On the dynamics of two efficient malaria vectors of the Afrotropical region: *Anopheles gambiae s.s.* and *Anopheles arabiensis*

Torleif Markussen Lunde

Dissertation for the degree of Philosophiae Doctor (PhD)

University of Bergen

2013

Preface

Six years ago, working with cod and salmon, the grand son of the meteorologist Vilhelm Bjerknes, Vilhelm Bjerknes (Jr.), gave me an advice; "Don't waste your time on caged fish. It is extremely boring.". Following his advice, I stared as a coordinator for the Nile basin Research Programme, where the topic was climate and malaria. Realizing the relationship between the two was poorly understood, I figured it could be worth spending some time digging into the topic. At that time, I got the impression from media, the magnitude of climate change was virtually certain, including the consequences it would lead to. While I was not too surprised the relationship between climate and malaria had some uncertainties, I was more puzzled when I realized how crude and insufficiently the description of African weather was, and even more disillusioned when I experienced the limitations of reanalysis. This meant I had to spend more time on understanding the weather and climate of Africa, which was not the plan in the first place.

Both supervisors, Bernt Lindtjørn and Asgeir Sorteberg, gave me enormous freedom to study the different sides of the relationship between malaria and climate. Bernt, always responding quickly, provided valuable input from the field in Ethiopia, and advising how the model could become readable for non-modellers. Asgeir, with a clear and simple language, guided me through the basics of atmospheric dynamics, and how climate models work. Both of them have always had time for questions and discussions, which I am grateful for.

Ellen Viste has an extraordinary ability to formulate questions in a pleasant way. Many times she has been able to formulate what I have been thinking, and ideas have been rejected after talking to her. The model development would not have been possible without our Ethiopian partners. Their knowledge about the weather, mosquitoes and malaria have been crucial in developing an understanding of the dynamics of malaria. On the scientific side, special thanks to Diriba Korecha, Meshesha Balkew, Eskindir Loha and Fekadu Massebo. I would also like to thank Dereje Tesfahun, Adugna Woyessa, Abebe Animut, and Wakgari Deressa for their hospitality, and helpful comments.

Putting together a malaria model is not only a scientific challenge, but also a technical one. I would like to thank the following people for valuable comments with respect to the implementation and integration with WRF: Alexander Oltu, Micheld d. S. Mesquita, Ulla Heikkilä, Mads Benstad, David Gill, Marius Jonassen, and Ingo Bethke.

On the social side there are many to thank: Marte Jurgensen and Nils Gunnar Sognstad which I shared office with for six months at Centre for International Health, and

at the Geophysical Institute, Iselin Medhaug, Mathew Reeve, Stephan Sobolowski, and Erik Kolstad. My office (and coffee) mates over the year Christophe Bernard, Marius Årthun, MaoLin Shen, Laura Ciasto, Erlend Knudsen and Justin Wettstein. I would also like to thank my wife, Luise, our kids, Brigt, Asgeir and Eivor Marie, my family and Logen for supporting me throughout the years.

Over the four years I have realized how little we know about African weather and global malaria, making this topic scientifically interesting.

Abstract

Weather and climate are only some of the factors influencing the dynamics of malaria. With the ongoing debate on the consequences of climate change, there is a need for models which are designed to address these questions. Historically, models have focused on the theoretical principles of eradication, with less emphasis on a changing environment. To estimate the potential impact of climate change on malaria, we need new models which consider a wider range of environmental variables.

In this thesis, we point at some factors which are important to robustly project the influence of climate and weather on malaria. These factors are described using a mathematical model which focus on the weather sensitive parts of malaria transmission; the mosquitoes and the parasites.

Mosquitoes transmitting malaria belong to the genus *Anopheles*. There are about 460 known anophelines, where 41 are considered to be dominant vectors of malaria. Each of these species have its own life history, and consequently weather and climate influence each species differently. In Africa, the public health impact of malaria is devastating, despite variable transmission. The most efficient mosquitoes are found in this continent: among them *Anopehels gambiae sensu stricto* and *Anopheles arabiensis*, which are considered to be of major importance.

In this thesis (Paper I) we describe a dynamical model which include these two species. Based on a literature review, we formulate a model which allows weather to influence each of the two species according to their life history. They compete over puddles, important for reproduction; *An. gambiae s.s.* mainly feed on humans opposed to *An. arabiensis* which feed on cattle and humans; they are allowed to disperse, meaning new ares can be occupied by the species; and as they become older, the daily probability of survival changes. Many of these factors are not important in a short time perspective. But, since climate change is slow process compared to the life of a singe mosquito, there is a need for additional complexity to study how a slowly changing environment influence the population dynamics of these malaria vectors.

To have confidence the model is realistic in the current climate we validated the model in paper II. To date, we constructed the most extensive database on the occurrence of the two mosquitoes. These data were used to validate the model described in paper I. We concluded the mosquito model produced comparable or better results than existing predictions of the two species under current climate.

An. arabiensis feed on humans and cattle. Since the density and distribution of

those are not static, but are changing over time, and the distribution of *An. arabiensis* is highly dependent on the density of cattle, there is a need to; 1. Document historical changes; 2. Understand how they are influenced by the environment. In paper III we reconstruct the cattle distribution and density in the 1960s, and show how climate variability influence the national cattle holdings. While climate variability has a minor influence in many countries, we also find variations in the climate can explain more than 40% of the national cattle holdings in some countries. The data developed in this paper can be used in the model described in paper I, as well as other studies where cattle is an important part of the system.

It has been claimed the optimal temperature for malaria transmission is between 30 to 32$^{\circ}C$, with the potential increasing linearly from 20 to 32$^{\circ}C$. With this claim, any warming in sub-Saharan Africa would potentially cause more malaria. Using the model developed in paper I, we show malaria transmission is most effective around 25$^{\circ}C$, with a decline in efficiency over end below this temperature (Paper IV). This disputes the theory claimed in previous papers. Any projections relating temperature and malaria should be interpreted with care.

The influence of climate change on malaria transmission is still uncertain. With this thesis, we have come a step further in understanding how the environment can alter malaria transmission. However, the future occurrence of malaria is dependent on many other factors, including malaria control measures, access to and usage of treatment, city planning, and immunity.

List of papers

1. TM Lunde, D Korecha, E Loha, A Sorteberg and B Lindtjørn. A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis, *Malaria Journal* 2013, **12**:28

2. TM Lunde, M Balkew, D Korecha, T Gebre-Michael, F Massebo, A Sorteberg and B Lindtjørn. A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Validation of species distribution and seasonal variations, *Malaria Journal*, **12**:78

3. TM Lunde and B Lindtjørn. Cattle and climate in Africa: How climate variability has influenced national cattle holdings from 1961-2008 *PeerJ* 55

4. TM Lunde, MN Bayoh, B Lindtjørn. How malaria models relate temperature to malaria transmission *Parasites & Vectors* 2013, **6**:20

Contents

List of Figures

Chapter 1

Introduction

> *... quartan fevers have the simpler characteristics. Nearly always they begin with shivering, then heat breaks out and the fever having ended, there are two days free; thus on the fourth day it recurs. But of tertian fever there are two classes. The one, beginning and desisting in the same way as quartan, has merely this distinction, that it affords one day free, and recurs on the third day. The other is far more pernicious; and it does indeed recur on the third day, yet out of fourty-eight hours, about thirty-six, sometimes less, sometimes more, are in fact occupied by the paroxysm, nor does the fever entirely cease in the remission, but only becomes less violent.*
>
> Celsus (25 BC–54 AD)

1.1 Background

Malaria has been around for thousands of years, and is still a major problem today. Despite efforts to eradicate malaria over the past 100 years, 149–274 million cases and 537,000–907,000 deaths from malaria occur in sub-Saharan Africa each year [5, 6]. Malaria is caused by microorganisms belonging to the genus *Plasmodium*, and can infect reptiles, birds and mammals. Of the more than 100 *Plasmodium* species, four of these infect humans. The transmission of the disease from one human to another involves mosquitoes of the genus *Anopheles*.

Female Anopheles mosquitoes require blood for egg production. If a Anopheles mosquito carrying malaria parasites bites a human, sporozoites can be transferred to the blood stream where they travel to the liver and enter liver cells. The coming days and months the sporozoites start asexually production of merozoites (tissue schizogony), which return to the blood. In the bloodstream the merozoites enter red blood cells where the production of merozoites continues. Some of the merozoites develop into the sexual form of Plasmodium parasites; gametocytes. When a new mosquito take a blood meal it may take up blood cells containing gametocytes. These develop into gametes, which in turn produce diploid zygotes developing into ookinetes. When the ookinetesenter the midgut wall of the mosquito, oocysts can be formed. The oocysts produce sporozoites which are released in the mosquitoes. The sporozoites then enter the salivary glands of the mosquito, and can again be transmitted to humans.

In 2000, Rogers and Randolph [7] discussed how malaria risk is modulated by climate, and suggested that the simplifications made in biological models means the models may not be suitable to assess the impact of climate change on malaria transmission. Classical biological models include the mosquito blood-feeding interval, mosquito longevity, and the development period of the malaria parasite inside the mosquito. All three factors are affected by temperature. If it is not too cold or hot for these factors, the number of mosquitoes is mainly driven by the formation of breeding sites. The formation of breeding sites is dependent on infiltration, convective rainfall, and evaporation. Infiltration is primarily driven by the soil type, rainfall intensity, and slope. Convective rainfall is dependent on the moist static energy and stability, and evaporation is dependent on temperature (due to increased kinetic energy), vapor content in the air, and air turbulence that can transport vapor away from the surface. These multiple factors illustrate some of the complexity of projecting future malaria transmission. Even though biologic models make many simplifications, they have been important in the control of malaria. MacDonald [8] used the concept of the basic reproductive number, R_0, to show that reducing the number of mosquitoes could be used to eradicate malaria locally. Equation 1.1 shows a classical formulation of R_0. R_0 must be greater or equal to one for disease persistence.

$$R_0 = \frac{1}{r} \frac{\frac{m}{h} a^2 b c p^{EIP}}{-\log p} \tag{1.1}$$

where r is the recovery rate of humans, m is the number of mosquitoes, h is the number of humans, a is the biting rate, b is the coefficient for transmission from human to mosquito, c is the coefficient for transmission from mosquito to human, EIP is the extrinsic incubation period, and p is the daily survival rate of the mosquitoes. The equation indicates that control of malaria can be achieved by increasing the recovery rate of humans (medical treatment), reducing the number mosquitoes, increasing the number of humans, increasing the mortality of mosquitoes, or lowering the biting rate.

In this thesis, we have studied the dynamics of some mosquito species involved in malaria transmission in sub-Saharan Africa, while including external factors that influence the distribution of species. One project describes a model including *Anopheles arabiensis* and *Anopheles gambiae s.s.*. In this paper we show how additional complexity in the model can alter the distribution of the two species. For example, by including an age dimension the optimal temperature for malaria transmission is 25 $^\circ C$, $2-3^\circ C$ lower than other models [9–11]. Concerns have been raised whether warmer temperatures will lead to more malaria. This is a complex question [12], and the answer might differ depending on how scientists phrase the question. The simplest question would be: what would happen to malaria if only temperatures increase and nothing else changes? This question is simplified by assuming the species composition and the distribution of mosquitoes remains the same. This simplified type of question can be answered using many models, including those of Ruiz et al. [13], Ermert et al. [10], and Parham et al. [11]. The advantage of asking the question this way is that the uncertainty of the answer is narrowed down; however, by simplifying the question, the answer might not be relevant to the real world. A more realistic question, still restricted to only looking at temperature, would be: how would the mosquito populations respond if, in addi-

tion to temperature, human settlement patterns and cattle stocks changed over the same period? This question requires a more dynamic model with more complexity, and by including changes in precipitation and winds, even more dynamics have to be included. In the first paper the aim was to build a model to explore mosquito dynamics, and the effect of these dynamics of malaria transmission. We show how the model can be used to look at one factor at a time, but also how the model behaves when more complexity is added.

In the other part of this thesis we ask whether our complex model is able to describe the species distribution of *An. gambiae s.s.* and *An. arabiensis*. Because these two vectors have different biting preferences, with *An. gambiae s.s.* biting more humans, and *An. arabiensis* feeding on both cattle and humans, it is relevant to model the distribution and relative fraction of each of the mosquito species. If these relative fraction of each of the species change with changing climate, this can also have an influence on the efficiency of malaria transmission. Several attempts have been made to map the distribution of *Anopheles gambiae s.s.* and *Anopheles arabiensis* [1–4, 14]. These papers all used statistical methods to map the distribution, and temporal changes of distributions are ignored. As such, the predictive usefulness of these models might be limited. However, if dynamic models can reproduce present-day temporal and spatial variability, we can have more confidence in their ability to make useful future projections.

Cattle populations can influence malaria transmission [15], but the true influence of cattle and other livestock is unknown. In the first two papers we show that the distribution of *Anopheles arabiensis* is closely linked to cattle distribution, but we emphasize that the cause of this relationship is unknown. In paper two we attributed the relationship to the mosquito's cost of finding a blood meal. If cattle is as important as we showed in the first two papers, a good understanding of changes in cattle distribution and density is needed. In the third paper we estimate the African cattle density in the 1960s. We also try to explain what causes variations in national cattle holdings. Understanding how cattle are influenced by the environment is important in and of itself, but our main motivation for this paper was to use cattle densities in the mosquito model that we described in the first paper.

The relationships between mosquito survival, biting rates, the number of days required to develop sporozoites, and malaria has been described in a number of models [9–11, 13, 16, 17]. These relationships are important to estimate the impact of global warming, as well as seasonal changes in temperature, on malaria transmission. In the fourth paper we show how mosquito mortality is described by different models, and evaluate the quality of these mosquito mortality models.

1.2 Aims

The overall aim of this thesis is to contribute to the understanding of the mechanisms that influence malaria transmission. The number of the *Anopheles gambiae* complex present at any given time and place is controlled by the temperature, the number of breeding sites, the number of hosts, the humidity, and intraspecific and interspecific

competition. The number of breeding sites is controlled by rainfall; breeding sites are in some cases formed as rivers recede. Rainfall is dependent on moisture in the air, how easily the moisture condenses (which depends on particles and temperature), the vertical profile of the atmosphere, and obstacles which can rise and cool the air. Humidity also influences the survival of mosquitoes; mosquitoes experiencing saturation deficits tend to live for shorter periods of time than mosquitoes in humid environments. In many sub-Saharan African countries, rainfall also influences humans and cattle densities and distribution, which could affect malaria transmission.

Temperature influences the longevity of mosquitoes, frequency of mosquito feeding and egg development, and controls the development of malaria sporozoites inside the mosquito. In this thesis, the main focus is on understanding how mosquitoes and malaria transmission are influenced by the environment, and to describe this in a dynamic model. Specifically, we have tried to answer or discuss the following questions:

- What are the main mechanisms controlling the number of *Anopheles gambiae* s.l. at any time and place (papers I and III)?
- What happens if one factor is not included, or if the parameterization schemes are changed (papers I and IV)?
- Are current models able to describe the mosquito component of malaria models adequately (papers I, II and IV)?
- What role do presence of humans and cattle have on the density of *Anopheles gambiae* s.l. (papers I, II and III)?
- Can cattle and human density be considered to be time invariant (papers I and III)?
- To what extent is our mosquito population model able to reproduce observations on continental scales (paper II)?

A more general perspective of this thesis is to update parameterization schemes, collect new data, and formulate a mathematical model that uses the new schemes to answer the following questions:

- What is the distribution of *Anopheles arabiensis* and *Anopheles gambiae s.s.* in a given time and place (paper II)?
- Do the assumptions made in current mosquito population models match with recent data (papers I and IV)?
- Is the added complexity of an age structure model worth the effort to better represent malaria transmission (papers I, II and IV)?

1.3 Outline

This thesis consists of a synthesis followed by four papers. In the synthesis we provide background information about topics that are relevant for the thesis, but which we did not find room for in the papers. An overview of the papers is provided in Chapter 2.

Chapter 2

Introduction to the papers

The aim of the papers in this thesis is to understand the role of *Anopheles gambiae s.l.* in malaria transmission, and how population dynamics of this mosquito are influenced by the environment. Originally the focus region was Ethiopia, and due to sparse data on *Anopheles arabiensis* compared with *Anopheles gambiae s.s.*, we created a model that included both. This approach allowed us to estimate the dynamics of *Anopheles arabiensis* based on the knowledge of its sibling species. The papers are presented in the chronological order in which they were written.

The first paper is a presentation of our mosquito population model, with examples of usage and sensitivity analysis included. In the second and third papers we validate our model, where we focus on model performance when all parameterization schemes are included in the second paper. In the third paper, we explore how national cattle holdings in Africa have responded to historical short term fluctuations in the climate and relate this to malaria transmission. In the fourth paper we specifically discuss the parameterization scheme for adult mosquito mortality.

2.1 Paper I: A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis

Paper I: *Lunde et al. (2013) A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis, Malaria Journal* ***12:****28*

.

In this paper we describe an age structured mosquito model with geographic dispersion of mosquitoes. The model described in this paper follows the thinking of Ross and MacDonald, but with more focus on application in the spatial domain and time varying parameters. To describe the life history of the mosquitoes we introduce new parameterization schemes that are derived statistically from observed data. The parameterization schemes require the input of near surface air temperature, near surface relative humidity, runoff, relative soil moisture, soil temperature, near surface wind speed and

direction, rivers and water bodies, human and bovine densities, and a landmask. For idealized studies, some of these variables can be omitted. In most of the results in this paper these variables are taken from a regional climate model (The Weather Research and Forecasting Model, WRF [18]). We present several arguments why a climate model is a better option than using observed weather data: weather data in the tropics have limited spatial generalizability, future weather and climate is projected using climate models, and historical weather records for parts of Africa are rare.

We performed several sensitivity tests. In the first experiment, we show that in addition to temperature, relative humidity and mosquito body size are also important factors related to malaria transmission. With respect to body size, this is in line with several studies [19–22], and our model is able to capture some of the aspects related to higher survival among larger mosquitoes.

In the second experiment, we explore how carrying capacity can influence the distribution of two competing species, *An. arabiensis* and *An. gambiae s.s.*. Our model shows that when there are ample puddles available, *An. gambiae s.s.* can be established, while in drier conditions *An. arabiensis* will dominate. These differences can be explained based on mosquito body size. *An. arabiensis*, which is generally larger than *An. gambiae s.s.*, can produce more eggs, which is an advantage when competing over a scarce resource.

The two proceeding experiments show how the two mosquito species are linked to their human and bovine hosts. Although it is uncertain if *An. arabiensis* prefers the environment shaped by cattle, or if the survival increases due to cattle blood, the parameter describing the probability of blood feeding, $P(B)$, seems to be important for the distribution of the two mosquito species. Therefore, there is a need for an accurate representation of bovine and human densities.

Dispersion and migration of mosquitoes has been discussed for a half century [23–27]. While it seems clear that dispersion does occur, there is no evidence for long distance migration of *Anopheles gambiae s.l.*. In the last experiment we show how the presence of one mosquito species can alter the dispersion pattern of another species. This might be important in the study of long-term changes in malaria transmission.

2.2 Paper II: A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. II. Validation of species distribution and seasonal variations

Paper II: *Lunde et al. (2013) A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. II. Validation of species distribution and seasonal variations, Malaria Journal* ***12:****78*

.

In paper I, we described a dynamic mosquito model including *An. arabiensis* and *An. gambiae s.s.*. The sensitivity to different parameters was investigated; we showed

how the distribution and density of the two species changed as we modified different model parameters. In paper II, we aimed to validate the model, focusing on the ability to separate the distribution of the two species. A total of 1,940 occurrence points were collected for *An. arabiensis*, 1,813 for *An. gambiae s.s.*, and 992 for *An. gambiae* Giles. Merging these data with three published databases [4, 28, 29] resulted in 2,926 occurrence points for *An. arabiensis*, 3,009 for *An. gambiae s.s.*, and 992 for *An. gambiae* Giles.

Several statistical models have been used to map the distribution of *Anopheles gambiae s.s.* and *Anopheles arabiensis* [1–4, 14]. With time, the predictions have improved, but the distribution in Central Africa remains uncertain due to lack of data for this region. While statistical methods have focused on the probability of mosquito occurrence, dynamic models also describe the density and temporal variations. These aspects are more difficult to validate, and for that reason any estimates of the occurrence of the two mosquito species using dynamic models should be considered as a best guess based on available data. Depending on which species is responsible for malaria transmission, the efficiency of interventions can vary. For example, the number of a species resting indoors can be efficiently reduced with indoor residual spraying, but when a species is resting outdoors, reducing the number of larva might be more efficient. It can be argued that any information about malaria vectors might become useful with time. For example, in independent studies on mosquito populations where each study only provides data on local conditions, taken together the studies can inform us about changes in vector composition and spatial distribution of each species.

The results from this paper add to previous studies mapping the range of *Anopheles gambiae s.s.* and *Anopheles arabiensis*. The accuracy in the spatial distribution of our model is about the same as Sinka et al. [4], with some improvements. Our model predicts that the density of *Anopheles gambiae s.s.* will be low when the human population density is low. Currently there are few studies confirming this, because mapping malaria vectors is mostly done in areas where malaria is a problem for humans, which tends to be in areas of higher human population density.

2.3 Paper III: Cattle and climate in Africa: How climate variability has influenced national cattle holdings from 1961-2008

Paper III: *Lunde et al. (2013) Cattle and climate in Africa: How climate variability has influenced national cattle holdings from 1961-2008, PeerJ 55*

.

In paper I we described a parameter, $P(B)$, which is related to the survival of *An. gambiae s.s.* and *An. arabiensis*. For the latter this parameter is dependent on the density of cattle, and the density of *An. gambiae s.s.* is influenced by the success of *An. arabiensis*. If it is true that this parameter partially controls the survival and reproductive success of *An. arabiensis*, it is highly relevant to know how the bovine density has changed over time, and what causes variability in bovine density.

In this paper we reconstruct the cattle density for two time periods. The 1960s

estimate is based on a map from Deshler [30], while the modern day equivalent is based on Gridded Livestock of the World [31]. We show that occasionally climate variability can explain more than 40% of the variance in national cattle holdings. In dry areas increased precipitation leads to more cattle, while in wet areas more precipitation leads to reductions. Because the model described in paper I is sensitive to $P(B)$, we aimed to address the need for reliable livestock distribution and density maps. We hope other authors and/or Food and Agriculture Organization of the United Nations (FAO) will reconstruct and project cattle densities, because these estimates are important to assess the impact of climate change.

2.4 Paper IV: How malaria models relate temperature to malaria transmission

Paper IV: *Lunde et al. (2013) How malaria models relate temperature to malaria transmission, Parasites & Vectors **6:**20*

.

When I started studying malaria, weather, and climate in 2008, I thought the relationship of rising temperatures resulting in more malaria was well established. As I started building the model, the I realized this model was very different from previous studies giving optimal transmission around 25 degrees Celsius. I did not want yet another long paper explaining all the parameterization schemes, and therefore I focused on the most important parameter; mosquito mortality.

Paper IV shows how mosquito mortality models influence our perception of the relationship between temperature and malaria. To simplify the interpretation, we excluded mosquito births and only included humans as a fraction of infectious humans, which were constant throughout the model integration. On the extreme, one of six models showed transmission would be most efficient at $27.5^{\circ}C$, while one model showed transmission would peak at $20.5^{\circ}C$. Thus, these models provide very different answers to the impact of a two degree warming on malaria transmission. However, the future occurrence of malaria is dependent on many other factors, including environment, malaria control measures, and immunity.

We also do a comparison between the models using the most extensive data set available on the relationship of temperature and mosquito survival. In this comparison, the model includes an age dimension, with varying mortality with age, producing results that are more in line with the observations. After we submitted paper IV, a work supporting our findings was published in Ecology Letters [32]. Like us, they find the optimal temperature for malaria transmission is around $25^{\circ}C$.

Chapter 3

Malaria

3.1 The role of models in understanding malaria transmission: historical outline

Several papers document the effect of malaria control interventions (for example [33–36]). However, insecticide resistance has been found to be a potential problem [5, 19, 20, 37–60].

At the turn of the 20th century the work of several researchers, including Battista Grassi and Ronald Ross, resulted in the discovery that mosquitoes of the *An.* genus transmit malaria [61, 62]. Over the next 20 years Ross, and later Lotka and Waite, developed mathematical models that became central in malaria control [63–69]. In the 1950s, George MacDonald refined these models and showed that DDT could be used to interrupt malaria transmission [70]. Equation 1.1 from Chapter 1, derived by Macdonald, indicates four ways to theoretically reduce R_0:

$$R_0 = \frac{1}{r} \frac{\frac{m}{h} a^2 b c p^{EIP}}{-\log p} \tag{3.1}$$

where r is the recovery rate of humans, m is the number of mosquitoes, h is the number of humans, a is the biting rate, b is the coefficient for transmission from human to mosquito, c is the coefficient for transmission from mosquito to human, EIP is the extrinsic incubation period, and p is the daily survival rate of the mosquitoes.

The first way to reduce R_0 is to reduce the number of days that a human is infectious, $1/r$. This can be achieved by clearing parasites (in particular, gametocytes) from the human body. Currently artemisinin-based combination therapy (ACT) is recommended for the treatment of uncomplicated *P. falciparum* malaria, and intravenous or intramuscular artesunate is recommended for severe malaria [71]. The second way to reduce R_0 is to reduce the number of mosquitoes per human, $\frac{m}{h}$, which can be achieved reducing the number of mosquitoes (indoor residual spraying or larva eradication), or increasing the human population density. Our impression is that the human population aspect has been overlooked when assessing historical changes in malaria. The third way to reduce R_0 is to decrease the human biting rate, a, by limiting human-vector contact. Insect repellents, bednets and covering house openings can be used to reduce a. This factor is also temperature dependent. The forth way to reduce R_0 is to decrease the daily survival probability of the mosquito vector, p. If a mosquito receives malaria

gametocytes but dies before the extrinsic incubation period has been completed, the mosquito will not be able to cause a secondary case. Indoor residual spraying and bednets impregnated with insecticides can potentially lower the daily survival probability.

The formulation of R_0 is elegant, showing the main principles of malaria transmission and control with a few symbols. According to Paul Reiter [12], both Ross and MacDonald were aware of the limitations of the simple model: immunity to malaria, mosquito and human behavior, and seasonality were not included, and mosquito mortality was assumed to follow an exponential curve. Therefore, forecasting, projections, and the study of seasonal and year-to-year variability need different model formulations.

In the 1970s Molineaux and Dietz et al. expanded the work of MacDonald by including several immunity classes (Garik project) [72, 73]. Of particular importance, they note:

> Previously, superinfections were assumed to have no influence on the recovery. This gave a good fit to the yearly average age distribution but was unable to reproduce simultaneously the seasonal fluctuations that later were observed in the two places selected for testing.

They also noted that a reduction in the vectorial capacity was necessary to go from a hyperendemic to a mesoendemic situation, and conclude the model worked well with the particular parasite strain and human population.

In the 2000s, Yang [74, 75] developed a model which included the effects of socioeconomic conditions, immunity and temperature. His theoretical model showed that socioeconomic conditions are more important than temperature. Further, he showed interventions targeted to protect the human population to be effective in all classes of malaria endemicity, while vector control will only be effective in areas with high transmission rates.

More recent models have increased in complexity and realism, by incorporating additional parameters such as puddle formation and data from climate models [10, 11, 76–78].

3.2 Transmission and course of infection

In humans, malaria is caused by five parasites of the genus *Plasmodium*; *P. vivax*, *P. falciparum*, *P. malariae*, *P. knowlesi*, and *P. ovale*. The life cycle of these parasites depends on two hosts, where the sexual development occurs in mosquitoes of the genus *Anopeheles*, and humans are the intermediate hosts. Female anophelines require blood for survival and egg development, and one blood source is humans. Transmission to the mosquito can occur when an anopheline bites a human carrying *Plasmodium* gametocytes, the first sexual stage of parasite development. After 8–20 days, *Plasmodium* sporozoites develop in the salivary gland of the mosquito. A mosquito biting a human at this time can result in *Plasmodium* sporozoites being released into the human blood stream, entering the liver where asexual reproduction occurs (tissue schizogony), differentiating into merozoites, and infecting red blood cells. In the red blood cells asexual

reproduction continues, and some of the *Plasmodium* merozoites become male and female gametocytes. These can then be transmitted anophelies.

Immunity can decrease malaria transmission, reducing the parasite rate outside early childhood. Metselaar and van Thiel proposed the following description of endemic classes [79] (after Guerra [80]):

Class	Parasite Rate (PR)	Description
Holoendemic	If the PR is constantly over 75% in children aged 1 year	Perennial, intense transmission resulting in a considerable degree of immunity outside early childhood. Stable malaria.
Hyperendemic	If the PR in children aged 2–10 years is constantly over 50%	Areas where transmission is intense but seasonal and immunity is insufficient in all age groups.
Mesoendemic	If the PR in children aged 2–10 years is between 11–50%. It may be higher for part of the year	Typically found among rural communities in subtropical zones when wide geographic variations in transmission risk exist. Can be regarded as unstable malaria in some cases, although epidemics are less severe than in hypoendemic areas.
Hypoendemic	If the PR in children aged 2–10 years is under 10%. It may be higher for part of the year	Areas where there is little transmission and the effects upon the general population during the average year are unimportant. Can be regarded as unstable malaria.

Table 3.1: Classes of malaria endemicity

3.3 Malaria situation in Africa

Annually there are 149–274 million cases and 537,000–907,000 deaths from malaria, with the majority of the cases occurring in sub-Saharan Africa [5, 6]. Hackett noted that *"Like chess, it [malaria] is played with a few pieces, but is capable of an infinite variety of situations."* [81] In the early 1900s malaria was widespread. Gething et al. [82] estimated 58% of the landmass had endemic malaria or a risk of malaria epidemics in the 1900s. In 2007 the percentage had decreased to 30%. It should be noted that for Africa these numbers are very uncertain. Figure 3.1 shows the standard deviation over the mean *Plasmodium falciparum* rate for children between 0 and 15 years in 2007 (the same data used by Gething et al. [82]). In large parts of Africa the standard deviation has the same magnitude as the mean. Gething et al. refer to a study by Lysenko that described malaria in the 1900s. The Gething study included both *Pl. falciparum* and *Pl. vivax*; therefore, the decrease is probably not as pronounced as they describe. Gething et al. [82] also argue that the relationship between climate and malaria has broken

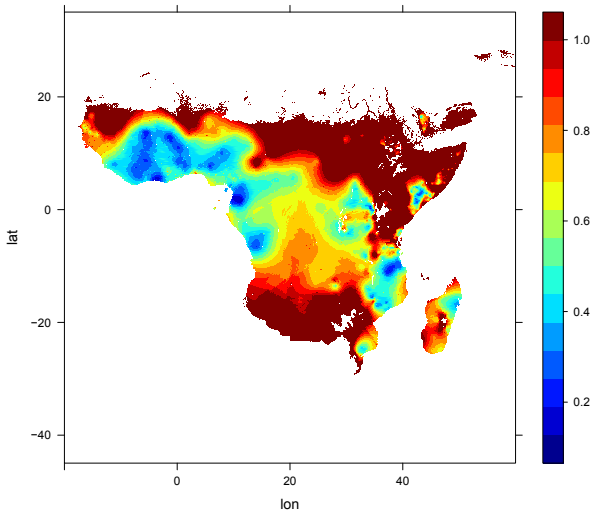

Figure 3.1: Standard deviation over mean *Plasmodium falciparum* rate for children between 0 and 5 years in 2007 plotted from 0 to 1 (0–100%).

down. This is a bold statement, because taking into account the data are not comparable and that the early 1900 data was constructed assuming a relationship between malaria and temperature. In his PhD thesis Guerra [80] describes the construction of Lysenko's map: *"These borders were refined using the theoretical distribution of malaria based on temperature requirements for sporogony of the four human malaria parasites and global isotherms."*. In light of this, it is no surprise that malaria appeared to be more strongly linked to climate in the 1900s.

To further highlight the uncertainty of the 2007 estimates for Africa, we plotted the standard deviation over the mean according to the endemicity class (Figure 3.2). In the epidemic/unstable and hypoendemic classes, the median sd/mean is greater than 100%, with the relative uncertainty dropping in the higher endemicity classes. This means that the uncertainty in areas with unstable malaria is generally larger than the uncertainty in areas with a large malaria burden.

In summary, both the historical and present-day estimates of malaria in Africa are highly uncertain, making malaria a difficult disease to study on continental scales.

3.4 Malaria control

Three main methods are used to control malaria today: insecticide-treated (ITNs) and untreated bed nets, indoor residual spraying (IRS), and Artemisinin-based Combination Therapy (ACT). Each of those methods has been reviewed in Cochrane Database Systematic Reviews.

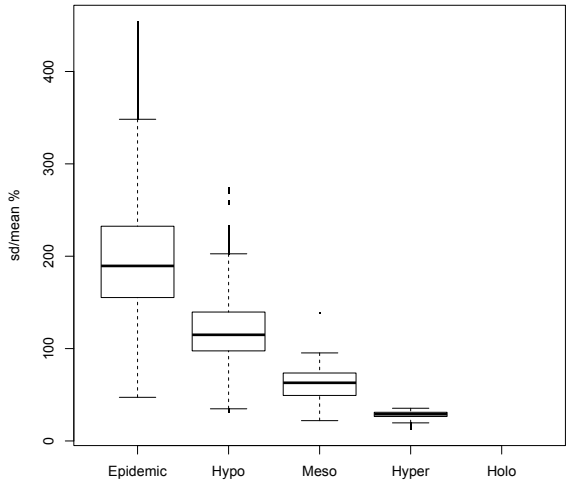

Figure 3.2: Standard deviation over mean *Plasmodium falciparum* rate for children between 0 and 5 years in 2007 according to the endemicity class).

3.4.1 Insecticide-treated bed nets

An initial study by Lengeler concluded that ITNs provide about 17% [95can be avoided by the use of an intervention [34]. Lengeler found that the incidence of uncomplicated malaria was reduced by 50% in areas with stable malaria, and reduced by 62% in areas of unstable malaria. A review by Eisele et al. later confirmed these results [83]. The review by Lengeler was based on 22 studies, where *An. gambiae s.l.* was one of the ten vectors studied, and the main vector in seven of the 22 studies. In two of these seven studies, *An. gambiae s.s.* was probably the only vector involved in malaria transmission, leaving five studies where both *An. gambiae s.s.* and *An. arabiensis* were potentially involved. Although several studies have been published since 2004, the latest review by Eisele et al. [83] is not able to address the effectiveness of bed nets given a specific vector; Bayoh et al. showed a rise in the fraction of *An. arabiensis* after introduction of ITNs [84], and in 2012 Loha et al. published a study on malaria in Chano, Ethiopia, where they showed ITNs gave individual protection, but not community protection. *An. arabiensis* was the main vector in the study area.

3.4.2 Indoor residual spraying

In 2010 Pluess et al. evaluated six studies of the effect of indoor residual spraying (IRS) on malaria. They found evidence that IRS reduced malaria in unstable malaria settings, but there are too few studies to make a conclusion about areas with stable malaria transmission [85]. *An. arabiensis* was studied in four of the publications and *An. gambiae s.s.* was studied in two of the publications. The only study where *An. arabiensis* was the sole vector was also the only study in a high transmission area. In this study, IRS had a positive protective effect (PE) on parasite incidence among children 1–5 years (14%), but no PE on older individuals (-2%). In general, IRS has been shown to be not as effective as ITNs. Insecticide resistance can reduce the effect

of both ITN and IRS [5, 19, 37–60].

3.4.3 Artemisinin-based combination therapy

Artemisinin-based combination (ACT) therapy does not (as far as we know) influence the biting rate or the survival of the vector. The most comprehensive review of this intervention included fifty studies. This review compared five ACTs, and found all of them to be effective. Because ACT is not the scope of this paper, nor the competence of the author, this intervention will not be described further.

3.4.4 Larvicides and draining breeding sites

The rate at which malaria spreads can be limited by reducing the density of vectors, which can be achieved by killing adult mosquitoes and/or targeting juveniles. While IRS mainly targets adult mosquitoes resting indoors, larval source management (LSM) could reduce the number of vectors feeding outdoors as well. Few studies have investigated interventions targeting juveniles alone, making it difficult to assess the impact of LSM. However, it is worth remembering that before MacDonald showed reducing the life span of adult mosquitoes was the most efficient way to reduce malaria, LSM was widely used, and with great success [86]. It has been argued LSM is only feasible where breeding sites are well defined. Fillinger and Lindsay argue that tire tracks, hoof prints, etc are of less importance compared to ponds near lakes and rivers. Therefore, LSM can be a viable approach in many situations [86]. The cost per person per year of LSM is about the same as IRS and ITNs.

An alternative to using human labor to treat breeding sites is to use adult mosquitoes to transfer a potent larvicide between resting and oviposition sites [87]. This methodology has several advantages: only sites used for oviposition are exposed to pesticides, and the more popular a site is, the more transfer events occur.

3.5 Vectors of malaria in Africa

3.5.1 The dominant vectors of human malaria in Africa

Sinka et al. [4] describe seven important malaria vectors in Africa, excluding *Anopheles pharoensis*. *Anopheles gambiae s.s.* and *An. arabiensis* are considered to be the most efficient vectors, although *An. funestus*, *An. nili* and *An. moucheti* are important in some regions [88–90]. *An. merus* is found on the coast of Eastern Africa, and is mainly associated with saline breeding places. *An. melas* is also associated with saline water, but is restricted to the western parts of West Africa. In high numbers these mosquitoes can be important vectors [88]. *Anopheles pharoensis* is the only African vector that is believed to migrate over longer distances [23].

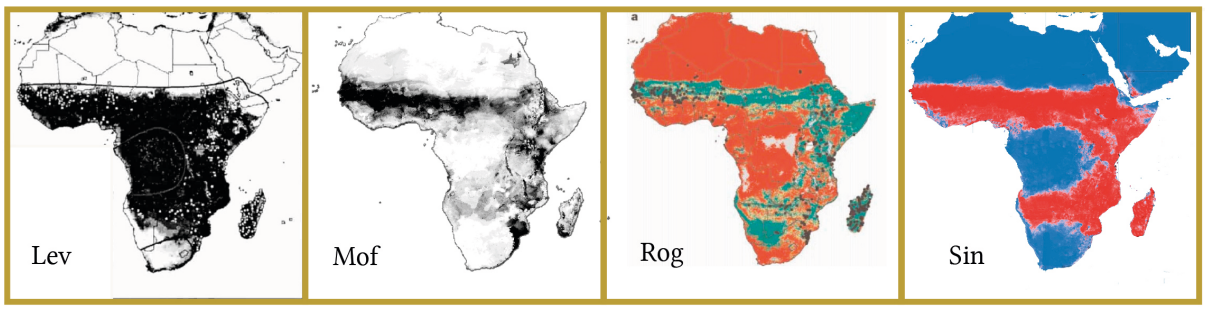

Figure 3.3: Predictions of the presence of *An. arabiensis*. Lev [1], Mof [2], Rog [3], and Sin [4].

3.5.2 Distribution of the Anopheles gambiae complex

The evolutionary relationship between the members of the *Anopheles gambiae* complex is poorly understood [91]. Understanding how species evolve is important to understand the life history of mosquitoes, which can be used in mosquito modeling and projections. While the coarse distribution of the members of the Anopheles gambiae complex is relatively well described [1–4, 14], there is still a lack of data on what causes the different distributions. Over the years the accuracy of predictive models has improved, aided by new studies on the occurrence of the species. Biophysical models attempt to explain the success of a species by taking a reverse-engineering approach; therefore, they can be valuable to test hypotheses and understand the life history of mosquitoes. Figure 3.3 shows the predicted distribution of *An. arabiensis* from four different models.

Chapter 4

Potential biases

4.1 Definition of bias

We define the final bias as the systematic difference between the observed data and modeled results. In a model, the potential known sources of the final bias will grow with the number of parameters in the model. Thus, a complex model will have more known bias, while a simple model will have more hidden bias related to the parameters which are not included. In this thesis we face three main types of bias. First, in the mathematical models and parameterization schemes we have bias due to the difference between an estimator's expectations and the true value of the parameter being estimated. Second, we have the omitted-variable bias, which arises when we omit a variable that should have been included in the model. Third, we have observational bias, which is related to the data used to formulate the conceptual framework of the model, the underlying data used to derive the parameterization schemes, and the data used for validation.

4.2 Bias related to the formulation of equations

All models have biases. The simplest model is the correlation between two variables, where our interpretation decides if variable x is influencing variable y, or variable y is influencing variable x, or the two variables by coincidence vary in the same pace. In a dynamic model, we write relationships as mathematical equations. An example could be the development from pupa to an adult mosquito. In real life this metamorphosis could be described by delay differential equations (equation 4.1), but for practical purposes (eg. numerical stability) they are often approximated and written as ordinary differential equations (ODEs, equation 4.2).

$$\begin{aligned} \frac{dP(t)}{dt} &= -P(t-\tau) \\ \frac{dA(t)}{dt} &= P(t-\tau) \end{aligned} \tag{4.1}$$

where τ is the number of days required to develop from pupa, P, to adult, A.

$$\begin{aligned} \frac{dP}{dt} &= -P \cdot r \\ \frac{dA}{dt} &= P \cdot r \end{aligned} \tag{4.2}$$

where A is the number of adults, P is the number of pupa, and r is the development rate from pupa to adult.

By deciding to use ODEs, we have introduced the first error into our model. For example, ODEs are capable of producing half a pupa and half a mosquito at a given time, and over time pupa converge towards zero. Consider a specific example. We start with two pupa, $P = 2$, and zero adults $A = 0$, neglecting mortality. Development from pupa to adult takes two days. The exact solution of this problem would be
$P(t = 0) = 2, P(t = 1) = 2, P(t = 2) = 0$ and
$A(t = 0) = 0, A(t = 1) = 0, A(t = 2) = 2$.
In the framework of ODEs the value of r would decide how fast development occurs. One method is to define rate as per day, day^{-1}. In this case $r = 1/2$. The exact solution using ODEs then becomes
$P(t = 0) = 2 \cdot e^{-r \cdot t} = 2.00, P(t = 1) = 1.21, P(t = 2) = 0.74$, and
$A(t = 0) = 2 - 2 \cdot e^{-r \cdot t} = 0, A(t = 1) = 0.79, A(t = 2) = 1.26$.

Another way to define the development rate, r, is to consider the fraction that have developed into adults at time t. Let us say 50% of the pupae developed into adults at the second day ($t = 1$), we could find an exact solution which satisfies this condition:

$$\begin{aligned} P(t = 1) &= P(t = 0) \cdot e^{-r \cdot 1} = P(t = 0) \cdot 0.5 \\ e^{-r} &= 0.5 \\ r &= -log(0.5) \end{aligned} \tag{4.3}$$

This approach then defines the development rate as the time it takes for 50% of the mosquitoes to develop from pupa to adult, $d(t)$, or more generally, $r = -log(0.5) \cdot d(t)^{-1}$.

4.3 Bias related parameterization schemes

In paper I, we define several parameterization schemes that describe factors such as mosquito mortality rates and mosquito development rates. These schemes are derived statistically from observed data (described in paper I). Statistical models have an error term that represents the unexplained variance, and this error is inherited but not accounted for in our biophysical model since only the estimate is used, –unless there are several realizations of the same model or randomness is assigned to a given parameter. The data that we used in the statistical model also have errors related to observations, data recording, and precision of the observations. One may wonder why we do not base the parameterization schemes on directly observed data. Doing so would remove the error term related to the statistical model, but would also mean we could only use the biophysical model at parameters (for example, temperature or humidity) where we have observed data. Therefore, the role of the statistically derived parameterization schemes is to extrapolate the observed data into temperature and humidity ranges where there are no data. In the case of the model described in paper I, this means we have more confidence in temperature ranges from $\approx 15 - 30°C$ and relative humidity over 40%, because these are ranges for which we have observed data. The parameterization of breeding sites is highly uncertain, because of lack of data for validation.

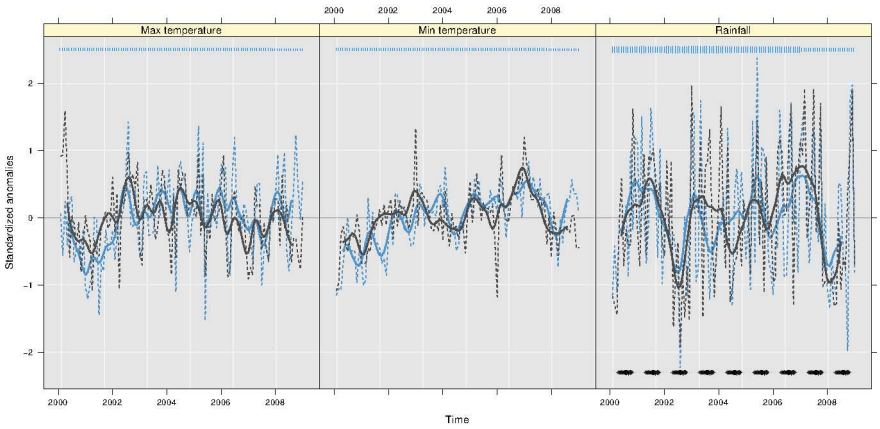

Figure 4.1: Observed and modeled anomalies of minimum and maximum temperature, and precipitation.

4.4 Biases in the driving data

Near surface temperatures, relative humidity, soil moisture, soil temperature, runoff, and other parameters included in the model can be taken from any source. In papers I and II we demonstrated how the model can be driven with data from a limited area model (Weather Research and Forecasting, WRF). To be confident that the models were able to reproduce the observed variability of precipitation and temperature, we constantly evaluated the biases of WRF. Because the original aim was to create a model that performed well for Ethiopia, we were concerned about reproducing temporal variability of observations in this region. Figure 4.1 shows the observed station anomalies and modeled anomalies for minimum temperature, maximum temperature, and precipitation. Weather stations that were used for validation are shown in Figure 4.2. To calculate the anomalies we picked the model point, m_i, which were closest to station, s_i. If station s_i had missing data, these were also removed from the corresponding model point, m_i. Standardized monthly anomalies were calculated for each station and model point (omitting missing data), and the area averaged anomalies according to time are shown in Figure 4.1. The model is able to capture the main tendencies in temperature and rainfall. However, it is apparent in Figure 4.1 that the use of model data will introduce additional biases, because the driving data will pass on errors. To keep the variables that feed into the mosquito model consistent, we did not bias correct the driving data.

In addition to Ethopia, we also simulated precipitation for the African continent as a whole. For this simulation, there are some biases in the model worth mentioning. East Africa, including Sudan, South-Sudan, Ethiopia, and Kenya are too wet compared to observations with the Kain-Fritsch [92] convective scheme, and West Africa is too dry. WRF includes a number of parameterization schemes, which approximate physical processes that cannot be resolved at coarser resolution. In the case of the 50 km simulations used in paper I and II we parameterized convection, microphysics, turbulence, gravity wave drag, and boundary layer height. In one particular study, we assessed the performance of the WRF model in simulating precipitation in the tropical region. The simulations were performed to test the combination of different microphysical and cu-

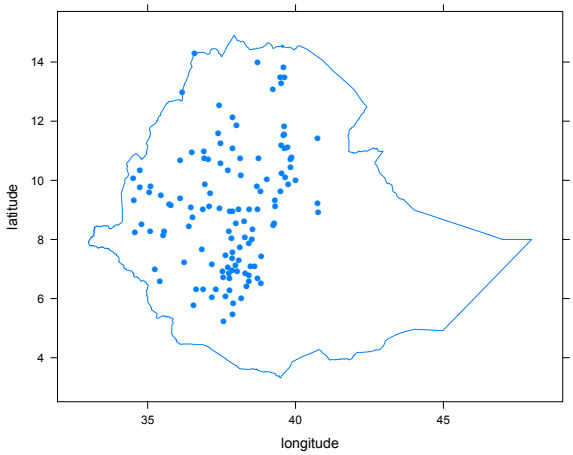

Figure 4.2: Stations used to validate regional climate simulations over Ethiopia

mulus parameterization schemes. Figure 4.3 shows the average precipitation from June to September 1990, while Figure 4.4 shows the difference between the WRF simulation and the observed data. Although not included in this thesis, we also found that the easterly waves were associated with West African precipitation. Our results showed different patterns of precipitation over Sahelian Africa, depending on the schemes chosen. These changes seemed to be associated with differences in the representation of the westward propagating waves ('easterly waves') in the model. Easterly waves are climatologically important because they may significantly alter the amount of rainfall in a region. We used the following microphysical schemes: MP3 [Hong, Dudhia and Chen (2004, MWR)]; MP6 [Hong and Lim (2006, JKMS)] and MP9 [Milbrandt and Yau (2005, JAS)], and two cumulus parameterization schemes: CP1 [Kain-Fritsch scheme] and CP2 [Betts-Miller-Janjic scheme].

4.4.1 Reanalysis data

A regional climate model is dependent on initial conditions and lateral boundary conditions from a model which covers a larger area than the domain in the regional climate model. For the simulations presented in paper I and II, we used data from NCEP [93] and ERA-interim [94]. NCEP is produced at a resolution of T42, while ERA-interim is at T255. Over time the observations going into the assimilation system have varied. Examples of data that can be assimilated are radiosonde data, satellite observed radiance (for example, AMSU), and satellite observed vapor (for example, Meteosat). An example of a change in the observational system is the introduction of Meteosat-8 over Africa in 2004. Because regional climate simulations use assimilated data, the inconsistencies in the assimilated data have consequences for the regional climate simulations. If, for example, the introduction of Meteosat-8 resulted in more moisture over the African region being measured, in the simulations this could lead to years after 2004 being wetter than the earlier periods. When a regional climate model is driven by

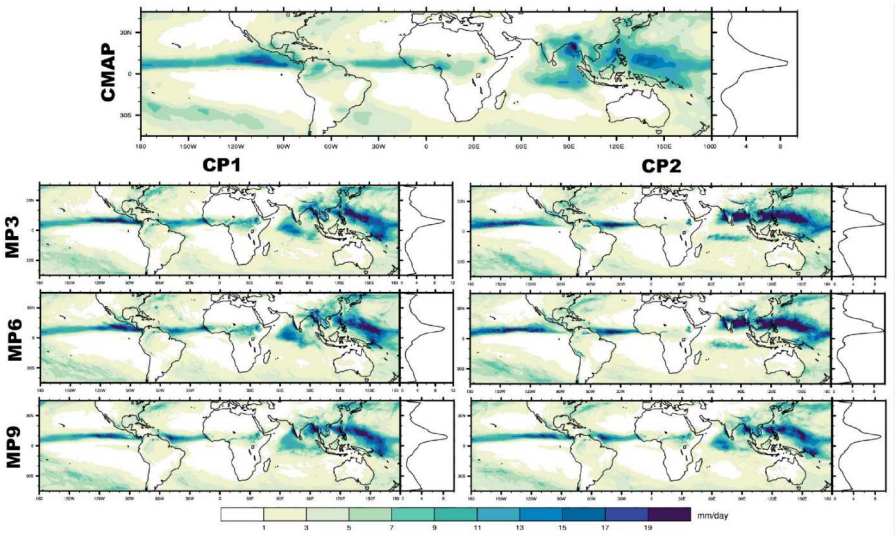

Figure 4.3: Average precipitation from June to September 1990. The plots indicate results from runs using different combinations of microphysical schemes (MP) and cumulus parameterization schemes (CP) in WRF. The CPC Merged Analysis of Precipitation (CMAP) precipitation is shown for comparison.

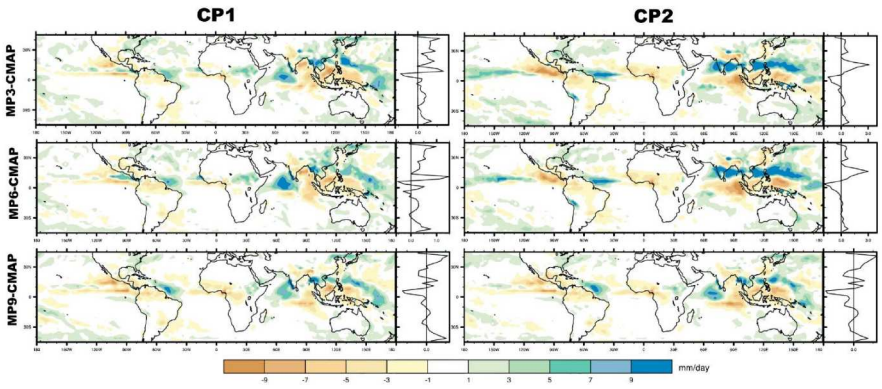

Figure 4.4: Same as Figure 4.3, but for the difference between the WRF simulation and the observations. The model data resolution was interpolated to that of CPC Merged Analysis of Precipitation (CMAP) for this difference plot.

global climate models this problem vanishes, because one of the main purposes of the global climate models is to balance the energy budget.

4.4.2 Human and bovine densities

In paper I we describe how the parameter $P(B)$, the probability of finding blood, is related to human and bovine density. Several data sets exist for human densities, including Gridded Population of the World (GPW) [95], AfriPop [96], UNEP/GRID - Sioux Falls Dataset [97], and LandScan Global Population Database [98]. Data sets for bovine densities include raw census data for 2002 [31] and the statistically derived (based on environmental variables) Gridded Livestock of the World [99].

The data sets that rely only on observations have coarse resolution, and have errors where the observational densities are low. The statistically derived data sets include an error term from the statistical model which are not reflected in the point estimate. These errors will be passed on to our mosquito population model described in paper I. We showed that the model is sensitive to the $P(B)$ parameter; therefore, the mosquito model error should be reduced with perfect observations of human and cattle densities. We try to address this issue in paper III.

4.4.3 Validation and calibration data

In papers I, II and IV we use data related to anophelines to develop, calibrate, and validate our mosquito population model. In the first paper we aim to develop parameterization schemes which describe the different aspects of the mosquito's life (calibration). In general we have only parameterized processes where there are available data, and most of the studies are based on laboratory experiments. The advantage of such studies is that the environment of the organisms can be controlled, making it possible to study a single process in the life history of the mosquitoes. The data we used to parameterize mortality and development rates in the aquatic stages were conducted under constant temperatures. This is very rare in nature, but laboratory studies with constant temperatures are required to isolate the effect of temperature or humidity and remove confounding variables. The weakest parameterization schemes in most malaria models are those related to puddle formation. Our scheme for puddle formation is somewhat different from the one used in the Liverpool Malaria Model [10], which only considered rainfall. In paper I we assume puddle formation is a function of relative soil moisture, or how close the soil is to saturation, with additional information on how likely it is the water would drain to a certain area (potential river length). The carrying capacity is based on data from Stephen Munga (in Kenya), where they counted the number of larvae in the dry and rainy season. From the same study area we estimated the potential river length and relative soil moisture in the two study periods, and approximated the carrying capacity. Ideally we would have had observations from several locations across Africa, but during model development we were not able to find such data.

For the validation in paper II we collected data on the occurrence of *An. arabiensis* and *An. gambiae s.s.* from available studies. The majority of the studies did not distinguish between the M and S form of *An. gambiae s.s.*, and we did not attempt to draw such distinctions. Some studies only provided maps of the observations, and in

these cases errors could be introduced during geo-referencing. We also discovered that several papers reported incorrect coordinates, and there is a chance we did not pick up all the inconsistencies in the reporting. To determine the distribution of *An. gambiae s.l.* in the DR Congo we also included observations of *An. gambiae* Giles, and classified those based on the expert opinion given in Sinka et al. [4]. These observations are less robust than newer observations where mosquitoes were classified using polymerase chain reaction, but might be valuable in areas where there are few observations.

In paper IV we use the same data to calibrate and validate the model. This approach means models which used these data to develop survival curves will have an advantage over models which used other data. Because three of the parameterization schemes in this paper used the validation data to calibrate the model, our model does not have an advantage over these three schemes. In the experiments in this paper we used constant temperatures. In nature mosquitoes experience temperature fluctuations, but because the curves presented in the paper are almost symmetrical, temperature fluctuations would probably have little effect on the optimal mean temperature [32].

Another aspect that might be poorly represented in our model is the dispersion of mosquitoes. First, the dispersion distance per day is uncertain. Some studies have shown the M form of *An. gambiae s.s.* fly downwind, while the S form of *An. gambiae s.s.* fly upwind, or that the behavior changes in the dry season compared to the wet season [100]. Second, the way we included mosquito dispersion into our model means that fractions of mosquitoes can disperse. We plan to implement a threshold such that the flux of mosquitoes over a boundary must be greater than one.

4.5 External validity

External validity refers to whether the model can be generalized to other areas or situations. As we mention in paper II, the different parameterization schemes have different validity. For example, the scheme for adult mortality specific for *An. arabiensis* is described as a scheme that is only suitable in cases where *An. arabiensis* is the only species. In paper I, we describe a puddle parameterization scheme that we believe will break down as the spatial resolution increases. Paper III describes how national cattle holdings have been influenced by climate variability. It is highly uncertain whether these findings can be extrapolated into the future. For each of the schemes in paper we have tried to highlight the assumed validity. In paper IV we use constant temperatures. Because constant temperatures are rare in nature, it should be explored if temperature variability give the same results. The influence of climate change on malaria transmission is still uncertain. The future occurrence of malaria is dependent on many other factors not currently included in our model, including malaria control measures, access to and usage of treatment, city planning, and immunity. Thus, to increase external validity, future models should attempt to accurately include these factors.

Chapter 5

Conclusions and perspectives

The main objective of this thesis was to develop a weather-driven malaria transmission model. Because weather primarily influences the vector and the parasite, the presentation of the model is related to those two components. To validate the model, knowing the true distribution of malaria vectors is important. As new papers on the life history of *Anopheles gambiae s.l.* appear in the literature, these data should be incorporated into models, to add value to the predictions of malaria transmission.

The majority of this thesis relates to how we describe the life history of *An. gambiae s.s.* and how it is different from *An. arabiensis*, but also how our model is different from other models. Whereas previous studies have neglected the age dependent life history of mosquitoes, survival related to the probability of finding blood, and mosquito size, we have incorporated these parameters into the model described in paper I. In papers II and IV we do the first tests to see if our approach adds value compared to other models, and there are indications that our approach provides comparable or better results than existing models (paper II and IV). As with any model, there are caveats in our model and its validation. We chose to drive the model with input data from WRF, and the results depend on the quality of this driving data. In addition, our validation of the temporal variability is only a collection of selected case studies, meaning the true capability of our model to represent variability in time has not been properly tested. In the bias section of this thesis we discuss how the choice of using ordinary differential equations (ODEs) introduces a bias. The main reason for using ODEs is the stability of the model, because we experienced numerical instabilities when using delay differential equations.

One of the aims of this thesis was to develop a model that can be generalized for sub-Saharan Africa. The results so far support that the model is able to separate the distribution of the two mosquito species, which makes us more confident that we have represented important life history characteristics. As we have discussed in earlier chapters, the parameter $P(B)$ seems important to describe the success of the two species, but we do not yet understand the true meaning of this parameter, or whether it is only a confounding factor.

When MacDonald developed his model of malaria transmission, he realized the representation of immunity was too simple. More complex models have since been

introduced to improve the representation of immunity and super infections [74, 101–103], but there has been less emphasis on describing the mosquitoes more accurately. The simple models have proved valuable to understand how the malaria burden can be reduced [8], but they are not suitable to accurately describe temporal and spatial changes in malaria burden. In paper IV we show with a simple example how the optimum temperature for malaria transmission can dramatically change our perception of the expected influences of climate change on malaria. The simple models were designed to understand the dynamics of malaria transmission, but more accurate models are needed for forecasting and projection applications. We believe the model presented in paper I can be simplified and parameterization schemes can be improved, but the presentation of a complex model can lead to further investigation of different parameters (like mosquito size). With time, some ideas presented in our model may be shown to be false, while other ideas might prove to be valuable. The type of studies presented in paper IV are useful to understand how single components of the model influence the dynamics of malaria transmission.

There is one important aspect related to future malaria projections that we do not cover at all: adaptation. Specifically, we did not discuss if mosquitoes are able to adapt to the environment, as long as there is water and blood available, or if the parasites are able to adapt as long as the mosquito density and transmission is high enough. The recent separation of the M- and S-form of *An. gambiae s.s.* suggest the vectors are able to adapt to new environments [104], in which case future projections of anophelines have limited value. Still, we do not know how fast the mosquitoes and parasites evolve, or if evolution occurs in jumps or is a gradual process. With these unknown factors, we have few other options than using the latest knowledge about the mosquitoes and parasites in dynamic or statistical models.

In these papers we have only briefly touched on how this mosquito model can be linked to a full malaria model. Coupling the mosquito model to a human component is rather simple, and we have a working version where we introduce a simple SEIR (susceptible, exposed, infectious, recovered) human component. In addition to only wanting to describe the weather sensitive component of malaria transmission, the sole focus on some of the vectors follows a reductionist thinking. With a full model, it would be harder to evaluate the model output. We believe a limited focus enables us to point at strengths and weaknesses in our model in a more rational way compared to a more complex model including humans. Human immunity against malaria is a separate study in and of itself, and we hope to properly represent this process in a future version of our model.

We also neglected the effect of interventions in our mosquito model, because the time and space invariant impacts of interventions are poorly described. For example, an accurate number describing the impact of ITNs is not available in the literature. Including IRS and ITNs interventions into our model would also challenge us by requiring equations for the development of insecticide resistance, which is thought to have been important in the reemergence of malaria in South Africa in the early 2000s. A natural next step is to collect data on these intervention parameters, to include the effects of interventions in our model.

It is important to remember that our mosquito model, and other models of the same type, are only suitable for scenarios; if temperature increases/decreases, and population growth is 5% per year, and the number of breeding sites is constant, we will get more/less malaria in the 2060s. For instance, to explore the consequences of mosquito populations and malaria transmission given warmer temperatures or rainfall anomalies, given the type of housing remains the same and 60% of the population uses bed nets every night. These models project what will happen under certain conditions, and are not meant to predict what will actually happen in the future.

Chapter 6

Scientific results

Paper I

6.1 A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis

Torleif Markussen Lunde, Diriba Korecha, Eskindir Loha, Asgeir Sorteberg and Bernt Lindtjørn

Malaria Journal 2013, **12**:28

Malaria Journal

BioMed Central
The Open Access Publisher

A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis

Malaria Journal 2013, **12**:28 doi:10.1186/1475-2875-12-28

Torleif Markussen Lunde (torleif.lunde@cih.uib.no)
Diriba Korecha (dkorecha@yahoo.com)
Eskindir Loha (eskindir_loha@yahoo.com)
Asgeir Sorteberg (asgeir.sorteberg@gfi.uib.no)
Bernt Lindtjørn (bernt.lindtjorn@cih.uib.no)

ISSN 1475-2875

Article type Research

Submission date 21 September 2012

Acceptance date 7 January 2013

Publication date 23 January 2013

Article URL http://www.malariajournal.com/content/12/1/28

A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis

Torleif Markussen Lunde[1,2,5*]
*Corresponding author
Email: torleif.lunde@cih.uib.no

Diriba Korecha[4,5]
Email: dkorecha@yahoo.com

Eskindir Loha[3]
Email: eskindir_loha@yahoo.com

Asgeir Sorteberg[2,5]
Email: asgeir.sorteberg@gfi.uib.no

Bernt Lindtjørn[1]
Email: bernt.lindtjorn@cih.uib.no

[1]Centre for International Health, University of Bergen, Bergen, Norway

[2]Bjerknes Centre for Climate Research, University of Bergen/Uni Research, Bergen, Norway

[3]Hawassa University, Hawassa, Ethiopia

[4]National Meteorological Agency of Ethiopia, Addis Ababa, Ethiopia

[5]Geophysical Institute, University of Bergen, Bergen, Norway

Abstract

Background

Most of the current biophysical models designed to address the large-scale distribution of malaria assume that transmission of the disease is independent of the vector involved. Another common assumption in these type of model is that the mortality rate of mosquitoes is constant over their life span and that their dispersion is negligible. Mosquito models are important in the prediction of malaria and hence there is a need for a realistic representation of the vectors involved.

Results

We construct a biophysical model including two competing species, *Anopheles gambiae s.s.* and *Anopheles arabiensis*. Sensitivity analysis highlight the importance of relative humidity and mosquito size, the initial conditions and dispersion, and a rarely used parameter, the probability of finding blood. We also show that the assumption of exponential mortality of adult mosquitoes does not match the observed data, and suggest that an age dimension can overcome this problem.

Conclusions

This study highlights some of the assumptions commonly used when constructing mosquito-malaria models and presents a realistic model of *An. gambiae s.s.* and *An. arabiensis* and their interaction. This new mosquito model, OMaWa, can improve our understanding of the dynamics of these vectors, which in turn can be used to understand the dynamics of malaria.

Keywords

Anopheles gambiae complex, Model, Malaria

Background

This is the first of two papers describing a dynamic model (Open Malaria Warning; OMaWa) of *Anopheles arabiensis* and *Anopheles gambiae s.s.* Our aims in this article are 1) to formulate recent research on the *Anopheles gambiae* complex in a mathematical framework, and 2) to show how the new formulations influence the dynamics of malaria and mosquito populations.

In this paper, we describe a model of the dynamics of the two species and then show how parameters can influence the success of the two species, and how temperature, humidity and mosquito size can influence malaria transmission.

Climate and malaria

Most of the 149-274 million cases and 537,000-907,000 deaths from malaria occur in sub-Saharan Africa [1, 2]. Climate has been one of the main drivers of this disease [3], governing the spatial extent and year-to-year variations. The pathway from climate to malaria goes through the parasite and the mosquito. Although it is well established [4] how parasite development is influenced by temperature [5], the vector's response to weather and climate is more complex. Mosquito density depends not only on temperature but also on the abundance of breeding sites (rainfall and evaporation) [6], desiccation (humidity) [7], and competition between mosquitoes [8]. In the past 20 years, a shift in the distribution of *An. arabiensis* and *An. gambiae s.s.* has been observed in Kenya [9], showing that the species composition is not static over time. In the context of climate change [10], variability in vector populations is a factor that has not been considered so far.

Malaria and mosquito models

At the turn of the 20th century the work of several researchers, including Battista Grassi and Ronald Ross, resulted in the discovery that mosquitoes of the *Anopheles* genus transmit malaria [11, 12]. Over the next 20 years, Ross, and later Lotka and Waite, developed mathematical models that became central in malaria control [13–19]. In the 1950s, George MacDonald refined these models and showed that DDT could be used to interrupt malaria transmission [20]. Since then, several modelers have followed in the footprints of Ross, Lotka, and MacDonald [21–30]. Some have designed models to show how temperature alone influences malaria transmission [31], while others have focused on the theoretical effect of bed nets [32], multiple interventions [33] or climate change [34–36]. There is also a growing

number of models that address the dynamics of immunity within individuals [37] and in communities [21, 38].

In 2011, The malERA Consultative Group on Modeling [39] provided a review of the current state of mathematical models and pointed to the importance of good mosquito models for assessing the impact of climate change on malaria.

Many traditional models rely on a threshold principle. The idea has been to find thresholds for longevity, number of bites or days to recovery that must be reduced to interrupt the transmission. With increased computational power it is now possible to make more complex models and hence explore a wider range for the dynamics of malaria and mosquito survival. By integrating the knowledge from simpler models into a complex system, it is possible to test if the assumptions are true over a wider geographical range. In addition, these complex models can make quantitative predictions about strategies for control [40].

Model summary and motivation

A model is mental copy that describes one possible representation of a system. We present an alternative formulation of the dynamics of *An. gambiae s.s.* and *An. arabiensis*. The model is a system of ordinary differential equations (ODEs) with three compartments: eggs, first to fourth instar larvae, and pupae; an age-structured formulation of adult mosquitoes; and size prediction for adult mosquitoes (measured as wing length in mm). This can be considered the skeleton of the model. As demonstrated later, the model structure can be simplified when mosquito size can be neglected or when we assume no births. The model can be run with a spatial structure in which we include or exclude mosquito dispersion, or as an idealized model in which the model is evaluated at a single point.
The ODEs parametrize daily mortality rates, which are size-dependent for adult mosquitoes; development rates in the aquatic stages; biting rates; fecundity; the probability of finding a blood meal; and mortality related to flushing of eggs, larva and pupa out of oviposition sites. These parametrization schemes are driven by air temperature, relative humidity, relative soil moisture, water temperature, and runoff. As already mentioned, the model can be applied in a spatial domain. In this case, temperature and other environmental data are taken from a regional climate model, the Weather Research and Forecasting Model (WRF) [41]. In the examples shown later, we run the model at a resolution of approximately 50 km and a temporal resolution of 5-20 years in steps of 3 h. In addition to weather data, human [42] and cattle [43, 44] densities are introduced to estimate the probability of feeding.

At this spatial resolution, the model should potentially be able to define larger foci of mosquito productivity, while the ability to identify hotspots will be limited [45]. However, 50 km is the standard for regional climate models addressing long-term changes in climate [46]. In addition, the true accuracy of historical cattle and human population density estimates for Africa in general is not likely to be greater than 50 km.

The mosquito model described here is designed to capture the spatial distribution and the time-dependent density of *An. gambiae s.s.* and *An. arabiensis*. If the model can capture the current distribution and density of the two species and how they are related to malaria, a future version of this model, including infections, could be used to explore the long-term impact of current interventions under a changing climate. To have confidence that the model has these abilities, several aspects not considered here should be evaluated (papers under preparation). In addition, if malaria modelers move towards the ensemble thinking widely adopted in the climate community, this model could be one representation of historical and future changes for malaria. The aim of such an ensemble would be to deal with uncertainties in the system. Ultimately, the goal would be to produce policy-relevant information including uncertainty.

We have chosen to represent the non-exponential mortality of *An. gambiae s.s.* and *An. arabiensis* as observed in laboratory settings [47], semi-field conditions [48], and in the field [49]. A common assumption is that in the field, mortality rates are constant with age because of predation [31]. To date, few studies have confirmed this, while there is field-based evidence of age-dependent *Ae. aegypti* mortality [49], which has implications for malaria transmission [50]. In the model, we also describe how mosquito size changes over the season. This might seem to be an overcomplication of the model. The motivation, however, is that we have observed substantial improvements for arid regions such the Sahel when we included mosquito size prediction. Fouet et al. reported that mosquito size is an important adaptation strategy in arid environments [51].

We do not claim that the additional complexity adds any value. Stating this before the model has been fully evaluated and compared to simpler models would be dangerous. The model is thus one possible way of describing the dynamics of *An. gambiae s.s.* and *An. arabiensis*. It is under continuous development, and we expect to add and alter components as new data become available.

To highlight some of the components that contribute to the dynamics of *An. gambiae s.s.* and *An. arabiensis* in the model, five sensitivity experiments focus on the effect of temperature, relative humidity and mosquito size on malaria transmission. We also show how *An. gambiae s.s.* and *An. arabiensis* respond to changes in the probability of finding blood, carrying capacity, initial conditions, and dispersion.

Material and methods: model description

Summary of the model

Figure 1 provides an overview of the model. In the following sections we present the ideas behind the model and its general structure, how a climate model is used to drive the mosquito model, and the parametrization schemes used in the model. It should be possible to read each part independently; for example, data from a climate model can be used to drive any malaria model; the parametrization scheme can be used in any malaria model; and the malaria model described here can be used with different parametrization schemes, with or without data from a climate model.

Figure 1 Overview of the mosquito model. A (regional) climate model is used to force the mosquito model. In addition, static and semi-static fields are used as part of the parametrization schemes. Human and bovine densities limit the availability of blood meals.

As mentioned above, the model comprises a system of ODEs for eggs, first to fourth instar larvae, and pupae; an age-structured formulation for adult mosquitoes; and size prediction for adult mosquitoes (measured as wing length in mm). The first limitation in the aquatic stage is the availability of ovipositing sites, which is parametrized in terms of relative soil moisture and the potential for puddle formation in a specific location. Once ovipositing sites have been formed, adult female mosquitoes are allowed to deposit eggs until the site is full, defined as the biomass relative to the carrying capacity for the location. To account for density-dependent mortality, first instar larvae can be preyed on by fourth instar larvae [52], and an extra density-dependent mortality term is added to account for prey-independent mortality [53]. The numbers of eggs, larvae and pupae are reduced when the precipitation rate exceeds the infiltration rate. The larval density in the aquatic habitat influences the size of adult mosquitoes [53]. We account for this by predicting mosquito size at emergence as a function of larval density. In addition to temperature and relative humidity [47], mosquito size influences the daily adult survival probabil-

ity [7,51,54,55] ([56], *Aedes aegypti*). We therefore describe an adult survival model that takes temperature, relative humidity and mosquito size into consideration. In addition, adult mortality and fecundity can increase if there are no or few sources of blood. This follows the idea that a mosquito living in an environment where much energy has to be used to find blood will do this at the cost of survival.

We adopt these general ideas for two species, *An. gambiae s.s.* and *An. arabiensis*. It should be noted that we have less confidence in the model for the *An. gambiae s.s.* M form, since aestivation (as documented by Lehmann et al. [57] and Adamou et al. [58]) is not included. In addition, there are some indications that the M form breeds in larger pools [59] and hence the puddle parametrization might have limited validity for this form.

In addition to time, the model can include two (three, since space is two-dimensional) additional dimensions, namely age and space. The space dimension allows dispersion of mosquitoes, meaning that (re)establishment through migration to areas that were previously free of *An. gambiae s.l.* is possible. The gradual invasion of Brazil by *An. arabiensis* in the 1930s [60] is one example of dispersion.

The ODEs were solved using the ODE solver lsoda [61–63]. The relative and absolute error tolerances were not modified from the original lsoda implementation ($1e^{-6}$). The model can be run either as a spatial model (with or without mosquito dispersion) or evaluated at a single point at which movement is neglected. A detailed overview of the possible model parameters can be found in Table 1.

Table 1 Model parameters

Variable	Description	Equation(s)/reference
T_{indoor}	Indoor temperature	36
T_{air}	Near surface temperature (2 m)	25, 26, 30, 36
ϵ	Potential number of new eggs	13
m_n	Number of mosquitoes in each age group	8
$P(B)$	Daily probability of getting a blood meal	41
T_{water}	Water temperature	14, 16, 18
T_{soil}	0-10 cm soil temperature	[91–94]
$\beta_{N,L}(T_{water})$	Natural mortiality rate, eggs, larva, and pupa	14, 1, 2, 3, 4, 5, 6
τ_{gamb}	*An. gambiae s.s.* development rate, aquatic stages	20
τ_{arab}	*An. arabiensis* development rate, aquatic stages	22
τ_E	*An. gambiae s.l.* development rate, eggs	[97] 1
$\tau_{L_{1-4}}$	*An. gambiae s.l.* development rate, instar 1-4	[97] 2, 3, 4, 5
τ_P	*An. gambiae s.l.* development rate, pupa	[97] 6
f_{arab}	Aquatic development rate modification *An. arabiensis*	[8]
f_{gamb}	Aquatic development rate modification *An. gambiae s.s.*	[8]
L_n	Number of larvae	21, 19
F_{arab}	Mortality rate modification	[72] 17
F_{gamb}	Mortality rate modification	[72] 15
S_f	scaling factor for wind dispersion	39
Fr_m	Flight range	41
E	Number of eggs	1
$G(T)$	Biting rate/gonotrophic cycle	26
t	time	
B_L	Larva biomass	1

$\beta_{I,x}$	Induced mortality in aquatic and adult stages	1, 2, 3, 4, 5, 6, 7, 8
SM_r	Dimensionless time varying water constant, or rate at which ovipositing sites are found	24
K	Carrying capacity	24
L_1	Number of 1^{st} instar larva	2
L_2	Number of 2^{nd} instar larva	3
L_3	Number of 3^{rd} instar larva	4
L_4	Number of 4^{th} instar larva	5
P	Number of pupa	6
C_{pred}	Predation constant. Currently set to 0	2
F_{gonot}	part of gonotrophic cycle formulation	26
D_d	Degree days	[108] , 26
T_c	Critical temperature	26
$\beta_{h,m}$	Adult mortality related to feeding	42
h	Number of humans	[42]
$\Phi_{\iota,J}$	flux	39
n	Dimension in age grid	
m_{size}	Size of newly emerged mosquitoes	9
m_{size_n}	Size of mosquitoes in age group n	12
L_{size}	Prediction of larva size	10
a_{spp}	Size constant	[22]
b_{spp}	Size constant	[22]
R_p	Potential river length in km	23
Ξ	Equally spaced river dataset resolution in degrees	23
ER	Earth radius in km (6371.22)	23
φ	latitude in radians	23
D	Diffusion coefficient	39
LT	Local time	37
κ	Diurnal modification for transport of mosquitoes	37
HBI	Human blood index	41, 42
$g\left(m_{size_n}\right)$	Size dependent mortality	28
$\beta_{N,m}$	Natural mortality of adult mosquitoes	32, 7, 8
$\varpi_{N,m}(\alpha,\zeta,a)$	Survival curve for adult mosquitoes	35, 31
α	Shape parameter for adult survival	33 30
T_{mod}	Sub-function for equation 33	34
$\rho_{bovine/cattle}$	Probability of finding cattle	41
ρ_{human}	Probability of finding humans	41

Description of model components

Differential equations for the aquatic compartment

The aquatic compartment consists of six stages: eggs (E), four larval stages (L_1, L_2, L_3, L_4), and pupae (P). Transitions between the different compartments can be expressed in terms of delayed equations. To simplify the solution and avoid numerical instabilities, we approximate the model as ODEs [21]. Lunde et al. reported on the errors introduced by this approximation [64].

New eggs added to the population depend on the number of adult mosquitoes (m), the size of adult mosquitoes (m_{size}), the inverse length of the gonotrophic cycle ($G(T)$), how much water is available (SM_r, dimensionless) and the larval biomass already present in puddles (B_L):

$$\frac{\delta E}{\delta t} = \epsilon(m, m_{size_n}) \cdot G(T) \cdot SM_r \cdot \left(1 - \frac{B_L}{K}\right) - (\beta_{N,E}(T) + \beta_{I,E} + \tau_E(T)) \cdot E, \tag{1}$$

where $\epsilon(m, m_{size_n})$ represents potential new eggs from each age group, $G(T)$ is either constant or dependent on temperature T, SM_r is a function of the relative soil moisture and the potential puddle formation area, K is the maximum larval biomass a grid cell can hold, $\beta_{N,E}(T)$ is natural mortality rate for eggs [Eqs. (16) and (18)], $\beta_{I,E}$ is the induced mortality rate for eggs (not specified) and τ_E is the inverse of development time from eggs to first instar larvae.

The term $1 - B_L/K$ is used as a scaling factor to modify the growth rate. When the population is low compared to the breeding sites available, its growth is high. As the population grows, there is more competition for food, predators become more abundant, and the growth slows. In the egg compartment this represents the idea that the mosquitoes will lay fewer eggs when breeding sites are already occupied [65].

First instar larvae (L_1) are added as eggs develop into larvae. Additional mortality is added in the transition stage in relation to how much biomass there already is in a given location [53]. This approximation of increased (density-dependent) mortality arises because of competition and predators; if a puddle already is full, the number of eggs developing to first instar larvae is reduced, whereas if a puddle is empty ($1 - B_L/K = 1$), no extra mortality occurs. Similar terms could have been added to the second, third and fourth instar larvae, but we assume that earlier life stages will be affected more by density-dependent competition and predation.

Shoukry looked at how fourth instar larvae of *An. pharoensis* prey on first instar larvae during a 24-h experiment [52]. Using these data, we add additional mortality for first instar larvae according to the density of fourth over first instar larvae. The constant C_{pred} is tunable to both limit the predation on L_1 and make it more specific to species in the future. At most temperatures, this constant does not influence the density of mosquitoes (Additional file 1).

The number of first instar larva is given by:

$$\frac{\delta L_1}{\delta t} = \tau_E(T) \cdot E \cdot \left(1 - \frac{B_L}{K}\right) - (\beta_{N,L}(T) + \beta_{I,L} + \tau_{L_1}(T)) \cdot L_1 - \frac{0.4465}{\left(\frac{L_4}{L_1} + 1\right)^{2.9891}} \cdot C_{pred}. \tag{2}$$

Second (L_2), third (L_3) and fourth instar larvae (L_4) and pupae (P) are controlled by the development rate τ and mortality β:

$$\frac{\delta L_2}{\delta t} = \tau_{L_1}(T) \cdot L_1 - \left(\beta_{N,L}(T) + \beta_{I,L} + \tau_{L_2}(T)\right) \cdot L_2 \tag{3}$$

$$\frac{\delta L_3}{\delta t} = \tau_{L_2}(T) \cdot L_2 - \left(\beta_{N,L}(T) + \beta_{I,L} + \tau_{L_3}(T)\right) \cdot L_3 \tag{4}$$

$$\frac{\delta L_4}{\delta t} = \tau_{L_3}(T) \cdot L_3 - \left(\beta_{N,L}(T) + \beta_{I,L} + \tau_{L_4}(T)\right) \cdot L_4 \tag{5}$$

$$\frac{\delta P}{\delta t} = \tau_{L_4}(T) \cdot L_4 - \big(\beta_{N,P}(T) + \beta_{I,P} + \tau_P(T)\big) \cdot P, \tag{6}$$

where β is the daily mortality rate, with the first subscript denoting natural (N) or induced (I) mortality and the second subscript denoting the aquatic stage. The subscript for the development rate, τ, corresponds to the aquatic stage. The parametrization schemes and data sources used to estimate the rate at which eggs are laid ($G(T)$ and ϵ), mortality (β) and the development rate (τ) are discussed later.

Differential equations for adult mosquitoes

The life history and mortality rate vary over the lifespan of a mosquito population. We formulated a model to account for this variation. Adult mosquitoes are denoted by m_n, where n indicates the age group; $n = 1$ is the youngest group and $n = 9$ refers to the oldest mosquitoes. The age groups in the model are $m_1 = [0, 1]$, $m_2 = (2, 4]$, $m_3 = (5, 8]$, $m_4 = (9, 13]$, $m_5 = (14, 19]$, $m_6 = (20, 26]$, $m_7 = (27, 34]$, $m_8 = (35, 43]$ and $m_9 = (44, \infty]$ days, with ageing coefficients a_n of 1.000, 0.500, 0.333, 0.250, 0.200, 0.167, 0.143, 0.125 and 0.067 for $n = 1, 2, \ldots, 9$, respectively. Mosquito ageing is represented by Ψ_n, where n denotes the age group. Ageing is time-invariant and is thus not related to the number of gonotrophic cycles.

Although there is no ageing from age group 9, the term Ψ_9 is included to limit the concentration of old mosquitoes. This is a user-specified variable and in the model results shown here we set this to $\frac{1}{15} day^{-1}$ for *An. arabiensis* and *An. gambiae s.s.*; this value should be set to ensure that mosquito populations can survive during dry periods [66, 67], but still hinder accumulation of old mosquitoes. This can be particularly useful if the mortality model described later is replaced with a model in which mortality is independent of age.

When m is written with subscripts $\imath$ and $\jmath$ in addition to n, this denotes inclusion of mosquitoes from neighboring areas. For example, subscript $\imath - 1$ indicates that mosquitoes to the west of the point of interest are interacting with the point of interest. The formulation presented here includes movement of mosquitoes, and where appropriate we denote mosquitoes by $m_{n,\imath,\jmath}$.

Again, β denotes mortality, with the first subscript denoting natural (N) or induced (I) mortality and the second subscript denoting the age group (m_n) of the mosquitoes. Φ represents the mosquito flux (transport) and subscripts $\imath$ and $\jmath$ define which boundaries are evaluated. This is discussed in the section "Movement of mosquitoes".

The number of adult mosquitoes of a specific age in a grid point is controlled by new mosquitoes from m_{n-1}, as well as the flux to and from the point of interest ($\sum_{\imath=-1}^{1} \sum_{\jmath=-1}^{1} \Phi_{\imath,\jmath} m_{n,\imath,\jmath}$), natural mortality β_{N,m_n}, induced mortality β_{I,m_n}, ageing to m_{n+1}, and mortality due to lack of food ($P(B)$). Parametrization schemes related to mortality are discussed later.

This results in the following equation for the first age group:

$$\frac{\delta m_1}{\delta t} = \tau_P(T) \cdot P + \sum_{\imath=-1}^{1} \sum_{\jmath=-1}^{1} \Phi_{\imath,\jmath} m_{1,\imath,\jmath} - (\beta_{N,m_1} + \beta_{I,m_1} + \Psi_1) \cdot m_1. \tag{7}$$

The equations for age groups $n = [2, 9]$ are

$$\frac{\delta m_n}{\delta t} = \Psi_{n-1} \cdot m_{n-1} + \sum_{\iota=-1}^{1} \sum_{\jmath=-1}^{1} \Phi_{\iota,\jmath} m_{n,\iota,\jmath} - (\beta_{N,m_n} + \beta_{I,m_n} + \Psi_n + \beta_{h,m}) \cdot m_n. \tag{8}$$

Differential equations predicting mosquito size

Mosquito size (m_{size}) is important for the efficiency of mosquito multiplication. There are also some indications that increased body size is a strategy for survival in arid environments [7]. In general, high larval density leads to a smaller body size as adults, and vice verse [68]. Where only one species is competing for a resource, such as in a small puddle, mosquito size, and hence the number of eggs laid by each mosquito, will be of less importance. If two species are competing for the same resource (e.g. *An. arabiensis* and *An. gambiae s.s.*), the trade off between development time and size can be important in competition for breeding sites. *An. gambiae s.s.* generally develop faster than *An. arabiensis*, but end up with a smaller body size. *An. arabiensis* spends more time in the aquatic stages and develops larger bodies, and can thus produce more eggs. Since our model includes competition between those species, we describe mosquito size as a function of competition for breeding sites. In theory this should improve our ability to separate geographical and seasonal distributions of *An. arabiensis* and *An. gambiae s.s.*

Since the size of *An. arabiensis* and *An. gambiae s.s.* stabilizes after approximately 4 days [7] and ovoposition does not start before this, it is not necessary to differentiate the maximum and minimum size depending on age to mimic changes in the number of eggs per mosquito with age. However, this may be required if mortality based on desiccation [7, 69] is used. Although mosquito size at a given time can be approximated using finite differences, we develop a different approach that is more efficient in terms of computational time in our model framework. Mosquito size for the first age group depends on larval size. Since the pupation time is short, this assumption is justified, although it might introduce minor errors. In a future version of the model, we plan to predict larval size dynamically. The limitations set on mosquito size (described in "Parametrization schemes in the aquatic stages") in this model might lead to *An. arabiensis* that are slightly too small compared size in the field study of Ye-Ebiyo et al. [70], but the size is in line with studies by Huestis et al. [71] and Kirby et al. [72]. Kirby et al. also noted that mixed populations of *An. arabiensis* and *An. gambiae s.s.* had a negative effect on mosquito size at some temperatures. This mechanism is not included in the current work. However, the most important aspect of modelling of mosquito size is to capture seasonal and spatial variations.

For size prediction we use the symbol m_{size_n}, where n is the age group as described above.

The size (wing length in mm)of newly emerged mosquitoes is approximated according to the linear relationship

$$m_{size_e} = 1.25 + 5 \cdot L_{size}, \tag{9}$$

where larva size L_{size} (in mg) is approximated as:

$$L_{size} = a_{spp} - b_{spp} \cdot min\left(\frac{B_L}{K}, 1\right). \tag{10}$$

The constants a_{spp} and b_{spp} are 0.45 and 0.12 for *An. arabiensis* and 0.383 and 0.147 for *An. gambiae s.s.*, respectively [22].

The size of mosquitoes in the first age group at any time is given by

$$\frac{\delta m_{size_1}}{\delta t} = min\left(max\left(\frac{\tau_P(T)\cdot P}{m_1}, 0\right), 1\right)\cdot log\left(\frac{m_{size_e}}{m_{size_1}}\right)\cdot m_{size_1}. \quad (11)$$

Therefore, the size of newly emerged mosquitoes (m_{size_1}) depends on the number of newly emerged pupae and the relative density of larva at the breeding site.

For the remaining age groups, size m_{size_n} is estimated as

$$\frac{\delta m_{size_n}}{\delta t} = min\left(\frac{\Psi_{n-1}\cdot m_{n-1}}{m_n}, 1\right)\cdot log\left(\frac{m_{size_{n-1}}}{m_{size_n}}\right)\cdot m_{size_n}. \quad (12)$$

Therefore, the size in age groups 2-9 only depends on the number of mosquitoes surviving from one age group to the next (m_{n-1}) and the size of mosquitoes in the younger age group ($m_{size_{n-1}}$).

Model forcing

To drive a dynamic malaria model it is necessary to have boundary conditions that are consistent over time and space. Temperature, relative humidity, and rainfall data from weather stations are point measures. Hence, they might not be representative of larger areas over shorter time scales. This is especially true in areas with varying topography or where convective rainfall is dominant [73–75]. Despite the limitations of rainfall stations, they can provide a robust estimate of large-scale events. By pooling data from several stations, the error for a single station is reduced and the data can provide a good estimate for dry and wet years, for example. Hence, weather stations are useful tools for validating climate models.

The problems of point measurements are described later, and represent one of the reasons why OMaWa is tightly linked to a climate model. As shown in sensitivity experiments, the model can also be run with constant forcing (e.g. temperature) or with data from weather stations.

Where we present results for Africa as a whole, OMaWa is driven by data from WRF 3.3.1. This realization (TC50), described in part two of this paper, has a tropical channel set-up in which set-up, the domain consists of boundaries above and below a certain latitude and no side boundaries. The model was run at 50-km resolution from January 1, 1989 to January 1, 2009. At the northern ($45°N$) and southern ($-45°N$) boundaries the model was driven by Era Interim. The Kain Frisch cumulus parametrization scheme was used [76, 77]. This experiment was not designed to reproduce observed year-to-year weather variability, but to assess the mean mosquito density and distribution. The driving experiment is described in the section on model validation.

Climate and weather models

Currently, our best guess of (future) climate at multidecadal time scales comes from general circulation models (GCMs). These models are designed to close the energy budget of the Earth and include an interactive representation of the atmosphere, ocean, land, and sea ice. A set of scenarios with different emissions describes how sensitive the climate is to atmospheric constituents (greenhouse gasses) [78]. While climate is the average weather over time and space, weather can change over minutes, hours, days and seasons. The same equations used to predict climate are used to predict weather. However, weather forecasts are more dependent on current observations of the atmosphere. Hence, weather predictions are initial value problems, whereas climate simulations are rather boundary value problems.

Both climate and weather models are mostly structured on a grid, with coordinates from west to east (x), north to south (y) and bottom to top (z). In the grid, one square (or polygon) represents the weather within that square. While climate models often have a horizontal resolution of more than 10000 km^2, operational weather models such as the European Centre for Medium-Range Weather Forecast (ECMWF) model are run at approximately $160km^2$. If the state of the atmosphere is observed correctly, higher resolution can lead to better local skill in predicting the weather. A hybrid between a weather model and a climate model is a limited-area model (LAM), which relies on initial and boundary conditions from a weather or climate model. Given these conditions (weather), the LAM can be run at a higher resolution over a limited area, which potentially improves the spatial accuracy of the coarse model [79]. The WRF model is a widely used LAM [41].

In tropical regions, most rainfall comes from convective clouds. This type of rainfall is generally intense and of short duration. The geographical extent of such rainfall episodes may be limited. Therefore, rainfall measurements in regions where convective rainfall is dominant should be handled with care [74, 75, 80, 81], especially when extrapolating station data to areas with no data. While station data are accurate at a specific point, climate models and satellite estimates give a more general description of the weather within a certain area; Chen and Knutson reviewed how models compare to observations at varying scales [82]. Since future climate is projected using climate models and considering the limitations of weather stations, construction of a mosquito/malaria model around a LAM is a good choice. The LAM will have higher resolution than most climate models, with higher-resolution orography, coastlines, and land use, but will still give a general description of the weather within a certain area.

Parametrization schemes in the aquatic stages

To relate a variable such as mortality to the physical environment, we need simplified equations that describe this relationship. An equation in which temperature influences mortality only states that there is a relationship between the two, but does not explain why temperature modifies mortality. In this paper we use parametrization schemes to represent the influence of the environment on mosquitoes. This section describes the aquatic parametrization schemes used, excluding water availability, which is discussed later.

The aquatic stages comprise eggs, four instar stages, and pupae. The number of eggs in a location at any time is controlled by the number of potential new eggs laid (ϵ), available water (K), natural and induced mortality ($\beta_{N/I,E/L/P}$) and movement from the E to the L_1 compartment. In addition, 20% instant mortality is introduced when rainfall exceeds the infiltration rate. This is in line with observations by Paaijmans et al. [83]. The number of new eggs is simplified to a function of the number of gravid mosquitoes in each age group and their size (measured as wing length) based on observations [55, 84–86]. The critical size is set to a wing length of 2.6 mm, which is less than that observed by Lyimo and Takken [85] but greater than observations by Yaro et al. [87]. Maximum wing length is set to 3.3 mm for *An. gambiae s.s.* [88, 89] and 3.7 mm for *An. arabiensis* [70]. The relationship between the number of eggs (ϵ) and wing length (m_{size_n}) is then approximated according to the linear relationship

$$\epsilon = \sum_{n=1}^{9} \left\{ \begin{array}{rl} (-433.3 + 166.7 \cdot m_{size_n}) \cdot m_n & \text{if } m_{size_n} > 2.6 \text{ mm} \\ 0 & \text{otherwise} \end{array} \right\}, \tag{13}$$

where m_n is the number of mosquitoes in age group n. Note that this limits the number of eggs laid by a single mosquito per gonotrophic cycle to approximately 184, which is somewhat less than the number observed by Yaro et al. [87], but in line with that reported by Howard et al. [90].

Estimation of water temperature

Using the 0-10-cm soil temperature (T_{soil}) from the NOAH land surface model [91–94] to approximate the mean water temperature (T_{water}) in larval habitats, we assume that evaporative cooling and heat fluxes at the water boundaries are negligible. Hence, the water temperature is equal to the top soil temperature. Paaijmans et al. showed that the 5-cm soil temperature represents the water temperature in small ponds reasonably well [95]. Therefore, the model will have limited validity in areas where larger puddles are the main breeding sites. There is also a chance that diurnal fluctuations will be slightly over- or underestimated. When a grid cell covers several km^2, this effect should be negligible, although we do not have data to support this. We hope to improve the prediction of water temperature in the future, either by modelling this explicitly or using a parametrized version based on data from Huang et al. [96].

Parametrization of mortality

We used two approaches to calculate mortality in the aquatic stages. In the simpler approach, we assume that mortality and development time in the aquatic stages are independent of the species. We also assume that the relationship between the mortality rate and temperature is the same for eggs, instars and pupae. In this method we do not consider competition effects as described by Paaijmans et al. [8]. This type of parametrization is suitable when the model is used for one species only (e.g. if the model represents an area where only one of the two species is present).

Species-independent mortality (BLL)

Data provided by Bayoh and Lindsay [97] were used to describe the mortality rate according to Eq. (14) ($p < 0.01, R^2 = 0.81$). We call this the BLL method. Mortality rate data are plotted in Figure 2b.

$$\beta_{N,L}(T_{water}) = \left(\frac{k_1}{T_{water}^{k_2}} + e^{k_3 \cdot T_{water} - k_4} \right) \cdot k_5 + \frac{k_6}{1 + k_7 \cdot e^{k_8 \cdot (T_{water} - k_9)}}, \tag{14}$$

where $\beta_{N,L}(T_{water}) = \beta_{N,E}(T_{water}) = \beta_{N,P}(T_{water})$ is the aquatic mortality rate per day and T_{water} is the water temperature ($^\circ C$). The constants k_n are given in Table 2.

Figure 2 Water temperature and mortality rates (day^{-1}) in the aquatic compartments. Blue points show data used to estimate the mortality curves. Blue lines indicate mortality without competition, while light blue to red shows mortality as competition increases. For reference, red points show data from Holstein [98].

Table 2 Constants for equation 14 and 33

Constant	Value	Equation
k_1	700000	14
k_2	8.4	14
k_3	.126	14
k_4	10.8	14
k_5	150	14
k_6	$-.08$	14
k_7	.1	14
k_8	$-.61$	14
k_9	33	14
c_1	0.1675256	33
c_2	0.0121402	33
c_3	0.1686	33
c_4	1.991	33
c_5	1.881	33
c_6	$4.641589e26$	33
c_7	250	33
c_8	23	33
c_9	12	33
c_{10}	100	33
c_{11}	3	33

Species-dependent mortality (KBLL)

Kirby et al. reported that the mortality rate of *An. gambiae s.s.* and *An. arabiensis* is modulated by the presence of each other in the temperature range $25 - 35°C$ [72]. To account for this we developed two mortality models, one for *An. gambiae s.s.* and one for *An. arabiensis*. We call this parametrization scheme KBLL. The mortality rates are based on data from Bayoh and Lindsay [97] and from Kirby et al. [72]. Although Holstein also reported larval mortality for (*An. gambiae s.s.*) when exposed to extreme low and high temperatures [98], we did not include these data when estimating the mortality curves. However, the data are plotted in Figure 2 for comparison. According to our curves, the *An. arabiensis* mortality rate will increase in the range $25 - 35°C$ as the relative presence of *An. gambiae s.s.* increases. Conversely, the mortality rate of *An. gambiae s.s.* will decrease as the proportion of *An. arabiensis* increases. The mortality rate $\beta_{N,L}$ is given by

$$F_{arab} = min\left(\frac{\sum_{n=1}^{4} L_{n,arab}}{\sum_{n=1}^{4} L_{n,gamb}}, 1\right) \tag{15}$$

$$\beta_{N,L,gamb}(T_{water}) = \begin{cases} 0.002404075 \cdot T_{water}^2 - 0.1127944 \cdot T_{water} + 1.337783 \\ \beta_{N,L}\left(T_{water}\right) \cdot (0.4 + 0.6 \cdot (1 + \sin\left(-10.9956 + 0.3142 \cdot T_{water}\right)))^{F_{arab}} \end{cases} \quad \text{if } 25 \leq T_{water} \leq 35 \tag{16}$$

and

$$F_{gamb} = min\left(\frac{\sum_{n=1}^{4} L_{n,gamb}}{\sum_{n=1}^{4} L_{n,arab}}, 1\right) \tag{17}$$

$$\beta_{N,L,arab}(T_{water}) = \begin{cases} 0.0006556736 \cdot T_{water}^2 - 0.02980226 \cdot T_{water} + 0.3587285 & \\ \beta_{N,L}(T_{water}) \cdot \left((2 + \cos(-18.8496 + 0.6283 \cdot T_{water}))^{0.9508002}\right)^{F_{gamb}} & \text{if } 25 \leq T_{water} \leq 35 \\ \beta_{N,L,gamb}(T_{water}) & \text{if } T_{water} \leq 21.91209. \end{cases} \tag{18}$$

F_{gamb} and F_{arab} are the ratio of *An. gambiae s.s.* to *An. arabiensis* larvae and *An. arabiensis* to *An. gambiae s.s.* larvae, respectively. At each time step, L_{size} is estimated as a function of B_L and K. As the density increases, there will be more competition and hence less food for each larva, which leads to smaller larvae.

Parametrization of the development rate

The rate of development between the different aquatic stages follows the corrected version of Bayoh and Lindsay [97]. Since these data are only valid for *An. gambiae s.s.*, we made a small modification to prolong the development times for *An. arabiensis*. Data from Kirby et al. [72] and Paaijmans et al. [8] suggest that time for development from a larva to an adult is approximately 5.5% longer for *An. arabiensis* than for *An. gambiae s.s.* Hence, we increased the development time for *An. arabiensis* by 5.5%. The reason for this longer development time is that *An. arabiensis* takes longer to develop a larger body. Curves of the development rate are shown in Figure 3.

Figure 3 Water temperature according to development time in days from first instar to adult. Left panel: ratio of *An. gambiae s.s.* to *An. arabiensis*. When greater numbers of *An. gambiae s.s.* are present, *An. arabiensis* develop more slowly. Right panel: ratio of *An. arabiensis* to *An. gambiae s.s.*. When greater numbers of *An. arabiensis* are present, *An. gambiae s.s.* develop more quickly.

The two previous studies also suggest that the development rate [8] and mortality [72] of the two species are modulated by the presence of each other, so we take account of this in out model. The development time for *An. arabiensis* is prolonged in the presence of *An. gambiae s.s.*, while the time is shortened for *An. gambiae s.s.* as the relative proportion of *An. arabiensis* increases. Using data from Paaijmans et al. [8], the development rate τ is modified according to

$$f_{arab} = min\left(100 \cdot \frac{\sum_{n=1}^{4} L_{n,arab}}{\sum_{n=1}^{4} L_{n,gamb} + \sum_{n=1}^{4} L_{n,arab}}, 75\right) \tag{19}$$

$$\tau_{gamb} = \tau_{gamb} \cdot (1 - f_{arab} \cdot 0.0008421)^{-1} \tag{20}$$

for *An. gambiae s.s.* and

$$f_{gamb} = min\left(100 \cdot \frac{\sum_{n=1}^{4} L_{n,gamb}}{\sum_{n=1}^{4} L_{n,gamb} + \sum_{n=1}^{4} L_{n,arab}}, 75\right) \tag{21}$$

$$\tau_{arab} = \tau_{arab} \cdot (1 + f_{gamb} \cdot 0.002138)^{-1} \tag{22}$$

for *An. arabiensis*. f_{arab} and f_{gamb} is the fraction of *An. arabiensis* and *An. gambiae s.s.*, respectively.

Parametrization of breeding sites

The formation of puddles can be described as a balance of runoff, infiltration, evaporation, and rainfall entering the puddle. The formulation of an idealized puddle can be found in Additional file 2.

Modelling of every single breeding site requires high enough resolution to resolve the puddle. In practice this is not possible and the problem has to be simplified.

Mushinzimana et al. described typical breeding sites in a Kenyan highland area [99]. Most of the puddles were located at less than 100 m from rivers, which means we can assume that semi-permanent puddles will mostly form in the proximity of rivers and lakes. They also found that the number of breeding sites was close to threefold higher in the rainy season compared to the dry season, and grouped breeding sites by surface area.

If we assume that breeding mainly occurs in the vicinity of potential rivers and lakes, the availability of breeding sites can be expressed as a function of potential river length and soil saturation. At high resolution this might not always be true [6], but since the model is designed to be applied to coarser grids, we believe the assumption is as reasonable as or more reasonable than the common assumption that puddle formation is only dependent on rainfall [29]. The newest version of the NOAH land surface model in WRF 3.4 also includes groundwater and dynamic vegetation, and future versions might change the way in which puddles are parametrized. In OMaWa we introduce a simple parametrization scheme to represent breeding sites.

The Hydrological Data and Maps based on SHuttle Elevation Derivatives at Multiple Scales (HydroSHEDS) 15s river data set from the US Geological Survey (USGS) [100] was used to derive the total potential river length within a grid cell. Since the algorithm used to develop this data set describes where water would collect if it were available within the catchment, it also represents a general description of the potential for water aggregation within an area. However, the validity might decrease on moving to finer scales [6].

Here we divide rivers into three different classes: perennial, intermittent and ephemeral streams. For each class, potential river length (R_p, km) within a grid is defined as

$$R_p = \sum \Xi \cdot \frac{2\pi ER}{360} \cdot \cos\varphi, \tag{23}$$

where Ξ is the equally spaced river data-set resolution in degrees, where $\Delta lon = \Delta lat$, ER is the radius of the Earth (6371.22 km) and φ is latitude in radians.

In a simplified model we estimate puddle volume as a function of river length and relative soil moisture. Although this is a very crude estimate, we compared this simple model with data from Mushinzimana et al. [99] and derived a simple expression for the carrying capacity in a grid cell:

$$K = \frac{B_{L,max}}{km_{river}} \cdot R_p \cdot SM_r, \tag{24}$$

where $\frac{B_{L,max}}{km_{river}}$ is the maximum larval biomass per km of river (2400 mg, estimated from data collected by Munga et al. [101]) and SM_r is the relative soil moisture content (fraction).

In the current implementation we do not distinguish between fast- and slow-flowing rivers. It should be noted that this way of approximating breeding sites has limited validity in areas with irrigation or around rivers where breeding sites could form as rivers recede [66, 67, 102]. Some special cases, such as along the River Nile in Sudan, where breeding sites form as a result of rainfall hundreds of kilometers away, will not be captured at all [103].

Parametrization of the gonotrophic cycle

The gonotrophic cycle depends on temperature and is important for the vectorial capacity of mosquitoes. Lardeux et al. studied the gonotrophic cycle for *An. pseudopunctipennis* [104]. We combine their data with other published studies on anophelines to estimate the length of the gonotropic cycle. There are few studies on *An. gambiae s.l.*, and hence we have to assume that other anophelines share the same physiology and strategy with respect to the gonotropic cycle. Ruiz et al. showed there are some differences [23], but until further evidence of the reproductive strategies of different members of *Anopheles* genera becomes available, we will not consider this effect. Studies used to develop the formula include those by Guillermo et al. ([105], *An. albimanus*), Afrane et al. ([106], *An. gambiae s.l.*), and Maharaj ([107], *An. arabiensis*). We also include the formula given by Hoshen and Morse [108]. Their model is based on degree days and is included according to Eq. (26). The gonotropic rate (day^{-1}) and data used to develop the formula are shown in Figure 4.

Figure 4 Inverse of the duration of the gonotropic cycle according to the mean daily temperature (in $^{\circ}C$). The solid black line shows Eq. 26 and the dashed line shows the formula given by Hoshen and Morse [108].

$$F_{gonot} = min\left(max\left(-\frac{2}{3} + \frac{1}{30}\cdot T_{air}, 0\right), .5\right) \tag{25}$$

$$G(T) = \left(1 + \frac{D_d}{T_{air} - T_c}\right)^{-1} \cdot F_{gonot} + \left(1.71 + 544347.6 \cdot T_{air}^{-3.93}\right)^{-1} \cdot \left(1 - F_{gonot}\right), \tag{26}$$

where T_{air} is the air temperature ($^{\circ}C$), D_d is degree days, and T_c is the critical temperature from Hoshen and Morse [108], with $D_d = 37$, and $T_c = 7.7$.

Parametrization of the age-dependent mortality of adult mosquitoes

The mortality of adult anophelines differs according to age and species [7, 107, 109]. This has often been overlooked in mosquito models [23, 110]. To show how this assumption can influence the stability of mosquito populations and malaria transmission, we use the mortality model of Martens [110] as a reference. We also plot Eq. 7 from Ermert [29] in Figure 5 to highlight the differences between this model and established models. For convenience, we repeated Marten's equation, as follows:

$$\beta_{N,m}(T) = 1 - e^{-\frac{1}{-4.4+1.31\cdot T-.03\cdot T^2}}. \tag{27}$$

Figure 5 Proportion of *An. gambiae s.s.* surviving at 60% relative humidity and mean temperature of 0, 10, 20, 30, and $40^{\circ}C$ (selected for clarity) according to time (in days). Dashed vertical lines indicate the different age groups in the model. The Survival curve panel shows Eq. 31, while the Numerical solution panel shows survival in the model when the age groups are split into nine classes. For reference we also show survival according to the Marten equation (27) and Eq. 7 from Ermert et al. (Bayoh scheme) [29]. The mean absolute error for all combinations of temperature and relative humidity was 73 for our model, 171 for the Marten model, and 129 for the Bayoh scheme.

Our new survival curves are based on unpublished data from Bayoh and Lindsay [47]. The validity ranges from 5 to 40°C by 5°C and $40-100\%$ by 20% relative humidity. We name the scheme BLLad (Bayoh-Lindsay-Lunde adult mortality). The data set and the curves are valid for *An. gambiae s.s.* The lowest agreement between the model and the data is at 40% relative humidity and 40°C. While the data suggest that all *An. gambiae s.s.* would be dead after approximately 2 days, the survival curve would result in no mosquitoes after approximately 4 days at 40% relative humidity and 40°C. To correct for this error, we include data from Kirby and Lindsay [111], who described the responses of *An. gambiae s.s.* and *An. arabiensis* to high temperatures. By assuming that maximum survival is 480 min for *An. gambiae s.s.* and 1440 min for *An. arabiensis* at temperatures greater than 40°C, we can set the mortality rate to $3day^{-1}$ and $1day^{-1}$, independent of age group. However, there are uncertainties at relative humidity below 40%. The lack of studies in this range is a limitation of this survival model, and could make the model less accurate for *An. gambiae s.l.* in some regions. The basic principle of these survival curves is that mortality will be low in the first few days after emergence. In addition, mosquitoes that survive up to a certain age have a higher survival probability (depending on T_{air} and relative humidity). In Figure 5, survival at 60% relative humidity and 0, 10, 20, 30, and 40°C is plotted.

Size affects the survival of adult mosquitoes [7, 51, 54, 55]([56], *Aedes aegypti*). If we assume that the major differences in mortality between *An. gambiae s.s.* and *An. arabiensis* can be attributed to mosquito size, we can modify α as a linear function of mosquito size. Here we subjectively choose reasonable constants for $h\left(m_{size}\right)$. T_{air} may be completely or partly replaced by indoor temperature (T_{indoor}, described later), depending on the proportion of mosquitoes indoors. In experiments covering the African domain, we assumed that 80% of *An. gambiae s.s.* and 20% of *An. arabiensis* are located indoors.

$$g\left(m_{size_n}\right) = 2.1731 - 0.3846 \cdot m_{size} \tag{28}$$

$$f(RH) = 6.48007 + 0.69570 \cdot (1 - e^{-0.06 \cdot RH}) \tag{29}$$

$$\alpha = \left(e^{(10+((1+\frac{(T_{air}+1)}{21})^{(2/3)})\cdot((1+\frac{(T_{air}+1)}{21})^2-(1+\frac{(T_{air}+1)}{21})\cdot 2-f(RH)))}\right) \cdot g\left(m_{size_n}\right) \tag{30}$$

$$\varpi_{N,m}(\alpha,\zeta,a) = \sum_{i=0}^{a}\left(\frac{\left((\alpha\cdot a)^{\sum_{i=0}^{n=(\zeta-1)} n}\right)}{\sum_{i=0}^{n=(\zeta-1)} n!}\right)\cdot e^{(-\alpha\cdot a)}, \tag{31}$$

where $\zeta = 6$, g is a function of mosquito size, and RH is relative humidity. The mortality rate for each age interval can then be approximated as

$$\beta_{N,m_n} = \left\{ \begin{array}{rl} \frac{\log\left(\frac{\varpi_{N,m_{t_2}}}{\varpi_{N,m_{t_1}}}\right)}{\Delta t} & \text{if } T < 40 \\ 3 & \text{otherwise} \end{array} \right\}. \tag{32}$$

If we assume that differences in adult mortality for *An. gambiae s.s.* and *An. arabiensis* can be explained by differences in body size, these BLLad curves can be used for both species. We explore this mortality model in [64].

AL adult mortality

A similar approach can be used for *An. arabiensis*. Using survival curves reported by Afrane et al. ([112], Figure two) (copyedited with g3data [113]), we can estimate mortality based on the daily maximum temperature. Because of the few data points, this approach is much more uncertain and should be considered experimental. The advantage of this mortality model is that the data are not estimated from

a laboratory setting. The maximum temperature reflects some aspects, such as radiation, albedo, and humidity, of the environment in which mosquitoes live. In some of the results presented in part two, we use this model for adult survival.

$$\alpha = c_1 - c_2 \cdot \left(c_9 + T_{mod}^{c_3} \cdot \left(T_{mod}^{c_4} - T_{mod} \cdot c_5 - c_8 \right) - c_7 \cdot \frac{c_{10}^{\frac{T_{max}}{c_{11}}}}{c_6} \right) + e^{-\left(\frac{T_{max}}{5}\right)} \cdot c_{11} \tag{33}$$

$$T_{mod} = 1 + \frac{T_{max} + 18}{11.10}. \tag{34}$$

Constants $c_{1,\ldots,11}$ are listed in Table 2. By setting $\zeta = 2$ we can simplify the survival curve for *An. arabiensis* to

$$\varpi_{N,m}(\alpha, \zeta, a) = \sum_{i=0}^{a} (1 + \alpha \cdot a) \cdot e^{-\alpha \cdot a}. \tag{35}$$

The corresponding curve is shown in Figure 6.

Figure 6 Proportion of *An. arabiensis* surviving at daily maximum temperatures. Estimated from Afrane et al. [112] (blue line). Dashed vertical lines indicate the different age groups in the model (grey lines).

Parametrization of air temperature

Paaijmans et al. discussed the importance of using indoor rather than outdoor temperature, to describe the environment for mosquitoes and parasites [114]. They included two studies that showed the relationship between indoor and outdoor temperature in Kenya [115] and Tanzania [116]. Here we add two additional studies, one from Kenya [48] and one describing the temperature in traditional and low-cost modern housing in the Eastern Cape, South Africa [117]. The data used to parametrize equation 36 came from; 1, Afrane et al. [48]; 2, Makaka and Meyer [117]; and 3, Paaijmans et al. [114–116] ($R^2 = 0.89$). It is clear that temperatures inside a house are more stable than outdoor temperatures. House type greatly influences daily temperature fluctuations [117, 118], and the model used here might not be valid for all house types. While some studies have assumed that houses are always hotter than the surroundings [119], we approximate the indoor temperature as

$$T_{indoor} = 10.33 + 0.58 \cdot T_{air}. \tag{36}$$

Since the data are based on maximum and minimum temperatures, the timing of the indoor temperature might be offset by a couple of hours. This is evident in a study by Makaka and Meyer [117], who delayed the maximum indoor temperature by a couple of hours compared to the environmental temperature. At present we do not account for this delay, since the diurnal temperature ranges will be correct even if we do not. The data and regression line are shown in Figure 7. Further studies on indoor compared to outdoor temperatures are needed to make this correction more accurate.

Figure 7 Relationship between outdoor and indoor temperatures. Numbers denote the study from which data were taken: 1, Afrane et al. [48]; 2, Makaka and Meyer [117]; and 3, Paaijmans et al. [114–116]. The blue area represents the 95% confidence interval, and the black line shows Eq. 36. $R^2 = 0.89$.

Hence, T_{air} can be partly or fully replaced by T_{indoor}, depending on the proportion of mosquitoes indoors.

It should be noted that we still do not include temperatures in resting places described by Holstein, such as holes in rocks and cracks in soil, covered pigsties, rabbit hutches, hen coops and dry wells [98], and by de Meillon ([120], under stones).

Approximation of mosquito movement

The role of diffusion and advection in vector borne diseases have been explored in several papers [102, 121–127]. Considering the gradual invasion of Brazil in the 1930s by *An. arabiensis* [60] it can be argued that movement of mosquitoes is important over decades. Here we include the active and passive transport of mosquitoes as fluxes across grid boundaries. Passive transport is movement of mosquitoes caused by wind, while active transport is movement due to flying. On shorter time scales the role of such movement will be limited. However, on long time scales it is necessary to allow mosquitoes to travel to allow them to establish in new locations.

Transport of mosquitoes is defined by fluxes (s^{-1}) at the grid boundaries. In the model we allow fluxes from the eight neighboring grid points. A special case is implemented when a neighbouring cell is water. In this case, fluxes to water are reduced to 0.1% of the original flux to avoid large losses of mosquitoes along the coastline. Given strong winds from land to the ocean, such an assumption could lead to accumulation of mosquitoes along the coast. Conversely, allowing free movement to the ocean could lead to undesired loss of mosquitoes.

Since the movement of mosquitoes has a high computational cost, the spatial fluxes do not change the size calculations. This will introduce some minor errors when the movement of mosquitoes is low compared to their density, with larger errors if many mosquitoes are moved relative to their density. When a cell free of mosquitoes is colonized, the size is set to 3.05 mm.

The possible flight range of anophelines varies with food availability [128]. We do not include vegetation types in the model and hence it is hard to justify differences in flight performance based on, for example, land use. The dispersion coefficient describes how far mosquitoes can move in a day. We assume that the dispersion coefficient D is constant, independent of geographical location. For *An. gambiae s.s.* and *An. arabiensis*, real flight performance outside the laboratory of only a few hundred meters per day (approx. 300-700 m) has been reported [102, 129, 130]. In this experiment we subjectively chose $D = 30mday^{-1}$ independent of age group. Anophelinae also travel with humans [131], which adds to the transport equation and makes the dispersion coefficient uncertain. Gillies noted that wind direction mostly has a minor effect on dispersal [129], while de Meillon [132] and Adams [133] reported distances of 2-4.5 miles (3-7 km) in the direction of the prevailing wind. Thus, it cannot be ruled out that wind plays a role on longer time scales. Hence, we express movement caused by wind as a function of 10-m zonal (u) and meridional (v) wind components (ms^{-1}). This can be understood by considering the following example. For a constant u-wind of $10ms^{-1}$ and v-wind set to 0, mosquitoes will be moved a distance related to a scale factor S_f, which is equal to the distance travelled at $20ms^{-1}$ to the east. For

example, with $S_f = 750 mday^{-1}$, the eastward distance traveled will be $S_f \cdot \frac{10 ms^{-1}}{20 ms^{-1}} 375$ m in 1 day, but since each mosquito is not modelled individually, it would be more natural to describe this as a fraction moving a certain distance. Different wind directions and speeds will result in other distances/fractions and directions. D and S_f are unknown tunable constants.

Since the species considered here are most active at night [22], movement will be suppressed between 06:00 and 18:00 h (local time) and amplified at night according to

$$\kappa = \frac{\left(\cos\left(LT \cdot \frac{\pi}{12}\right) + 1\right)^2}{1.506925}, \tag{37}$$

where LT is local time, $\int_0^{24} \kappa \approx 1$ and

$$LT = \begin{cases} UTC_{time} + \frac{longitude}{15} - 24 & \text{if } UTC_{time} + \frac{longitude}{15} >= 24 \\ UTC_{time} + \frac{longitude}{15} + 24 & \text{if } UTC_{time} + \frac{longitude}{15} <= 0 \\ UTC_{time} + \frac{longitude}{15} & \text{otherwise.} \end{cases}$$

Transport of mosquitoes and mosquito sizes inside and outside a grid are defined by

$$\frac{\delta m_n}{\delta t} = \sum_{\iota=-1}^{1} \sum_{\jmath=-1}^{1} \Phi_{\iota,\jmath} m_{n,\iota,\jmath}. \tag{38}$$

More specifically, during a time Δt, movement can be calculated as follows. On a day with no wind, transport is equal in all directions, $D = 30\ mday^{-1}$, and the flux at a boundary is defined as

$$\Phi_{\iota,\jmath} = \kappa_{\iota,\jmath} \Delta t \cdot \frac{D}{\Delta d_{\iota,\jmath} \cdot 24 \cdot 60 \cdot 60}, \iota = \{-1, 1\}, \jmath = \{-1, 1\} \tag{39}$$

and transport $\eta_{\iota,\jmath n}$ is then equal to

$$\eta_{\iota,\jmath n} = m_{\iota,\jmath n} \cdot \Phi_{\iota,\jmath}. \tag{40}$$

In the presence of wind, we obtain additional transport as a function of zonal and meridional wind components.

Mortality related to feeding

One factor that is often overlooked in malaria (mosquito) models is survival related to food availability ($P(B)$). Ye-Ebiyo et al. reported that maize pollen availability has a positive effect on larval (and hence mosquito) fitness [70, 134]. Creating maps of plant types is beyond the scope of this study, and hence we chose not to account for mortality related to crops. However, we performed initial tests in which we included GlobCover Land Cover version V2.2 (European Space Agency [135]) to give a rough estimate of regions where increased fitness could be expected. The other source of food for female anophelines is blood. Compared to a starved mosquito, a mosquito that has had access to blood on days 1-3 has a theoretical flight distance that is increased by a factor of 6-7 [128]. Therefore, it is plausible that the higher (lower) the probability of finding a blood meal ($P(B)$), the higher (lower) is survival in the early life stages of adult mosquitoes. Bouma and Rowland reported higher parasite prevalence among children of families who kept cattle compared to those who did not [136], which can indicate either higher survival (older mosquitoes) or simply that some anophelines are attracted to cattle. If we assume

that a newly emerged mosquito has a flight range of $Fr_m = 0.5\, km^2 day^{-1}$, the daily probability of finding a blood meal can be calculated as

$$P(B) = \left\{ \begin{array}{ll} HBI \cdot \rho_{human} \cdot Fr_m + (1 - HBI) \cdot \rho_{bovine} \cdot Fr_m & \text{if } P(B) < 1 \\ 1 & \text{otherwise} \end{array} \right\}, \tag{41}$$

where ρ_{human} and ρ_{bovine} is the probability of finding a human and bovine source, respectively. ρ_{humans} is defined as the human population density per km^2 multiplied by 0.1 (since a smaller area on a human is accessible) and ρ_{bovine} is defined as the bovine density per km^2, each with a user-defined threshold at which the density is so low that $P(B)$ is virtually zero. Since $P(B)$ is a conceptual parameter, it can be tuned.

Since blood meals, besides sugar meals, are important for the mobility [128] and survival of female anophelines [137], the success of a species is likely to be linked to the presence of the preferred host. The dominant blood source for *An. arabiensis* is bovine and human blood, while it is human blood for *An. gambiae s.s.* [138]. In reality there are strong indications that the human blood index is a dynamic quantity rather than a constant [139–142]. In the current implementation, *HBI* is a static number and hence there are probably errors related to this term. To find the probability of feeding on humans at each time step, we combine two data sets. Between 2000 and 2010 we use population densities from the Gridded Population of the World (GPW) [42], and for before 2000 and after 2010 we use growth rates from the Population Division of the Department of Economic, and Social Affairs of the United Nations Secretariat [143]. Since there are no projections of cattle densities, this quantity is time-invariant and based on Food and Agriculture Organization (FAO) 2005 estimates [44]. We are currently working to include time-varying cattle densities.

In the model, mortality caused by food limitations is a function of how many humans or cattle are available per mosquito and the human blood index. We assume that *HBI* is time- and space-invariant, and only depends on the species. For simplicity we chose available humans to be humans who are not sleeping under a bed net. In the simulations presented here, we set bed net usage to zero, and hence the results represent mosquito distribution without interventions. Bayoh et al. hypothesized that the survival of the different species is related to the availability of the preferred host [9]. The daily mortality rate caused by limited human blood is expressed as

$$\beta_{h,m} = max\left(1 - \frac{\frac{30}{HBI}\sum h}{\sum_{n=2}^{n=\infty} m_n}, 0\right). \tag{42}$$

The functional form of of equation 42 can be seen in Additional file 3.

Figure 8 shows the probability of finding a blood meal for the sibling species on January 1, 1999.

Figure 8 Probability of finding a blood meal for *An. arabiensis* ($HBI = 0.4$) and *An. gambiae s.s.* ($HBI = 0.95$) with zero bed net coverage.

Results and discussion

Sensitivity experiments

Sensitivity experiments are useful in understanding which parameters are important for the success of *An. arabiensis* and *An. gambiae s.s.* and which are important for malaria transmission. Classical sensitivity analysis investigates the robustness of a study when parameters are estimated from statistical modelling. Our model uses parametrization schemes to represent the influence of the environment on the two species. We show how the model responds to changing temperature, humidity, mosquito size, dispersion and the probability of finding blood. This approach does not allow us to directly measure the robustness of each parametrization scheme, but gives us an insight into which external factors influence the model and where it is of importance to have improved parametrization schemes. We use the term sensitivity experiments for this analysis.

Settings

To demonstrate some of the capabilities of the model, we set up a series of experiments. Some aspects are best visualized as a one-dimensional model (time and age), while other features are shown using a spatial domain (time, age, and space). For the one-dimensional experiments, the water temperature is set to the air temperature, except for temperature greater than $33°C$, for which we set temperature to $33°C$. This modification is required since pupae and fourth instar larvae will not develop below $18°C$ or above $34°C$ [144]. The results are therefore less robust when temperature is greater than $33°C$. Unless otherwise stated, we use size-dependent mortality, correction for indoor temperature, the KBLL method to estimate mortality in the aquatic stages, correction for the development rate in the aquatic stages depending on the ratio of each species, and movement of mosquitoes (in the spatial cases).

Sensitivity to temperature, relative humidity and mosquito size (TempHumSize)

The age-dependent mortality is influenced by temperature, relative humidity and mosquito size [Eq. (32)]. This experiment explores how the dynamics of malaria is sensitive to temperature, relative humidity and mosquito size (measured as mm). We assume that no births occur to isolate the effect of the transmission process, and consequently constant mosquito body size in the course of integration, but include mortality and the biting rate. In this experiment we assume that only one species is present (since the main competition occurs in the aquatic stages). This sensitivity test is designed to observe how the proportion of mosquitoes becomes infected as a function of temperature, relative humidity and mosquito size, given that we start with 1000 newly emerged mosquitoes, with $m_1 = 1000$ and $m_{2-9} = 0$ as the initial conditions. In this experiment, 1% of the human population is infectious for *Plasmodium falciparum*. Mosquitoes are infected with an efficiency of 100%, meaning that biting an infectious human results in gametocyte transmission to the mosquito. In practice, this would be the same as saying that 10% of humans were infectious and gametocyte transmission had an efficiency of 10%. We also neglect the effect of heterogeneous biting. This is the only experiment in which we model the proportion of infectious mosquitoes explicitly. The modified equations describing the transmission process are described in [64].

The rate of sporozoite development within mosquitoes is expressed as [5]

$$pf = \left(a + \frac{b}{e^{(T_{air})^c - d}} \right)^{-1}, \tag{43}$$

where a = 9.5907, $b = 0.0051029$, $c = 0.7349$, and $d = 17.0325$. This expression was derived from the figure in MacDonald page 119 [5] using g3data [113], and fitted using non-linear least-squares [145].

The gonotrophic cycle and biting rate are defined in Eq. (26).

The integrations are repeated with different combinations of temperature and relative humidity. This is a simple representation of gametocyte transmission to mosquitoes and is an idealized approach for exploring the proportion of mosquitoes (of the original 1000) that would become infected under different temperature, RH and mosquito size. Figure 9 shows how the percentage of infectious mosquitoes changes with temperature, RH and mosquito size. Lyimo and Koella reported that the largest mosquitoes were less likely to have sporozoites, but had more oocysts than smaller mosquitoes [54]. They attributed this to increased mortality in the presence of many oocysts, an effect that is not included in our model. Figure 9 shows that the potential percentage of infected mosquitoes is sensitive to all three parameters in the model. Although higher survival has been attributed to body size in dry [7, 51] and semi-arid environments [55], the advantage or disadvantage of a larger body has been poorly described in saturated environments. Therefore, the sensitivity to body size at 80% RH should be interpreted with care. According to the model, temperature is not the only factor that governs the transmission of malaria (in areas with no interventions); humidity and how mosquitoes adapt to dehydration stress are also important factors. The most efficient transmission, expressed as the integral, with respect to days, occurs at $25°C$ at 40% and 80% RH, and at $24.5°C$ at 10% RH, independent of mosquito size.

Figure 9 Percentage of 1000 mosquitoes that are infectious after x days. The y-axis represents temperature in degrees centigrade. The model is integrated at two mosquito sizes (2.8 and 3.2 mm for wing length) and three relative humidity values.

These results should be viewed in light of recent findings by Paaijmans et al. that optimal transmission occurs at lower temperatures [4].

Sensitivity to temperature and carrying capacity (TempCar)

The aim of this sensitivity test was to investigate how carrying capacity and temperature determine the relative proportion of *An. arabiensis* and *An. gambiae s.s.* We set the relative humidity to 80% and the probability of getting a blood meal to one. We assumed that the soil was saturated and we varied the temperature between 16 and $38°C$ (with corrections over $33°C$ for water temperature) and the carrying capacity between 0.0625 and $125 mgkm^{-2}$.

Carrying capacity in the aquatic stages influences larval growth and adult survival. While *An. arabiensis* invests more time in growth than *An. gambiae s.s.*, the former develops a larger body, and consequently has the potential to oviposit more eggs than the latter. If the two species experience the same mortality rate in the aquatic stages, more *An. gambiae s.s.* will emerge, but over time *An. arabiensis* can face this challenge by outnumbering the eggs of *An. gambiae s.s.* in the habitat. Thus, we are interested in testing how the carrying capacity in the aquatic stages alters the relative proportion of each of the adult species. In this model we only consider the competition between these two species, and hence neglect other competing species [146].

As observed in Figure 10, *An. gambiae s.s.* dominates between 27 and $30°C$. This is the effect of the development rate modifications described by Kirby et al. [72] and Paaijmans et al. [8] (Figure 2 and "Species-dependent mortality (KBLL)"). Interestingly, the dominance of *An. arabiensis* is most

pronounced in the drier simulations, meaning that high competition, compared to adult survival, is favourable for this species. This can be attributed to the strategy of larger body size and higher egg production. Lehmann et al. found that *An. arabiensis* dominated during the dry season, while *An. gambiae s.s.* dominated in the rainy season [57]. The advantage of *An. arabiensis* in crowded breeding places might be one factor contributing to the shift in species composition as the surface area of puddles starts to shrink.

Figure 10 Fraction of *An. arabiensis* as a function of air temperature and carrying capacity. The water temperature is set to the same value as the air temperature unless the temperature is greater than $33°C$ (at which most pupae would not develop into adults [144]). In this case the water temperature is set to $33°C$, but the productivity will remain low. The fraction of *An. gambiae s.s.* is one minus the fraction for *An. arabiensis*.

Sensitivity to temperature and the probability of finding blood (pBlood1D)

This experiment shows how the model responds to changes in the probability of finding a blood meal, which influences the rate at which mosquitoes can oviposit and increases energy consumption if hosts are hard to locate. If, for example, cattle are easier to find compared to humans, *An. arabiensis* will potentially use less energy per batch of eggs and will also be able to utilize breeding sites at a higher rate than *An. gambiae s.s.* It is also possible that *An. arabiensis* uses cattle for navigation [147]. Over time, such differences might lead to dominance by one species. In this experiment, we varied the probability of finding blood, $P(B)$, for *An. arabiensis* from zero to one, as well as varying the temperature as described for TempCar.

We set the probability of finding blood to one for *An. gambiae s.s.*, independent of the probability of *An. arabiensis* finding a blood meal. This is a purely theoretical experiment designed to demonstrate a concept. The probability of finding blood is varied between zero and one for *An. arabiensis*. The scenario in which $P(B) = 1$ for *An. gambiae s.s.* and zero for *An. arabiensis* is not a realistic scenario, but the difference in $P(B)$ is grounded in differences in their feeding behaviour, whereby *An. arabiensis* can utilize cattle more efficiently than *An. gambiae s.s.*, for example.

Figure 11 shows the relative fraction of *An. arabiensis*. In addition to the pattern observed in Figure 10, it is also evident that if $P(B)$ is low for *An. arabiensis*, *An. gambiae s.s.* dominates. $P(B)$ can be interpreted as a parameter that describes both the probability of finding blood for reproduction and survival, and the energy spent in the search for a blood meal. For example, easy access to cattle might give *An. arabiensis* an advantage in exploiting breeding sites, which could lead to suppression of the number of *An. gambiae s.s.* if increased use of bed nets reduces the effective human population or causes higher mortality of anthropophilic species. This mechanism might help to explain the decline in *An. gambiae s.s.* observed by Bayoh et al. [9].

Figure 11 Fraction of *An. arabiensis* as a function of air temperature and probability of finding a blood meal. The fraction of *An. gambiae s.s.* is one minus the fraction for *An. arabiensis*.

Sensitivity to the probability of finding blood in a spatial domain (pBlood2D)

This experiment is similar to pBlood1D, but this time we integrate the model for 5 years over the African domain. The experiment consists of two runs, for which the first has $P(B)$ similar to Figure 8 and the second has $P(B) = 1$ over all land areas for both species. The population density is space-invariant at 400 $humans/km^2$ (remember that the number of mosquitoes is limited by the number of hosts). Thus, the only limitation in this experiment is the physical environment (air and water temperatures, relative humidity, wind and run-off), which is updated every 3 h. The initial conditions for the mosquito populations were the same for the two runs.

Even though we have stated that the probability of finding blood $P(B)$ is an expression of the cost of finding a host, it might well be that $P(B)$ also includes a component that describes the environment shaped by cattle and humans. Therefore, it should be noted that it is difficult to distinguish between the true probability of finding blood and the environmental changes caused by the presence of humans or cattle.

Under the scenario of equal probability of finding blood for the two species, *An. gambiae s.s.* loses the competition after 5 years (Figures 12 and 13), probably because of the greater reproductive potential of *An. arabiensis*. The only strongholds left for this species are DRC, Congo, and Gabon. Hence, the strategy of *An. arabiensis* to develop a larger body, produce more eggs, and possibly reduce adult mortality at the cost of spending more time in the aquatic stages is successful when access to blood is unlimited. *An. arabiensis* has extended its distribution as far north as the southern tip of Western Sahara. While the original set-up of the model (P0) limits the distribution of *An. gambiae s.l.* to approximately $17°N$ in the Sahel, the experiment with $P(B) = 1$ (P1) has a distribution up to $22°N$ in Mali, Niger, Chad and Sudan. This is in line with observations of the northerly limit of *An. gambiae s.l.* [148–150]. The lack of *An. gambiae s.l.* north of 17° in the original set-up (P0) might be a result of the way the model is formulated. The population density is calculated within a box of approximately $50km \times 50km$. It might well be the case that pockets of higher population/cattle densities within this box could sustain a mosquito population. This is not resolved in the model. It is also worth mentioning the study by De Meillon [151] of the anophelines of Namibia, which revealed that *An. gambiae s.l.* is present in large parts of the country. The original set-up (P0) allows sustainable mosquito populations in Namibia, while the density of *An. gambiae s.l.* in P1 is more comparable to the observations of De Meillon. The problems of capturing the distribution of *An. gambiae s.l.* in Namibia may originate from the problems of resolving pockets of high host density or changes in cattle density and distribution at the time of the study compared to the present day [43, 44, 152].

Figure 12 Square root of number of *An. arabiensis* per km^2 in the two pBlood2D experiments. In P0 we used realistic values of the probability of finding blood, $P(B)$, while in P1 the probability of finding blood was set to 1, independent of the location. See the text for further details.

Figure 13 Square root of the number of *An. gambiae s.s.* per km^2 in the two pBlood2D experiments. In P0 we used realistic values of the probability of finding blood, $P(B)$, while in P1 the probability of finding blood was set to 1, independent of the location. See the text for further details.

It is also worth mentioning that the density of *An. gambiae s.l.* in South Africa is not very sensitive to the probability of finding a blood meal. Hence, the distribution of *An. gambiae s.l.* is mainly restricted by climate according to the model.

Figures 12 and 13 show the distribution and density of *An. arabiensis* and *An. gambiae s.s.* under realistic (P0) and space-invariant (P1) $P(B)$ after 6, 12, and 18 months. The integration was started on January 1 and the model was run for 5 years.

Mosquito transport (mosqTran)

The purpose of this experiment was to demonstrate how the initial conditions and competition influence the distribution of *An. gambiae s.s.* and *An. arabiensis*. To explore the theoretical dispersion distance and the influence of the initial conditions, we set up a simple experiment. In mosqTran(a) the model was initialized with *An. arabiensis* at $-4.494381°E$, $14.0154°N$ (Sahel), and *An. gambiae s.s.* at $-4.494381°E$, $6.502846°N$ (Côte d'Ivore, Ivory Coast) on January 1, 1989. The second experiment, mosqTran(b), had the same setup, but without *An. arabiensis*.

The purpose of this demonstration was to show the importance of mosquito movement and how new areas can or cannot be colonized. In a model in which movement is restricted, the vector range would also be restricted by the initial model conditions. For example, if only one point was specified for mosquitoes at the beginning of the integration, only the same point would have mosquitoes after 100 years. With dynamic movement the mosquitoes could colonize new areas if the environmental conditions, or the probability of finding blood, change over time.

Figure 14 shows the relative difference in *An. gambiae s.s.* distribution in the two experiments. It is evident that in the presence of *An. arabiensis*, *An. gambiae s.s.* fails to colonize large parts of Mali and Burkina Faso. It can be argued that this is not a result of the initial conditions, but of competition. Additional file 4 illustrates why this is indeed a result of the initial conditions, although the initial conditions would not play a role in the absence of competition.

Figure 14 Relative change in dispersal (mean over 5 years) for *An. gambiae s.s.* with (mosqTran(a)) and without (mosqTran(b)) *An. arabiensis*. The black solid circle and triangle indicate the initial position of *An. arabiensis* and *An. gambiae s.s.*, respectively.

Figure 15 shows the number of months required to reach a density of $20 mosquitoes/km^2$. It is interesting to note that dispersal occurs in pulses. The dispersal of *An. arabiensis* is slower than that of *An. gambiae s.s.*, probably because of the drier conditions in the Sahel and *An. gambiae s.s.* reached the area before *An. arabiensis* (Figure 15). The simulations show that establishment in an already occupied area is a much slower process compared to the case of no competition. From the simulations we can also speculate on whether the dominance of one species can act as a barrier to genetic flow, like a mountain range or dessert. This also raises some questions regarding whether hibernation or dispersal is the mechanism behind the dominance of the *An. gambiae s.s.* M form in parts of Mali. Although there are strong indications that the *An. gambiae s.s.* M undergoes aestivation during the dry season [57, 58, 71], it is also possible that the persistence of the *An. gambiae s.s.* M form in the Niono district in Mali can serve as a refuge during the dry season [153]. In both cases the M form receives a kick-start at the beginning of the rainy season, and might slow down the dispersal of *An. arabiensis* and the S form of *An. gambiae s.s.* A similar mechanism could contribute to the dominance of *An. arabiensis* in Ethiopia in the Turkana district, where the presence of *An. arabiensis* prevents rapid invasion by *An. gambiae s.s.*

Figure 15 Number of months required to reach a density of $20 mosquitoes/km^2$**.** Panel 1 (left to right) represents *An. arabiensis* in experiment mosqTran(a), panel 2, *An. gambiae s.s.* with the presence of *An. arabiensis* (mosqTran(a)), and panel 3, *An. gambiae s.s.* with no competition (mosqTran(b)). The red solid circle and triangle indicate the initial position of *An. arabiensis* and *An. gambiae s.s.*, respectively.

Conclusions

We developed a model to predict the presence and abundance of *An. arabiensis* and *An. gambiae s.s.* The model is age-structured and includes mosquito dispersal.

Sensitivity tests showed that as well as temperature, relative humidity and mosquito size are important factors in malaria transmission. The result for body size is in line with several studies [7, 51, 54, 55, 88, 154] and thus the model captures some of the aspects related to higher survival among larger individuals. Note that we have not accounted for the higher metabolism in large mosquitoes [71], which might reduce survival under warm and dry conditions. There are also contrasting results with respect to body size and egg production [155]. It is likely that there is an optimum size that depends on the environment and is a function of temperature and humidity. Currently there are few results to back up this statement. However, Sanford et al. found significant differences in *Anopheles gambiae s.s.* wing length between Mali and Guinea-Bissau [156].

We show that relative humidity can be important for malaria transmission. Several models have neglected the role of (relative) humidity [29, 157] and it is true that desiccation might not be a driver of mortality at moderate humidity (>70%?). The main argument for leaving out this parameter is the corresponding reduction in model complexity. As long as rainfall drive the carrying capacity, mosquito numbers will be restricted at lower humidity (no rain), and as a consequence the resulting number of mosquitoes can be limited for the wrong reasons, but with the correct result. For example, Ermert et al. [28] handle this deficiency by reducing vector survival during dry atmospheric conditions, defined as a function of 10-day accumulated rainfall. More studies on the survival of *An. gambiae s.l.* in relation to size and relative humidity in the range 5-40% are needed for more confidence in the role of humidity in the survival of *An. gambiae s.l.*

Assumption of exponential mortality has several advantages (see Figure 5 for examples of models in which exponential mortality is used). The model becomes fast to solve and it is easier to analyse the equations analytically. However, several studies have shown that mortality of *An. gambiae s.l.* is not exponential, and that inclusion of an age dimension alters the expected outcome of interventions targeted to reduce the vector population [50]. Therefore, we believe that models in which age-dependent mortality is assumed should be further explored. The sensitivity tests also suggest that carrying capacity within a restricted area plays a role in the distribution of *An. arabiensis* and *An. gambiae s.s.* The true carrying capacity is hard to estimate on a continental scale and thus relies on qualified guesswork taking into account rainfall, groundwater and soil saturation, for example. Carrying capacity influences not only the relative distribution of the two species but also the total number of mosquitoes. To correctly estimate the biting rate, a correct estimate of carrying capacity is required, and thus more work is needed to parametrize puddle formation. It should also be noted that no current large-scale models can describe the formation of puddles as rivers retreat, as described by Animut et al. [158].

Experiment pBlood2D showed how the model responds to the parameter $P(B)$, the probability of finding a blood meal. $P(B)$ is important in describing a realistic distribution of *An. arabiensis* and *An. gam-*

biae s.s. Thus, we hypothesize that the large-scale distribution of bovines is key to the success of *An. arabiensis*. Likewise, large-scale human density favours the presence of *An. gambiae s.s.*

Finally, experiment mosqTran showed how the initial conditions influence the dispersal of *An. gambiae s.s.* (and *An. arabiensis*). The distribution of *An. gambiae s.s.* changes dramatically with the presence of *An. arabiensis*, and thus the initial model conditions are highly relevant for correct description of the distribution of the two species. When rainfall is highly seasonal, the first come, first served principle seems to be important for the success of a species in drier conditions. Whether or not this plays a role in the evolution of aestivation in *An. gambiae s.s.* M form [57] is a question that should be further investigated.

The strong influence of initial conditions on dispersal of the *An. gambiae* complex is not irrelevant when assessing the impact of climate change, since vectorial capacity varies between species.

The availability of mosquito models allows researchers to build on and improve our understanding of the role of the *An. gambiae* complex in malaria transmission. We hope to refine the model as new data on mosquito biology become available, and to incorporate the effects of interventions.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The work presented here was carried out as a collaboration between all the authors. BL and AS defined the research theme. TML designed methods and mosquito experiments, performed the model runs, analysed the data, interpreted the results and wrote the paper. DK, AS and TML designed and evaluated the regional climate simulations. EL provided input with respect to the malaria situation in Ethiopia, which in turn was used in the model formulation. All the authors have contributed to, seen and approved the manuscript.

Acknowledgements

We are grateful to the National Center for Atmospheric Research (NCAR) for making their WRF model available in the public domain. We also thank the Bergen Centre for Computational Science for computational and other resources provided during this study. This work was made possible by grants from The Norwegian Programme for Development, Research and Education (NUFU) and the University of Bergen. We thank Steve Lindsay for providing data on the survival of *An. gambiae s.s.* under different temperatures and relative humidities. We thank two anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

References

1. Greenwood B, Mutabingwa T: **Malaria in 2002.** *Nature* 2002, **415:**670–672. [http://www.ncbi.nlm.nih.gov/pubmed/11832954]

2. World Health Organization: *World Malaria Report 2011*. Switzerland: World Health Organization; 2011.

3. Rogers DJ, Randolph SE: **The global spread of malaria in a future, warmer world.** *Science* 2000, **289:**1763–1766. [http://www.ncbi.nlm.nih.gov/pubmed/10976072]

4. Paaijmans KP, Blanford S, Chan BHK, Thomas MB: **Warmer temperatures reduce the vectorial capacity of malaria mosquitoes.** *Biol Lett* 2012, **8:**465–468. [http://www.ncbi.nlm.nih.gov/pubmed/22188673]

5. MacDonald G: *Dynamics of Tropical Disease*. London: Oxford University Press; 1973.

6. Bomblies A, Duchemin JB, Eltahir EAB: **Hydrology of malaria: Model development and application to a Sahelian village.** *Water Resour Res* 2008, **44:**W12445.

7. Gray EM, Bradley TJ: **Physiology of desiccation resistance in *Anopheles gambiae* and *Anopheles arabiensis*.** *Am J Trop Med Hyg* 2005, **73:**553–559. [http://www.ajtmh.org/cgi/content/abstract/73/3/553]

8. Paaijmans KP, Huijben S, Githeko AK, Takken W: **Competitive interactions between larvae of the malaria mosquitoes *Anopheles arabiensis* and *Anopheles gambiae* under semi-field conditions in western Kenya.** *Acta Trop* 2009, **109:**124–130. [http://www.ncbi.nlm.nih.gov/pubmed/18760989]

9. Bayoh MN, Mathias DK, Odiere MR, Mutuku FM, Kamau L, Gimnig JE, Vulule JM, Hawley WA, Hamel MJ, Walker ED: ***Anopheles gambiae:* historical population decline associated with regional distribution of insecticide-treated bed nets in western Nyanza Province, Kenya.** *Malar J* 2010, **9:**62. [http://www.ncbi.nlm.nih.gov/pubmed/20187956]

10. Intergovernmental Panel on Climate Change: *Fourth Assessment Report: Climate Change 2007: Working Group I Report: The Physical Science Basis*. Geneva: IPCC; 2007. [http://www.ipcc.ch/ipccreports/ar4-wg1.htm]

11. Capanna E: **Grassi versus Ross: who solved the riddle of malaria?** *Int Microbiol* 2006, **9:**69–74. [http://www.ncbi.nlm.nih.gov/pubmed/16636993]

12. Ross R: **Life-History of the parasites of malaria.** *Nature* 1899, **60:**322–324.

13. Waite H: **Mosquitoes and malaria. A study of the relation between the number of mosquitoes in a locality and the malaria rate.** *Biometrika* 1910, **7:**421–436. [http://www.jstor.org/stable/2345376]

14. Lotka AJ: **Contribution to the analysis of malaria epidemiology. I. General part.** *Am J Epidemiol* 1923, **3:**1–36. [http://aje.oxfordjournals.org]

15. Lotka AJ: **Contribution to the analysis of malaria epidemiology. II. General part (continued). Comparison of two formulae given by sir Ronald Ross.** *Am J Epidemiol* 1923, **3:**38–54. [http://aje.oxfordjournals.org]

16. Lotka AJ: **Contribution to the analysis of malaria epidemiology. III. Numerical part.** *Am J Epidemiol* 1923, **3:**55–95. [http://aje.oxfordjournals.org]

17. Lotka AJ, Sharpe FR: **Contribution to the analysis of malaria epidemiology. IV. Incubation lag.** *Am J Epidemiol* 1923, **3:**96–112. [http://aje.oxfordjournals.org]

18. Ross R: *Report on the Prevention of Malaria in Mauritius*. London: Waterlow and Sons Limited; 1908.

19. Ross R: *The Prevention of Malaria*. London: Murray; 1911.

20. MacDonald G: *The Epidemiology and Control of Malaria*. London: Oxford University Press; 1957.

21. Yang HM: **Malaria transmission model for different levels of acquired immunity and temperature-dependent parameters (vector).** *Rev Saude Publica* 2000, **34:**223–231. [http://www.ncbi.nlm.nih.gov/pubmed/10920443]

22. Depinay JMO, Mbogo CM, Killeen G, Knols B, Beier J, Carlson J, Dushoff J, Billingsley P, Mwambi H, Githure J, Toure AM, McKenzie FE: **A simulation model of African *Anopheles* ecology and population dynamics for the analysis of malaria transmission.** *Malar J* 2004, **3:**29. [http://www.ncbi.nlm.nih.gov/pubmed/15285781]

23. Ruiz D, Poveda G, Velez ID, Quinones ML, Rua GL, Velasquez LE, Zuluaga JS: **Modelling entomological-climatic interactions of *Plasmodium falciparum* malaria transmission in two Colombian endemic-regions: contributions to a National Malaria Early Warning System.** *Malar J* 2006, **5:**66. [http://www.ncbi.nlm.nih.gov/pubmed/16882349]

24. Pascual M, Ahumada JA, Chaves LF, Rodo X, Bouma M: **Malaria resurgence in the East African highlands: temperature trends revisited.** *Proc Natl Acad Sci U S A* 2006, **103:**5829–5834. [http://www.ncbi.nlm.nih.gov/pubmed/16571662]

25. Bomblies A, Duchemin JB, Eltahir EAB: **A mechanistic approach for accurate simulation of village scale malaria transmission.** *Malar J* 2009, **8:**223. [http://www.ncbi.nlm.nih.gov/pubmed/19799793]

26. Parham PE, Michael E: **Modeling the effects of weather and climate change on malaria transmission.** *Environ Health Perspect* 2010, **118:**620–626. [http://www.ncbi.nlm.nih.gov/pubmed/20435552]

27. Hoshen MB, Morse AP: **A weather-driven model of malaria transmission.** *Malar J* 2004, **3:**32. [http://www.ncbi.nlm.nih.gov/pubmed/15350206]

28. Ermert V, Fink AH, Jones AE, Morse AP: **Development of a new version of the Liverpool Malaria Model. II. Calibration and validation for West Africa.** *Malar J* 2011, **10:**62. [http://www.ncbi.nlm.nih.gov/pubmed/21410939]

29. Ermert V, Fink AH, Jones AE, Morse AP: **Development of a new version of the Liverpool Malaria Model. I. Refining the parameter settings and mathematical formulation of basic processes based on a literature review.** *Malar J* 2011, **10:**35. [http://www.ncbi.nlm.nih.gov/pubmed/21314922]

30. White LJ, Maude RJ, Pongtavornpinyo W, Saralamba S, Aguas R, Van Effelterre T, Day NPJ, White NJ: **The role of simple mathematical models in malaria elimination strategy design.** *Malar J* 2009, **8:**212. [http://www.ncbi.nlm.nih.gov/pubmed/19747403]

31. Mordecai EA, Paaijmans KP, Johnson LR, Balzer C, Ben-Horin T, de Moor E, McNally A, Pawar S, Ryan SJ, Smith TC, Lafferty KD, Thrall P: **Optimal temperature for malaria transmission is dramatically lower than previously predicted.** *Ecol Lett* 2012. [http://www.ncbi.nlm.nih.gov/pubmed/23050931]

32. Smith DL, Hay SI, Noor AM, Snow RW: **Predicting changing malaria risk after expanded insecticide-treated net coverage in Africa.** *Trends Parasitol* 2009, **25**(11):511–516. [http://www.ncbi.nlm.nih.gov/pubmed/19744887]

33. Eckhoff PA: **A malaria transmission-directed model of mosquito life cycle and ecology.** *Malar J* 2011, **10:**303. [http://www.ncbi.nlm.nih.gov/pubmed/21999664]

34. Parham PE, Michael E: **Modeling the effects of weather and climate change on malaria transmission.** *Environ Health Perspect* 2010, **118:**620–626. [http://www.ncbi.nlm.nih.gov/pubmed/20435552]

35. Ermert V, Fink AH, Morse AP, Paeth H: **The impact of regional climate change on malaria risk due to greenhouse forcing and land-use changes in tropical Africa.** *Environ Health Perspect* 2012, **120:**77–84. [http://www.ncbi.nlm.nih.gov/pubmed/21900078].

36. Tanser F, Sharp B: **Global climate change and malaria.** *Lancet Infect Dis* 2005, **5:**256–258. [http://www.ncbi.nlm.nih.gov/pubmed/15854877]

37. Gurarie D, Karl S, Zimmerman PA, King CH, St Pierre TG, Davis TME: **Mathematical modelling of malaria infection with innate and adaptive immunity in individuals and agent-based communities.** *PLoS One* 2012, **7**(3):e34040. [http://www.ncbi.nlm.nih.gov/pubmed/22470511]

38. Ducrot A, Sirima SB, Somé B, Zongo P: **A mathematical model for malaria involving differential susceptibility, exposedness and infectivity of human host.** *J Biol Dyn* 2009, **3:**574–598. [http://www.ncbi.nlm.nih.gov/pubmed/22880962]

39. malERA Consultative Group on Modeling: **A research agenda for malaria eradication: modelling.** *PLoS Med* 2011, **8:**e1000403. [http://www.ncbi.nlm.nih.gov/pubmed/21283605]

40. Boni MF, Buckee CO, White NJ: **Mathematical models for a new era of malaria eradication.** *PLoS Med* 2008, **5:**e231. [http://www.ncbi.nlm.nih.gov/pubmed/19067482]

41. Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Wang W, Powers JG: **A Description of the Advanced Research WRF Version 2**. Tech. rep., The National Center for Atmospheric Research 2005.

42. Center for International Earth Science Information Network (CIESIN) - Columbia University; and Centro Internacional de Agricultura Tropical (CIAT): **Gridded Population of the World Version 3 (GPWv3): population density grids**. Tech. rep., Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University 2005. [http://sedac.ciesin.columbia.edu/gpw]

43. Robinson T, Fao's Animal Production and Health Division: **Observed livestock densities.** 2011, [http://www.fao.org/AG/againfo/resources/en/glw/GLW_dens.html]

44. Wint W, Robinson T: **Gridded livestock of the world - 2007.** Tech. rep., Food and agriculture organization of the United Nations, Rome 2007.

45. Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, Ghani A, Drakeley C, Gosling R: **Hitting hotspots: spatial targeting of malaria for control and elimination.** *PLoS Med* 2012, **9:**e1001165. [http://www.ncbi.nlm.nih.gov/pubmed/22303287]

46. World Climate Research Programme: **Cordex.** 2012, [http://wcrp.ipsl.jussieu.fr/cordex/about.html]

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

48. Afrane YA, Zhou G, Lawson BW, Githeko AK, Yan G: **Effects of microclimatic changes caused by deforestation on the survivorship and reproductive fitness of Anopheles gambiae in western Kenya highlands.** *Am J Trop Med Hyg* 2006, **74:**772–778. [http://www.ncbi.nlm.nih.gov/pubmed/16687679]

49. Harrington LC, Vermeylen F, Jones JJ, Kitthawee S, Sithiprasasna R, Edman JD, Scott TW: **Age-dependent survival of the dengue vector *Aedes aegypti* (Diptera: Culicidae) demonstrated by simultaneous release-recapture of different age cohorts.** *J Med Entomol* 2008, **45:**307–313. [http://www.ncbi.nlm.nih.gov/pubmed/18402147]

50. Bellan SE: **The importance of age dependent mortality and the extrinsic incubation period in models of mosquito-borne disease transmission and control.** *PLoS One* 2010, **5:**e10165. [http://www.ncbi.nlm.nih.gov/pubmed/20405010]

51. Fouet C, Gray E, Besansky NJ, Costantini C: **Adaptation to aridity in the malaria mosquito *Anopheles gambiae:* chromosomal inversion polymorphism and body size influence resistance to desiccation.** *PLoS One* 2012, **7:**e34841. [http://www.ncbi.nlm.nih.gov/pubmed/22514674]

52. Shoukry A: **Cannibalism in Anopheles-Pharoensis Theo.** *Experientia* 1980, **36:**308–309.

53. Muriu SM, Coulson T, Mbogo CM, Godfray HCJ: **Larval density dependence in *Anopheles gambiae* s.s., the major African vector of malaria.** *J Anim Ecol* 2012. [http://www.ncbi.nlm.nih.gov/pubmed/23163565]

54. Lyimo EO, Koella JC: **Relationship between body size of adult *Anopheles gambiae* s.l. and infection with the malaria parasite *Plasmodium falciparum*.** *Parasitology* 1992, **104**(t 2):233–237. [http://www.ncbi.nlm.nih.gov/pubmed/1594289]

55. Ameneshewa B, Service MW: **The relationship between female body size and survival rate of the malaria vector *Anopheles arabiensis* in Ethiopia.** *Med Vet Entomol* 1996, **10:**170–172. [http://www.ncbi.nlm.nih.gov/pubmed/8744710].

56. Maciel-De-Freitas R, Codeco CT, Lourenco-De-Oliveira R: **Body size-associated survival and dispersal rates of *Aedes aegypti* in Rio de Janeiro.** *Med Vet Entomol* 2007, **21:**284–292. [http://www.ncbi.nlm.nih.gov/pubmed/17897370]

57. Lehmann T, Dao A, Yaro AS, Adamou A, Kassogue Y, Diallo M, Sékou T, Coscaron-Arias C: **Aestivation of the African malaria mosquito, *Anopheles gambiae* in the Sahel.** *Am J Trop Med Hyg* 2010, **83:**601–606. [http://www.ncbi.nlm.nih.gov/pubmed/20810827]

58. Adamou A, Dao A, Timbine S, Kassogue Y, Yaro AS, Diallo M, Traore SF, Huestis DL, Lehmann T: **The contribution of aestivating mosquitoes to the persistence of *Anopheles gambiae* in the Sahel.** *Malar J* 2011, **10:**151. [http://www.ncbi.nlm.nih.gov/pubmed/21645385]

59. Gimonneau G, Pombi M, Choisy M, Morand S, Dabiré RK, Simard F: **Larval habitat segregation between the molecular forms of the mosquito *Anopheles gambiae* in a rice field area of Burkina Faso, West Africa.** *Med Vet Entomol* 2011. [http://www.ncbi.nlm.nih.gov/pubmed/21501199]

60. Parmakelis A, Russello MA, Caccone A, Marcondes CB, Costa J, Forattini OP, Sallum MAM, Wilkerson RC, Powell JR: **Historical analysis of a near disaster: *Anopheles gambiae* in Brazil.** *Am J Trop Med Hyg* 2008, **78:**176–178. [http://www.ncbi.nlm.nih.gov/pubmed/18187802]

61. Soetaert K, Petzoldt T, Setzer RW: **Solving Differential Equations in R: Package deSolve**. *J Stat Software* 2010, **33:**1–25. [http://www.jstatsoft.org/v33/i09]

62. Hindmarsh AC: *ODEPACK, A Systematized Collection of ODE Solvers*. Amsterdam: Scientific Computing, North-Holland; 1983.

63. Petzold LR: **Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations.** *SIAM J Sci Stat Comput* 1983, **4:**136–148.

64. Lunde TM, Bayoh MN, Lindtjørn B: **How malaria models relate temperature to malaria transmission.** *Parasit Vectors* 2012, **6:**20.

65. Munga S, Minakawa N, Zhou G, Barrack OOJ, Githeko AK, Yan G: **Effects of larval competitors and predators on oviposition site selection of *Anopheles gambiae* sensu stricto.** *J Med Entomol* 2006, **43:**221–224. [http://www.ncbi.nlm.nih.gov/pubmed/16619602]

66. Omer SM, Cloudsley-Thompson JL: **Survival of female *Anopheles gambiae* Giles through a 9-month dry season in Sudan.** *Bull World Health Organ* 1970, **42:**319–330. [http://www.ncbi.nlm.nih.gov/pubmed/5310144]

67. Jawara M, Pinder M, Drakeley CJ, Nwakanma DC, Jallow E, Bogh C, Lindsay SW, Conway DJ: **Dry season ecology of *Anopheles gambiae* complex mosquitoes in The Gambia.** *Malar J* 2008, **7:**156. [http://www.ncbi.nlm.nih.gov/pubmed/18710559]

68. Gilles JRL, Lees RS, Soliban SM, Benedict MQ: **Density-dependent effects in experimental larval populations of *Anopheles arabiensis* (Diptera: Culicidae) can be negative, neutral, or overcompensatory depending on density and diet levels.** *J Med Entomol* 2011, **48:**296–304. [http://www.ncbi.nlm.nih.gov/pubmed/21485365]

69. Lindsay SW, Parson L, Thomas CJ: **Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae* sensu stricto and *An. arabiensis*, using climate data.** *Proc Biol Sci* 1998, **265:**847–854. [http://www.ncbi.nlm.nih.gov/pubmed/9633110]

70. Ye-Ebiyo Y, Pollack RJ, Kiszewski A, Spielman A: **Enhancement of development of larval *Anopheles arabiensis* by proximity to flowering maize (Zea mays) in turbid water and when crowded.** *Am J Trop Med Hyg* 2003, **68:**748–752. [http://www.ncbi.nlm.nih.gov/pubmed/12887038]

71. Huestis DL, Yaro AS, Traore AI, Adamou A, Kassogue Y, Diallo M, Timbine S, Dao A, Lehmann T: **Variation in metabolic rate of *Anopheles gambiae* and *A. arabiensis* in a Sahelian village.** *J Exp Biol* 2011, **214:**2345–2353. [http://www.ncbi.nlm.nih.gov/pubmed/21697426]

72. Kirby MJ, Lindsay SW: **Effect of temperature and inter-specific competition on the development and survival of *Anopheles gambiae* sensu stricto and *An. arabiensis* larvae.** *Acta Trop* 2009, **109:**118–123. [http://www.ncbi.nlm.nih.gov/pubmed/19013420]

73. Reichardt, K, Angelocci LR, Bacchi OOS, Pilotto JE: **Daily rainfall variability at a local scale (1,000 ha), in Piracicaba, SP, Brazil, and its implications on soil water recharge**. *Scientia Agricola* 1995, **52:**43–49. [http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-90161995000100008&nrm=iso]

74. Bitew MM, Gebremichael M: **Spatial variability of daily summer rainfall at a local-scale in a mountainous terrain and humid tropical region.** *Atmospheric Research* 2010, **98:**347–352. [http://www.sciencedirect.com/science/article/pii/S0169809510001833]. [International Conference on Nucleation and Atmospheric Aerosols (Part 1) - ICNAA 2009].

75. Sivakumar MVK, Hatfield JL: **Spatial variability of rainfall at an experimental station in Niger, West Africa.** *Theor Appl Climatology* 1990, **42:**33–39. [http://dx.doi.org/10.1007/BF00865524]. [10.1007/BF00865524].

76. Kain J, Fritsch J: **Convective parameterization for mesoscale models: The Kain–Fritsch scheme.** *Representation Cumulus Convection Numerical Models Metor Monogr, Am Meteor Soc* 1993, **46:**165–170.

77. Kain J: **The Kain–Fritsch convective parameterization: An update**. *J Appl Meteorolgy* 2004, **43:**170–181.

78. Taylor KE, Stouffer RJ, Meehl GA: **An overview of CMIP5 and the experiment design.** *Bull Am Meteor Soc* 2011, **93:**485–498. doi: 10.1175/BAMS-D-11-00094.1.

79. Feser F, Rockel B, von Storch H, Winterfeldt J, Zahn M: **Regional climate models add value to global model data: a review and selected examples.** *Bull Am Meteor Soc* 2011, **92:**1181–1192. doi: 10.1175/2011BAMS3061.1.

80. Sharon D: **The spatial pattern of convective rainfall in Sukumaland, Tanzania – a statistical analysis.** *Theoretical Appl Climatology* 1974, **22:**201–218. [http://dx.doi.org/10.1007/BF02243468]. [10.1007/BF02243468]

81. Fiener P, Auerswald K: **Spatial variability of rainfall on a sub-kilometre scale.** *Earth Surf Processes Landforms* 2009, **34:**848–859. [http://dx.doi.org/10.1002/esp.1779]

82. Chen CT, Knutson T: **On the verification and comparison of extreme rainfall indices from climate models.** *J Climate* 2008, **21:**1605–1621. [http://journals.ametsoc.org/doi/abs/10.1175/2007JCLI1494.1]

83. Paaijmans KP, Wandago MO, Githeko AK, Takken W: **Unexpected high losses of *Anopheles gambiae* larvae due to rainfall.** *PLoS One* 2007, **2:**e1146. [http://www.ncbi.nlm.nih.gov/pubmed/17987125]

84. Ramasamy, Srikrishnaraj, Hadjirin, Perera, Ramasamy: **Physiological aspects of multiple blood feeding in the malaria vector *Anopheles tessellatus*.** *J Insect Physiol* 2000, **46:**1051–1059. [http://www.ncbi.nlm.nih.gov/pubmed/10802118]

85. Lyimo EO, Takken W: **Effects of adult body size on fecundity and the pre-gravid rate of *Anopheles gambiae* females in Tanzania.** *Med Vet Entomol* 1993, **7:**328–332. [http://www.ncbi.nlm.nih.gov/pubmed/8268486]

86. Russell TL, Lwetoijera DW, Knols BGJ, Takken W, Killeen GF, Ferguson HM: **Linking individual phenotype to density-dependent population growth: the influence of body size on the population dynamics of malaria vectors.** *Proc Biol Sci* 2011, **278:**3142–3151. [http://www.ncbi.nlm.nih.gov/pubmed/21389034]

87. Yaro AS, Dao A, Adamou A, Crawford JE, Traore SF, Toure AM, Gwadz R, Lehmann T: **Reproductive output of female *Anopheles gambiae* (Diptera: Culicidae): comparison of molecular forms.** *J Med Entomol* 2006, **43:**833–839. [http://www.ncbi.nlm.nih.gov/pubmed/17017216]

88. Koella JC, Lyimo EO: **Variability in the relationship between weight and wing length of *Anopheles gambiae* (Diptera: Culicidae).** *J Med Entomol* 1996, **33:**261–264. [http://www.ncbi.nlm.nih.gov/pubmed/8742532]

89. Lehmann T, Dalton R, Kim EH, Dahl E, Diabate A, Dabire R, Dujardin JP: **Genetic contribution to variation in larval development time, adult size, and longevity of starved adults of *Anopheles gambiae*.** *Infect Genet Evol* 2006, **6:**410–416. [http://www.ncbi.nlm.nih.gov/pubmed/16524787]

90. Howard A, Adongo E, Vulule J, Githure J: **Effects of a botanical larvicide derived from *Azadirachta indica* (the neem tree) on oviposition behaviour in *Anopheles gambiae* ss mosquitoes.** *J Med Plants Res* 2011, **5:**1948–1954.

91. Chen F, Mitchell K, Schaake J, Xue Y, Pan H, Koren V, Duan Q, Ek M, Betts A: **Modeling of land surface evaporation by four schemes and comparison with FIFE observations.**. *J Geophysical Res-Atmospheres* 1996, **101:**7251–7268.

92. Koren V, Schaake J, Mitchell K, Duan Q, Chen F, Baker J: **A parameterization of snowpack and frozen ground intended for NCEP weather and climate models.** *J Geophysical Res-Atmospheres* 1999, **104:**19569–19585.

93. Betts A, Chen F, Mitchell K, Janjic Z: **Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta Model using FIFE data.** *Mon Weather Rev* 1997, **125:**2896–2916.

94. Ek M, Mitchell K, Lin Y, Rogers E, Grunmann P, Koren V, Gayno G, Tarpley J: **Implementation of Noah land surface model advances in the national centers for environmental prediction operational mesoscale Eta model.** *J Geophysical Res-Atmospheres* 2003, **108:**GCP 12 1–16.

95. Paaijmans KP, Heusinkveld BG, Jacobs AFG: **A simplified model to predict diurnal water temperature dynamics in a shallow tropical water pool.** *Int J Biometeorol* 2008, **52:**797–803.

96. Huang J, Walker ED, Vulule J, Miller JR: **Daily temperature profiles in and around Western Kenyan larval habitats of *Anopheles gambiae* as related to egg mortality.** *Malar J* 2006, **5:**87. [http://www.ncbi.nlm.nih.gov/pubmed/17038186]

97. Bayoh MN, Lindsay SW: **Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito *Anopheles gambiae* in the laboratory.** *Med Vet Entomol* 2004, **18:**174–179. [http://www.ncbi.nlm.nih.gov/pubmed/15189243]

98. Holstein M: *Biology of Anopheles gambiae: Research in French West Africa*, *Volume 9*. World Health Organization, Monograph Series 1954.

99. Mushinzimana E, Munga S, Minakawa N, Li L, Feng CC, Bian L, Kitron U, Schmidt C, Beck L, Zhou G, Githeko AK, Yan G: **Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands.** *Malar J* 2006, **5:**13. [http://www.ncbi.nlm.nih.gov/pubmed/16480523]

100. Lehner B, Verdin K, Jarvis A: **New global hydrography derived from spaceborne elevation data**. *Eos, Trans, Am Geophysical Union* 2008, **89:**93–94.

101. Munga S, Yakob L, Mushinzimana E, Zhou G, Ouna T, Minakawa N, Githeko A, Yan G: **Land use and land cover changes and spatiotemporal dynamics of anopheline larval habitats during a four-year period in a highland community of Africa.** *Am J Trop Med Hyg* 2009, **81:**1079–84, [http://www.ncbi.nlm.nih.gov/pubmed/19996440]

102. Baber I, Keita M, Sogoba N, Konate M, Diallo M, Doumbia S, Traore S, Ribeiro J, Manoukis NC: **Population size and migration of *Anopheles gambiae* in the Bancoumana Region of Mali and their significance for efficient vector control.** *PLoS One* 2010, **5:**e10270. [http://www.ncbi.nlm.nih.gov/pubmed/20422013]

103. Ageep TB, Cox J, Hassan MM, Knols BGJ, Benedict MQ, Malcolm CA, Babiker A, El Sayed BB: **Spatial and temporal distribution of the malaria mosquito *Anopheles arabiensis* in northern Sudan: influence of environmental factors and implications for vector control.** *Malar J* 2009, **8:**123. [http://www.ncbi.nlm.nih.gov/pubmed/19500425].

104. Lardeux FJ, Tejerina RH, Quispe V, Chavez TK: **A physiological time analysis of the duration of the gonotrophic cycle of *Anopheles pseudopunctipennis* and its implications for malaria transmission in Bolivia.** *Malar J* 2008, **7:**141. [http://www.ncbi.nlm.nih.gov/pubmed/18655724]

105. Rua GL, Quinones ML, Velez ID, Zuluaga JS, Rojas W, Poveda G, Ruiz D: **Laboratory estimation of the effects of increasing temperatures on the duration of gonotrophic cycle of *Anopheles albimanus* (Diptera: Culicidae).** *Mem Inst Oswaldo Cruz* 2005, **100:**515–520. [http://www.ncbi.nlm.nih.gov/pubmed/16184229]

106. Afrane YA, Lawson BW, Githeko AK, Yan G: **Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of *Anopheles gambiae* (Diptera: Culicidae) in western Kenya highlands.** *J Med Entomol* 2005, **42:**974–980. [http://www.ncbi.nlm.nih.gov/pubmed/16465737]

107. Maharaj R: **Life table characteristics of *Anopheles arabiensis* (Diptera: Culicidae) under simulated seasonal conditions.** *J Med Entomol* 2003, **40:**737–742. [http://www.ncbi.nlm.nih.gov/pubmed/14765646]

108. Hoshen M, Morse A: **A model structure for estimating malaria risk**. In *Environmental Change and Malaria Risk: Global and Local Implications*, *Volume 9*. Edited by Takken W, Martens P, Bogers RJ. 2005:10.

109. Dawes EJ, Churcher TS, Zhuang S, Sinden RE, Basanez MG: ***Anopheles* mortality is both age- and *Plasmodium*-density dependent: implications for malaria transmission.** *Malar J* 2009, **8:**228. [http://www.ncbi.nlm.nih.gov/pubmed/19822012]

110. Martens W: *Health Impacts of Climate Change and Ozone depletion. An Eco-Epidemiological Modelling Approach.* The Netherlands: Maastricht University Press; 1997.

111. Kirby MJ, Lindsay SW: **Responses of adult mosquitoes of two sibling species, *Anopheles arabiensis* and *A. gambiae* s.s. (Diptera: Culicidae), to high temperatures.** *Bull Entomol Res* 2004, **94:**441–448. [http://www.ncbi.nlm.nih.gov/pubmed/15385063]

112. Afrane YA, Zhou G, Lawson BW, Githeko AK, Yan G: **Life-table analysis of *Anopheles arabiensis* in western Kenya highlands: effects of land covers on larval and adult survivorship.** *Am J Trop Med Hyg* 2007, **77:**660–666. [http://www.ncbi.nlm.nih.gov/pubmed/17978067]

113. Frantz J: **g3data.** [http://www.frantz.fi/software/g3data.php]

114. Paaijmans KP, Thomas MB: **The influence of mosquito resting behaviour and associated microclimate for malaria risk.** *Malar J* 2011, **10:**183. [http://www.ncbi.nlm.nih.gov/pubmed/21736735]

115. Bodker R, Akida J, Shayo D, Kisinza W, Msangeni HA, Pedersen EM, Lindsay SW: **Relationship between altitude and intensity of malaria transmission in the Usambara Mountains, Tanzania.** *J Med Entomol* 2003, **40:**706–717. [http://www.ncbi.nlm.nih.gov/pubmed/14596287]

116. Minakawa N, Omukunda E, Zhou G, Githeko A, Yan G: **Malaria vector productivity in relation to the highland environment in Kenya.** *Am J Trop Med Hyg* 2006, **75:**448–453. [http://www.ncbi.nlm.nih.gov/pubmed/16968920]

117. Makaka G, Meyer E: **Temperature stability of traditional and low-cost modern housing in the Eastern Cape, South Africa.** *J Build Phys* 2006, **30:**71.

118. Okech BA, Gouagna LC, Knols BGJ, Kabiru EW, Killeen GF, Beier JC, Yan G, Githure JI: **Influence of indoor microclimate and diet on survival of *Anopheles gambiae* s.s. (Diptera: Culicidae) in village house conditions in western Kenya.** *Int J Trop Insect Sci* 2004, **24**(3):207–212.

119. Worrall E, Connor SJ, Thomson MC: **A model to simulate the impact of timing, coverage and transmission intensity on the effectiveness of indoor residual spraying (IRS) for malaria control.** *Trop Med Int Health* 2007, **12:**75–88. [http://www.ncbi.nlm.nih.gov/pubmed/17207151]

120. de Meillon B: **Observations on *Anopheles funestus* and *Anopheles gambiae* in the Transvaal**. *South Afr Inst Med Res* 1934, **32:**195.

121. Lou Y, Zhao XQ: **The periodic Ross–Macdonald model with diffusion and advection.** *Applicable Anal Int J* 2010, **89:**1067–1089.

122. Gourley SA, Liu R, Wu J: **Some vector borne diseases with structured host populations: Extinction and spatial spread.** *Siam J Appl Math* 2006, **67:**408–433.

123. Lewis M, Renclawowicz J, Van den Driessche P: **Traveling waves and spread rates for a West Nile virus model.** *Bull Math Biol* 2006, **68:**3–23.

124. Thomson MC, Connor SJ, Quinones ML, Jawara M, Todd J, Greenwood BM: **Movement of *Anopheles gambiae* s.l. malaria vectors between villages in The Gambia.** *Med Vet Entomol* 1995, **9:**413–419. [http://www.ncbi.nlm.nih.gov/pubmed/8541594]

125. Garrett-Jones C: **A dispersion of mosquitoes by wind.** *Nature* 1950, **165:**285. [http://www.ncbi.nlm.nih.gov/pubmed/15405802]

126. Ming JG, Jin H, Riley JR, Reynolds DR, Smith AD, Wang RL, Cheng JY, Cheng XN: **Autumn southward 'return' migration of the mosquito *Culex tritaeniorhynchus* in China.** *Med Vet Entomol* 1993, **7**(4):323–327. [http://www.ncbi.nlm.nih.gov/pubmed/8268485]

127. Garrett-Jones C: **The possibility of active long-distance migrations by *Anopheles pharoensis* Theobald.** *Bull World Health Organ* 1962, **27:**299–302.

128. Kaufmann C, Briegel H: **Flight performance of the malaria vectors *Anopheles gambiae* and *Anopheles atroparvus*.** *J Vector Ecol* 2004, **29:**140–153. [http://www.ncbi.nlm.nih.gov/pubmed/15266751]

129. Gillies MT: **Studies on the dispersion and survival of *Anopheles gambiae* Giles in East Africa, by means of marking and release experiments.** *Bull Entomol Res* 1961, **52:**99–127. [http://dx.doi.org/10.1017/S0007485300055309]

130. Costantini C, Li SG, Della Torre A, Sagnon N, Coluzzi M, Taylor CE: **Density, survival and dispersal of *Anopheles gambiae* complex mosquitoes in a west African Sudan savanna village.** *Med Vet Entomol* 1996, **10:**203–219. [http://www.ncbi.nlm.nih.gov/pubmed/8887330]

131. Gillies MT, Meillon BD: *The Anophelinae of Africa South of the Sahara (Ethiopian zoogeographical region)*. Johannesburg: The South African Institute for Medical Research; 1968.

132. de Meillon B: **A note on *Anopheles gambiae* and *Anopheles funestus* in Northern Rhodesia.** *South Afr Ins Med Res* 1937, **40:**306–317.

133. Adams P: **Some observations on the flight of stained anophelines at Nkana, Northern Rhodesia.** *Ann Trop Med Parasitol* 1940, **34:**35–43.

134. Ye-Ebiyo Y, Pollack RJ, Spielman A: **Enhanced development in nature of larval *Anopheles arabiensis* mosquitoes feeding on maize pollen.** *Am J Trop Med Hyg* 2000, **63:**90–93. [http://www.ncbi.nlm.nih.gov/pubmed/11358003]

135. Bicheron P, Defourny P, Brockmann C, Schouten L, Vancutsem C, Huc M, Bontemps S, Leroy M, Achard F, Herold M, Ranera F, Arino O: **GLOBCOVER: Products Description and Validation Report**. Tech. rep., MEDIAS France, Toulouse 2008.

136. Bouma M, Rowland M: **Failure of passive zooprophylaxis: cattle ownership in Pakistan is associated with a higher prevalence of malaria.** *Trans R Soc Trop Med Hyg* 1995, **89:**351–353. [http://www.ncbi.nlm.nih.gov/pubmed/7570859]

137. Okech BA, Gouagna LC, Killeen GF, Knols BGJ, Kabiru EW, Beier JC, Yan G, Githure JI: **Influence of sugar availability and indoor microclimate on survival of *Anopheles gambiae* (Diptera: Culicidae) under semifield conditions in western Kenya.** *J Med Entomol* 2003, **40:**657–663. [http://www.ncbi.nlm.nih.gov/pubmed/14596279]

138. Kiszewski A, Mellinger A, Spielman A, Malaney P, Sachs SE, Sachs J: **A global index representing the stability of malaria transmission.** *Am J Trop Med Hyg* 2004, **70:**486–498. [http://www.ncbi.nlm.nih.gov/pubmed/15155980]

139. Mahande A, Mosha F, Mahande J, Kweka E: **Feeding and resting behaviour of malaria vector, *Anopheles arabiensis* with reference to zooprophylaxis.** *Malar J* 2007, **6:**100. [http://www.ncbi.nlm.nih.gov/pubmed/17663787]

140. Muriu SM, Muturi EJ, Shililu JI, Mbogo CM, Mwangangi JM, Jacob BG, Irungu LW, Mukabana RW, Githure JI, Novak RJ: **Host choice and multiple blood feeding behaviour of malaria vectors and other anophelines in Mwea rice scheme, Kenya.** *Malar J* 2008, **7:**43. [http://www.ncbi.nlm.nih.gov/pubmed/18312667]

141. Kibret S, Alemu Y, Boelee E, Tekie H, Alemu D, Petros B: **The impact of a small-scale irrigation scheme on malaria transmission in Ziway area, Central Ethiopia.** *Trop Med Int Health* 2010, **15:**41–50. [http://www.ncbi.nlm.nih.gov/pubmed/19917039]

142. Caputo B, Nwakanma D, Jawara M, Adiamoh M, Dia I, Konate L, Petrarca V, Conway DJ, della Torre A: ***Anopheles gambiae* complex along The Gambia river, with particular reference to the molecular forms of *An. gambiae* s.s.** *Malar J* 2008, **7:**182. [http://www.ncbi.nlm.nih.gov/pubmed/18803885]

143. Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat: **World Population Prospects: The 2004 Revision**. Tech. rep., United Nations 2004.

144. Bayoh MN, Lindsay SW: **Effect of temperature on the development of the aquatic stages of *Anopheles gambiae* sensu stricto (Diptera: Culicidae).** *Bull Entomol Res* 2003, **93:**375–381. [http://www.ncbi.nlm.nih.gov/pubmed/14641976]

145. R Development Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2011. [http://www.R-project.org/]. [ISBN 3-900051-07-0].

146. Kweka EJ, Zhou G, Beilhe LB, Dixit A, Afrane Y, Gilbreath TM, Munga S, Nyindo M, Githeko AK, Yan G: **Effects of co-habitation between *Anopheles gambiae* s.s. and *Culex quinquefasciatus* aquatic stages on life history traits.** *Parasit Vectors* 2012, **5:**33. [http://www.ncbi.nlm.nih.gov/pubmed/22321562]

147. Tirados I, Gibson G, Young S, Torr SJ: **Are herders protected by their herds? An experimental analysis of zooprophylaxis against the malaria vector *Anopheles arabiensis*.** *Malar J* 2011, **10:**68. [http://www.ncbi.nlm.nih.gov/pubmed/21435266]

148. Lewis D: **The anopheline mosquitoes of Sudan**. *Bull Entomol Res* 1956, **47:**475–494.

149. Rioux JA: **Contribution a l'etude des culicides (Diptera-Culicidae) du Nord-Tchad**. *Mission Epidemiologique au Nord Tchad* 1960, :53–92.

150. Coetzee M, Craig M, le Sueur D: **Mapping the distribution of members of the *Anopheles gambiae* complex in Africa and adjacent islands.** *Parasitol Today* 2000, **16:**74–77.

151. Meillon BD: **Malaria survey of South-West Africa.** *Bull World Health Organ* 1951, **4:**333–417.

152. Deshler W: **Cattle in Africa: distribution, types, and problems.** *Geogr Rev* 1963, **53:**52–58.

153. Diuk-Wasser MA, Toure MB, Dolo G, Bagayoko M, Sogoba N, Traore SF, Manoukis N, Taylor CE: **Vector abundance and malaria transmission in rice-growing villages in Mali.** *Am J Trop Med Hyg* 2005, **72:**725–731. [http://www.ncbi.nlm.nih.gov/pubmed/15964957]

154. Charlwood JD, Pinto J, Sousa CA, Ferreira C, Petrarca V, do E Rosario V: **'A mate or a meal'–pre-gravid behaviour of female *Anopheles gambiae* from the islands of São Tomé and Príncipe, West Africa.** *Malar Journal* 2003, **2:**9. [http://www.ncbi.nlm.nih.gov/pubmed/12801421]

155. Hogg JC, Thomson MC, Hurd H: **Comparative fecundity and associated factors for two sibling species of the *Anopheles gambiae* complex occurring sympatrically in The Gambia.** *Med Vet Entomol* 1996, **10**(4):385–391. [http://www.ncbi.nlm.nih.gov/pubmed/8994142]

156. Sanford MR, Demirci B, Marsden CD, Lee Y, Cornel AJ, Lanzaro GC: **Morphological differentiation may mediate mate-choice between incipient species of *Anopheles gambiae* s.s.** *PLoS One* 2011, **6**(11):e27920. [http://www.ncbi.nlm.nih.gov/pubmed/22132169]

157. Martens W: **Health impacts of climate change and ozone depletion: an eco-epidemiological modelling approach.** *PhD thesis*, Maastricht University, Maastricht, Netherlands 1997.

158. Animut A, Gebre-Michael T, Balkew M, Lindtjørn B: **Abundance and dynamics of anopheline larvae in a highland malarious area of south-central Ethiopia.** *Parasit Vectors* 2012, **5:**117. [http://www.ncbi.nlm.nih.gov/pubmed/22695178]

Additional file

Additional_file_1 as PDF
Additional file 1. Density of mosquitoes under different predation regimes and temperatures

Additional_file_2 as PDF
Additional file 2. Idealized puddle model

Additional_file_3 as PDF
Additional file 3. The functional form of of equation 42

Additional_file_4 as PDF
Additional file 4. A note on how initial conditions can influence the spatial distribution of *An. gambiae s.l.*

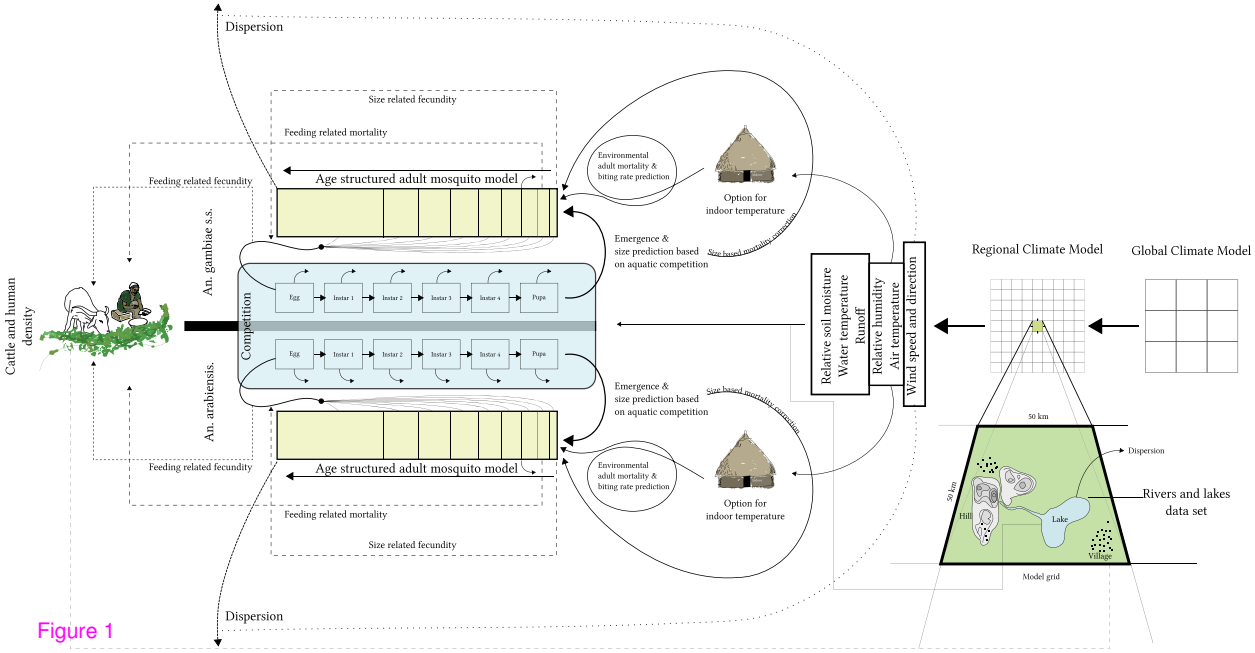
Dispersion
Size related fecundity
Feeding related mortality
Feeding related fecundity
Age structured adult mosquito model
Environmental adult mortality & biting rate prediction
Option for indoor temperature
Emergence & size prediction based on aquatic competition
Size based mortality correction
An. gambiae s.s.
An. arabiensis.
Competition
Egg
Instar 1
Instar 2
Instar 3
Instar 4
Pupa
Cattle and human density
Relative soil moisture
Water temperature
Runoff
Relative humidity
Air temperature
Wind speed and direction
Regional Climate Model
Global Climate Model
50 km
Dispersion
Rivers and lakes data set
Hill
Lake
Village
Model grid
Dispersion
Size related fecundity
Feeding related mortality
Feeding related fecundity
Age structured adult mosquito model
Environmental adult mortality & biting rate prediction
Option for indoor temperature
Emergence & size prediction based on aquatic competition
Size based mortality correction
An. arabiensis.
An. gambiae s.s.
Competition
Egg
Instar 1
Instar 2
Instar 3
Instar 4
Pupa
Cattle and human density
Dispersion

Figure 1

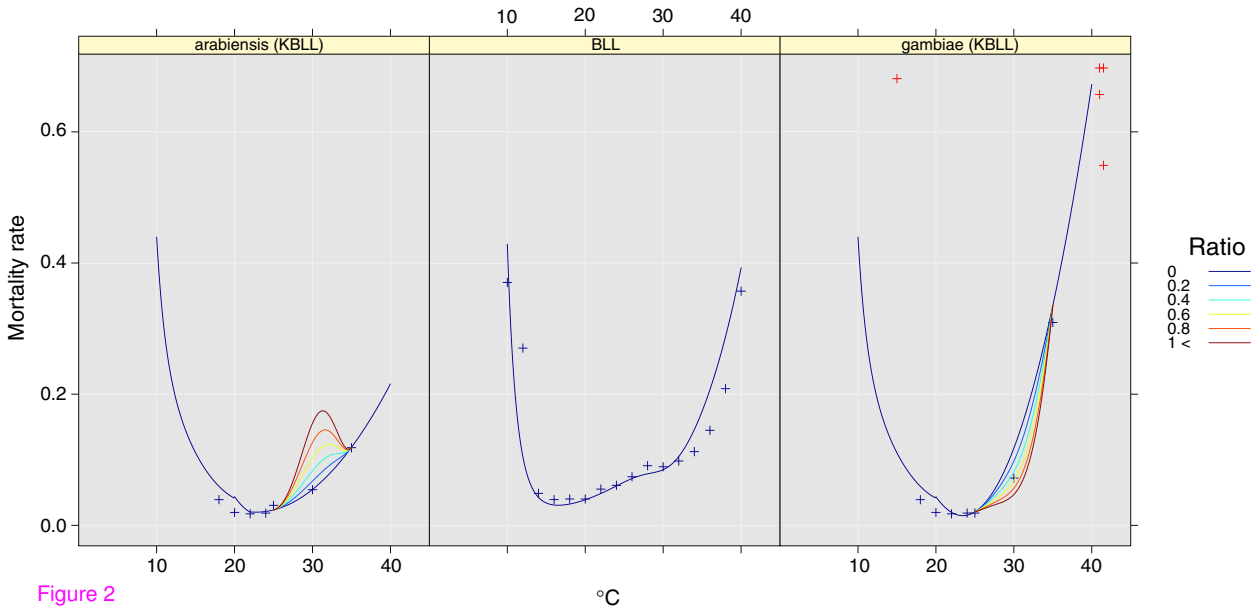
arabiensis (KBLL)
BLL
gambiae (KBLL)
Mortality rate
0.0
0.2
0.4
0.6
10
20
30
40
°C
Ratio
0
0.2
0.4
0.6
0.8
1 <

Figure 2

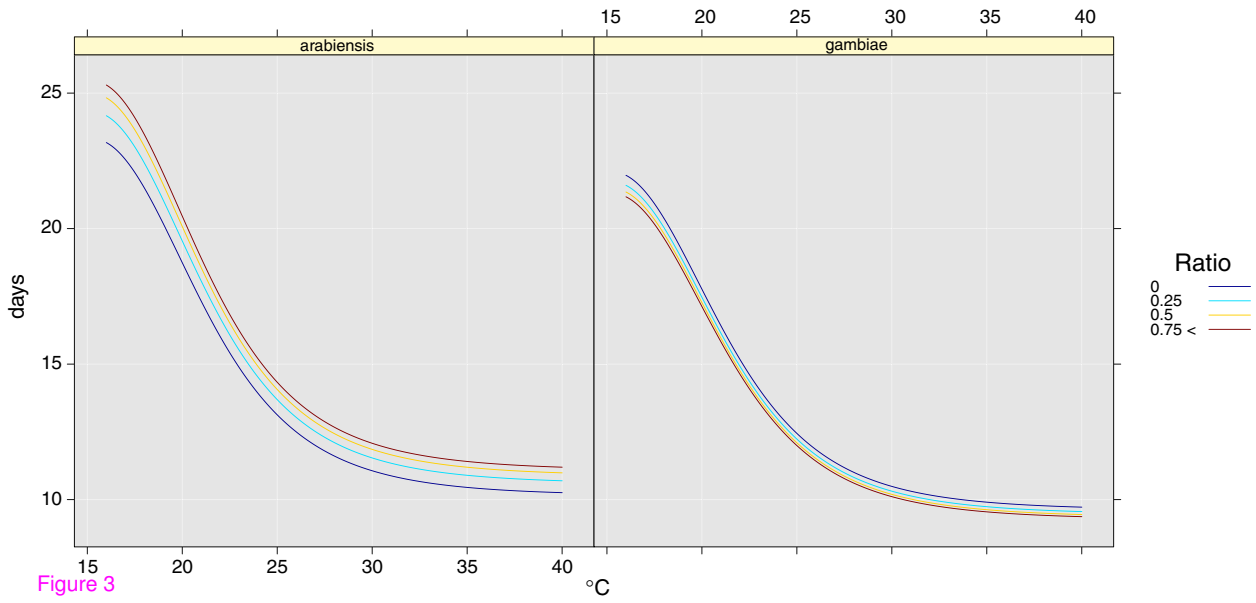

Figure 3

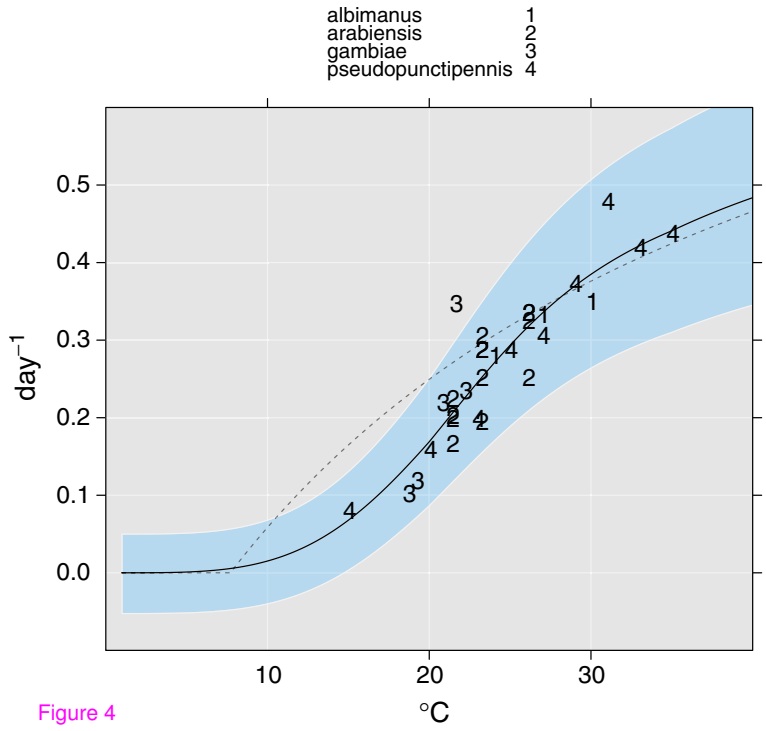
albimanus 1
arabiensis 2
gambiae 3
pseudopunctipennis 4
day^{-1}
0.0
0.1
0.2
0.3
0.4
0.5
10
20
30
°C

Figure 4

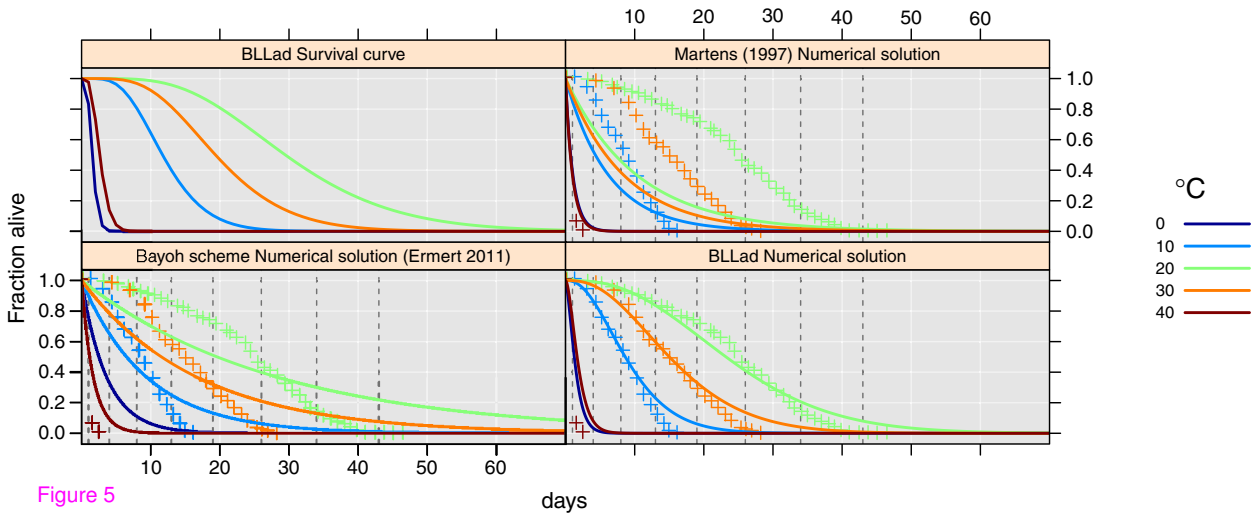
BLLad Survival curve
Martens (1997) Numerical solution
Bayoh scheme Numerical solution (Ermert 2011)
BLLad Numerical solution
Fraction alive
days
°C
0
10
20
30
40

Figure 5

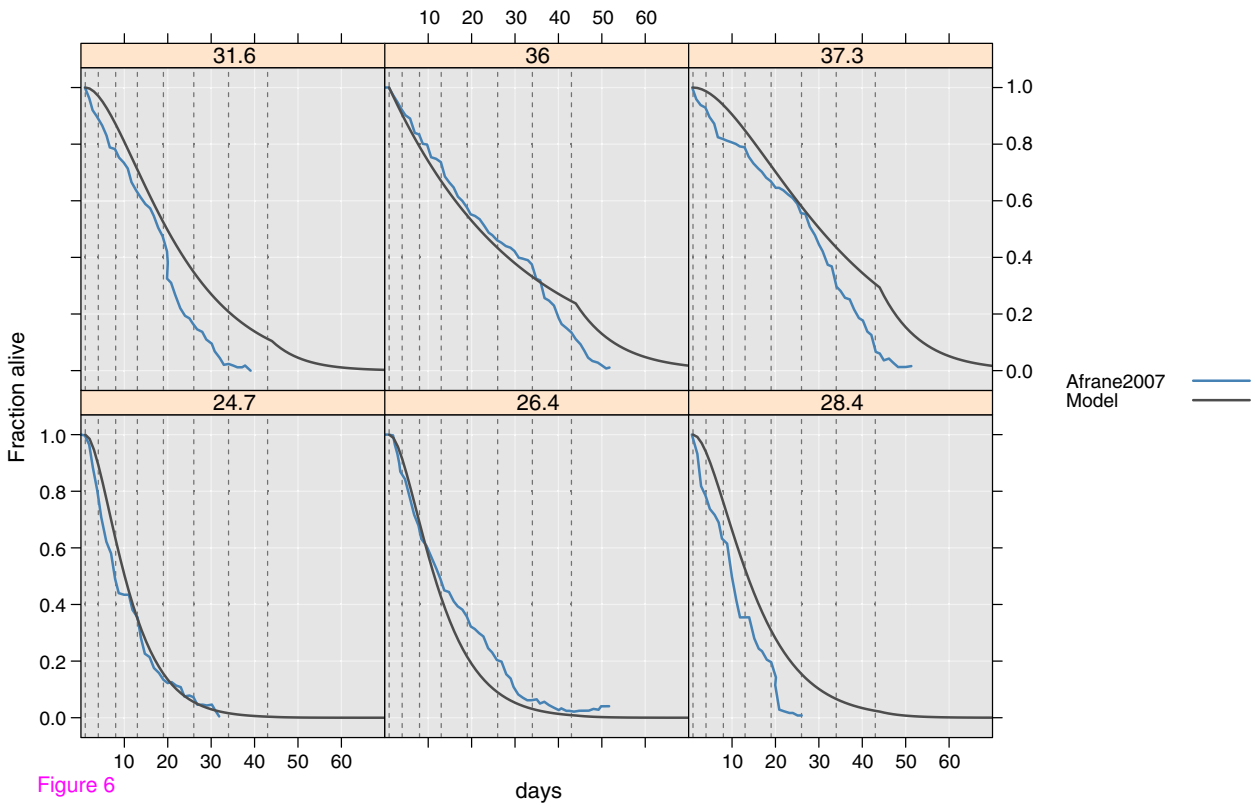

Figure 6

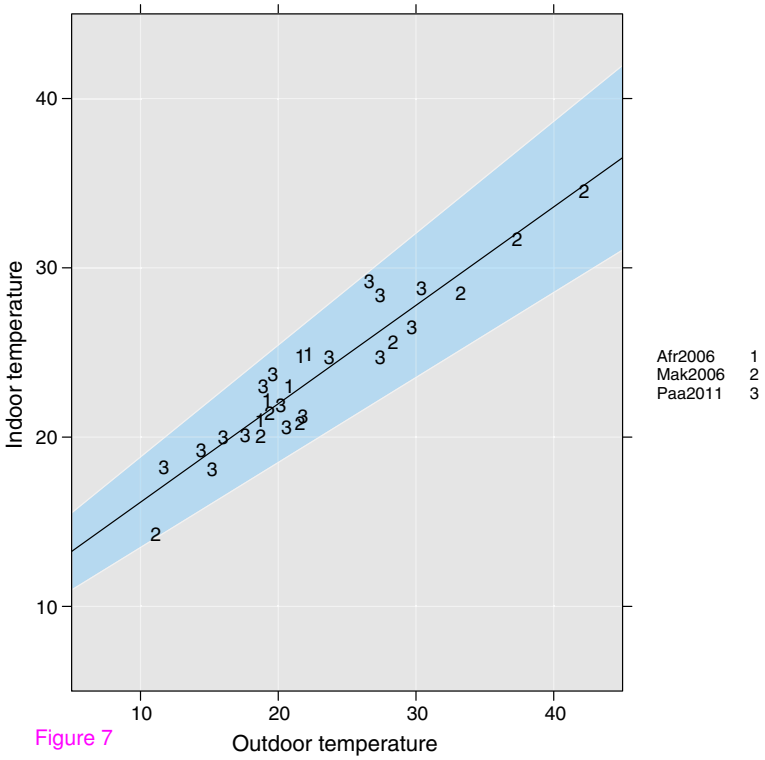
Indoor temperature
Outdoor temperature
10
20
30
40
Afr2006 1
Mak2006 2
Paa2011 3

Figure 7

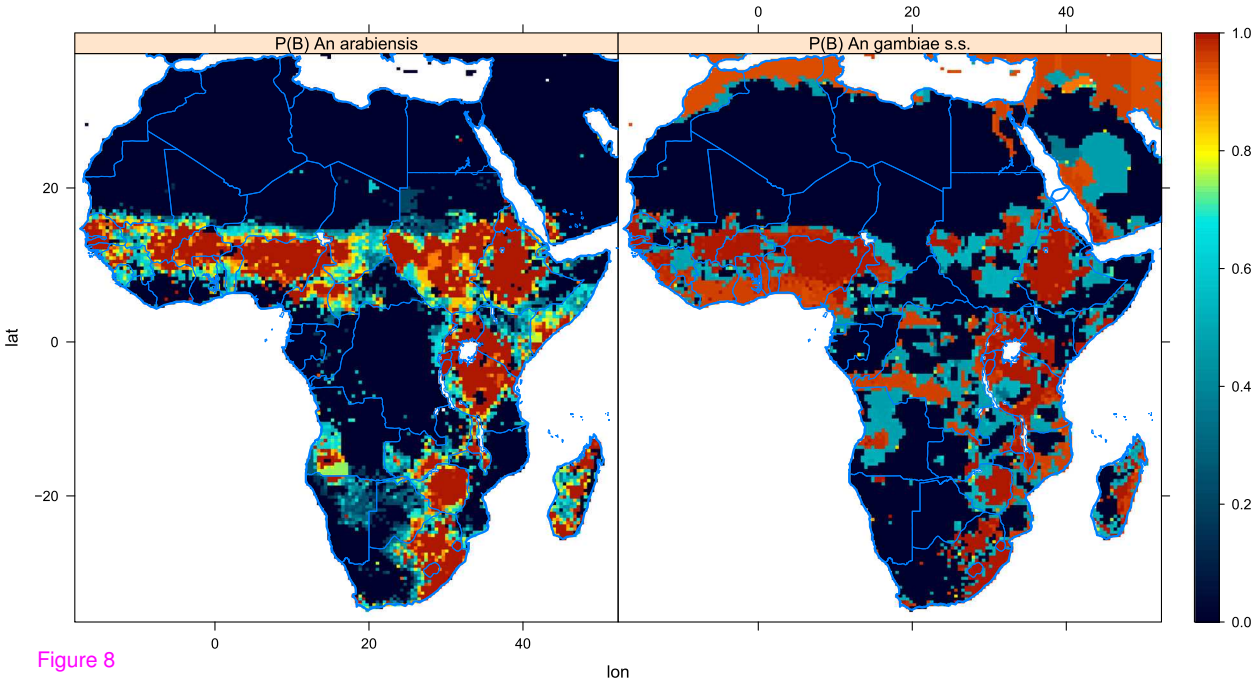
P(B) An arabiensis
P(B) An gambiae s.s.
lat
lon
20
0
−20
0
20
40
1.0
0.8
0.6
0.4
0.2
0.0

Figure 8

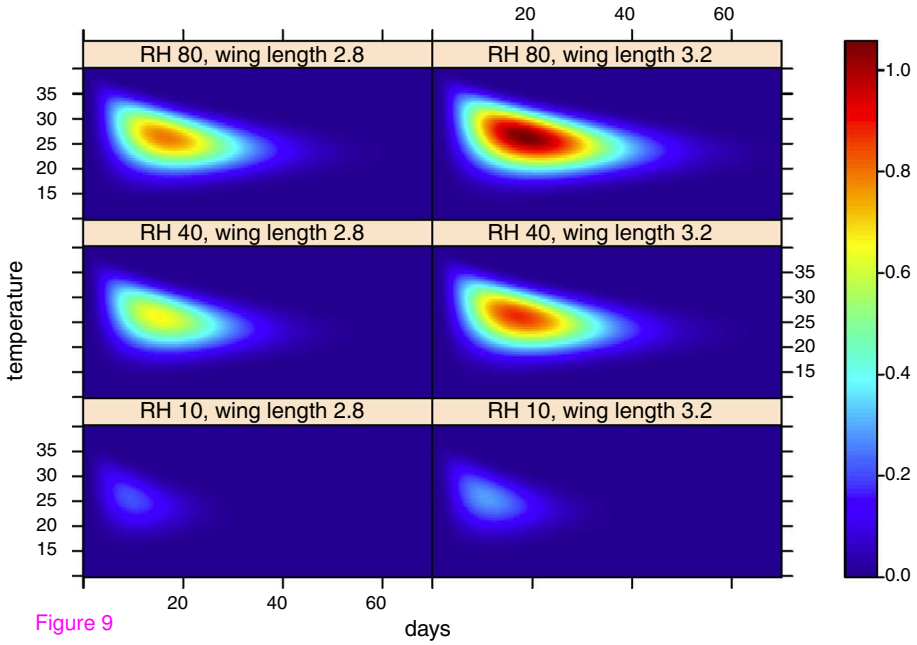
RH 80, wing length 2.8
RH 80, wing length 3.2
RH 40, wing length 2.8
RH 40, wing length 3.2
RH 10, wing length 2.8
RH 10, wing length 3.2
temperature
days
20
40
60
35
30
25
20
15
1.0
0.8
0.6
0.4
0.2
0.0

Figure 9

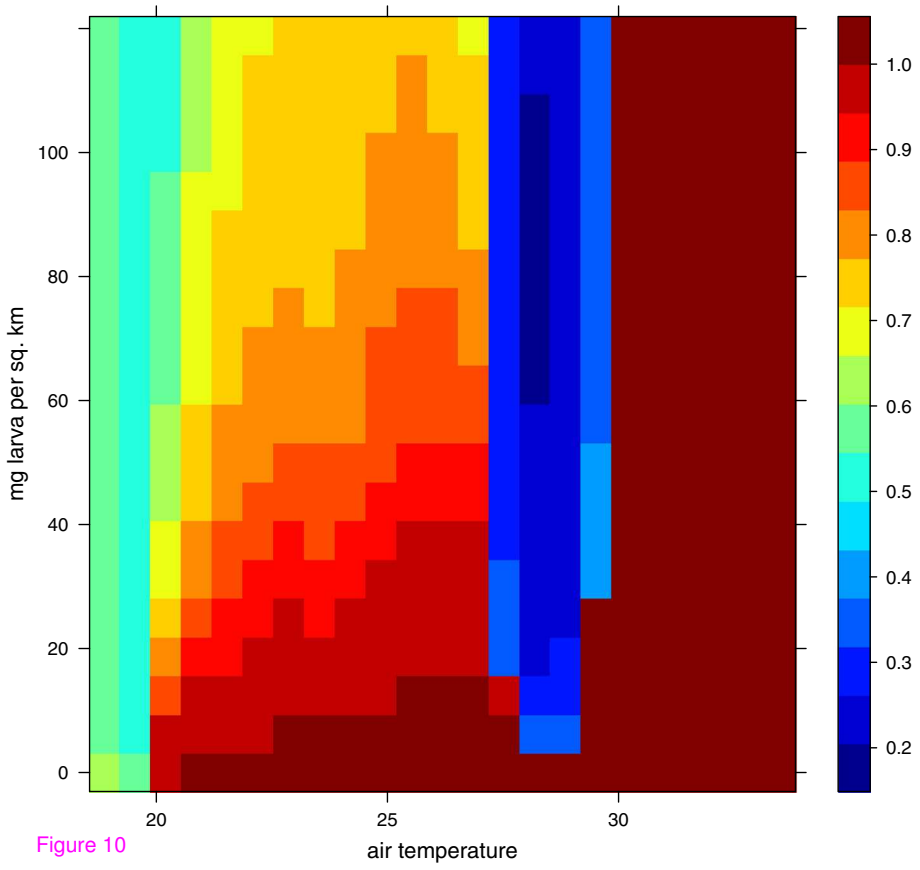
mg larva per sq. km
0
20
40
60
80
100
20
25
30
air temperature
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2

Figure 10

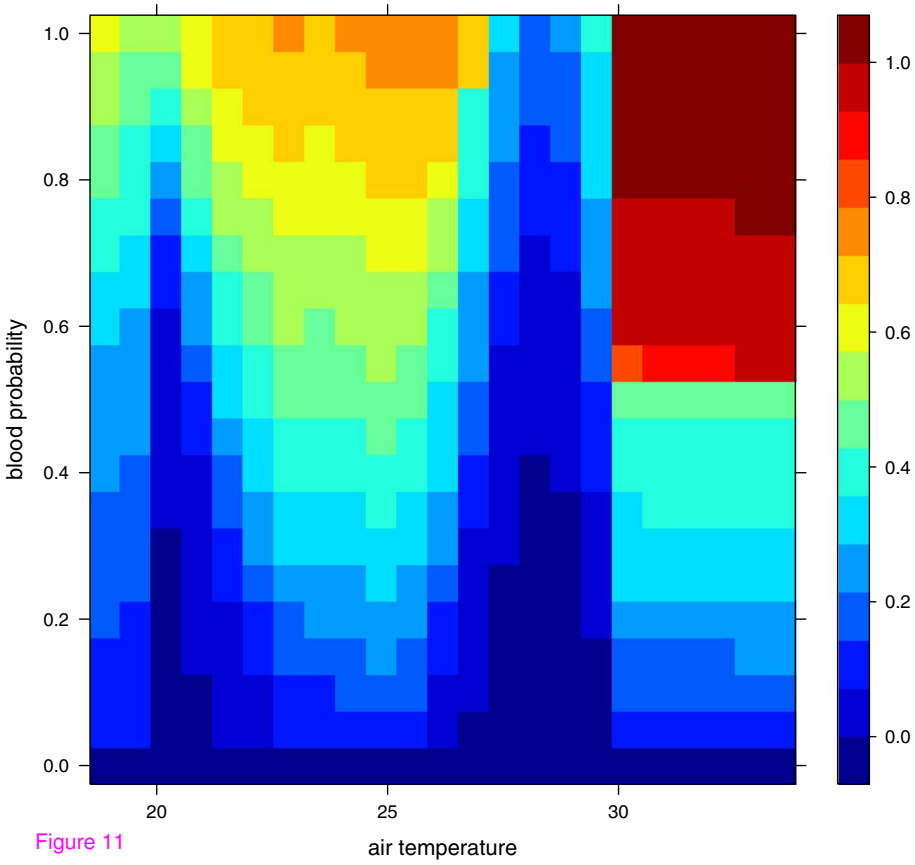
blood probability
1.0
0.8
0.6
0.4
0.2
0.0
20
25
30
air temperature
1.0
0.8
0.6
0.4
0.2
0.0

Figure 11

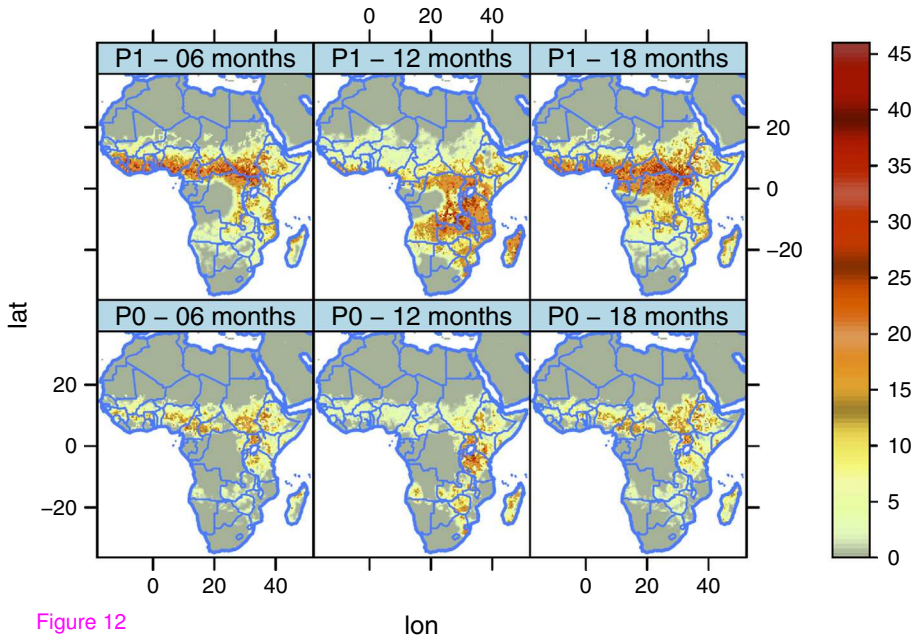
P1 – 06 months
P1 – 12 months
P1 – 18 months
P0 – 06 months
P0 – 12 months
P0 – 18 months
lat
lon
0
20
40
−20
45
40
35
30
25
20
15
10
5
0

Figure 12

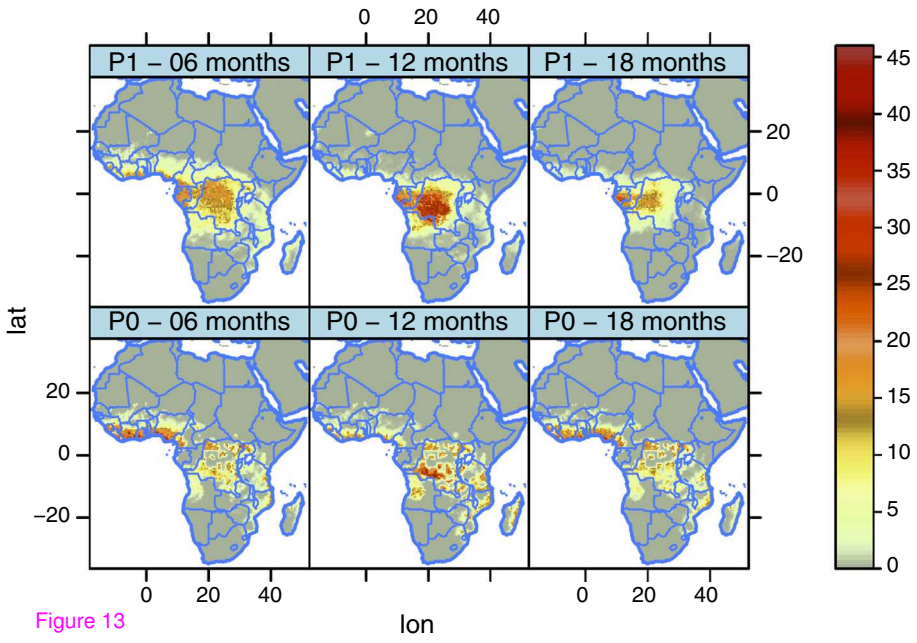
P1 – 06 months
P1 – 12 months
P1 – 18 months
P0 – 06 months
P0 – 12 months
P0 – 18 months
lat
lon
0
20
40
20
0
−20
45
40
35
30
25
20
15
10
5
0

Figure 13

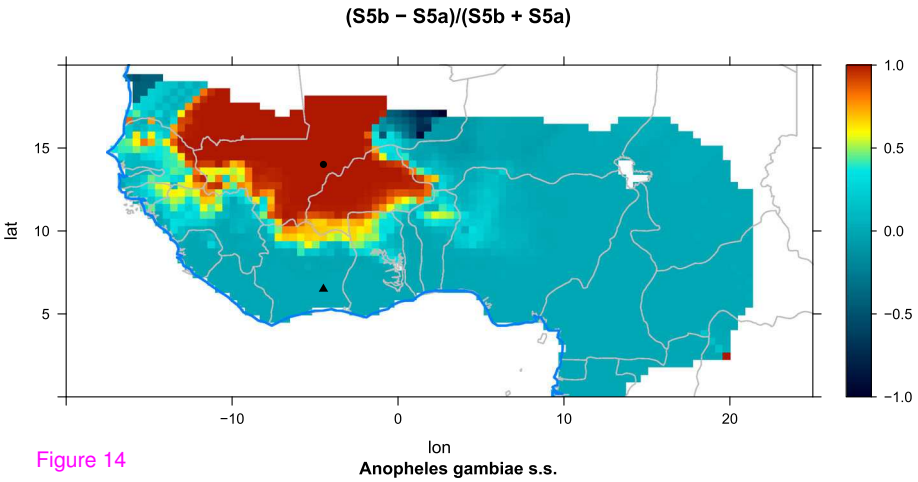
(S5b − S5a)/(S5b + S5a)
lat
15
10
5
−10
0
10
20
lon
1.0
0.5
0.0
−0.5
−1.0
Anopheles gambiae s.s.

Figure 14

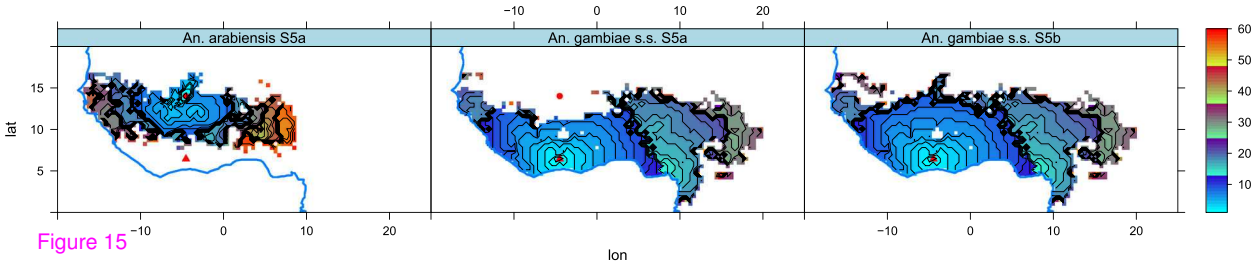
An. arabiensis S5a
An. gambiae s.s. S5a
An. gambiae s.s. S5b
lat
lon
−10
0
10
20
5
10
15
60
50
40
30
20
10

Figure 15

Additional files provided with this submission:

Additional file 1: 1858695900808695_add1.pdf, 41K
http://www.malariajournal.com/imedia/2023521800900915/supp1.pdf
Additional file 2: 1858695900808695_add2.pdf, 11K
http://www.malariajournal.com/imedia/1723306409900915/supp2.pdf
Additional file 3: 1858695900808695_add3.pdf, 475K
http://www.malariajournal.com/imedia/1441340604900916/supp3.pdf
Additional file 4: 1858695900808695_add4.pdf, 20K
http://www.malariajournal.com/imedia/1354779937900916/supp4.pdf

Paper II

6.2 A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. II. Validation of species distribution and seasonal variations

Torleif Markussen Lunde, Meshesha Balkew, Diriba Korecha, Teshome Gebre-Michael, Fekadu Massebo, Asgeir Sorteberg and Bernt Lindtjørn

Malaria Journal 2013, **12**:78

Malaria Journal

BioMed Central
The Open Access Publisher

A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. II. Validation of species distribution and seasonal variations

Malaria Journal 2013, **12**:78 doi:10.1186/1475-2875-12-78

Torleif M Lunde (torleif.lunde@cih.uib.no)
Meshesha Balkew (meshesha_b@yahoo.com)
Diriba Korecha (dkorecha@yahoo.com)
Teshome Gebre-Michael (teshomemg@yahoo.com)
Fekadu Massebo (fekimesi@yahoo.com)
Asgeir Sorteberg (asgeir.sorteberg@gfi.uib.no)
Bernt Lindtjørn (Bernt.Lindtjorn@cih.uib.no)

ISSN 1475-2875

Article type Research

Submission date 21 September 2012

Acceptance date 18 February 2013

Publication date 25 February 2013

Article URL http://www.malariajournal.com/content/12/1/78

A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. II. Validation of species distribution and seasonal variations

Torleif Markussen Lunde[1,2,6*]
*Corresponding author
Email: torleif.lunde@cih.uib.no

Meshesha Balkew[3]
Email: meshesha_b@yahoo.com

Diriba Korecha[4,6]
Email: dkorecha@yahoo.com

Teshome Gebre-Michael[3]
Email: teshomemg@yahoo.com

Fekadu Massebo[5]
Email: fekimesi@yahoo.com

Asgeir Sorteberg[2,6]
Email: asgeir.sorteberg@gfi.uib.no

Bernt Lindtjørn[1]
Email: bernt.lindtjorn@cih.uib.no

[1]Centre for International Health, University of Bergen, Bergen, Norway

[2]Bjerknes Centre for Climate Research, University of Bergen/Uni Research, Bergen, Norway

[3]Aklilu Lemma Institute of Pathobiology, Addis Ababa University, Addis Ababa, Ethiopia

[4]National Meteorological Agency of Ethiopia, Addis Ababa,

[5]Arba Minch University, Arba Minch, Ethiopia

[6]Geophysical Institute, University of Bergen, Bergen, Norway

Abstract

Background

The first part of this study aimed to develop a model for *Anopheles gambiae s.l.* with separate parametrization schemes for *Anopheles gambiae s.s.* and *Anopheles arabiensis*. The characterizations were constructed based on literature from the past decades. This part of the study is focusing on the model's ability to separate the mean state of the two species of the *An. gambiae* complex in

Africa. The model is also evaluated with respect to capturing the temporal variability of *An. arabiensis* in Ethiopia. Before conclusions and guidance based on models can be made, models need to be validated.

Methods

The model used in this paper is described in part one (Malaria Journal 2013, 12:28). For the validation of the model, a data base of 5,935 points on the presence of *An. gambiae s.s.* and *An. arabiensis* was constructed. An additional 992 points were collected on the presence *An. gambiae s.l.*. These data were used to assess if the model could recreate the spatial distribution of the two species. The dataset is made available in the public domain. This is followed by a case study from Madagascar where the model's ability to recreate the relative fraction of each species is investigated. In the last section the model's ability to reproduce the temporal variability of *An. arabiensis* in Ethiopia is tested. The model was compared with data from four papers, and one field survey covering two years.

Results

Overall, the model has a realistic representation of seasonal and year to year variability in mosquito densities in Ethiopia. The model is also able to describe the distribution of *An. gambiae s.s.* and *An. arabiensis* in sub-Saharan Africa. This implies this model can be used for seasonal and long term predictions of changes in the burden of malaria. Before models can be used to improving human health, or guide which interventions are to be applied where, there is a need to understand the system of interest. Validation is an important part of this process. It is also found that one of the main mechanisms separating *An. gambiae s.s.* and *An. arabiensis* is the availability of hosts; humans and cattle. Climate play a secondary, but still important, role.

Keywords

Anopheles gambiae complex, Model, Malaria

Background

Several attempts have been made to map the distribution of *Anopheles gambiae s.s.* and *Anopheles arabiensis* [1–5], two of the most important vectors of human malaria in sub-Saharan Africa. MacDonald [6] showed that limiting the human-vector contact reduces malaria transmission, and that the most efficient control measure is to increase the mortality rate of the involved mosquitoes. His thinking has been adopted in current malaria control efforts. Two of the most common interventions today are indoor residual spraying (IRS) [7] and insecticide-treated bed nets (ITNs) [8]. Often, there is no detailed understanding of the life history, behaviour and species composition where the interventions are applied [3].

Anopheles arabiensis inhabits areas from South Africa in the south to Mauritania and Sudan in the north. In Central-West Africa there is a pocket with very few observations of *An. arabiensis*. The border of this pocket is formed by Angola, Zambia, Burundi, Rwanda, Uganda, South-Sudan, Central African Republic, Congo, Gabon, and Equatorial Guinea. *Anopheles gambiae s.s.* is currently separated into five chromosomal forms: Forest, Bamako, Savanna, Mopti and Bissau [9], and two molecular forms:

M and S [10, 11]. It is distributed from South Africa to Mauritania and northern Mali, but is absent in Ethiopia and Northern Sudan. The species is considered the most efficient malaria vector in Africa [12].

Recent studies have shown that interventions aimed to prevent malaria has an impact on balance between *An. gambiae s.s.* and *An. arabiensis* [13]. The relative fraction of each species can vary from month to month, and year to year [14]. In Tanzania it has been shown that multi-decadal changes in the species composition can influence malaria transmission [15]. Given the observed changes in species composition, and their different capacity as vectors of malaria, it is highly relevant to have models which include several species when assessing the impact of climate variability and climate change.

This paper is the second of two describing and validating a new model of the dynamics of *An. gambiae s.s.* and *An. arabiensis* The model, which is described in part one [16], is a biophysical model driven by output from a climate model. Biophysical models seek to understand what drives a certain biological process, and to describe this with mathematical equations. Unlike statistical models, which often rely on observations to predict species presence and absence, biophysical models can be run with no information with respect to observed distribution and densities, and base the model equations on laboratory studies aiming to isolate different aspects of the life history of the mosquitoes. The role of field observations on the presence or absence of a species in the case of biophysical models, is to validate the model after an experiment has been completed. In some studies observations are used to reduce the uncertainty of unknown parameters [17].

In addition to predicting the current distribution, these type of models can be used to project changes in the historical and future density and distribution of these species. They can describe changes from day-to-day, month-to-month, year-to-year, and decade-to-decade. The model, named Open Malaria Warning (OMaWa) [16], includes several components, describing the mosquito's life from the aquatic stages to adult. In the aquatic stages, life history varies for eggs, larvae and pupae. As adults the life history changes with age. OMaWa is driven with air temperature, relative humidity of the air, wind speed and direction, soil temperature, relative soil moisture, and runoff from a climate model. These variables are used to parametrize mortality, rate at which eggs are laid, biting rate, development rate in the aquatic stages, and dispersion (spread) of mosquitoes. In part one, it was shown how the model responded to different forcings, and focused on its sensitivity to temperature, humidity, mosquito size, the probability of finding blood, and dispersion. Thus the results presented here should be seen in light of the sensitivity analysis. A full description of the model used here can be found in part one [16].

This is the first time a biophysical model has been used to model the relative density of *An. gambiae s.s.* and *An. arabiensis*, with simulations covering an entire continent. It is also the first time age dependent life history and mosquito dispersion (spread of mosquitoes) has been included in a continental analysis. The model is validated against 6,927 presence/absence points of the two species, and a more detailed analysis is carried out for Madagascar. The data is freely available to the public [18]. This study has also evaluated the ability to model the temporal variability, using case studies for Ethiopia.

Methods

Occurrence and distribution of *An. gambiae s.l.* in Africa

Continental validation

To date there are three data sets describing the occurrence of *An. arabiensis* and *An. gambiae s.s.* [3, 19, 20]. Additional online resources have been described by Hay *et al* [21]. To compliment and

extend these databases, a systematic search was conducted. A total of 1,940 occurrence points were collected for *An. arabiensis*, 1,813 for *An. gambiae s.s.*, and 992 for *An. gambiae*. Merging these data with the three databases [3, 19, 20] result in 2,926 occurrence points for *An. arabiensis*, 3,009 for *An. gambiae s.s.*, and 992 for *An. gambiae* [18]. Three methods were used to geo-reference the points. In papers where coordinates were given, these coordinates were used. If possible they were cross checked against given place names. In cases where only place name, and a description of the place were given, the locations were searched up using Google Maps/Earth. Where only a map was provided, the map was imported to qgis and geo-referenced [22], and occurrence points were manually extracted.

The database containing *An. gambiae* was mainly used to estimate the occurrence of *An. gambiae s.l.* in Namibia, DRC, South Sudan, Angola, Congo, and northern South Africa. To classify the points the expert opinion polygons from Sinka *et al* [3] was used. A point falling within the *An. arabiensis* polygon only was classified as *An. arabiensis*, points falling within the *An. gambiae s.s.* polygon only as *An. gambiae s.s.*, and points falling within both polygons were assigned both species. To classify true presence/absence points the data described previously was used. Observations of *An. gambiae s.s.* were classified as presence for this species. Absence points for *An. gambiae s.s.* were those where *An. arabiensis* had been recorded, and no *An. gambiae s.s.* had been observed within a radius of 100 km. The same approach was used for *An. arabiensis*.

This model (OMaWa) was compared with species predictions from four other models, as well as the expert opinion from Sinka *et al* [3]. The first was the paper by Rogers *et al* [1] where they used satellite data to predict the presence of *An. arabiensis* and *An. gambiae s.s.*. To reproduce the images in the paper the figures were geo-referenced, and polygons were drawn based on the 0.65-1 probability. The selection was based on the colouring they used in the figure. Next a 50 by 50 km grid was overlaid with the polygons, and points falling within the polygons were classified as presence points. Points falling outside were classified as absence. The second paper is by Levine *et al* [2]. They used a genetic algorithm to predict the presence of the two species. As before, the images were geo-referenced, and polygons were constructed based on dark grey to black shading. Next, absence and presence was constructed as for Rogers *et al* [1]. The third paper is a recent paper by Sinka *et al* [3]. Since this is a three band RGB (Red-Green-Blue) raster, the pixel values were first converted to a one band raster: $1-(0.299\cdot R+0.587\cdot G+0.114\cdot B)/255$. This new raster image was then gridded to a 50 by 50 km grid. Presence was defined as probability greater than approximately 0.4. As for Rogers *et al* [1], this threshold was selected based on the colouring in the figure (and it must be assumed the authors chose the colours based on what they thought to be realistic classifications). Where applicable, the weighted absolute mean error was also calculated based where weights were equal to the probability given in the maps. The fourth paper is by Moffet *et al* [5]. The same methodology as for Sinka *et al* [3] was used to construct a comparable map. For the expert opinion, presence/absence points were constructed with the same methodology used for Levine [2] and Rogers [1]. The extracted data and scripts are available upon request. The mosquito density from OMaWa was classified as present if the 19 year mean was greater than 0.004 mosquitoes per square kilometre, and absent if less. Quality of the models were estimated as mean absolute error (MAE), which is recommended over the root mean square when comparing model performance [23].

Relative fraction of each species, Madagascar

To investigate if the model is able to estimate the relative fraction of *An. gambiae s.s.* and *An. arabiensis* data from Pock Tsy's *et al* [24] and Chauvet's [25] article describing the fraction of each species in Madagascar was used. In total these two data sets consist of 275 observations, and should thus be suitable to give a rough idea about the relative fraction of the two members. Different measures were given to evaluate the model skill:

a) For each observation there are information about the month of collection as well as longitude and latitude. From the model data, covering the period 1990-2008, the closest point to each observation in the month of collection is selected, and the yearly monthly mean is calculated. These data were used to make box plots, weighting for the number of observations in each point, comparing the observations with the model.

b) From the data produced in a, maps were created using a distance weighted kernel with cut off at 100 km. Hence observations further away than 100 km were not included, and closer points will be given more weight.

c) The distance to the closest wrong (difference in fraction greater than 0.2) and correct (difference smaller than 0.2) prediction will be indexes for the spatial accuracy. A non-parametric test like the Wilcoxon rank sum test with continuity correction (Mann-Whitney) test can then be used to test if the two indexes differ by a location shift of zero, and the alternative is that they differ by some other location shift.

Temporal variability

Model setup

In addition to looking at the spatial patterns, it is of interest how the model reproduce temporal variability in mosquito numbers. Originally, this model was developed to increase the understanding of malaria epidemiology in Ethiopia. The motivation of introducing *An. gambiae s.s.* was to test if the model had a general validity, not limited to Ethiopia. Two high resolution runs only covering Ethiopia were done; one at 30 km, covering the period from 2000 to 2006 (Eth30), and one at 18 km (Eth18) covering the period from 2008 to 2011. These two runs differ from the one covering all of Africa in the way that the weather simulations were forced to follow the observed weather pattern. The technique used to accomplish this is called spectral nudging. In the African run (TC50) the intention was not to reproduce the exact year to year variability, but the interest was to reproduce reasonable weather in a reasonable climate, and thus no nudging was used. To validate the ability to reproduce seasonal variations data from Eth30 and Eth18 to drive OMaWa was used.

For simulations driven by Eth30 the model was run without dispersion, BLL aquatic mortality, development rate with no species correction, default gonotrophic cycle, and AL adult mortality. TC50 and Eth18 were run with the following parametrization: with dispersion, KBLL aquatic mortality, development rate with species correction, default gonotrophic cycle, and BLLad adult mortality. All results are based on single realizations of the model, and error bars are therefore not reported.

Validation data

There are few papers describing the year to year, and seasonal variations in mosquito numbers in Ethiopia. In the validation process three papers were used, one master thesis, and field data from Chano Mille, Arba Minch describing mosquito seasonality.

The first, a paper by Kenea *et al* [26], is describing *An. arabiensis* larva density in the vicinity of six villages in central Ethiopia, December 2007 to June 2008. The second paper is by Taye *et al* [27] and is reporting bi-monthly (October 2001 to August 2002) adult *An. arabiensis* numbers in Sille (Southern Ethiopia). The third paper is by Yemane Ye-Ebiyo *et al* [28], where they report larva density in seven naturally formed puddles, in Ziway. Since this paper does not report density in the area as a whole, the

data might not be directly comparable to the modelled ones. To overcome this problem all time series were scaled, both observations and model results, as standardized anomalies:

$$\frac{x - \frac{1}{n}\sum x}{\sqrt{\frac{\left(x - \frac{1}{n}\sum x\right)^2}{n}}} \tag{1}$$

To compare the absolute density, it would be required that the papers reported the larva/mosquito density per square kilometre over a larger area. Since this is not the case, scaling is necessary. The last study is by Balkew, where the seasonality of *An. arabiensis* in Awash Valley, Ethiopia, was described [29]. The study locations are plotted in Figure 1.

Figure 1 Areas used for validation of the model. KEN11 (blue circle), TAY2006 (green circle), YE2003 (orange circle), BAL2001 (red circle), and FEK2012 (grey circle).

In addition to the published data, Fekadu Massebo collected one year (May 2009 to April 2010) of mosquito densities in Chano Mille, Ethiopia. The study site is described in [30, 31]. To see if the model was able to reproduce the mosquito densities, Eth18 was used to drive OMaWa.

Validation statistics

All correlations (Pearson) are calculated from the values reported in the papers [26–29], and a similar time series (sampled the same month as the observations) is constructed from the model averaging the four closest model points:

$$cor\left(\frac{x_{obs} - mean(x_{obs})}{sd(x_{obs})}, \frac{x_{mod} - mean(x_{mod})}{sd(x_{mod})}\right) \tag{2}$$

Climate model realizations

The simulations in this paper was driven by three different realizations of a limited area climate model. The first realization (Eth30), carried out in 2009, comes from WRF model version 3.1.1 [32]. It was run at 90 km resolution using a tropical channel set up. In this type of setup, the domain consists of the boundaries above and below certain latitude and no side boundaries. This process allows the interaction from the extra-tropics through the north-and-south boundaries. In addition, it allows the generated waves to propagate around the globe more naturally – as in the real world and in global models. The meridional boundary conditions were specified using six-hourly National Centers for Environmental Prediction (NCEP) Reanalysis 2 (T42) data. The runs have meridional boundaries at $45°S$ and $37°N$, with 27 vertical levels, ranging from the surface to pressure p = 10 hPa. Inside the channel, a domain with 30 km resolution was set up. This domain has boundaries at $25.56°E$, $53.18°E$, $0.24°N$, and $19.29°N$. To ensure the model reproduced the observed year to year anomalies, the model was nudged, using spectral nudging, against waves (wind, pressure, and temperature) longer than 1,000 km in both domains. The Kain Frisch cumulus parametrization scheme was used [33, 34].

The second realization (TC50), carried out in 2011, had again a tropical channel set up. The model was run at 50 km resolution from January 1 1989 to January 1 2009. At the north and southern boundaries the model was driven by Era Interim. The Kain Frisch cumulus parametrization scheme was used [33, 34]. No nudging was used, and therefore it is less probable the model would reproduce year to year variability in the weather. This run was used to assess the mean state of mosquito density and distribution.

In the third experiment (Eth18), done in 2012, WRF 3.3.1 was used with the Tiedtke cumulus parametrization scheme [35, 36]. The model was run at 18 km resolution from January 1 2008 to August 1 2011, with data from Era Interim at the boundaries. Outside the planetary boundary layer the same type of spectral nudging as described earlier was applied. The domain had boundaries at 30.57° N, 50.99° N, 1.45° S, and 18.97° E.

The Regional Committee for Medical and Health Research Ethics, Western Norway, and the Ethical Committee of the Faculty of Medicine of Addis Ababa University and The National Health Research Ethics Review Committee (NERC) of Ethiopia granted ethical approval for the study.

Results and discussion

Distribution of Anopheles gambiae s.l.

Occurence of Anopheles gambiae s.l. in Africa (TC50)

Figure 2 is showing the presence data collected as part of this work. Data collection on *An. gambiae* was focused on areas where little information about the occurrence of *An. arabiensis* and *An. gambiae s.s.* was available. Figure 3 shows the modelled mean density of *An. arabiensis* and *An. gambiae s.s.*. The white contours are indicating the presence of each species. The pattern is consistent with the general perception of the species range [3]. This is the first time a model [1–4] has been able to reproduce the absence of *An. gambiae s.s.* in Ethiopia. Still there are some unresolved issues. To date there are no records of *An. arabiensis* in Côte d'Ivoire; no models, this included, have been able to model the absence of this species in Côte d'Ivoire. A look at the figure also reveals some probable inconsistencies with respect to the species distribution in southern Chad where *An. arabiensis* should be dominating [37]. In South-Africa the distribution is consistent with observations from 1958 [38]. There are however no recent available surveys of *An. gambiae s.l.* in the states of Gauteng, North West or South Western Limpopo. In Namibia, where *An. gambiae s.l.* has been observed as far south as $-23.7°N$ [39], the model limits the range to approximately $-21°N$. Since there are no available data on the recent distribution of this complex in Namibia, it is difficult to know whether the model is correct or wrong. The model also suggests *An. gambiae* is absent in large parts of Gabon. Previous studies have found *An. gambiae* in Lambarene [40] and Moyen-Ogooue [41], while Mouchet only found this species in Libreville of twelve sites sampled [42]. It should be noted that Mourou *et al* later found *An. gambiae* in Port-Gentil [43], as predicted by the model, while Mouchet [42] did not record this species 26 years earlier. Elissa *et al* [44] also found low concentrations of *An. gambiae s.s.* in Haut-Ogooué, which was also predicted by the model. In the north-eastern part of Gabon it has not been possible to find any recent mosquito surveys, and it is therefore hard to conclude if the predicted absence of *An. gambiae* in this region is correct.

Figure 2 Presence points for *An. arabiensis*, *An. gambiae s.s.* and *An. gambiae*.

Figure 3 Mean density of *An. arabiensis* and *An. gambiae s.s.*, 1990-2008. White contours show where the species were present during the simulation.

To evaluate the quality of the model with respect to classifying the presence and absence of the species the methodology described previously was used. Table 1 shows the mean absolute error for the four papers [1–3, 5], expert opinion and this model. For reference, a MAE of 1 would be equivalent to completely wrong predictions, and 0 would be perfect. While the genetic algorithm of Levine [2] and the predictions based on satellite imagery by Rogers [1] show poor skill, the recent papers by Moffet *et al* [5] Sinka *et al* [3] are great improvements compared to those. Still, they have less skill than the expert opinion if comparing to the unweighed MAE. This model (OMaWa) has lower MEA than all the models included in this analysis, and including weights in the MEA makes it superior even to the expert opinion. The occurrence data suggest the expert opinion for *An. arabiensis* is wrong over West Africa and Southern Cameroon. A mosquito survey in Namibia, and north-eastern Gabon, would also clarify the present-day species composition in these countries.

Table 1 Mean absolute error species presence/absence (Weighted mean absolute error)

	Model	**MAE**
1	Levine	0.33 (NA)
2	Rogers	0.29 (NA)
3	Moffet	0.20 (0.07)
4	Expert Opinion	0.07 (NA)
5	Sinka	0.13 (0.05)
6	OMaWa	0.07 (0.01)

Relative fraction of each species, Madagascar (TC50)

Since Madagascar has a sharp separation between *An. arabiensis* and *An. gambiae s.s.*, the island is well suited to address whether the model is able to reproduce the relative fraction of each species.Three measures to evaluate the model was defined. For method a) the mean absolute error was 0.22. The box plot in Figure 4 show the fraction of *An. arabiensis* from the model, grouped by the fraction in the observations. It is clear, while capturing the main tendencies well, the model has problems with the exact separation between the two species. In the mixed group, the model tends to let one species dominate over the other, possibly letting *An. arabiensis* dominate too easily.

Figure 4 Box-plot of fraction *An. arabiensis*, Madagascar. Blue is the fraction from the model, while red is observations. The arabiensis group is where observations showed more than 85% *An. arabiensis*, gambiae is where observations showed less than 15% *An. arabiensis*, and mixed is the remaining data. Dot/triangle indicate the median.

Figure 5, created using method b), shows the fraction of *An. arabiensis* as modelled, and observed. An eyeball comparison shows the separation is shifted westward in the model, and a bias in the South-Eastern tip of Madagascar. Whether this is a result of (climate) model resolution, failing to accurately separating the west/east gradient in topography, or the biological parametrization being inaccurate is hard to quantify. It is hoped this can be tested in a future analysis with higher model resolution.

Figure 5 Fraction of *An. arabiensis*. Model 1990-2008, and observations smoothed with a squared inverse distance weighted kernel with cut-off at 100 km.

Table 2 shows the distance to the closest model point, distance to the closest model point with correct prediction, and distance to the closest point with wrong prediction as described in c). At all quantiles

the distance to the closest correct prediction is 1.5 to 7 smaller than the closest wrong prediction. A Mann-Whitney test with confidence level of 0.99 shows the difference in location between wrong and correct predictions is 9.84 (5.07 25.68) km ($p < .0001$). Thus, although with biases, it is concluded that distance to closest correct prediction and closest wrong prediction are non-identical populations.

Table 2 Distance to closest correct and wrong prediction

	0%	**10%**	**25%**	**50%**	**75%**	**90%**	**100%**
Distance to closest point	1.16	2.63	3.62	4.96	8.12	10.42	25.80
Distance to closest correct prediction	1.16	3.03	4.05	6.39	10.51	43.81	275.96
Distance to closest wrong prediction	1.63	4.06	5.55	28.05	73.72	112.66	311.81

Distance to the closest model point, distance to the closest model point with correct prediction, and distance to the closest point with wrong prediction. Model vs. observations.

Temporal variability

It is important that mosquito models reproduce the seasonal cycle correctly, since this will be an indication of the sensitivity to climate. Here results from the model are compared to a number of observational studies. The comparison with each individual study might not have much information, but it is recommended that readers look at the results as a whole, having in mind the continental analysis showing the model is able to separate the distribution of *An. arabiensis* and *An. gambiae s.s.*. These results are meant to complement the continental analysis. Eth30 and Eth18 refers to the weather data used to drive OMaWa.

KEN11: 2007-2008 larva density in central Ethiopia (Eth30)

In this study [26] Kenea *et al* reported the *An. arabiensis* larva density in six locations in central Ethiopia, December 2007 to June 2008. Five of the sites followed the same seasonality, while one had the highest density before the rainy season started. The model is not designed to capture such local variations, but is rather aiming to describe the median, or sometimes mean, state within a certain area. In their study all anopheline positive habitats present within a 500 m radius of each irrigated village/town and 700m along the major drainages (lake or river) were sampled. This means that the data should be comparable to what is modelled. The seasonality of larva density, $l_{sum} = \sum_{\iota=1}^{4} l_\iota$, per puddle area, A_p, is then calculated as $C_l \frac{l_{sum}}{A_p[m^2]}$, where C_l is a dimensionless constant. Correlations with the median relative seasonality, model vs. Kenea et al., is $0.97(0.81, 0.99)$, and mean relative seasonality $0.92(0.55, 0.98)$. The observations and modelled results can be seen in Figure 6.

Figure 6 Scaled variations over time of six locations (dashed grey line), and the median seasonality (solid grey line) in Central Ethiopia [26] (data from KEN11). Blue solid line shows modelled relative seasonality in the same area.

TAY2006: 2001 mosquito catch Sille, Ethiopia (Eth30)

In 2001-2002 Aseged Taye *et al* [27] recorded number of man biting *An. arabiensis* in Sille, Ethiopia. For simplicity it is assumed the human biting rate is independent of temperature and availability of breeding sites. This means the relative monthly mean sum of mosquitoes from the model should be directly comparable with the records from the paper. The model seems to under-predict the relative abundance of *An. arabiensis* in October 2001, and over-predict the rise in mosquito numbers in February. Otherwise the modelled number of mosquitoes seems comparable to what was observed by Taye *et al*.

The correlation between observations and model (2001-2002) is $0.91(0.36, 0.99)$. The observations and model results are shown in Figure 7.

Figure 7 Scaled variations over time of *An. arabiensis* in Sille, Ethiopia (data from TAY2006), observed (grey solid line), model 2001-2002 (solid blue line), model multi-year monthly mean (dashed blue line).

YE2003: 2001 3 month larva variability in Zwai (Eth30)

If it is assumed larva per dip has units $LPD = C\frac{larva}{m^2}$, where C is a constant, and that the samples are representative for a larger area, the relative number of larva in that area can be estimated as $LPD \cdot W_a$, where W_a is the mean water area in m^2. This way it is assumed the number of puddles is constant from July to September, and that the puddles only change their surface area. These values are roughly comparable to the modelled number of larva. Since only the latitude (and not the longitude) is reported in the paper, and Zwai is not located at latitude $9°N$, model data between longitudes 38.69 to $39.23°E$ and latitudes 7.88 to $8.42°N$, an area covering Zwai, were selected. Using this method correlation is $0.99(0.321.00)$. Confidence interval is estimated using 1,000 random samples of the points within the bounding box, and the 2.5% and 97.5% quantiles of the correlations is reported. Since the sample size is small and the data might not be directly comparable, the correlation should be interpreted with care. The data from the observations and the model can be seen in Figure 8.

Figure 8 Scaled variations over time of *An. arabiensis* larva in Zwai, Ethiopia (YE2003). Observed (grey solid line), and model (solid blue line).

BAL2001: 1999-2000 mosquito catch Awash, Ethiopia (Eth30)

This study was carried out in 1999-2000 in Metehara at longitudes 39.50 to $40.00°E$ and latitudes 8.75 to $8.92°N$. The data are based indoor space spray collections. Since the malaria model was not run for 1999, and 2000 is considered as a spin-up year, the multi-year monthly mean for the years 2001-2006, and 2008-2009 was used (since the climate model was done as two separate runs, one starting January 2000, and one starting January 2007). The observations are compared to the scaled sum of mosquitoes of all age groups, which should be comparable to what was reported in the thesis. Correlations in Buse + Gelcha (two locations described in the thesis) was $0.75(0.1, 0.95)$, $0.79(0.27, 0.95)$ for Sugar Estate, and 0.76 for Metehara Town. Confidence intervals are not reported for Metehara Town since the number of observations are low. The data can be seen in Figure 9.

Figure 9 Scaled variations over time of adult *An. arabiensis* in Awash, Ethiopia (BAL2001). Observed (grey solid line), and model (solid blue line).

FEK2012: 2009-2010 mosquito catch Chano Mille, Ethiopia (Eth18)

As seen in Figure 10, and correlations in Table 3, the model corresponds well with the observations in Chano, 2009-2010. While the weather station in Arba Minch recorded some heavy rainfall events in

October/November 2009 the regional climate model did not capture these events, or did not dump the precipitation in the right location [45]. In general the driving model (WRF) was too wet in spring 2009, and too dry in autumn 2009. This might be the reason for the slight mismatch in mosquito numbers in these seasons. To have confidence in malaria/mosquito models at these fine scales, there is a need for a better representation of precipitation in the climate models. The differences between the trapping methods also highlight the uncertainty of related to data collection, especially when the number of mosquitoes is low. From December 2009 to March 2010 the observed number of *An. arabiensis* was very low (Figure 10). It is interesting that despite of this, malaria started to rise in these months [30].

Figure 10 Modeled and observed variations in *An. arabiensis*. The left panel shows catches broken down to catch method (grey dotted lines), and modelled *An. arabiensis*. The right panel shows modelled mosquitoes and total twice monthly catches (data from FEK2012).

Table 3 Correlations for model and mosquitoes captured in Chano Mille

	correlation
Twice monthly	0.80 (0.58, 0.91)
Space Spray	0.83 (0.49, 0.95)
Pit Shelter	0.80 (0.43, 0.94)
CDC Light Trap	0.83 (0.48, 0.95)

Twice monthly shows correlation with total catch of *An. arabiensis* and modelled numbers. The Space Spray, Pit Shelter, and CDC Light Trap show correlation (95% confidence interval) with modelled number of mosquitoes and monthly catches using three different methods

Summary of temporal variability analysis

Each of the five case studies consist of short time series, with different observational methodologies. It was attempted to show how the model results can be compared to the different type of observations, and in general the model is in good agreement with the observations. Since none of the studies cover several years, it was only possible to validate whether the model captured the seasonal cycle in mosquito numbers. The good agreement with all of the five case studies, means the model probably responds correctly to the environment, and thus it is likely OMaWa can reproduce year-to-year variability as well.

Conclusions

In this paper, the model has been validated using independent data. The model was designed to have a general validity, not being restricted to a specific locality. The study shows the model can capture the distribution and density of *An. gamibiae s.s.* and *An. arabiensis* across Africa, and that it is able to model the seasonal and year-to-year variations in mosquito densities. While the results are robust with respect to the mean distribution and density, there is a sampling bias related to the recent distribution in DR Congo, northern South Africa, southern Namibia, eastern Angola, Central African Republic, eastern Gabon, eastern Chad, South Sudan, and Somalia. This implies models can not be robustly validated in these regions, and that long term changes in the species composition can not be addressed. For the temporal variability, the model has only been validated for Ethiopia, using short time series. Although the model matches well with the observations, most of the time series are short, implying the ability to reproduce year-to-year variations has not been fully addressed.

The results suggest sufficiently high bovine density influences the large-scale distribution of *An. arabiensis*. Similarly, the presence of *An. gambiae s.s.* is linked to the presence of humans, modulated by the density of *An. arabiensis*. Water and air temperature, and availability of breeding sites play secondary roles for the continental distribution of these species, but might be locally important in margin zones. The recent distribution shifts in species composition observed in Kenya [13, 46] might be partially explained by increased mortality of *An. gambiae s.s.* due to interventions like IRS and LLINs. An alternative explanation might be the competitive advantage of *An. arabiensis* efficiently feeding on cattle, and thus suppressing the number of *An. gambiae s.s.* through easier access to blood, and thus reproducing at a higher rate. Over time, these interventions mainly reduce the human biting rate, and not necessarily the longevity of mosquitoes; the most efficient measure in MacDonald's formula of the basic reproductive number. Next, it can be challenged if a reduction of the number of breeding sites, lowering the number of adult mosquitoes per human, would be as efficient, and cost-effective, as IRS and LLINs over time. Studies on the long-term effect of interventions on the mortality rate of mosquitoes is needed to evaluate how these interventions work in practice. The large scale distribution of *An. arabiensis*, and its relation to cattle distribution, also rises the question of this species is using the odour of bovine to navigate, and if this causes of the observed coexistence of *An. arabiensis* and cattle. If this applies on large scales, there are reasons to believe the same mechanisms manifest themselves on small scales. In that case, keeping cattle separate from humans should further reduce the human biting rate in areas where *An. arabiensis* is the dominant species.

Several studies have found out the gene flow of *An. gambiae s.s.* and *An. arabiensis*, and how the species have spread, and is evolving, across Africa. In the model, which gives a good representation of the distribution of the two species, *An. gambiae s.s.* is spreading most efficient on surfaces with continuous human populations, while *An. arabiensis* disperse more easily on surfaces with continuous cattle populations. It is hypothesized the lack of such a human surface between Kenya and Ethiopia can explain the absence of *An. gambiae s.s.* in Ethiopia; - to spread to Ethiopia, there is a need of a more or less continuous human population cover from Lake Victoria to southern Ethiopia, sufficient breeding sites, and temperatures which are not too extreme. Thus, not only climate control the presence and absence of these species, but also the availability of hosts. This has implications for the ability to project the future distribution of the two species.

Before models can be related to improving human health, or guide which interventions are to be applied where, there is a need to understand the system of interest. Validation is an important part of this process. Concluding based on too little data, and basing projections of for example the effects of climate change on models which have not been validated, is dangerous [47], might mislead the public, and lead to less confidence in science. The way forward would be to include effects on interventions. This would allow us to understand how residual spraying and bed nets influence the mosquito populations, and in turn malaria. Incorporating interventions in a continental model requires a) spatial data describing which interventions were applied when, and b) the long term effect of these interventions. Currently such data might exist, but have not been systematized for use by the research community. In these two papers, the focus has deliberately been on the mosquito population. By looking at each component involved in malaria transmission separately, the understanding of the dynamics of malaria can be improved. This process is crucial to robustly estimate how a changing environment and society, has changed, and will change, the premises for malaria transmission.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The work presented here was carried out in collaboration between all authors. BL and AS defined the research theme. TML designed methods and mosquito experiments, did the model runs, analysed the data, interpreted the results and wrote the paper. DK, AS and TML designed the regional climate simulations, and evaluated those. FM, MB and TGM collected data for validation, and contributed with comments on the biology of *An. gambiae s.l.*. All authors have contributed to, seen and approved the manuscript.

Acknowledgements

We are grateful for National Center for Atmospheric Research (NCAR) for making their WRF model available in the public domain. We also thank the Bergen Centre for Computational Science for the computational and other resources provided during this study. This work was made possible by the grants from The Norwegian Programme for Development, Research and Education (NUFU) and the University of Bergen. Thanks to Anne Sinka for making the presence/absence points for *An. gambiae s.l.* available.

References

1. Rogers DJ, Randolph SE, Snow RW, Hay SI: **Satellite imagery in the study and forecast of malaria.** *Nature* 2002, **415:**710–715. [http://www.ncbi.nlm.nih.gov/pubmed/11832960]

2. Levine RS, Peterson AT, Benedict MQ: **Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm.** *Am J Trop Med Hyg* 2004, **70:**105–109. [http://www.ncbi.nlm.nih.gov/pubmed/14993618]

3. Sinka ME, Bangs MJ, Manguin S, Coetzee M, Mbogo CM, Hemingway J, Patil AP, Temperley WH, Gething PW, Kabaria CW, Okara RM, Van Boeckel T, Godfray HCJ, Harbach RE, Hay SI: **The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic precis.** *Parasit Vectors* 2010, **3:**117. [http://www.ncbi.nlm.nih.gov/pubmed/21129198]

4. Lindsay SW, Parson L, Thomas CJ: **Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae* sensu stricto and *An. arabiensis*, using climate data.** *Proc Biol Sci* 1998, **265:**847–854. [http://www.ncbi.nlm.nih.gov/pubmed/9633110]

5. Moffett A, Shackelford N, Sarkar S: **Malaria in Africa: vector species' niche models and relative risk maps.** *PLoS One* 2007, **2:**e824. [http://www.ncbi.nlm.nih.gov/pubmed/17786196]

6. MacDonald G: *The Epidemiology and Control of Malaria*. London: Oxford University Press; 1957.

7. Pluess B, Tanser FC, Lengeler C, Sharp BL: **Indoor residual spraying for preventing malaria.** *Cochrane Database Syst Rev* 2010, CD006657. [http://www.ncbi.nlm.nih.gov/pubmed/20393950]

8. Noor AM, Mutheu JJ, Tatem AJ, Hay SI, Snow RW: **Insecticide-treated net coverage in Africa: mapping progress in 2000-07.** *Lancet* 2009, **373:**58–67. [http://www.ncbi.nlm.nih.gov/pubmed/19019422]

9. Coluzzi M, Petrarca V, Dideco M: **Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae***. *B Zool* 1985, **52:**45–63.

10. della Torre A, Tu Z, Petrarca V: **On the distribution and genetic differentiation of *Anopheles gambiae s.s.* molecular forms.** *Insect Biochem Mol Biol* 2005, **35:**755–769. [http://www.ncbi.nlm.nih.gov/pubmed/15894192]

11. Favia G, della Torre A, Bagayoko M, Lanfrancotti A, Sagnon N, Touré YT, Coluzzi M: **Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation.** *Insect Mol Biol* 1997, **6:**377–383. [http://www.ncbi.nlm.nih.gov/pubmed/9359579]

12. Coetzee M: **Distribution of the African malaria vectors of the *Anopheles gambiae* complex.** *Am J Trop Med Hyg* 2004, **70:**103–104. [http://www.ncbi.nlm.nih.gov/pubmed/14993617]

13. Bayoh MN, Mathias DK, Odiere MR, Mutuku FM, Kamau L, Gimnig JE, Vulule JM, Hawley WA, Hamel MJ, Walker ED: ***Anopheles gambiae*: historical population decline associated with regional distribution of insecticide-treated bed nets in western Nyanza Province, Kenya.** *Malar J* 2010, **9:**62. [http://www.ncbi.nlm.nih.gov/pubmed/20187956]

14. Adamou A, Dao A, Timbine S, Kassogue Y, Yaro AS, Diallo M, Traore SF, Huestis DL, Lehmann T: **The contribution of aestivating mosquitoes to the persistence of *Anopheles gambiae* in the Sahel.** *Malar J* 2011, **10:**151. [http://www.ncbi.nlm.nih.gov/pubmed/21645385]

15. Derua YA, Alifrangis M, Hosea KM, Meyrowitsch DW, Magesa SM, Pedersen EM, Simonsen PE: **Change in composition of the *Anopheles gambiae* complex and its possible implications for the transmission of malaria and lymphatic filariasis in north-eastern Tanzania.** *Malar J* 2012, **11:**188. [http://www.ncbi.nlm.nih.gov/pubmed/22681999]

16. Lunde TM, Korecha D, Loha E, Sorteberg A, Lindtjørn B: **A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis.** *Malar J* 2013, **12:**28. [http://www.ncbi.nlm.nih.gov/pubmed/23342980]

17. Ermert V, Fink AH, Jones AE, Morse AP: **Development of a new version of the Liverpool Malaria Model. II. Calibration and validation for West Africa.** *Malar J* 2011, **10:**62. [http://www.ncbi.nlm.nih.gov/pubmed/21410939]

18. **OMaWa** [ftp://ftp.uib.no/pub/gfi/tlu004/OMaWa/]

19. Coetzee M, Craig M, le Sueur D: **Mapping the distribution of members of the *Anopheles gambiae* complex in Africa and adjacent islands**. *Parasitol Today* 2000, **16:**74–77.

20. **Biodiversity occurrence data published by: Walter Reed Biosystematics Unit Smithsonian Institution, GBIF Data Portal** [data.gbif.org].

21. Hay SI, Sinka ME, Okara RM, Kabaria CW, Mbithi PM, Tago CC, Benz D, Gething PW, Howes RE, Patil AP, Temperley WH, Bangs MJ, Chareonviriyaphap T, Elyazar IRF, Harbach RE, Hemingway J, Manguin S, Mbogo CM, Rubio-Palis Y, Godfray HCJ: **Developing global maps of the dominant *Anopheles* vectors of human malaria.** *PLoS Med* 2010, **7:**e1000209. [http://www.ncbi.nlm.nih.gov/pubmed/20161718]

22. Jack R, Dubinin M, Massing M, Luthman L: **Georeferencer GDAL**. Tech. rep., http://www.gdal.org/ 2011. [http://gis-lab.info/qa/qgis-georef-new-eng.html]

23. Willmott CJ, Matsuura K: **Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance**. *Clim Res* 2005, **30:**79–82. [http://www.int-res.com/abstracts/cr/v30/n1/p79-82/]. [10.3354/cr030079].

24. Pock Tsy JML, Duchemin JB, Marrama L, Rabarison P, Le Goff G, Rajaonarivelo V, Robert V: **Distribution of the species of the *Anopheles gambiae* complex and first evidence of *Anopheles merus* as a malaria vector in Madagascar.** *Malar J* 2003, **2:**33. [http://www.ncbi.nlm.nih.gov/pubmed/14609436]

25. Chauvet G: **Répartition et écologie du complex *Anopheles gambiae* à Madagascar**. *Cah ORSTOM sér Ent Méd* 1969, **VII:**235–278.

26. Kenea O, Balkew M, Gebre-Michael T: **Environmental factors associated with larval habitats of anopheline mosquitoes (Diptera: Culicidae) in irrigation and major drainage areas in the middle course of the Rift Valley, central Ethiopia.** *J Vector Borne Dis* 2011, **48:**85–92. [http://www.ncbi.nlm.nih.gov/pubmed/21715730]

27. Taye A, Hadis M, Adugna N, Tilahun D, Wirtz RA: **Biting behavior and *Plasmodium* infection rates of *Anopheles arabiensis* from Sille, Ethiopia.** *Acta Trop* 2006, **97:**50–54. [http://www.ncbi.nlm.nih.gov/pubmed/16171769]

28. Ye-Ebiyo Y, Pollack RJ, Kiszewski A, Spielman A: **Enhancement of development of larval *Anopheles arabiensis* by proximity to flowering maize (Zea mays) in turbid water and when crowded.** *Am J Trop Med Hyg* 2003, **68:**748–752. [http://www.ncbi.nlm.nih.gov/pubmed/12887038]

29. Balkew M: **Studies on the anopheline mosquitoes of Metehara and surrounding areas in relation to malaria transmission**. *Master's thesis*, Department of Biology, Addis Ababa University 2001.

30. Loha E, Lindtjorn B: **Predictors of *Plasmodium falciparum* Malaria incidence in Chano Mille, South Ethiopia: a longitudinal study.** *Am J Trop Med Hyg* 2012, **87:**450–459. [http://www.ncbi.nlm.nih.gov/pubmed/22826493]

31. Massebo F, Balkew M, Gebre-Michael T, Lindtjørn B: **Blood meal origins and insecticide susceptibility of *Anopheles arabiensis* from Chano in South-West Ethiopia.** *Parasit Vectors* 2013, **6:**44.

32. Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Wang W, Powers JG: **A description of the advanced research WRF version 2**. Tech. rep., The National Center for Atmospheric Research 2005.

33. Kain J, Fritsch J: **Convective parameterization for mesoscale models: The Kain–Fritsch scheme.** *Representation Cumulus Convection Numerical Models Metor Monogr, Am Meteor Soc* 1993, **24:**165–170.

34. Kain J: **The Kain–Fritsch convective parameterization: An update**. *J Appl Meteorology* 2004, **43:**170–181.

35. Zhang C, Wang Y, Hamilton K: **Improved Representation of Boundary Layer Clouds over the Southeast Pacific in ARW-WRF Using a Modified Tiedtke Cumulus Parameterization Scheme**. *Monthly Weather Rev* 2011, **139:**3489–3513. doi:10.1175/MWR-D-10-05091.1.

36. Tiedtke M: **A comprehensive mass flux scheme for cumulus parameterization in large-scale models**. *Monthly Weather Rev* 1989, **117:**1779–1800.

37. Kerah-Hinzoumbe C, Peka M, Nwane P, Donan-Gouni I, Etang J, Same-Ekobo A, Simard F: **Insecticide resistance in *Anopheles gambiae* from south-western Chad, Central Africa.** *Malar J* 2008, **7:**192. [http://www.ncbi.nlm.nih.gov/pubmed/18823537]

38. Brink C: **Malaria control in the Northern Transvaal**. *S Afr Med J* 1958, **32:**800–809.

39. Meillon BD: **Malaria survey of South-West Africa**. *Bull World Health Organ* 1951, **4:**333–417.

40. Service MW: **Contribution to the knowledge of the mosquitoes (Diptera, Culicdae) of Gabon**. *Cah ORSTOM sér Ent Méd* 1976, **3:**259–263.

41. Sylla EH, Kun JF, Kremsner PG: **Mosquito distribution and entomological inoculation rates in three malaria-endemic areas in Gabon.** *Trans R Soc Trop Med Hyg* 2000, **94:**652–656. [http://www.ncbi.nlm.nih.gov/pubmed/11198649]

42. Mouchet J: **Survey of Potential Yellow Fever Vectors in Gabon and Chad**. *World Health Organ* 1970, **VBC71.279:**1–10.

43. Mourou JR, Coffinet T, Jarjaval F, Pradines B, Amalvict R, Rogier C, Kombila M, Pages F: **Malaria transmission and insecticide resistance of *Anopheles gambiae* in Libreville and Port-Gentil, Gabon.** *Malar J* 2010, **9:**321. [http://www.ncbi.nlm.nih.gov/pubmed/21070655]

44. Elissa N, Karch S, Bureau P, Ollomo B, Lawoko M, Yangari P, Ebang B, Georges AJ: **Malaria transmission in a region of savanna-forest mosaic, Haut-Ogooué, Gabon.** *J Am Mosq Control Assoc* 1999, **15:**15–23. [http://www.ncbi.nlm.nih.gov/pubmed/10342264]

45. Stensrud D: *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*. Cambridge University Press; 2007. [http://books.google.no/books?id=lMXSpRwKNO8C]

46. Mwangangi JM, Mbogo CM, Orindi BO, Muturi EJ, Midega JT, Nzovu J, Gatakaa H, Githure J, Borgemeister C, Keating J, Beier JC: **Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years.** *Malar J* 2013, **12:**13. [http://www.ncbi.nlm.nih.gov/pubmed/23297732]

47. Lunde TM, Bayoh MN, Lindtjørn B: **How malaria models relate temperature to malaria transmission.** *Parasit Vectors* 2013, **6:**20. [http://www.ncbi.nlm.nih.gov/pubmed/23332015]

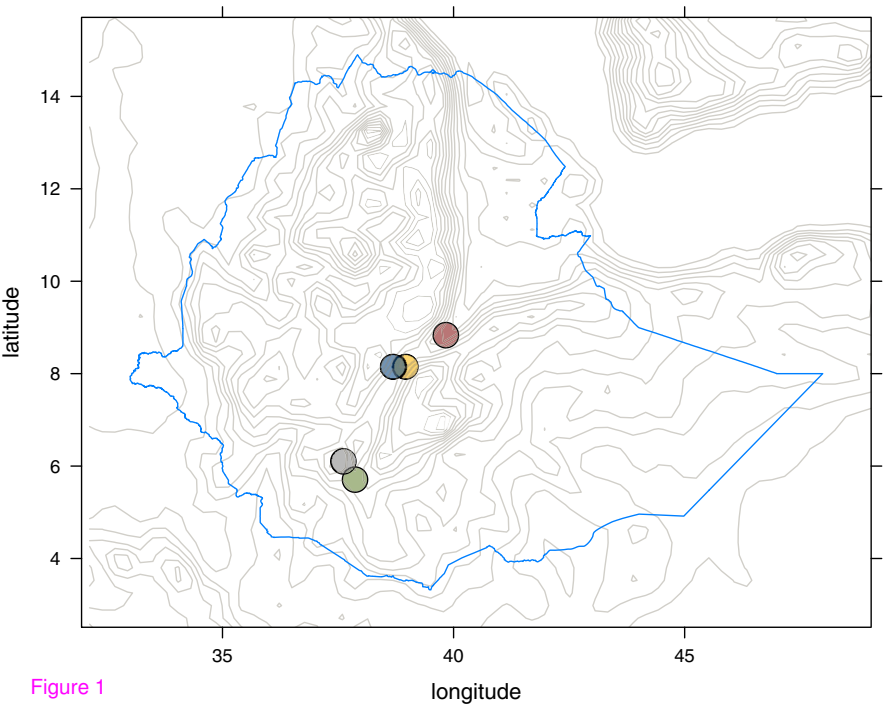

Figure 1

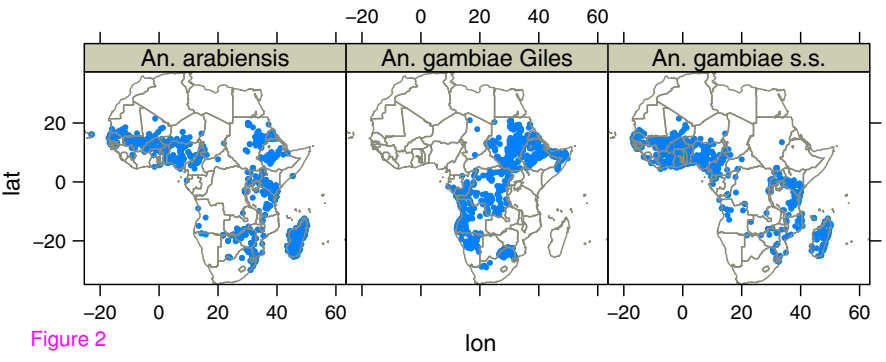
An. arabiensis
An. gambiae Giles
An. gambiae s.s.
lat
lon
20
0
−20
−20
0
20
40
60

Figure 2

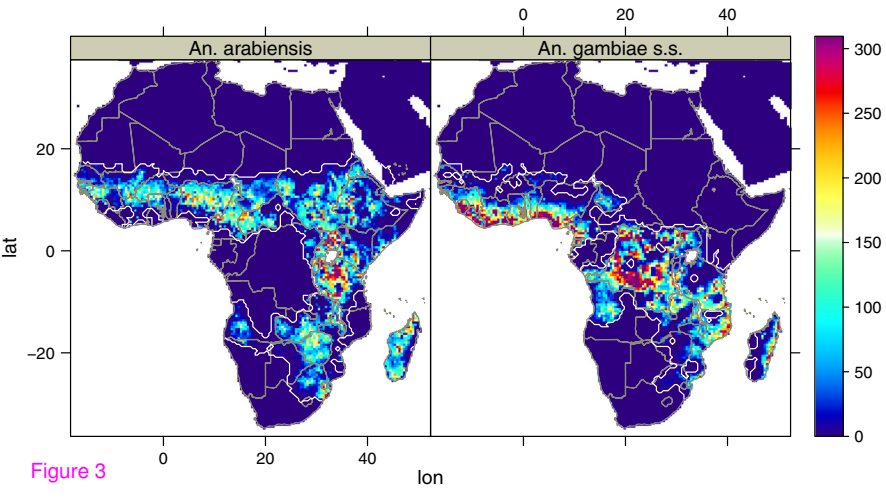
An. arabiensis
An. gambiae s.s.
lat
lon
20
0
−20
0
20
40
300
250
200
150
100
50
0

Figure 3

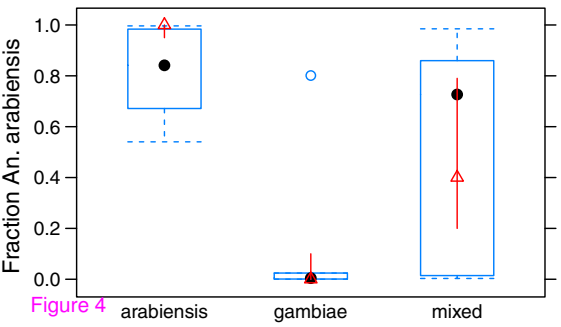
Fraction An. arabiensis
1.0
0.8
0.6
0.4
0.2
0.0
arabiensis
gambiae
mixed

Figure 4

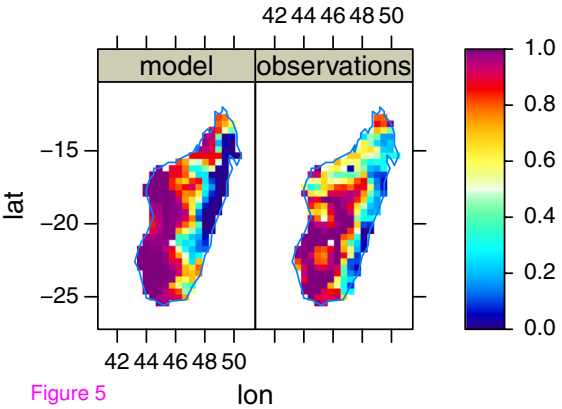
model
observations
42 44 46 48 50
−15
−20
−25
lat
42 44 46 48 50
lon
1.0
0.8
0.6
0.4
0.2
0.0

Figure 5

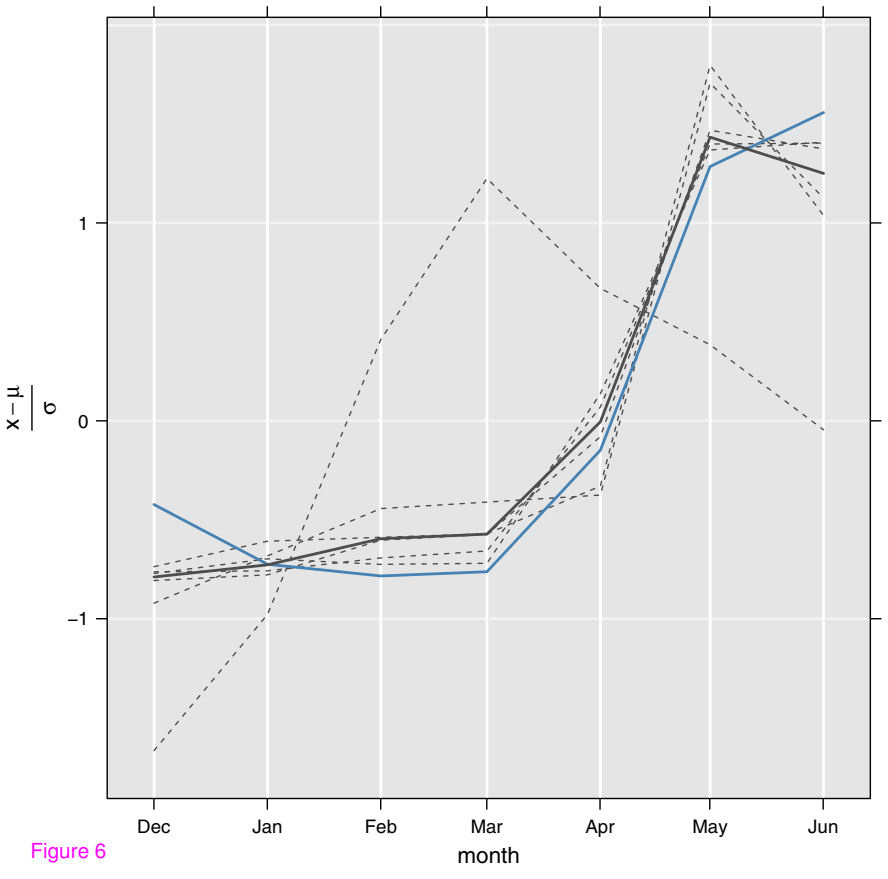

Figure 6

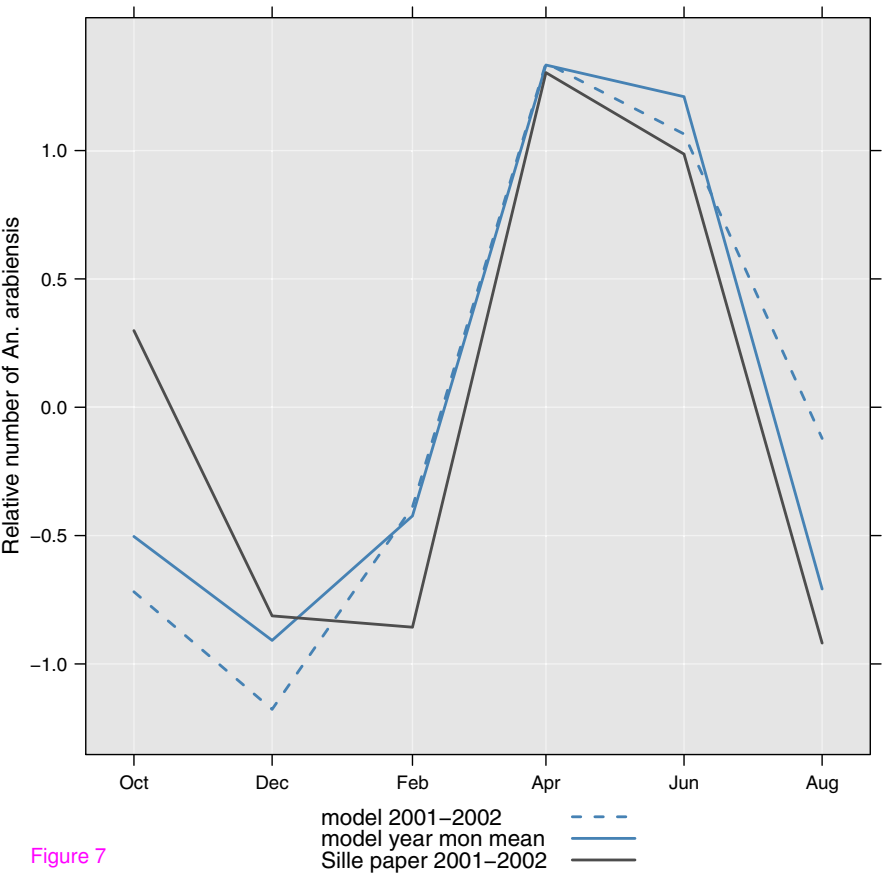
Relative number of An. arabiensis
1.0
0.5
0.0
−0.5
−1.0
Oct
Dec
Feb
Apr
Jun
Aug
model 2001−2002
model year mon mean
Sille paper 2001−2002

Figure 7

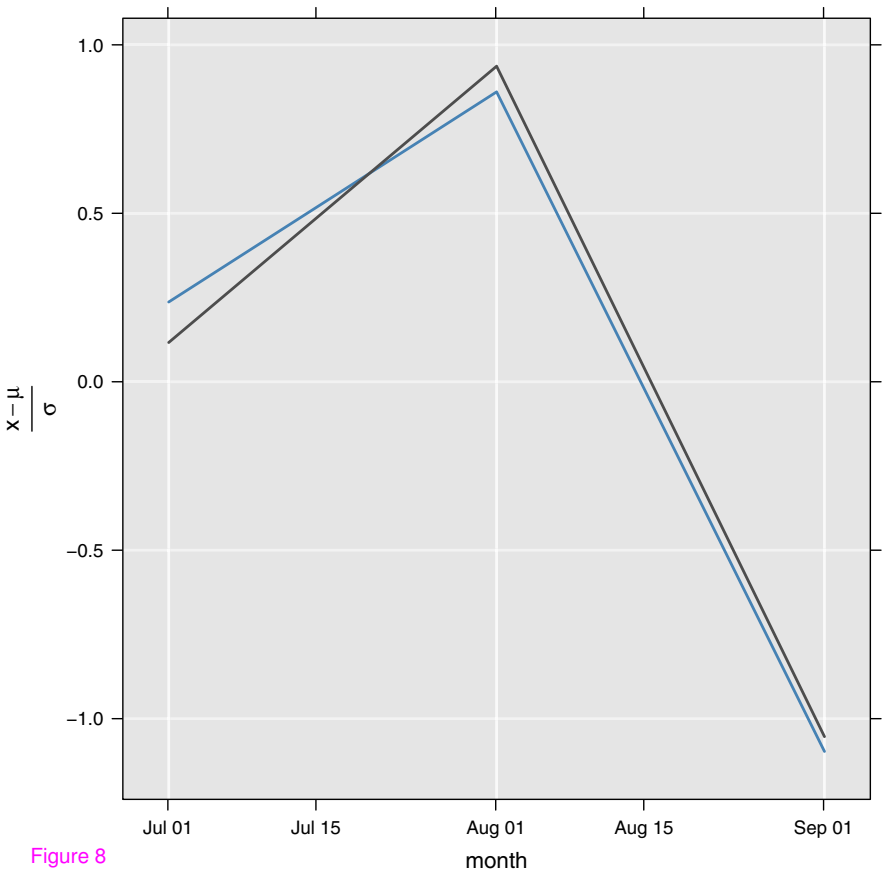
1.0
0.5
0.0
−0.5
−1.0
$\frac{x-\mu}{\sigma}$
Jul 01
Jul 15
Aug 01
Aug 15
Sep 01
month

Figure 8

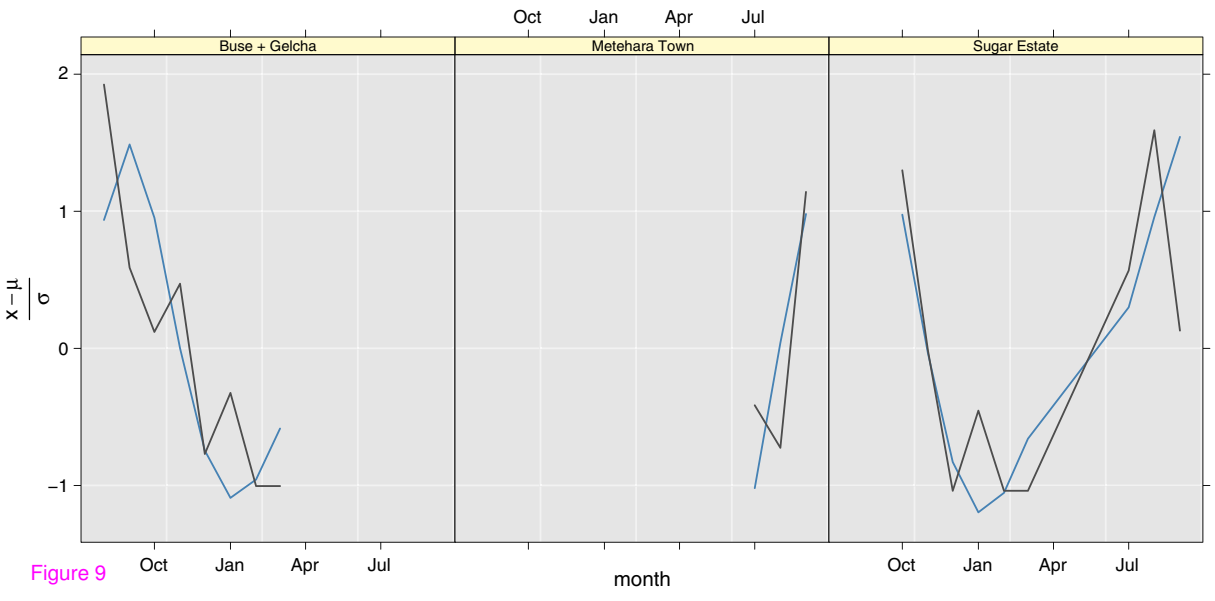
Oct
Jan
Apr
Jul
Buse + Gelcha
Metehara Town
Sugar Estate
2
1
0
−1
x − μ / σ
Oct
Jan
Apr
Jul
month
Oct
Jan
Apr
Jul

Figure 9

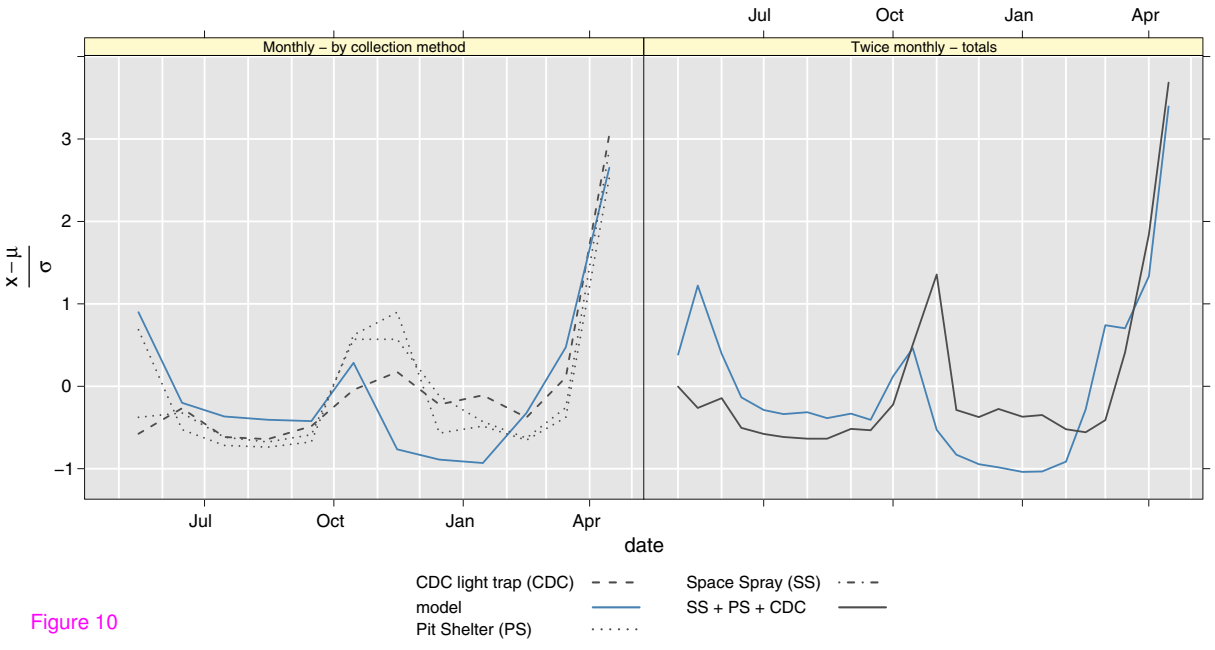
Jul
Oct
Jan
Apr
Monthly – by collection method
Twice monthly – totals
3
2
1
0
−1
x − μ / σ
Jul
Oct
Jan
Apr
date
CDC light trap (CDC)
model
Pit Shelter (PS)
Space Spray (SS)
SS + PS + CDC

Figure 10

Paper III

6.3 Cattle and climate in Africa: How climate variability has influenced national cattle holdings from 1961-2008

Torleif Markussen Lunde and Bernt Lindtjørn

PeerJ 2013 55

PeerJ

Cattle and climate in Africa: How climate variability has influenced national cattle holdings from 1961–2008

Torleif Markussen Lunde[1,2] and Bernt Lindtjørn[2]

[1] Bjerknes Centre for Climate Research, University of Bergen/Uni Research, Norway
[2] Centre for International Health, University of Bergen, Norway

ABSTRACT

The role of cattle in developing countries is as a source of high-quality food, as draft animals, and as a source of manure and fuel. Cattle represent important contribution to household incomes, and in drought prone areas they can act as an insurance against weather risk. So far, no studies have addressed how historical variations in temperature and rainfall have influenced cattle populations in Africa. The focus of this study is to assess the historical impact of climate variability on national cattle holdings. We reconstruct the cattle density and distribution for two time periods; 1955–1960 and 2000–2005. Based on estimates from FAO and official numbers, we generated a time series of cattle densities from 1961–2008, and compared these data with precipitation and temperature anomalies for the same period. We show that from 1961–2008 rainfall and temperature have been modulating, and occasionally controlling, the number of cattle in Africa.

Subjects Agricultural Science, Ecosystem Science, Entomology, Global Health, Infectious Diseases
Keywords Cattle, Climate, Africa, Malaria, Precipitation, Temperature

Submitted 9 February 2013
Accepted 3 March 2013
Published 19 March 2013

Corresponding author
Torleif Markussen Lunde,
torleif.lunde@cih.uib.no

Academic editor
Jianhua Xu

Additional Information and Declarations can be found on page 15

DOI **10.7717/peerj.55**

INTRODUCTION

Background

Since the 1960s there has been a period of climatic change with global land-surface temperature rising about 0.5–0.6 °C (*Hansen, Sato & Ruedy, 2012*). Although some studies have addressed effect of climate variability on cattle populations (*Angassa & Oba, 2007*; *Cossins & Upton, 1988*; *Desta & Coppock, 2002*; *Oba, 2001*), no studies have described how historical variations in temperature and rainfall influence cattle populations in Africa.

FAO estimates the number of cattle in Africa during the period 2001 to 2010 is twice the estimates for the years 1961–1970. But, how variations in the climate influence cattle depend on the ecological setting, and how variations in cattle influence the population, depend on the availability of alternative energy sources as well as the cultural setting. The role of cattle in developing countries is as a source of high-quality food, as draft animals, and as a source of manure and fuel (*Scoones, 1992*; *Taddesse et al., 2003*). Cattle represent important contribution to household incomes (*Seo & Mendelsohn, 2006*), and in drought prone areas they can act as an insurance against weather risk (*Fafchamps & Gavian, 1997*).

Most of the cattle in Africa are in arid and semiarid areas. In the forested humid areas of humid West-Africa, as well as Democratic Republic of Congo, the tsetse tolerant N'Dama and West African Shorthorn breeds are common. The most common cattle breed is, however, the East and West African Zebu, which make up the majority of African cattle (*Deshler, 1963*). Close to Lake Chad, the heat tolerant Kuri breed can be found, although the density has declined since the 1950s (*Tawah, Rege & Aboagye, 1997*), and in East Africa, the Sanga can be found on the western branch of the Great Rift Valley. In South Africa the Afrikander is common. The different cattle types probably represent mixtures of breeds introduced at various times (*Deshler, 1963*).

Many production systems supply water from ponds and rivers during the wet season, and the need for watering increases with higher temperatures (*Seif, Johnson & Lippincott, 1979*). The IPCC 2007 report concluded that changes in range-fed livestock numbers in any African region will be directly proportional to changes in annual precipitation (*Intergovernmental Panel on Climate Change, 2007*).

Rötter and van de Geijn (1999) discussed how changes in climate potentially can influence livestock;

- Feed production (Direct effect of CO_2, temperature and precipitation)
- Animal health (Direct effect of feed production, heating through temperature, and watering via precipitation and evaporation)
- Diseases (Indirect effect of stress, parasites, and vector borne diseases).

The coming century, it is virtually certain temperatures will increase, and that the intensity of precipitation will change (*Min et al., 2011*). How the cattle has been, and will be influenced directly through climate variability, and indirectly through parasites and vector borne diseases is still uncertain. The lack of certainty in projected absolute changes in precipitation amounts and how cattle respond to climate, makes it difficult to to predict impacts of climate change. It is therefore necessary to understand the historical impact of climate on cattle before projecting future impacts and developing adaptation strategies (*Hoffmann, 2010*).

Feed production and consumption

The growth of plants are directly influenced by the atmospheric CO_2 concentrations, with two metabolic pathways; C3 and C4 (*Stokes et al., 2010*). While the productivity is expected to increase for C3 plants, quality, productivity and digestibility is expected to decrease with increasing CO_2 concentrations. C4 plants are probably less affected (*Stokes et al., 2010*). In the subtropical Australia it has been hypothesized that lower precipitation can be compensated for by the benefits of increased CO_2 (*Henry et al., 2012*). The compensating effect in tropical Africa is uncertain.

A study by Seif et al. (*Seif, Johnson & Lippincott, 1979*) showed Zebu water consumption increased by 58% when temperature increased from 10 to 31 °C. This is only 2.8% increase per degree, if we consider this as a linear process. At higher temperatures, feed

consumption decreases (*Seif, Johnson & Lippincott, 1979*) and fertility increases (*Jöchle, 1972*). Lower food consumption can be of importance in the dry season. This is in line with the perception of farmers in the savanna zone of central Senegal, who say low temperatures may lead to fodder shortage (*Mertz et al., 2009*).

Indirect consequences of climate

Although climate influence the vegetation in these semiarid environments, the coupling between climate and animal numbers might not be as straight forward as grass production and the need for watering; in tsetse infested areas, high temperatures might reduce the vector populations and cause a reduction in animal trypanosomiasis (*Hall et al., 1984*; *Terblanche et al., 2008*), in Nigeria increased rainfall has been linked to outbreaks of blackquarter (*Bagadi, 1978*), and in the eastern part of Africa, where east coast fever prevail, climate variability can be related to the survival and reproductive success of the tick *Rhipicephalus appendiculatus* (*Branagan, 1973*), and as the development of the *Theileria parasite* (*Young & Leitch, 1981*). Livestock also play a role in malaria transmission by creating favourable environments and blood meals for *Anopheles arabiensis*. We have previously shown that understanding fluctuations in cattle populations is important to assess the historical and future distribution of two of the most efficient vectors of malaria in Africa (*Lunde et al., 2013a*; *Lunde et al., 2013b*). Tirados et al. (*2011*) showed that a cattle herd of 20 heads outside a house reduced the number of *Anopheles arabiensis* landing on humans by 50%. It has also been speculated that certain malaria epidemics in India and Somalia can be explained by herds of livestock being decimated during drought years (*Choumara, 1961*; *Cragg, 1923*).

To quantify the historical impact of climate variability on cattle in Africa, we construct a statistical model which include precipitation (P), temperature during the rainy season (T_w), and temperature during the dry season (T_d). We also adjust for armed conflicts and include a sigmoid-shaped Gompertz curve which represents an increase in infrastructure over time (number of herders/farmers, provision of wells/water stations, veterinary services) that allows an increased carrying capacity over time without population density-dependence.

METHODS

The main aim of this paper is to quantify the effect climate variability has had on national cattle holdings from 1961–2008. To do so we specify a linear model under the assumption of normally distributed errors and constant variance:

$$W_{n_{y,c}} = \beta + m_1 \cdot G_e(a,b) + m_2 \cdot CF_{y,c} + m_3 \cdot T_{d_{y,c}} + m_4 \cdot T_{w_{y,c}} + m_5 \cdot P_{y,c} + \epsilon \quad (1)$$

where

β = Intercept

$G_e(a,b)$ = Gompertz function with parameters a and b

$CF_{y,c}$ = Armed conflict weighted by cattle density within a country

$T_{d_{y,c}}$ = Five year weighted mean temperature anomalies in the dry season, spatially weighted by cattle density within a country

$T_{w_{y,c}}$ = Five year weighted mean temperature anomalies in the wet season, spatially weighted by cattle density within a country

$P_{y,c}$ = Five year weighted annual mean monthly precipitation anomalies, spatially weighted by cattle density within a country

ϵ = Error.

In the following sections we explain how the spatial weights are constructed, the data sources, and corrections done to the data.

Construction of spatial weights

In 1963 Walter Deshler published a map of cattle distribution in Africa. The map is complete, except from two countries with large cattle populations; Ethiopa and Upper Volta (Burkina Faso). Data was also missing from Gabon and Spanish Sahara (Western Sahara), but these territories were probably empty of cattle. For Ethiopia and Upper Volta (Burkina Faso) we used FAO's estimate of 2005 cattle density and adjust the totals to Faostat's estimate for 1961. This process is described later.

We geo-referenced the raster map published by Deshler to a Miller Oblated Stereographic projection. Thereafter, the country borders, coastlines and rivers were manually removed, only leaving the dots in the maps. One point in the original map represents 5000 cattle (heads). In the rasterized version of the map, one point would consist of a group of pixels. The geo-referenced raster is a one band grayscale raster with values from 0 (black) to 255 (white). First, pixels with values grater than 200 were removed. Such a high threshold was chosen based on manually checking the distribution of representative dots. The remaining points could now be treated as probable candidates of being an observation of 5000 cattle. To automatically identify groups of points, we applied the Partitioning Around Medoids (PAM) algorithm (*Kaufman and Rousseeuw, 1990*). Since we knew the approximate total number of cattle in each country, and we also knew each point represent 5000 heads, the expected number of clusters was $\approx FAO_{tot,country} \cdot 5000^{-1}$, where $FAO_{tot,country}$ is the FAO estimate of national cattle holdings. To speed up the algorithm we split the computation for each country in hexagonal tiles. After running the PAM-algorithm for all countries, except Ethiopia, Upper Volta, Gabon, and Spanish Sahara, we manually removed or added points which either were duplicates, or were not detected by the algorithm.

After the raster map had been converted to clean points, we used a spherical non-parametric estimator method to calculate point densities. Such kernel estimators were developed to omit problems with discontinuities of the estimates dependent on the bin positions. In this work we used a spherical kernel developed by Kevin Hodges (*Hodges, 1996*) (with power $m = 1$). This is a computational efficient kernel designed to derive storm track statistics. It is defined locally so that the influence of a point is restricted to a local region.

To choose a global smoothing parameter we maximize the cross-validation function suggested by Diggle and Fisher (*Diggle & Fisher, 1985*):

$$\Gamma^d(C_n) = -\frac{1}{n}\sum_{i=1}^{n}\log_e[\hat{f}_{-i}(X_i, C_n)] \tag{2}$$

where

$$\hat{f}_{-i}(X_i, C_n) = \frac{1}{n-1}\sum_{j\neq i}^{n} K(X_j \cdot X_i, C_n). \tag{3}$$

Still the greatest value of the local, and hence global, smoothing parameter which is described later, is restricted by the grid spacing. If the spherical cap is too small, some points will not be included in the density estimation, and κ must therefore be restricted.

For the maps produced in this paper the value of $\kappa = 21907.45(\tilde{\kappa} = 1.000046)$ which is equivalent to an arc bandwidth radius of 0.55°. This parameter is then adaptively modified based on the ideas of a pilot density estimate and cross validation as described by Hodges (*Hodges, 1996*). If the smoothing parameter is

$$\kappa(\tilde{\kappa} = 1 + 1/\kappa) \tag{4}$$

the local smoothing parameter is determined as:

$$\kappa_{N,i} = \kappa_N\left(\frac{\hat{f}_p(X_i)}{g}\right)^{\gamma} \tag{5}$$

where κ_N is the global smoothing parameter, $\hat{f}_p(X_i)$ is the pilot estimate at each point X_i, g is the geometric mean of the pilot densities. The γ parameter is subjectively chosen to be 0.5 which Abramson (*Abramson, 1982*) showed (in the Cartesian domain) give lower bias than normal fixed bandwidth estimates.

After smoothing the cattle observation we normalize the densities to match $5000 \cdot n$.

To estimate a comparable cattle density around year 2000 we converted the FAO observed bovine density (census data) (*Robinson & Fao's Animal Production and Health Division, 2011*) to points, each point equal to 5000 animals. First, the FAO raster was converted to polygons using the Geospatial Data Abstraction Library (GDAL) (*The Open Source Geospatial Foundation*). In cases where the modulus of the sum inside the polygon is non-zero, the probability of sampling an additional point (Z_{h_i+1}) is the modulus divided by 5000. Next, we construct 50 realizations of the maps. Each time we sample n_i completely spatial random points (*Bivand, Pebesma & Gomez-Rubio, 2008*; *Pebesma & Bivand, 2005*) within each polygon, and estimate the density as described earlier. Mostly, the observations from 2000 are aggregated to district level, and hence the observations do not have the same quality with respect to spatial distribution as those of Walter. The global smoothing parameter is held constant, while the local smoothing parameter will vary for each of the 50 estimates.

This method is used to provide a best-guess estimates of the cattle densities around 1960 and 2000 without making any assumptions about dependencies on land use or climate. There are two good reasons for doing this. First of all we do not know how the cattle distribution is related to climate within individual countries. Secondly; if we had already assumed that cattle distribution and density was dependent on climate or land use it would be hard to justify relating this data set to those variables.

Time series of national cattle holdings and spatially weighted time series of climate

FAOstat (*FAO, 2011*) reports the estimated number of cattle heads within a country from 1961. We relate this to the annual mean temperature and precipitation from University of Delaware air temperature and precipitation and repeat the same analysis with CRU v3.1. The data sets were interpolated to the same grid as the cattle densities using distance weighted interpolation. It should be noted that for example Madagascar, Somalia, and Ethiopia have very few weather observations. In countries with few observations, the results are less robust. Since the data from FAOstat is reported on national scale we need to aggregate the temperatures and precipitation to the same levels. To do this we use the newly constructed cattle densities. Each value inside the country (c) boundaries are given a weight ($W_{i,y,c}$) based on the cattle density.

$$W_{i,y,c} = \frac{X_{i,y,c}}{\sum_{i=1}^{n} X_{i,y,c}} \tag{6}$$

where the cattle density in year (y) is linearly interpolated between 1960 and 2000.

The weighted mean temperature anomalies (T) or precipitation anomalies (P) for each country is then (given for T here):

$$T_{y,c} = \sum_{i=1}^{n} T_{i,y,c} \cdot W_{i,y,c}. \tag{7}$$

Standardized anomalies can be calculated from the actual temperature or precipitation by dividing the difference from the mean on the standard deviation, or more specifically (x is actual temperature and n is the number of observations):

$$T_{y,c} = \frac{x - \frac{1}{n}\sum x}{\sqrt{\frac{\left(x - \frac{1}{n}\sum x\right)^2}{n}}}. \tag{8}$$

To account for the weather the past years we do an additional time smoothing with a kernel, K

$$K = [0.016, 0.127, 0.265, 0.327, 0.265]. \tag{9}$$

And the new $T_{y,c}$ becomes

$$T_{y,c} = \sum_{i=0}^{4} T_{y-i,c} \cdot K[5-i]. \tag{10}$$

Armed conflicts

To adjust for conflicts (CF) which might have influenced the cattle densities (*Brück & Schindler, 2009*), we use the armed conflict site data set from UCDP/PRIO. This data set contains year, coordinates (L) and radius in km (r) of conflicts (CF) from 1946 to 2005. On the same grid we define $CF_{i,j}$ as a function of distance (D) from L and r.

$$CF_{i,j} = \frac{D^4}{r^4} \tag{11}$$

where $CF_{i,j} > 0$.

Allowing increased carrying capacity over time without population density-dependence

We introduce a sigmoid-shaped growth curve which represents an increase in infrastructure (number of herders/farmers, provision of wells/water stations, veterinary services). This function allows an increased carrying capacity over time without population density-dependence. We use a Gompertz function, and adjust the time and scale of the data. A description of the procedure is following in the next lines.

We normalize time (t_n) from -2 to 2 (so that $1961 = -2$ and $2008 = 2$). This normalization is done based on the properties of the Gompertz function. The cattle numbers (W) from Faostat are also normalized (W_n) to range from $\min(W_n) = 0$ to $\max(W_n) = 1$:

$$W_n(t) = (\max(W) - \min(W))^{-1} \cdot W(t) + \left(1 - \frac{\max(W)}{\min(W)}\right)^{-1} \tag{12}$$

where $W(t)$ is the number of cattle at time t, and $W_n(t)$ is the scaled number of cattle at time t.

Next, we estimate a and b using nonlinear weighted least-squares to optimize the function:

$$G_e(a,b) = a \cdot e^{(b \cdot e^{(-t_n)})} \tag{13}$$

and

$$W_n = G_e(a,b) + \epsilon. \tag{14}$$

Depending on the country, the cattle numbers reported by FAO might be based on estimates. Since these estimates are more unreliable than actual observations we want to give less weight to those. To define the weights we apply a two way search to find the minimum number of years since the last observation (Ω). For example if there were

observations in 1999 and 2003, but not in 2000–2002, the weights for 1999, 2000, 2001, 2002 and 2003 would be $1^{-1}, 2^{-1}, 3^{-1}, 2^{-1}, 1^{-1}$.

Using Eq. (1) we use stepwise model selection by Bayesian information criterion (BIC) to estimate the model which explain most of the variance. A few cases suggested that war had a positive effect on cattle numbers. Since we believe this is unreasonable, war having a positive effect on national cattle holdings was not allowed in the model.

We assume errors follow a normal distribution, $\epsilon \sim \mathcal{N}(0, \sigma^2)$, and test this assumption by applying a Shapiro-Wilk test of normality, as well as investigating the normal QQ plot of the residuals. To test for heteroscedasticity, we applied a Breusch-Pagan (Cook and Weisberg) test.

Data corrections

As mentioned, 1960 data was missing from Gabon, Spanish Sahara (Western Sahara), Ethiopia and Upper Volta (Burkina Faso). For Ethiopia and Upper Volta (Burkina Faso) we use FAO's estimate of 2005 cattle density (Gridded livestock of the world (GLW) (*Wint & Robinson, 2007*)), and adjust the totals to Faostat's estimate for 1961. For these countries. Since GLW was released, additional data has become available for Afder, Gode, Korahe, Warder, Fik, Degehabur, and Shiniele in Ethiopia (*Central Statistical Authority, 2004*). GLW is updated with this information. This data set should roughly give an estimate of the cattle distribution and density for 2000–2005. Since Ethiopia was classified as Ethiopia PDR in 1961 we used the total of Ethiopia and Eritrea in 2000 to match the 1961 Ethiopia PDR total. To make pseudo points for the four countries we randomly sampled (*Bivand, Pebesma & Gomez-Rubio, 2008*; *Pebesma & Bivand, 2005*) nearest integer of administrative zone totals divided by 5000 points in each zone.

For the present day estimates it should be noted that data for Mauritania was missing. FAO does report the estimated total, and to estimate the density for Mauritania we distribute the total in the areas which are not reported as zero. There are two major areas in Mauritania which are likely to have cattle. The major area is to the south, while a smaller area is located around 21.5 North–6.6 degrees East. In the latter area we assume the density to be approximately equal to the density on the Mali side of the border, while the remaining is equally distributed in the Southern area.

Non-technical summary of the methods

We estimated a continuous surface of cattle densities and distribution from point observations. From this data we calculated annual mean precipitation, and dry and wet season temperature anomalies where the cattle were present. These time series were correlated with the official cattle holdings for each country using a linear model, giving more weight to actual observations, and less weight to estimates. In addition, we have included the Gompertz function to account for adaptation and growth. We also adjust for armed conflicts which has been important for cattle numbers in, for example, Mozambique (*Brück & Schindler, 2009*).

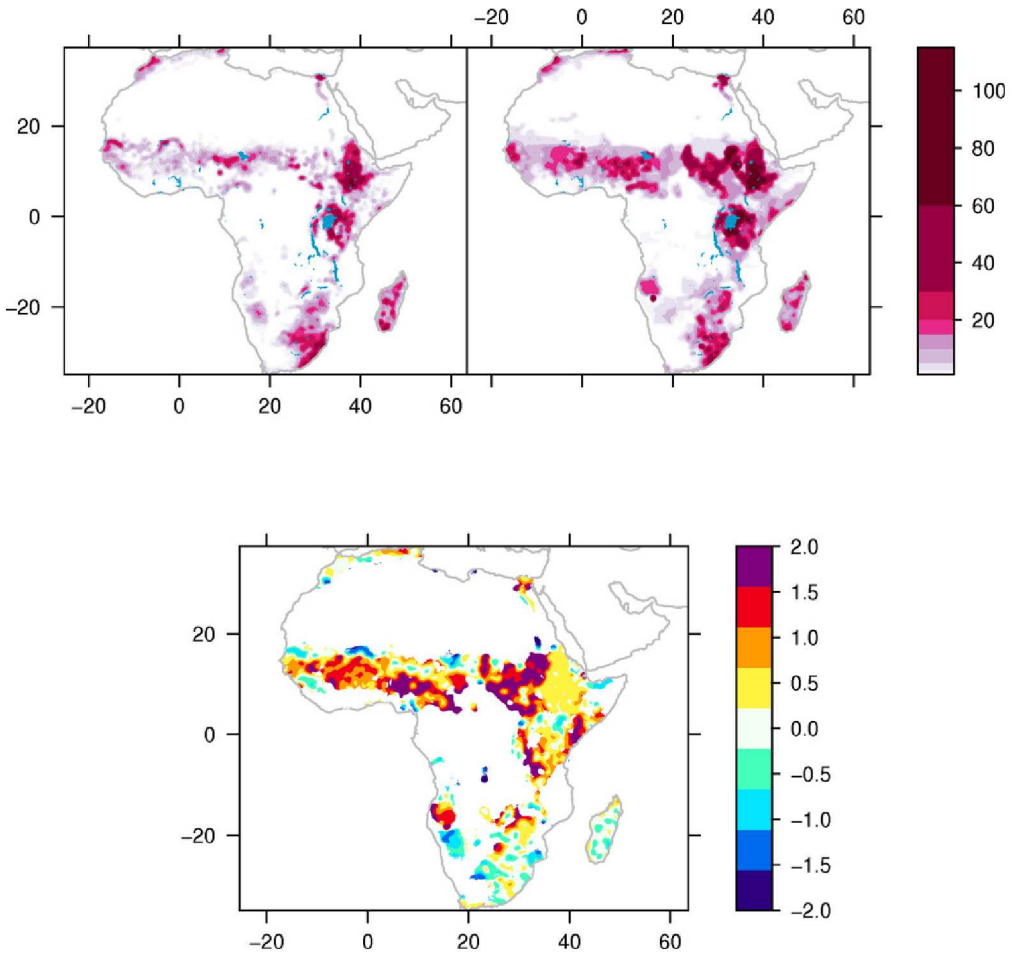

Figure 1 Cattle density and distribution 1960s and 2000s. Estimated cattle (heads per square kilometre) density around 1960 (upper left), 2000 (upper right) and difference (2000–1960, bottom) relative to the mean.

RESULTS

Figure 1 shows the estimated cattle distribution and density in the 1960s and 2000s, and their difference over the mean density. In general, most areas have more cattle today than in the early phase of global warming. But there are some exceptions. The declines are mainly observed in Northern Somalia, and Northern Kenya, around the Niger river, Mauritania, parts of South Africa, Mozambique, Namibia, and Madagascar. Although there are large variations, the main signal is that dry areas have seen a reduction in the amount of cattle opposite to wetter areas, which have seen an increase.

Figure 2 shows the sign of the slope for precipitation (P), wet season temperature (T_w), and dry season temperature (T_d). Only values where the degree of confidence is greater than 95%, and the model could explain more than 30% of the variance are shown. In most of the countries where precipitation is significantly correlated with cattle numbers, increased annual rainfall is associated with increased cattle numbers. The exceptions,

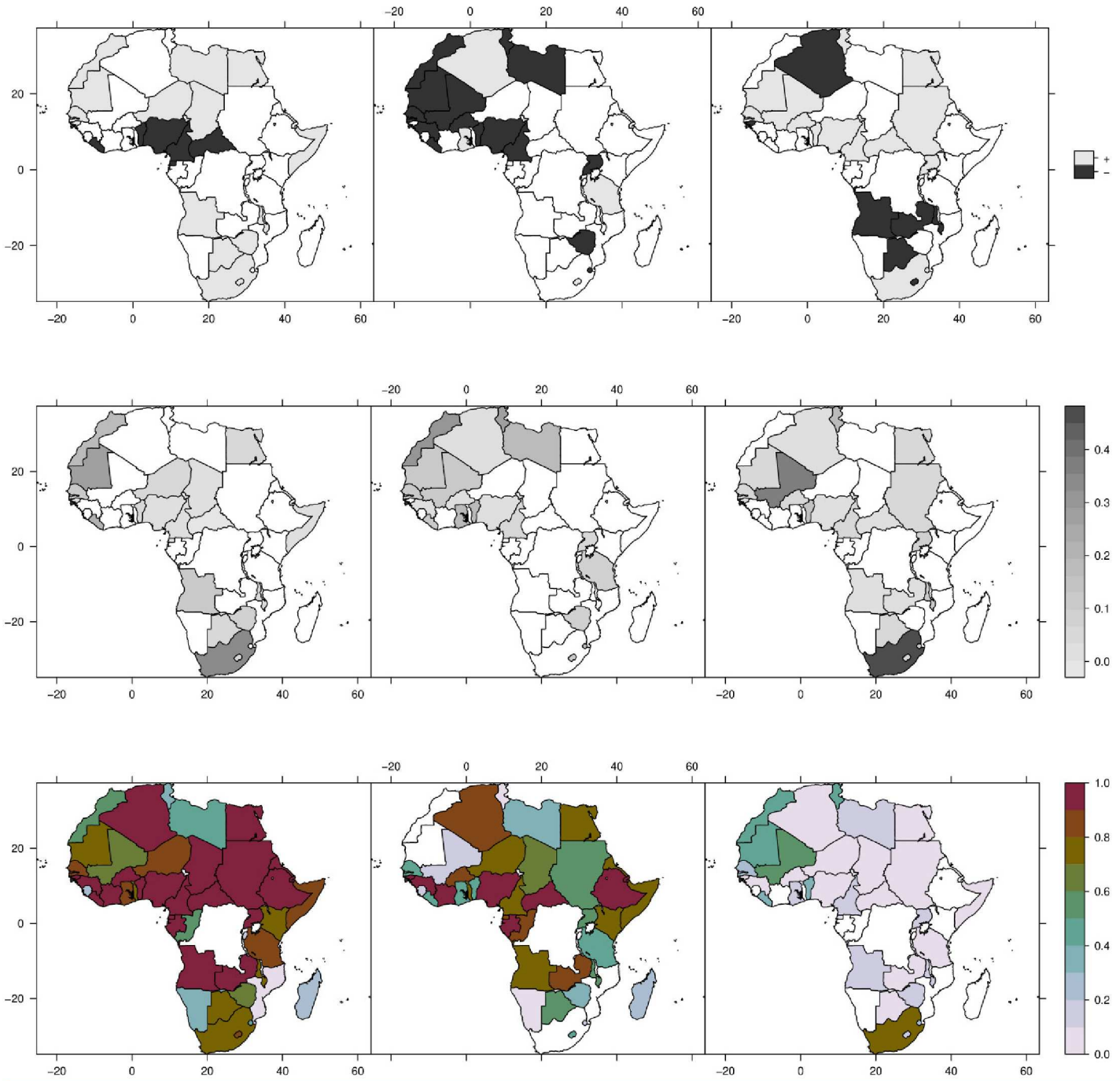

Figure 2 Effect of climate variability on cattle holdings. Upper panel (left to right): Sign of slope for precipitation (P), wet season temperature (T_w), and dry season temperature (T_d). Positive values means increased precipitation/temperature is associated with increased number of cattle and vice versa. Middle panel (left to right): Percent variance explained by precipitation, wet season temperature and dry season temperature. Lower panel (left to right): Total variance explained by the model, variance explained by variability in Gompertz function, and variance explained by climate variability.

which have notable cattle populations, are Cameroon, Nigeria, and Benin, countries which have warm and relatively humid climates.

On the other hand, the influence of temperature during the wet season demonstrate a more diverse pattern. Temperature can influence cattle through direct heating, through vector borne diseases, and by modulating evaporation. However, the most dominant factors controlling evaporation is the vapor content in the air and the turbulence which can transport vapor away from the surface. There are eleven countries where temperature explain more than 10% of the variance, where warmer wet season temperatures have a positive impact in Lesotho (10%) and Ghana (17%), and a negative impact in Zimbabwe (7%), Uganda (5%), Benin (17%), Mali (13%), Senegal (11%), Mauritania (13%), Libya (17%), Morocco(/Western Sahara) (30%), Liberia (10%), Gambia (10%), and Tunisia (21%).

The pathway from temperature to cattle in these countries is probably diverse. For example in Tunisia and Morocco theileriosis is influenced by temperature. In the drier and warmer countries, the need for water increases as the temperature increases.

Outside the rainy season, temperature can interact differently. At higher temperatures, feed consumption decreases (*Seif, Johnson & Lippincott, 1979*) and fertility increases (*Jöchle, 1972*). Lower food consumption can be of importance in the dry season. According to the model, Western Sahel seems to be especially sensitive to the dry-season temperature. This is in line with the perception of farmers in the savanna zone of central Senegal, who say cold temperatures may lead to fodder shortage (*Mertz et al., 2009*).

In the least climate sensitive countries, the Gompertz model can explain most of the variance. We interpret the greatest climate sensitivity will be seen where resource limits are reached, and the resource limits are then modulated by climate (Fig. 3A). Thus, we expect the present day climate insensitive countries to be more vulnerable to climate variability as the cattle populations converge toward the carrying capacity limit; One way to adapt to climate variability, from season-to-season and year-to-year, is to move cattle to new areas. As the number of cattle increases, there will be more competition for food and water, and the strategy of moving cattle might be less successful if most areas are already occupied. In this case, use of concentrate feeds might be a viable alternative, while the access to water might still make the cattle populations vulnerable.

Although temperature is important, precipitation show a more consistent pattern, with a positive effect in the drier countries, and a negative effect in wetter countries (Fig. 3B). We obtained similar results as the ones shown in Fig. 3B when using a mixed effect model (*Bates, Maechler & Bolker, 2011*). In this analysis, countries were classified based on the UNEP index of aridity (*Middleton, United Nations Environment Program & Thomas, 1997*), taking into account the potential evapotranspiration and average annual precipitation. The mixed linear model, with the same form as the linear model described earlier, was repeated for cattle belonging to each of aridity classes with country as a random variable. In this analysis, the fixed effect of one standard deviation increase in precipitation were; Arid: 0.60 (CI 95% 0.47, 0.74), Semi-arid: 0.11 (0.01, 0.22), Dry subhumid: 0.14

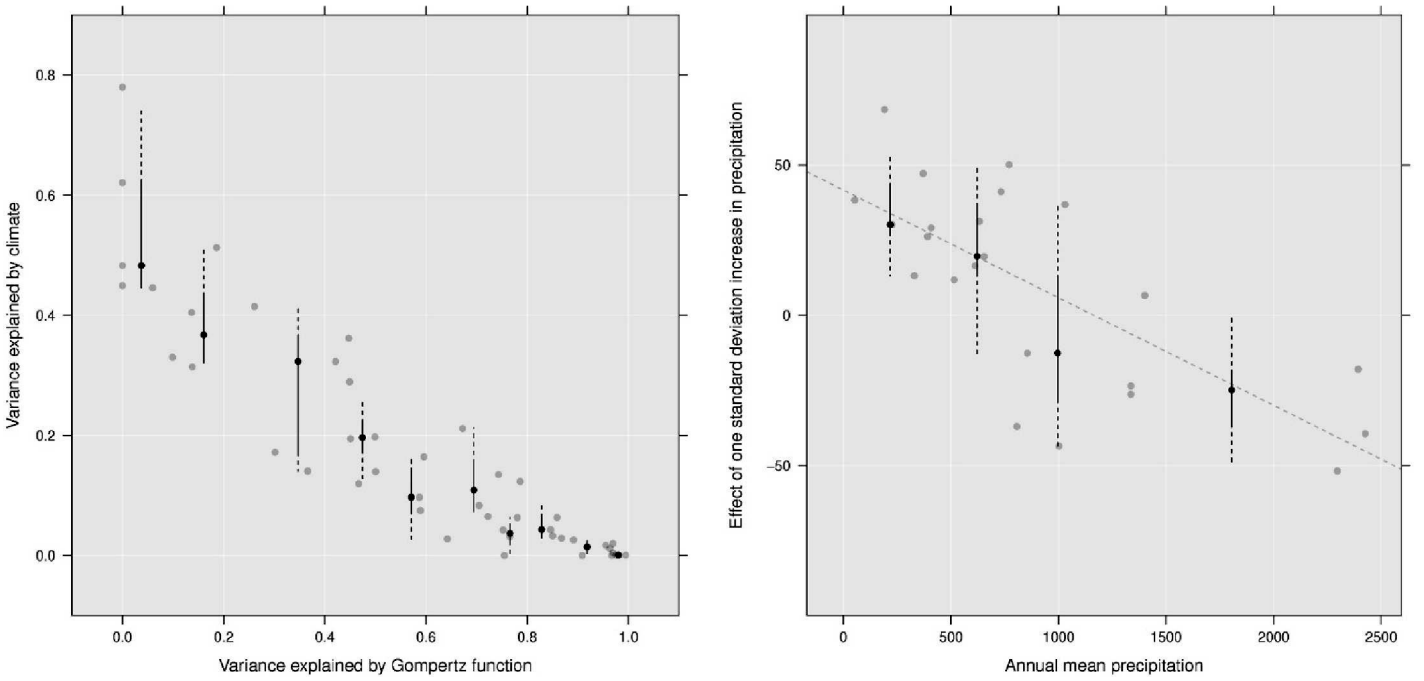

Figure 3 Climate sensitivity. Left: Relationship between the variance explained by the Gompertz function and climate. In countries where the number of cattle fit well with the Gompertz curve there is little climate sensitivity. Where the carrying capacity is close to its limit the cattle populations will fluctuate around the mean and hence the Gompertz function will not explain the variance. In these countries climate variability is the major source of variation. Vertical lines are box-and-whisker diagrams plotted at quantiles 0–0.1 to 0.9–1 by 0.1. Black dot indicate median for each quantile. Right: The effect of one standard deviation increase in precipitation vs annual mean rainfall (1961–1990). *Y*-axis is in %, and show the change in cattle numbers relative to the mean number of cattle (1961–1990). A value of 50 on the *y*-axis would indicate one standard deviation of rainfall (over five years) would increase the national cattle holdings by 50% relative to the mean. A value of −50 would indicate a reduction of 50%. Vertical lines are box-and-whisker diagrams plotted at quantiles 0–0.25 to 0.75–1 by 0.25. Black dot indicate median for each quantile.

(0.07, 0.22), Subhumid: −0.01 (−0.07, 0.05). No clear associations were found between temperature anomalies and national cattle holdings.

DISCUSSION

The cattle populations which are most sensitive to climate variability are located in arid regions. In the IPCC 2007 report it was stated that "… changes in range-fed livestock numbers in any African region will be directly proportional to changes in annual precipitation." This is only true in dry environments, and that in wetter environments, increased precipitation has no, or negative impacts on the national cattle stocks.

The observed negative association between precipitation and national cattle holdings in humid countries could be the effect of cross-border movement of livestock. We show that, while the Sahelian nations Niger and Chad show the expected positive relationship between rainfall and cattle numbers, their southern neighbours (Benin, Nigeria, Cameroon and CAR) have negative relationships. Figure 1 shows that the majority of cattle both these groups of countries are close to the borders between the former and latter groups. Since nomadic cattle herding or transhumance is the dominant form of cattle herding in this region, a likely explanation is that cattle are driven south during periods

of drought in the Sahel to areas with higher rainfall. In Mali, cattle are less frequently herded across the southern borders but instead are driven from the Sahel into higher rainfall regions further south within the country. This may explain in part why Mali has no clear relationship with rainfall, as the main effect of drought may be a shift in cattle further south within the country. Unfortunately the data presented here do not allow quantifying cross-border movement. It is also possible that the negative response to precipitation in humid areas can be explained by precipitation driven variability in the tsetse-fly populations, with more efficient transmission of trypanosomes in wet years. To establish any associations between variability in the number of tsetse flies and cattle, there is a need for historical datasets describing how the number of tsetse flies has varied in space and time.

Due to lack of consistent weather records for parts of Africa, and uncertainty in the estimated cattle densities, the absolute estimated effect of preceding precipitation and temperature anomalies on national cattle holdings should be interpreted with care. The consistent response in dry and humid regions does however make physical sense, with long wet or dry periods being the main factor controlling water and food availability, and in the end, cattle. We have not addressed if the declines and increases in cattle numbers are reflecting planning, by for example slaughtering and selling more cattle in dry periods, or if the declines in cattle numbers in dry periods can be attributed to natural mortality due to lack of food and water. These are important aspects which should be further investigated and documented, since the response to long term (30 years) changes in the climate might be different from the responses to short term fluctuations (2–3 years).

Since this is a statistical model with its limitations, it is therefore optimistic to extrapolate these relationships into the future. Instead we will show the expected changes in precipitation in the next century, and combine these maps with information about where cattle was present in the 1960s. We look at three of the Representative Concentration Pathways (RCP), where RCP 2.6 (Image, 14 models) showing future climate with strong mitigation, RCP 4.5 (miniCam, 18 models) is an in-between scenario and RCP 8.5 (Message, 16 models) assumes no mitigation. Figure 4 shows the expected response in precipitation under the four different scenarios. To minimize the effect of multidecadal variability we have used three averaging periods, baseline (1961–2000), near future (2006–2050), and distant future (2051–2100). Under the scenario where no mitigation takes place (Message), the Southern part of Africa will have a reduction in mean annual precipitation of 0.1–0.2 standard deviations in 2051–2100. Under the Image scenario the signal is weakened and the agreement between the models is lower. The miniCam scenario lies in between. The Southern part of Africa is one of the areas with a relatively high sensitivity to precipitation, and it is very likely that the cattle populations in this region will be negatively affected if no mitigation takes place. However, the increased CO_2 can reduce the impact of decreased rainfall, by increasing the soil nutrient availability. It is also likely that mean precipitation will decrease in parts of Mauritania, Mali and Senegal. The opposite is true for Eastern Africa and Eastern Sahel. These areas have shown little sensitivity to precipitation the last 50 years, and it is likely to very likely that the annual

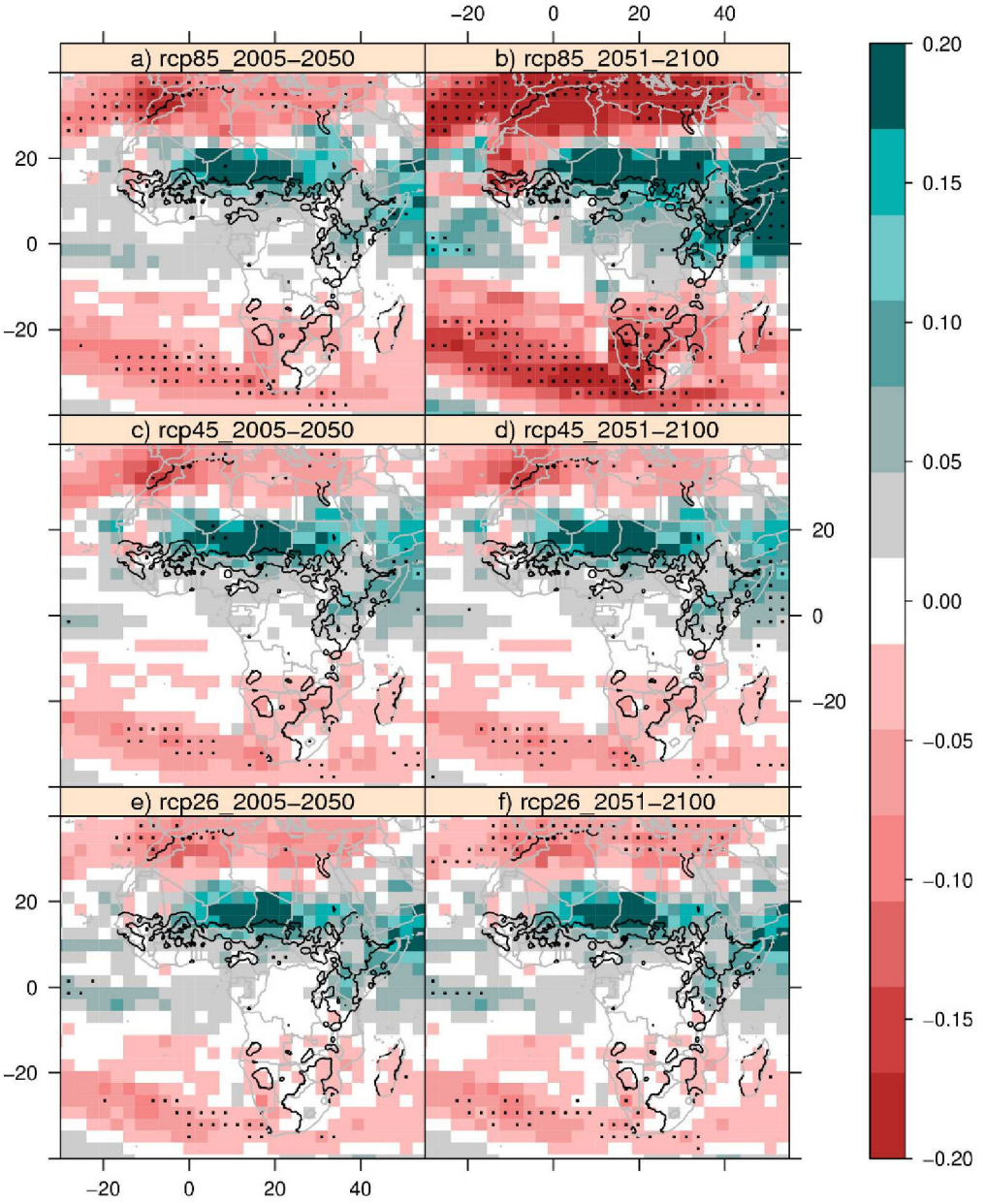

Figure 4 Future precipitation patterns. The mean expected changes from 1961–2000 in precipitation under three different climate scenarios. Shading indicate standard deviations, black contours show the presence of cattle in the 1960s. Only values where more than 66% of the models agree are shown. Black dots indicate more than 90% of the models agree on the sign of change.

precipitation will increase with no mitigation. The signal is weaker under the Image and miniCam scenarios.

It is interesting to note that the East African countries are less sensitive to climate variability. If new areas are utilized for cattle production over time, like the Ethiopian case (*Funk et al., 2012*), the resulting increased carrying capacity of a country will wash

out the effects of climate variability in the data. In the analysis presented here we suggest the national cattle holdings in Ethiopia have not been influenced by climate variability. However, the cattle populations in the Borana pastoral production system in Ethiopia have been strongly influenced by rainfall variability and trends (*Angassa & Oba, 2007*). It is therefore important to remember our analysis is restricted to nations.

The concept of carrying capacity is briefly addressed in the 2007 IPPC report, part III (*Intergovernmental Panel on Climate Change, 2007*). Carrying capacity is a key to understand how cattle will be influenced by changes in the climate. Adaptation on national scales can happen through utilization of new areas, but at some point there will be few new areas to use, and the vulnerability to climate will most likely increase.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was made possible by the grants from The Norwegian Programme for Development, Research and Education (NUFU) and the University of Bergen. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
The Norwegian Programme for Development, Research and Education (NUFU).
University of Bergen.

Competing Interests

The authors declare that they have no competing interests

Author Contributions

- Torleif Markussen Lunde conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper.
- Bernt Lindtjørn conceived and designed the experiments, wrote the paper.

Data Deposition

The following information was supplied regarding the deposition of related data:
ftp://ftp.uib.no/pub/gfi/tlu004/OMaWa/

REFERENCES

Abramson IS. 1982. On bandwidth variation in kernel estimates-a square root law. *The Annals of Statistics* **10(4)**:1217–1223 DOI 10.1214/aos/1176345986.

Angassa A, Oba G. 2007. Relating long-term rainfall variability to cattle population dynamics in communal rangelands and a government ranch in southern Ethiopia. *Agricultural Systems* **94(3)**:715–725 DOI 10.1016/j.agsy.2007.02.012.

Bagadi HO. 1978. The relationship between the annual rainfall and outbreaks of blackquarter of cattle in northern Nigeria. *Tropical Animal Health and Production* **10(2)**:124–126 DOI 10.1007/BF02235324.

Bates D, Maechler M, Bolker B. 2011. lme4: Linear mixed-effects models using S4 classes, R package version 0.999375-38.

Bivand RS, Pebesma EJ, Gomez-Rubio V. 2008. *Applied spatial data analysis with R*. NY: Springer.

Branagan D. 1973. The developmental periods of the ixodid tick rhipicephalus appendiculatus neum. under laboratory conditions. *Bulletin of Entomological Research* **63(01)**:155–168 DOI 10.1017/S0007485300050951.

Brück T, Schindler K. 2009. Smallholder land access in post-war northern Mozambique. *World Development* **37(8)**:1379–1389 DOI 10.1016/j.worlddev.2008.08.016.

Central Statistical Authority. 2004. Livestock aerial survey in the Somali region. Technical report, Central Statistical Authority, Ministry of Finance and Economic Development, Ethiopia, 02.

Choumara R. 1961. Notes sur le paludisme au Somaliland. *Rivista di Malariologia* **40**:9–34.

Cossins NJ, Upton M. 1988. The impact of climatic variation on the Borana pastoral system. *Agricultural Systems* **27(2)**:117–135 DOI 10.1016/0308-521X(88)90025-X.

Cragg CY. 1923. The zoophilism of *Anopheles* in relation to the epidemiology of malaria in India. *Indian Journal of Medical Research* **4**:962–964.

Deshler W. 1963. Cattle in Africa: distribution, types, and problems. *Geographical Review* **53(1)**:52–58 DOI 10.2307/212808.

Desta S, Coppock DL. 2002. Cattle population dynamics in the southern *Ethiopian* rangelands, 1980–97. *Journal of Range Management* **55(5)**:439–451 DOI 10.2307/4003221.

Diggle PJ, Fisher NI. 1985. Sphere: a contouring program for spherical data. *Computers & Geosciences* **6**:725–766 DOI 10.1016/0098-3004(85)90015-9.

Fafchamps M, Gavian S. 1997. The determinants of livestock prices in Niger. *Journal of African Economies* **6(2)**:255–295 DOI 10.1093/oxfordjournals.jae.a020928.

FAO 2011. Faostat live animals. http://faostat.fao.org/site/573/DesktopDefault.aspx?PageID=573, September.

Funk C, Rowland J, Eilerts G, Kebebe E, Biru N, White L, Galu G. 2012. A climate trend analysis of Ethiopia. *U.S. Geological Survey, Fact Sheet* 3053.

Oba G. 2001. The effect of multiple droughts on cattle in Obbu, northern Kenya. *Journal of Arid Environments* **49(2)**:375–386 DOI 10.1006/jare.2000.0785.

Hall MJ, Kheir SM, Rahman AH, Noga S. 1984. Tsetse and trypanosomiasis survey of southern Darfur province, Sudan. II. Entomological aspects. *Tropical Animal Health and Production* **16(3)**:127–40 DOI 10.1007/BF02252777.

Hansen J, Sato M, Ruedy R. 2012. Perception of climate change. *Proceedings of the National Academy of Sciences of the United States of America* **109(37)**:E2415–E2423 DOI 10.1073/pnas.1205276109.

Henry B, Charmley E, Eckard R, Gaughan JB, Herarty R. 2012. Livestock production in a changing climate: adaptation and mitigation research in Australia. *Crop & Pasture Science* **63**:191–202 DOI 10.1071/CP11169.

Hodges KI. 1996. Spherical nonparametric estimators applied to the UGAMP model integration for AMIP. *Monthly Weather Review* **124**:2914–2932 DOI 10.1175/1520-0493(1996)124<2914:SNEATT>2.0.CO;2.

Hoffmann I. 2010. Climate change and the characterization, breeding and conservation of animal genetic resources. *Animal Genetics* **41(Suppl 1)**:32–46 DOI 10.1111/j.1365-2052.2010.02043.x.

Intergovernmental Panel on Climate Change. 2007. Fourth assessment report: climate change 2007: working Group II: impacts, adaptation and vulnerability. Geneva: IPCC.

Jöchle W. 1972. Seasonal fluctuations of reproductive functions in Zebu cattle. *International Journal of Biometeorology* **16(2)**:131–144 DOI 10.1007/BF01810284.

Kaufman and Rousseeuw. 1990. *Finding groups in data*. New York: Wiley.

Lunde TM, Balkew M, Korecha D, Gebre-Michael T, Massebo F, Sorteberg A, Lindtjørn B. 2013a. A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. ii. validation of species distribution and seasonal variations. *Malaria Journal* **12(1)**: 78 DOI 10.1186/1475-2875-12-78.

Lunde TM, Korecha D, Loha E, Sorteberg A, Lindtjørn B. 2013b. A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis. *Malaria Journal* **12(1)**: 28 DOI 10.1186/1475-2875-12-28.

Mertz O, Mbow C, Reenberg A, Diouf A. 2009. Farmers' perceptions of climate change and agricultural adaptation strategies in rural Sahel. *Environmental Management* **43(5)**:804–16 DOI 10.1007/s00267-008-9197-0.

Middleton N, United Nations Environment Program, Thomas DSG. 1997. *World atlas of desertification*. Arnold: Hodder Arnold Publication.

Min SK, Zhang X, Zwiers FW, Hegerl GaC. 2011. Human contribution to more-intense precipitation extremes. *Nature* **470(7334)**:378–81 DOI 10.1038/nature09763.

Pebesma EJ, Bivand RS. 2005. Classes and methods for spatial data in R. *R News* **5(2)**:9–13.

Robinson T, Fao's Animal Production and Health Division. 2011. Observed livestock densities. *Available at http://www.fao.org/AG/againfo/resources/en/glw/GLW_dens.html*.

Rötter R, van de Geijn SC. 1999. Climate change effects on plant growth, crop yield and livestock. *Climatic Change* **43**:651–681 DOI 10.1023/A:1005541132734.

Scoones I. 1992. The economic value of livestock in the communal areas of southern Zimbabwe. *Agricultural Systems* **39**:339–359 DOI 10.1016/0308-521X(92)90074-X.

Seif SM, Johnson HD, Lippincott AC. 1979. The effects of heat exposure (31 °C) on Zebu and Scottish highland cattle. *International Journal of Biometeorology* **23(1)**:9–14 DOI 10.1007/BF01553372.

Seo S, Mendelsohn R. 2006. The impact of climate change on livestock management in Africa: a structural ricardian analysis. *CEEPA Discussion Paper* **23**:1–48.

Stokes CJ, Crimp S, Giffird R, Ash AJ, Howden SM. 2010. *Adapting agriculture to climate change: preparing australian agriculture, forestry and fisheries for the future*. Melbourne: CSIRO Publishing.

Taddesse G, Peden D, Abiye A, Wagnew A. 2003. Effect of manure on grazing lands in Ethiopia, east African highlands. *Mountain Research and Development* **23(2)**:156–160 DOI 10.1659/0276-4741(2003)023[0156:EOMOGL]2.0.CO;2.

Tawah CL, Rege JEO, Aboagye GS. 1997. A close look at a rare african breed — the Kuri cattle of lake Chad basin: origin, distribution, production and adaptive characteristics. *South African Journal of Animal Science* **27(2)**:31–40.

Terblanche JS, Clusella-Trullas S, Deere JA, Chown Steven L. 2008. Thermal tolerance in a south-east African population of the tsetse fly *Glossina* pallidipes (diptera, glossinidae):

implications for forecasting climate change impacts. *Journal of Insect Physiology* **54(1)**:114–127 DOI 10.1016/j.jinsphys.2007.08.007.

The Open Source Geospatial Foundation. Geospatial data abstraction library (gdal). GDAL version 1.8.1 (accessed Oct 26 2011).

Tirados I, Gibson G, Young S, Torr SJ. 2011. Are herders protected by their herds? An experimental analysis of zooprophylaxis against the malaria vector *Anopheles* arabiensis. *Malaria Journal* **10**:68 DOI 10.1186/1475-2875-10-68.

Wint W, Robinson T. 2007. Gridded livestock of the world - 2007. Technical report, Food And Agriculture Organization Of The United Nations, Rome.

Young AS, Leitch BL. 1981. Epidemiology of east coast fever: some effects of temperature on the development of *Theileria* parva in the tick vector, Rhipicephalus appendiculatus. *Parasitology* **83(pt 1)**:199–211 DOI 10.1017/S0031182000050162.

Paper IV

6.4 How malaria models relate temperature to malaria transmission

Torleif Markussen Lunde, Mohamed Nabie Bayoh and Bernt Lindtjørn

Parasites & Vectors 2013, **6**:20

Parasites & Vectors

RESEARCH Open Access

How malaria models relate temperature to malaria transmission

Torleif Markussen Lunde[1,2,4*], Mohamed Nabie Bayoh[3] and Bernt Lindtjørn[2]

Abstract

Background: It is well known that temperature has a major influence on the transmission of malaria parasites to their hosts. However, mathematical models do not always agree about the way in which temperature affects malaria transmission.

Methods: In this study, we compared six temperature dependent mortality models for the malaria vector *Anopheles gambiae* sensu stricto. The evaluation is based on a comparison between the models, and observations from semi-field and laboratory settings.

Results: Our results show how different mortality calculations can influence the predicted dynamics of malaria transmission.

Conclusions: With global warming a reality, the projected changes in malaria transmission will depend on which mortality model is used to make such predictions.

Keywords: *Anopheles gambiae* sensu stricto, Climate, Temperature, Mathematical model

Background

Since the 1950s, near-surface global temperatures have increased by about 0.5-0.6°*C* [1], and it is likely that temperatures will continue to increase over the next century [2]. Model predictions, reported widely in climate policy debates, project that a warmer climate could increase malaria caused by the parasites *Plasmodium falciparum* and *P. vivax* in parts of Africa [3]. Malaria is transmitted by mosquitoes of the *Anopheles* genus, with *Anopheles gambiae s.s.*, *An. arabiensis* and *An. funestus* being the dominant vector species in Africa [4,5].

These projections rely on knowledge about how the malaria parasite and anopheline vectors respond to changes in temperature. While a lot is known [6] about how parasite development is influenced by temperature [7], the same cannot be said for mosquitoes. In addition to temperature, humidity [8,9], breeding site formation [10], and competition between mosquitoes [11,12] are important factors controlling the number of vectors at any time.

Climate predictions about humidity and precipitation are more uncertain than temperature projections. Therefore, it is of interest to see if a consensus exists between different malaria models about how temperature alone influences malaria transmission. In the past, studies have suggested that the optimal temperature for malaria transmission is between 30 and 33°*C* [13-15].

Here, we compare six mortality models (Martens 1, Martens 2, Bayoh-Ermert, Bayoh-Parham, Bayoh-Mordecai and Bayoh-Lunde) to reference data (control) for *Anopheles gambiae s.s.*, and show how these models can alter the expected consequences of higher temperatures. The main purpose of the study is to show if there are any discrepancies between the models, with consequences for the ability of projecting the impact of temperature changes on malaria transmission.

We have focused on models that have been designed to be used on a whole continent scale, rather than those that focus on local malaria transmission [10,16,17].

Methods

Survival models

Six different parametrization schemes have been developed to describe the mortality rates for adult *An. gambiae s.s.*. These schemes are important for estimating the

*Correspondence: torleif.lunde@cih.uib.no
[1]Bjerknes Centre for Climate Research, University of Bergen, Norway
[2]Centre for International Health, University of Bergen, Norway
Full list of author information is available at the end of the article

temperature at which malaria transmission is most efficient. The models can also be used as tools to describe the dynamics of malaria transmission. In all of the equations presented in this paper, temperature, T and T_{air} are in °C.

Martens 1

The first model, which is called Martens scheme 1 in Ermert *et al.* [18], and described by Martens *et al.* [19-21], is derived from three points, and shows the relationship between daily survival probability (p) and temperature (T). This is a second order polynomial, and is, mathematically, the simplest of the models.

$$p(T) = -0.0016 \cdot T^2 + 0.054 \cdot T + 0.45 \quad (1)$$

Martens 2

In 1997 Martens [21] described a new temperature-dependent function of daily survival probability. This model has been used in several studies [13,14,22,23]. In the subsequent text this model is named Martens 2. Numerically, this is a more complex model than Martens 1, and it increases the daily survival probability at higher temperatures.

$$p(T) = e^{-\frac{1}{-4.4+1.31 \cdot T - .03 \cdot T^2}} \quad (2)$$

Bayoh-Ermert

In 2001, Bayoh carried out an experiment where the survival of *An. gambiae s.s.* under different temperatures (5 to 40 in 5°C steps) and relative humidities (RHs) (40 to 100 in 20% steps) was investigated [24]. This study formed the basis for three new parametrization schemes. In the naming of these models, we have included Bayoh, who conducted the laboratory study, followed by the author who derived the survival curves.

In 2011, Ermert *et al.* [18] formulated an expression for *Anopheles* survival probability; however, RH was not included in this model. In the text hereafter, we name this model Bayoh-Ermert. This model is a fifth order polynomial.

Overall, this model has higher survival probabilities at all of the set temperatures compared with the models created by Martens.

$$\begin{aligned} p(T) = &- 2.123 \cdot 10^{-7} \cdot T^5 + 1.951 \cdot 10^{-5} \cdot T^4 \\ &- 6.394 \cdot 10^{-4} \cdot T^3 + 8.217 \cdot^{-3} \cdot T^2 \\ &- 1.865 \cdot 10^{-2} \cdot T + 7.238 \cdot 10^{-1} \end{aligned} \quad (3)$$

Bayoh-Parham

In 2012, Parham *et al.* [25] (designated Bayoh-Parham in subsequent text) included the effects of relative humidity and parametrized survival probability using the expression shown below. This model shares many of the same characteristics as the Bayoh-Ermert model. The mathematical formulation is similar to the Martens 2 model, but constants are replaced by three terms related to RH ($\beta_0, \beta_1, \beta_2$).

$$p(T, RH) = e^{-\left(T^2 \cdot \beta_2 + T \cdot \beta_1 + \beta_0\right)^{-1}} \quad (4)$$

where $\beta_0 = 0.00113 \cdot RH^2 - 0.158 \cdot RH - 6.61$, $\beta_1 = -2.32 \cdot 10^{-4} \cdot RH^2 + 0.0515 \cdot RH + 1.06$, and $\beta_2 = 4 \cdot 10^{-6} \cdot RH^2 - 1.09 \cdot 10^{-3} \cdot RH - 0.0255$.

For all models reporting survival probability, we can rewrite p to mortality rates, β according to:

$$\beta = -\ln(p) \quad (5)$$

Bayoh-Mordecai

Recently, Mordecai *et al.* [26] re-calibrated the Martens 1 model by fitting an exponential survival function to a subset of the data from Bayoh and Lindsay [24]. They used the survival data from the first day of the experiment and one day before the fraction alive was 0.01. Six data points were used for each temperature.

$$p(T) = -0.000828 \cdot T^2 + 0.0367 \cdot T + 0.522 \quad (6)$$

Bayoh-Lunde

From the same data [24], Lunde *et al.* [27], derived an age-dependent mortality model that is dependent on temperature, RH, and mosquito size. This model assumes non-exponential mortality as observed in laboratory settings [24], semi-field conditions [28], and in the field [29]. In the subsequent text we call this model Bayoh-Lunde. The four other models use the daily survival probability as the measure, and assume that the daily survival probability is independent of mosquito age. The present model calculates a survival curve (ϖ) with respect to mosquito age. Like the Bayoh-Parham model, we have also varied the mosquito mortality rates according to temperature and RH.

Because mosquito size is also known to influence mortality [8,9,30,31], we applied a simple linear correction term to account for this. In this model, the effect of size is minor compared with temperature and relative humidity. The survival curve, ϖ, is dependent on a shape and scale parameter in a similar manner as for the probability density functions. The scale of the survival function is dependent on temperature, RH, and mosquito size, while the scale parameter is fixed in this paper.

The mortality rate, $\beta_n(T, RH, size)$ (equation 7) is fully described in Additional file 1, with illustrations in Additional files 2, and 3.

$$\beta_n(T, RH, size) = \frac{\ln\left(\frac{\varpi_{N,mt_2}}{\varpi_{N,mt_1}}\right)}{\Delta t} \quad (7)$$

Biting rate and extrinsic incubation period

The equations used for the biting rate, $G(T)$, and the inverse of the extrinsic incubation period (EIP, pf) are described in Lunde et al. [27]. For convenience,

these equations and their explanations are provided in Additional file 1. The extrinsic incubation period was derived using data from MacDonald [7], while the biting rate is a mixture of the degree day model by Hoshen and Morse [32], and a model by Lunde *et al.* [27]. Since our main interest in this research was to examine how mosquito mortality is related to temperature in models, we used the same equation for the gonotrophic cycle for all of the mortality models. If we had used different temperature-dependent gonotrophic cycle estimates for the five models, we would not have been able to investigate the effect of the mortality curves alone.

Malaria transmission

We set up a system of ordinary differential equations (ODEs) to investigate how malaria parasites are transmitted to mosquitoes. Four of the mortality models (equations 1, 2, 3, and 4) are used in a simple compartment model that includes susceptible (S), infected (E) and infectious mosquitoes (I) (equation 8):

$$\begin{aligned}\frac{dS}{dt} &= -(\beta + G(T)\cdot H_i)\cdot S \\ \frac{dE}{dt} &= (G(T)\cdot H_i)\cdot S - (\beta + pf)\cdot E \\ \frac{dI}{dt} &= pf\cdot E - \beta\cdot I\end{aligned} \tag{8}$$

where H_i is the fraction of infectious humans, which was set to 0.01. $G(T)$ is the biting rate, and pf is the rate at which sporozoites develop in the mosquitoes. The model is initialized with $S = 1000$, $E = I = 0$ and integrated for 150 days with a time step of 0.5. As the equations show, there are no births in the population, and the fraction of infectious humans is held constant during the course of the integration. This set-up ensures that any confounding factors are minimized, and that the results can be attributed to the mortality model alone.

Because the Lunde *et al.* [27] (Bayoh-Lunde) mortality model also includes an age dimension, the differential equations must be written taking this into account. Note that the model also can be used in equation 8 if we allow β to vary with time.

We separate susceptible (S), infected (E) and infectious (I), and the subscript denotes the age group. In total there are 25 differential equations, but where the equations are similar, the subscript n has been used to indicate the age group.

Formulating the equation this way means we can estimate mosquito mortality for a specific age group. We have assumed that mosquito biting behaviour is independent of mosquito age; this formulation is, therefore, comparable to the framework used for the exponential mortality models.

The number of infectious mosquitoes is the sum of I_n, where $n = 2, \ldots, 9$.

$$\begin{aligned}\frac{dS_1}{dt} &= -(\beta_1 + a_1)\cdot S_1 \\ \frac{dS_n}{dt} &= a_{n-1}\cdot S_1 - (\beta_n + a_n + G(T)\cdot H_i)\cdot S_n \\ & \quad n = 2, 3, ..., 9 \\ \frac{dS_n}{dt} &= G(T)\cdot H_i\cdot S_2 - (\beta_2 + a_2 + pf)\cdot E_2 \\ \frac{dE_n}{dt} &= G(T)\cdot H_i\cdot S_n + a_{n-1}\cdot E_{n-1} \\ & \quad - (\beta_n + a_n + pf)\cdot E_n \\ & \quad n = 3, 4, ..., 9 \\ \frac{dI_2}{dt} &= pf\cdot E_2 - (\beta_2 + a_2)\cdot I_2 \\ \frac{dI_n}{dt} &= pf\cdot E_n + a_2\cdot I_{n-1} - (\beta_n + a_n)\cdot I_n \\ & \quad n = 3, 4, ..., 9\end{aligned} \tag{9}$$

Age groups for mosquitoes (m) in this model are $m_1 = [0, 1]$, $m_2 = (2, 4]$, $m_3 = (5, 8]$, $m_4 = (9, 13]$, $m_5 = (14, 19]$, $m_6 = (20, 26]$, $m_7 = (27, 34]$, $m_8 = (35, 43]$, $m_9 = (44, \infty]$ days, and coefficients a_n, where $n = 1, 2, \ldots, 9$, are 1.000, 0.500, 0.333, 0.250, 0.200, 0.167, 0.143, 0.125, 0.067. The rationale behind these age groups is that as mosquitoes become older, there is a greater tendency of exponential mortality compared to younger mosquitoes.

This model has initial conditions $S_1 = 1000$, and all other 0.

A note on the use of ODEs and rate calculations can be found in Additional file 4.

Validation data

To validate the models, we used the most extensive data set available on mosquito survival [24] under different temperatures (5 to 40 by 5°C) and RHs (40 to 100 by 20%) [24]; it is the same data that the Bayoh-Ermert, Bayoh-Parham and Bayoh-Lunde models were derived from. These data describe the fraction of live mosquitoes (f_a) at time t, which allows us to validate the models over a range of temperatures. Because three of the models used the Bayoh and Lindsay data to develop the survival curves, this comparison is unrealistic for Martens models.

Hence, to account for this we have used three independent data sets to validate the fraction of infectious mosquitoes and the mosquito survival curves.

Scholte *et al.* (Figure two in [33]) published a similar data set, but this was based on a temperature of 27 ± 1°C

and a RH of 80 ± 5%, whereas Afrane *et al.* (Figure two in [28]) used mean temperatures of 21.5 to 25.0 and RHs of 40-80%. Use of these data sets will allow us to complement the validation to determine if the patterns of malaria transmission are consistent with that of the control (Table 1). In addition to the data from Scholte *et al.* [33], we also found the following data set, which is suitable for validation of the survival curves but not the transmission process itself, because the data does not show the survival curve until all of the mosquitoes are dead [Kikankie, Master's thesis (Figures three to eight, chapter 3, 25°*C*, 80% RH) [34]]. These results are also shown in Table 1. The additional validation only gives information about the model quality between 21, and 27°*C*; however, it serves as an independent model evaluation to determine if the results are consistent and independent of the data set used to validate the models.

Using the data from Bayoh and Lindsay, Afrane *et al.* or Scholte *et al.* [33], we can calculate the fraction of mosquitoes that would become infectious at time t, using equation 8. We replace β with the time-dependent $\beta(t)$, which is a time varying mortality rate. This approach was used for the data from [24] and [33].

$$\beta(t) = -ln\left(\frac{f_a\left(t+\frac{1}{2}\right)}{f_a\left(t-\frac{1}{2}\right)}\right) \tag{10}$$

$\beta(t)$ is linearly interpolated at times with no data. The reference data from Bayoh and Lindsay [24] are hereafter designated as the control data in the subsequent text, whereas data from Scholte *et al.* [33] is called Scholte in Table 1. Table 1 also shows the skill scores of the mortality model alone (for the figures in Additional file 3).

Because some of the schemes do not include RH, we have displayed the mean number of infectious mosquitoes, I, for schemes that do include it. For the validation statistics, RH has been included. However, for schemes where the RH has not been taken into account, single realization at all humidities has been employed.

Validation statistics

Skill scores (S) are calculated following Taylor [35]:

$$S_s = \frac{4\cdot(1+r)^4}{(\hat{\sigma}_f + 1/\hat{\sigma}_f)^2\cdot(1+r_0)^4} \tag{11}$$

where r is the Pearson correlation coefficient, $r_0 = 1$ is the reference correlation coefficient, and $\hat{\sigma}_f$ is the variance of the control over the standard deviation of the model (σ_f/σ_r). This skill score will increase as a correlation increases, as well as increasing as the variance of the model approaches the variance of the model.

The Taylor diagram used to visualize the skill score takes into account the correlation (curved axis), ability to represent the variance (x and y axis), and the root mean square.

Another important aspect is determining at which temperatures transmission is most efficient. If mosquitoes have a peak of infectiousness at, for example, 20°*C* in one model, temperatures above this will lead to a smaller fraction of mosquitoes becoming infectious. A different model might set this peak at 27°*C*, so that at temperatures from 20-27°*C*, the fraction of infectious mosquitoes will increase, followed by a decrease at higher temperatures. Isolating the point at which the mosquitoes are the most efficient vectors for malaria parasites is important for assessing the potential impacts of climate change. To show the differences between the models, we report the temperature where the maximum efficiency for producing infectious mosquitoes was observed. This can be done by maximizing equation 12.

$$\arg\max_{T\in[10,40]}\int_{t=0}^{\infty} Idt \tag{12}$$

For the transmission process we also report Akaike information criterion (AIC) [36] from a generalized linear model with normal distribution. Since the observations are not independent, and residuals do not follow a normal distribution, we sample 100 values from the simulations 1000 times. We set the probability of sampling $y_{i,j}$ equal

Table 1 Skill scores

	Control	AIC Control	Scholte	AF	BL mortality model	SK mortality model
Martens 1	0.01	76 (56, 96)	0.00	0.03	0.36	0.25
Martens 2	0.38	9 (-14, 30)	0.55	0.37	0.54	0.45
Martens 3	0.65	-38 (-75, -9)	0.53	0.77	0.65	0.52
Bayoh-Ermert	0.27	30 (1, 58)	0.16	0.43	0.79	0.56
Bayoh-Parham	0.16	26 (-11, 55)	0.05	0.31	0.79	0.59
Bayoh-Lunde	0.90	-111 (-148, -81)	0.83	0.94	0.90	0.81
Bayoh-Mordecai	0.62	-53 (-82, -29)	0.58	0.70	0.57	0.49

Skill scores as defined in equation 11. "Control" represents the validation of infectious mosquitoes using the data from Bayoh and Lindsay [24], "Scholte" [33] represents the validation of infectious mosquitoes using the data from Scholte, "AF" represents the validation of infectious mosquitoes using the data from Afrane, "BL mortality model" represents the validation of the mortality model using the data from Bayoh and Lindsay [24], and "SK mortality model" represents the validation of the mortality model using data from Scholte [33] and Kikankie [34]. "AIC control" is Akaike information Criterion for "Control" (95 % confidence interval).

to normalized (sum = 1) fraction of infected mosquitoes of the control. This method allow us to generate a model with normally distributed, non-correlated errors. Median AIC, with 95% confidence intervals are reported in Table 1.

Results

Figure 1 shows the percentage of infectious mosquitoes plotted against time (days) (x) and temperature (y). The control shows that the most efficient transmission occurs at 25°*C*, while the maximum percentage of infectious mosquitoes at any time is 1.1. We found that the Martens 1 and 2 models both underestimate the fraction of infectious mosquitoes, while the Bayoh-Ermert and Bayoh-Lunde models had comparable values. While the Bayoh-Parham model affords similar values at 40% RH, it overestimates the fraction of infectious mosquitoes at higher RHs (Additional file 3). There are also substantial differences at which the temperatures for transmission are most efficient.

While Martens 1 has the most efficient transmission at 20.4°C, Martens 2 and Bayoh-Ermert show the transmission efficiency peaking at 26.8 and 27.5°*C*. Both the control and Bayoh-Lunde models peak at 25°*C*, as measured according to equation 12, Bayoh-Parham peaks at 26.3°*C*, and Bayoh-Mordecai peaks at 24.4°*C* (Figure 2).

The numerical solution of the Bayoh-Ermert mortality model also reveals that it has problems related to enhanced mosquito longevity at all of the selected temperatures; this effect was especially pronounced around 20°*C*. We also found that the Bayoh-Parham model has issues with prolonged mosquito survival.

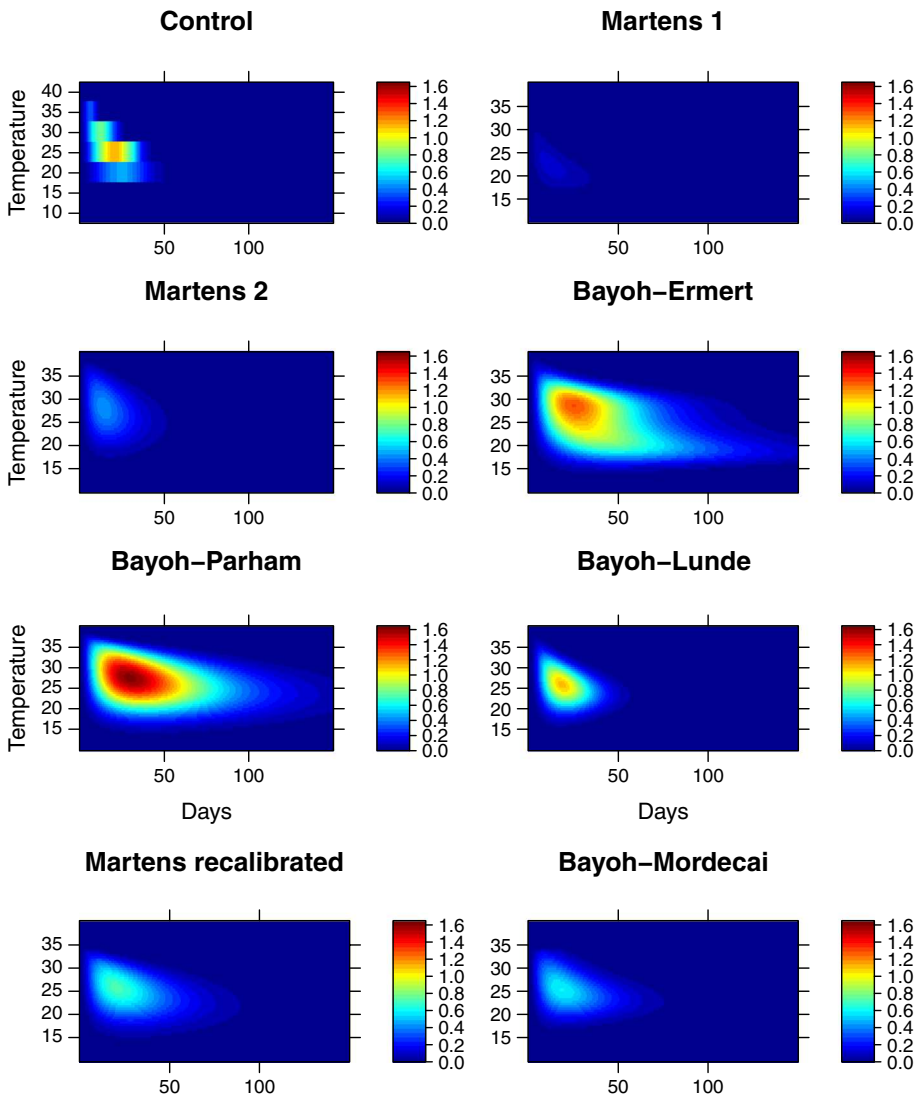

Figure 1 The percentage of infectious mosquitoes over time and temperature.

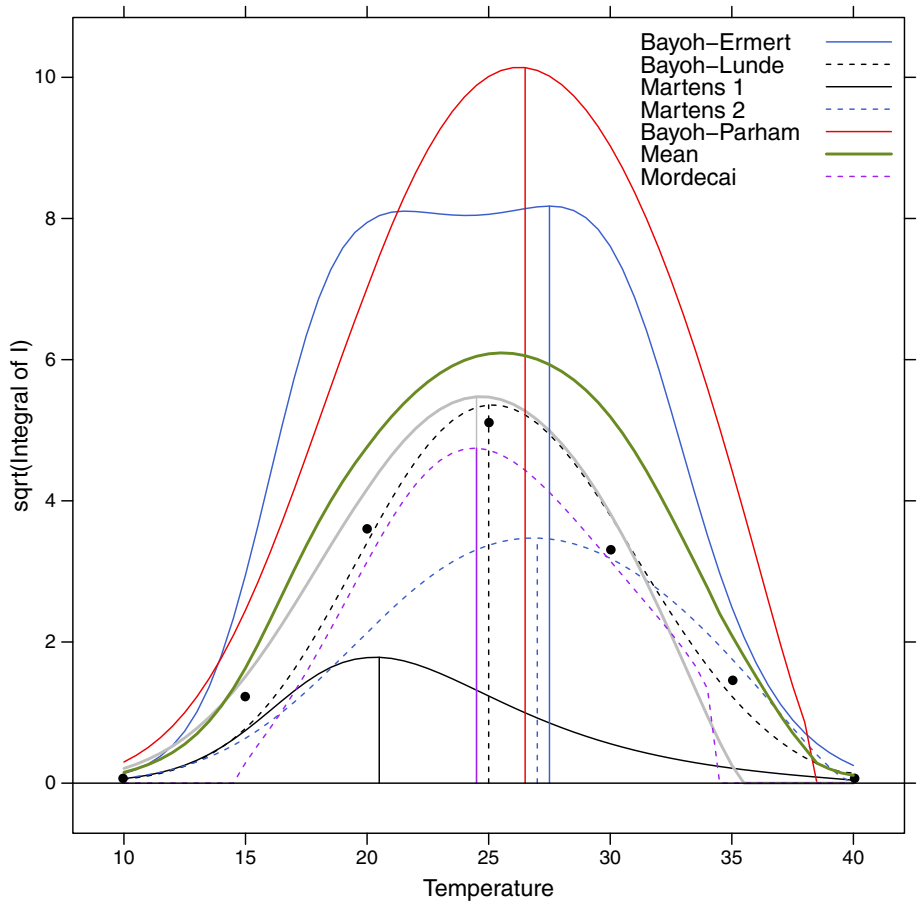

Figure 2 Integral of infectious mosquitoes over temperature. Models: Bayoh-Ermert (blue solid line), Martens 1 (black solid line), Martens 2 (blue dashed line), Martens 3 (grey solid line), Bayoh-Parham (red solid line), Lunde (black dashed line), and the mean value of the five models (green thick solid line). Black dots indicate the results for the control, and vertical lines show the temperature at which the maximum can be found (equation 12).

To evaluate the skill of the models, with emphasis on spatial patterns and variance, we investigated the skill score that was defined in equation 11. The standard deviation, root mean square and correlation coefficient are summarized in a Taylor diagram (Figure 3). Skill scores closer to 1 are a sign of better performance from a model (Table 1).

When validating the transmission process using the data from Bayoh and Lindsay (Table 1, column 1), the majority of the penalty for the Martens 1 and 2 models was due to the low variance, indicating that the mortality is set too high compared with the reference. Further analysis found that the Bayoh-Ermert model correlated poorly with the reference, and the variance, $\hat{\sigma}_f$, was too high. The Bayoh-Parham model also suffered from low correlation, as well as too high variance. Overall, the Bayoh-Lunde model has the highest skill score, followed by the Bayoh-Mordecai model. The patterns are consistently independent of the data used to validate the models with respect to the malaria transmission process. Validation of the survival curves alone, and their relationship with the transmission process, is discussed in the next section.

The relatively simple Martens 2 model ranked third among the models. We re-calibrated [37,38] the model using the data from Bayoh and Lindsay. The re-calibrated model (equation 13) generated a skill score of 0.65 (for the transmission process). In addition, Martens 2 was most efficient at 24.5°*C*. The Martens 3 model can be used for temperatures between 5 and 35°*C*.

$$p(T) = e^{-\frac{1}{-4.31564+2.19646\cdot T-0.058276\cdot T^2}} \tag{13}$$

The newly calibrated Martens 2 model (hereafter called Martens 3), can be seen in Figure 2; the skill scores are reported in Table 1.

To investigate how sensitive the results of the Mordecai *et al.* [26] analysis are to the choice of mortality model, we calculated the optimal temperature for malaria transmission using their full temperature-sensitive malaria R_0 model (equation 2 in [26]). The mortality rate, $\mu(T)$, was

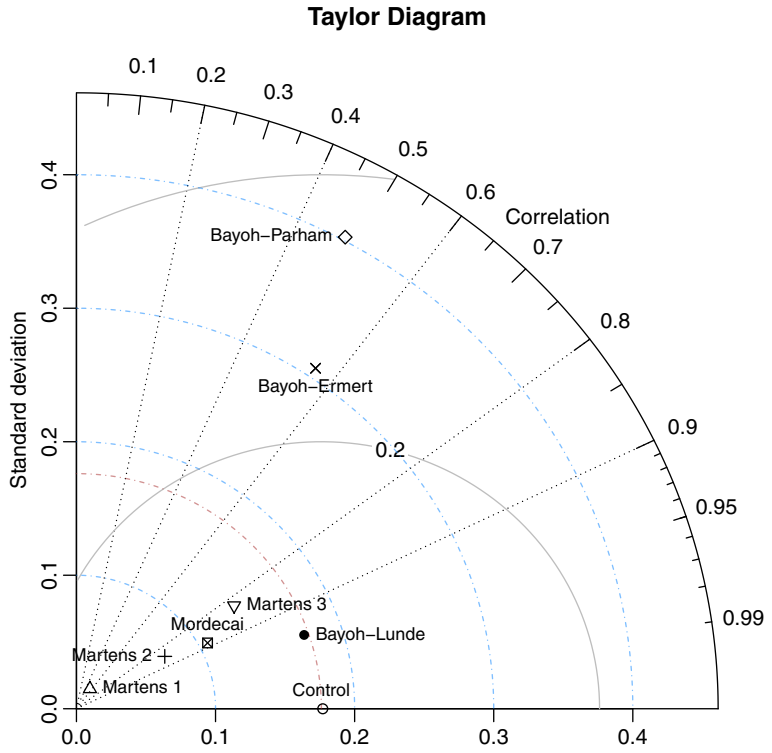

Figure 3 Taylor diagram. The model names are indicated next to the symbols. The x and y axes represent the standard deviations, the curved grey lines are the root mean square, while the dashed lines represent the Pearson correlation coefficient.

replaced with $-ln(p(T))$ from the exponential models. Population density (N), and recovery rate, r, were set to 1, since these do not influence the optimal temperature for malaria transmission. The results can be seen in Table 2. Relative differences between the two methods is in the range from 1–11% (Table 2). Figure 4 shows R_0 according to temperature (with $N = 1, r = 1$) for the exponential models. The maximum R_0 ranges from 10 (Martens 1) to 206 (Bayoh-Parham).

Table 2 Optimal temperature for malaria transmission

	This paper	R_0 from Mordecai *et al.*	Relative difference %
Martens 1	20.4	23.0	11.98
Martens 2	26.8	27.0	0.74
Martens 3	24.7	26.0	5.13
Bayoh-Ermert	27.5	27.2	1.10
Bayoh-Parham	26.3	26.9	2.26
Bayoh-Lunde	25.2		
Bayoh-Mordecai	24.4	25.6	4.80

Optimal temperature calculated using the methods in this paper, and by using methodology in Mordecai *et al.* [26].

Discussion and conclusions

The relationship between sporozoite development and the survival of infectious mosquitoes at different temperatures is poorly understood; therefore, any model projections relating the two should be interpreted with care. The Martens 2 and Bayoh-Ermert models suggest that areas of the world where temperatures approach 27°C could experience more malaria. Martens 3, Bayoh-Mordecai, and our model (Bayoh-Lunde) suggests that transmission is most efficient at around 25°C. The Martens 1 model peaks at 20.4°C, and Bayoh-Parham at 26.3°C (Figure 1). None of the models, except Bayoh-Lunde, capture all of the characteristics of the reference data, however.

Table 1 also shows the skill score for the mortality model alone. Both the Bayoh-Parham and the Bayoh-Ermert models have good representations of the survival curves. However, the nature of the exponential mortality curves gives them the choice of rapid mortality giving a reasonable, but underestimated, transmission process (Martens 2), or a good fit to the survival curves, which in turn makes the mosquitoes live too long, resulting in a poor transmission process (Bayoh-Parham and Bayoh-Ermert). Because the Bayoh-Lunde model offers a fair description of the survival curves as well as an age structure in the differential equations, we consider that the transmission process is well described. The Martens

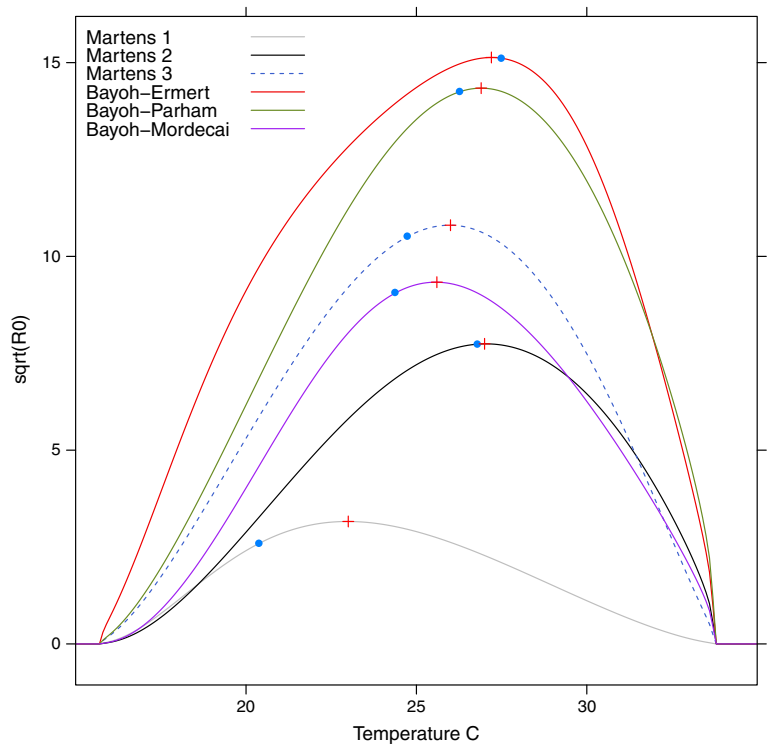

Figure 4 R_0 as a function of temperature calculated using equation 2 in Mordecai *et al.* [26], but with different mortality models. Blue dots represent optimal temperatures using the methods in this paper, and red crosses is the optimal temperature using the methods from Mordecai *et al.* [26].

1 and 2, Bayoh-Ermert, Bayoh-Mordecai and Bayoh-Parham models all assume constant mortality rates with age, and would, therefore, not benefit from being solved in an age-structured framework.

The Martens 1 model has been used in several studies [19-21], with the latest appearance by Gething *et al.* in this journal [39]. Considering the poor skill of the Martens 1 model, the validity, or etiology, of results presented in these papers should be carefully considered.

It is likely that regions with temperatures below 18°*C*, as is typical for the highland areas of East and Southern Africa, which are too cold for malaria transmission, might experience more malaria if their temperatures increase. However, malaria transmission in the future will be dependent on many other factors such as poverty, housing, access to medical care, host immunity and malaria control measures.

Most countries in Sub-Saharan Africa have annual mean temperatures between 20 and 28°*C*. In these areas, linking past and future temperature fluctuations to changes in malaria transmission is challenging. Our data suggest that one way to reduce this uncertainty is to use age-structured mosquito models. These models produce results that agree with the observed data, and non-exponential mosquito mortality has been demonstrated in several studies [33,40-42], although the true nature of mosquito survival in the field is not fully elucidated. The newly calibrated Martens 2 model described here also produces acceptable results. If simplicity is a goal in itself [43], models that assume exponential mortality will still have utility. To believe in projections of the potential impact of long-term, large-scale climate changes, it is crucial that models have an accurate representation of malaria transmission, even at the cost of complexity. For studies of malaria transmission at village level, other approaches might be more suitable [10,16,44,45].

Additional files

Additional file 1: Details of the Bayoh-Lunde model, mosquito biting rate, and parasite extrinsic incubation period.

Additional file 2: This file shows how ζ can be used to change the shape of the Bayoh-Lunde survival curve. The black line is the reference data, while the red line represents the Bayoh-Lunde survival curve. Temperature, relative humidity (as a fraction from 0 to 1), and ζ are given in the panel strips.

Additional file 3: Survival curves for all of the models investigated in this study plotted at different temperatures and relative humidities. The figure on page two shows the legend as well as an example of non-exponential mortality.

Additional file 4: A note on the use of ordinary differential equations, age structure (with an example), and rate calculations.

Abbreviations
BL: Bayoh and Lindsay; EIP: Extrinsic incubation period; ODEs: Ordinary differential equations; SK: Scholte and Kikankie.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
The work presented here was carried out in collaboration between all of the authors. BL, MNB and TML defined the research theme. MNB provided the data for the control. TML designed the methods and experiments, did the model runs, analysed the data, interpreted the results, and wrote the paper. All authors read and approved the final version of the manuscript.

Acknowledgements
This work was made possible by grants from The Norwegian Programme for Development, Research and Education (NUFU) and the University of Bergen. Our thanks go to Asgeir Sorteberg for commenting on the manuscript, and three anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

Author details
[1]Bjerknes Centre for Climate Research, University of Bergen, Norway. [2]Centre for International Health, University of Bergen, Norway. [3]KEMRI/CDC Research and Public Health Collaboration, Kisumu, Kenya. [4]Bjerknes Centre for Climate Research, Uni Research, Norway.

Received: 5 October 2012 Accepted: 15 January 2013
Published: 18 January 2013

References

1. Hansen J, Sato M, Ruedy R: **Perception of climate change.** *Proc Natl Acad Sci USA* 2012, **109:**E2415—2423.
2. Intergovernmental Panel on Climate Change: *Fourth Assessment Report: Climate Change 2007: Working Group I Report: The Physical Science Basis*. Geneva: IPCC; 2007.
3. Gething PW, Smith DL, Patil AP, Tatem AJ, Snow RW, Hay SI: **Climate change and the global malaria recession.** *Nature* 2010, **465:**342–345.
4. Sinka ME, Bangs MJ, Manguin S, Rubio-Palis Y, Chareonviriyaphap T, Coetzee M, Mbogo CM, Hemingway J, Patil AP, Temperley WH, Gething PW, Kabaria CW, Burkot TR, Harbach RE, Hay SI: **A global map of dominant malaria vectors.** *Parasit Vectors* 2012, **5:**69.
5. Sinka ME, Bangs MJ, Manguin S, Coetzee M, Mbogo CM, Hemingway J, Patil AP, Temperley WH, Gething PW, Kabaria CW, Okara RM, Van Boeckel T, Godfray HCJ, Harbach RE, Hay SI: **The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic precis.** *Parasit Vectors* 2010, **3:**117.
6. Paaijmans KP, Blanford S, Chan BHK, Thomas MB: **Warmer temperatures reduce the vectorial capacity of malaria mosquitoes.** *Biol Lett* 2012, **8:**465–468.
7. MacDonald G: *Dynamics of Tropical Disease*. London: Oxford University Press; 1973.
8. Gray EM, Bradley TJ: **Physiology of desiccation resistance in *Anopheles gambiae* and *Anopheles arabiensis*.** *Am J Trop Med Hyg* 2005, **73:**553–559.
9. Fouet C, Gray E, Besansky NJ, Costantini C: **Adaptation to aridity in the malaria mosquito Anopheles gambiae: chromosomal inversion polymorphism and body size influence resistance to desiccation.** *PLoS One* 2012, **7:**e34841.
10. Bomblies A, Duchemin JB, Eltahir EAB: **Hydrology of malaria: Model development and application to a Sahelian village.** *Water Resour Res* 2008, **44:**W12445.
11. Paaijmans KP, Huijben S, Githeko AK, Takken W: **Competitive interactions between larvae of the malaria mosquitoes *Anopheles arabiensis* and *Anopheles gambiae* under semi-field conditions in western Kenya.** *Acta Trop* 2009, **109:**124–130.
12. Kweka EJ, Zhou G, Beilhe LB, Dixit A, Afrane Y, Gilbreath TM, Munga S, Nyindo M, Githeko AK, Yan G: **Effects of co-habitation between Anopheles gambiae s.s. and Culex quinquefasciatus aquatic stages on life history traits.** *Parasit Vectors* 2012, **5:**33.
13. Craig MH, Snow RW, le Sueur D: **A climate-based distribution model of malaria transmission in Sub-Saharan Africa.** *Parasitol Today* 1999, **15:**105–111.
14. Parham PE, Michael E: **Modeling the effects of weather and climate change on malaria transmission.** *Environ Health Perspect* 2010, **118:**620–626.
15. Martens W: *Health impacts of climate change and ozone depletion. An eco-epidemiological modelling approach*. The Netherlands: Maastricht University Press; 1997.
16. Depinay JMO, Mbogo CM, Killeen G, Knols B, Beier J, Carlson J, Dushoff J, Billingsley P, Mwambi H, Githure J, Toure AM, McKenzie FE: **A simulation model of African Anopheles ecology and population dynamics for the analysis of malaria transmission.** *Malar J* 2004, **3:**29.
17. White MT, Griffin JT, Churcher TS, Ferguson NM, Basanez MG, Ghani AC: **Modelling the impact of vector control interventions on Anopheles gambiae population dynamics.** *Parasit Vectors* 2011, **4:**153.
18. Ermert V, Fink AH, Jones AE, Morse AP: **Development of a new version of the Liverpool Malaria Model. I. Refining the parameter settings and mathematical formulation of basic processes based on a literature review.** *Malar J* 2011, **10:**35.
19. Martens W, Jetten T, Rotmans J, Niessen L: **Climate change and vector-borne diseases: A global modelling perspective.** *Global Environmental Change* 1995, **5:**195–209.
20. Martens WJ, Niessen LW, Rotmans J, Jetten TH, McMichael AJ: **Potential impact of global climate change on malaria risk.** *Environ Health Perspect* 1995, **103:**458–464.
21. Martens W: *Health impacts of climate change and ozone depletion: an eco-epidemiological modelling approach. PhD thesis*. Maastricht, Netherlands: Maastricht University; 1997.
22. Parham PE, Michael E: **Modeling the effects of weather and climate change on malaria transmission.** *Environ Health Perspect* 2010, **118:**620–626.
23. Ruiz D, Poveda G, Velez ID, Quinones ML, Rua GL, Velasquez LE, Zuluaga JS: **Modelling entomological-climatic interactions of Plasmodium falciparum malaria transmission in two Colombian endemic-regions: contributions to a National Malaria Early Warning System.** *Malar J* 2006, **5:**66.
24. Bayoh N: *Studies on the development and survival of Anopheles gambiae sensu stricto at various temperatures and relative humidities. PhD thesis*: University of Durham; 2001.
25. Parham PE, Pople D, Christiansen-Jucht C, Lindsay S, Hinsley W, Michael E: **Modeling the role of environmental variables on the population dynamics of the malaria vector *Anopheles gambiae* sensu stricto.** *Malar J* 2012, **11:**271.
26. Mordecai EA, Paaijmans KP, Johnson LR, Balzer C, Ben-Horin T, de Moor E, McNally A, Pawar S, Ryan SJ, Smith TC, Lafferty KD, Thrall P: **Optimal temperature for malaria transmission is dramatically lower than previously predicted.** *Ecol Lett* 2013, **16:**22–30.
27. Lunde TM, Korecha D, Loha E, Sorteberg A, Lindtjørn B: **A dynamic model of some malaria-transmitting anopheline mosquitoes of the Afrotropical region. I. Model description and sensitivity analysis.** *Malaria J* 2013. in press.
28. Afrane YA, Zhou G, Lawson BW, Githeko AK, Yan G: **Effects of microclimatic changes caused by deforestation on the survivorship and reproductive fitness of *Anopheles gambiae* in western Kenya highlands.** *Am J Trop Med Hyg* 2006, **74:**772–778.
29. Harrington LC, Vermeylen F, Jones JJ, Kitthawee S, Sithiprasasna R, Edman JD, Scott TW: **Age-dependent survival of the dengue vector Aedes aegypti (Diptera: Culicidae) demonstrated by simultaneous release-recapture of different age cohorts.** *J Med Entomol* 2008, **45:**307–413.
30. Lyimo EO, Koella JC: **Relationship between body size of adult *Anopheles gambiae s.l.* and infection with the malaria parasite *Plasmodium falciparum*.** *Parasitology* 1992, **104:**233–237.
31. Ameneshewa B, Service MW: **The relationship between female body size and survival rate of the malaria vector Anopheles arabiensis in Ethiopia.** *Med Vet Entomol* 1996, **10:**170–172.
32. Hoshen M, Morse A: **A model structure for estimating malaria risk.** In *Environmental Change and Malaria Risk: Global and Local Implications, Volume 9*. Edited by Takken W, Martens P, Bogers RJ. Dordrecht, The Netherlands: Springer; 2005:10.

33. Scholte EJ, Knols BGJ, Takken W: **Infection of the malaria mosquito *Anopheles gambiae* with the entomopathogenic fungus *Metarhizium anisopliae* reduces blood feeding and fecundity.** *J Invertebr Pathol* 2006, **91:**43–49.
34. Kikankie C: *Susceptibility of laboratory colonies of members of the Anopheles gambiae complex to entomopathogenic fungi Beauveria bassiana. Master's thesis*. Johannesburg: University of the Witwatersrand; 2009.
35. Taylor KE: **Summarizing multiple aspects of model performance in a single diagram.** *J Geophys Res* 2001, **106:**7183–7192.
36. Akaike H: **A new look at the statistical model identification.** *IEEE Trans Autom Control* 1974, **19:**716–723.
37. Byrd RH, Lu-Chen PH, Nocedal J, Zhu CY: **A limited memory algorithm for bound constrained optimization.** *Siam J Scientific Comput* 1995, **16:**1190–1208.
38. R Development Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2011.
39. Gething PW, Van Boeckel TP, Smith DL, Guerra CA, Patil AP, Snow RW, Hay SI: **Modelling the global constraints of temperature on transmission of Plasmodium falciparum and P. vivax.** *Parasit Vectors* 2011, **4:**92.
40. Dong Y, Morton JC, Ramirez JL, Souza-Neto JA, Dimopoulos G: **The entomopathogenic fungus Beauveria bassiana activate toll and JAK-STAT pathway-controlled effector genes and anti-dengue activity in *Aedes aegypti*.** *Insect Biochem Mol Biol* 2012, **42:**126–132.
41. Hardstone MC, Huang X, Harrington LC, Scott JG: **Differences in development, glycogen, and lipid content associated with cytochrome P450-mediated permethrin resistance in *Culex pipiens quinquefasciatus* (Diptera: Culicidae).** *J Med Entomol* 2010, **47:**188–198.
42. Glunt KD, Thomas MB, Read AF: **The effects of age, exposure history and malaria infection on the susceptibility of *Anopheles* mosquitoes to low concentrations of pyrethroid.** *PLoS One* 2011, **6:**e24968.
43. White LJ, Maude RJ, Pongtavornpinyo W, Saralamba S, Aguas R, Van Effelterre T, Day NPJ, White NJ: **The role of simple mathematical models in malaria elimination strategy design.** *Malar J* 2009, **8:**212.
44. Killeen GF, Smith TA: **Exploring the contributions of bed nets, cattle, insecticides and excitorepellency to malaria control: a deterministic model of mosquito host-seeking behaviour and mortality.** *Trans R Soc Trop Med Hyg* 2007, **101:**867–880.
45. Saul AJ, Graves PM, Kay BH: **A cyclical feeding model for Pathogen transmission and its application to determine vectorial capacity from vector infection rates.** *J Appl Ecol* 1990, **27:**123–133.

doi:10.1186/1756-3305-6-20
Cite this article as: Lunde *et al.*: How malaria models relate temperature to malaria transmission. *Parasites & Vectors* 2013 **6**:20.

Bibliography

[1] Rebecca S Levine, A Townsend Peterson, and Mark Q Benedict. Geographic and ecologic distributions of the anopheles gambiae complex predicted using a genetic algorithm. *Am J Trop Med Hyg*, 70(2):105–9, February 2004. (document), 1.1, 2.2, 3.3, 3.5.2

[2] Alexander Moffett, Nancy Shackelford, and Sahotra Sarkar. Malaria in africa: vector species' niche models and relative risk maps. *PLoS One*, 2(9):e824, 2007. (document), 1.1, 2.2, 3.3, 3.5.2

[3] David J Rogers, Sarah E Randolph, Robert W Snow, and Simon I Hay. Satellite imagery in the study and forecast of malaria. *Nature*, 415(6872):710–5, February 2002. (document), 1.1, 2.2, 3.3, 3.5.2

[4] Marianne E Sinka, Michael J Bangs, Sylvie Manguin, Maureen Coetzee, Charles M Mbogo, Janet Hemingway, Anand P Patil, Will H Temperley, Peter W Gething, Caroline W Kabaria, Robi M Okara, Thomas Van Boeckel, H Charles J Godfray, Ralph E Harbach, and Simon I Hay. The dominant anopheles vectors of human malaria in africa, europe and the middle east: occurrence data, distribution maps and bionomic precis. *Parasit Vectors*, 3:117, 2010. (document), 1.1, 2.2, 3.5.1, 3.3, 3.5.2, 4.4.3

[5] Brian Greenwood and Theonest Mutabingwa. Malaria in 2002. *Nature*, 415(6872):670–2, February 2002. 1.1, 3.1, 3.3, 3.4.2

[6] World Health Organization. *World malaria report 2011*. World Health Organization, Switzerland, 2011. 1.1, 3.3

[7] D J Rogers and S E Randolph. The global spread of malaria in a future, warmer world. *Science*, 289(5485):1763–6, September 2000. 1.1

[8] G. MacDonald. *Dynamics of Tropical Disease*. Oxford University Press, London, 1973. 1.1, 5

[9] M. H. Craig, R. W. Snow, and D. le Sueur. A climate-based distribution model of malaria transmission in sub-saharan africa. *Parasitology Today*, 15(3):105–111, 1999. 1.1

[10] Volker Ermert, Andreas H Fink, Anne E Jones, and Andrew P Morse. Development of a new version of the liverpool malaria model. i. refining the parameter settings and mathematical formulation of basic processes based on a literature review. *Malar J*, 10(1):35, 2011. 1.1, 3.1, 4.4.3

[11] Paul E Parham, Diane Pople, CÃ©line Christiansen-Jucht, Steve Lindsay, Wes Hinsley, and Edwin Michael. Modeling the role of environmental variables on the population dynamics of the malaria vector anopheles gambiae sensu stricto. *Malar J*, 11(1):271, August 2012. 1.1, 3.1

[12] Paul Reiter. Global warming and malaria: knowing the horse before hitching the cart. *Malar J*, 7 Suppl 1:S3, 2008. 1.1, 3.1

[13] Daniel Ruiz, German Poveda, Ivan D Velez, Martha L Quinones, Guillermo L Rua, Luz E Velasquez, and Juan S Zuluaga. Modelling entomological-climatic interactions of plasmodium falciparum malaria transmission in two colombian endemic-regions: contributions to a national malaria early warning system. *Malar J*, 5:66, 2006. 1.1

[14] S W Lindsay, L Parson, and C J Thomas. Mapping the ranges and relative abundance of the two principal african malaria vectors, anopheles gambiae sensu stricto and an. arabiensis, using climate data. *Proc Biol Sci*, 265(1399):847–54, May 1998. 1.1, 2.2, 3.5.2

[15] Inaki Tirados, Gabriella Gibson, Stephen Young, and Stephen J Torr. Are herders protected by their herds? an experimental analysis of zooprophylaxis against the malaria vector anopheles arabiensis. *Malar J*, 10:68, 2011. 1.1

[16] W.J.M. Martens. *Health impacts of climate change and ozone depletion. An eco-epidemiological modelling approach*. Maastricht University Press, The Netherlands, 1997. 1.1

[17] Paul Edward Parham and Edwin Michael. Modeling the effects of weather and climate change on malaria transmission. *Environ Health Perspect*, 118(5):620–6, May 2010. 1.1

[18] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers. A description of the advanced research wrf version 2. Technical report, The National Center for Atmospheric Research, 2005. 2.1

[19] Caroline Fouet, Emilie Gray, Nora J Besansky, and Carlo Costantini. Adaptation to aridity in the malaria mosquito anopheles gambiae: chromosomal inversion polymorphism and body size influence resistance to desiccation. *PLoS One*, 7(4):e34841, 2012. 2.1, 3.1, 3.4.2

[20] Emilie M. Gray and Timothy J. Bradley. Physiology of desiccation resistance in anopheles gambiae and anopheles arabiensis. *Am J Trop Med Hyg*, 73(3):553–559, 2005. 2.1, 3.1

[21] E O Lyimo and J C Koella. Relationship between body size of adult anopheles gambiae s.l. and infection with the malaria parasite plasmodium falciparum. *Parasitology*, 104 (Pt 2):233–7, April 1992. 2.1

[22] B Ameneshewa and M W Service. The relationship between female body size and survival rate of the malaria vector anopheles arabiensis in ethiopia. *Med Vet Entomol*, 10(2):170–2, April 1996. 2.1

[23] C GARRETT-JONES. The possibility of active long-distance migrations by anopheles pharoensis theobald. *Bulletin World Health Organization*, 27(2):299–302, 1962. 2.1, 3.5.1

[24] G B White. Evidence for anopheles squamosus migration? *Nature*, 227(5259):739–40, August 1970. 2.1

[25] J G Ming, H Jin, J R Riley, D R Reynolds, A D Smith, R L Wang, J Y Cheng, and X N Cheng. Autumn southward 'return' migration of the mosquito culex tritaeniorhynchus in china. *Med Vet Entomol*, 7(4):323–7, October 1993. 2.1

[26] Ibrahima Baber, Moussa Keita, Nafomon Sogoba, Mamadou Konate, M'Bouye Diallo, Seydou Doumbia, SF Traore, JM Ribeiro, and Nicholas C Manoukis. Population size and migration of anopheles gambiae in the bancoumana region of mali and their significance for efficient vector control. *PLoS One*, 5(4):e10270, 2010. 2.1

[27] Abdoulaye Adamou, Adama Dao, Seydou Timbine, Yaya Kassogue, Alpha Seydou Yaro, Moussa Diallo, Sekou F Traore, Diana L Huestis, and Tovi Lehmann. The contribution of aestivating mosquitoes to the persistence of anopheles gambiae in the sahel. *Malar J*, 10:151, 2011. 2.1

[28] M. Coetzee, M.H. Craig, and D. le Sueur. Mapping the distribution of members of the anopheles gambiae complex in africa and adjacent islands. *Parasitology Today*, 16:74–77, 2000. 2.2

[29] Biodiversity occurrence data published by: Walter reed biosystematics unit smithsonian institution (accessed through gbif data portal, data.gbif.org, 2011-09-07). 2.2

[30] Walter Deshler. Cattle in africa: Distribution, types, and problems. *Geographical Review*, 53(1):52–58, 1963. 2.3

[31] Timothy Robinson and Fao's Animal Production and Health Division. Observed livestock densities, 6 2011. 2.3, 4.4.2

[32] Erin A Mordecai, Krijn P Paaijmans, Leah R Johnson, Christian Balzer, Tal Ben-Horin, Emily de Moor, Amy McNally, Samraat Pawar, Sadie J Ryan, Thomas C Smith, Kevin D Lafferty, and Peter Thrall. Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecol Lett*, October 2012. 2.4, 4.4.3

[33] Corine Karema, Maru W Aregawi, Alphonse Rukundo, Alain Kabayiza, Monique Mulindahabi, Ibrahima S Fall, Khoti Gausi, Ryan O Williams, Michael Lynch, Richard Cibulskis, Ngabo Fidele, Jean-Pierre Nyemazi, Daniel Ngamije, Irenee Umulisa, Robert Newman, and Agnes Binagwaho. Trends in malaria cases, hospital admissions and deaths following scale-up of anti-malarial interventions, 2000-2010, rwanda. *Malar J*, 11(1):236, July 2012. 3.1

[34] C Lengeler. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev*, (2):CD000363, 2004. 3.1, 3.4.1

[35] Eskindir Loha and Bernt Lindtjorn. Predictors of plasmodium falciparum malaria incidence in chano mille, south ethiopia: A longitudinal study. *Am J Trop Med Hyg*, July 2012. 3.1

[36] Devanand Moonasar, Tej Nutulaganti, Philip S Kruger, Aaron Mabuza, Eric S Raswiswi, Frew G Benson, and Rajendra Maharaj. Malaria control in south africa 2000-2010: beyond mdg6. *Malar J*, 11(1):294, August 2012. 3.1

[37] Meshesha Balkew, Muntaser Ibrahim, Lizette L Koekemoer, Basil D Brooke, Howard Engers, Abraham Aseffa, Teshome Gebre-Michael, and Ibrahim Elhassen. Insecticide resistance in anopheles arabiensis (diptera: Culicidae) from villages in central, northern and south west ethiopia and detection of kdr mutation. *Parasit Vectors*, 3(1):40, 2010. 3.1, 3.4.2

[38] P Carnevale, J C Toto, P Guibert, M Keita, and S Manguin. Entomological survey and report of a knockdown resistance mutation in the malaria vector anopheles gambiae from the republic of guinea. *Trans R Soc Trop Med Hyg*, 104(7):484–9, July 2010. 3.1, 3.4.2

[39] F Chen, K Mitchell, J Schaake, YK Xue, HL Pan, V Koren, QY Duan, M Ek, and A Betts. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES*, 101(D3):7251–7268, MAR 20 1996. 3.1, 3.4.2

[40] M. Coleman, B. Sharp, I. Seocharan, and J. Hemingway. Developing an evidence-based decision support system for rational insecticide choice in the control of african malaria vectors. *Journal of Medical Entomology*, 43(4):663–668, 2006. 3.1, 3.4.2

[41] Nelson Cuamba, Kwang Shik Choi, and Harold Townson. Malaria vectors in angola: distribution of species and molecular forms of the anopheles gambiae complex, their pyrethroid insecticide knockdown resistance (kdr) status and plasmodium falciparum sporozoite rates. *Malar J*, 5:2, 2006. 3.1, 3.4.2

[42] Cyrille Czeher, Rabiou Labbo, Gaelle Vieville, Ibrahim Arzika, HervÃ© Bogreau, Christophe Rogier, Laure Diancourt, Sylvain Brisse, FrÃ©dÃ©ric Ariey, and Jean-Bernard Duchemin. Population genetic structure of anopheles gambiae and anopheles arabiensis in niger. *J Med Entomol*, 47(3):355–66, May 2010. 3.1, 3.4.2

[43] K R Dabira, A Diabata, F Agostinho, F Alves, L Manga, O Faye, and T Baldet. Distribution of the members of anopheles gambiae and pyrethroid knock-down resistance gene (kdr) in guinea-bissau, west africa. *Bull Soc Pathol Exot*, 101(2):119–23, April 2008. 3.1, 3.4.2

[44] Joaquim DaSilva, Brad Garanganga, Vonai Teveredzi, Sabine Marx, Simon Mason, and Stephen Connor. Improving epidemic malaria planning, preparedness and response in southern africa. *Malaria Journal*, 3(1):37, 2004a. 3.1, 3.4.2

[45] Jean-Marc O Depinay, Charles M Mbogo, Gerry Killeen, Bart Knols, John Beier, John Carlson, Jonathan Dushoff, Peter Billingsley, Henry Mwambi, John Githure, Abdoulaye M Toure, and F Ellis McKenzie. A simulation model of african anopheles ecology and population dynamics for the analysis of malaria transmission. *Malar J*, 3:29, July 2004. 3.1, 3.4.2

[46] Luc Djogbenou, Roch Dabire, Abdoulaye Diabate, Pierre Kengne, Martin Akogbeto, Jean Marc Hougard, and Fabrice Chandre. Identification and geographic distribution of the ace-1r mutation in the malaria vector anopheles gambiae in south-western burkina faso, west africa. *Am J Trop Med Hyg*, 78(2):298–302, February 2008. 3.1, 3.4.2

[47] Rousseau F Djouaka, Adekunle A Bakare, Honore S Bankole, Julien Mc Doannio, Ousmane N Coulibaly, Hortense Kossou, Manuele Tamo, Harcourt I Basene, O K Popoola, and Martin C Akogbeto. Does the spillage of petroleum products in anopheles breeding sites have an impact on the pyrethroid resistance? *Malar J*, 6:159, 2007. 3.1, 3.4.2

[48] K. L. Ebi, J. Hartman, N. Chan, J. McConnell, M. Schlesinger, and J. Weyant. Climate suitability for stable malaria transmission in zimbabwe under different climate change scenarios. *Climatic Change*, 73(3):375–393, 2005. 3.1, 3.4.2

[49] Brian M. Greenwood, Kalifa Bojang, Christopher J. M. Whitty, and Geoffrey A. T. Targett. Malaria. *The Lancet*, 365(9469):1487–1498, 2005. 3.1, 3.4.2

[50] D. L. Hartl. The origin of malaria: Mixed messages from genetic diversity. *Nature Reviews Microbiology*, 2(1):15–22, 2004. 3.1, 3.4.2

[51] Y El-S Himeidan, M Y Dukeen, El-A El-Rayah, and I Adam. Anopheles arabiensis: abundance and insecticide resistance in an irrigated area of eastern sudan. *East Mediterr Health J*, 10(1-2):167–74, 2004. 3.1, 3.4.2

[52] Rh Hunt, M Edwardes, and M Coetzee. Pyrethroid resistance in southern african anopheles funestus extends to likoma island in lake malawi. *Parasit Vectors*, 3:122, 2010. 3.1, 3.4.2

[53] Hitoshi Kawada, Gabriel O Dida, Kazunori Ohashi, Osamu Komagata, Shinji Kasai, Takashi Tomita, George Sonye, Yoshihide Maekawa, Cassian Mwatele, Sammy M Njenga, Charles Mwandawiro, Noboru Minakawa, and Masahiro Takagi. Multimodal pyrethroid resistance in malaria vectors, anopheles gambiae s.s., anopheles arabiensis, and anopheles funestus s.s. in western kenya. *PLoS One*, 6(8):e22574, 2011. 3.1, 3.4.2

[54] Hitoshi Kawada, Kyoko Futami, Osamu Komagata, Shinji Kasai, Takashi Tomita, George Sonye, Cassian Mwatele, Sammy M Njenga, Charles Mwandawiro, Noboru Minakawa, and Masahiro Takagi. Distribution of a knockdown resistance mutation (l1014s) in anopheles gambiae s.s. and anopheles arabiensis in western and southern kenya. *PLoS One*, 6(9):e24323, 2011. 3.1, 3.4.2

[55] Clement Kerah-Hinzoumbe, Mallaye Peka, Philippe Nwane, Issa Donan-Gouni, Josiane Etang, Albert Same-Ekobo, and Frederic Simard. Insecticide resistance in anopheles gambiae from south-western chad, central africa. *Malar J*, 7:192, 2008. 3.1, 3.4.2

[56] Christophe K Kikankie, Basil D Brooke, Bart G J Knols, Lizette L Koekemoer, Marit Farenhorst, Richard H Hunt, Matthew B Thomas, and Maureen Coetzee. The infectivity of the entomopathogenic fungus beauveria bassiana to insecticide-resistant and susceptible anopheles arabiensis mosquitoes at two different temperatures. *Malar Journal*, 9:71, 2010. 3.1, 3.4.2

[57] malERA Consultative Group on Modeling. A research agenda for malaria eradication: modeling. *PLoS Med*, 8(1):e1000403, 2011. 3.1, 3.4.2

[58] Derrick K Mathias, Eric Ochomo, Francis Atieli, Maurice Ombok, M Nabie Bayoh, George Olang, Damaris Muhia, Luna Kamau, John M Vulule, Mary J Hamel, William A Hawley, Edward D Walker, and John E Gimnig. Spatial and temporal variation in the kdr allele l1014s in anopheles gambiae s.s. and phenotypic variability in susceptibility to insecticides in western kenya. *Malar J*, 10:10, 2011. 3.1, 3.4.2

[59] Givemore Munhenga. *Characterisation of resistance mechanisms in the major malaria vector Anopheles arabiensis from southern Africa*. PhD thesis, Faculty of Science, University of the Witwatersrand, Johannesburg, 2010. 3.1, 3.4.2

[60] Givemore Munhenga, Hieronymo T Masendu, Basil D Brooke, Richard H Hunt, and Lizette K Koekemoer. Pyrethroid resistance in the major malaria vector anopheles arabiensis from gwave, a malaria-endemic area in zimbabwe. *Malar J*, 7:247, 2008. 3.1, 3.4.2

[61] Ernesto Capanna. Grassi versus ross: who solved the riddle of malaria? *Int Microbiol*, 9(1):69–74, March 2006. 3.1

[62] R. Ross. Life-history of the parasites of malaria. *Nature*, 60(1553):322–324, 1899. 3.1

[63] Alfred J. Lotka. Contribution to the analysis of malaria epidemiology. i. general part. *Am. J. Epidemiol.*, 3(supp1):1–36, 1923. 3.1

[64] Alfred J. Lotka. Contribution to the analysis of malaria epidemiology. ii. general part (continued). comparison of two formulae given by sir ronald ross. *Am. J. Epidemiol.*, 3(supp1):38–54, 1923. 3.1

[65] Alfred J. Lotka. Contribution to the analysis of malaria epidemiology. iii. numerical part. *Am. J. Epidemiol.*, 3(supp1):55–95, 1923. 3.1

[66] Alfred J. Lotka and F. R. Sharpe. Contribution to the analysis of malaria epidemiology. iv. incubation lag. *Am. J. Epidemiol.*, 3(supp1):96–112, 1923. 3.1

[67] Ronald Ross. *Report on the prevention of Malaria in Mauritius*. Waterlow and Sons Limited, London, 1908. 3.1

[68] R. Ross. *The prevention of malaria*. London, Murray, 1911. 3.1

[69] H. Waite. Mosquitoes and malaria. a study of the relation between the number of mosquitoes in a locality and the malaria rate. *Biometrika*, 7(4):421–436, 1910. 3.1

[70] G. MacDonald. *The epidemiology and control of malaria*. Oxford University Press, London, 1957. 3.1

[71] World Health Organization. *Guidelines for the treatment of malaria - 2nd edition*, volume 2. World Health Organization, Geneva, 2010. 3.1

[72] K Dietz, L Molineaux, and A Thomas. A malaria model tested in the african savannah. *Bull World Health Organ*, 50(3-4):347–57, 1974. 3.1

[73] L Molineaux, K Dietz, and A Thomas. Further epidemiological evaluation of a malaria model. *Bull World Health Organ*, 56(4):565–71, 1978. 3.1

[74] H M Yang. Malaria transmission model for different levels of acquired immunity and temperature-dependent parameters (vector). *Rev Saude Publica*, 34(3):223–31, June 2000. 3.1, 5

[75] H. M. Yang. A mathematical model for malaria transmission relating global warming and local socioeconomic conditions. *Revista De Saude Publica*, 35(3):224–231, 2001. 3.1

[76] Arne Bomblies, Jean-Bernard Duchemin, and Elfatih A. B. Eltahir. Hydrology of malaria: Model development and application to a sahelian village. *Water Resour. Res.*, 44(12):W12445, 2008. 3.1

[77] Hong-Wei Gao, Li-Ping Wang, Song Liang, Yong-Xiao Liu, Shi-Lu Tong, Jian-Jun Wang, Ya-Pin Li, Xiao-Feng Wang, Hong Yang, Jia-Qi Ma, Li-Qun Fang, and Wu-Chun Cao. Change in rainfall drives malaria re-emergence in anhui province, china. *PLoS One*, 7(8):e43686, 2012. 3.1

[78] Paul Edward Parham and Edwin Michael. Modeling the effects of weather and climate change on malaria transmission. *Environ Health Perspect*, 118(5):620–6, May 2010. 3.1

[79] D Metselaar and PH van Thiel. Classification of malaria. *Tropical Geographical Medicine*, 11:157–161, 1959. 3.2

[80] Carlos A. Guerra. *Mapping the contemporary global distribution limits of malaria using empirical data and expert opinion*. PhD thesis, Linacre College, University of Oxford, 2007. 3.2, 3.3

[81] Lewis Hackett. *Malaria in Europe, an ecological study*. London: Oxford University Press, 1937. 3.3

[82] Peter W. Gething, David L. Smith, Anand P. Patil, Andrew J. Tatem, Robert W. Snow, and Simon I. Hay. Climate change and the global malaria recession. *Nature*, 465(7296):342–345, May 2010. 3.3

[83] Thomas P Eisele, David Larsen, and Richard W Steketee. Protective efficacy of interventions for preventing malaria mortality in children in plasmodium falciparum endemic areas. *Int J Epidemiol*, 39 Suppl 1:i88–101, April 2010. 3.4.1

[84] M Nabie Bayoh, Derrick K Mathias, Maurice R Odiere, Francis M Mutuku, Luna Kamau, John E Gimnig, John M Vulule, William A Hawley, Mary J Hamel, and Edward D Walker. Anopheles gambiae: historical population decline associated with regional distribution of insecticide-treated bed nets in western nyanza province, kenya. *Malar J*, 9:62, 2010. 3.4.1

[85] Bianca Pluess, Frank C Tanser, Christian Lengeler, and Brian L Sharp. Indoor residual spraying for preventing malaria. *Cochrane Database Syst Rev*, (4):CD006657, 2010. 3.4.2

[86] Ulrike Fillinger and Steven W Lindsay. Larval source management for malaria control in africa: myths and reality. *Malar J*, 10:353, 2011. 3.4.4

[87] Gregor J Devine and Gerry F Killeen. The potential of a new larviciding method for the control of malaria vectors. *Malar J*, 9:142, 2010. 3.4.4

[88] M. T. Gillies and Botha De Meillon. *The Anophelinae of Africa South of the Sahara (Ethiopian zoogeographical region)*. Johannesburg: The South African Institute for Medical Research, 1968. 3.5.1

[89] M. T. Gillies and M. Coetzee. *A supplement to the Anophelinae of Africa south of the Sahara (Afrotropical region)*. The South African Institute for Medical Research, 1987. 3.5.1

[90] M Coetzee and D. Fontenille. Advances in the study of anopheles funestus, a major vector of malaria in africa. *Insect Biochem Mol Biol*, 34:599–605, 2004. 3.5.1

[91] B.J. White, FH Collins, and NJ Besansky. Evolution of anopheles gambiae in relation to humans and malaria. *Annu Rev Ecol Evol Sys*, 42, 2011. 3.5.2

[92] J.S. Kain and J.M. Fritsch. Convective parameterization for mesoscale models: The kain–fritsch scheme. *The Representation of Cumulus Convection in Numerical Models. Metor. Monogr., Amer. Meteor. Soc.*, (46):165–170, 1993. 4.4

[93] M. Kanamitsu, W. Ebisuzaki, J. Woollen, S. K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter. NCEP-DEO AMIP-II reanalysis (r-2). *Bulletin of the American Meteorological Society*, 83(11):1631–1643, 2002. 4.4.1

[94] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The era-interim reanalysis: configuration and performance of the data

assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, 2011. 4.4.1

[95] Center for International Earth Science Information Network (CIESIN) - Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). Gridded population of the world version 3 (gpwv3): Population density grids. Technical report, Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University, 2005. 4.4.2

[96] Andrew J Tatem, Susana Adamo, Nita Bharti, Clara R Burgert, Marcia Castro, Audrey Dorelien, Gunther Fink, Catherine Linard, John Mendelsohn, Livia Montana, Mark R Montgomery, Andrew Nelson, Abdisalan M Noor, Deepa Pindolia, Greg Yetman, and Deborah Balk. Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation. *Popul Health Metr*, 10(1):8, May 2012. 4.4.2

[97] Andy Nelson. Unep/grid - sioux falls dataset. 4.4.2

[98] Eddie A. Bright, Phil R. Coleman, Amy N. Rose, and Marie L. Urban. Landscan 2011, 2012. 4.4.2

[99] William Wint and Timothy Robinson. Gridded livestock of the world - 2007. Technical report, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, Rome, 2007. 4.4.2

[100] Nicholas C Manoukis, Ibrahima Baber, Moussa Diallo, Nafomon Sogoba, and JosÃ© M C Ribeiro. Seasonal climate effects anemotaxis in newly emerged adult anopheles gambiae giles in mali, west africa. *PLoS One*, 6(11):e26910, 2011. 4.4.3

[101] J. Tumwiine, J.Y.T. Mugisha, and L.S. Luboobi. A mathematical model for the dynamics of malaria in a human host and mosquito vector with temporary immunity. *Applied Mathematics and Computation*, 189(2):1953–1965, 2007. 5

[102] David Gurarie, Stephan Karl, Peter A Zimmerman, Charles H King, Timothy G St Pierre, and Timothy M E Davis. Mathematical modeling of malaria infection with innate and adaptive immunity in individuals and agent-based communities. *PLoS One*, 7(3):e34040, 2012. 5

[103] A Ducrot, S B Sirima, B Somé, and P Zongo. A mathematical model for malaria involving differential susceptibility, exposedness and infectivity of human host. *J Biol Dyn*, 3(6):574–98, November 2009. 5

[104] Tovi Lehmann, Adama Dao, Alpha Seydou Yaro, Abdoulaye Adamou, Yaya Kassogue, Moussa Diallo, TraorÃ© SÃ©kou, and Cecilia Coscaron-Arias. Aestivation of the african malaria mosquito, anopheles gambiae in the sahel. *Am J Trop Med Hyg*, 83(3):601–6, September 2010. 5

Errata

August 6, 2013

In paper I equation 32 should read:

$$\varpi_{N,m}(\alpha,\zeta,a)=\sum_{i=0}^{a}\left(\frac{\left((\alpha\cdot i)^{\sum_{j=0}^{n=(\zeta-1)}j}\right)}{\sum_{j=0}^{n=(\zeta-1)}j!}\right)\cdot e^{(-\alpha\cdot a)}, \tag{1}$$

where a is days since emergence. Equation 35 should read:

$$\varpi_{N,m}(\alpha,\zeta,a)=\sum_{i=0}^{a}(1+\alpha\cdot i)\cdot e^{-\alpha\cdot a}. \tag{2}$$