Malaria Population Dynamics and Climate Forcing Under Agricultural Development

By

Andres Baeza

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in the University of Michigan
2013

Doctoral Committee:

Professor Mercedes Pascual, Chair

Professor Bobbi Low

Professor Pejman Rohani

Professor John Vandermeer

Dedication

A mis mujeres y a mi padre

Acknowledgements

I want to thank my adviser, Mercedes Pascual, for all the guidance, dedication, encouragement, and support she provided during my time at Michigan. She has been involved in every stage of this dissertation, providing scientific insight to generate high quality science. I hope that some of these qualities of hers will someday be reflected in my own career. Being in Mercedes' laboratory has been an incredible experience, and it was truly an honor to be her student.

I'm also thankful for every member of my dissertation committee. Professors Low, Rohani, and Vandermeer were all always willing to discuss and provide feedback on this work, and they improved it with valuable comments and suggestions.

In addition, I want to thank my collaborators, especially Menno Bouma. He has been involved in every chapter of this dissertation, providing valuable comments and guidance. His knowledge about malaria and the particular situation in India is enormous. He has been very generous in sharing his knowledge with me. Also, he was a great companion on our trip together to India, where we shared many shots of whiskey in our hotel rooms while exchanging ideas, some of them now incorporated in this work. I would also like to thank Andy Dobson for providing ideas and code, and for the opportunity to visit him at Princeton to develop together the mosquito model we put in Chapter II. In addition, I am grateful to Ramesh Dhiman from the Malaria Research Institute in Delhi, India, for facilitating the collaboration and collection of the epidemiological and mosquito control data used in Chapters 2, 3, and 4, Annemarie Ter veen for kindly providing the historical data set of malaria in Mississippi, and Pietro Ceccato from the International Research Institute at Columbia University for the time he kindly spent with me when I visited him. My Labmates, specially Ed Baskerville for all the code and friendship during this period.

Finally, I want to thank my family, especially my wife Brieta, for all the support, patience, and understanding during all these good years in Ann Arbor.

Table of Content

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

The scope of this dissertation

Malaria is among the deadliest diseases of all times, having caused more deaths and misery than all the wars in the history of humanity combined. Malaria distinguishes itself from other diseases in its remarkably complex and well-adapted pathogen, with a variety of strategies to colonize human hosts in many different environments and with a long evolutionary history with humans, their ancestors, and the ancestors of their ancestors. As a consequence malaria's parasites have influenced human development and prosperity more than any other species since the beginning of civilization.

An important component of this co-evolutionary history between humans and malaria has been the development of agriculture and food production. Historically, farming settlements have influenced transmission and virulence of the malaria pathogen by providing suitable environments with multiple hosts in close proximity, and today, most malaria affliction takes place in rural agricultural communities, primarily in Africa and India.

The scope of this dissertation is to investigate the ecological associations between malaria population dynamics and economic and social prosperity in agrarian communities, as well as to understand the implication of land-use changes related to agricultural and human behavior on the persistence and distribution of the disease. The approach this dissertation takes is to combine theoretical ecology on

the disease with statistical analysis and natural history, to shed light on the feedbacks that operate between agriculture, malaria and climate, at multiple spatial and temporal scales.

The concept of a feedback loop refers to "a situation in which two or more dynamical systems are connected together such that each system influences the other and their dynamics are thus strongly coupled" (Astrom & Murray, 2008). Since the seminal work of Lotka and Volterra, these feedback loops have been known to be an important condition in regulating population growth; however, it was not until recently that they have also been recognized as an important component of the structure and stability of coupled human and natural systems (Lui et al., 2007) and specifically, of the relationship between disease persistence and economic development (Bonds and Rohani, 2010). Throughout this work the implications of this feedback are studied.

The results of this dissertation should be of relevance to public health authorities and policy makers, as they provide insights into the processes through which the economy and poverty alleviation affect malaria transmission, at the same time, that they address the implications of malaria for economic development.

Malaria natural history

Malaria Evolution and its relationship with agriculture and human behavior

Malaria is a disease caused by parasites of the genus *Plasmodium*, transmitted through a female mosquito of the genus *Anopheles*, and characterized by sexual and asexual stages in the life cycle of the parasites. When an infected mosquito bites a human, sporozoites enter the human body directly via the liver. After many cellular and nuclear divisions in the liver, the sporozoites produce approximately 15,000 to 30,000 merozoites that colonize the blood cells after a couple of days. Once

inside the blood cells, merozoites consume the hemoglobin and reproduce asexually, with the subsequent destruction of the host blood cells and colonization of new ones. It is at this stage that the actual symptoms of the disease are observed. Some of the merozoites differentiate into sexual gametocyte cells that stay in the bloodstream until a female mosquito bites the infected human, passing the gametocytes into the vector. This gametocyte then matures inside the mosquito, reproducing sexually to produce more sporozoites that eventually infect another human host.

Four species of *Plasmodium* malaria exist for the human host: *P. vivax*, *P. malariae*, *P. ovale*, and *P. falciparum*. Phylogenetic analyses have shown that the four species of human malaria are only weakly related to one another, and all have been linked to primate hosts long before humans populated the planet (Escalante and Ayala, 1995). Among the four species, *P. falciparum* is the most virulent, and is responsible for most deaths worldwide, especially in Africa.

The relationship between malaria burden and agriculture development has existed for a long time. Anthropological (Livingstone, 1957) and molecular (Rich et al., 2009) studies have suggested that the origin of the actual lineage of *P. falciparum* coincided with the development of agricultural settlement in Africa around 10,000 years ago. *P. falciparum* shows high virulence, with great production of merozoites that usually produce severe effects on the human host, including death. In addition, *P. falciparum* does not survive for long in the human host. These two characteristics suggest that a sufficiently large number of human hosts and mosquito vectors were needed for *P. falciparum* to survive in ancient times. As agricultural settlements expanded the human population, this resulted in changes in the environment influencing the ecology of the mosquito vectors, producing more mosquito-breeding sites and increasing its human-biting preference (Coluzzi, 1999). As the human population increased and mosquitoes adapted to human meals, *P. falciparum* also adapted to this new

setting, thereby selecting for strains of high virulence (Livingstone, 1957; Carter and Mendis, 2002; Rich et al., 2009).

With this new parasite, the human population also evolved genetic mechanisms to counteract the burden of the disease. The most striking adaptation to *P. falciparum* is sickle cell anemia, which resulted from a single point mutation in the gene for the beta chain of hemoglobin, which in heterozygotes provides protection against *P. falciparum*, but produces another disease, anemia, in the homozygote (Carter and Mendis, 2002). This reflects the strength of the burden of the disease in the early stages of human societies. In this context, it is likely that malaria affected the economy of the earliest farming populations (Packard, 2007). In some areas malaria was probably a barrier to the development of agriculture, decreasing the chances of survival and a sedentary lifestyle. This situation probably influenced the migration of people to areas that were free from malaria, for example to locations at high altitudes. In others areas, human adaptations counteracted the burden of the disease with selective mutations and protective immunity, helping humans to increase productivity and thus population size. In both cases, it is clear that the interaction among vector, parasite, and human host has shaped the present distribution and burden of malaria around the world (Hume et al., 2003).

Since the discovery of the *Plasmodium* parasite that causes the disease, and the contribution of the *Anopheles* vector to transmitting the parasite from human to human at the beginning of the XX century, great progress has been made in understanding the intrinsic and extrinsic factors affecting the distribution of the disease under different ecological scenarios. The population size of vector mosquitoes determined by climatic conditions, the type of *Plasmodium* species, and the different degrees of immunity of the population shape the dynamics of the disease.

Among the most important figures in malaria research in the early days was Sir Ronald Ross. He was one of the first (among others) to propose the vector transmission route, and he was a strong and early supporter of the importance of mosquito control for economic reasons, even before the mosquito hypothesis was completely established (Ross, 1905). Besides demonstrating the mosquito hypothesis himself (Ross, 1897), he also contributed significantly to the quantitative theory of malaria. In 1904 he published his first mathematical model to explain mosquito movement and the scale of larvae control, and he published his first malaria transmission model by 1908 (Ross, 1911). His work motivated and influenced the new generation of malariologists in the use of quantitative analyses to understand disease transmission and the causes of epidemics (Smith et al., 2012).

The climate influence

The recognition of the influence of climatic factors, and specifically rainfall, in determining epidemic seasonality and interannual variability was rapidly recognized soon after the discovery of the transmission route. This knowledge was quickly implemented by malaria public health officers who promptly absorbed this biological knowledge to generate the prediction of epidemic events. Among the pioneers in this area was Sir R. Christopher. While working as an officer in the Medical and Sanitary Department in the Punjab, Christopher witnessed the worst famines that occurred in the area during the late XIX century, including the largest malaria epidemics on record in the region (deZulueta et al., 1994). This experience shaped his perspective on the causes and the routes by which climatic factors, specifically monsoon rainfall and the lack of it, could influence the magnitude of the epidemics. In particular he was among the first to generate predictions of malaria epidemics based on climatic information, and to develop an early warning system in the region (Swarrop, 1949; Bouma and vanderKaay, 1996). In addition, the drought-driven famines helped him to recognize the importance of agriculture and food security in shaping malaria epidemics in India. In fact, his early

predictions included not just rainfall, but also the prices of the important crops as indicators of human susceptibility (Christopher, 1911).

The eradication era

Another important figure that shaped the malaria research agenda during the second part of the XX century was G. Macdonald. His work, summarized in his book "The Epidemiology and Control of Malaria" (Macdonald, 1957), is still a fundamental piece in our understanding of malaria population dynamics. This work set the stage for most of the subsequent efforts on malaria epidemiology up to the present day, and its mathematical formulation pioneered the use of compartmental mathematical and computational models, now widely used not just in malaria research but also for many other diseases (Smith et al., 2010). From a practical perspective, Macdonald's theory helped to define the scientific basis for the design of the Global Malaria Eradication Program (GMEP) between 1955 and 1969. First, the theory provided rigorous quantification of the limits of transmission through the well-known number R_0, and second, it generated predictions of the effects that different methods of control would have on R_0, making an strong case for interventions that would decrease the contact rate between mosquitoes and humans, a.

With Macdonald's mathematical theory of malaria control, and the massive arsenal of insecticides produced after World War II, over-optimism that the disease would be eradicated within a relatively short period of time was generated in the public health sector and governments of affected regions (Najera, 2010). The GMEP global eradication campaign was launched by the World Health Organization, and DDT-based mosquito control intervention was the main tool used to fulfill the goal of eradication. The strategy of the campaign was simple: a short war against mosquitoes that would last less than 20 years (Najera, 2010). The campaign was divided in different stages, from evaluation to confirmation of elimination. A large surveillance and monitoring system was put in

place in the different countries whose function was to help define the areas most in need of intervention, and to evaluate the success of the phases of the campaign in order to plan subsequent stages. During the following decades, most of the financial efforts for malaria were put into the campaign in Africa and Asia and limited funds were spent on research, leaving an immense gap in knowledge that, in retrospect, had severe consequences for the control of the disease in the following decades (Packard, 2007).

The resurgence

By the mid-1970, mosquito resistance to DDT was reported in many countries, and by the end of the decade the resurgence of malaria was evident (Packard, 2007). It was clear by then that the GMEP was not reaching the objectives proposed, and funding for the campaign shrunk considerably. Eventually the GMEP was abandoned, creating a considerable backlash in malaria research and its control. India, for example, had considerable success in decreasing malaria mortality 10 fold during the GMEP, but experienced large epidemics in the 1970s, causing a considerable burden to the country that persists today (Sharma, 1999).

It was clear by the 1980s that a different approach for fighting malaria was needed in light of the global resurgence. With a new focus on controlling the disease rather than achieving its eradication, new malaria control programs were put in place in many countries, and new global initiatives were created to generate and coordinate actions at a global scale. Consequently, a new research agenda was proposed to confront the resistance of the mosquito, the parasite's evolutionary way to escape detection by the human immune system, and to consider the increased recognition of the influence of socioeconomic factors. In addition, a revitalization of the importance of including climate factors, the ecology of the mosquito, and the imminent threat of global warming are now part of the new agenda. Today, initiatives such as the Roll Back Malaria partnership has attracted the participation of

more than 500 partners, including endemic countries, private sector nongovernmental and community-based organizations, and the attention of large foundations and research institutions.

Poverty and malaria

One of the lessons learned after the GMEP, which has been applied to this new strategy, is the need to consider the complex ecological and evolutionary roots of the disease and its interwoven connection with human population activities, specifically the interface between poverty, agricultural practices, and malaria risk (Kitron, 1987). An indisputable current malaria pattern around the world is the relationship between malaria and poverty distribution at the global scale. More than 70% of all cases occur in rural communities usually living under extreme economic vulnerability, and its burden is suggested to be the cause of a decrease of more than 3% of GDP in malaria-ridden countries (Sachs and Malaney, 2002).

While this striking relationship between malaria and rural poverty has been long recognized, the influence of agricultural development and land-use changes on the dynamics of malaria and the contribution of malaria to maintaining poverty are still poorly understood, mostly because of the complex interaction with other extrinsic factors such as rainfall and temperature, affecting both the vector and the parasite as well as food production, and the limited long-term surveillance data in both malaria epidemiology and agricultural dynamics.

A particularly important factor in the malaria-agricultural system is water distribution for crop production and the modification of the environment due to infrastructure to generate irrigation potential. Many entomological and parasitological studies have addressed this issue and, not surprisingly, the literature is in tune with this dichotomy: some studies show a positive relationship between irrigation, vector population size, and malaria occurrence, while others report that an increment in irrigation regimes increase vector carrying capacity, though not necessarily increase

malaria occurrence. This apparent dilemma between irrigation and malaria risk, defined as the *paddy paradox* (Ijumba and Lindsay, 1996), remains poorly understood.

Outline of the dissertation

In this dissertation I present a series of papers in which I try to disentangle the contribution of land-use changes related to agricultural and irrigation development to malaria population dynamics and its elimination, under the underlying influence of climate forcing. The main goal of these projects was to provide ecological insight on the feedback mechanisms that operate among agriculture, socioeconomic development, human behaviour, and malaria, in order to gain an understanding of their dynamical consequences on the spatial and temporal distribution of the disease. Chapters II, III, and IV consider malaria in Northwest India, in the States of Gujarat and Rajasthan, in the last two decades. Chapter V presents a historical study of malaria in the United States, in the State of Mississippi between 1917 and 1940, and relies on data from the elimination era.

Chapter two: Malaria and climate forcing. The Effect of irrigation

Malaria transmission in Northwest India is epidemic and unstable, and usually cases occur only for a few months after the monsoon season. In this area the influence of climate, particularly rainfall, has been well-known since the beginning of the twentieth century, and this connection provides valuable information for an early warning of epidemic risk. However, large scale land-use changes in irrigation and agricultural development are shifting the semi-desert environment to one in which water is more accessible and provides increasing suitable areas for mosquito breeding. The influence of these changes on the malaria-climate relationship remains poorly understood. In this chapter I provide evidence for the decoupling effect of irrigation on the predictability skill of malaria epidemics based on large-scale regional climate rainfall. In addition I present evidence for more

endemic malaria in irrigated areas. Finally, I suggest that the decoupling of the malaria-rainfall association is due to mosquito control intervention.

Chapter three: Malaria and its control in highly variable environments

A legacy from the GMEP in India is the use of the Annual Parasitic Incidence (API, cases/1000 population) to determine and prepare subsequent area interventions. Today, in accordance with the technical criteria of the national Vector-borne Disease Control program, only those villages that report 2 API or more, qualify for mosquito control intervention based on insecticides. The implication of the use of this policy in the population dynamics of malaria epidemics is still not well understood. In Chapter III, I explore the dynamical consequence of control policy and the vector-control application in India under high environmental variability. This chapter shows how including this mechanistic feedback between previous malaria incidence, the perception of the risk, and the implementation of control influences the effectiveness of the control strategy and the efficacy of reaching elimination. It is this feedback related to human intervention and not immunity that generate these interannual cycles.

Chapter four: Long-lasting transition toward malaria elimination under irrigation development

Irrigation and water storage infrastructure is known to increase malaria risk by increasing mosquito breeding sites. However, over a longer time scale, irrigation provides food security and an increase in income that eventually generates the possibility for individuals to control the disease. In Chapter IV, I examine simultaneous changes in malaria risk and irrigation-related land-use changes over more than three decades. In this chapter, I describe and characterize a transitional stage in malaria risk that occurs in the early stages of irrigation. I provide evidence to suggest this transition is long lasting and temporary state of heightened risk that is followed by a decrease and eventual elimination of risk when socioeconomic health is improved by the benefits provided by irrigation. The finding of this

work provides a strong case for including health impact assessment in irrigation projects and water storage infrastructure for longer periods of time.

Chapter five: A time series model to test the double causality between malaria and income in Mississippi, United States

Malaria today is a disease of poverty and its distribution is more common among populations with low socioeconomic conditions (Gallup and Sachs, 2001). This poverty-malaria pattern remains the subject of debate in the scientific community: while some authors argue that malaria is a major determinant of population wealth (Malaney et al., 2004, Gallup and Sachs 2001), others claim that socioeconomic conditions are instead what drives disease levels (Packard, 2007). A better understanding of the route by which disease and socio-economic conditions influence each other and the main direction of causation has implications for the implementation of policies and plans for malaria control and elimination. In Chapter V, I investigated retrospectively the malaria dynamics in Mississippi, United States, during the first part of the XX century when the disease was on the path to elimination. I statistically tested the double causality hypothesis between malaria and socioeconomic conditions. This work is a first attempt to statistically and dynamically test the poverty trap hypothesis. Results provide evidence for (1) the larger effect of income and purchase power for malaria control on malaria's burden, and (2) the relatively low influence of malaria on the cotton-based productivity of Mississippi.

References

Astrom, K.J., Murray, R., 2008. Feedback Systems: an introduction for scientists and engineers. Princeton University Press, Princeton, NJ.

Bonds, M.H., Keenan, D.C., Rohani, P., Sachs, J.D., 2010. Poverty trap formed by the ecology of infectious diseases. P R Soc B 277, 1185-1192.

Bouma, M.J., vanderKaay, H.J., 1996. The El Nino Southern Oscillation and the historic malaria

epidemics on the Indian subcontinent and Sri Lanka: An early warning system for future epidemics? Trop Med Int Health 1, 86-96.

Carter, R., Mendis, K.N., 2002. Evolutionary and historical aspects of the burden of malaria. Clin Microbiol Rev 15, 564.

Christophers, R., 1911. Malaria in the Punjab, in: India, G.o. (Ed.), Scientific memoirs by officers of the medical and sanitary departments. Superintendent Government Printing, Calcutta.

Coluzzi, M., 1999. The clay feet of the malaria giant and its African roots: hypotheses and inferences about origin, spread and control of Plasmodium falciparum. Parassitologia 41, 277-283.

de Zulueta, J., Mujtaba, S.M., Shah, I.H., 1980. Malaria control and long-term periodicity of the disease in Pakistan. Trans R Soc Trop Med Hyg 74, 624-632.

Escalante, A.A., Barrio, E., Ayala, F.J., 1995. Evolutionary Origin of Human and Primate Malarias - Evidence from the Circumsporozoite Protein Gene. Mol Biol Evol 12, 616-626.

Hume, J.C.C., Lyons, E.J., Day, K.P., 2003. Human migration, mosquitoes and the evolution of Plasmodium falciparum. Trends Parasitol 19, 144-149.

Hume, J.C.C., Lyons, E.J., Day, K.P., 2003. Malaria in antiquity: a genetics perspective. World Archaeol 35, 180-192.

Ijumba, J.N., Lindsay, S.W., 2001. Impact of irrigation on malaria in Africa: paddies paradox. Med Vet Entomol 15, 1-11.

Kitron, U., 1987. Malaria, Agriculture, and Development - Lessons from Past Campaigns. Int J Health Serv 17, 295-326.

Livingstone, F.B., 1957. Sickling and Malaria. Brit Med J 1, 762-763.

Liu, J.G., Dietz, T., Carpenter, S.R., Alberti, M., Folke, C., Moran, E., Pell, A.N., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C.L., Schneider, S.H., Taylor, W.W., 2007. Complexity of coupled human and natural systems. Science 317, 1513-1516.

Macdonald, G., 1957. The Epidemiology and Control of Malaria. Oxford University Press, London.

MacDonald, G., Cuellar, C., Foll, C., 1968. The Dynamics of Malaria. Bulletin of the World Health Organization 38, 743-755.

Najera, J.A., 1989. Malaria and the work of WHO. Bulletin of the World Health Organization 67, 229-243.

Najera, J.A., Gonzalez-Silva, M., Alonso, P.L., 2011. Some lessons for the future from the Global Malaria Eradication Programme (1955-1969). PLoS Med 8, e1000412.

Packard, R.M., 2007. The Making of a Tropical Disease, a Short History of Malaria. Johns Hopkins University Press, Baltimore.

Packard, R.M., 2009. "Roll Back Malaria, Roll in Development"? Reassessing the Economic Burden of Malaria. Popul Dev Rev 35, 53.

Reiner, R.C., Perkins, T.A., Barker, C.M., Niu, T.C., Chaves, L.F., Ellis, A.M., George, D.B., Le Menach, A., Pulliam, J.R.C., Bisanzio, D., Buckee, C., Chiyaka, C., Cummings, D.A.T., Garcia, A.J., Gatton, M.L., Gething, P.W., Hartley, D.M., Johnston, G., Klein, E.Y., Michael, E., Lindsay, S.W., Lloyd, A.L., Pigott, D.M., Reisen, W.K., Ruktanonchai, N., Singh, B.K., Tatem, A.J., Kitron, U., Hay, S.I., Scott, T.W., Smith, D.L., 2013. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010. J R Soc Interface 10.

Rich, S.M., Leendertz, F.H., Xu, G., LeBreton, M., Djoko, C.F., Aminake, M.N., Takang, E.E., Diffo, J.L., Pike, B.L., Rosenthal, B.M., Formenty, P., Boesch, C., Ayala, F.J., Wolfe, N.D., 2009. The origin of malignant malaria. P Natl Acad Sci USA 106, 14902-14907.

Ross, R., 1897. On some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood. Brit Med J 2, 1786-1788.

Ross, R., 1905. The logical basis of the sanitary policy of mosquito reduction. Science 22, 689-699.

Ross, R., 1911. Some quantitative studies in epidemiology. Nature 87, 466-467.

Sachs, J., Malaney, P., 2002. The economic and social burden of malaria. Nature 415, 680-685.

Sharma, V.P., 1996. Re-emergence of malaria in India. Indian J Med Res 103, 26-45.

Smith, D.L., Battle, K.E., Hay, S.I., Barker, C.M., Scott, T.W., McKenzie, F.E., 2012. Ross, Macdonald, and a Theory for the Dynamics and Control of Mosquito-Transmitted Pathogens. Plos Pathog 8.

Swaroop, S., 1949. Forecasting of Epidemic Malaria in The Punjab, India. The American Journal of Tropical Medicine and Hygiene 29, 1-17.

Ulrich, J.N., Naranjo, D.P., Alimi, T.O., Muller, G.C., Beier, J.C., 2013. How much vector control is needed to achieve malaria elimination? Trends Parasitol 29, 104-109.

Chapter 2

Malaria and climate forcing: The Effect of irrigation

Abstract

Background: Rainfall variability and associated remote sensing indices for vegetation are central to the development of early warning systems for epidemic malaria in arid regions. The considerable change in land-use practices resulting from increasing irrigation in recent decades raises important questions on concomitant change in malaria dynamics and its coupling to climate forcing. Here, the consequences of irrigation level for malaria epidemics are addressed with extensive time series data for confirmed Plasmodium falciparum monthly cases, spanning over two decades for five districts in north-west India. The work specifically focuses on the response of malaria epidemics to rainfall forcing and how this response is affected by increasing irrigation.

Methods and Findings: Remote sensing data for the Normalized Difference Vegetation Index (NDVI) are used as an integrated measure of rainfall to examine correlation maps within the districts and at regional scales. The analyses specifically address whether irrigation has decreased the coupling between malaria incidence and climate variability, and whether this reflects (1) a breakdown of NDVI as a useful indicator of risk, (2) a weakening of rainfall forcing and a concomitant decrease in epidemic risk, or (3) an increase in the control of malaria transmission. The predictive power of NDVI is compared against that of rainfall, using simple linear models and wavelet analysis to study

the association of NDVI and malaria variability in the time and in the frequency domain respectively.

Conclusions: The results show that irrigation dampens the influence of climate forcing on the magnitude and frequency of malaria epidemics and, therefore, reduces their predictability. At low irrigation levels, this decoupling reflects a breakdown of local but not regional NDVI as an indicator of rainfall forcing. At higher levels of irrigation, the weakened role of climate variability may be compounded by increased levels of control; nevertheless this leads to no significant decrease in the actual risk of disease. This implies that irrigation can lead to more endemic conditions for malaria, creating the potential for unexpectedly large epidemics in response to excess rainfall if these climatic events coincide with a relaxation of control over time. The implications of our findings for control policies of epidemic malaria in arid regions are discussed.

Introduction

The response of epidemic malaria to large-scale change in land-use practices related to irrigation and agriculture in arid regions remains poorly understood (Tyagi, 2004). In the last three decades, for example, the expansion of a large network of irrigation canals has supplied an important source of freshwater for agriculture in many arid regions of India; in so doing, it has also contributed to the economic development of these regions. More generally, change in irrigation schemes, and associated agricultural practices, are considered among the potential drivers underlying malaria's increasing global burden (Sachs and Malaney 2002), but their consequences remain poorly understood given the complexity of their effects on transmission via human wealth and vector ecology. In particular, it is not clear how irrigation is modifying the coupling of epidemic malaria to rainfall variability in arid regions.

The population dynamics of malaria at the edge of its distribution, in either deserts or highlands, where rainfall and temperature respectively limit transmission, are characterized by strong seasonality and significant variation in the size of outbreaks from year to year (Yacob and Swaroop 1949; Christophers, 1911; Thomson et al., 2006). In these regions the role of climate forcing is potentially central to the prediction of inter-annual variability of epidemics. The high variability in the number of cases between years challenges public health efforts, as severe intermittent epidemics can strain medical facilities.

In the north-west of India, there has been a long-standing interest in the development of early-warning systems based on rainfall (Gill, 1923; Swaroop, 1949) and economic conditions (Zurbrigg, 1994); this has regained significance in the last decades following the failed eradication attempts in the 1960s and '70s. Epidemics have re-emerged in the desert states of Rajasthan and Gujarat, and have once again motivated interest in climate forcing (Bouma and vanderKaay 1996; Akhtar and McMichael, 1996) and its interplay with socio-economic factors. An increment in burden has been attributed to the extension of the canal network that provides water for regional agriculture (Tyagi and Yadav, 2001).

Despite the potential of remote sensing for the generation of early-warning systems (Rogers et al., 2002), efforts have been largely focused on defining malaria's spatial niche and seasonal timing (Thomson et al., 1999; Hay et al., 1998) rather than on predicting the seasonal burden in areas of unstable malaria (but see (Ceccato et al., 2005) for dry regions of Eritrea). Because vegetation can be used as a proxy for the amount of water in the ground and, therefore, for the humidity of the environment, the Normalized Difference Vegetation Index (NDVI) provides a spatially-explicit link between rainfall and malaria from local to regional scales. Thus NDVI patterns are especially relevant for investigating the relationship between epidemic events and regional climatic drivers in desert regions [Connor et al., 1998; Thomson et al., 1997]. However, any vegetation index will be

susceptible to landscape changes due to irrigation and agricultural practices. The extensive time series of malaria cases in desert and semi-desert districts of Rajasthan and Gujarat provide an opportunity to examine how NDVI-malaria associations change across regions that represent a gradient in levels of irrigation.

This paper addresses how the level of irrigation modifies the dynamics of malaria across such a gradient and how irrigation affects the predictability of epidemics based on climatic variables. It further investigates whether a decoupling of rainfall forcing and malaria epidemics with increased levels of irrigation reflects (1) a real reduction in the risk of transmission, (2) more effective control efforts but no reduction in disease risk, or (3) a breakdown of NDVI as a sensitive indicator of rainfall's inter-annual variability. These different mechanisms have very different implications not just for prediction but for the possibility of unexpected epidemics when control efforts fail. Statistical analyses are used to address these different hypotheses, and the findings are discussed in light of a simple dynamical model of mosquito abundance and irrigation. The data provide additional empirical evidence for the possibility of surprises in years when intervention and monsoon rains vary simultaneously but in opposite directions. This suggests that control policies based on residual insecticide spraying, whose planning is "reactive" to disease levels in the previous season, need to be modified given the known consequences of past failures and relaxation of control in irrigated areas. The implications of these findings for forecasting and control policies in other arid regions with epidemic malaria are also discussed.

Methods

Malaria and remote sensing data

Epidemiological data

The epidemiological data consists of time series of monthly confirmed cases of Plasmodium falciparum from five districts in north-west India (Fig. 2.1): Bikaner and Barmer in Rajasthan, and Kutch, Kheda, and the combined area of Banaskantha, Mehsana, and Patan (hereafter, BMP) in Gujarat state. The data from BMP was combined into one dataset to circumvent problems related to modifications of boundaries over the time of this study, and the resulting separation into different districts of what was originally a single administrative unit at the beginning of the data collection (1976). For Bikaner, BMP and Kheda, the data represent 30 years of observations (1976-2009), and for Barmer, 24 years (1985-2009 with a gap between 2004 and 2006). For each time series the monthly cases were divided by the total yearly human population of the respective district, and all analyses relied on this normalized incidence data (See Fig. S2.1). The blood samples were collected by passive surveillance of patients that visited their local health facility, and by active surveillance of patients with fevers in house-to-house visits. These data were obtained from the offices of the Joint Director, Vector Borne Diseases, Commissionaire of Health Rajasthan and Gujarat. Information on the level of irrigation was extracted from statistical abstract books from the states of Gujarat and Rajasthan between 1979 and 2006. For each district the percentage of agriculture under some type of irrigation was computed by dividing the number hectares of agriculture by the number of hectares of land under irrigation (Fig. 2.1). These quantities were then ranked by district based on these irrigation levels. The resulting ranking remains unchanged for the whole time period covered by this study.

The Normalized Difference Vegetation Index

NDVI is defined as the difference in radiation reflected by any surface in two bands of the energy spectrum - the infrared and the red band. From this index, which ranges from -1 to 1, it is possible to discriminate between radiation reflected by vegetation and other surfaces. Values greater than 0.2 quantify vegetation greenness from different sources and/or from different seasons. The NDVI

data utilized in this analysis was obtained from two sources that cover respectively two different time periods. The first one is the Global Inventory Modeling and Mapping Studies (GIMMS; Tucker et al., 2005). The original source of this product is a combination of observations made by different NOAA missions carrying the Advanced Very High Resolution Radiometer (AVHRR). This dataset has a temporal extension ranging from 1982 to 2006 on a bi-monthly basis, and a spatial extension covering the entire globe at a resolution of 8 km. The data have been corrected to avoid distortions and to show only positive values (Tucker et al., 2005). A window area that includes the five districts from which we extracted the vegetation information (19.18°N - 29.15°N and 66.98°E - 78.55°E; Fig. 2.1) was defined for the years with epidemiological information up to 2006. The second source corresponds to the NDVI product from the Moderate Imaging Spectroradiometer (MODIS) in the TERRA satellite. The data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC) located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center. Specifically, all the analyses with MODIS rely on the monthly average product at 1 km resolution for the month of September from 2000 to 2009.

The statistical analyses focused on the longer NDVI dataset from NOAA (referred to hereafter as NDVI) and corroborated that similar results for the spatial correlation maps (see below) were obtained with the shorter but more recent MODIS NDVI.

Rainfall Data

Monthly accumulated rainfall records were acquired to compare predictability based on rainfall to that based on NDVI. The data, which span 20 years, were obtained from one local station within each district and supplied by the Indian Meteorological Department in Pune (India).

Statistical and numerical analysis

The Signature of NDVI

Irrigation increases water available for agriculture, making yearly multi-crop rotations possible. The differences in water supply and multi-crop rotations, therefore, should be reflected in the amount of vegetation and in the signature of the NDVI time series. In highly irrigated areas we should observe less inter-annual variability than in non-irrigated areas. At the same time, the seasonality of NDVI should also vary across districts as different peak times can result from different crop seasons supported by irrigation. Here, both the coefficient of variation of NDVI for each grid-point (pixel) in the study area and the month at which NDVI shows the largest average value (peak) were calculated as a way to quantify the regularity and seasonality over time for each location in the study area. The coefficient of variation (CV) is a dimensionless measure of the variability of a quantity with respect to its mean value. Therefore, CV allows us to compare different districts independently from their mean value.

Correlation Maps

Since malaria time series exhibit strong seasonality with a peak of cases between October and December (Fig. S2.2), the association between malaria and NDVI was examined for the monthly cases of each of the months of the epidemic season (October, November and December) and NDVI at a given preceding month for each year. Spearman rank correlation was used as a non-parametric measure of association between incidence in a given month (say October) and NDVI (say in September). A map of the NDVI/malaria correlation was obtained by computing this correlation repeatedly for each 8 × 8 km grid point in the study area. These maps provide one way to discern the hypothesis that rainfall, as a regional phenomena, no longer acts as a driver from the alternative that it continues to do so, but it is poorly reflected in the local NDVI. For this purpose, a large regional box overlapping the Thar desert and the area within the boundaries of a given district are relevant. The former provides information on NDVI at a large regional level (and, therefore, climate), and the latter corresponds to a more local level under the influence of district land-use and

irrigation practices. The resulting correlation map shows the correlation coefficient only for those grid points in which its value is statistically significant ($p < 0.05$ for NOAA and $p < 0.1$ for MODIS). These maps were obtained for both NDVI products.

In addition, time series were constructed for each district for the average NDVI over an area of approximately 576 km2, selected because it exhibited the highest rank correlation within the given district. These time series were then used to fit parametric linear regression models of malaria cases for a given epidemic month as a function of NDVI in the preceding months. Since rainfall has been of interest as a predictor variable in early-warning systems for desert malaria, the proportion of the variance explained by NDVI was compared here to that explained by accumulated rainfall, for every month preceding the epidemic season. This allowed a comparison of the predictive skill of NDVI to that of rainfall, and an examination of the respective delays between NDVI, rainfall, and epidemics. Rainfall was accumulated over the monsoon months based on previous results indicating the usefulness of the resulting quantity in malaria transmission models for this region (Laneri et al., 2010).

Wavelet analysis

In order to further investigate the relationship between NDVI time series and malaria, the spectral signature of both times series was obtained using wavelet analysis. Wavelet analysis is particularly well suited for studying the dominant periodicities of epidemiological time series because of the non-stationary nature of disease dynamics (Cazelles et al., 2007; Pascual et al., 2008). By contrast to the standard Fourier power spectrum, which provides a global analysis (over time) of the dominant frequencies of a time series, the wavelet spectrum is local in time, providing the additional information of when a specific frequency is present (for a detailed explanation of the method and the interpretation of the results see (Laneri et al., 2010; Cazelles et al., 2007) and (Pascual et al., 2008) for an epidemiological application). Here the implementation developed by Cazelles and

collaborators (Cazelles et al., 2007) was used which includes the assessment of statistical significance based on bootstraps methods. For each district the wavelet power spectrum was calculated for both the NDVI and the malaria time series.

Model and simulations

In order to interpret and discuss the results of the statistical analyses, a simple model of malaria risk was also developed that encapsulates basic elements of the relationship between rainfall, irrigated agriculture and mosquito population abundance (Appendix 2.1: A simple model of mosquito population dynamics, rainfall and irrigation). The model consists of a set of differential equations describing how rainfall water is allocated to different compartments of the landscape (See Fig. S2.3). Simulations of the model are used to examine the consequences of increasing irrigated area for (1) the seasonal and inter-annual correlation between rainfall and mosquito abundance and (2) disease risk measured in terms of mosquito abundance.

Results

The analyses of the variability and seasonality of the spatially explicit NDVI time series show that NDVI inside BMP and Kheda have a very low coefficient of variation (CV) for a large fraction of grid points, and that the average peak of NDVI occurred more often in the months of January and February (reflecting irrigation for the winter crops). By contrast Barmer and Bikaner both show a higher CV for most locations, and the peak vegetation months fall after the monsoon rains and preceding the epidemic season (in September and October). Thus, the temporal dynamics of NDVI differ across the districts and are altered in the presence of irrigation and associated agriculture (See Fig. S2.4).

Spatial correlation maps (Fig. 2.2) show that the prevalence of malaria in Barmer and Bikaner, the two districts with the lowest irrigation values, has a strong and significant positive association with September's NDVI for a large region including parts of the Thar Desert. Thus, high values of NDVI in the arid zone precede the observation of an important portion of cases in the consecutive months. In the adjacent region with higher levels of irrigation, BMP, only a small part of the district shows a positive and significant correlation with NDVI; although an association can still be observed with regional NDVI from areas outside the district. This suggests that NDVI has weakened as an indicator of rainfall variability but that this variability continues to act to some extent as a driver of malaria at inter-annual time scales. At the highest level of irrigation, Kheda showed a different pattern: no significant correlation with NDVI is present for any location of the study area, inside or outside the district. In this district, rainfall variability appears to no longer act as a driver of epidemic size. Similar results are obtained with the more recent NDVI product, MODIS (See Fig. S2.5). The parametric linear models show similar results. Barmer, Bikaner and Kutch have a high and significant correlation with a coefficient of determination $R2>0.7$, whereas for Kheda and BMP, this association weakens considerably (Fig. 2.3).

A similar picture emerges from the analysis in the frequency domain. Figure 2.4 shows the power spectrum for both the malaria and NDVI time series for the two districts corresponding to the extremes of the irrigation spectrum. Results show that for Barmer, malaria and NDVI both show strong and significant power in the 2 and 4-year periods. On the other extreme of the irrigation intensity, the spectrum for Kheda shows that most of the variance is concentrated in the one year period, as expected for seasonal malaria (epidemics), and in the five year period with this multiyear cycle absent in NDVI. NDVI shows some power for period 4, but for a small window of time (between 1992 and 1998). A general pattern observed in the power spectrum of both NDVI and malaria for most all the districts is that as irrigation increases, the time interval during which the

seasonal signal (at period one) predominates increases with irrigation (see Fig. S2.7). In Barmer, for example, the annual signal is absent for considerable extents of time, with the exception of a few years in the early nineties. By contrast, this signal is present in Kheda almost uniformly over time as the dominant scale of variability. Concurrently, the times with variance concentrated around period 2 decrease as the extent of agricultural land under irrigation expands. This is also true for NDVI with the exception of Barmer, which shows less activity than the other districts in this particular frequency. Although the variance of Kheda is concentrated in the seasonal cycle, the average seasonal pattern of malaria incidence in this district also shows a less pronounced seasonality, in the sense that the troughs, in the inter-epidemic months, exhibit higher values than in less irrigated districts (See Figure S2.2). At the same time, on average, the seasonal incidence during epidemic months in Kheda is higher (and not lower). Thus, when epidemics occur, disease burden tends to be larger than in the less irrigated districts. This is the case both in earlier and more recent times (See Fig. 2.1 and Table S2.1 for the large outbreaks of 2004).

A comparison between NDVI and rainfall predictability shows that NDVI is a better explanatory variable at a lead time of one month (See Fig. S2.7). In August, the best predictor month for rainfall, NDVI performed equally well for Barmer, Bikaner, and Kutch. This comparison also shows that for BMP, rainfall is a better predictor than NDVI, and that in Kheda neither rainfall nor NDVI can predict the size of epidemic events, confirming the decoupling of the malaria system from the annual effect of rainfall in the number of cases.

Consistent with the above results, a dynamical model (see Methods and Fig. S2.3) illustrates that mosquito abundance increases as the area of the landscape under irrigated agriculture increases (See Fig. S2.8, panel B). However, this increment is only observed for the months in which irrigation is supplied to the agriculture (the driest months) and not during the epidemic season. This would result

in less pronounced troughs between epidemics and more sustained transmission. It also follows from this that irrigation cannot modify the inter-annual correlation between rainfall and mosquito abundance in the model (See Fig. S2.9, panel D). Thus, disease risk remains associated with excess rainfall, in the absence of additional factors such as intervention efforts or the increase in levels of population herd immunity with higher transmission.

Discussion

The resurgence of malaria in desert and semi-desert areas of Rajasthan and Gujarat over the last three decades once again underscores the potential relevance of an early-warning system based on climate variability. Any new predictive tool must now also consider the considerable changes in land-use patterns that arise from irrigation and agriculture. More generally, the long-term surveillance programmes in India provided an opportunity to address climate forcing in the context of changing irrigation patterns, a problem of relevance to desert malaria in other continents. The results suggest an increasing decoupling of malaria with local and then regional rainfall variability that possibly reflect more effective control measures, rather than a real reduction in potential risk. This has clear implications for control policies as discussed below.

This work specifically examined the relationship between a remote sensing index of vegetation (NDVI) and malaria in districts that differ considerably in levels of irrigation. September NDVI was shown to provide a reliable predictor of malaria prevalence during the epidemic period (October, November and December) both within the district and in the larger region; however, this prediction is strongest in districts where irrigation is low or completely absent. These results provide a bridge between predictions at a somewhat longer lead time, in July and August, based on rainfall and the actual peak of cases in October, November and December. For the non-irrigated districts,

September NDVI explained the inter-annual variation in the size of malaria epidemics better than the accumulated rainfall of August and September. The coherence of malaria and NDVI in the frequency domain, with 2 and 4-year cycles, further confirms the importance of the monsoon rains as a driver of malaria's inter-annual cycles in these arid and un-irrigated regions.

In contrast, in the highly irrigated districts, Kheda and in the combined Banaskantha, Mehsana, and Patan (BMP), the association between NDVI and malaria weakens and NDVI provides a poor predictor of the magnitude of an outbreak. For BMP this is only the case for NDVI within the district and not at larger regional scales. This suggests a persistent role of rainfall forcing on malaria transmission, with irrigation and agriculture mainly compromising how well NDVI reflects this variability within the district. For Kheda, the more extreme breakdown of the association with NDVI both inside and outside the district indicates that rainfall no longer determines the size of epidemics from year to year.

Two hypotheses arise as possible explanations of this breakdown: 1) an increase in wealth with irrigation would underlie more effective intervention measures based on residual insecticide spraying, which in turn would prevent epidemics under anomalous and large monsoons. More irrigated, wealthier districts would therefore experience a greater decoupling than poorly or non-irrigated ones. Or 2) the importance of rainfall has diminished due to changes in either the entomological and ecological conditions underlying the vectors' dynamics, or climatic conditions. In particular, multi-crop systems supported by irrigation and rainfall can continually provide the water necessary for a suitable mosquito environment. For example, a study in Pakistan's Punjab region (Johansson et al., 2009), where malaria used to exhibit pronounced monsoon driven epidemics, described how irrigation now provides perennial sites for mosquito breeding, with different vector species breeding at different seasons during the year. On longer time scales, both mechanisms might be at play, with the development of irrigation ultimately resulting in more extensive ecological and socio-economic

change and a more permanent reduction of malaria risk (Zurbrigg, 1994). This implies that similar effects may occur in Kheda, which currently exhibits similar epidemic timing, but with a higher disease incidence (when epidemics occur) than in other districts, suggesting that disease risk following anomalous monsoons has not decreased with irrigation, but that intervention measures have prevented its manifestation. Indoor Residual Spray (IRS) in these districts is planned at the village level based on the incidence of malaria in the previous year. All the dwellings in villages whose annual parasite incidence exceeds 2 per thousand population are targeted for spray in the next year and preceding the epidemic season.

Higher levels of irrigation and associated wealth would underlie more effective implementation of planned IRS, and in so doing keep malaria transmission in check at the same time that malaria risk itself would remain responsive to climate variability (and even increase). This creates the possibility of surprises, with unexpected large epidemics when a temporary failure of control coincides with an anomalous monsoon year. An example of this is suggested by data on residual insecticide spraying in Kheda (See Table S2.1). In 2004, the number of cases increases dramatically following a series of years of low incidence and low insecticide application. This year also exhibits an excess of rainfall. This large epidemic was then followed by an increase in the level of insecticide spraying and, correspondingly, by low levels of malaria cases, which in turn gave way to low levels of control up to the present.

These findings have implications for control policies and the relevance of a climate-based early-warning system to control itself. Control in these regions is explicitly reactive in the sense that cases in the previous year determine the allocation and planning of resources for insecticide spraying in the next season within the districts. Even in the absence of such policy, the implementation and planning of control in epidemic regions is likely to be implicitly 'reactive'. This is because a feedback is likely to emerge from past disease levels to future human and economic resources allocated to

reducing vector abundance and malaria transmission. Whether or not an explicit reactive control policy exists, this feedback can operate at time scales longer than one year, depending in a complex way on the perception of the problem and availability of resources. Thus, our findings have implications beyond the specific district of Kheda to other regions of epidemic malaria where irrigation and its concomitant development underlie more effective control measures. They suggest that inter-annual variability in control levels can set the stage for temporal windows of high susceptibility to anomalous weather conditions. They also indicate that remote sensing in indicator regions (such as the Thar desert) can be used to forecast the potential risk of an outbreak given regional climate variability, especially when control levels fall to low levels.

The dynamical model of mosquito abundance in an irrigated landscape further illustrates the persistent role of climate forcing for disease risk under irrigation. Under irrigation, the seasonal relationship between rainfall and mosquitoes is altered, but only in the non-epidemic months (of low rainfall). In the epidemic months, mosquito abundance continues to respond to rainfall events and, therefore, the inter-annual variability of mosquito abundance after the rainfall season remains unaltered (See Fig. S2.9, Panel D). This is consistent with the scenario of no decrease in the risk of large epidemics in Kheda with increased irrigation, and the persistence of climatic variability as an important factor, now modulated by levels of control.

Entomological studies may also shed light on this conclusion. *Anopheles culicifacies*, the principal malaria vector in rural areas of India, is mainly present in riverine and canal areas, where two peaks in vector abundance are observed yearly, one in the monsoon season (Jul-Aug) and another in March-April associated with irrigated rice fields (Mukhtar et al., 2003). Although temperatures in March increase enough to support mosquito development, temperatures become very high in May-June, leading to the decline of vectors' abundance and lifespan. The short vector longevity during winter leads to a low sporogony rate, and therefore, a low transmission rate for this period. In the

monsoon season, however, optimum temperatures and humidity, and extensive areas for breeding, generate suitable conditions for parasite development and large vector populations.

Finally, this study has considered the particular scale of the district, at which the epidemiological data were aggregated. This may not match well the scales at which irrigation influences socioeconomic conditions. In this regard, the complexity of the interaction between malaria dynamics and land-use change, and the long-term consequences of the interaction between socio-economic determinants and malaria dynamics, deserve a closer examination. At local scales, irrigation can be associated with high mosquito populations, but not necessarily with high incidence, a phenomenon known as the paddies paradox (Kant and Pandey, 1999; Ijumba et al., 2001; Ijumba et al., 2002). Irrigation also may provide new ways to enhance wealth that subsequently improve education levels, housing conditions, or other types of protective measures taken by the individuals in a community or population (Yasuoka et al., 2006; Ng'ang'a et al., 2009). Therefore, in addition to control measures implemented at larger spatial scales, the observed decoupling between climate variability and malaria can also reflect long-term changes in socio-economic drivers such as better coverage by health facilities, self protective measures, and house improvements. Regardless of the specific mechanism, in areas where the risk itself persists for anomalous climatic conditions, it would be beneficial to incorporate predictions of this risk based on remote sensing tools in the planning of spray interventions.

Further work is needed to understand the connection between agriculture, mosquitoes, human behaviour, and wealth in human-modified landscapes at different spatial and temporal scales. At the large scale of districts, these findings underscore the differential effects of irrigated landscapes on malaria's risk and predictability, including the possibility of unexpected epidemics that are more difficult to predict because of the complex interaction between climate forcing and control efforts. At lower scales, to understand the role of remote sensing in predicting epidemic risk in different

agricultural landscapes and how this risk changes in space and time, more research is clearly needed.

Conclusions

Remote sensing (the vegetation index known as NDVI) provides a useful predictor of malaria epidemics in regions with low levels of irrigation. Increased irrigation modifies the coupling between climatic forcing and malaria's inter-annual variability. This decoupling appears to reflect the effect of control measures rather than a reduction in disease risk. Thus, early-warning systems based on remote sensing in regional indicator regions remain of value to control itself and to the preparedness for public health responses. In addition, reactive control policies may lead to unexpected large epidemics in areas with increased irrigation, when anomalous rainfall coincides with relaxations of control. Prediction efforts coupled to non-reactive control would be of particular value in the transition stage from largely rainfall-driven epidemics to a more permanent reduction of the malaria risk that would accompany socio-economic development and increased irrigation.

Figures

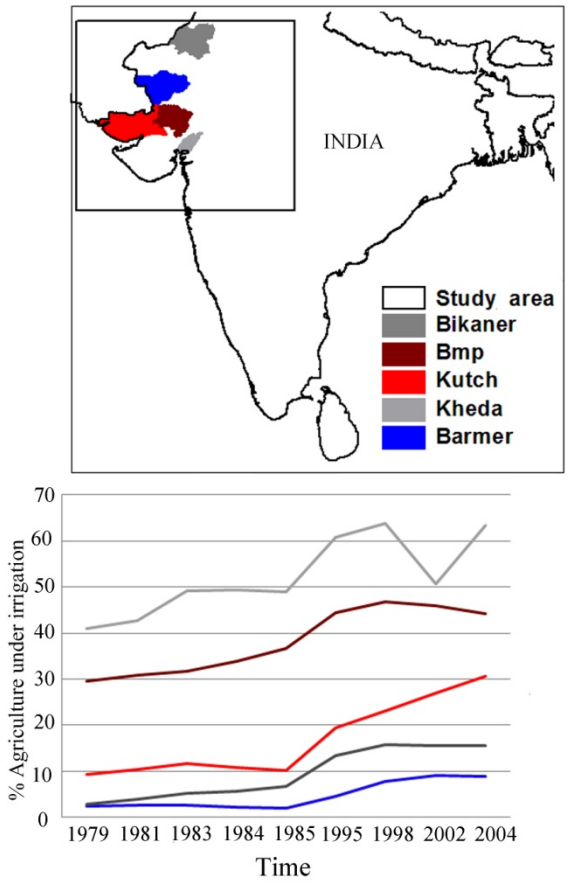

Figure 2. 1 Study area in north-west India and the level of irrigation of each district

Each line represents a time series of the percentage of agricultural land under some source of irrigation (source: district statistical books, Gujarat and Rajasthan).

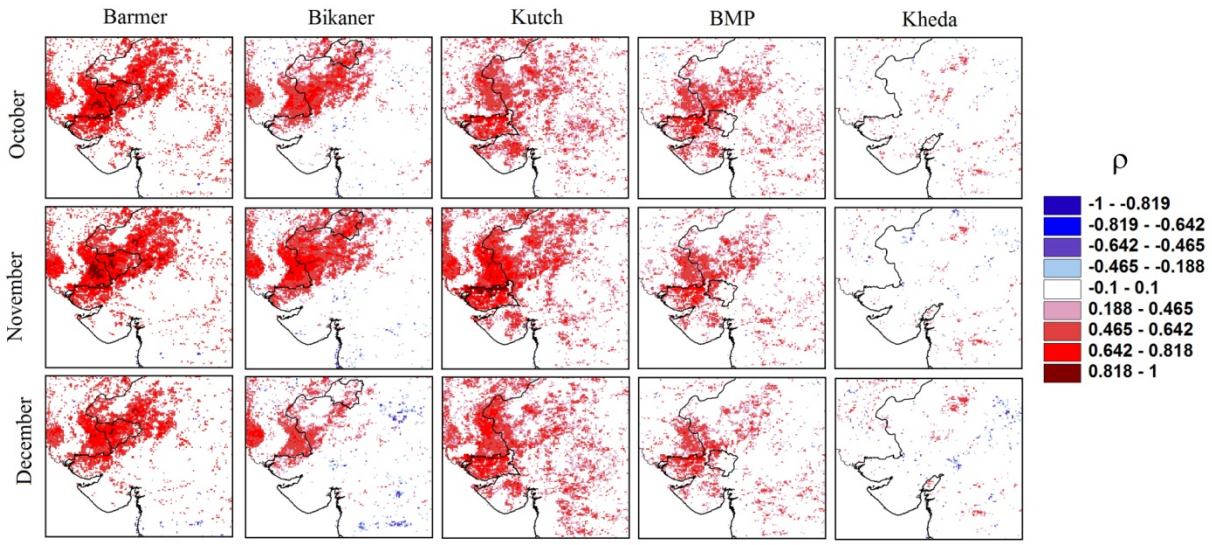

Figure 2.2 Correlation maps

Spearman rank correlation (ϱ) between malaria incidence from October to December and September NDVI, for each location (8 × 8 km grid point) of the study area. Each location (pixel) then represents the correlation in time between NDVI at this location and malaria incidence from a specific district. This boundary of this district is indicated inside each map. A high spatial correlation is observed over a large regional area (including the Thar desert), especially for the driest and weakly irrigated districts.

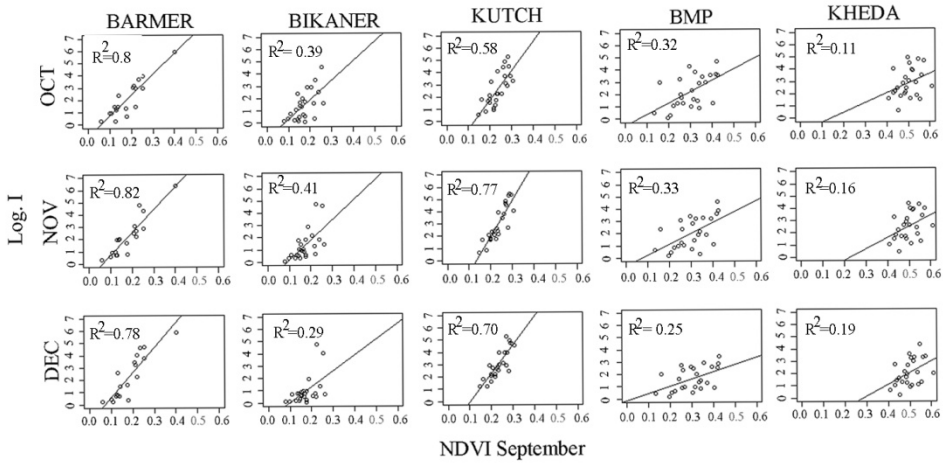

Figure 2.3 Linear regression plots

September NDVI is the predictor of malaria incidence in October, November and December.

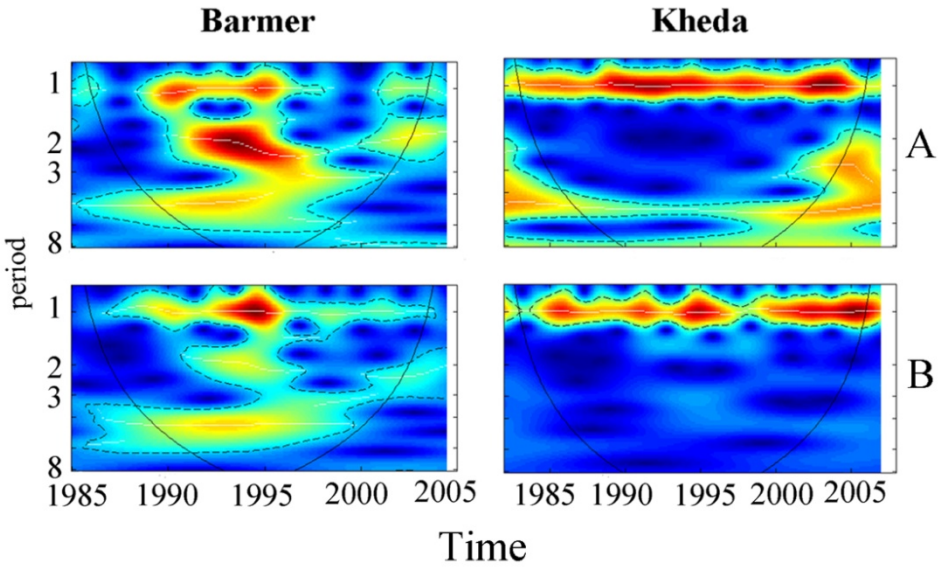

Figure 2.4 Wavelet power spectrum for malaria incidence and NDVI

The wavelet power spectra of the cases (panel A) and NDVI (panel B) are shown for Barmer and Kheda, the two districts at the extremes of the irrigation gradient (for the rest of the districts see Figure S2.6). The wavelet spectrum shows the variance (technically the power) for different periods (y-axis) and for different years (x-axis). The scale ranges from blue to red, with red indicating high power at a particular year and period. As irrigation increases, the 1-year period becomes stronger and the 2 and 4-year periods become weaker.

Appendices

Appendix 2.1: A simple model of mosquito population dynamics, rainfall and irrigation.

A relatively straightforward model can be constructed that examines the underlying dynamic interactions between rainfall, irrigated agriculture, and mosquito abundance. We use this model to interpret and illustrate some of the patterns observed in the statistical analysis of the empirical data.

In order to keep the model simple, we consider the dynamics of the mosquito population in a landscape composed of agricultural and non agricultural land (See Figure S2.3). Specifically, we define *D* the total proportions of a district's land as follows:

$$D = 1 = a + p.$$

where a is the proportion of the total area that is designated for agriculture and p is the proportion designated for other uses. We also differentiate between seasonal agriculture and agriculture under irrigation, such that

$$a = i + n$$

where i and n are the proportions of land covered by irrigated and non-irrigated agriculture, respectively.

Five equations describe the system:

$$\frac{dP}{dt} = p * f(\pi(- e * P - \frac{d}{e} * P * i$$
$$\frac{dW}{dt} = \left(\frac{d}{e} * P - c * e * W - f(\psi) \right) * i$$
$$\frac{dA_n}{dt} = f(\pi) * y_n * n - h * A_n$$
$$\frac{dA_i}{dt} = f(\psi) * y_i * i - h * A_i$$
$$\frac{dM}{dt} = b * \left(\rho * P + \alpha_n * A_n + \alpha_i * A_i + \omega * W \right) - \mu * M$$

P is the total volume of water that stands in p and W is the total volume of water stored in the canal network. A_n and A_i denote the total agriculture yield obtained in a (respectively, irrigated and non-irrigated), and M is the mosquito abundance.

Here $f(\pi)$ is the bi-annual rain fall cycle, which we characterize using

$$r_1*(1 + r_0*\sin(2\pi f_r)) + m_1*(1 + m_0*\sin(2\pi f_m)).$$

m_1 is the average extra water due to monsoon events with frequency $f_m = 1/24$ months, and r_1 is the average annual rainfall with $f_r = 1/12$. e is the evaporation rate of water in puddles, d is the rate at which water is drained into W, c is the relative rate at which irrigation evaporates (we assume a value around 0.1), and $f(\psi)$ is the rate at which water storage in W is supplied to irrigated agriculture. $f(\psi)$ is also a sine wave function with a period of one year and a maximal peak lagging 6 months after that of the rainfall season. y_n and y_i are the conversion constants from water to seasonal and irrigated crop production, and h is the crop harvesting rate. Although crop yield most likely differs for irrigated versus non-irrigated systems, and this may influence mosquito breeding preference, we set $y_n = y_i$ to focus on the large scale effect of increasing the area designated for irrigated agriculture. $\omega*W$ and $\rho*P$ correspond to the contribution of the canal network and accumulated water on p to

mosquito abundance, respectively, and $\alpha_n A_n$ and $\alpha_i A_i$ correspond to the contribution of non-irrigated (seasonal) and irrigated agriculture. In our simulations we assume that $\alpha_n < \alpha_n$.

The total yield of non-irrigated agriculture is assumed to be proportional to the amount of rain that falls directly in n. This is denoted by $f(\pi)*n$, where $f(\pi)$ is the total rain that falls onto D.

Mosquito birth rate is assumed to be entirely dependent upon available water in the district (agriculture, non-agriculture land, and canal network), but the potential for mosquitoes to breed varies depending on land-use. We ignore larval stages and associated time delays and simply assume that mosquitoes die at a constant rate μ. It would be relatively straightforward to add additional stages for mosquito biting rates and the abundance of infected and susceptible hosts, but the details of this would obscure the main points we wish to make, so we simply assume that mosquito abundance is a good index of transmission potential for malaria.

We initially used the model to examine the seasonal (annual) and inter-annual correlations between rainfall (measured by P) and mosquito abundance M. Figure S2.9 illustrates the correlations with and without irrigation: With no irrigation, there is a strong relationship between seasonal rainfall and mosquito abundance (panel A). When we increase the area of irrigated agriculture, then we see a weaker correlation between rainfall and mosquito abundance (Figure S2.9: panel B). However, irrigation does not affect the inter-annual correlation, as illustrated in panels C and D.

We also examine the long-term consequences for epidemic risk of increasing the area designated for agriculture under irrigation. We simulated the model for a range from almost no irrigated land $i = 0.01$, to thirty percent irrigated landscape, and measure the total mosquito abundance in a year and the maximal and minimal values (See Figure S2.8, panels A and B). These simulations show that

irrigation leads to an increase in the mosquito abundance (panel B), but it does not affect the maximum, which is still controlled by the amount of rainfall (panel A).

Appendix 2.2: Supporting Figures

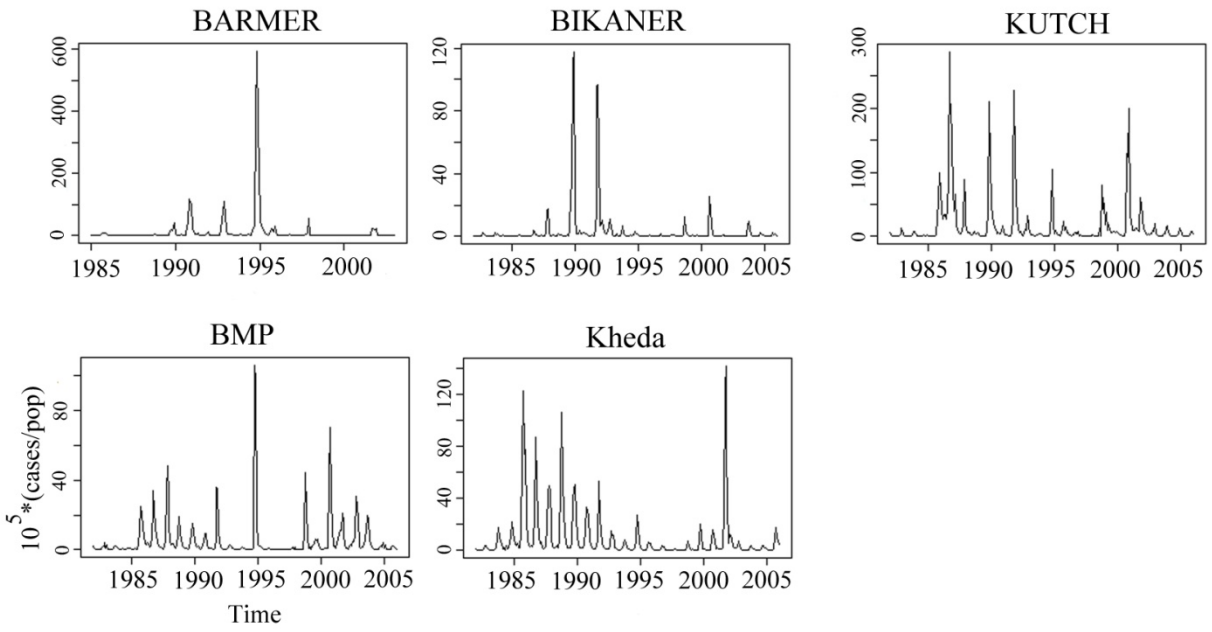

Figure S2.1 Time series of malaria incidence

The y-axis represents the monthly number of cases per 100,000 people. Note that the range in the y-axis varies across districts. For comparison purposes, see Additional file 2.2.

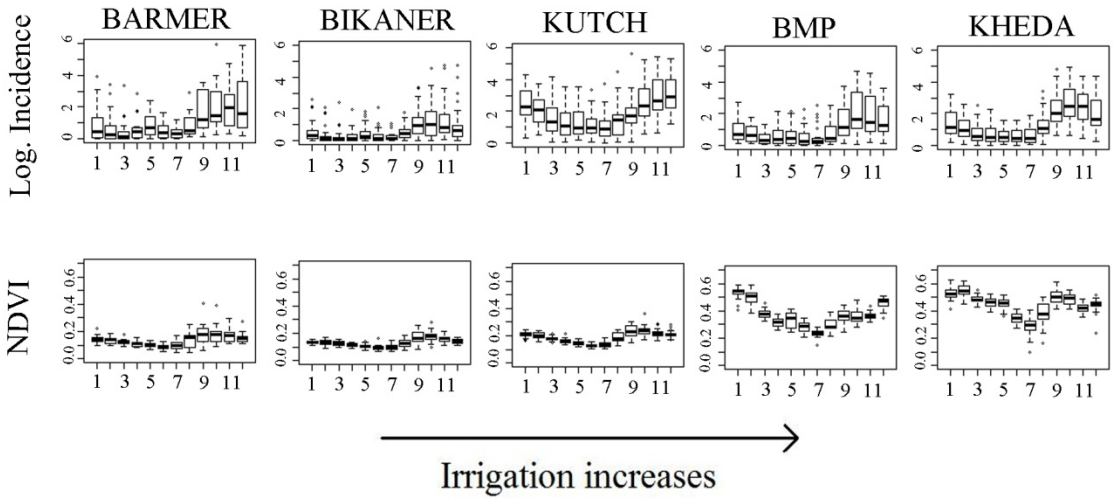

Figure S2.2 Box-plots of malaria incidence and NDVI

The first row shows the average and the range of anomalies of cases (in logarithmic scale) for each district in a gradient of irrigation intensity. The second row shows NDVI from the time series inside the districts.

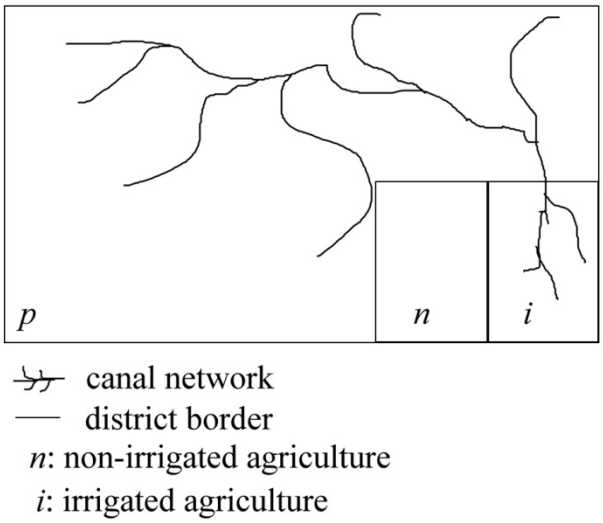

Figure S2.3 Graphical representation of the model.

The land inside the district is divided into irrigated and non-irrigated agriculture (*i* and *n*) and into other uses (*p*). A network of canals drains the water that precipitates on *p* to supply the production of irrigated agriculture.

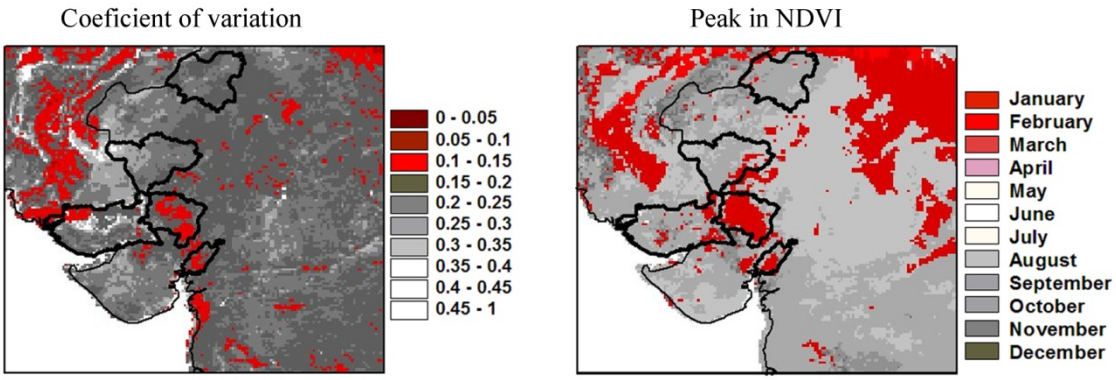

Figure S2.4 Coefficient of variation and seasonality of NDVI

BMP and Kheda both exhibit low coefficients of variation (red colors) and a peak in NDVI in the month of January. (In the left side of both figures, the red colored areas delineate the irrigation tract associated with the Indus River in neighbouring Pakistan.)

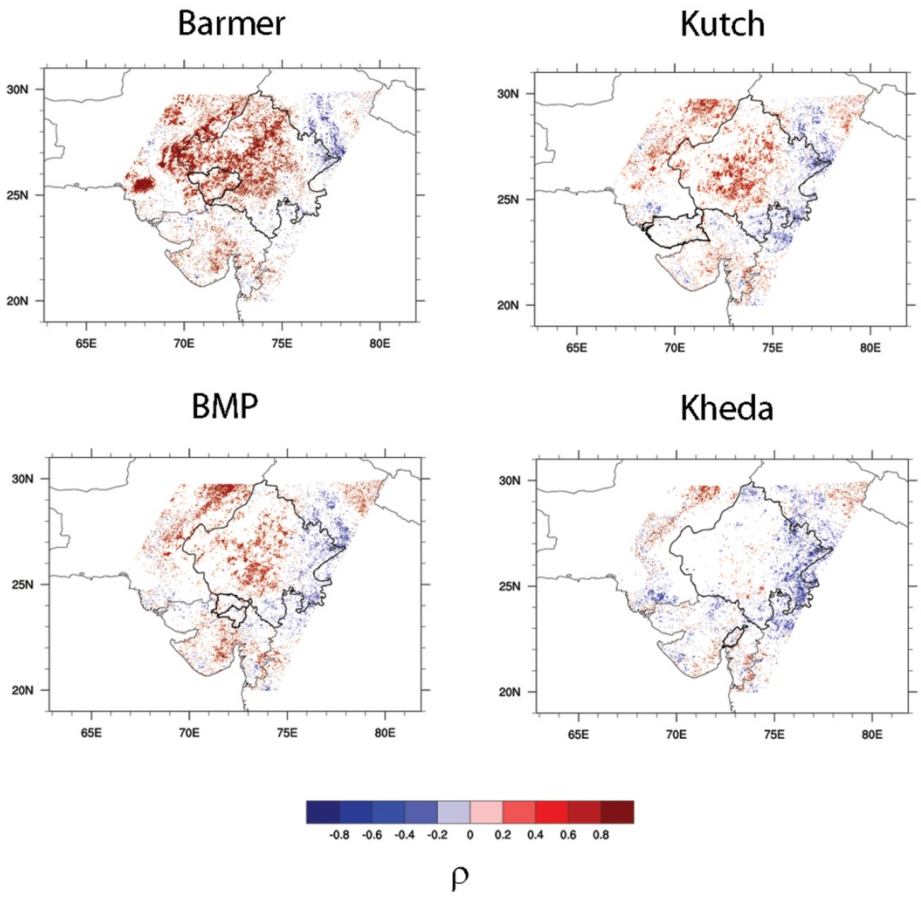

Figure S2.5 Correlation maps using MODIS images.

Spearman rank correlation between September NDVI from MODIS and malaria incidence for a specific district in the epidemic season (the sum of the cases for October, November and December). Note that the dataset consists of ten years (and only 7 years for Barmer). At a significant level of 0.1, evidence for an association between malaria and NDVI at the regional level is present for both Barmer and Kutch. This pattern is less pronounced, however, than for the NOAA NDVI data because of the shorter length of the time series for MODIS NDVI.

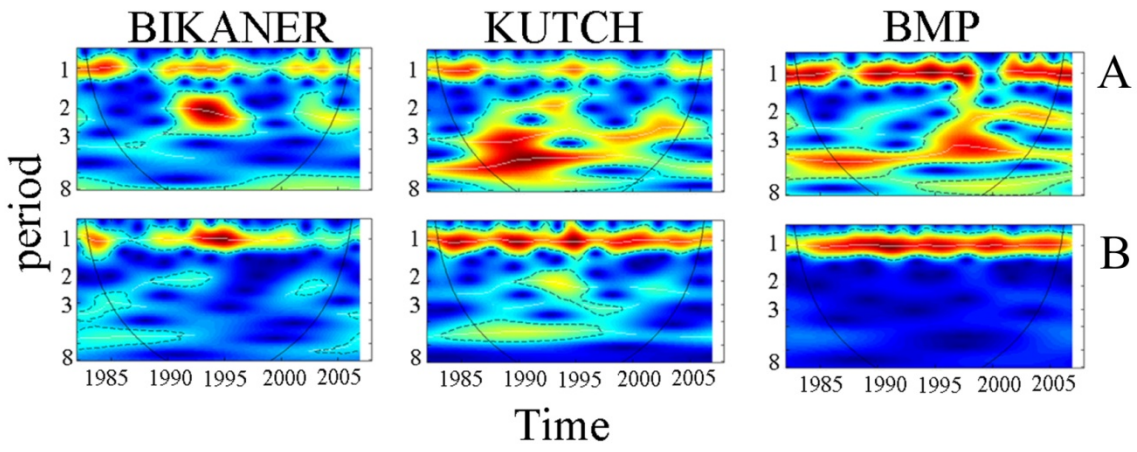

Figure S2.6 Wavelets analysis (continuation)

Similar to Figure 2.4, but for the three districts in the middle part of the irrigation gradient. The picture shows that as irrigation intensified, the 1 year signal became stronger over longer periods of time, both for incidence (Panel A) and NDVI (Panel B).

year	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
rainfall	860	1455.6	915.6	83	459.5	700.7	132.1	50	935.7	454.44	454.58	-	-	-	-
prop. population covered	0.299847	0.13544	0.198545	0.194474	0.107693	0.024417	0.045025	0.060538	0.041295	0.392468	0.372458	0.202045	0.064263	0.064187	0.036892
cases	1117	2418	1077	631	153	663	1748	1954	16855	2519	432	770	574	423	581

Table S2.1Total rainfall, insecticide application and number of cases recorded in Kheda

The amount of insecticide use corresponds to the proportion of the state population covered by spray activity in that particular year. Rainfall and cases are the total values for the year. Note that in 2004 the large number of cases coincides with anomalous conditions of rainfall and a relatively low level of insecticide application, not just that year but also for a number of previous years. In 1997 and 2001 similar anomalous rainfall conditions did not produce this large number of cases. (Insecticide use data were missing for 2003; the value in the table was predicted by a linear regression of insecticide use at the district level as a function of insecticide use for 10 talukas, administrative units within the district; regression coefficient 0.979). No rainfall data after 2006 are available at this point.

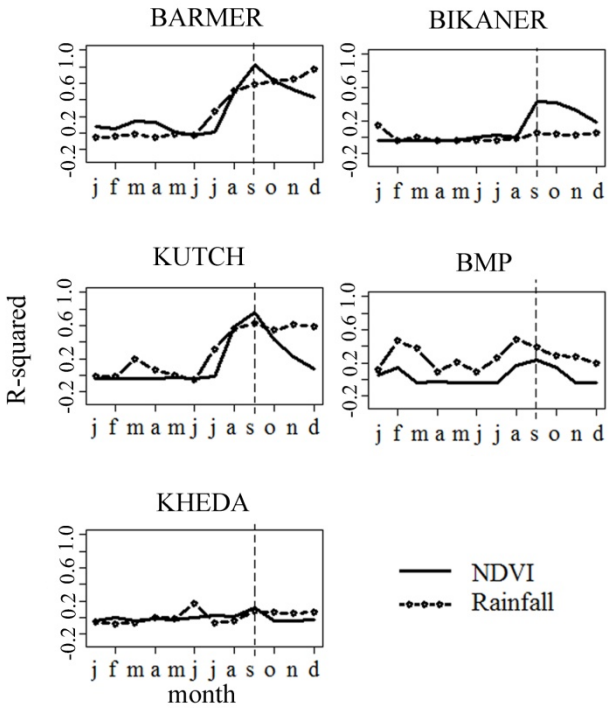

Figure S2.7 Malaria predictability based on NDVI v/s rainfall

The x-axis shows the month of the year used to fit a linear model of the number of cases in the epidemic season (October to December). The y-axis shows the corresponding R-squared value. NDVI is a better predictor than rainfall one month prior (September; dashed line) to the epidemic season (October-November-December) for Barmer, Bikaner and Kutch. For BMP, rainfall from Banaskantha is a better predictor. For Kheda, neither NDVI, nor rainfall, are good predictors of epidemics.

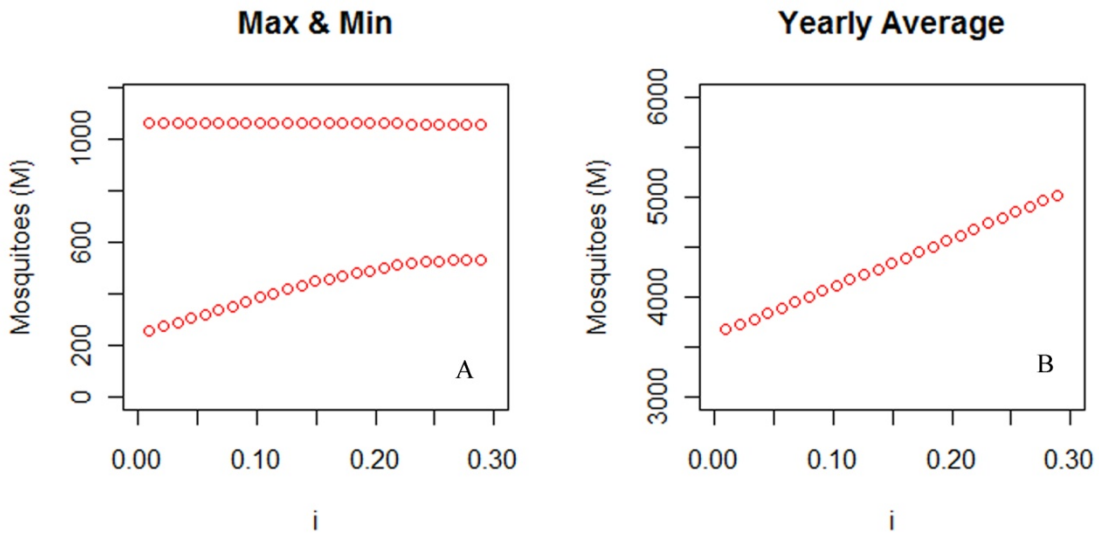

Figure S2.8 Maximum, minimum and yearly average mosquito abundance

Panel A shows that the minimum mosquito abundance increases as the total area under irrigation increases, however its maximum does not change. Panel B shows that mosquito abundance increases linearly with the proportion of land under irrigation (*i*).

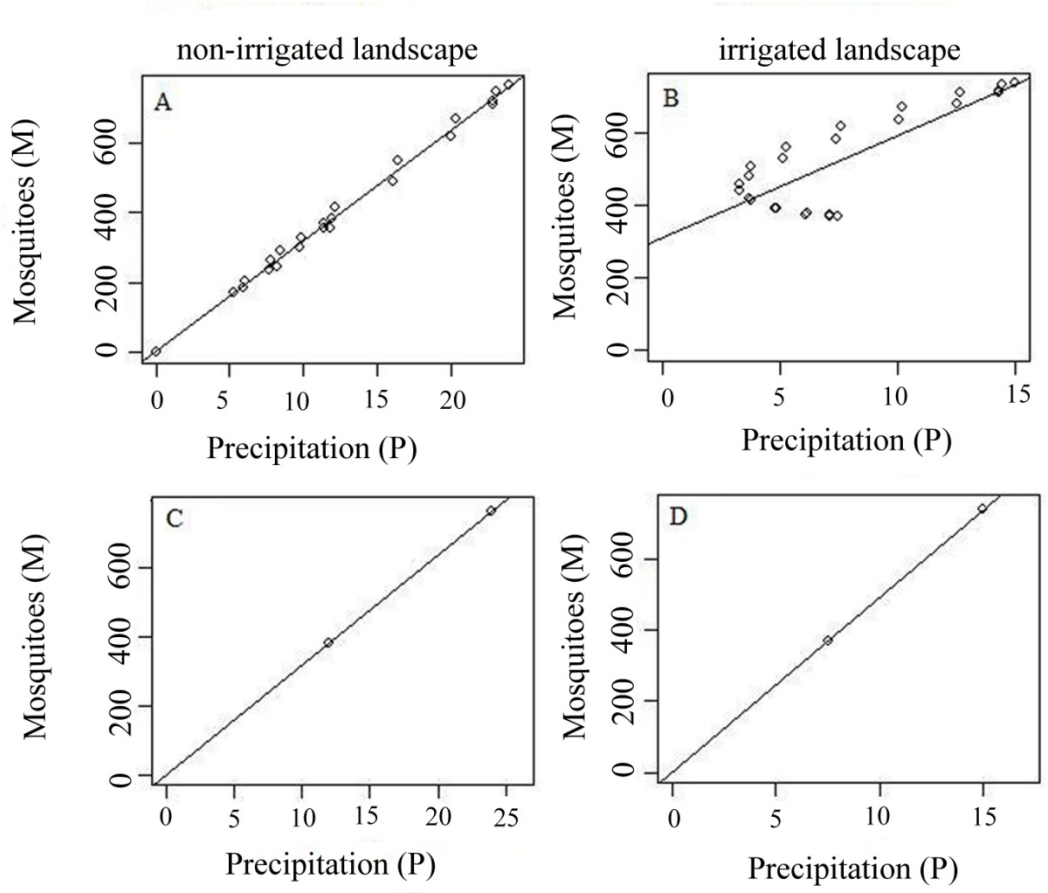

Figure S2.9 Seasonal and inter-annual correlation

Correlation between mosquito (M) and precipitation (*P*) with non-irrigated agriculture (*i*= 0; panel, A and C), and with 30 percent of the landscape under irrigated agriculture (*i*= 0.3; panels B and D). Seasonal correlation in panels A and B and inter-annual correlation in panels C and D. The values for the rest of the parameters are: $n= 0.1$; $p= (1- n – i)$; $e= 30$; $d= 200$; $c= 0.1$; $b= 120$;$\mu= 18$;$\rho= 0.8$;$\omega= 0.1$;$f_n=f_i= 3$; $r_1= 200$; $r_0= 0.99$; $m_1= 200$; $m_0= 0.99$; $h= 5$; $\alpha_n = 2$;$\alpha_i= 3$. The annual cycle leads to the change in correlation, but not the actual inter-annual variability.

References

Akhtar, R., McMichael, A.J., 1996. Rainfall and malaria outbreaks in western Rajasthan. Lancet, 348,1457-1458.

Bouma, M.J., vanderKaay, H.J., 1996. The El Nino Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: An early warning system for future epidemics? Trop Med Int Health. 1,86-96.

Cazelles, B., Chavez, M., de Magny, G.C., Guegan, J.F., Hales, S., 2007. Time-dependent spectral analysis of epidemiological time-series with wavelets. J R Soc Interface. 4, 625-636.

Christophers, R., 1911. Malaria in the Punjab. In: *Dept Gov India (New Series) no 46*. Edited by Sanit SMOM. Calcutta, India: Superintendent Government Printing.

Ceccato, P., Ghebremeskel, T., Jaiteh, M., Graves, P.M., Levy, M., Ghebreselassie, S., Ogbamariam, A., Barnston, A.G., Bell, M., del Corral, J., Connor, S.J., Fesseha, I., Brantly, E.P., Thomson, M.C., 2007. Malaria stratification, climate, and epidemic early warning in Eritrea. Am J Trop Med Hyg. 77,61-68.

Connor, S.J, Thomson, M.C, Flasse, S.P., Perryman, A.H., 1998. Environmental information systems in malaria risk mapping and epidemic forecasting. Disasters. 22,39-56.

Gill, C.A., 1923. Malaria in the Punjab. The malaria forecast for the year 1922. Indian J Med Res. 661-666.

Hay, S.I., Snow, R.W., Rogers, D.J., 1998. Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. Trans R Soc Trop Med Hyg. 92,12-20.

Ijumba, J.N., Lindsay, S.W., 2001. Impact of irrigation on malaria in Africa: paddies paradox. Med Vet Entomol. 15,1-11.

Ijumba, J.N., Shenton, F.C., Clarke, S.E., Mosha, F.W., Lindsay, S.W., 2002. Irrigated crop production is associated with less malaria than traditional agricultural practices in Tanzania. Trans R Soc Trop Med Hyg. 96,476-480

Johansson, M.A., Cummings, D.A.T., Glass, G.E., 2009. Multiyear climate variability and dengue-El Nino Southern Oscillation, weather, and dengue incidence in Puerto Rico, Mexico, and Thailand: A longitudinal data analysis. *PLos Med.* 6,11.

Kant, R., Pandey, S.D., 1999. Breeding preferences of Anopheles culicifacies in the rice agro-ecosystem in Kheda district, Gujarat. Indian Journal of Malariology. 36,53-60.

Laneri, K., Bhadra, A., Ionides, E.L., Bouma, M.J,. Dhiman. R.C., Yadav, R.S., Pascual, M., 2010. Forcing versus Feedback: Epidemic Malaria and Monsoon Rains in Northwest India. PLos Comput Biol. 6,9.

LP DAAC : ASTER and MODIS Land Data Products and Services [https://lpdaac.usgs.gov/]

Mukhtar, M., Herrel, N., Amerasinghe, F.P., Ensink, J., van der Hoek, W., Konradsen, F., 2003. Role of wastewater irrigation in mosquito breeding in south Punjab, Pakistan. Southeast Asian J Trop Med Publ Health. 34, 72-82.

Ng'ang'a, P.N., Jayasinghe, G., Kimani, V., Shililu, J., Kabutha,. C., Kabuage, L., Githure, J., Mutero, C., 2009. Bed net use and associated factors in a rice farming community in Central Kenya. Malar J. 8.

Thomson, M.C., Doblas-Reyes, F.J., Mason, S.J., Hagedorn, R., Connor, S.J., Phindela, T., Morse, A.P., Palmer, T.N., 2006. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. Nature. 439,576-579.

Pascual, M., Cazelles, B., Bouma, M.J., Chaves, L.F., Koelle, K., 2008. Shifting patterns: malaria dynamics and rainfall variability in an African highland. Proc R Soc B-Biol Sci. 275,123-132.

Rogers, D.J., Randolph, S.E., Snow, R.W., Hay, S.I., 2002. Satellite imagery in the study and forecast of malaria. Nature. 415,710-715.

Sachs, J., Malaney, P., 2002. The economic and social burden of malaria. Nature. 415,680-685.

Swaroop, S., 1949. Forecasting of epidemic malaria in the Punjab, India. Am J Trop Med Hyg. 29,1-16.

Thomson, M.C., Connor, S.J., D'Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M., Greenwood, B., 1999. Predicting malaria infection in Gambian children from satellite data and bed net use surveys: The importance of spatial correlation in the interpretation of results. Am J Trop Med Hyg. 61,2-8.

Thomson, M.C., Connor, S.J., Milligan, P., Flasse, S.P., 1997. Mapping malaria risk in Africa: What can satellite data contribute? Parasitol Today. 13,313-318.

Tucker, C.J., Pinzon, J.E., Brown, M.E., Slayback, D.A., Pak, E.W., Mahoney, R., Vermote, E.F., El Saleous, N., 2005. An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. International Journal of Remote Sensing. 26,4485-4498.

Tyagi, B.K., Yadav, S.P., 2001. Bionomics of malaria vectors in two physiographically different areas of the epidemic-prone Thar Desert, north-western Rajasthan (India). J Arid Environ. 47,161-172.

Tyagi, B.K., 2004. A review of the emergence of *Plasmodium falciparum*-dominated malaria in irrigated areas of the Thar Desert, India. Acta Trop. 89(2),227-239.

Yacob, K.B., Swaroop, S., 1946. Malaria and Rainfall in the Punjab. Journal of the Malaria Istitute of India. 6.

Yasuoka, J., Mangione, T.W., Spielman, A., Levins, R., 2006. Impact of education on knowledge, agricultural practices, and community actions for mosquito control and mosquito-borne disease prevention in rice ecosystems in Sri Lanka. Am J Trop Med Hyg. 74,1034-1042.

Zurbrigg, S., 1994. Re-thinking the" human factor" in malaria mortality: the case of Punjab, 1868-1940. Parassitologia. 36,121-135.

Chapter 3

Malaria and its control in highly variable environments

Abstract

In areas of the world where malaria prevails under unstable conditions, attacking the adult vector population through insecticide-based Indoor Residual Spraying (IRS) is the most common method for controlling epidemics. Defined in policy guidance, the use of Annual Parasitic Incidence (API) is an important tool for assessing the effectiveness of control and for planning new interventions. To investigate the consequences that a policy based on API in previous seasons might have on the population dynamics of the disease and on control itself in regions of low and seasonal transmission, we formulate a mathematical malaria model that couples epidemiologic and vector dynamics with IRS intervention. This model is parameterized for a low transmission and semi-arid region in northwest India, where epidemics are driven by high rainfall variability. We show that this type of feedback mechanism in control strategies can generate transient cycles in malaria even in the absence of environmental variability, and that this tendency to cycle can in turn limit the effectiveness of control in the presence of such variability. Specifically, for realistic rainfall conditions and over a range of control intensities, the effectiveness of such 'reactive' intervention is compared to that of an alternative strategy based on rainfall and therefore vector variability. Results show that the efficacy of intervention is strongly influenced by rainfall variability and the type of policy implemented. In particular, under an API 'reactive' policy, high vector populations can coincide more frequently with low control coverage, and in so doing generate large unexpected epidemics and

decrease the likelihood of elimination. These results highlight the importance of incorporating information on climate variability, rather than previous incidence, in planning IRS interventions in regions of unstable malaria. These findings are discussed in the more general context of elimination and other low transmission regions such as highlands.

Introduction

Malaria transmission is low and highly seasonal at the edge of the distribution of the disease where climate variables, either temperature or rainfall, limit vector abundance and parasite development. Thus, in these regions, control efforts unavoidably operate in highly variable environments where malaria dynamics are known as 'unstable' or 'epidemic'. It is the interplay of control and climate variability in one such environment, desert fringes, that interests us here, especially for the kind of dynamic intervention policy that would prove most effective.

Historically, the discovery of effective and long lasting residual insecticides, such as DDT, contributed to a significant extent to WHO's global initiative to eradicate malaria initiated in 1955. These campaigns dramatically reduced India's estimated annual malaria deaths from an estimated million (Russell, 1936) to only a few hundred. When malaria eradication was no longer considered feasible at a global scale, policy shifted to treatment of the disease to limit clinical burden, and the use of Indoor Residual Spray (IRS here on) was mostly abandoned and viewed as an expensive short term eradication tool. However, with the dramatic resurgence of malaria in South Asia in the mid 1970s, IRS made a comeback as a proven method to reduce morbidity, but without a careful appraisal of spraying methods previously used as a long-term control. The legacy of the eradication efforts in India are still visible in the form of an impressive network of diagnostic services (microscopy) and the continued dependence on indoor spraying, that proved so effective in large

parts of the country where vectors are rather inefficient because they are short-lived and exhibit zoophilic behavior.

Another legacy from the eradication era is the use of the Annual Parasite Incidence rate (API), usually expressed as cases per 1000 per year, to evaluate the effectiveness of interventions and plan the subsequent phases of the campaign (Najera et al., 2011). In India, areas with API > 2.0 cases per 1000 population in the preceding 2 years qualified for intervention (NMEP, 1983). Although the absolute number justifying IRS intervention has been modified over the years, the reactive nature of insecticide use has remained unchanged until the present. In addition to the inherent delays in the purchasing and delivery of commodities and supplies, reasons for a response that applies in the following transmission season include the practical requirements of a national scale and of a rigorous implementation that considers for example the timing of the transmission season after the monsoon and the duration of insecticidal activity. The uniform approach that characterized the eradication design did not take into account however the dynamic differences of malaria within the country such as those between endemic and epidemic regions, and the potential implications of a delayed response if elimination were not to be achieved in the short term.

Thus, an important but often unrecognized consequence of relying on these indices for the planning of subsequent interventions is the potentially reactive nature of the public health response. We specifically refer here to this kind of response as 'reactive' to describe a delayed response whose level depends on the disease burden in the past over some temporal window of time, for example cases in the last season. This effectively establishes a dynamic feedback between past incidence levels and current control efforts. Such feedback can arise either from explicit control policies or from the myriad processes underlying the allocation of intervention efforts and the perception of risk in public health systems operating under limited resources. Regardless of the actual mechanisms, case-

detection would trigger control intervention in affected areas leading to the subsequent decrease in incidence. This reduction in the number of cases would in turn result in a decreased perception of risk leading to the relaxation of control before actual elimination is achieved. Because malaria incidence in the absence of intervention would then return to previous levels defined by environmental and socioeconomic conditions, a reactive intervention has the potential to generate recurrent disease cycles and unexpected epidemics.

These dynamics are of particular relevance where transmission is 'unstable' or 'epidemic' and under the influence of highly variable environmental drivers such as rainfall. In these regions, climate variability is known to generate strong interannual variation in the number of cases (e.g. Laneri et al., 2010) but its interaction with vector control is poorly understood. Although a vast body of work has addressed and compared the effects of particular interventions (Goodman et al., 2001; Guyatt et al., 2002; Smith et al., 2009; Pedercini et al. 2011; White et al., 2011; Kigozi et al., 2012; Hamusse et al. 2012), there has not been much work exploring the effectiveness of different strategies under high environmental variability (but see Worrall et al., 2007). Here we hypothesize that in these areas reactive control policies can generate long cycles between IRS interventions and epidemics, and also influence the efficiency of the allocation of resources and the risk of malaria in the long run.

The long-term malaria control program in the arid northwestern states of India provides a unique opportunity to simultaneously follow malaria epidemiology and IRS intervention in a seasonal and low-transmission region under strong rainfall forcing, where control policy explicitly defines target areas based on the number of recorded cases in previous years. To investigate the dynamical consequences of this policy, a coupled human-mosquito transmission model is parameterized for this semi-desert region. With this model, the effectiveness of reactive control is also compared against that of an alternative strategy based on rainfall variability. Results show that transient

multiannual cycles can emerge in the absence of climate variability. When driven by observed daily precipitation, the model exhibits intermittent unexpected epidemics corresponding to large rainfall events and situations of unpreparedness and inefficient control. These patterns are consistent with observations in time series of cases and IRS coverage in the region. Our findings further suggest that control efforts whose level is determined based on rainfall itself provide a more effective and efficient strategy.

Background and Motivation

Malaria situation and large scale control strategy in the states of Gujarat and Rajasthan, India.

Gujarat and Rajasthan are the western states of India and their combined population exceeds 120 million people in an area of 196.000 Km^2. In most northern and western districts, total rainfall does not exceed 200 mm per year, while in the southernmost districts more than 1000 mm can be recorded. High interannual variation in rainfall intensity is observed over the region. As a consequence, large fluctuations in the abundance of the mosquito population from year to year can be generated, and these fluctuations in vector abundance are reflected in the pronounced interannual variability of the seasonal malaria outbreaks (Laneri et al., 2010). The principal species of malaria are *Plasmodium falciparum* and *Plasmodium vivax,* and the principal malaria vector is *Anopheles culicifacies* (Barik et al., 2009).

Since the resurgence of malaria in the 1970s, several malaria control intervention strategies have been implemented to decrease the burden of malaria in the region (Sharma and Sharma 1989, Sharma et al., 1991). The malaria endemicity of an area (village, talukas or districts) is estimated based on the number of cases detected by the surveillance system, consisting of clinical malaria cases

diagnosed in health centers (passive case detection) and cases brought to light by outreach workers (active case detection). Based on these estimations, the modified plan of operation (MPO) since 1977 consisted of the use of IRS in rural areas recording an API > 2 from the previous 2 malaria seasons (NMEP, 1983). In the last 15 years, this threshold is calculated based on the previous 3 years (NMEP, 1995). This policy is reflected in the delayed increase in control levels that follows large epidemic years (Fig 3.1).

Methods

The malaria-mosquito model

We developed a mathematical model to investigate the temporal dynamics generated under this mechanistic feedback between intervention and incidence, and their consequences for the interannual variability of seasonal outbreaks and control itself. The model is a coupled mosquito-malaria model that explicitly considers a control intervention that increases the mortality of the adult mosquito population. It is written as a system of differential equations and organized into three modules: The first one is the epidemiological module which divides the total human population, N, into 4 classes for susceptible (S), exposed (E), infected (I) and recovered (R) individuals respectively. The transitions between classes and the structure of the model are represented in the diagram of Figure 3. 2, and a more detailed explanation of the model is given in section 8 (Appendix 3.1).

In particular, the force of infection β or per capita rate at which susceptible individuals become infected, depends on the environmental factors that influence the density of the adult mosquito population and the mosquito vectorial capacity, as well as the human and mosquito behavior that influence their encounter rate and the rate at which mosquitoes bite human hosts. In the model, the

force of infection is then expressed as the product of the human biting rate a, the probability that an infectious bite becomes an infection b, the ratio M/N of mosquitoes (M) to humans (N), and the infectious mosquito fraction W/M.

The second module of the model tracks the dynamics of the mosquito population (Fig. 3.2) and follows closely the general representation in Alonso et al. (2010). This subdivides the vector population into larval (L) and adult stages (M) and further subdivides the adult mosquitoes into three classes for uninfected X, infected V and infectious W individuals, respectively, so that $M = X+V+W$. The larval birth rate depends on the intrinsic growth rate of larvae F, the density of adult mosquitoes M, and the carrying capacity K. The parameter K is defined as the total area available for mosquito growth and depends on the amount of water in the environment (P) and the number of mosquitoes supported by a unit area (K_r).

Larvae develop into adults at rate d_L and die at a constant rate δ_L. The larvae that survive per unit time $d_L L$ become uninfected adult mosquitos X. These mosquitoes become infected at rate acy where c is the probability that a bite of an infected human results in an infected mosquito, and y is the proportion of infected humans $(y = I/N)$. We further consider that infected mosquitoes V do not transmit the parasite until sporozoites develop in the salivary glands of the mosquitoes. Thus, infected mosquitoes become infectious W at the sporogony rate γ.

A large entomological literature has accumulated for the most important mosquito vectors in India (Ramachandra Rao, 1984). This literature provides most of the information needed to parameterize the entomological submodel described above for this region based on the most common malaria vector, *Anopheles culicifacies*. For a few parameters, unavailable for this species, we relied instead on studies of other *Anopheles* species (Appendix 3.2). For the human submodel, we adopt the parameter

estimates previously obtained by Laneri et al. (2010) for the district of Kutch (Gujarat) with a recently developed statistical method for parameter inference from time series data.

One important aspect of unstable malaria transmission in semi-desert environments is the constraint of mosquito population growth imposed by water availability in the environment. This is incorporated in a third module of the model via a simple representation of the process of water accumulation in the landscape from precipitation and its effect on larval intraspecific competition. To estimate the parameters of this conversion of rainfall into mosquito larvae, and in particular the carrying capacity, we maximize the likelihood of the model assuming only measurement error (Appendix 3.2). Table 3.1 summarizes the parameters of the model as well as their corresponding source.

Control-induced mosquito mortality

Now we turn our attention to mosquito adult mortality δ_M and the implementation in the model of the reactive policy defined above. In the model we specifically incorporate the effect of IRS intervention as an additional mortality rate δ_C, over the natural mortality δ_0, which therefore shortens the average lifespan of an adult mosquito, with the total mortality now given by $\delta_M = \delta_0 + \delta_C$.

To implement reactive control, we consider δ_C a function of cases in the previous seasons given by

$$I_P = \int_{t-\tau_c}^{t} E_{(t)} \mu_{EI} dt$$

μ_{EI} represents the per-capita rate at which individuals in class E move to the class I. That is, I_P corresponds to the total number of cases added over a given interval of time $[t-\tau_C, t]$ in the past. This

expression emulates the API policy described in the background and motivation section. The shape of the control function δ_C is shown in Figure 3.2 and is based on the observed patterns in the time series data (See results).

We compare this strategy against a different one based on the observed rainfall rather than cumulated previous incidence. Now $\delta_M = \delta_0 + \delta_R$ with δ_R, the control-induced mosquito mortality based on rainfall, given by

$$\delta_R = \frac{\delta_{\max} R}{h_R + R},$$

where δ_{max} denotes the maximal extra mortality possible, h_R is a half saturation constant, and R is the total cumulative rainfall observed over the previous month,

$$R = \int_{\tau_R} r_{(t)} dt .$$

Thus, the magnitude of this response saturates with extreme events. Furthermore, intervention in this case only occurs during the rainy season, between June and August for reactive control and between July and September for rainfall-based control.

Assessment of intervention and control effort

In general it is expected that as the resources and effort put into control measures increase, the intensity of the intervention also increases and the number of cases should go down. Thus the effectiveness of a control strategy can be measured by the total reduction in risk and will depend on the intensity of the intervention or total control effort over the whole window of time considered. Here we define the total control effort (TCE) as the sum of the control induced mortality rate (δ_C)

over the period of time ΔT for which the intervention was in place. This quantity can then be interpreted as the total budget and resources needed to maintain the extra mosquito mortality over the intervention time (Appendix 3.3).

The assessment of the effectiveness of IRS interventions can be measured in three different ways: the direct estimation of the effectiveness of the insecticide, confirmatory entomological assessment and the epidemiological assessment of the intervention (Detinova, 1962). The first basically consists of checking that the insecticide is working properly, namely that the insecticide is effectively killing mosquitoes. For our purposes, we assumed 100 percent effectiveness without loss of generality. Confirmatory entomological assessment consists of estimating the total mosquito population that was effectively eliminated after the intervention. Of higher relevance to public health officers is the epidemiological assessment of the intervention quantified here with the Total Epidemiological or Disease Efficacy (TDE) and measured as the total number of cases over the period in which the intervention was applied. This is calculated by integrating the total number of new cases $(E\mu_{EI})$ over a period of time ΔT (Appendix 3.3).

The relationship between TCE and TDE can be depicted by a simple malaria intervention cost-benefit curve (MICB; Fig. S3.1). Importantly, the use of this curve allows us to compare the two strategies for similar control efforts over the whole period of intervention (e.g. ten years). In this way, differences in their effect do not depend on short-term, year to year, differences in allocation determined by climate variability.

Numerical Simulations

To incorporate climate variability, we forced the model with daily rainfall intensity (mm/day) from 6 climatic stations in the district of Barmer (Rajasthan). We randomly sampled from this dataset a subset of 45 series of 13 consecutive years each, beginning each series and corresponding simulation

with a different starting location and time point. For each rainfall scenario, we ran the model for the two control strategies over a set of parameter values to cover a large range of control intensities (90 simulations per rainfall scenario, each one a combination of parameters for the control functions δ_C and δ_R). Each simulation was run for 13 years, but control was applied only after an initial period of 3 years to let the effect of initial conditions decrease and disease levels settle around a mean determined by rainfall conditions. Only the final 10 years were considered in the estimation of effectiveness and effort. For the reactive strategy we consider three possibilities for the previous cases I_P calculated over for 1, 2, and 3 years (τ_C) respectively.

We also ran simulations for the same parameter values (Table 1) but under constant rainfall to establish the baseline temporal dynamics that are generated by reactive control in the absence of external forcing. A range of values was considered that encompasses observed median rainfall for the semi-arid districts of Kutch (Gujarat) and Barmer (Rajasthan) for the last ten years. Because we are interested in examining the initial, transient, dynamics but also the long-term temporal patterns, these simulations were run for a long time (~2000 years).

Results

Control under a constant environment

Figure 3.3 shows the population covered by IRS intervention as a function of the number of cases in the previous two years. An increasing and nonlinear relation is observed that saturates for high incidence. We note that the period over which this relationship is most evident does not need to correspond exactly with that of the policy, especially at the larger scale of the whole district vs. that of the local villages. This is because the actual manifestation of the policy depends also on the

allocation of human and economic resources and the perception of the problem at multiple levels of the administration.

Based on this pattern, we define the mortality due to reactive control in the model, δ_C, as an increasing function of the number of cases observed in previous years with a saturation form given by

$$\delta_C = \frac{\delta_{\max} I_P}{h + I_P},$$

where δ_{max} denotes the maximal mortality rate and h the half saturation constant. This constant specifically represents the reactiveness of the intervention by determining how fast the additional vector mortality rate rises as a function of I_P, the cases observed in the past (section 3.1.1).

Not surprisingly, in the absence of environmental variability, the long-term temporal dynamics of the model exhibits a one year cycle corresponding to the period of the control forcing. In the initial transient dynamics, however, a longer inter-annual cycle is observed (Figure 3.4A). This cycle has a long period of 10 years for the parameters considered here, as shown in Figure 3.4B. Although its amplitude decreases in time, so that this oscillation does not persist in the long-term, it does so slowly, over multiple decades. This tendency to cycle is important because it can interact with rainfall variability in ways that lead to intermittent and unexpected large epidemics, as evident in our comparisons next of the two control strategies under variable rainfall.

Control under environmental variability

Without control but under the observed rainfall forcing, the model exhibits high variability in the number of cases with patterns similar to those observed in the incidence time series from the surveillance program in this region. For similar values of annual and monsoon rainfall (JJA), the

number of cases over the same period can vary significantly depending on the specific pattern of rainfall variability within and between years (Fig. S3.2A and Fig. S3.2B). Importantly, this variability influences the effectiveness and results of the comparison of the two control policies, as we show next.

A reactive policy is typically less effective in decreasing the total malaria burden, independently from control effort. This is shown in Figure 3.5, top row, for results that are representative of most of the simulations and comparisons, and correspond to a high number of cases in response to rainfall time series with high variability from one year to the next. More importantly, under high control effort, the reactive policy always misses some epidemics with the consequent decrease in the likelihood of elimination.

There are exceptions however to this typical outcome and these are represented in Figure 3.5, second row, where the reactive policy performs better than that based on rainfall. This pattern is observed whenever the particular rainfall time series has a succession of consistently high years followed by a series of low ones. Thus, a number of successive epidemic events are followed by multiple years with low incidence. Here, the superposition of low transmission and forceful control over a sustained period of time following the large number of cases effectively drives the disease to elimination.

Another uncommon pattern is illustrated at the bottom row of Figure 3.5. This corresponds to situations in which the reemergence of the disease is observed, in a pattern of rainfall essentially opposite to the previous one. Here, rainfall starts with a sequence of low years and therefore low transmission. Thus, there is no intervention and also almost no recorded cases over the first 10 years. After this time, large rainfall events drive two large epidemics. The reactive policy is unable to prevent this kind of re-emergence, whereas the rainfall based control does. Finally, we also note that

even under high levels of control in some scenarios the reactive strategy cannot generate full elimination.

Discussion

In many developing countries that suffer a burden of malaria and other infectious diseases, public health authorities face the dilemma of how best to allocate limited resources so that social, ecological and economic benefits are maximized. Judicious allocation of scarce disease control resources in space and time may therefore improve cost-efficiency. In malarious regions of the world where spraying of insecticides is the preferred method of vector control, IRS usually constitutes a major part of the malaria budget, and its targeted use could apart from reducing expenditure, delay developing resistance of vectors. Although originally designed for a short "all out war" against the malaria parasite, the practices of reactive insecticide application, that is, the practice of spraying in the following transmission season in areas where malaria has exceeded a certain threshold, has been preserved in countries on the Indian subcontinent.

Emphasis on evaluating the efficiency of IRS has been placed on testing the susceptibility of the local vector(s). However, its suitability as long term malaria control strategy for a wide range of malaria conditions has never been assessed. In this paper we demonstrated that in a dry region of India with epidemic prone (unstable) malaria, this policy can prove less effective and more effort-costly than a policy based on risk monitoring using rainfall itself. Specifically, our model simulations show that under a reactive policy, the feedback mechanism between planned intervention and previous incidence generates gaps in control protection as the result of short term achievement followed by favorable climatic conditions (Baeza et al., 2011). The combination of variability in a climate driver with that of intervention facilitates the temporary breakdown of control and the reemergence of the disease within a few years. This can prevent complete elimination even under

intense control efforts. Our modeled results are likely to underestimate the true efficacy of such an alternative policy, as complacency during periods of low or absent malaria interferes with adequate diagnosis and reporting of cases. The frequently reduced availability, and, through lack of practice, competence of microscopists, undermines the reliability and validity of diagnostic services that are the foundation of a reactive intervention policy. In addition, the inadequate stock management of anti-malarial drugs is usually responsible for the high burden in regions of occasional epidemic malaria.

In the past, public health authorities have noted that one of the most common causes of the failure of malaria eradication programs in India has been the lack of appreciation by the authorities of the re-emergence of malaria as a problem (NMEP, 1983) and in a recent review on the resurgence of malaria around the world, Cohen et al. (2012) have concluded that in most cases these patterns follow the relaxation of intervention with budget constraints and resource limitations. Our results underscore that this phenomenon can also occur on shorter time scales of years and in a recurrent fashion under climate variability. Preventing this feedback from operating would be important in the planning for long-term campaigns, whether it originates from an explicit policy or a suite of implicit mechanisms related to the allocation of resources and the perception of risk, especially as the critical point of elimination is approached.

In our model, a rainfall based control strategy, contrary to a reactive one, creates a coupling between the natural periodicity of the disease and its control application. Including this mechanism in the intervention ensures a more efficient and effective targeting of the mosquito population in periods when transmission intensifies. Mosquito/malaria control programs that include the natural inter-annual variability of the disease might increase the epidemiological efficiency of these

interventions in the long term, in support of the proposed importance for long term control strategies of incorporating a better understanding of the ecology of the vector (Ferguson, 2010).

Rainfall has been used in the first half of the 20th century in the epidemic prone Punjab as a guide to rationally distribute scarce resources of quinine and aid its swift availability in districts most at epidemic risk (Swaroop, 1949). For climate based predictions to be operational today, however, these must be obtained with a sufficient lead time relative to the epidemic season, as the logistical planning of vector control requires preparing for the mobilization of large quantities of insecticide, workers and information. Our alternative rainfall based IRS policy would have faced serious practical limitations in the past with the limited number of rainfall gauges deployed, reporting delays, and the requirement of completing the campaigns before the transmission season takes off. Quantitative data of rainfall and, more importantly, its precise distribution can now be monitored almost in real-time using remote sensing. To extend lead times of early-warnings, teleconnections are also sought that have a basis in physical atmospheric and oceanic mechanisms, and take advantage of ocean regions that remotely influence the interannual variability of regional climate drivers of the disease in the area of interest (Cash et al., 2013). For example, Sea Surface Temperatures in the Tropical South Atlantic can help anticipate malaria prone conditions for the arid Northwest India, with a lead time a couple of months longer than that of local rainfall itself (Cash et al., 2013). A similar approach motivated the use of Sea Surface Temperatures in the Tropical Pacific to anticipate anomalous cholera outbreaks in Bangladesh, lengthening the lead time to nine months in this case (Cash et al., 2008; Cash et al., 2009; Cash et al., 2010). The large variability in the relationship between the monthly rainfall distribution and the number of cases observed here in different simulations with different rainfall time series suggests that in addition to interannual variability, short-term (daily) fluctuations also matter (Fig. S3.2B). The consideration of daily temperature

variation has already been emphasized for climate change and malaria (Paaijmans, 2010), and similarly a better understanding of intra-annual rainfall variability is also lacking.

Historically, rainfall related cycles of malaria have been evident in the semi-arid regions of India since the mid nineteenth century. These followed a cycle of 7 years bearing the signature of ENSO (Bouma and van der Kaay, 1996). Interestingly, the potential cycling of malaria in our model results exclusively from a reactive vector control policy. Although this tendency to cycle over ten years post-dates the historical epidemics, it may have contributed to the complexity of the periodic exacerbations of malaria observed in the second half of the 20th century.

Epidemic preparedness has improved in Western India in the last decade with more frequent electronic reporting of malaria cases and rapid selective spraying as response to outbreak within the same year, changes that have strengthened the traditional reactive policy. However, modern rainfall monitoring and rainfall predictions with a longer lead-time may further aid malaria forecasting and preparedness also in the context of fast-response interventions.

In a semi-desert epidemic prone region and under low transmission, the guidance of IRS intervention by local environmental drivers proved to be a more efficient strategy than a delayed "fire-fighting" policy. This conclusion might also apply to other low transmission regions, such as highlands, where malaria dynamics are also highly seasonal and under strong climate forcing. For the particular parameters considered here, our malaria transmission model generated transient multi-year cycles in the size of seasonal outbreaks. These transient cycles are the basis for the observed failure of the reactive control when coincident with some of the large rainfall events. An analytical study of the model dynamics will follow to investigate the generality of this cyclical behavior under different transmission intensities and under different environmental drivers.

Figures

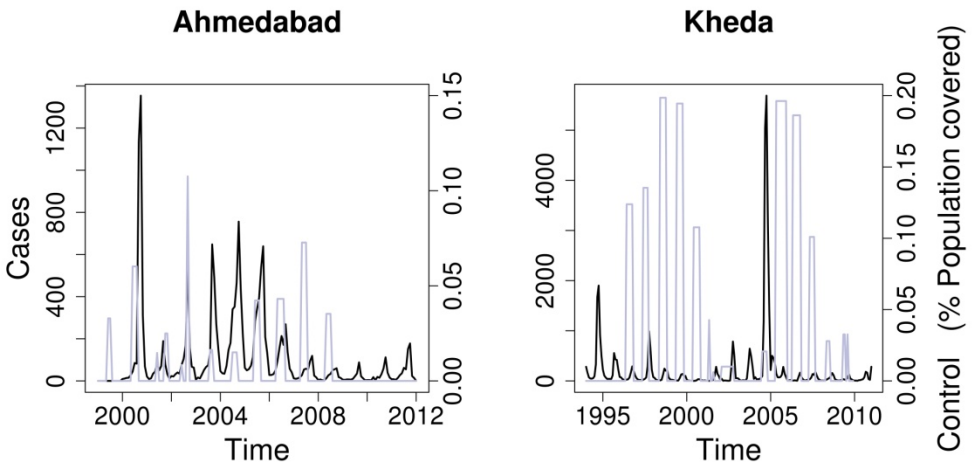

Figure 3. 1. Cases and IRS intervention for the districts of Ahmedabad (2000-2008) and Kheda (1995-2009) in the state of Gujarat.

The grey lines represent the percentage of population covered by IRS intervention over the same period. Cases for Kheda correspond to *Plasmodium falciparum*, whereas those for Ahmedabad are the sum of infections by *Plasmodium falciparum* and *Plasmodium vivax*.

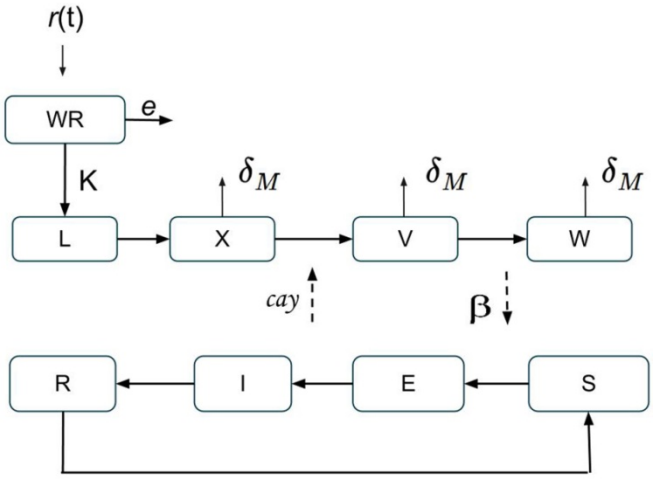

Figure 3. 2 Diagram of the coupled-mosquito malaria transmission model.

Rainfall provides suitable areas for the development of larvae that generate adult mosquitoes. These mosquitoes can be uninfected X, infected V, or infectious W. Transmission occurs at rate β from

mosquitoes to humans and acy from humans to mosquitoes. Total adult mosquito mortality δ_M is a combination of natural and control induced mortality.

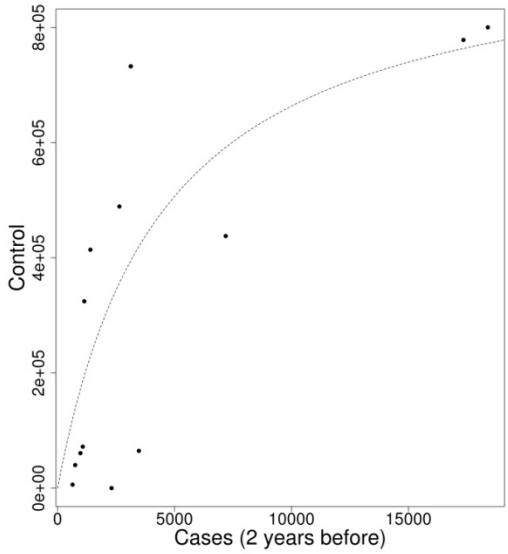

Figure 3.3 Cases vs. Control.

The nonlinear relationship between the number of observed cases reported in the two previous years (x-axis) and the level of population covered by the subsequent IRS intervention (y-axis) for the district of Kheda. The dashed line is a fitted curve of the form $y=Ax/(B+x)$. This functional form is used in our model to simulate the 'reactive' control function δ_C.

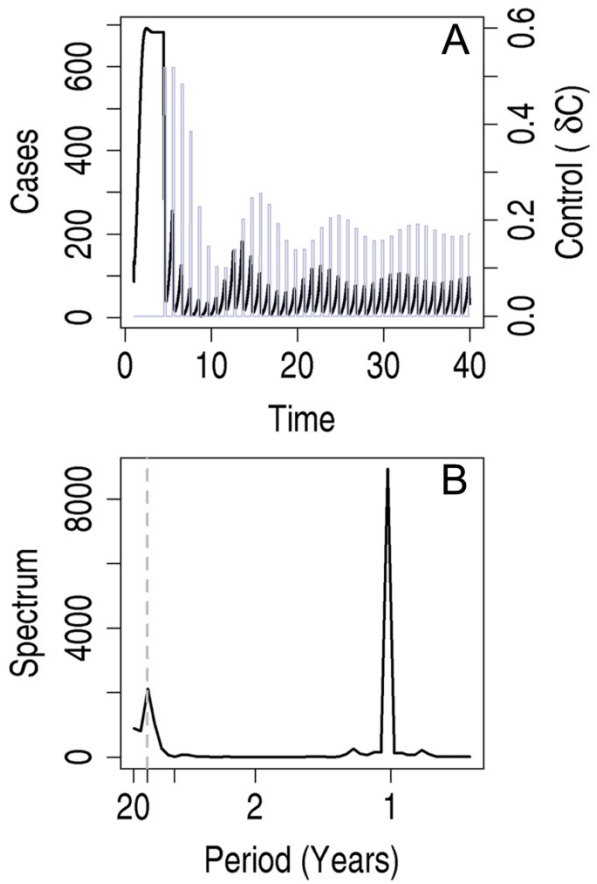

Figure 3.4 Malaria and control dynamics without environmental variability.

The top panel shows the time series of new infected individuals ($\mu_{EI}E_{(t)}$; darker line) and control application in terms of the control-induced mosquito mortality (δ_C; lighter line) at the beginning (transient) phase of the simulation. The graph below shows the periodogram power spectrum for the same time series, that is the power (or variance) corresponding to each frequency. A peak in the spectrum indicates a characteristic cycle in the data. Although an annual cycle of period 1 corresponding to the control-forcing dominates the dynamics, a longer interannual cycle with a 10 years period (grey dashed line) is also present in the transients. This cycle eventually decays, leaving only the one year periodicity, but does so over a relatively large period of time. This figure illustrates that even in the absence of environmental forcing, multi-annual cycles can be generated following a perturbation by reactive control.

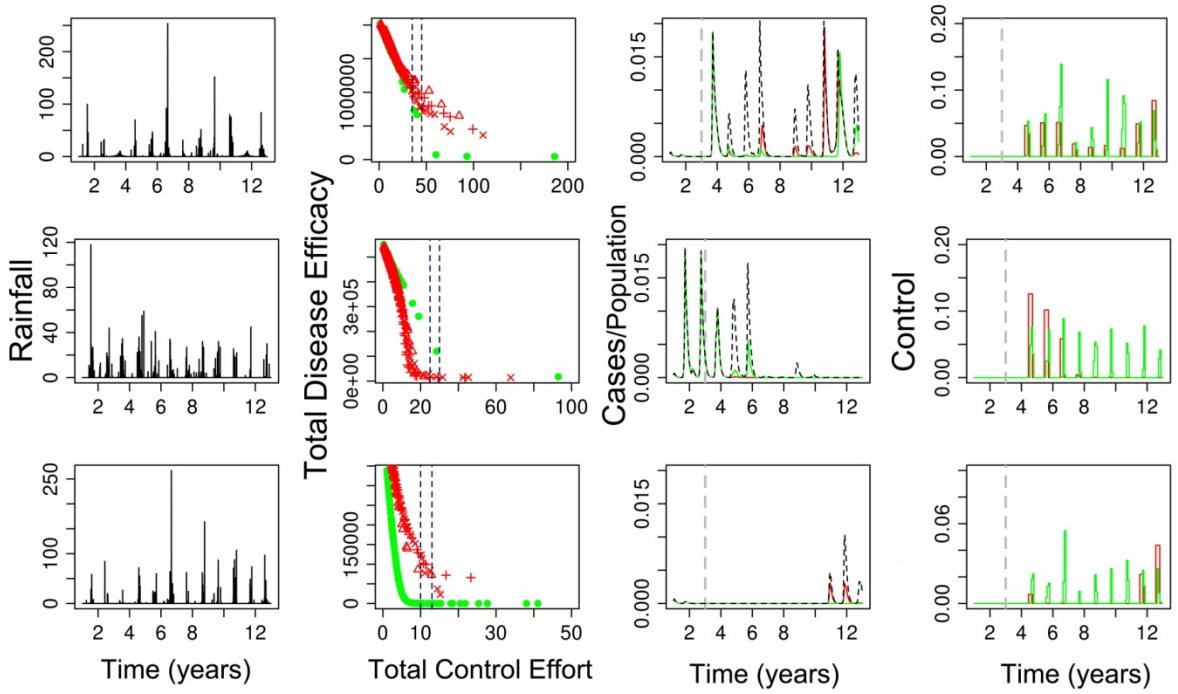

Figure 3.5 Comparison of control strategies.

The first column on the left shows three examples of daily rainfall time series. For each of these, we simulated the model over a range of control intensities (by varying δ_{max} and h) and compared the effect of the two different control strategies, based respectively on (1) total incidence over a window of time in previous years (in red) and (2) total rainfall in the previous month (in green). The plots in the second column summarize results from an ensemble of simulations in the form of a malaria intervention cost-benefits curves. The curves are created by plotting each simulation as a point for the Total Control Effort (TCE; x-axis) and the Total Epidemiological/Disease Efficacy (TDE, y-axis) as defined in the text. Triangles, sums and crosses symbols correspond to different numbers of years for cumulating incidence in the reactive response, respectively 1, 2, and 3 years. Inside the dashed lines are simulations in which the control effort TCE is the same for both strategies. The third column shows the time series of new cases (mid-right column) observed under such similar control intensity. The black dashed line represents the new cases observed in the absence of control. The left-most column shows the control intensity (δ_C) time series for the scenarios described above. The example on the top row illustrates the most common outcome of the ensemble of outcomes observed in the simulations, and corresponds to situations with the larger number of cases in the absence of intervention. In this case rainfall-based control outperforms the reactive strategy. Exceptions to this pattern are illustrated with the cases in the second and third row (as described in the text).

Sub-model	Parameter name	Symbol/Function	Parameter value	Units	References

M	Human blood index	HBI	0.032	Human bood/total blood	DeZoysa et al. (1998); Reisen and Boreham (1982)
M	Gonotrophic cycle	g	2.2	days	Sloof and Herath (1980)
M & H	Human biting rate	$a=HBI/g$	0.015	human bites/day	Calculated from HBI and g; Reisen and Boreham (1982)
M	Probability ($M \rightarrow H$)	b	0.64		Alonso et al. (2010)
M	Average Temperature	Te	28	degree Celcius (°C)	Fan and vanden Dool (2008)
M	Critical developmental temperature for *P. falciparum*	Te_c	16	degree Celcius (°C)	Detinova (1962)
M	Degree days	D	111	days	Detinova (1962)
M	Duration of sporogony	$st=D/(Te-Te\_c)$	9.25	days	Detinova (1962)
M	Sporogony rate	$\gamma=1/st$	0.108	1/day	Calculated from st
M	Probability ($H \rightarrow M$)	c	0.3		Alonso et al. (2010)
M	Per-capita fertility	f	56	L/M	Ramachandra Rao (1984)
M	Larvae intrinsic growth rate	$F=f/g$	25.45	L/M*day	Calculated from f and g
M	Larvae to adult survival probability at Te	s	0.6		Bayoh and Linday 2003
M	Larval developmental time at Te	t_d	10	days	Bayoh and Linday 2003
M	Larval develomental rate at Te	$d_L=s/t_d$	0.06	1/day	Calculated from s and t_d
M	Larval mortality rate	δ_L	0.01	1/day	Russell and Ramachandra Rao (1942b) & Ramachandra Rao (1984)

M	Adults natural mortality	δ_0	0.08	1/day	Russell and Ramachandra Rao (1942a)
M	Local carrying capacity	K_r	6.33	L/m2	Estimated
H	Human lifespan	$1/\delta_H$	55	years	Laneri et al., 2010
H	Human birth rate	$1/br$	55	years	Laneri et al., 2010
H	Per-capita epidemiological transition rates $\{S,E,I,R\}$	μ_{ij}		1/year	Laneri et al. (2010)
H	Human Population	N	203150	person	Taluka (sub-district) average from Census
W	Evaporation rate	e	2.63	1/day	Estimated
W	Surface area	A	3523	Km2	Average taluka area

Table 3.1 Parameter values and their corresponding source for the coupled mosquito-human transmission model.

Those parameters inferred by fitting the model to the time series data (Section 3) are indicated as 'estimated'.

Appendices

Appendix 3.1: Mathematical description of the model

Entomological submodel and water dependence

The model includes a structured mosquito population composed of larvae and three adult stages for uninfected X, infected V and infectious W mosquitoes.

The rate of change in the mosquito population for the different classes is described by the set of differential equations

$$\begin{aligned}
\frac{dP}{dt} &= r_{(t)}A - eP \\
\frac{dL}{dt} &= FM\left(1-\frac{L}{K}\right) - \delta_L L - dL \\
\frac{dX}{dt} &= dL - acyX - \delta_M X \\
\frac{dV}{dt} &= acyX - \gamma V - \delta_M V \\
\frac{dW}{dt} &= \gamma V - \delta_M W
\end{aligned} \quad .$$

Our model includes a water reservoir P based on the rain $r_{(t)}$ that falls in the area A and its evaporation at rate e. This water in the system represents the limiting resource for larvae to grow. In the model we assume that the mosquito larvae population grows logistically with a carrying capacity K proportional to P, the total water in the landscape. $K = K_r P$, where K_r is the maximum number of larvae than can be supported per square meter of surface of the water bodies. Thus, the number of

larvae that can recruit to the system will change according to the availability of water in the landscape.

The mosquito intrinsic growth rate F is a function of the per-capita fertility f, the number of new larvae per female oviposition per unit of time, and the gonotrophic cycle g (Appendix 3.2). Larvae will develop into adults at rate d, and those that do not complete the process will die at rate δ_L.

Epidemiological submodel

This system follows Laneri et al (2010) and is an extension of the classic Ross-Macdonald model that follows the number of individuals in each of four different classes, for susceptible S, exposed E, infected I, and recovered R respectively. The temporal changes in the number of individuals in each class is described by the set of differential equations

$$\frac{dS}{dt} = \mu_{RS}R - \beta S + b_r N - \delta_H S$$
$$\frac{dE}{dt} = \beta S - \mu_{EI}E - \delta_H E$$
$$\frac{dI}{dt} = \mu_{EI}E - \mu_{IR}I - \delta_H I$$
$$\frac{dR}{dt} = \mu_{IR}I - \mu_{RS}R - \delta_H R$$

where μ_{ij} represents the per-capita rate at which individuals in class i move to the class j. The model assumes a loss of immunity in the host at rate $\mu_{RS}>0$. The parameter b_r corresponds to the natural birth rate of the population whose total number of individuals is N (where $N = S+E+I+R$). δ_H denotes the natural mortality of the human host. We assume that the human population remains constant so that $b_r = \delta_H$, and that there is no disease induced mortality. The force of infection, the rate at which susceptible individuals become infected, is defined as

$$\beta = ab\frac{M}{N}\frac{W}{M}$$

where a is the vector's biting rate, the number of mosquito bites per human per day, and b is the probability that a bite by an infectious mosquito will produce an infection in the human host. Thus, aM/N represents the encounter rate and bW/M, the probability that this encounter results in an exposed individual.

Appendix 3.2: Model parameterization

Entomological parameterization

We based our choice of parameters on the entomological literature on the most common malaria vector in this area of India, Anopheles culicifacies. Where parameters were not found for this species, we relied on studies of other Anopheles species.

Biting rate

Compared to other mosquito species, *A. culicifacies* is a relatively poor vector due to its zoophilic behavior. Using the derivation calculation of Reisen and Boreham (1982), the human biting rate of *A. culicifacies* can be estimated as

$$a=HBI/g$$

where HBI is the anthropophagic or human blood index. This is the proportion of human-positive blood meals among all over positive reactive mosquito meals. g is the gonotrophic cycle. This is the number of days that it takes a mosquito to bite, lay the eggs and find the next victim. Estimation of the human biting index has been conducted by Reisen and Boreham (1982) in the province of Punjab and by DeZoysa et al. (1998) in Sri Lanka. Sloof and Herath (1980) and Ramachandra Rao (1984) have provided estimates of the gonotrophic cycle for A. culicifacies. They found values of 1.7 and 2.2 days. In this work we use the maximum of this range.

Mosquito growth rate

The population growth rate of larvae depends on their intrinsic growth rate and the death rate. The intrinsic growth rate F is defined as

$$F=f/g$$

Where f denotes the per-capita adult fecundity (new larvae per adult mosquito oviposition) and $1/g$ is the length of the gonotrophic cycle. f has been estimated by Russell and Ramachandra Rao (1942a). The death rate δ_L was estimated by Russell and Ramachandra Rao (1942b).

Parasite developmental rate in the mosquito

The rate at which sporozoites develop depends on the *Plasmodium* species as well as on the temperature of the area. The derivation of this rate for *Plasmodium falciparum* was based on Detinova (1962), which defines the days needed for sporozoites to develop as follows:

$$1/\gamma = D/(T-T_C)$$

Where D measures the number of days it takes the parasite to develop under conditions of 1 degree above the critical temperature T_C, the value below which sporozoites will not form. T is the temperature of the environment. We maintain this constant throughout the experiment at a value of 28 degrees. This is the average surface temperature in the district of Kutch during the last ten years. Table 1 summarizes each parameter, as well as the reference from which the respective parameters were obtained.

Statistical estimation of unknown parameters

We consider that the incidence values observed by the surveillance system are those produced by the model but with a measurement error. Specifically, let y be the total number of new cases recorded by the surveillance system in a particular month. We model these cases as

$$yt \sim negbin(Ct, \sigma),$$

with

$$Ct = \rho \int_{t-1}^{t} \mu_{EI} E_{(t)} dt .$$

ρ gives the reporting rate of the cases C_t, and $negbin(a,b)$ indicates a negative binomial distribution with mean a and variance $a + a^2b$. This distribution allows for overdispersion of the error and maintains positive and discrete counting of cases. Our estimation approach compares the model predictions against the monthly cases reported in the district of Barmer between the years 1994 and 2010, and maximizes a likelihood defined as the probability of observing the data under this model. We choose this particular district of Rajasthan because of negligible control levels, and therefore presumably no interference by control in the estimation of the particular parameters. The maximum likelihood estimation of the parameter K_r, e, ρ, and the size parameter of the binomial distribution function were obtained using the Nelder-Mead algorithm in the function Optim, in R (Bolker, 2008). This method estimates parameter values that generate similar dynamics to those observed in the data (Fig. S3.3).

References for model parameterization

Alonso, D., Bouma, M.J., Pascual, M., 2011. Epidemic malaria and warmer temperatures in recent decades in an East African highland. Proc Biol Sci 278, 1661-1669.

Bayoh, M.N., Lindsay, S.W., 2004. Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito Anopheles gambiae in the laboratory. Med Vet Entomol 18, 174-179.

Bolker, B., 2008. Ecological Models and Data in R. Princeton University Press, Princeton, NJ.

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.

Dezoysa, A.P.K., Herath, P.R.J., Abhayawardana, T.A., Padmalal, U.K.G.K., Mendis, K.N., 1988. Modulation of Human Malaria Transmission by Anti-Gamete Transmission Blocking Immunity. T Roy Soc Trop Med H 82, 548-553.

Fan, Y., van den Dool, H., 2008. A global monthly land surface air temperature analysis for 1948-present. J Geophys Res-Atmos 113.

Laneri, K., Bhadra, A., Ionides, E.L., Bouma, M., Dhiman, R.C., Yadav, R.S., Pascual, M., 2010. Forcing versus feedback: epidemic malaria and monsoon rains in northwest India. PLoS Comput Biol 6, e1000898.

Ramachandra Rao, T., 1984. The Anophelines of India. Indian Council of Medical Research, New Delhi.

Reisen, W.K., Boreham, P.F.L., 1982. Estimates of Malaria Vectorial Capacity for Anopheles culicifacies and Anopheles stephensi in Rural Punjab Province, Pakistan. J Med Entomol 19, 98-103.

Russell, P.F., Ramachandra Rao, T., 1942. On the ecology of larvae of Anopheles culecifacies Giles, in borrowpits. Bulletin of Entomological Research 32, 341-361.

Russell, P.F., Ramachandra Rao, T., 1942. A study of the density of Anopheles culicifacies in relation to malaria endemisity. American Journal of Tropical Medicine and Hygiene 22, 535-558.

Slooff, R., Herath, P.R., 1980. Ovarian development and biting frequency in Anopheles culicifacies Giles in Sri Lanka. Trop Geogr Med 32, 306-311.

Appendix 3.3: Quantification of the effort and effectiveness assessment of control intervention

In the model the direct quantity of the intervention effort per-day is the vector mortality induced by the control. On a per-capita basis, this corresponds to $\delta_C \Delta t$. Thus, during a period of time the total control effort (TCE) can be defined as

$$TCE = \sum \delta c \Delta t$$

This quantity measures the total effort applied to reduce the mosquito population for the total duration of control (Δt; e.g. 10 years). TCE can also be viewed as reflecting the total resources needed for vector control over this time.

A more intuitive way to quantify this effort would be to quantify the average total population that is effectively covered per year (Supplementary data, Figure 3.4). We consider that over a season, and therefore the yearly round of intervention, the rate of population coverage per day, denoted by λ, is proportional to the average control-induced mortality rate, so that

$$\alpha \frac{\sum \delta c}{T} = \lambda$$

where the sum is over the season of $T=90$ days, and α is the proportionality constant.

Thus the temporal change in the population that remains unprotected by IRS intervention follows an exponential model of the form

$$\frac{dN}{dt} = -\lambda N$$

and the total population covered can be calculated as

$$N_C = N(1 - e^{-\lambda t})$$

To set α, we consider that a constant application of the maximum control rate results in a total population covered after $T=90$ days of N_{max}. It follows that

$$\frac{N \max}{N} = e^{\frac{-\alpha \sum \delta \max}{T} T}$$

and

$$\alpha = -\frac{\sum \delta_{\max}}{\log(\frac{N \max}{N})}$$

The epidemiological assessment of a control intervention is defined as the decrease in the total number of new cases over the period in which the intervention was applied. The new cases at each time step are $\mu_{EI}E_{(t)}dt$. Integrating these cases over the intervention period ΔT we obtain,

$$\int_{\Delta t} \mu_{EI} E(t) dt$$

the Total Disease Efficiency (TDE).

Appendix 3.4 Supportive figures

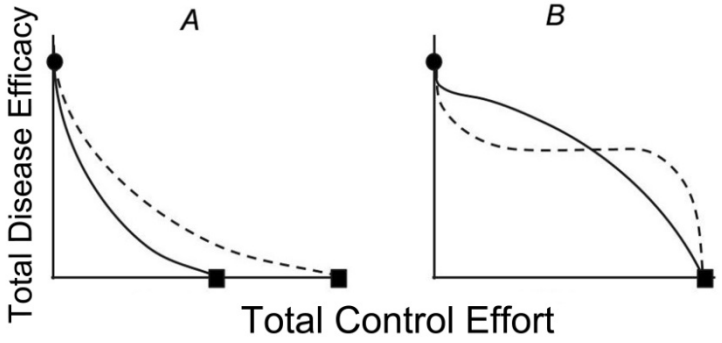

Figure S3 1 Malaria intervention cost-benefit curve

The curve is constructed by calculating the total disease efficacy (TDE, y axis) for a particular total control effort (TCE, x axis). The figure depicts two situations that can arise when comparing two hypothetical strategies under the same conditions: In scenario A, one of the strategies (dashed line) is always more efficient than the other (solid line), no matter how much effort is applied, and elimination is reached at a low level of control effort. Scenario B depicts a situation in which one strategy is better than the other under low intensity, but the opposite ranking happens under high intensity. Thus control strategies may vary in the shape of the MICB curve, which allows for comparison between different strategies. We used this curve to quantify the two strategies of control intervention over a range of situations dictated by observed patterns of rainfall variability.

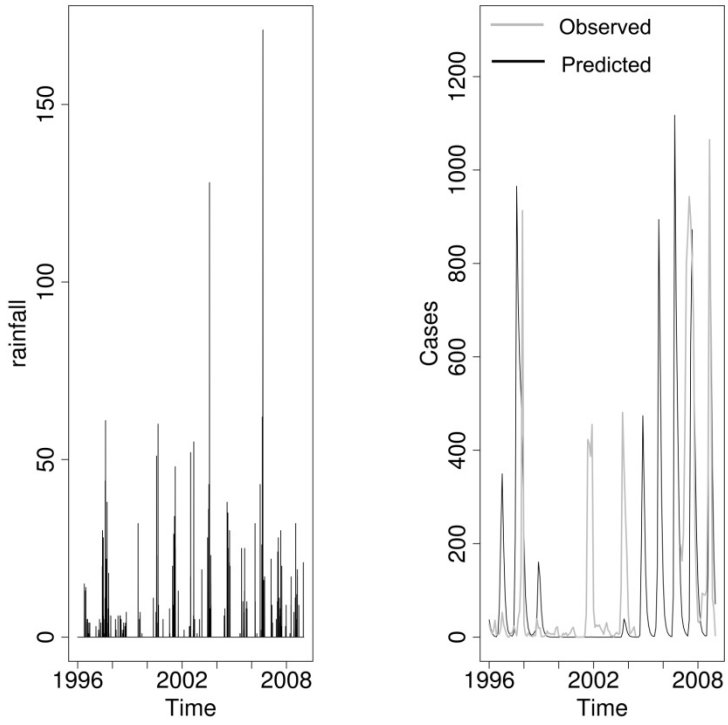

Figure S3.2 Maximal likelihood estimation of parameter Kr and e

To the left the daily rainfall record for the station in Barmer (Log likelihood = -2721.047). The plot on the right shows the observed (grey line) and estimated (black line) new cases of Plasmodium falciparum, using the "best" parameters. The parameter values obtained for Kr and e where 6.339 and 2.6 respectively. No data between 2004 and 2006 was possible to obtain. The model does not capture the 2 epidemics of 2001 and 2003, a relative dry period according to the rainfall record for this station.

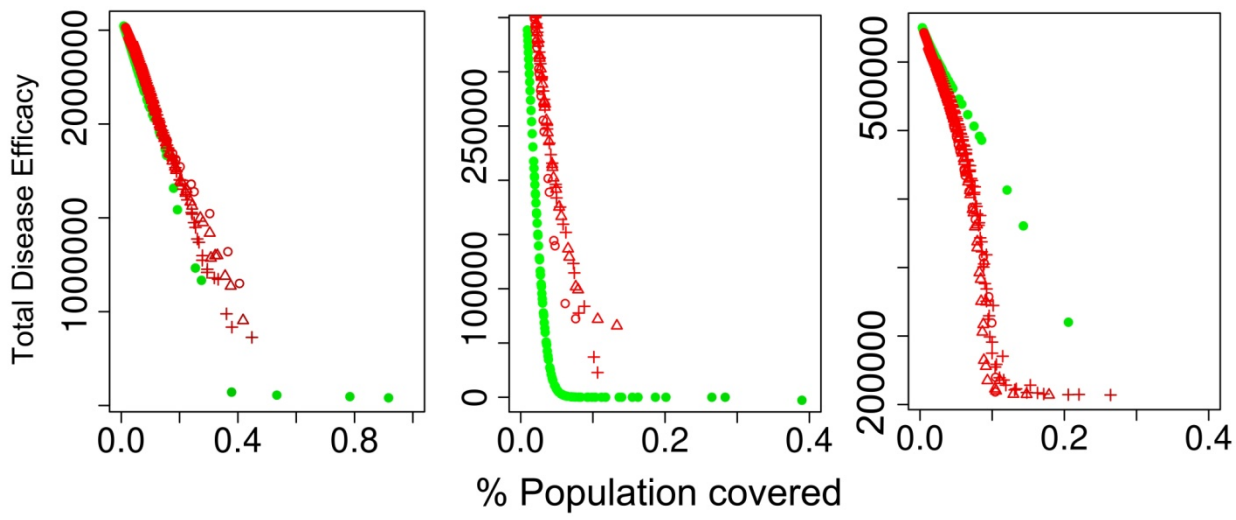

Figure S3.3 Comparison of MICB curves for total population covered

Total population covered by intervention (Appendix 3.2) replaces here the measurement of total control effort in the x-axis of these curves. To calculate the conversion factor α we assumed a maximal population covered over a season of 90%. Otherwise, the simulations are the same as those of Figure 3.5. The leftmost plot exemplifies the most common outcome observed in the ensemble of simulations. Towards the right, the second and third plots correspond to the two other scenarios presented in figure 3.5.

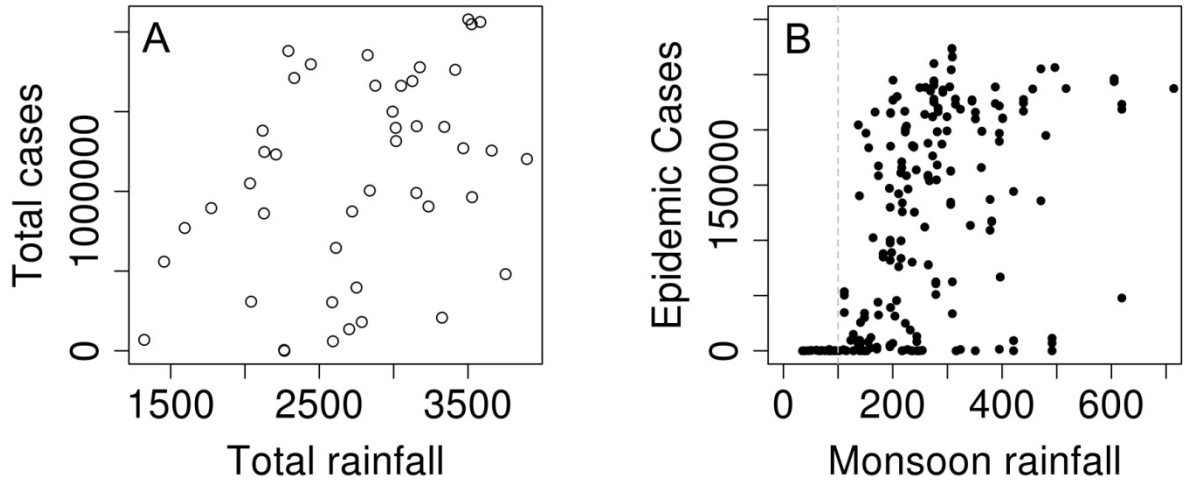

Figure S3.4 Rainfall variability and malaria epidemics

Panel A depicts the relationship between total rainfall over a period of 13 years (x-axis) and the total cases observed (with no control) over the same period (y-axis). Each point represents one of the rainfall scenarios in which control interventions were applied. Panel B shows monsoon rainfall (June-August) and epidemic cases (Sep-Dec). Each point is now one year. The grey dashed line shows an arbitrary rainfall threshold of 100mm. Bellow this threshold no epidemic were observed. Both plots exemplify the importance of the distribution of rainfall events on the magnitude and the occurrence of epidemic events.

References

Alonso, D., Bouma, M.J., Pascual, M., 2011. Epidemic malaria and warmer temperatures in recent decades in an East African highland. Proc Biol Sci 278, 1661-1669.

Baeza, A., Bouma, M.J., Dobson, A.P., Dhiman, R., Srivastava, H.C., Pascual, M., 2011. Climate forcing and desert malaria: the effect of irrigation. Malaria J 10, 190.

Barik, T.K., Sahu, B., Swain, V., 2009. A review on Anopheles culicifacies: From bionomics to control with special reference to Indian subcontinent. Acta Trop 109, 87-97.

Bouma, M.J., van der Kaay, H.J., 1996. The El Nino Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics? Trop Med Int Health 1, 86-96.

Cash, B., Rodo, X., Ballaster, J., Bouma, M., Baeza, A., Dhiman, R.C., Pascual, M., 2013. Malaria epidemics and the influence of the Tropical South Atlantic on the Indian monsoons. Nat Clim Chang In press.

Cash, B.A., Rodo, X., Kinter, J.L., 2008. Links between tropical pacific SST and cholera incidence in Bangladesh: Role of the eastern and central tropical Pacific. J Climate 21, 4647-4663.

Cash, B.A., Rodo, X., Kinter, J.L., 2009. Links between Tropical Pacific SST and Cholera Incidence in Bangladesh: Role of the Western Tropical and Central Extratropical Pacific. J Climate 22, 1641-1660.

Cash, B.A., Rodo, X., Kinter, J.L., Yunus, M., 2010. Disentangling the Impact of ENSO and Indian Ocean Variability on the Regional Climate of Bangladesh: Implications for Cholera Risk. J Climate 23, 2817-2831.

Cohen, J.M., Smith, D.L., Cotter, C., Ward, A., Yamey, G., Sabot, O.J., Moonen, B., 2012. Malaria resurgence: a systematic review and assessment of its causes. Malaria J 11, 122.

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.

Ferguson, H.M., Dornhaus, A., Beeche, A., Borgemeister, C., Gottlieb, M., Mulla, M.S., Gimnig, J.E., Fish, D., Killeen, G.F., 2010. Ecology: a prerequisite for malaria elimination and eradication. PLoS Med 7, e1000303.

Goodman, C.A., Mnzava, A.E., Dlamini, S.S., Sharp, B.L., Mthembu, D.J., Gumede, J.K., 2001. Comparison of the cost and cost-effectiveness of insecticide-treated bednets and residual house-spraying in KwaZulu-Natal, South Africa. Trop Med Int Health 6, 280-295.

Guyatt, H.L., Kinnear, J., Burini, M., Snow, R.W., 2002. A comparative cost analysis of insecticide-treated nets and indoor residual spraying in highland Kenya. Health Policy Plann 17, 144-153.

Hamusse, S.D., Balcha, T.T., Belachew, T., 2012. The impact of indoor residual spraying on malaria incidence in East Shoa Zone, Ethiopia. Global Health Action 5, 1-8.

Kigozi, R., Baxi, S.M., Gasasira, A., Sserwanga, A., Kakeeto, S., Nasr, S., Rubahika, D., Dissanayake, G., Kamya, M.R., Filler, S., Dorsey, G., 2012. Indoor Residual Spraying of Insecticide and Malaria Morbidity in a High Transmission Intensity Area of Uganda. Plos One 7.

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 Comput Biol. 6,9.

Najera, J.A., Gonzalez-Silva, M., Alonso, P.L., 2011. Some lessons for the future from the Global Malaria Eradication Programme (1955-1969). PLoS Med 8, e1000412.

NMEP, 1983. Malaria and its Control in India, in: Programe, N.M.E. (Ed.). Goverment of India, Delhi.

NMEP, 1995. Operational manual for malaria action programme, in: National malaria eradication programme (Ed.). Ministry of Health and Family Welfare,Government of India, New Delhi.

Paaijmans, K.P., Blanford, S., Bell, A.S., Blanford, J.I., Read, A.F., Thomas, M.B., 2010. Influence of climate on malaria transmission depends on daily temperature variation. P Natl Acad Sci USA 107, 15135-15139.

Pedercini, M., Blanco, S.M., Kopainsky, B., 2011. Application of the Malaria Management Model to the Analysis of Costs and Benefits of DDT versus Non-DDT Malaria Control. Plos One 6.

Ramachandra Rao, T., 1984. The Anophelines of India. Indian Council of Medical Research, New Delhi.

Russell, P.F., 1936. Malaria in India: Impressions from a tour. American Journal of Tropical Medicine 16, 653-664.

Sharma, R.C., Gautam, A.S., Bhatt, R.M., Gupta, D.K., Sharma, V.P., 1991. The Kheda Malaria Project - the Case for Environmental-Control. Health Policy Plann 6, 262-270.

Sharma, V.P., Sharma, R.C., 1989. Community Based Bioenvironmental Control of Malaria in Kheda District, Gujarat, India. J Am Mosquito Contr 5, 514-521.

Smith, T., Maire, N., Ross, A., Penny, M., Chitnis, N., Schapira, A., Studer, A., Genton, B., Lengeler, C., Tediosi, F., De Savigny, D., Tanner, M., 2008. Towards a comprehensive simulation model of malaria epidemiology and control. Parasitology 135, 1507-1516.

Swaroop, S., 1949. Forecasting of epidemic malaria in the Punjab, India. Am J Trop Med Hyg 29, 1-17.

White, M.T., Conteh, L., Cibulskis, R., Ghani, A.C., 2011. Costs and cost-effectiveness of malaria control interventions - a systematic review. Malaria J 10.

WHO, 2012. World Malaria Report 2012. World Health Organization Geneva.

Worrall, E., Connor, S.J., Thomson, M.C., 2007. A model to simulate the impact of timing, coverage and transmission intensity on the effectiveness of indoor residual spraying (IRS) for malaria control. Tropical Medicine & International Health 12, 75-88.

Chapter 4

Long-lasting transition toward malaria elimination under irrigation development

Abstract

In arid areas, people living in the proximity of irrigation infrastructure are potentially exposed to a higher risk of malaria due to changes in ecohydrological conditions that lead to increased vector abundance. However, irrigation provides a pathway to economic prosperity that over longer time scales is expected to counteract these negative effects. A better understanding of this transition between increased malaria risk and regional elimination, in particular whether it is slow or abrupt, is relevant to sustainable development and disease management. By relying on space as a surrogate for stages of time, we investigate this transition in a semidesert region of India where a megairrigation project is underway and expected to cover more than 1,900 million hectares and benefit around 1 million farmers. Based on spatio-temporal epidemiological cases of Plasmodium vivax malaria and land-use irrigation from remote sensing sources, we show that this transition is characterized by an enhanced risk in areas adjacent to the trunk of the irrigation network, despite a forceful and costly insecticide-based control. Moreover, this transition between climate-driven epidemics and sustained low risk has already lasted a decade. Given the magnitude of these projects, these results suggest that increased health costs have to be planned for over a long time horizon. They further highlight the need to integrate assessments of both health and environmental impacts to guide adaptive mitigation

strategies. Our results should help to define and track these transitions in other arid parts of the world subjected to similar tradeoffs.

Introduction

In agricultural economies, food insecurity imposes a strong pressure to extend agriculture to marginal areas. In low-rainfall regions, irrigation offers considerable rewards, creating water resources for irrigation and other usage. On either side of the border of India and Pakistan, an extensive arid region is intersected by large rivers carrying water from the Himalayan glaciers and rainwater from a short rainy season. Hundreds of millions of people depend on the southwestern monsoon for their survival. Over the centuries, its periodic failure and severe ensuing drought and famine conditions have provided a strong incentive to develop and extend irrigation. The continuing expansion of the Indian population in the 21st century adds further pressure to optimize the country's water resources for agriculture, fisheries, industrial and general usage.

The development of water resources, in the context of malaria and dengue, exemplifies a central challenge in sustainability science: how can we achieve socioeconomic development based on land-use transformations, with concomitant increases in human well-being, when the transformations can compromise ecosystem services and human health for present and future generations?

For arid regions, concerns have been raised about the consequences for malaria epidemiology resulting from ecological changes subsequent to the arrival of irrigation, with several studies reporting local increases in prevalence and parasitemia (Jayaraman, 1982; Yohannes et al., 2005; Yewhalaw et al., 2009; Srivastava et al., 2009). By increasing surface water levels, irrigation modifies eco-hydrological conditions of the landscape, creating more standing bodies of water for longer

periods of time (Sharma; 2001), thereby increasing the abundance of mosquito breeding sites and adult vector populations (Yadav et al., 1989; Amerasinghe et al., 1992; Amerasinghe and Indrajith, 1994; Konradsen et al., 1998; Herrel et al., 2001; Claborn et al., 2002; Dia et al., 2010; Kibret et al., 2010). In addition, agricultural development can, increase the frequency of human-vector contact, when human labor and mosquito breeding seasons are synchronized (Doannio et al., 2002), and promotes migration to newly irrigated areas (Shah and Singh, 2004), thus changing the spatial scale of malaria transmission. The global population at risk of contracting malaria due to proximity to irrigation infrastructure has been estimated at around 800 million, which represents approximately 12% of the global malaria burden (Keiser et al., 2005).

In these dry, fragile ecosystems, where increase in water availability from rainfall is the limiting factor for malaria transmission, irrigation infrastructure can drastically alter mosquito population abundance to levels above the threshold needed to maintain malaria transmission. In northwestern India, an increase in domestic and para-domestic water storage support *Anopheles stephensi*, India's urban malaria vector. In the desert areas of Rajasthan, a rise in malaria associated with *A. culicifacies* has been reported following large-scale irrigation development by the Indira Gandhi canal (Tyagi and Yadav, 2001; Tyagi, 2004). Currently, seasonal epidemics of mainly *Plasmodium vivax* occur in these semi-deserts, at the edge of the geographic distribution of the disease. At these fringes, *P.vivax,* with its relapses (White, 2011), has a competitive advantage over *P. falciparum* the more malignant form of malaria that also occurs in India. *P. falciparum*, less able to persist in unfavorable transmission conditions, is less consistently present in this region, and displays more interannual fluctuations. Because we are mostly concerned with the seminal changes of malaria resulting from irrigation in the area, we have in this study focused on changes of *P. vivax*.

Despite these environmental changes favoring transmission, historical studies in the semi-arid regions of Pakistan and the former Punjab Province have also documented irrigation-based development having the opposite effect. With the arrival of irrigation in the latter half of the 19th century, large-scale migration and colonization followed, and the regional malaria burden initially rose dramatically (Darling, 1925; de Zulueta et al., 1980). However, the Eastern and Western Punjab of former British India, part of India and Pakistan, respectively, since 1947, are now rich agricultural regions with low malaria prevalence.

Apparently, ecological and socioeconomic factors alter the dynamics and distribution of the parasite, the vector and human susceptibility, from the arrival of irrigation to its long-term stabilization. Although previous studies taken together suggest that different effects operate over different time horizons (Klinkenberg et al., 2004), observations typically correspond to different stages of the developmental process, local in time or in space. No study to date has followed remotely sensed irrigation characteristics and malaria levels simultaneously over a period of time that encompasses these different stages, or over large regions whose irrigation gradient provides a surrogate for these temporal stages.

The long-term malaria surveillance program in arid northwestern India provides a unique spatio-temporal data set for considering just such a gradient in irrigation intensification over the last fifteen years. By following the changes in malaria incidence, vegetation and socioeconomic data at the level of sub-districts, we identify a transition phase toward sustainable low risk (elimination) lasting for more than a decade, and characterized by an enhanced environmental malaria risk despite intensive mosquito control efforts. This protracted phase highlights the need for considering health impacts in the long-term planning, assessment and mitigation of projects related to water resources.

Results

Study area

This research was conducted in a semi-arid area of the northeast part of the state of Gujarat, in the districts of Kutch, Banas Kantha and Patan (Fig. S1). These districts are divided into 25 sub-districts, or *talukas,* for which the epidemiological surveillance data is aggregated (Fig. 4.1 Panel A). Due to the influence of the southwest monsoon, rainfall is extremely variable from year to year in the region. Annual rainfall ranges from 120 in western part of Kutch to 600 mm in the eastern border of Banas Kantha, and is concentrated between the months of June and September. This climatic pattern creates strong seasonal malaria and high variability between years (Laneri et al., 2010; Baeza et al., 2011). The peak of the epidemics usually varies from August to November, depending on the parasite species and the timing and length of the rainfall season (Bhatt et al., 2008).

Spatial-temporal patterns of malaria and irrigation

Strong and long-lasting differences exist in malaria population dynamics between the talukas located in the eastern and the western parts of the study area. We identified these two main regions (depicted in red and green colors in Fig. 4.1 Panel A) using a Bayesian statistical method that identifies groups of spatial locations (*talukas*) whose temporal disease dynamics are similar (Baskerville et al., 2013; Methods). These differences, observed for *P. vivax* reflect distinct overall incidence levels, as illustrated by the time series of the monthly cases accumulated for each group (Fig. 4.1 Panel B). The identified grouping was robust to changes in modeling assumptions (Methods). A similar pattern was also observed for *P. falciparum*, the species with the lower and less consistent regional presence. Throughout the entire region a slow declining trend is also apparent, presumably as the result of the intense level of mosquito control intervention in the area (Dattani et al., 2009). However, due to the dynamic interplay between rainfall and control intervention in the

region, which can cause a tendency to cycle at decadal time scales (Baeza et al., 2013), caution is needed in extrapolating this trend to the future.

The identified differences in malaria population dynamics are strongly coherent with long-term irrigation patterns. Figure 4.1 Panel C shows a map of irrigated areas for the year 2009 obtained from remote sensing information and the spectral signature of important crops in the area (Methods). In this figure, two regions are also apparent and closely map onto the malaria clusters. Whereas the westernmost sub-districts in Banas Kantha and Patan are intensively irrigated (mainly from deep-wells) and have been for over 30 years (Fig. S4.2), the eastern ones, in Kutch and parts of Patan and Banas Kantha, have little irrigation. Thus, we refer to these two regions as "mature irrigated" and "low-irrigated", respectively. Figure 4.1 Panel D shows the epidemic vulnerability of the low irrigated areas after above normal monsoon rains for the year 2003. The incidence of malaria recorded in the low-irrigated zone (red dots) was more than twice the burden observed in the mature irrigated one (green dots). This particular large epidemic followed a very dry year (2002) with little malaria, and a reduced (re-active) insecticide coverage response in 2003 (see Methods for insecticide application policy). Thus, although this particular year exhibits an increase in incidence at every location, the ranking in epidemic size is consistent with the two spatial clusters defined over the whole period of study, despite differences in local climatic conditions or control intervention.

Are these differences in malaria risk between the mature-irrigated and low-irrigated areas associated with the overall level of development and wealth of these two main regions? Table 4.1 summarizes the results from a statistical comparison between high and low malaria risk zones in terms of socioeconomic indicators for the year 2001 (Methods). In general, high-risk talukas had a lower proportion of literate people and more limited access to sources of improved drinking water. In addition, no significant differences were observed between the percentage of people with access to

state-supplied public health and medical facilities, and education. The differences between the proportion of people with access to credit from agricultural societies, however, are pronounced, with 80% of the farming communities living in the low-risk malaria area having access to credit for improving agricultural practices, compared with only 60% of those living in the high-risk area.

Change in irrigation and malaria risk in last decade

Currently, a large irrigation project is under expansion at the edge of the mature irrigated and low-irrigated sub-districts (Fig. 4.2 Panel A). Most of the low-irrigated territory is expected to receive the complete intended water supply for agriculture and human consumption by the year 2014. Since we do not have direct information on yearly irrigation, and therefore, on the annual change that has occurred during the same period of time for which the epidemiological and control data were obtained, we estimated these changes by relying on the yearly variation in vegetation coverage during the peak of the irrigation (*Rabi*) season (January). During this period without rain, most of the satellite-observed vegetation in this arid environment should be the result of irrigated crops, an assumption that is supported by the observation of the seasonality of the Normalized Difference Vegetation Index (NDVI) in the area (Fig. S4.3), and the spatial clustering of the vegetation outside the rainy season. Taking advantage of this seasonality and the map of true irrigation from the year 2009, we develop a classification based on a threshold value of NDVI to separate mature irrigated and low-irrigated locations (pixels) (Methods). Figure 4.2 Panel A shows the areas classified as irrigated and non-irrigated for the years 2001, and Panel B shows in red the areas classified as irrigated in 2009 but not in 2001. This latter map highlights that most of the change in irrigation (outside the monsoon season) during the last decade took place in the fringe zone between the mature irrigated and low-irrigated regions. These ecological changes occurring on the border match the path of the main irrigation canal of this mega-irrigation project taking place in the area of study (Fig. 4.2 Panel A). This increment in vegetation was especially pronounced in the southernmost

talukas where the principal canal first arrived more than a decade ago (Fig. 4.2 Panel B, and fig S4.4), but the canal is expanding to the north, to the un-irrigated districts of Rajasthan, with a side-branch to the talukas in Kutch (Fig. 4.2 Panel A).

How do these changes in irrigation within the last decade correlate to the changes in malaria risk during the same period? The spatial distribution of malaria burden, especially within the most malarious and low-irrigated talukas, can be examined in more detail by dividing the time series into two periods (2000–2005 and 2006–2010), accumulating the incidence during the two periods and then normalizing them by the corresponding value for the sub-district with the highest burden. In this way we can compare the spatial distribution of the cases independently from the yearly variation in malaria due to climate conditions or control application. These maps of spatial relative risk (Fig. 4.3, Panel A for 2006–2010 and Fig. S4.6 for 2000–2005) also highlight this same boundary region in the middle of the mature irrigated and low-irrigated clusters. In this fringe zone, the incidence of *Plasmodium vivax* was higher than in the low-irrigated more western talukas of Kutch, particularly in the last part of the decade. *P. falciparum* does not show this clear distinction especially for the more recent years (Fig. S4.5) given its low incidence.

Strikingly, this transition region with highest levels of malaria is also observed in the efforts to control the disease. Figure 4.3 (Panel B) shows the percentage of population covered by mosquito indoor residual spray application (IRS) in each sub-district between 2006 and 2010 (see fig. S4.7 for 2000 and 2005). In the transition zone up to 80% of the population qualifies for spraying; signifying the raised levels of Public Health efforts to address increased levels of malaria. This contrasts with the low-irrigated regions to the West, and particularly with the mature-irrigated areas that required least intervention. This clearly highlights that this zone is epidemiologically different from the other two regions previously described.

Based on the incidence and control of malaria and the ecological changes observed, three main eco-epidemiological zones can be recognized (Fig. 4.3 Panel C): (1) an area of low disease burden and low requirement of control coverage, corresponding to sub-districts that have been irrigated over a long period of time (several decades; Fig. S4.2); (2) a transition region with high incidence despite high control coverage (IRS coverage of 80–90% of the targeted population) in the talukas adjacent to the advancing irrigation project; and (3) and a low-irrigation area in Kutch with variable rainfall-dependent seasonal outbreaks and intermediate (variable) levels of required intervention.

The transition zone in malaria risk, between the "mature-irrigated" low-malaria risk and the "low-irrigated" and high-risk talukas, is characterized by an increased environmental risk that has now already lasted for at least a decade.

Discussion

We have shown that enhanced disease risk despite heightened intervention is concentrated in the sub-districts adjacent to the main canal that have experienced the most pronounced change in irrigation levels in the last decade. By contrast, a sustained low disease burden, not requiring high coverage with vector control, is found in neighboring sub-districts that have been irrigated for at least three decades. This long-lasting transition phase is consistent with the historical changes reported for the Punjab, once the center of some of the most devastating malaria epidemics on record (Swaroop, 1949; Bouma and van der Kaay, 1996) and today one of the more prosperous food-producing parts of India, with low endemic levels of the disease. These historical changes took place in a period in excess of half a century and their dynamics remain only partially understood.

A better understanding of socioeconomic and ecological differences between recently irrigated and mature irrigation areas could provide the means to reduce the malaria burden and shorten the "transition phase" (Russell and Ramachandra Rao, 1942). On the environmental side, changes in vectors' ecology following increases in surface water levels and soil salinity have been proposed for the decrease in malaria risk in the Punjab (Herrel et al., 2004; Klinkenberg et al., 2004). Historically an enhanced malaria risk has also been related to construction activities, such as the local production of bricks and road works, that create vector's habitats by altering the landscape and fall under the "tropical aggregation of labor" (Molineaux, 1998). For the expanding population in the command area of the Sardar Sarovar Project in Gujarat this has been recognized as a problem (Srivastava et al., 2009), along with the seepage of water from improperly constructed and maintained irrigation structures.

On the socioeconomic side, our observations based on the 2001 census data, show that the in the cluster of sub-districts irrigated for at least three decades, farmers had easier access to agricultural credit, and populations benefited from higher literacy levels and better access to clean water. (Similar analyses for the 2011 census would be of interest when these data become available). By extending periods of water availability beyond the rainy season, irrigation creates the possibility of multi-crop rotations and also facilitates the use of high yield varieties with superior economic return (FAO, 2003). Over the years, these changes should ensure food security and more stable income, leading to improved socioeconomic conditions and the ability of the population to seek health care and afford preventive measures, eventually spiraling out of the "malaria poverty trap" (Bonds et al., 2010; Sachs and Malaney, 2002).

Regardless of specific mechanisms, this long-lasting transition from high risk to low disease prevalence is usually accompanied by a resource-intensive vector control operation mainly based on

the use of insecticides. Given the timescale of this transition, efficient and long term policies and sustainable intervention capability is required, especially in usually poor semi-arid areas undergoing intensive and rapid ecological change, where resources to support long-term control interventions can be limited. The exacerbation of malaria risk combined with a deceleration of economic development or even the temporary relaxation of control (Cohen et al., 2012) may lengthen the transition phase and allow for epidemic surprises in years of anomalous high rainfall (Baeza et al., 2013), in a regression back to climate-driven dynamics. This is a concern especially in areas where groundwater extraction surpasses the current capacity of water sources (Rodell et al., 2009).

Based on the high cost of interventions and the large areas involved, Environmental Impact Assessments (EIA) in irrigation developments should include Health Impact Assessments (HIA) at all different phases of the projects (WHO, 2012). Ongoing monitoring, surveillance, and adaptive mitigation of the negative consequences on water-borne and vector-borne diseases are needed as the projects evolve over time, as well as provision for the cost of these activities over a significant time horizon (NTPC, 2005). During the construction phase of the Sardar Sarovar Project (1979-1985), malaria incidence increased significantly (Srivastava et al., 2009). Later, during the developmental activities in the command area, considerations regarding water- and vector-borne diseases were incorporated and mitigation measures implemented. Even though HIA is generally included in EIA for large-scale developmental projects, the sustained implementation of recommended measures to decrease these impacts is often incomplete (Caussy et al., 2003; Garg, 2009). The situation is generally more critical for small- to medium-scale developmental projects in South-East Asia, such as the construction of smaller irrigation canals and ponds for water collection and storage. In these settings, HIA, and specifically monitoring and surveillance, are more limited, despite their potential to affect extensive areas and large populations.

The observations on the long transition phase in Gujarat reinforce the insight that development of water resources requires a strong binding commitment to finance and implement projects which maintain public health and safety (Birley, 1991). Environmental management methods for sustainable disease control are strongly needed. Several of these methods have proved to be cheaper, more effective, and feasible to implement at local scales, including those that manipulate water flow in irrigation systems, such as intermittent irrigation or canal flushing (Keiser et al., 2005). The challenge ahead, then, will be to apply these methods over extensive regions and maintain them for long enough periods. Already in the early 1980s the interdisciplinary and multi-sectorial nature of health problems related to irrigation were recognized. A joint WHO/FAO/UNEP Panel of Experts on Environmental Management was established which produced guidelines and raised main issues at stake with policy-setting institutions worldwide (Slooff, 1990). The likely decline of global food security with predicted climate, and the rise of food prices in recent years, should be an incentive to reinvigorate initiatives to develop methods that mobilize global water resources without compromising human and environmental health.

Methods

Study area and epidemiological data

The western district of Kutch contains 9 sub-districts or talukas. In the eastern districts of Banas Kantha and Patan, 16 talukas complete our data set. Taluka Bhabhar was part of taluka Deodhar until the year 2000. For this analysis cases for both talukas were merged and included as a single unit.

The epidemiological data consists of time series of microscopically confirmed cases of *Plasmodium falciparum* and *Plasmodium vivax* from rural areas for the 25 talukas in these three districts. These

monthly time series span a period of 12-15 years, from 1997 to 2011 for Kutch and Banas Kantha, and between 2000 and 2011 for Patan. The malaria data based on active surveillance (health workers' home visits twice a month) and passive surveillance (self-reporting at public health facilities), were obtained from the office of the Joint Director (Malaria), Commissionerate of Health & Family Welfare, Gandhinagar (Gujarat). All the analyses were conducted using incidence rates (cases/1000 population). Rural population for each taluka data were obtained by linear interpolation from three national district decadal population census reports for 1991, 2001 and 2011, by removing large urban agglomeration and cities.(for the robustness of the results to other means of estimating population numbers see Figs. S4.6 and S4.7). Socioeconomic data were obtained from the District Census Handbook of concerned district for the year 2001 from Directorate of Census Operations, Gujarat (Table 4.1).

Indoor residual spraying data and control in the region

Indoor residual spray (IRS) with an effective insecticide is applied to control malaria in high-risk villages and to contain malaria outbreaks. The national policy requires IRS application in high-risk villages only, for which villages are stratified by risk at the beginning of each calendar year (high risk reported in the previous year: villages reporting ≥2 malaria cases/1000 population; those reporting confirmed malaria death or cumulative increase in *falciparum* in the past three years). Usually beginning in late May/early June, two rounds of IRS three months apart with a pyrethroid insecticide or three rounds 45-days apart with malathion are applied. For this study, we obtained IRS data of each taluka for 2001 through 2010 from the office of the Joint Director (Malaria) Gandhinagar.

In outbreak years IRS intervention efforts are complemented, through intensive active case detection and prompt treatment and better coverage by insecticides. Until recently chloroquine has

been the first line of treatment for *P. falciparum and P. vivax* cases. Recent changes in control efforts for the district of Kutch include the introduction of deltamethrin impregnated nets at a small scale since 2004, of long lasting nets after our study period, and of Artemisinin based combination therapy (ACT) for the treatment of *P. falciparum* cases in the year 2010.

Remote sensing data and NDVI-based irrigation product

Irrigation map for the year 2009

A procedure to discriminate irrigated from non-irrigated crops using remote sensing information is to compare the empirical spectral signature of crops, obtained by ground control measurements, against the spectral signature of vegetation in the area of interest gathered by satellite sensors (Dheeravath et al., 2010). The irrigation map presented in Figure 4.1 panel C was developed in this way by the Bhaskaracharya Institute for Space Application and Geo-Informatics (BISAG) Gandhinagar, Gujarat, and constructed using the Advanced Wide Field Sensor (AWiFS), at a resolution of 56 meters and with ground control points from the BISAG data library.

Reconstructing irrigation maps and estimating change in irrigated areas based on remote sensing

Agricultural seasons can be divided into Kharif and Rabi, with crops in the former usually planted in June/July and consisting of a mixture of irrigation and rain-fed crops, and those in the latter (mainly wheat) planted in October/November and harvested in March/April, relying mostly on irrigation.

To reconstruct areas of irrigated agriculture outside the year of 2009, we used the 250 m resolution Normalized Difference Vegetation Index (NDVI) images from the Moderate-resolution Imaging Spectroradiometer instrument (MODIS) aboard the Terra satellite. The data were provided by the Land Processes Distributed Active Archive Center located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (https://lpdaac.usgs.gov). Our method is based on the observation of the seasonal temporal pattern of vegetation in the region: in highly-

irrigated areas, NDVI time series show a bimodal seasonal pattern with one peak in September (end of Kharif crop) and another in January (middle of Rabi crop) reflecting irrigation (SI, Fig. S4.3). Thus, the NDVI image for the month of January 2009 shows a spatial pattern similar to that of the irrigated map for that same year (Fig, S4.8). Based on the comparison of NDVI values in January 2009 with the irrigation map, we developed a classification that differentiates irrigated and non-irrigated pixels based on a cut-off value in NDVI. Thus, by collecting images from the MODIS instrument, we generated a set of Boolean images between 2001 and 2013 (Fig. 4,2 Panel A for 2001, and Fig. S4.8 for other years, including 2009) that were then used to quantify the change in the proportion of irrigated land in each sub-district (Fig. 4.2 Panel B for 2001-2009 and Fig. S4.9 for other years). The best cut-off value was obtained when the classification performance was the most accurate in predicting true occurrence of irrigated and non-irrigated locations simultaneously.

Specifically, let Pi and Pn be univariate density functions of NDVI-MODIS images for irrigated and non-irrigated classes, respectively, and Fi and Fn, their corresponding cumulative distribution functions at different cut-off values of $\mathrm{NDVI} \in [0,1]$. For our purposes, Fi is the empirical cumulative distribution function of the predicted probability for pixels of MODIS-NDVI images when irrigation actually occurred (BISAG map), and Fn is the empirical cumulative distribution function for areas where irrigation did not occur. We determined the value of NDVI that maximizes the Kolmogorov-Smirnov distance $D(\tau)=max|Fi - Fn|$, where τ is the value of NDVI with the greatest distance D between the two curves. A threshold of $\mathrm{NDVI} = 0.34$ (F-statistic= 0.8) was found as the best value for this binary classification.

Group inference via a Markov transition model

In order to identify groups of locations with similar dynamics, we employed a nonparametric Markov transition model (Reiner et al., 2012) in a Bayesian framework (Baskerville et al., 2013).

Under this model, the data is discretized into a set of finite levels by putting all zeros in the lowest level and then dividing the remaining data into observed quantiles of cases per capita. Transitions between levels over time are described by a Markov transition matrix. Locations are assigned to different groups, and within a group each location's time series is assumed to follow the same transition matrix. Good groupings thus have locations with similar dynamics assigned to the same groups, and the goal is to identify the best grouping, defined as the one with the highest marginal likelihood. Transition-matrix rows were assigned non-informative Jeffreys priors (Jeffreys, 1946), and the marginal likelihood was evaluated analytically for all $224 = 16{,}777{,}216$ different arrangements into one or two groups. We checked results for robustness to changes in transition-matrix priors, to the number of disease levels, and to changes in the number of groups. An equivalent maximum-likelihood analysis also yielded identical results.

Figures

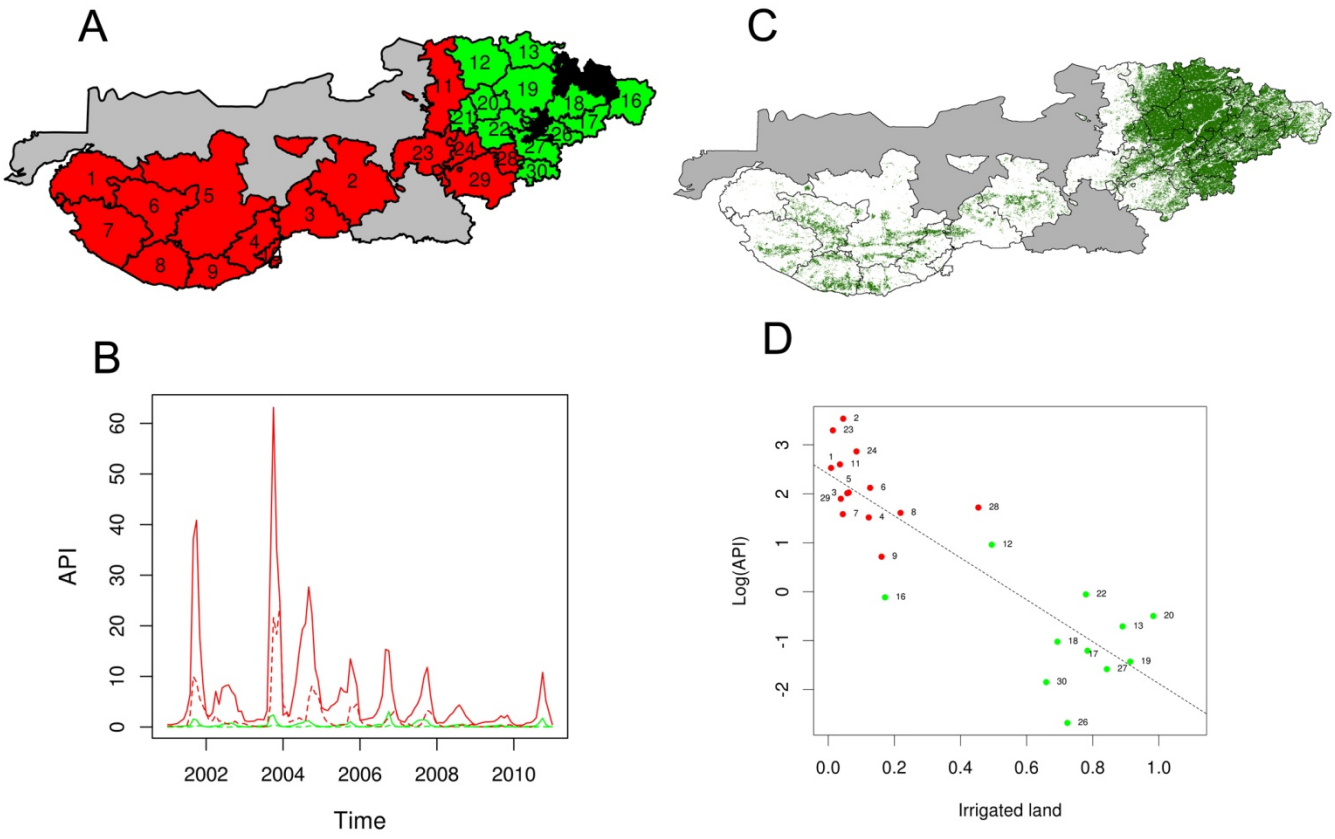

Figure 4.1 Spatial-temporal pattern in malaria population dynamics and its relationship to irrigation development.

Panel A shows the two groups in the configuration of malaria risk obtained by the Bayesian grouping algorithm. Areas of high risk are colored in red and those of low risk areas, in green. (The overall results on groupings are robust to the choice of parasite species, as well as to the number of levels for the different quintile divisions of the epidemiological data). In B, the time series of accumulated cases are shown for the two groups (red and green) and for the two malaria species (solid line: *P. vivax* and dashed line: *P. falciparum*). Most of the malaria burden in the region corresponds to *P. vivax*. For comparison with A, panel C shows the irrigation pattern for the year 2009, for which a detailed irrigation map was available. This comparison shows that the high and low malaria groups map respectively onto the non- (or low-) irrigated (in white) and irrigated areas (in green) respectively. Although the map is for 2009, a similar broad pattern of irrigation holds across years (Fig. S4.8), and the eastern region has been irrigated for at least three decades (Fig. S4.2). The detailed pattern of how irrigation has changed in the more recent decade is addressed in Figure 4.2. An example of the malaria-irrigation relationship is presented in panel D, for the large epidemic of the year 2003. In this figure, each point corresponds to incidence (1000*cases/pop in

log scale) in a particular taluka, during the epidemic season (Sep-Dec), as a function of the proportion of the land classified as irrigated, based on January's vegetation from remote sensing (Methods). The numbers in A correspond to the name of a taluka as follows: 1 Lakhpat, 2 Rapar, 3, Bhachau 4, Anjar & Gandihan, 5 Bhuj, 6 Nakhatrana, 7 Abdasa, 8 Mandavi, 9 Mundra, 11 Cac, 12 Tharad, 13 Dhanera, 16 Danta, 17 Vadgam, 18 Palanpur, 19 deesa, 20 Deodar, 21 Bhabhar, 22 Kankrej, 23 Santalpur, 24 Radhanpur, 26 Sidhpur, 27 Patan, 28 Harij, 29 Sami, 30 Chanasma.

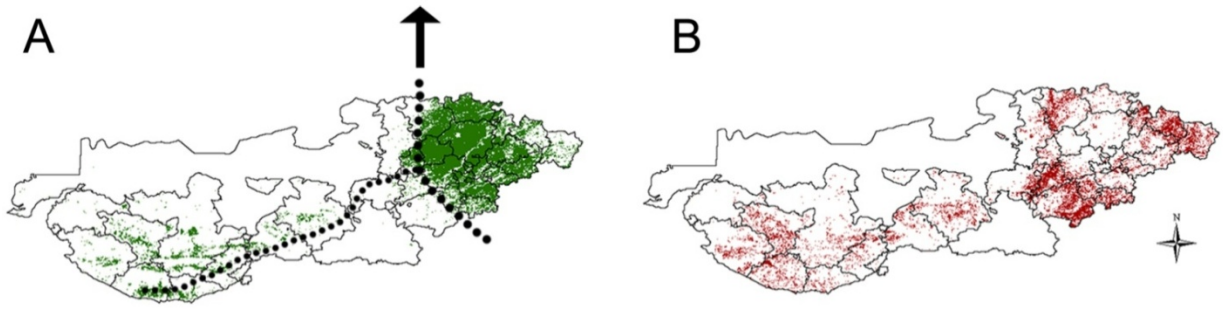

Figure 4.2 Irrigation pattern and its change over time.

Panel A shows the area classified as irrigated agriculture (in green) for 2001, based on the NDVI classification of irrigated pixels outside the monsoon season (at a 250 m2 resolution), as described in the Methods. The black dotted lines in panel A represent the position of the trunk of the canal and its main branch to the west. Panel B shows the areas that have experienced the most pronounced variation in irrigation levels between 2001 and 2009. Specifically, areas in red correspond to those classified as irrigated in 2009 but non-irrigated in 2001 (see also figures S4.8 and S4.9).

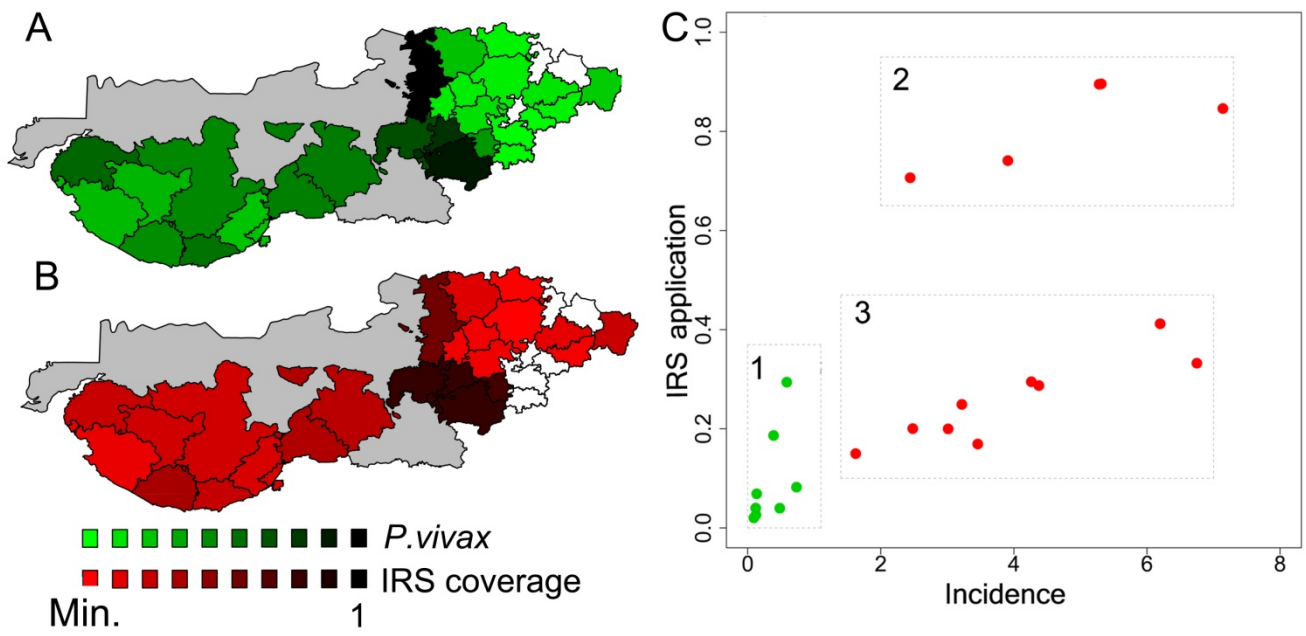

Figure 4.3 Spatial distribution of relative malaria risk and control application.

The spatial distribution of total incidence for *Plasmodium vivax* relative to its maximum value is shown in panel A for the period of 2006 to 2010. The scale of green tones corresponds to values of malaria incidence relative to this maximum and spans values of 'risk' from the lowest (=0.006, in light green) to the highest (up to 1, in dark green). Panels B shows a similar spatial distribution of the IRS effort based on the mean proportion of the population covered, again relative to the maximum value observed between the period of 2006-2010. Colors from light to dark red represent population covered with IRS according to the national policy of insecticide house-spraying, from a low (=0.0004) to a high percentage, respectively. Talukas with no information are shown in white; gray areas represent the un-inhabited Rann of Kutch. (Similar spatial patterns are seen for the period of 2000-2005 in Figs. S4.6 and S4.7). Panel C shows the three malaria epidemiological zones defined based on incidence, control and irrigation**.** The mean proportion of population covered by IRS is shown as a function of median incidence from 2000 to 2010. Green dots correspond to the talukas that have been irrigated for a long time and present low malaria risk and low IRS coverage (zone 1). Differences in the level of IRS coverage for areas of high risk (red dots) can be observed between the talukas in Kutch with relative low control but high risk (zone 3) and those in Banas Kantha and Patan, the transition zone, that exhibit both high IRS application and high incidence (zone 2), with coverage reaching values of around 80%. Despite the downwards trend in cases, the existence of these three zones persist throughout the decade (Fig. S4.10).

Variable	High risk	Low risk	t_value	df	p.value	W	W p.value
Literates/Illiterates	0.7577	0.8935	-0.8827	16.1796	0.3903	71	0.7675
Improved drinking water	0.9227	0.9903	-2.2142	13.9951	0.0439	36.5	0.0282
Agricultural credit societies	0.6106	0.8183	-2.7805	18.2579	0.0122	18	7.00E-04
Banks access	0.2565	0.2464	0.1966	19.1836	0.8462	83	0.7675
Education access	0.9946	0.9975	-1.0629	22.9806	0.2989	51	0.1492
Medical access	0.7705	0.6615	1.5383	22.9944	0.1376	102	0.1832
% Irrigated land	0.1615	0.7172	-9.1045	21.25	<0.0001	2	<0.0001
IRS control	0.4558	0.069	4.6783	16.176	<0.0001	146	2.00E-04

Table 4.1Statistical analysis of differences between mature irrigated and nonirrigated areas

These variables consist of the ratio between literate and illiterate population, the proportion of the population in each taluka with access to potable water, agricultural credit societies, banks, education and public health and medical facilities. Education facilities encompass all primary and secondary schools. Medical access corresponded to the Community Health Centers, Primary Health Centers, sub-centers, and hospitals available in each taluka. We tested if these socioeconomic indicators are significantly different between irrigated and non-irrigated talukas. A two-sample location unpaired Welch's t-test, with the level of malaria risk obtained from the cluster algorithm as a categorical variable, was used to test if the socioeconomic indicators from the talukas were sampler from two different normal distributions based on the level of the malaria risk. We also performed a non-parametric Wilcoxon test that does not assume normality. We applied these tests for IRS control and for the percentage of the talukas under irrigation to support our findings.

Appendices

Appendix 4.1: Supporting Figures

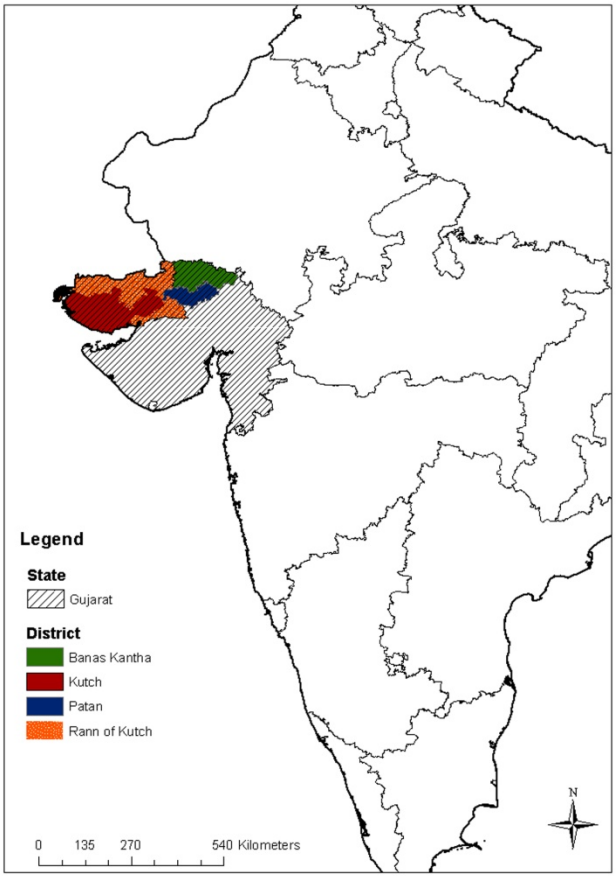

Figure S4.1 Study area

The study is conducted in the state of Gujarat (dashed area) in the districts of Kutch (red), Banas Kantha (green) and Patan (blue). The un-inhabited Rann of Kutch is also shown (in orange).

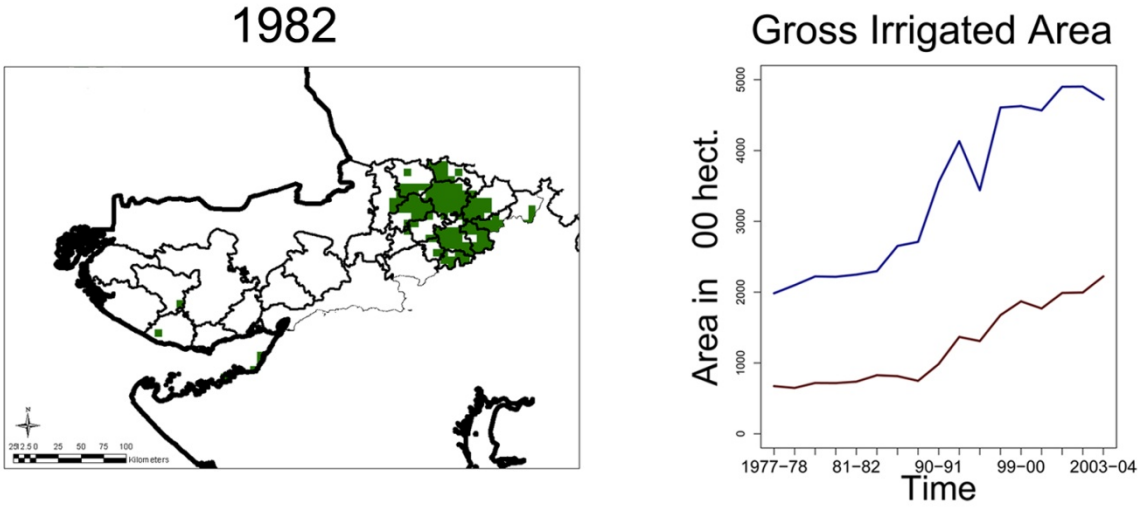

Figure S4.2 Historical trends in irrigation in the districts of Kutch and Banas Kantha

To complement our spatial-temporal data, we obtained from the literature historical data on irrigation growth in the region. The map on the left panel shows an NDVI image for January 1982 at an 8 km resolution, obtained from the AVHRR instrument (Tucker et al., 2005). NDVI values > 0.38 are plotted in green, to highlight pixels classified as irrigated (Methods). (Because this threshold value was obtained for MODIS NDVI, we transformed it from MODIS to AVHRR measurements, by fitting a non-linear regression of the form $y \sim a * x^b$, with $x = NDVI\text{-}MODIS$ and $y = NDVI\text{-}AVHRR$, and both variables corresponding to the taluka average for the month of January during the years in which MODIS and AVHRR overlapped, 2001-2006). The right panel shows gross irrigated area at the district level for Kutch (red line) and Banas Kantha (blue line) between 1977 and 2004. This additional information on irrigation in the region indicates that the eastern group of talukas, whose spatial extent maps onto the detailed irrigation map of 2009 (Figure 4.1 and Methods), have been irrigated at least since the beginning of the 1980s.

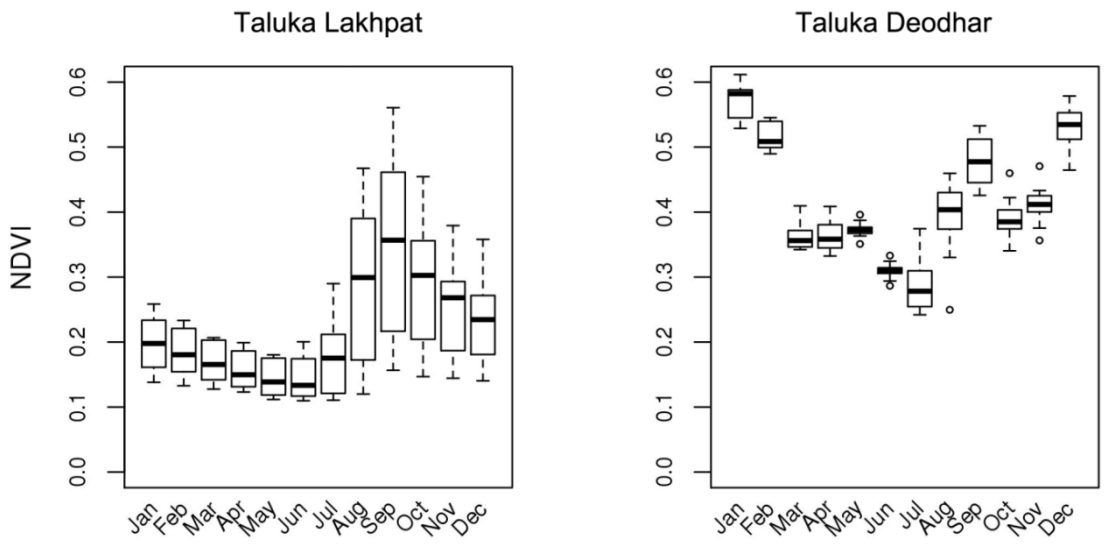

Figure S4.3 NDVI seasonality

The boxplots show the seasonality of average NDVI for the two talukas at the extremes of the irrigation gradient. At the dry end, Lakhpat (taluka #1, Fig. 4.1 panel A) exhibits only one peak in the vegetation index for the month of September, following the monsoon season. At the other extreme, the highly irrigated taluka of Deodhar (taluka #20, Fig. 4.1 panel A) displays a bimodal seasonality with a much larger peak in January and much less interannual monthly variation. This January peak is clearly outside the monsoon season and most likely reflects the effect of irrigation on vegetation. Thus, we consider the values of the NDVI vegetation index in that month to formulate a classification of irrigated and non-irrigated land (Methods).

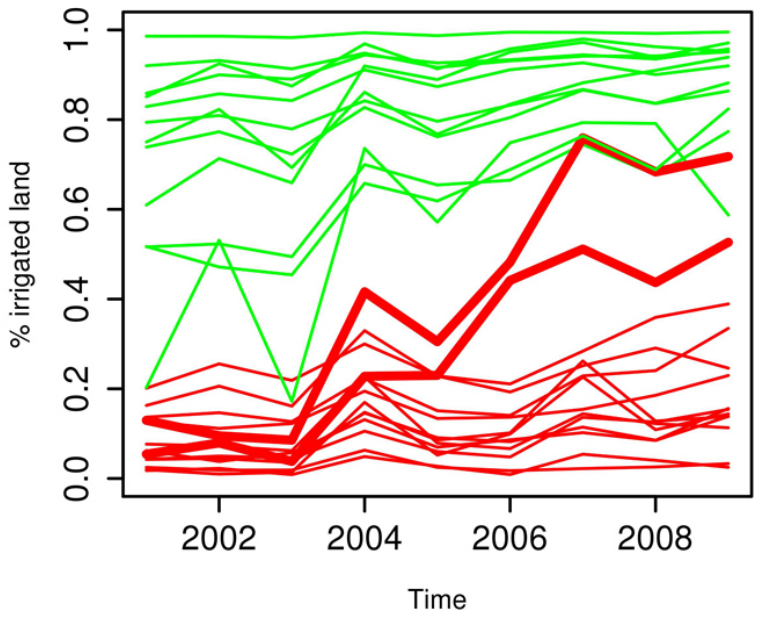

Figure S4.4 Temporal changes in estimated irrigation

Each line represents the annual proportion of the taluka area classified as irrigated between 2001 and 2009, based on the classification procedure using NDVI images (Methods).The green and red lines correspond to talukas in the low and high risk zone according to the results obtained using the grouping algorithm (in Fig. 4.1 Panel A). The two thicker red lines highlight the talukas where most of the change occurred during the decade; these talukas are located in the southernmost part of the study area (Figure 4.2 Panel B) and at the boundary of the two regions identified in Figure 4.1.

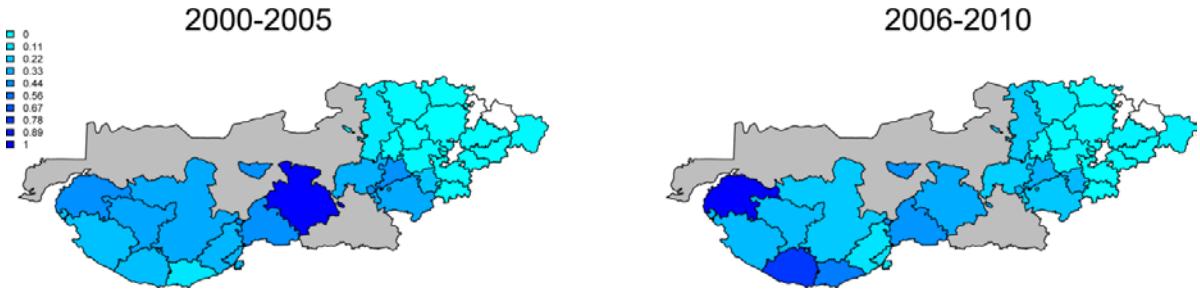

Figure S4.5 Relative risk of Plasmodium falciparum malaria between the years 2000-2010

Colors from light to dark blue correspond to the level of P. falciparum incidence relative to their maximal value. Because the number of cases for this parasite are low (Fig. 4.1 Panel B), especially in the second half of the decade, the pattern can be quite noisy.

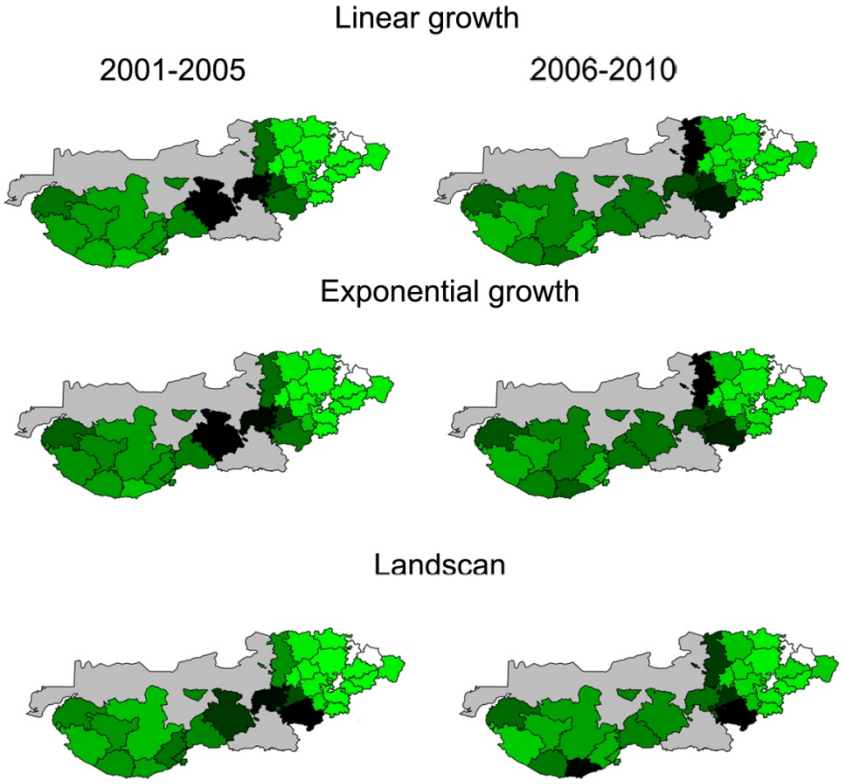

Figure S4.6 Comparison of results on the spatial distribution of malaria risk for three different ways to estimate population values in between the two censuses of 2001 and 2011

As a way to corroborate the robustness of our findings to the linear population estimation approach, we also extrapolated the population changes based on exponential growth. To take into consideration possible yearly variation in population we also used the Landscan grid population product from 2001-2011 (Bright, et al., 2002 & Bright, et al., 2012). Because of changes in the methodology from one year to the next, it has been noted that this product should not be used in a time series fashion. Thus, we relied on Landscan data only for results concerning spatial patterns of prevalence in this figure, and IRS coverage in Fig. S4.7). See the caption of Figure 4.3 for further details. To facilitate comparison, the maps of that figure (obtained with the linear interpolation) are repeated here and included in the top row. The two other methods give similar spatial distributions; in particular, the high risk area remains, at the boundary of the "mature-irrigated" and "low-irrigated" groups of talukas (Figure 4.1). This is a robust feature with the exception of one additional 'dark' taluka on the coast of Kutch for the Landscan's population estimates (Methods).

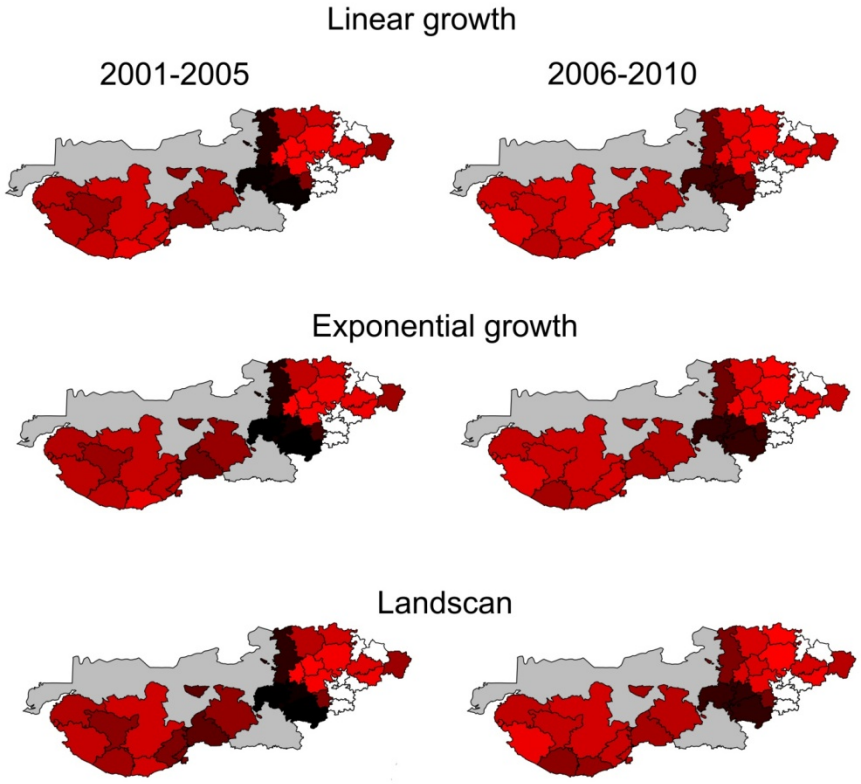

Figure S4.7 Comparison of results on the spatial distribution of IRS coverage (relative to its maximum value), for three different ways to estimate population values in between the two censuses of 2001 and 2011

Linear interpolation, exponential interpolation and the gridded product known as Landscan. See the caption of Figures S4.6 and 4.3 for more details. Maps of the top row are those of that figure. The high coverage area at the boundary of the eastern and western groups of talukas (Figure 4.1) is a robust feature.

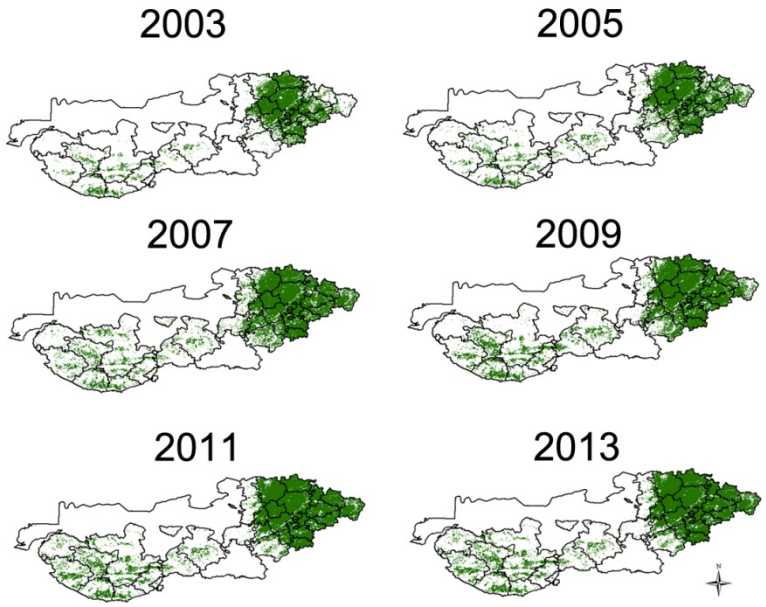

Figure S4.8 Reconstruction of irrigation between 2003 and 2013 based on MODIS NDVI images

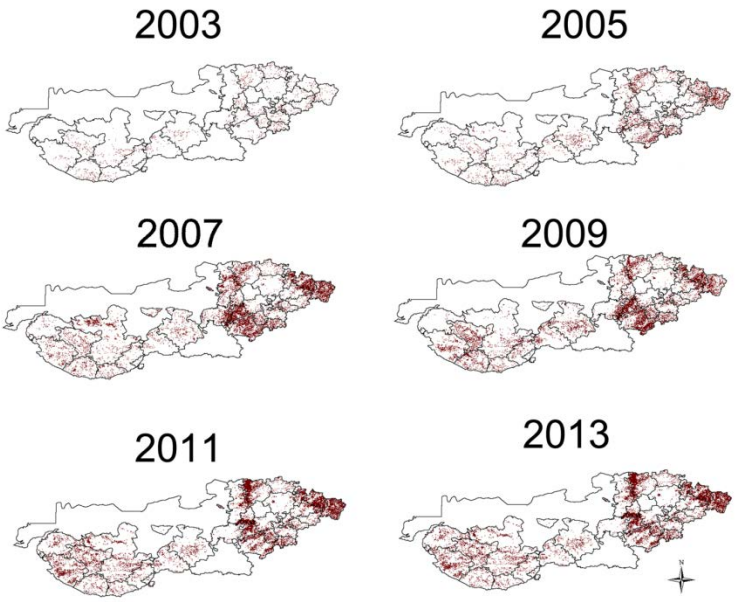

Figure S4.9 Spatial change in irrigation between 2001 and subsequent years (2003, 2005, 2007, 2009, 2011, and 2013)

A pronounced change is observed for the talukas at the boundary between the low (or no) irrigation region in the west and the mature irrigated region in the east.

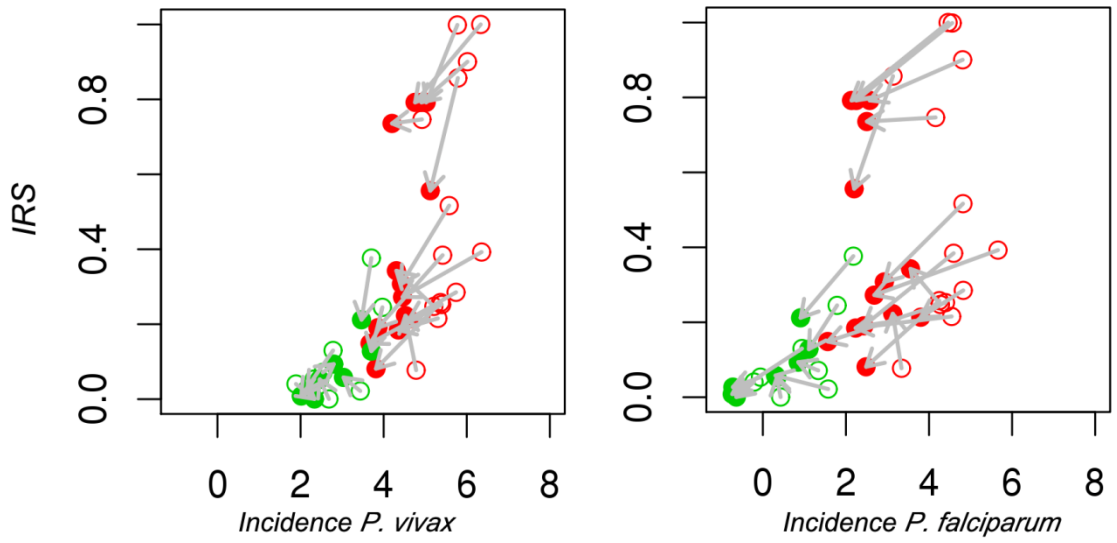

Figure S4.10 Trend trajectory of the malaria system and persistence through time of the three identified zones for IRS population coverage and malaria risk, described in the main text

Each dot in the plot shows the value of control Intervention in terms of percent population covered by IRS as a function of the total incidence per ten thousand people, in logarithmic scale. Empty dots represent the period between 2000 and 2005 and filled dots that between 2006 and 2011. Colors are similar to those in figure 4.3 Panel C (and 1 A). Despite a downward trend in incidence, the three main regimes depicted in Figure 4.3 Panel C remain present.

References

Amerasinghe, P.H., Amerasinghe, F.P., Wirtz, R.A., Indrajith, N.G., Somapala, W., Pereira, L.R.., Rathnayake, A.M.S., 1992. Malaria transmission by Anopheles subpictus (Diptera: Culicidae) in a new irrigation project in Sri Lanka. J Med Entomol 29(4),577-581.

Amerasinghe, FP, Indrajith, NG., 1994. Post irrigation breeding patterns of surface-water mosquitos in the Mahaweli Project, Sri Lanka, and comparisons with preceding developmental phases. J Med Entomol 31(4),516-523.

Baeza, A., Bouma, M.J., Dobson, A.P., Dhiman, R., Srivastava, H.C., Pascual, M., 2011. Climate forcing and desert malaria: the effect of irrigation. Malaria J 10,190.

Baskerville, E.B., Bedford, T, Reiner, R.C., Pascual M., 2013. Nonparametric Bayesian grouping methods for spatial time-series data. arXiv, 1306-5202.

Baeza, A, Bouma, M.J, Dhiman, R, Pascual, M., 2013. Malaria control under unstable dynamics: Reactive vs. climate-based strategies. Acta Trop.(*In press*)

Bhatt, RM, Shrivastava HC, Rajnikant HM., Yadav, RS., 2008. Dynamics of *Anopheles culicifacies* transmitted malaria in the absence of effective zoopro-phy-laxis in a riverine settlement in Gujarat, India. Curr Sci 95,82-87

Birley, M.H., 1991. Methods of Forecasting the Vector-Borne Disease Implications in Development of a Water-Resources Project. Techniques for Environmentally Sound Water Resources Development. 50-63.

Bonds, M.H, Keenan, DC., Rohani, P., Sachs, J.D., 2010. Poverty trap formed by the ecology of infectious diseases. P Roy Soc B-Biol Sci. 277(1685),1185-1192.

Bouma, M.J., van der Kaay H.J., 1996. The El Niño Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics? Trop Med Inter Health 1(1),86-96.

Caussy, D., Kumar, P., Sein, U.T., 2003. Health impact assessment needs in south-east Asian countries. B World Health Organ. 81(6),439-443.

Claborn, D.M, Hshieh, P., Roberts, D., Klein, T., Zeichner, B., Andre, R., 2002. Environmental factors associated with larval habitats of malaria vectors in northern Kyunggi Province, Republic of Korea. J Am Mosquito Contr Assoc. 18(3),178-185.

Cohen, J.M., Smith, D.L., Cotter, C., Ward, A., Yamey, G., Sabot, O.J., Moonen, B., 2012. Malaria resurgence: a systematic review and assessment of its causes. Malaria J. 11, 122.

Darling, ML., 1925. The Punjab peasant in prosperity and debt (Oxford University Press, London).

Dattani, M., Prajapati, P., Raval, D., 2009. Impact of indoor residual spray with synthetic pyrethroid in gandhinagar district, gujarat. Indian J Community Med. 34(4),288-292.

de Zulueta, J., Mujtaba, S.M., Shah, I.H., 1980. Malaria control and long-term periodicity of the disease in Pakistan. T Roy Soc Trop Med H. 74(5),624-632.

Dheeravath, V., Thenkabail, P.S., Chandrakantha, G., Noojipady, P., Reddy, G.P.O., Biradar, C.M., Gumma, M.K., Velpuri, M., 2010. Irrigated areas of India derived using MODIS 500 m time series for the years 2001-2003. Isprs J Photogramm. 65(1),42-59.

Dia, I., Samb, B., Konate, L., Fontenille, D., 2010. Population structure of newly established Anopheles funestus populations in the Senegal River basin using paracentric chromosomal inversions. Acta Trop. 115(1-2),90-94.

Doannio, J.M., Dossou-Yovo, J., Diarrassouba, S., Rakotondraibé, M.E., Chauvancy, G., Chandre, F., Rivière, F., Carnevale, P., 2002. Malaria transmission in the rice growing area of Kafine village, Cote d'Ivoire. B Soc Pathol Exot. 95(1),11-16.

FAO, 2003. Rethinking the approach to Ground Water and Food Security in Water Reports. Rome, Food and Agriculture Organization of the United Nations. available at: http://www.fao.org/docrep/005/y4495e/y4495e00.htm (accessed 15 March 2013).

Garg, A., Dhiman, R.C., Bhattacharya, S., Shukla, P.R., 2009. Development, malaria and adaptation to climate change: a case study from India. Environ Manage. 43(5),779-789.

Herrel N, et al. (2001) Breeding of Anopheles mosquitoes in irrigated areas of South Punjab, Pakistan. Med Vet Entomol 15(3):236-248.

Herrel, N., Amerasinghe, F.P., Ensink, J., Mukhtar, M., van der Hoek, W., Konradsen, F., 2004. Adult anopheline ecology and malaria transmission in irrigated areas of South Punjab, Pakistan. Med Vet Entomol. 18(2),141-152.

Jayaraman, T.K., 1982. Malarial impact of surface irrigation projects: a case study from Gujarat, India. Agriculture and Environment. 7(1),23-34.

Jeffreys, H., 1946. An Invariant Form for the Prior Probability in Estimation Problems.Proc R Soc Lon Ser-A. 186(1007),453-461.

Keiser, J., Singer, B.H., Utzinger, J., 2005. Reducing the burden of malaria in different eco-epidemiological settings with environmental management: a systematic review. Lancet Infect Dis. 5(11),695-708.

Keiser, J., De Castro, M.C., Maltese, M.F., Bos, R., Tanner, M., Singer, B.H., Utzinger, J. 2005. Effect of irrigation and large dams on the burden of malaria on a global and regional scale. Am J Trop Med Hyg. 72(4),392-406.

Kibret, S., Alemu, Y., Boelee, E., Tekie, H., Alemu, D., Petros, B., 2010. The impact of a small-scale irrigation scheme on malaria transmission in Ziway area, Central Ethiopia. Trop Med Int Health. 15(1),41-50.

Klinkenberg, E., Konradsen, F., Herrel, N., Mukhtar, M., van der Hoek, W., Amerasinghe, F.P., 2004. Malaria vectors in the changing environment of the southern Punjab, Pakistan. T Roy Soc Trop Med H. 98(7),442-449.

Klinkenberg, E., van der Hoek, W., Amerasinghe, F.P., 2004. A malaria risk analysis in an irrigated area in Sri Lanka. Acta Trop. 89(2),215-225.

Konradsen, F., Stobberup, K.A., Sharma, S.K., Gulati, O.T., van der Hoek, W., 1998. Irrigation water releases and Anopheles culicifacies abundance in Gujarat, India. Acta Trop. 71(2),195-197.

Laneri, K., Bhadra, A., Ionides, E.L., Bouma, M., Dhiman, R.C., Yadav, R.S., Pascual, M., 2010. Forcing versus Feedback: Epidemic Malaria and Monsoon Rains in Northwest India. PLos Comput Biol. 6,9.

Molineaux, L., 1998. The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. Malaria: principles and practice of malariology, Wernsdorfer & McGregor eds. (Churchill Livingstone, Edinburgh).

NTPC., 2005. Project Implementation Plan. Chapter 2: Health Programs. (Nam Theun 2 Power Company) available at: http://www.namtheun2.com/images/stories/pip/PIP%20Final%20-%20Part%20B%20Chapter%202_Health_050527.pdf (accessed July 2013).

Reiner, R., King, A., Emch, M., Yunus, M., Faruque, A., Pascual, M., 2012. Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. P Natl Acad Sci USA 109(6),2033-2036.

Rodell, M., Velicogna, I., Famiglietti J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature 460(7258),999-1002.

Russell, P.F., Ramachandra Rao, T. 1942. A study of the density of Anopheles culicifacies in relation to malaria endemicity. Am J Trop Med Hyg. 22:535-558.

Sachs, J., Malaney, P., 2002. The economic and social burden of malaria. Nature. 415(6872),680-685.

Shah, T., Singh, O.P., 2004. Irrigation Development and Rural Poverty in Gujarat, India: A Disaggregated Analysis. Water International. 29(2),167-177.

Sharma, K.D., 2001. Indira Gandhi Nahar Pariyojana - lessons learnt from past management practices in the Indian arid zone. Iahs-Aish P. (268),49-55.

Slooff, R., 1990. Towards healthier water resources management. Waterlines. 9(2),2-2.

Srivastava, H.C., Bhatt, R.M., Kant, R., Yadav, R.S., 2009. Malaria associated with the construction of the Sardar Sarovar Project for water-resources development, in Gujarat, India. Annals Trop Med Parasitol. 103(7),653-657.

Swaroop, S., 1949. Forecasting of epidemic malaria in the Punjab, India. The American Journal of Tropical Medicine and Hygiene. 29(1),1-17.

Tyagi, B.K., Yadav, S.P., 2001. Bionomics of malaria vectors in two physiographically different areas of the epidemic-prone Thar Desert, north-western Rajasthan (India). J Arid Environ. 47(2),161-172.

Tyagi, B.K., 2004. A review of the emergence of Plasmodium falciparum-dominated malaria in irrigated areas of the Thar Desert, India. Acta Trop. 89(2),227-239.

Yadav, R.S., Sharma, R.C., Bhatt, R.M., Sharma, V.P., 1989. Studies on the anopheline fauna of Kheda district and species-specific breeding habitats. Indian J Malariol. 26(2),65–74.

Yewhalaw, D., Legesse, W., Van Bortel, W., Gebre-Selassie, S., Kloos, H., Duchateau, L., Speybroeck, N., 2009. Malaria and water resource development: the case of Gilgel-Gibe hydroelectric dam in Ethiopia. Malaria J. 8.

Yohannes, M., Haile, M., Ghebreyesus, T.A., Witten, K.H., Getachew, A., Byass, P., Lindsay, S.W., 2005. Can source reduction of mosquito larval habitat reduce malaria transmission in Tigray, Ethiopia? Trop Med Int Health 10(12),1274-1285.

White, N.J., 2011. Determinants of relapse periodicity in Plasmodium vivax malaria. Malaria J 10,297.

WHO, 2012. health 2020: a European policy framework supporting action across government and society for health and well-being. Geneva, World health Organization, available at: 222.euro.who.int/_data/assets/pdf_file/0009/169803/RC62wd09_Eng.pdf

Chapter 5

A time series model to test the double causality between malaria and income in Mississippi, United States

Abstract

In agricultural communities, malaria and socioeconomic conditions are intimately connected. Malaria incidence can affect income level by decreasing labor productivity, and low income level can in turn affect prevention and control affordability, resulting in the exacerbation of disease burden, which reinforces the negative feedback loop between poverty and disease. Theoretical studies of the population dynamics of infectious diseases have proposed a 'poverty trap' resulting from this double causality or feedback loop, when considering an aggregated level of infection due to multiple pathogens. Although multiple empirical studies have addressed these linkages for one direction of influence or the other, no study has considered the explicit dynamics of this epidemiological-economic system, especially when both effects are acting in concert for a given pathogen of significant health impact. An understanding of the relative importance of these two effects in concert and in a dynamic framework is relevant to the population dynamics of malaria, to its control, public health policy and eventual elimination in developing countries.

By relying on epidemiological and economic data from the eradication era (1914-1927) in the highly malarious state of Mississippi (1914-1927), we statistically test a dynamic poverty trap hypothesis

between malaria incidence, labor productivity, and income with a time series model. We compare this hypothesis against others that specifically disconnect part of the feedback loop, thereby changing the causality pathways, via a likelihood approach. Results show that while the full model with explicit feedback loops among malaria, cotton productivity, and income better explain the data than partial parts of the system, malaria burden had only a weak effect on cotton productivity, and causality appears dominated by the direction from income to malaria in a non-linear manner. Furthermore, sensitivity analyses of parameter values suggest that the elimination of the disease was most likely driven by macro-level determinants related to the overall improvement of socioeconomic conditions, bottom-up control methods, and the affordability of the latter. We discuss our findings for the historical United States and its implications for the current malaria situation in other parts of the world under similar endemic conditions but on the path to elimination.

Introduction

Malaria is one of the most deadly diseases of all times, currently responsible for more than 200 million clinical episodes every year. Most of these cases are concentrated in regions of the world with the lowest average GDP: Ninety percent of all malaria cases occur in sub-Saharan Africa (Teklehaimanot and Mejia, 2008), in India (Sharma, 1999), and in different countries in South America. Malaria-ridden countries exhibit on average five times less income than countries without the disease, and tend to grow at a rate that is 1.3% slower (Gallup and Sachs, 2001).

While this association between malaria and poverty has long been recognized (Sachs and Malaney, 2002; Laxminarayan, 2004; Malaney et al., 2004; Chuma et al., 2006; Teklehaimanot and Mejia, 2008), the direction of this causality and the relative importance of the potential feedbacks in the two opposite directions remain the subject of debate (Packard, 2009 & 2007).

When the parasite in its merozoite stage colonizes blood cells and consumes hemoglobin, fever episodes and anemia are experienced by symptomatic hosts. These episodes not surprisingly reduce the time and energy that an individual spends in work-related activities. Under recurrent malaria episodes, therefore, long-term human productivity is expected to decrease, thereby slowing the economic return of a household (Guiguemde et al., 1997; Chuma et al., 2006) and eventually, the prosperity of a village or entire country. In the opposite direction, several studies have shown that socioeconomic factors such as income (Filmer, 2005), state-level resources for disease control (White, 2011), education (Yasuoka et al., 2006), household construction (Hulden & Hulden, 2009), and agricultural development (Baeza et al., In press) can influence malaria epidemiological patterns. If malaria burden is a dominant factor in shaping the economy of an individual, a community, or a nation, measures that directly decrease the burden of the disease will have indirect benefits for the alleviation of poverty. Alternatively, if decreasing poverty level dominates the risk of contracting malaria, then maintaining the downward trend in poverty level may collaterally decrease the burden of malaria without a top-down intervention. Thus a better understanding of what direction of causality is most important has clear implications for malaria control and elimination policy, and can shed light on the conditions needed to maintain a stable elimination state (Chiyaka et al., 2013).

Despite its relevance, however, limited work has been done to understand the dynamical consequences of this double causality. In particular, most studies on malaria and socioeconomic conditions assume that causation runs in one specific direction. Theoretical studies, however, that do include the full feedback system in epidemiological models have proposed a poverty trap hypothesis of infectious diseases and income to explain the relationship between aggregated disease prevalence and GDP (Bonds and Rohani, 2010). Although a few recent studies have proposed testing this double causality hypothesis using spatial data (Bond et al., 2010, Somi et al, 2007), to our knowledge

no study has established this hypothesis based on longitudinal, temporal observations and their statistical confrontation to dynamical models.

The historical data set of malaria in the cotton-driven economy of Mississippi between 1917-1950, in the period that immediately precedes the eradication of the disease in the United States, provides an opportunity to address these questions; specifically, how important were income and human productivity in driving malaria patterns in the region and, simultaneously, to what extent did malaria incidence influence the labor productivity of the poor sharecroppers. We are particularly interested here in identifying the path of causality and the strength of causation. Using annual State-level records of malaria cases, income, and human productivity, we statistically confronted the feedback hypothesis against alternatives that disconnect specific pathways using likelihood inference. We show that even though the poverty trap hypothesis cannot be rejected, its influence in driving the disease to elimination was relatively low and much of the variability in the system can be explained by large-scale economic drivers, such as cotton prices and mosquito-related control affordability. Moreover, our results suggest that a hypothesis in which the causality runs from socioeconomics to malaria seems to be better supported by the data. In light of these results, we discuss their implications for endemic malaria in current times and in other parts of the world on the path to poverty alleviation and malaria elimination.

Methods

Malaria and socio-economic conditions in historical United States (1900-1940)

Plasmodium malaria arrived in North America with the early colonists and the first Africans brought to the continent as slaves. By the 1875, malaria-related fevers in the United States were observed in the northernmost states, such as Michigan and Illinois, where disease prevalence started a steady

decline. By the beginning of the nineteen hundreds, the distribution of the disease had shrunk to only the southernmost states of Florida, Louisiana, South Carolina, and Mississippi, with its final elimination achieved by 1950 (Ter Veen, 2005).

This study focuses on the State of Mississippi during the first part of the 20th century for several reasons: First, Mississippi was the state with the most reliable malaria data in the region (Ter Veen, 2005). Second, by the time malaria was approaching elimination, the economy of Mississippi was mostly based on cotton production (Schulman, 1994). Third, this cotton production was intense in human labor with very low levels of mechanization and technology. And fourth, most of the farms were large properties owned by a few landowners who rented the land to sharecroppers, most of them originating from poor black communities. Therefore, the connection between the disease and the economy, regardless of direction, should be most evident in this state, and thus, if the poverty trap exists, it should be detectable and significant in this region.

All the analyses presented here were conducted on the data between 1918 and 1938 (Fig. 5.1, dashed lines) in order to avoid the effect of World War II and the technification and diversification of the Mississippi economy right before the sharecropper system collapsed in 1939 (Schulman, 1994).

Epidemiological and socioeconomic data

The epidemiological data consists of annual morbidity cases of malaria in the state of Mississippi reported by the Mississippi State Board of Health between 1916 and 1945 (Fig. 5.1 Panel A). The data show a downwards trend with the final elimination of the disease by 1950. During the years of this study, three large epidemics are apparent. A time series of cotton productivity per person was constructed based on cotton production per hectare, by multiplying this value by total area of cotton harvested and then dividing it by the rural population (Fig. 5.1 Panel B). Cotton production and

cotton prices were reported by the Mississippi Department of Agriculture. Income data were obtained for the same period and correspond to total yearly ($/year) taxable income reported by the annual edition of the Statistical Abstract of the United States, corrected according to the price index of the year to avoid changes due to inflation (Fig. 5.1 Panel C).

To correct for the possible effect of population migration, malaria cases were transformed to incidence (cases per 1000 person) by dividing the time series by the yearly rural population obtained from annual editions of the Statistical Abstracts of the United States. Income and productivity were also converted to per-person using similar rural population figures. Without loss of generality all the statistical inference analyses were conducted on the anomalies of the incidence and cotton-related human productivity computed as $x\_a=[x-mean(x)]/sd(x)$, where x is the variable in the untransformed scale and x_a, the corresponding anomaly.

Statistical model and likelihood-based inference method

In order to statistically test the possibility of a poverty trap created by the interactions between malaria, income, and cotton productivity, a time series model relating the these three variables was constructed. Specifically, the model consists of a system of linear and non-linear regressions that explain the variability observed in malaria incidence (M), income (I), and cotton yield human productivity (Y) simultaneously. These variables, defined as the state variables S, with $S=\{M,I,Y\}$, are estimated based on a set of external and internal predictors (Fig. 5.2). External predictors, or drivers D (Fig. S5.3), are variables that are known to explain variation in S but are not influenced by them. Internal predictors, on the contrary, are the state variables themselves that may or may not be a predictor of another state variable (Fig. 5.2).

Let us define M, Y, and I as random variables, and m_t, y_t, and i_t as random realizations at time t from the distributions $p(M)$, $p(Y)$, and $p(I)$, respectively. $p(M)$ and $p(Y)$ are assumed normal with means μ_Y and μ_M, and $p(I)$, gamma with shape parameters $=\mu_I^2/V$ and $scale=V/\mu_I$.

The mean of these distributions, μ_M, μ_Y, and μ_M follow regression equations of the form

$$\mu^I{}_t = a + bCT_t + c\hat{Y}_{t-\Delta} + \varepsilon_I$$

$$\mu^M{}_t = d + \frac{f(Py_t)}{\hat{I}_{t-\Delta}}(\hat{I}_{t-\Delta})^{\alpha} + gH + \varepsilon_M$$

$$\mu^Y{}_t = i + j\hat{M}_t + kF + \varepsilon_Y,$$

where a, d, and i are the intercepts of the regressions and ε_I, ε_M, and ε_Y correspond to white, observational or process noise, depending on the estimation method. State variables are presented with the accent symbol and Δ to symbolize the temporal delay hypothesis (See Temporal formulation of the model).

External predictors for each regression

In this section we present a detailed explanation of the model and the drivers. In Figure S5.3, the time series of these drivers is shown.

Income equation ($\mu^I{}_t$): Because the per-capita income of Mississippi at the time was mainly based on a cotton economy, the linear regression, μ_I, includes cotton price per pound per year, CT_t, and cotton-related human productivity Y (Fig. 5.2). The model includes these effects in an additive way. We also tested the model with the interaction term $C*Y$, but its inclusion did not improve the model performance (results not shown).

Malaria equation ($\mu^M{}_t$): The malaria (anomaly) estimation at time *t*, m_t, is an additive function of the price of pyrethrum flowers, *Py*, and the number of people per house, *H*. Pyrethrum flowers form the active component of the insecticide used as a repellent for mosquitos in this area. This quantity represents therefore the price of a control intervention. We test the effect of *I* on malaria by dividing *Py* by *I,* and then multiply it by the term I^{α}. The justification of this model comes from previous studies suggesting a similar relationship between quinine (control) and an income proxy (cotton prices). This quantity was defined as the affordability index of malaria (Terr Veen, 2005). Using the parameter α, we explicitly test here the affordability index effect of income in the model: When $\alpha=1$ income cancels out and is not a factor explaining malaria variability, and only the price of the control intervention acts as an external driver. If $\alpha=0$, we have *Py/I*, a quantity analogous to the affordability index described by Terr Veen. Given this functional form, if income modulates the affordability of control, this effect would be a non-linear one.

Per-household population was included as another external predictor. This driver has been proposed in the literature as a "proxy" for a large scale socioeconomic factor indirectly related to transmission but with relation to aggregated index for overall prosperity.

No climatic variable was included directly in the malaria regression, as previous studies have suggested that most of the influence of climate drivers, specifically maximum and minimum temperature, is only present on the seasonality of the disease, but not on its interannual variability (Terr Veen, 2005).

Human productivity equation ($\mu^Y{}_t$)*:* As mentioned in the data description, cotton in Mississippi was mostly produced by human labor, and the quantity that may reflects the only technological method used by the sharecroppers at the time to improve cotton production per unit of area is

fertilized. Thus, cotton-related human productivity, μ_Y, was a linear function of total imported fertilizer per acre per year, F. In addition to fertilizer another factor that may explain the variation of cotton productivity is malaria anomalies. This internal predictor was included as another linear factor in the regression.

Statistical inference, parameter estimation, and temporal formulation of the model

A feedback in the model occurs only when parameters $c \neq 0$, $\alpha \neq 1$, and $j \neq 0$ (Fig. 5.2). We named this model the feedback hypothesis H_0. Thus, we explicitly tested for the disconnection of the feedback loop by considering the case of parameters c or $j = 0$ or $\alpha = 1$ (Fig. 5.3). In this way, four hypotheses were compared: H_1, when $c=0$ but $\alpha \neq 1$ and $j \neq 0$, we have a model in which the direction of causality goes from Income (I) to Malaria (M) to Yield (Y). This model is called the IMY model. H_2, when $\alpha=1$ but $c \neq 0$ or $j \neq 0$, then the causality goes from M to Y to I (the MYI model). H_3, when $j=0$ but $\alpha \neq 1$ and $c \neq 0$, the causality goes from Yield to Income to Malaria (YIM model). And finally H_4, when $c=0$, $j=0$, and $\alpha = 1$. In this case the null hypothesis was considered.

The time dimension is included in the model by setting one of the internal predictors one time step in the past. Specifically, two models were constructed: Model 1 with $M_t = f(I_{t-1})$, $I_t = f(Y_t)$, and $Y_t = f(M_t)$, and Model 2 with $M_t = f(I_t)$, $I_t = f(Y_{t-1})$, and $Y_t = f(M_t)$ (Fig. 5.3, vertical dashed lines). We did not test a model of the form $Y_t = f(M_{t-1})$, as we found no evidence for infections in previous years influencing the following year's production, at least directly.

By using this temporal structure, a chain of causality can be constructed for every yearly step, and the model can be analyzed assuming a first order Markov process (Pawitan, 2001). Thus causality can be statistically tested using standard likelihood theory. Specifically, the likelihood of observing i_t, m_t, and y_t at time step t given the data i_t, y_t, and m_t is defined as

$$L_M(\theta) = \prod_{t=2}^{n} p(y_t|\hat{m}_t)p(m_t|\hat{i}_{t-1})p(i_{t-1}|\hat{y}_{t-1}),$$

for the 1-step-ahead estimation method, and

$$L_M(\theta) = \prod_{t=2}^{n} p(y_t|\dot{m}_t)p(m_t|i_{t-1})p(i_{t-1}|\dot{y}_{t-1}),$$

for the trajectory matching.

The tilde "$\hat{}$" represents the data point observed in the time before, and "$\dot{}$" the simulated value from the time step before.

We estimated the model parameters using a one step-ahead and trajectory matching estimation (Bolker, 2008). When implementing point estimation, the parameters of each sub-model are estimated based on the data from the time step before (Fig. S5.1). Trajectory matching estimation considers instead the actual dynamics of the system on time by simulating the trajectory of the system from a given set of initial conditions and then calculating the likelihood using the entire simulated time series (Fig S5.1). The likelihood function L is maximized using the package Optim with the Nelder-Mead algorithm. Because the model hypotheses are nested, a likelihood ratio test was used to compare model performances based on the χ^2 test. The confident intervals of the likelihood functions are constructed to using $+-\frac{\chi^2{}_1(1-\alpha)}{2}$. Because we are interested on confronting the feedback hypothesis, only the confidence intervals to the side of the parameter values that generate the other models were calculated.

Sensitivity Analysis

As a way to quantify the relative contribution of each driver, we carried out a sensitivity analysis using the analytical form of the model and the best parameters obtained from the inference analysis. Specifically, by setting $I_t=I_{t-1}=I>0$ and $\alpha=0$, an analytical solution of the form $AM+BM^2+C=0$ can be obtained (Appendix 5.1: Equilibrium). Using the solution of this quadratic equation, and setting all the other drivers to their observed mean, we evaluated the equilibrium (M^*, I^*) when one driver at a time was varied from the minimum to the maximum value observed.

Results

Table 5.1 shows the results of the likelihood ratio test for the trajectory matching inference, comparing the feedback model against the competing hypotheses described in the Methods section. This table shows that the feedback model performs significantly better than those that disconnect either income from malaria (model MYI) or human productivity from income (model IMY). By contrast, when the effect of malaria on human productivity was disconnected (model YIM), the maximum likelihood did not decrease significantly. The neutral model, the one that disconnects the feedback completely, exhibited the worst performance, confirming the importance of some relationship between the internal predictors in the dynamics. All the models compared in Table 5.1 are those with a one year delay in the effect of income on malaria, as these models show the best likelihoods (Fig. S5.2).

Figure 4 shows the likelihood profiles of parameters c, α, and j, respectively. In the three panels the maximum likelihood estimates are highlighted. For the three parameters, this maximum is outside the values corresponding to the alternative hypothesis (Fig. 5.4). Panel B shows that the non-linear effect of income on malaria is only significant at the 95% but not at the 99% level. Furthermore, Panel C shows that the value of $j = 0$ falls inside both the 95% and 99% CI. Similar results hold from the LRT based on the one-step-ahead estimation, except that in this case, the YIM model is also significantly different from the feedback model (Fig. S5.4). Thus while we cannot rule out completely the possibility of the feedback model, the importance of malaria on cotton productivity appears relatively weak, compared to the influence of income on malaria.

In addition, results from the sensitivity analysis further support the causality pathway of the YMI model. Figure 5.5 shows a phase diagram of the change in the equilibrium point for malaria and

income when one parameter value was changed at a time. This figure illustrates the considerable importance of cotton prices (Fig. 5.5, black dots) in driving both malaria and income in the system. Additionally, while household size H did in fact influence malaria, it did not influence income (Fig. 5.5, green dots). A similar effect is observed for Pyrethrum *Py* as well (Fig. 5.5, blue dots). These results suggest that the causality was stronger in the direction of socio-economic drivers on malaria rather than the other way around.

Finally, figure 5.6 illustrates a simulation using the best estimates for the parameters. These simulations were run for similar initial conditions and without noise. The figure shows that the dynamics of the system are poorly represented when the drivers are not included.

Discussion

With a series of nested time series model and state level data of malaria incidence, per-capita income, and cotton-related human productivity, we investigated the evidence for a poverty trap effect in the State of Mississippi between 1917 and 1938. During this period a decrease in the number of cases was evident in all of the United States, with the final elimination of the disease observed around 1950 in the southernmost states. Results from the statistical analyses provide evidence against the poverty trap hypothesis. They also suggest that the dynamics of the system are mostly dominated by the external conditions imposed on the region at the time. A sensitivity analysis further suggests that causality in the system was dominated in the direction from socio-economic drivers to malaria.

The elimination of malaria in the southern United States was a process lasting for at least 50 years, and the factors influencing this decrease remain only partially understood. Several hypotheses have been proposed to explain this decline in incidence (Humphreys 2001; Terr Veen, 2005). These

include changes in the ecology and behavior of Anopheles vectors, improvements in malaria medication, migration of the population from rural areas to urban centers, control of larval and adult mosquitoes by local authorities, and better economic conditions present in the south after the Civil War (Andrews, 1948; Andrews et al., 1950). Yet few studies have examined the data with rigorous statistical analyses.

Our findings are consistent with the hypothesis that suggest that most of the interannual variation in malaria during this time correlating with the possible reduction of domestic densities of Anopheles due to the use of insecticide. They further confirm the important role of income and cotton prices in modulating the affordability of the control (Terr Veen, 2005). However, most of the decline in incidence was explained by the number of people per household. This variable was used here as a proxy for an external large-scale socioeconomic factor indirectly related to malaria transmission. Its ecological implications for directly influencing mosquito biting behavior, and thus transmission intensity, cannot be ruled out completely, however, as the decrease in people per household had been proposed to explain the decline in malaria from areas of low transmission (Hulden & Hulden, 2009). Most likely the decreasing trend in malaria can also be attributed to other social factors such as education, house improvement, agricultural indemnification, and modernization (Bradley, 1966).

Further studies need to address higher resolution data (monthly county level observation, for example) and more mechanistic models to gain a better picture of the processes by which income and socioeconomic characteristics interacted with ecological factors in the region. In addition, socioeconomic data that more accurately represents the financial status of poor families should be obtained. The type of income data used here was limited in this regard as taxable income was reported by landowners, and therefore did not directly represent the income of the sharecroppers.

The name "sharecroppers" was given to the families renting the land and was derived from the way in which the rental agreements and payments were conducted: A person or a household paid the rent by "sharing" the harvest. Because the percentage was negotiated in advance of the next season, only information on the production from the previous season was available to the sharecroppers, and this was mostly dominated by cotton prices. Therefore it is not surprising that the model most supported by the data was the one of the form $M=f(I_{t-1})$, which suggests that the income of the sharecroppers (influenced by the affordability of malaria control) was mostly reflected in the taxable income from one year prior.

The analyses presented here are not just relevant from a historical perspective, but also from a practical point of view. Today, the tenant sharecropping system is a common practice in many places with malaria, for example in India (Deiniger et al., 2013). If similar tenant systems exist in different parts of the world, it is possible that also similar socioeconomic drivers are behind the population dynamics of the disease. Therefore, an understanding the connection between agriculture and income variability, and malaria population dynamics can help identify best policies and practices for the control of malaria. It can also lead to the consideration of large-scale economic predictions in the prediction of malaria incidence itself (Zurbrigg, 1994).

There is still a critical lack of studies in the literature that empirically address the dynamical consequences of the double causality between malaria and poverty. Although further research is needed, this work provides a quantitative framework to analyze dynamical processes linking disease ecology and economic determinants simultaneously. Such a framework can contribute insight into the current trajectory of malaria in areas of the world on the path to eliminating poverty. Coupling the intrinsic and extrinsic forces driving the biology of malaria with economic variables at different scales of analysis can raise insights into the critical paths for intervention to be effective and

efficient, and will eventually help to develop a formal integrated theory coupling ecological and socioeconomic systems (Bonds and Rohani, 2010). To achieve this goal, it is fundamental to rely on long-term time series of relevant ecological and economic data, as well as new statistical tools, to confront the double causality hypothesis.

The interaction among agricultural practices, malaria, and its vector has been tightly interwoven since the beginning of agriculture (Carter and Mendis, 2002), and most malaria cases today are still found in poor communities working the land for agricultural purposes. The results presented here suggest that the most effective route for malaria control over time will be one that integrates not just environmental and public health considerations, but also socioeconomic factors, including those that are not directly related to malaria control interventions.

Figures

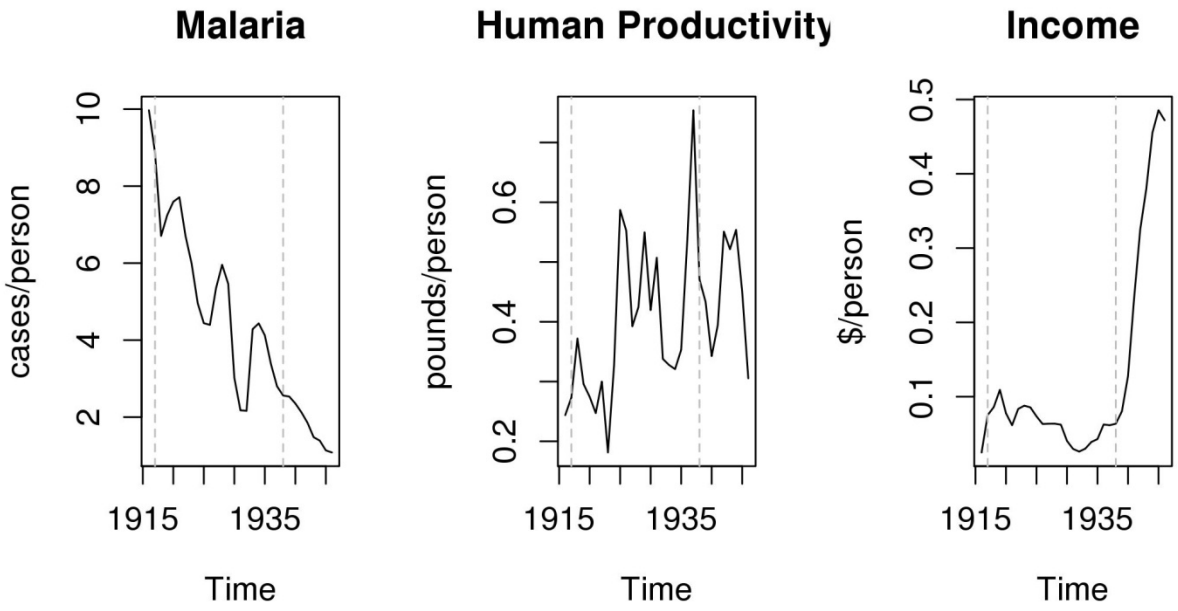

Figure 5. 1 Yearly records of malaria incidence, cotton productivity per person (human productivity), and per capita income between 1916 and 1946 for the State of Mississippi

The dashed lines represent the years 1917 and 1938, and bracket the period of time for which the poverty trap hypothesis was tested.

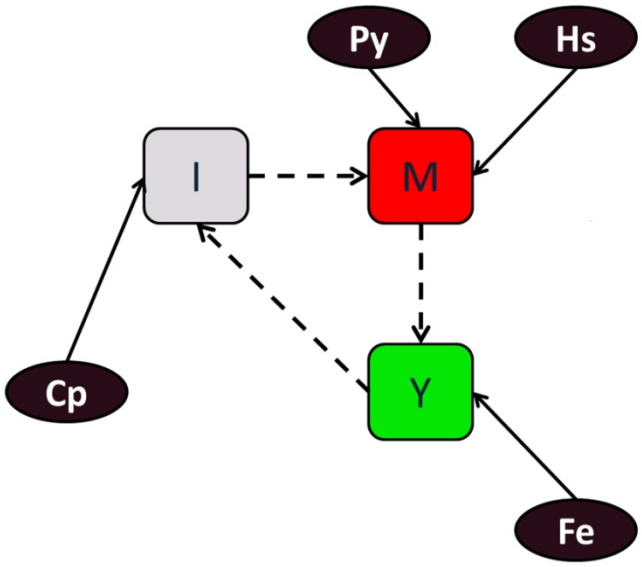

Figure 5.2 The feedback hypothesis

The three state variables are depicted in color. Malaria (*M*) in red, cotton yield/human productivity (*Y*) in green, and per-capita income (*I*) in gray. The external predictors, or drivers, of the system are

shown in black. The dashed lines represent the mechanistic feedback loop between the state variables.

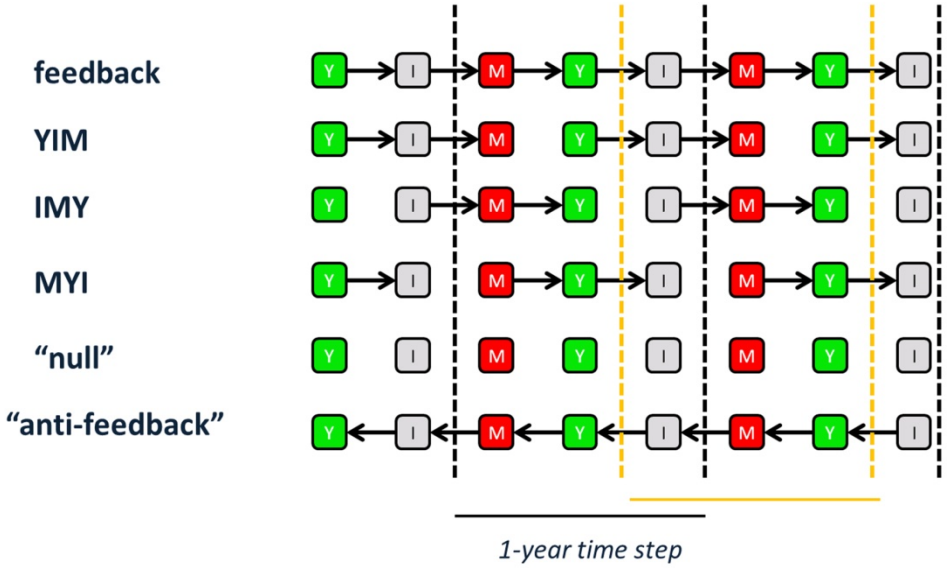

Figure 5.3 The feedback model and alternative hypotheses over time

The top row represents the feedback model in which the causality runs in both directions. Squares represent the state variables *M* (malaria), *Y* (productivity), and *I* (income). Each state variable is an internal predictor of the other variable as represented by the black arrows. The feedback hypothesis is confronted with several alternatives in which a particular mechanistic link is disconnected (second, third, and fourth rows; see Methods). A model in which the system is completely disconnected is shown in the fifth row ("null model"). For completeness, a model in which the direction of causality was completely reversed was also tested (bottom row). The dashed vertical lines correspond to two different hypotheses concerning the time delay involved in cause and effect. Specifically, black vertical lines illustrate a model of the form $M_t = f(I_{t-1})$, $I_t = f(Y_t)$ and $Y_t = f(M_t)$, in which malaria is influenced by income of the previous year, whereas income depends on productivity, and similarly malaria affects productivity, in the same year. Orange lines illustrate instead a model of the form $M_t = f(I_t)$, $I_t = f(Y_{t-1})$ and $Y_t = f(M_t)$, where now malaria and income are related with no delay, but yield in the previous year influences income.

	logLik	AIC	df	Chisq	Pr(>Chisq)
feedback	37.27	-48.53	13	-	-

YIM	35.94	-47.88	12	2.65	0.1
IMY	28.73	-33.47	12	17.07	3.61E-05
MYI	35.04	-46.09	12	4.45	0.03
NULL	23.95	-27.89	10	26.64	6.99E-06

Table 5.1 Likelihood ratio test results

The feedback model (top row) is best supported by the data based on both the log likelihood (second column) and the Akaike information criterion (AIC; third column) which corrects for the different degrees of freedom (df). The likelihood ratio test shows that this model is significantly better than the model in which human labor did not influence income (IMY), and the one in which income did not influence malaria (MYI). However, a model in which malaria does not influence human labor (YIM) did not differ significantly from the feedback model, suggesting that the causality from socioeconomic conditions to malaria was stronger than the other way around. The completely disconnected (Null) model performed the worst.

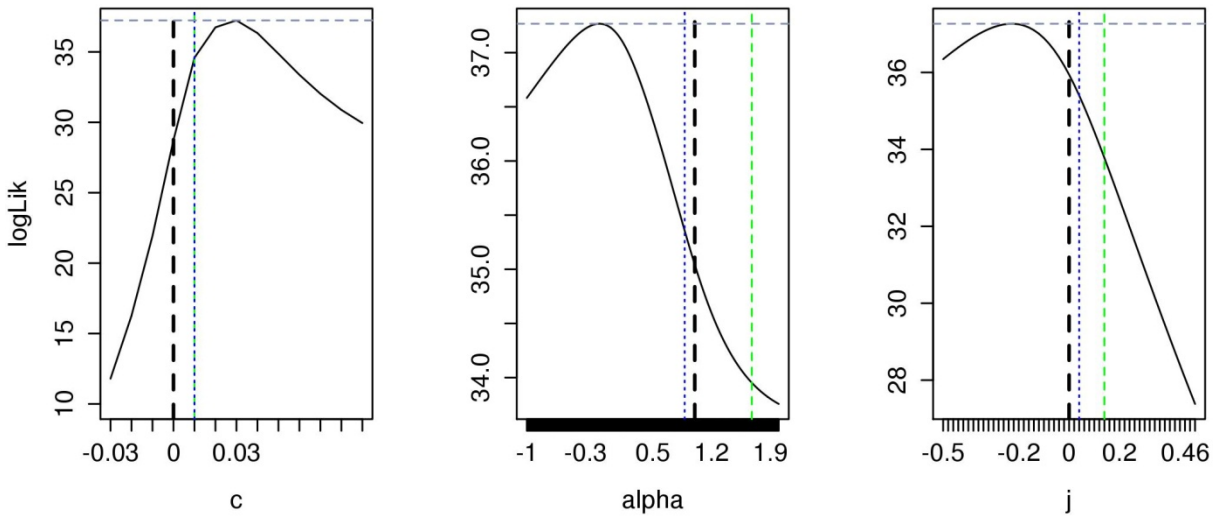

Figure 5.4 Likelihood profile for the parameters testing the poverty trap hypothesis

This figure shows the profile likelihood for the following three parameters: c, connecting human productivity to income (panel A); α, connecting income to malaria (Panel B); and *j*, connecting the effect of malaria on human productivity (Panel C). This figure shows that the maximal likelihood estimator (MLE) of the feedback model (vertical dashed and gray lines) lies outside the parameter values when this is disconnected, as defined in the Methods section and represented here by the vertical dashed black lines ($c=0$ in A, $\alpha=1$ in B, and $j=0$). The vertical blue and green lines represent the 95% and 99% confidence intervals, respectively, from the likelihood ratio test (Table 1).

Parameter c is positive and significantly different from 0 at 95%, as well as 99% CI, suggesting a positive relationship between income and the productivity per person. Furthermore, the MLE estimator for parameter α (Panel B) is significantly different from the model's 95% but not at 99%. Finally, while the MLE for j (Panel C) is negative, this value is not significantly different from the model in which the effect of malaria is not included ($j=0$).

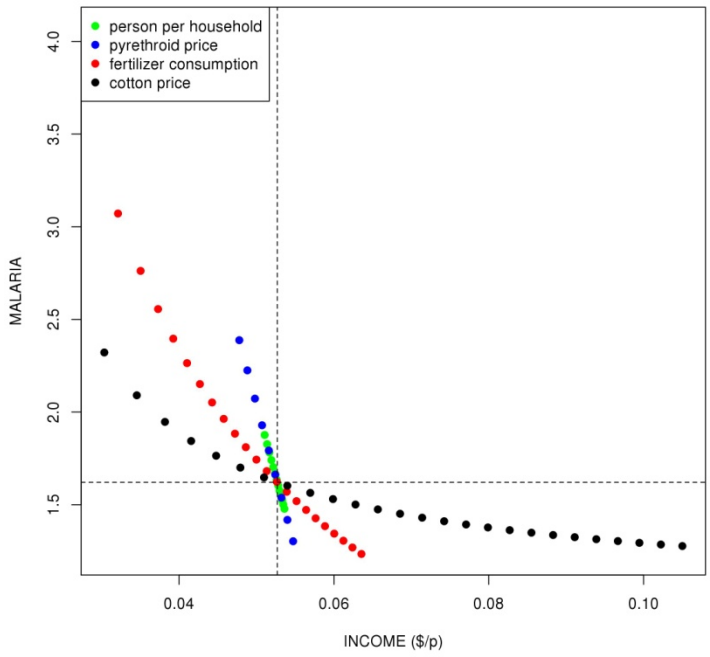

Figure 5.5 Phase plot of the malaria-income equilibrium for different values of the drivers

The vertical dashed lines in the figure display the equilibrium point (I^*, M^*) for the analytical model (SI, Text 1) using the best parameters estimated from the feedback model (Fig. 5.2) and the average value of the drivers. Each colored point represents a combination of malaria and income values for a particular value of a given driver when all others are kept constant at their arithmetic mean value. A large influence of socio-economic drivers on malaria's deviation from the mean can be observed. On the contrary, a relatively low effect of malaria-related drivers on income is observed (household size and pyretroid flower prices, represented by the green dots and blue dots, respectively). Malaria and income are non-linearly related in the direction from income to malaria.

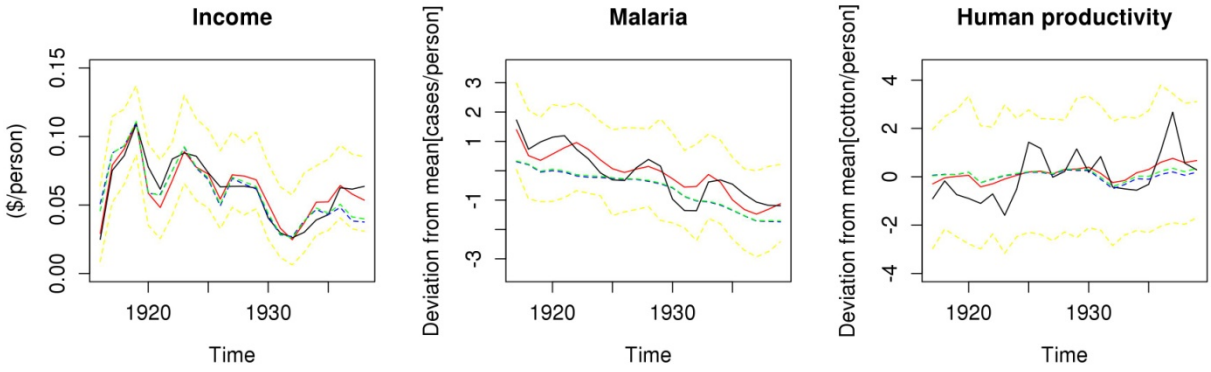

Figure 5. 6 Time series of simulation results

Yearly estimation of per capita income (Panel A), malaria anomalies (Panel B), and human productivity anomalies in pounds of cotton per person (Panel C) using the parameters estimated with the feedback model (Table 1). Black lines represent the observations and green lines, the predictions obtained by simulation from specified initial conditions. Blue and green dashed lines correspond respectively to the simulations for the parameters at their 95% and 99% confidence interval limits (See Fig. 5.4). IN yellow the max and min value observed of 1000 random draws.

Appendices

Appendix 5.1: Equilibrium

$$I_t = a + bCT_t + cY_{t-\Delta} + \varepsilon_I$$

$$M_t = d + \frac{f(P_t)}{\hat{I}_{t-\Delta}}(I_{t-\Delta})^{\alpha} + gH + \varepsilon_M$$

$$Y_t = i + jM_t + kF + \varepsilon_Y$$

I_t, M_t and Y_t are per-capita income, malaria prevalence and per-capita productivity respectively. CT_t is the cotton price at year t, P_t pyrethroid flowers import price, $\mathrm{H_t}$ dwelling population per house and $\mathrm{F_t}$ pounds per hectare of fertilize imported per year. Setting $\alpha = 1$ and the drivers D to it arithmetic mean $\overline{\mathrm{D}}$, at equilibrium $I_t = I_{t-\Delta} = I^* > 0$

$$M^*I^* - I^*[d + gH] = fP$$

with

$$M^*I^* = M^* * [a + bCT + c(i + jM^* + kF)]$$

and

$$-I^*[d + gH] = -[a + bCT + c(i + jM^* + kF)][d + gH]$$

By simple algebra we get a quadratic equation of the form

$$A + BM^* + CM^{*2} = 0$$

with

$$A = -(a + bCT + cid + cigH + cdkF + ckgFH + fP)$$

$$B = a + CT + ci + ckF - cjd - cjgH$$

$$C = cj.$$

And the solution

$$M^* = \frac{\pm B + \sqrt{B - 4AC}}{2A}$$

$$I^* = a + bCT + c(i + jM^* + kF).$$

The solution (I^*, M^*) is plotted fir different values of a D driver Di from the min(D) to max(D) (Fig. 5.5).

Appendix 5.2: Supporting Figures

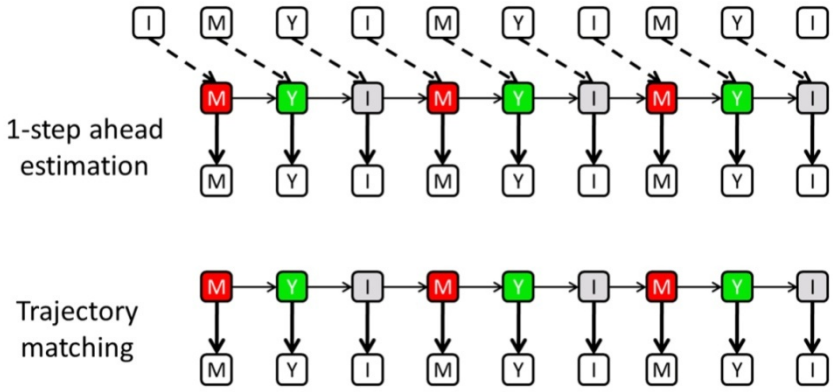

Figure S5.1 Graphical representation of the estimation methods

Data on malaria incidence (*M*), human productivity (*Y*), and income (*I* white squares) were used to estimate the relationship between the state variables represented in red, green, and gray, respectively. On top, the 1-step-ahead estimation method used data from previous observations to estimate the model. Panel B, on the bottom, represents the trajectory matching method based on shutting a simulated trajectory from initial conditions. The likelihood of the simulated trajectory (color squares) is computed using the observed data, represented in white at the bottom of the color squares.

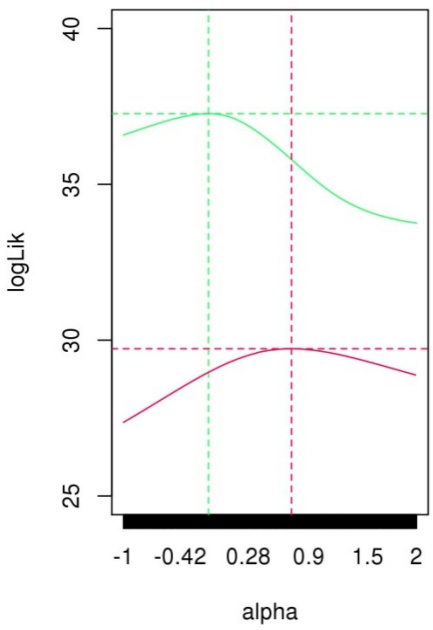

Figure S5.2 Likelihood profile comparison of parameter α

The figure displays the profile likelihood of the parameter α for the model $M_t = f(I_{t-1})$, $I_t = f(Y_t)$ and $Y_t = f(M_t)$ (solid green line) and the model $M_t = f(I_t)$, $I_t = f(Y_{t-1})$ and $Y_t = f(M_t)$ (solid red line). Vertical dashed and horizontal lines represent the parameter values and the likelihood values of the corresponding MLEs for comparison. The temporal causality for the model with the lag from income to malaria is clearly better supported by the data.

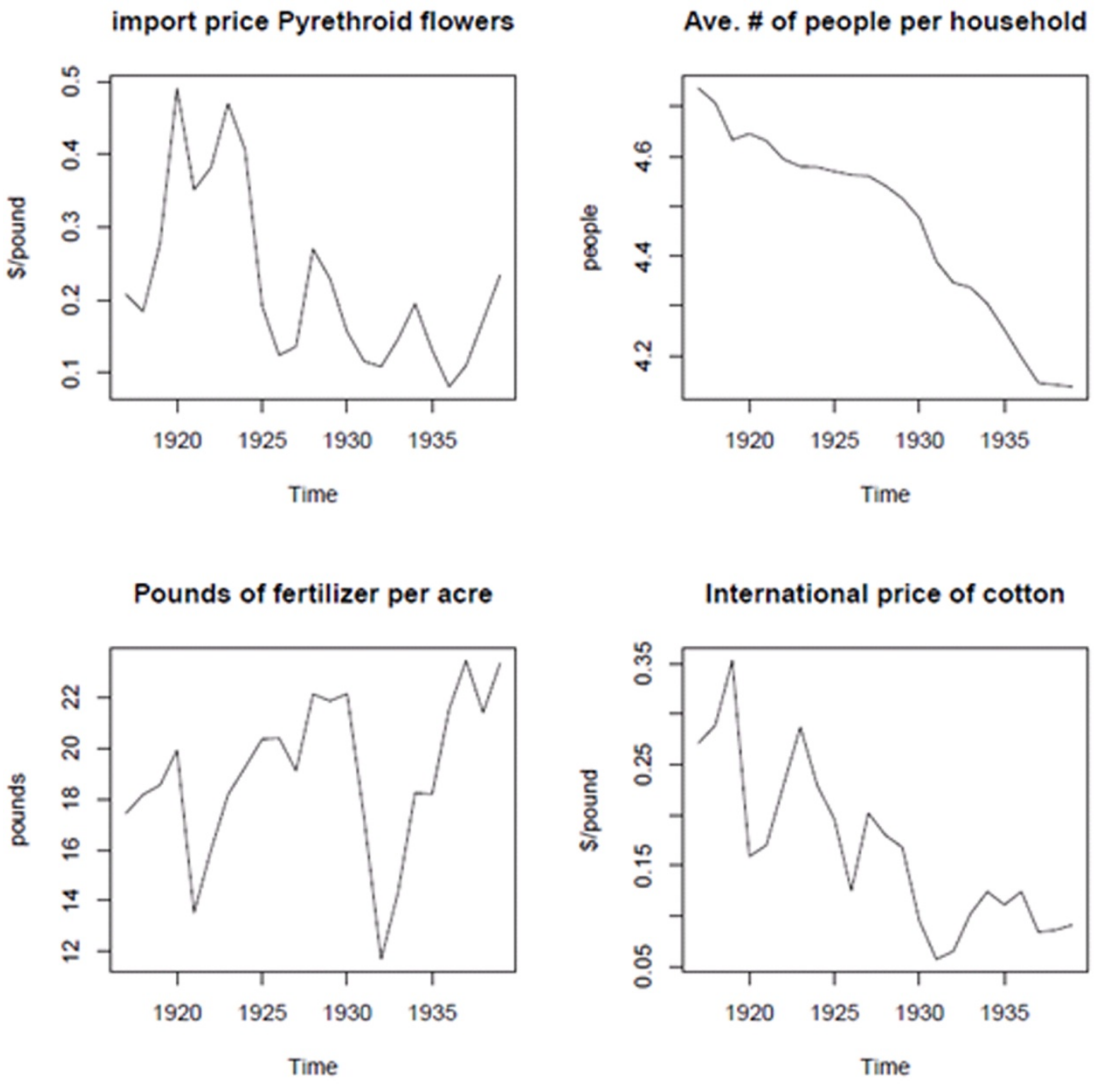

Figure S5.3 Time series of the drivers or external predictors of the model.

The two top plots show the drivers of the malaria equation. On the left, the price of importation of pyrethroid flowers is shown in dollars per pound; on the right, the average number of people per household in Mississippi. On the bottom row, the driver of human productivity, the total weight of fertilizer per acre per year, is shown (left). The international price of cotton, on the right, was used as the driver in the income equation.

	logLik	AIC	df	Chisq	Pr(>Chisq)
feedback	37.27	-48.53	13	-	-
YIM	35.94	-47.88	12	2.65	0.1
IMY	28.73	-33.47	12	17.07	3.61E-05
MYI	35.04	-46.09	12	4.45	0.03
NULL	23.95	-27.89	10	26.64	6.99E-06

Table S5.1 Model comparison

The table shows the results of the model comparison based on the maximal log-Likelihood (logLik) for the trajectory matching estimation method. A Likelihood Ratio test (LRT) was conducted using an ANOVA table (Chisq and Pr). In addition, the Akaike Information Criteria (AIC) was used to correct for the number of parameters (df).

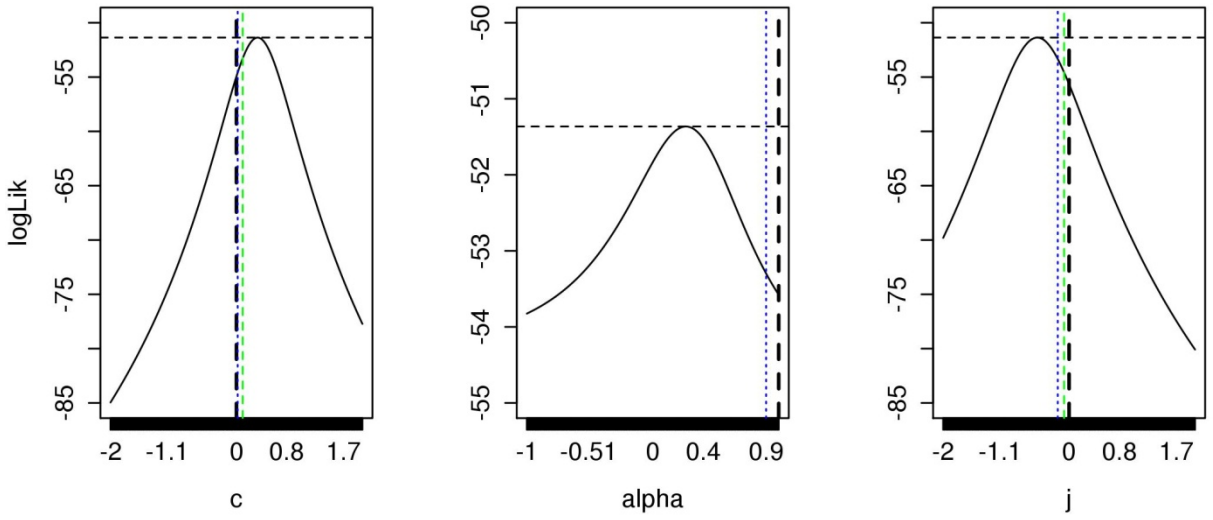

Figure S5.4 Likelihood profile of the poverty trap parameters using a 1-step-ahead estimation method.

The figure depicts the likelihood profile in a similar way than that presented in Figure 5.4, with the exception that in this case the effect of malaria is significant in explaining cotton-related human productivity.

References

Andrews, J.A., Quinby, G.E., Langmuir, A.D., 1950. Malaria Eradication in the United States. Am J Public Health N 40, 1405-1411.

Andrews, J.M., 1948. What's Happening to Malaria in the USA? Am J Public Health N 38, 931-942.

Bolker, B., 2008. Ecological Models and Data in R. Princeton University Press, Princeton, NJ.

Bonds, M.H., Keenan, D.C., Rohani, P., Sachs, J.D., 2010. Poverty trap formed by the ecology of infectious diseases. P R Soc B 277, 1185-1192.

Bonds, M.H., Rohani, P., 2010. Herd immunity acquired indirectly from interactions between the ecology of infectious diseases, demography and economics. J R Soc Interface 7, 541-547.

Bradley, G.H., 1966. A review of malaria control and eradication in the United States. Mosquito News 26, 462-470.

Carter, R., Mendis, K.N., 2002. Evolutionary and historical aspects of the burden of malaria. Clin Microbiol Rev 15, 564-+.

Chiyaka, C., Tatem, A.J., Cohen, J.M., Gething, P.W., Johnston, G., Gosling, R., Laxminarayan, R., Hay, S.I., Smith, D.L., 2013. The Stability of Malaria Elimination. Science 339, 909-910.

Chuma, J.M., Thiede, M., Molyneux, C.S., 2006. Rethinking the economic costs of malaria at the household level: Evidence from applying a new analytical framework in rural Kenya. Malaria J 5.

Deininger, K., Jin, S.Q., Yadav, V., 2013. Does Sharecropping Affect Long-term Investment? Evidence from West Bengal's Tenancy Reforms. Am J Agr Econ 95, 772-790.

Filmer, D., 2005. Fever and its treatment among the more and less poor in sub-Saharan Africa. Health Policy Plann 20, 337-346.

Gallup, J.L., Sachs, J.D., 2001. The economic burden of malaria. Am J Trop Med Hyg 64, 85-96.

Guiguemde, T.R., Coulibaly, N., Coulibaly, S.O., Ouedraogo, J.B., Gbary, A.R., 1997. A precise method for estimating the economic costs of malaria: Application of the method in a rural area in Burkina Faso (West Africa). Trop Med Int Health 2, 646-653.

Hulden, L., Hulden, L., 2009. The decline of malaria in Finland - the impact of the vector and social variables. Malaria J 8.

Humphreys, M., 2001. Malaria. Poverty, Race, and Publich Health in the United States. John Hopkins University Press, Baltimore.

Laxminarayan, R., 2004. Does reducing malaria improve household living standards? Trop Med Int Health 9, 267-272.

Lisansky, E., 1958. Eradication of Malaria as an Endemic Disease in the United-States. Ann Intern Med 48, 428-438.

Malaney, P., Spielman, A., Sachs, J., 2004. The malaria gap. American Journal of Tropical Medicine and Hygiene 71, 141-146.

Packard, R.M., 2007. The Making of a Tropical Disease. A Short History of Malaria. Johns Hopkins University Press.

Packard, R.M., 2009. "Roll Back Malaria, Roll in Development"? Reassessing the Economic Burden of Malaria. Popul Dev Rev 35, 53-+.

Pawitan, Y., 2001. In All Likelihood. Statistical modelling and Inference Using Likelihood. Oxford University press, Oxford.

Plowright, R.K., Sokolow, S.H., Gorman, M.E., Daszak, P., Foley, J.E., 2008. Causal inference in disease ecology: investigating ecological drivers of disease emergence. Front Ecol Environ 6, 420-429.

Sachs, J., Malaney, P., 2002. The economic and social burden of malaria. Nature 415, 680-685.

Schulman, B., 1994. From Cotton Belt to Sunbelt. FEderal Policy, Economic Development, and the Tranformation of the South, 1938-1980. Duke University Press, Durham.

Sharma, V.P., 1999. Current scenario of malaria in India. Parassitologia, Vol 41, Nos 1-3, September 1999, 349-353.

Somi, M.F., Butler, J.R., Vahid, F., Njau, J., Kachur, S.P., Abdulla, S., 2007. Is there evidence for dual causation between malaria and socioeconomic status? Findings from rural Tanzania. Am J Trop Med Hyg 77, 1020-1027.

Teklehaimanot, A., Mejia, P., 2008. Malaria and poverty. Ann Ny Acad Sci 1136, 32-37.

Terr Veen, J.L.M., 2005. Determinants of malaria transmission in the United States between 1900 and 1946, School of Hygiene & Tropical Medicine. University of London, London.

White, M.T., Conteh, L., Cibulskis, R., Ghani, A.C., 2011. Costs and cost-effectiveness of malaria control interventions - a systematic review. Malaria J 10.

Worrall, E., Basu, S., Hanson, K., 2005. Is malaria a disease of poverty? A review of the literature. Trop Med Int Health 10, 1047-1059.

Yasuoka, J., Mangione, T.W., Spielman, A., Levins, R., 2006. Impact of education on knowledge, agricultural practices, and community actions for mosquito control and mosquito-borne disease prevention in rice ecosystems in Sri Lanka. American Journal of Tropical Medicine and Hygiene 74, 1034-1042.

Chapter 6

Conclusions

Summary of major findings

This dissertation addressed several aspects the dynamical consequences of including agriculture and socioeconomic factors to explain the population distribution of malaria in space and time. Feedbacks operating in both directions where demonstrated, especially in the context of climate variability. Together the different chapters present several explanations and possibilities for how these feedbacks affect the inter-annual cycles of malaria.

Chapter II showed how the signal from climate forcing in the inter-annual variability of malaria in semi-desert India is decoupled in highly irrigated landscapes. We presented evidence to suggest that mosquito control intervention is the most likely driver behind this decoupling. In addition, I offered data supporting the persistence of more endemic malaria in highly irrigated areas. Finally, by using remote-sensing information about the vegetation, an operational regional index for forecasting environmental malaria risk in areas of low irrigation is now available.

In Chapter III I explored further the implications of mosquito control intervention and malaria population dynamics in highly variable environments. With a coupled malaria-mosquito model parameterized for the Northwest states of India, I showed how the policy and the implementation of vector control in Gujarat can generate a mechanistic feedback between disease and intervention levels. This feedback can in turn generate multi-annual cycles in malaria incidence in areas of low

rainfall and low vectorial capacity. In the presence of rainfall variability, these cycles can modify the effectiveness and the efficacy of intervention, and subsequently the likelihood of malaria elimination. Including climatic information in the malaria control program for vector-borne diseases would facilitate synchronization of the natural cycles of both disease and intervention, and in so doing, prove more efficient in the long term.

Chapter IV synthesized the changes in malaria risk that occur under irrigation intensification and socioeconomic development over time. A transitional zone in malaria risk, characterized by high risk despite high mosquito control, was described for newly irrigated areas. This transitional zone was shown to be protracted, having lasted already over a decade. It is positioned between a climate-driven high malaria risk zone, with low or nil irrigation development and poor socioeconomic conditions, and a low-malaria risk zone where control is no longer required and overall economic conditions have significantly improved. These results highlight the need for including Health Impact Assessments into water and irrigation developmental projects over long-term horizons.

Chapter V explored the last stage of malaria before elimination in the United States during the first part of the twentieth century, with an emphasis on testing the poverty trap hypothesis between income and malaria. A time series model, parameterized using yearly economic and epidemiological data, and connecting human productivity with malaria, indicated that the influence of income and cotton price on malaria incidence was stronger than the effect that malaria itself on human productivity. The influence of income on the purchase power for control measures appears to be non-linear. The combination of high prices for control and low income due to low cotton price seems to have been an important factor behind the amplitude of malaria epidemics. Elimination seems to have been mostly driven by large-scale socioeconomic factors related to the overall improvement in the economy of the United States.

Future directions

Climate forcing

In semi-desert environments climate forcing continues to be an important factor in the dynamics of malaria transmission and mosquito ecology, even in the presence of large-scale land-use changes. This influence, however, seems to operate via more channels when those land changes occurred: by influencing mosquito ecology directly, but also by influencing crop production and thus the economic vulnerability of the population. Several questions remain open for future research in order to understand the implications of climate variability and climate change for epidemic vulnerability and preparedness.

First, most of the analyses presented in this work addressed the magnitude of epidemics between years by accumulating cases over a season or the entire year. Because of this, climate influences were also aggregated, and rainfall related data were also accumulated over a season (e.g., monsoon rainfall, NDVI September). It is not clear, however, how the accumulation of rainfall within a season may influence the phases of an epidemic, especially at the beginning of the rainy season. It is likely that to generate transmission in such a short window of time, changes in mosquito abundance must be greatly amplified by the frequency and the temporal correlation of rainfall events. Detailed models connecting rainfall, surface water dynamics, mosquito ecology, and malaria transmission are a critical tool needed for our understanding of malaria epidemics within a season. Similarly, changes in maximal temperature can influence mosquito longevity (Russell, 1942) that, in turn, might impact the vectorial capacity, and recent findings suggest that daily temperature fluctuations can control the intensity of malaria transmission (Paaijmans et al., 2010). At the moment, we lack the knowledge on how the interaction between those two climatic drivers might influence mosquito ecology and

ultimately the transmission of the parasite. This information is critical for the parameterization of more accurate mosquito models that can provide more reliable representation of the timing of the epidemics. Through a better understanding of these climate-related mechanisms on mosquito ecology, especially on intraseasonal time scales, we will be able to investigate other strategies for control that can be both more effective and sustainable over the long term.

In addition, climate factors also have an impact on crop productivity, which can then influence mechanisms by which food availability modifies human susceptibility and transmission in years of food scarcity, a topic which is still poorly understood. Similarly, the hypothesis of the contribution of famine to the spread of epidemics of malaria remains controversial (Zurbrigg, 1994). This opens the doors for fascinating ecological questions: For example, how do changes in food availability over time influence the immune response of a population under high and seasonal levels of parasitemia? How is the evolution of the parasite shaped by large-scale climatic patterns, considering the large diversity of strategies that the parasites have to avoid detection by the immune system? And how do mechanisms at these disparate two scales interact over time? A multi-sectorial and multi-scale perspective, involving geneticists, ecologists, climatologists, and economists, is needed to face the continuing challenges of malaria control and elimination.

Malaria population dynamics and agriculture development

Future scenarios of water availability for India indicate that for the next decades this resource will be scarcer (Postel 1993; Rodell et al., 2009). In the opposite direction, food consumption, and therefore production, is expected to increase in the near future (Poster, 1998), increasing the pressure towards more infrastructure for storage and its distribution to vaster areas (Jha, 2013). Thus, water-related infrastructure projects will most likely continue to increase in areas where water availability is the limiting factor for development, and the confrontation of those two opposite trends will contribute

to an increase in the risk of malaria and other zoonotic diseases by generating unoccupied and sometimes abandoned irrigated infrastructure with the potential for increased vector populations. In addition practices of water storage common in urban and peri-urban areas of India were another malaria vector (Anopheses spefensis) thrives are likely to increase as well, generating even more habitat for mosquitoes.

In this scenario, a new approach to agricultural development, better interwoven with vector-borne disease control, is needed to tackle these potential threats and to generate long-term sustainable solutions (Batterman et al., 2009). Several methods to decrease vector populations without compromising food productivity have been tested before, most of them before the GMEC, and some of them have been successfully applied to malaria reduction locally (see intermittent irrigation, for example Keiser et al., 2002). It is not clear, however, how to extrapolate these methods to large-scale irrigation projects in such a way that regional risk can be quantified and compared. On this front, the development of models combining water distribution in irrigation channels at multiple scales with mosquito population dynamics (Otta and Kaga, 2013) and their relationship to vectorial capacity and transmission would provide valuable information on the ways in which water can be efficiently distributed to the fields and at the same time decrease the potential risk for diseases. These models would also provide a way to include in early stages of the projects estimation about the possible health impact and the costs under different regimes of water distribution. Interdisciplinary collaboration between hydrologists, engineers, entomologists, ecologists, and geographers, among others, would be crucial to generate such models.

No other human-nature system would provide a better example than malaria, of the importance of considering feedback mechanisms, given its entrenched historical relationship to human prosperity. Due to the inherent evolutionary link between agriculture, mosquito ecology, and malaria, it is likely

that future changes in land-use will continue to have consequences for environmental risk of this and other vector-borne diseases. Moreover, since most of the poorest, most isolated and most vulnerable populations in the world are rural communities, the burden of malaria will most likely continue to affect them. Most of the people in these communities practice subsistence agriculture, exchanging crops for cash each season. Usually, these populations are also without access to education and seldom have public health access, suggesting that the socioeconomic conditions in which these communities are living today are very similar to those present at the dawn of agriculture. The complexity of malaria's dynamics and the elusive nature of the malaria vaccines create an important challenge for ecologists and policymakers to generate solutions that can systematically reduce both disease risk and poverty. This challenge is even greater in endemic areas of transmission where the elimination of the disease is unlikely to be achieved by only malaria-specific control methods.

References

Batterman, S., Eisenberg, J., Hardin, R., Kruk, M.E., Lemos, M.C., Michalak, A.M., Mukherjee, B., Renne, E., Stein, H., Watkins, C., Wilson, M.L., 2009. Sustainable Control of Water-Related Infectious Diseases: A Review and Proposal for Interdisciplinary Health-Based Systems Research. Environ Health Persp 117, 1023-1032.

Jha, M., 2013. Sustainable Management of Groundwater Resources in Developing Countries: Constraints and Challenges, in: Ramkumar, M. (Ed.), On a Sustainable Future of the Earth's Natural Resources. Springer Berlin Heidelberg, pp. 325-348.

Keiser, J., Utzinger, J., Singer, B.H., 2002. The potential of intermittent irrigation for increasing rice yields, lowering water consumption, reducing methane emissions, and controlling malaria in African rice fields. J Am Mosquito Contr 18, 329-340.

Ohta, S., Kaga, T., 2013. Effect of irrigation systems on temporal distribution of malaria vectors in semi-arid regions. Int J Biometeorol, 1-11.

Paaijmans, K.P., Blanford, S., Bell, A.S., Blanford, J.I., Read, A.F., Thomas, M.B., 2010. Influence of climate on malaria transmission depends on daily temperature variation. Proc Natl Acad Sci U S A 107, 15135-15139.

Postel, S., 1993. The Politics of Water. World Watch 6, 10-18.

Postel, S.L., 1998. Water for food production: Will there be enough in 2025? Bioscience 48, 629-637.

Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature 460, 999-1002.

Zurbrigg, S., 1994. Re-thinking the "human factor" in malaria mortality: the case of Punjab, 1868-1940. Parassitologia 36, 121-135.