahan, G., Songsore, J. and Surjadi, C. (2010) The citizens at risk: from urban sanitation to sustainable cities. Routledge.

141. Ahern J. Urban landscape sustainability and resilience: The promise and challenges of integrating ecology with urban planning and design. Landsc Ecol. 2013;28: 1203-1212. doi:10.1007/s10980-012-9799-z

142. Pickett STA, Cadenasso ML, Rosi-Marshall EJ, Belt KT, Groffman PM, Grove JM, et al. Dynamic heterogeneity: a framework to promote ecological integration and hypothesis generation in urban systems. Urban Ecosyst. 2017;20. doi:10.1007/s11252-016-0574-9

143. Zhou W, Pickett STA, Cadenasso ML. Shifting concepts of urban spatial heterogeneity and their implications for sustainability. Landsc Ecol. 2017;32: 15-30. doi:10.1007/s10980-016-0432-4

144. Bolay JC. Slums and urban development: Questions on society and globalisation. Eur J Dev Res. 2006;18: 284-298. doi:10.1080/09578810600709492

145. Mitlin, D., and Satterthwaite, D. (201) Urban Poverty in the Global South: Scale and Nature. Routledge, London.

146. Zhao L, Lee X, Smith RB, Oleson K. Strong contributions of local background climate to urban heat islands. Nature. 2014;511: 216-219. doi:10.1038/nature13462

147. Mishra V, Ganguly AR, Nijssen B, Lettenmaier DP. Changes in observed climate extremes in global urban areas. Environ Res Lett. 2015;10. doi:10.1088/1748-9326/10/2/024005

148. Donnelly MJ, McCall PJ, Lengeler C, Bates I, DAlessandro U, Barnish G, et al. Malaria and urbanization in sub-Saharan Africa. Malar J. 2005;4: 12. doi:10.1186/1475-2875-4-12

149. Omumbo J a., Hay SI, Snow RW, Tatem a. J, Rogers DJ. Modelling malaria risk in East Africa at high-spatial resolution. Trop Med Int Heal. 2005;10: 557-566. doi:10.1111/j.1365-3156.2005.01424.x

150. Murdock CC, Blanford S, Luckhart S, Thomas MB. Ambient temperature and dietary supplementation interact to shape mosquito vector competence for malaria. J Insect Physiol. 2014;67: 37-44. doi:10.1016/j.jinsphys.2014.05.020

151. Ahern M, Kovats RS, Wilkinson P, Few R, Matthies F. Global health impacts of floods: Epidemiologic evidence. Epidemiol Rev. 2005;27: 36-46. doi:10.1093/epirev/mxi004

152. Rezza G, Nicoletti L, Angelini R, Romi R, Finarelli A, Panning M, et al. Infection with chikungunya virus in Italy: an outbreak in a temperate region. Lancet. 2007;370: 1840-1846. doi:10.1016/S0140-6736(07)61779-6

153. Smith DL, Dushoff J, McKenzie FE. The risk of a mosquito-borne infection in a heterogeneous environment. PLoS Biol. 2004;2: e368. doi:10.1371/journal.pbio.0020368

154. Acevedo MA, Prosper O, Lopiano K, Ruktanonchai N, Caughlin TT, Martcheva M, et al.Spatial heterogeneity, host movement and mosquito-borne disease transmission. PLoS One. 2015;10: e0127552. doi:10.1371/journal.pone.0127552

155. Cosner C, Beier JC, Cantrell RS, Impoinvil D, Kapitanski L, Potts MD, et al. The effects of human movement on the persistence of vector-borne diseases. J Theor Biol.2009;258: 550-560. doi:10.1016/j.jtbi.2009.02.016

156. Watts N, Adger WN, Agnolucci P, Blackstock J, Byass P, Cai W, et al. Health and climate change: Policy responses to protect public health. The Lancet. 2015. pp. 1861-1914. doi:10.1016/S0140-6736(15)60854-6

157. Georgescu M, Morefield PE, Bierwagen BG, Weaver CP. Urban adaptation can roll back warming of emerging megapolitan regions. Proc Natl Acad Sci. 2014;111: 2909-2914.doi:10.1073/pnas.1322280111

158. Trusilova K, Frh B, Brienen S, Walter A, Masson V, Pigeon G, et al. Implementation of an urban parameterization scheme into the regional climate model COSMO-CLM. J Appl Meteorol Climatol. 2013;52: 2296-2311. doi:10.1175/JAMC-D-12-0209.1

159. Potter KA, Arthur Woods H, Pincebourde S. Microclimatic challenges in global change biology. Glob Chang Biol. 2013;19: 2932-2939. doi:10.1111/gcb.12257

160. Pincebourde S, Murdock CC, Vickers M, Sears MW. Fine-Scale Microclimatic Variation Can Shape the Responses of Organisms to Global Change in Both Natural and Urban Environments. Integr Comp Biol. 2016;56: 45-61. doi:10.1093/icb/icw016

161. Afrane YA, Klinkenberg E, Drechsel P, Owusu-Daaku K, Garms R, Kruppa T. Does irrigated urban agriculture influence the transmission of malaria in the city of Kumasi, Ghana? Acta Trop. 2004;89: 125-134. doi:10.1016/j.actatropica.2003.06.001

162. Murdock CC, Sternberg ED, Thomas MB. Malaria transmission potential could be reduced with current and future climate change. Sci Rep. 2016;6. doi:10.1038/srep27771

163. Baker JL. Urban Poverty : A Global View. World Bank Gr. 2008; v-27.

164. Crane R, Daniere A. Measuring access to basic services in global cities: Descriptive and behavioral approaches. J Am Plan Assoc. 1996;62: 203-221.

165. Peng G, Li J, Chen Y, Norizan AP, Tay L. High-resolution surface relative humidity computation using MODIS image in Peninsular Malaysia. Chinese Geogr Sci. 2006;16: 260-264. doi:10.1007/s11769-006-0260-6

166. Anselin L. Spatial Econometrics: Methods and Models [Internet]. Operational Regional Science Series. 1988. doi:10.1007/978-94-015-7799-1

167. Anselin L. The Moran Scatterplot as an ESDA tool to assess local instability in spatial association. Spatial analytical perspectives on GIS. 1996. pp. 111-125.

168. Anselin L. Spatial econometrics. A companion to Theor Econom. 2001; 310-330. doi:10.1002/9780470996249

169. Anselin L. Local Indicators of Spatial AssociationLISA. Geogr Anal. 1995;27: 93-115. doi:10.1111/j.1538-4632.1995.tb00338.

170. Longini IM, Koopman JS, Haber M, Cotsonis GA. Statistical inference for infectious diseases. Risk-specific household and community transmission parameters. Am J Epidemiol. 1988;128: 845-59. Available: http://www.ncbi.nlm.nih.gov/pubmed/3421247

171. Viboud C, Bjrnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science (80-). 2006;312: 447-451. doi:10.1126/science.1125237

172. Bertuzzo E, Mari L, Righetto L, Gatto M, Casagrandi R, Blokesch M, et al. Prediction of the spatial evolution and effects of control measures for the unfolding Haiti cholera outbreak. Geophys Res Lett. 2011;38. doi:10.1029/2011GL046823

173. Lowe R, Cazelles B, Paul R, Rod X. Quantifying the added value of climate information in a spatio-temporal dengue model. Stoch Environ Res Risk Assess. 2016;30: 2067-2078. doi:10.1007/s00477-015-1053-1

174. Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. J Stat Softw. 2005;12: 1-16. doi:10.1103/PhysRevLett.44.1404

175. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol. 2002;64: 583-616.

176. Kramer M, Service BC, Beltsville a RS. R2 Statistics for Mixed Models. Proc Conf Appl Stat Agric. 2005; 148-160.

177. Magee L. R2measures based on wald and likelihood ratio joint significance tests. Am Stat. 1990;44: 250-253. doi:10.1080/00031305.1990.10475731

178. Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. Chapman Texts in Statistical Science Series. 2004. doi:10.1007/s13398-014-0173-7.2

179. Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Stat Sci. 1992;7: 457-472. doi:10.1214/ss/1177011136

180. Parham PE, Michael E. Modeling the effects of weather and climate change on malaria transmission. Environ Health Perspect. 2010;118: 620-626. doi:10.1289/ehp.0901256

181. Besag J, York J, Molli A: Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math. 1991, 43: 1-20. 10.1007/BF00116466.

182. Salje H, Lessler J, Paul KK, Azman AS, Rahman MW, Rahman M, et al. How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study. Proc Natl Acad Sci. 2016;113: 13420-13425.

183. Romeo-Azna, Vr and Pascual, M. Linking vector abundance to human density: a modified model of vector-borne transmission. Ecology Letters. 2018.

184. Siraj AS, Bouma MJ, Santos-Vega M, Yeshiwondim AK, Rothman DS, Yadeta D, et al. Temperature and population density determine reservoir regions of seasonal persistence in highland malaria. Proc R Soc B Biol Sci. 2015;282: 20151383. doi:10.1098/rspb.2015.1383

185. Arnfield AJ. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. International journal of climatology. 2003;23(1):1-26.

186. Edmondson JL, Stott I, Davies ZG, Gaston KJ, Leake JR. Soil surface temperatures reveal moderation of the urban heat island effect by trees and shrubs. Sci Rep. 2016;6. doi:10.1038/srep33708

187. Pascual M, Rod X, Ellner SP, Colwell R, Bouma MJ (2000) Cholera dynamics and El Nio-Southern Oscillation. Science 289(5485):1766-1769.

188. Zhou G, Minakawa N, Githeko AK, Yan G (2004) Association between climate variability and malaria epidemics in the East African highlands. Proc Natl Acad Sci USA 101(8):2375-2380

189. Cazelles B, Chavez M, McMichael AJ, Hales S (2005) Nonstationary influence of El Nio on the synchronous dengue epidemics in Thailand. PLoS Med 2(4):e106

190. Alonso, D., Bouma, M. J., and Pascual, M. (2011). Epidemic malaria and warmer temperatures in recent decades in an East African highland.Proceedings of the Royal Society B: Biological Sciences,278(1712), 1661-1669. http://doi.org/10.1098/rspb.2010.2020

191. Nagao Y, Thavara U, Chitnumsup P, Tawatsin A, Chansang C, Campbell-Lendrum D. Climatic and social risk factors forAedesinfestation in rural Thailand. Trop Med Int Health. 2003;8(7):650-9. pmid:12828549

192. Paaijmans KP, Imbahale SS, Thomas MB, Takken W. Relevant microclimate for determining the development rate of malaria mosquitoes and possible impliations of climate change. Malar J. 2010;9:196. pmid:20618930

193. Pascual, M. 2015. Climate and population immunity in malaria dynamics: harnessing information from endemicity gradients. Trends Parasitol. 31: 532-534.

194. Siraj AS, Santos-Vega M, Bouma MJ, Yadeta D, Carrascal DR, Pascual M. Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. Science. 2014;343(6175):1154-8. pmid:24604201

195. Laneri K, Bhadra A, Ionides EL, Bouma M, Dhiman RC, Yadav RS, Pascual M (2010) Forcing versus feedback: epidemic malaria and monsoon rains in northwest India. PLoS Comp Biol 6: e1000898.

196. Baeza A, Bouma MJ, Dhiman R, Pascual M (2014) Malaria control under unstable dynamics: reactive vs. climate-based strategies. Acta Tropica 129:42-51

197. Murdock CC, Blanford S, Luckhart S, Thomas MB. Ambient temperature and dietary supplementation interact to shape mosquito vector competence for malaria. J Insect Physiol. 2014;67: 3744. doi:10.1016/j.jinsphys.2014.05.020

198. Bayoh MN: Studies on the development and survival of Anopheles gambiae sensu stricto at various temperatures and relative humidities. 2001, Durham: PhD thesis. University of Durham

199. Mayne B. A study of the influence of relative humidity on the life and infectibility of the mosquito. Indian J Med Res. 1930;17: 11191137.

200. Detinova, T.S., 1962. Age-grouping methods in Diptera of medical importance with special reference to some vectors of malaria. Monogr Ser World Health Organ 47, 13-191.

201. Bayoh, M. N. and Lindsay, S. W. (2003) Bull. Entomol. Res. 93, 375-381.

202. Cazelles, B., Chavez, M., McMichael, A. J. and Hales, S. 2005 Nonstationary influence of El Nino on the synchronous dengue epidemics in Thailand. PLoS Med. 2, 313-318

203. Chaves, L. F. and Pascual, P. 2006 Climate cycles and forecasts of Cutaneous Leishmaniasis, a non-stationary vector borne disease. PLoS Med. 3, 1320-1328.

204. Bhadra A, Ionides EL, Laneri K, Pascual M, Bouma M, Dhiman RC (2011) Malaria in northwest India: data analysis via partially observed stochastic differential equation models driven by Lvy noise. J Am Stat Assoc 106: 440-451.

205. Ionides EL, Bret C, King AA (2006) Inference for nonlinear dynamical systems. Proc Natl Acad Sci USA 103: 18438-18443.

206. King AA, et al. (2015) pomp: Statistical Inference for Partially Observed Markov Processes (R package, version 1.0.0.0). Available at pomp.r-forge.r-project.org.

207. Bret C, He D, Ionides EL, King AA (2009) Time series analysis via mechanistic models. Ann App Stat 3: 319-348.

208. He D, Ionides EL, King AA (2010) Plug-and-play inference for disease dynamics: measles in large and small towns as a case study. J R Soc Interface 7: 271-283.

209. Parham PE, Michael E: Modelling the Effects of Weather and Climate Change on Malaria Transmission. Environ Health Perspect. 2009, 118 (5): 620-626.

210. Lunde TM, Bayoh MN, Lindtjrn B: How malaria models relate temperature to malaria transmission. Parasit Vectors. 2013, 6 (1): -10. 10.1186/1756-3305-6-1.

211. Yamana and E. A. Eltahir, Incorporating the effects of humidity in a mechanistic model of Anopheles gambiae mosquito poulation dynamics in the Sahel region of Africa, Parasites and Vectors, vol. 9, article 235, 2013.

212. Wigglesworth V: The Principles of Insect Physiology. 1939, London: Methuen and Co. Ltd.

213. Liu K, Tsujimoto H, Cha SJ, Agre P, Rasgon JL: Aquaporin water channel AgAQP1 in the malaria vector mosquito Anopheles gambiae during blood feeding and humidity adaptation. Proc Natl Acad Sci U S A. 2011, 108 (15): 6062-10.1073/pnas.1102629108.

214. IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment of the Intergovernmental Panel on Climate Change. 446: 727728. Cambridge and New York: Cambridge University Press.

215. Zhao L, Lee X, Smith RB, Oleson K. Strong contributions of local background climate to urban heat islands. Nature. 2014;511: 216219. doi:10.1038/nature13462

216. Zhao L, Lee X, Smith RB, Oleson K. Strong contributions of local background climate to urban heat islands. Nature. 2014;511: 216219. doi:10.1038/nature13462

217. Singh et al., 2014 D. Singh, M. Tsiang, B. Rajaratnam, N.S. Diffenbaugh Observed changes in extreme wet and dry spells during the south Asian summer monsoon season Nat. Clim. Change, 4 (2014), pp. 456-461

218. Dai, 2006 A. Dai Precipitation characteristics in eighteen coupled climate models J. Clim., 19 (2006), pp. 4605-4630

219. Mukherjee S, Aadhar S, Stone Dahiti. 2018. Increase in extreme precipitation events under anthropogenic warming in India. Weather and Climate Extremes

220. Vittal et al., 2016 H. Vittal, S. Ghosh, S. Karmakar, A. Pathak, R. Murtugudde Lack of dependence of Indian summer monsoon rainfall extremes on temperature: an observational evidence Sci. Rep., 6 (2016)