

## RESEARCH

## Open Access



# Environmental determinants of malaria transmission in African villages

Noriko Endo\* and Elfatih A. B. Eltahir

## Abstract

**Background:** Malaria transmission is complex, involving a range of hydroclimatological, biological, and environmental processes. The high degree of non-linearity in these processes makes it difficult to predict and intervene against malaria. This study seeks both to define a minimal number of malaria transmission determinants, and to provide a theoretical basis for sustainable environmental manipulation to prevent malaria transmission.

**Methods:** Using a field-tested mechanistic malaria model, HYDREMATS, a theoretical study was conducted under hypothetical conditions. Simulations were conducted with a range of hydroclimatological and environmental conditions: temperature ( $t$ ), length of wet season ( $T_{wet}$ ), storm inter-arrival time ( $T_{int}$ ), persistence of vector breeding pools ( $T_{on}$ ), and distribution of houses from breeding pools and from each other ( $X_{dist}$  and  $Y_{dist}$ , respectively). Based on the theoretical study, a malaria time scale,  $T_{or}$ , and a predictive theory of malaria transmission were introduced. The performance of the predictive theory was compared against the observational malaria transmission data in West Africa. Population density was used to estimate the scale that describes the spatial distribution of houses.

**Results:** The predictive theory shows a universality in malaria endemic conditions when plotted using two newly-introduced dimension-less parameters. The projected malaria transmission potential compared well with the observation data, and the apparent differences were discussed. The results illustrate the importance of spatial aspects in malaria transmission.

**Conclusions:** The predictive theory is useful in measuring malaria transmission potential, and it can also provide guidelines on how to plan the layout of human habitats in order to prevent endemic malaria. Malaria-resistant villages can be designed by locating houses further than critical distances away from breeding pools or by removing pools within a critical distance from houses; the critical distance is described in the context of local climatology and hydrology.

**Keywords:** Malaria transmission, Hydrology, Spatial impact, Characteristic time scale

## Background

The transmission dynamics of malaria are complex. Climatological, hydrological, and biological factors interact non-linearly at various stages of the transmission cycle and shape malaria transmission dynamics. Another key element of malaria transmission dynamics is environmental factors. The large dimension and non-linearity of those factors make it difficult to predict and intervene against malaria.

Malaria can be sustainably prevented through environmental modification approaches. Climatological, hydrological, and biological factors are hardly controllable; malaria treatment interventions require continuous effort to attenuate or avoid the impact of those factors. Environmental factors, on the other hand, can be manipulated; well-designed planning of environmental conditions could work as a long-term malaria prevention approach. Facing the low momentum in the public health arena of malaria control, due to fatigue of donors, spread of drug-resistance, fragile health infrastructures etc., prevention approaches, as opposed to treatment approaches,

\*Correspondence: [enori@mit.edu](mailto:enori@mit.edu)

Ralph M. Parsons Laboratory, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

could be more sustainable and cost-effective in fighting malaria.

Planning of house locations relative to vector breeding sites and removal of pools near houses could be effective malaria prevention techniques. The importance of the relative distributions of houses and pools has been reported both in simulation studies [1, 2] and in observational studies [3–6], spelling out the large risk of contracting malaria when houses are located near vector breeding pools. The threshold distance for high malaria transmission risk is reported to be some hundreds of meters [3, 4, 6]. This paper will demonstrate that the threshold distance to prevent malaria is not universal, but significantly influenced by climatology, hydrology, and biology of local vectors, and that those factors interact in a complex manner. Planning of malaria-resistant village, thus, may not be straightforward.

This study seeks both to reduce the dimension of malaria transmission determinants, and to aid environmental manipulation to prevent malaria. The objective of this paper is to provide a tool to analyse malaria transmission potential at various environments, including different house distribution conditions. The analyses are conducted for conditions where mosquito populations are limited primarily by water availability, such as the fringe of Sahel. Based on theoretical studies with a wide range of forcings and parameters, a predictive theory of malaria transmission potential will be presented. The theory will then be tested for West Africa. In testing the theory, the impact of house distribution is emphasized.

## Methods

### Malaria transmission simulation model

This study was conducted using a mechanistic malaria transmission model, HYDREMATS—HYDRology, Entomology, and Malaria Transmission Simulator. HYDREMATS is an agent-based model calibrated for relatively dry regions in West Africa [7–9]. HYDREMATS features the explicit representation of the location and the persistence of vector breeding pools, and the location of human houses. The spatial and temporal settings of the simulations are explained below.

### Spatial settings of simulation domain

Simulations were conducted under hypothetical settings. The spatial setup was a 1000 by 1500 m domain, having at the center a 20 m-wide breeding pool expanding longitudinally (Fig. 1a). The pool appears and disappears depending on the hydrological setting assumed, which is described in the following section. Houses are located at  $X_{\text{dist}}$  m away from the pool transversely at both sides of the pool, and  $Y_{\text{dist}}$  m away from each other longitudinally.

### Temporal settings of pools

The temporal settings employed are described in Fig. 1b. This study assumed that a year has a clear wet season for  $T_{\text{wet}}$  days and a dry season for the rest of the year. Pools appear intermittently at an interval of  $T_{\text{int}}$  days during a rainy season of  $T_{\text{wet}}$  days.  $T_{\text{int}}$  represents the storm inter-arrival time. Out of  $T_{\text{int}}$  days, pools persist for  $T_{\text{on}}$  days, gradually decreasing the probability of mosquitoes laying eggs and imposing persistence-dependent larval mortality, representing progressive disappearance of pools. Each experiment was run for the simulation period of four years, repeating the temporal setting described above.

A certain minimum number of mosquitoes were allowed to exist throughout simulations in order to seed mosquito population dynamics for the first and the subsequent rainy seasons after long dry periods. The parameter values used in the simulations are summarized in Table 1.

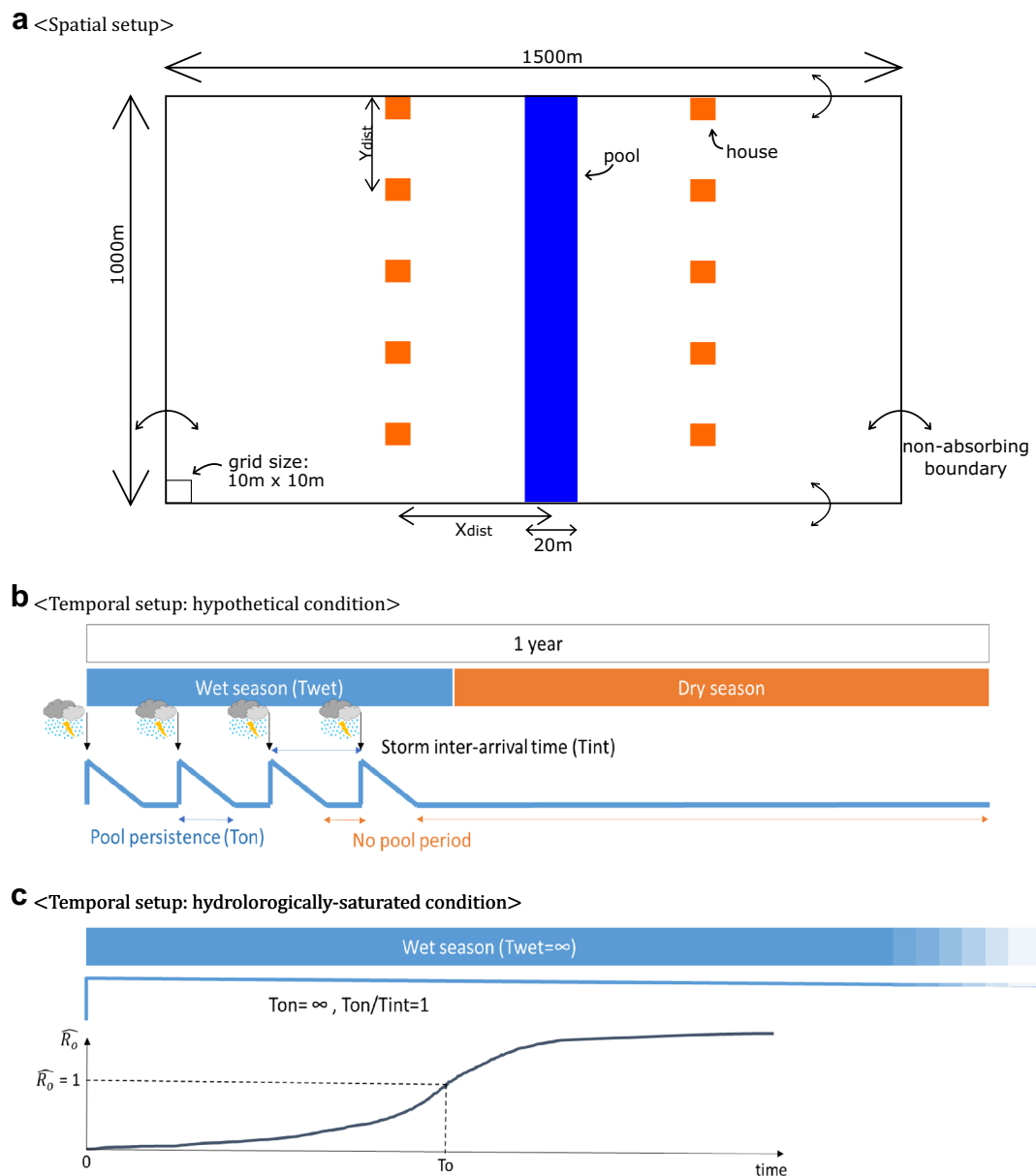
### Basic reproduction rate, $R_0$

The evaluation of the malaria transmission potential under specified conditions was based on the basic reproduction rate ( $R_0$ ), which is defined as the number of infections expected to be generated by a single infectious person in a totally susceptible population. The mechanistic structure of HYDREMATS is most suitable for estimating  $R_0$  under dynamic conditions; conventional Ross-MacDonald formula [10] assumes static conditions.  $R_0$  values for a given condition were estimated in HYDREMATS to follow the exact definition of  $R_0$ . HYDREMATS can simulate the status of each human, as well as each mosquito, and track infections (which human infects which mosquito, and which mosquito infects which human; hence it can calculate how many people were infected from an infectious person). In this way, HYDREMATS can calculate  $R_0$  for any dynamic environmental conditions and mosquito populations.

Long-term average of  $R_0$  (hereafter  $\hat{R}_0$ ) was calculated as follows, where malaria endemics was defined as  $\hat{R}_0 \geq 1$ .

$$\hat{R}_0 = \sum_{i=1}^{N_{\text{inf}}} R_0^i / N_{\text{inf}},$$

where  $N_{\text{inf}}$  is the number of infectious people simulated over the simulation period of four years, and  $R_0^i$  is the number of infections generated from the  $i$ -th infectious person, assuming the population is susceptible. Because the calculation of  $R_0$  requires the assumption of fully susceptible population, any dual infection was counted towards  $R_0^i$ . The number of simulated malaria cases originated from the  $i$ -th infectious person, thus, differs from  $R_0^i$ . Note that  $N_{\text{inf}}$  increases with time and that  $R_0^i$  is dictated by the time-varying vector population and



**Fig. 1** Spatial and temporal setup of the model. **a** Spatial setting in the simulations. Simulations were conducted under a 1000 m × 1500 m domain with 10 m × 10 m grids. A 20 m-width pool appears at the center of the domain. Houses are located  $X_{dist}$  away from the pool, and  $Y_{dist}$  away from each other. **b** Temporal setting: hypothetical condition. The hypothetical condition used in this study defines a clear wet and dry season in a year. During a wet season ( $T_{wet}$ ), pools are created at every storm inter-arrival time ( $T_{int}$ ), and persist for  $T_{on}$  days. Simulations were conducted for 4 years repeating the specified hypothetical hydrological conditions. **c** Temporal setting: hydrologically-saturated condition. The hydrologically saturated condition defines the condition where  $T_{wet} = \infty$  and  $T_{on}/T_{int} = 1$ . The persistent pool is assumed to be created at time zero. Under this condition, the time required to reach  $\hat{R}_0 = 1$  is defined as  $T_0$

environmental conditions.  $\hat{R}_0$  can also be defined as the average of  $R_0^i$  values.

**Time scale for stable malaria transmission,  $T_0$**

This study introduces a new malaria time scale,  $T_0$ .  $T_0$  (t) is defined as the length of the rainy season required

to reach  $\hat{R}_0 = 1$  under the condition of permanent water body (hereafter hydrologically-saturated condition) at a given temperature  $t$  and a given spatial condition (Fig. 1c). The hydrologically-saturated condition assumes a persistent pool (i.e.,  $T_{wet} = \infty$ ,  $T_{on} = \infty$  and  $T_{on}/T_{int} = 1$ ) that is created at the beginning of the simulation

**Table 1 Model parameters**

Variable	Parameter values	Additional values in sensitivity study	Unit
$X_{\text{dist}}$	50, 100, 150, 200, 250, 300	–	m
$Y_{\text{dist}}$	50	70, 30, 10	m
$T_{\text{wet}}$	1, 2, 3, 4, 5, 6, 8, 12	–	month
$T_{\text{on}}$	6, 9, 12, 15, 18, 21, 27	(depend on $T_{\text{int}}$ )	day
$T_{\text{int}}$	28	21, 14	day
Temperature (t)	19, 21, 23, 25, 27, 29, 31, 33, 35	(21, 27, 33 only)	°C

period.  $T_o$  is thus not a function of hydrological variables:  $T_{\text{wet}}$ ,  $T_{\text{int}}$ ,  $T_{\text{on}}$ . Under this condition,  $\hat{R}_o = 1$  keeps rising until the system becomes stable. The stability point may or may not exceed  $\hat{R}_o = 1$ . If the system stabilizes at  $\hat{R}_o < 1$ ,  $T_o$  was defined as infinite.  $T_o$  may or may not exceed 1 year.

With  $t^{\text{lr}}$ ,  $t^{\text{ad}}$ , and  $t^{\text{para}}$  being *temperatures* that influence larval development time, adult life span, and parasite development time, respectively,  $T_o(t)$  was defined as

$$T_o(t) = T_o(t^{\text{lr}}, t^{\text{ad}}, t^{\text{para}}),$$

where this study assumes  $t = t^{\text{lr}} = t^{\text{ad}} = t^{\text{para}}$ . One can also define  $t^{\text{lr}}$  as water temperature, and  $t^{\text{ad}}$  and  $t^{\text{para}}$  as air temperature. In this study, water and air temperatures were assumed to be equal.

In order to analyse the contributions of larval development time, adult life span, and parasite development time at different temperatures,  $T_o^{\text{lr}}(t)$ ,  $T_o^{\text{ad}}(t)$ ,  $T_o^{\text{para}}(t)$  were introduced:

$$T_o^{\text{lr}}(t) = T_o(t, t_{\text{ref}}, t_{\text{ref}}),$$

$$T_o^{\text{ad}}(t) = T_o(t_{\text{ref}}, t, t_{\text{ref}}),$$

$$T_o^{\text{para}}(t) = T_o(t_{\text{ref}}, t_{\text{ref}}, t),$$

where  $t_{\text{ref}}$  is a reference temperature set at 27 °C.

### Two dimensionless parameters

Two dimensionless parameters are introduced:  $D_1 = T_{\text{wet}}/T_o$  and  $D_2 = T_{\text{on}}/T_{\text{int}}$ .

$D_1$  measures the relative length of the wet season,  $T_{\text{wet}}$ , compared to the malaria time scale,  $T_o$ . Relatively small  $D_1$  indicates a wet season that is too short to create the critical size of mosquito population that may sustain malaria transmission.

$D_2$  describes the variability in hydrologic conditions determining the intermittency of pools. It indicates the proportion of time that pools exist during a rainy season, and also the deviation between the condition during a rainy season and under the hydrologically-saturated condition.  $D_2$  nears one when pools are persistent, and

zero when pools are short-lasting. As  $D_2$  reaches one, the condition becomes closer to the hydrologically-saturated condition ( $D_2 = 1$  deviates slightly from the hydrologically-saturated condition, where  $T_{\text{on}} = \infty$ , because finite  $T_{\text{on}}$  values impose gradual reduction in oviposition probability and the additional larval mortality).

### West Africa as a testing site

The results of the theoretical work were tested against observational data over West Africa. West Africa was selected as a testing site for the following three reasons. First, the largest burden of malaria is in Africa (e.g., 90% of deaths), of which sub-Saharan West Africa experiences disproportionately high mortality and morbidity [11]. Second, West Africa features strong North–South gradients in temperature and rainfall, resulting in a range of climate conditions and malaria vulnerability [9, 12, 13] (Additional file 1). Finally, detailed hydrological data are available for West Africa from Yamana's study [9]. HYDREMATS has been calibrated successfully in previous studies [7–9] over the semi-arid climate in West Africa.

### Estimation of model parameters for West Africa

Hydroclimatological data and population data were used to estimate the dimensionless parameters,  $D_1$  and  $D_2$ , which were then applied to estimate malaria transmission potential. Hydroclimatological data used were monthly temperature and rainfall data from 2001 to 2010 from CRU TS 3.21 [14]. Population density ( $PD$ ) data were obtained from Gridded Population of the World, Version 3 (GPWv3) for the year 2000 [15].

The parameters used for estimating  $D_1$  and  $D_2$  are summarized in Table 2.  $D_1$  was obtained using  $T_{\text{wet}}$  and  $T_o$ .  $T_{\text{wet}}$  is the length of the rainy season. In the observational data, a rainy season was defined as the period from when annual cumulative rainfall exceeds 10% until when it reaches 95%. This definition works well for West Africa because this region experiences just one distinct rainy season per year (Additional file 2). In estimating the malaria time scale,  $T_o$ , information on temperature,  $X_{\text{dist}}$ ,

**Table 2 Model parameter estimation**

Dimension-less number	Environmental parameters	Parameter estimation
D <sub>1</sub>	T <sub>wet</sub>	Rainy season length, during which annual cumulative rainfall is >10% and <95%
	T <sub>o</sub>	Rainy season average temperature
	t	
	X <sub>dist</sub>	Implied from population density
	Y <sub>dist</sub>	Set at constant
D <sub>2</sub>	T <sub>on</sub>	Estimated as T <sub>on</sub> /T <sub>int</sub> from the empirical relationship
	T <sub>int</sub>	

and Y<sub>dist</sub> are required. The temperature used was the average temperature over the rainy season, because this is the major period when mosquito population dynamics and malaria transmission occur.

Estimation of X<sub>dist</sub> and Y<sub>dist</sub> is not easy, because houses do not line up as assumed in the hypothetical setting in this study. For simplicity, Y<sub>dist</sub> was set at a fixed value of 50 m. X<sub>dist</sub> was estimated by comparing the observational PD (Additional file 3) and the PD implied from X<sub>dist</sub> (PD<sub>imp</sub>(X<sub>dist</sub>)). The implied relationship between PD<sub>imp</sub> and X<sub>dist</sub> was derived assuming the following: (a) five people live in a house, (b) household density of a village is 1/X<sub>dist</sub><sup>2</sup> (per m<sup>2</sup>), and (c) a fraction (1/α) of an area is populated by villages; PD<sub>imp</sub> = 5 × X<sub>dist</sub><sup>-2</sup> × α<sup>-1</sup> × 10<sup>-6</sup> (per km<sup>2</sup>). α was set at 15, so that the distribution of PD values can be captured more accurately within the X<sub>dist</sub> values used in this study. Using the observational PD, X<sub>dist</sub> was inferred (Additional file 3):

$$X_{dist} = \begin{cases} 50 & \text{if } PD \geq PD_{imp}(X_{dist} = 50) \\ 100 & \text{if } PD \geq PD_{imp}(100) \text{ and } PD < PD_{imp}(50) \\ 150 & \text{if } PD \geq PD_{imp}(150) \text{ and } PD < PD_{imp}(100) \\ 200 & \text{if } PD \geq PD_{imp}(200) \text{ and } PD < PD_{imp}(150) \\ 250 & \text{if } PD \geq PD_{imp}(250) \text{ and } PD < PD_{imp}(200) \\ 300 & \text{if } PD \geq PD_{imp}(300) \end{cases}$$

D<sub>2</sub> is the proportion of pool persistence (T<sub>on</sub>) to the storm inter-arrival time (T<sub>int</sub>). In obtaining D<sub>2</sub>, instead of estimating T<sub>on</sub> and T<sub>int</sub> separately, the values of T<sub>on</sub>/T<sub>int</sub> were estimated. Although pool persistence was simulated explicitly for West Africa in Yamana’s study [9], determining representative T<sub>on</sub> values is challenging because natural climate forcing is neither regular nor periodic, and because T<sub>on</sub> values vary significantly for each storm event. However, the total of T<sub>on</sub> over the rainy season, T<sub>on</sub><sup>tot</sup>, can be obtained easily without any need to differentiate storm events. The total of T<sub>int</sub> over the rainy

season, on the other hand, equals T<sub>wet</sub> (Fig. 1b). Thus, in the hypothetical simulation setting, the proportion of T<sub>on</sub><sup>tot</sup> to T<sub>wet</sub> equals the proportion of T<sub>on</sub> to T<sub>int</sub> (i.e., T<sub>on</sub><sup>tot</sup>/T<sub>wet</sub> = T<sub>on</sub>/T<sub>int</sub>). Moreover, although T<sub>on</sub><sup>tot</sup> data are available only for the twelve sites used in Yamana’s study [9], a strong correlation between T<sub>on</sub><sup>tot</sup>/T<sub>wet</sub> and *annual-rainfall*/T<sub>wet</sub> was found (Additional file 4). T<sub>on</sub><sup>tot</sup>/T<sub>wet</sub> was found to scale empirically with *annual-rainfall*/T<sub>wet</sub> with a factor of proportionality equal to 0.14.

Thus, D<sub>2</sub> can be estimated from a globally available rainfall data set. The empirical equation relating D<sub>2</sub> and rainfall data reads:

$$D_2 = T_{on}/T_{int} = T_{on}^{tot}/T_{wet} = \textit{annual-rainfall}/T_{wet} \times 0.14$$

Although information on T<sub>on</sub> and T<sub>int</sub> is not usually available separately, Yamana’s study [9] provides reliable estimates for T<sub>on</sub>/T<sub>int</sub> or D<sub>2</sub>.

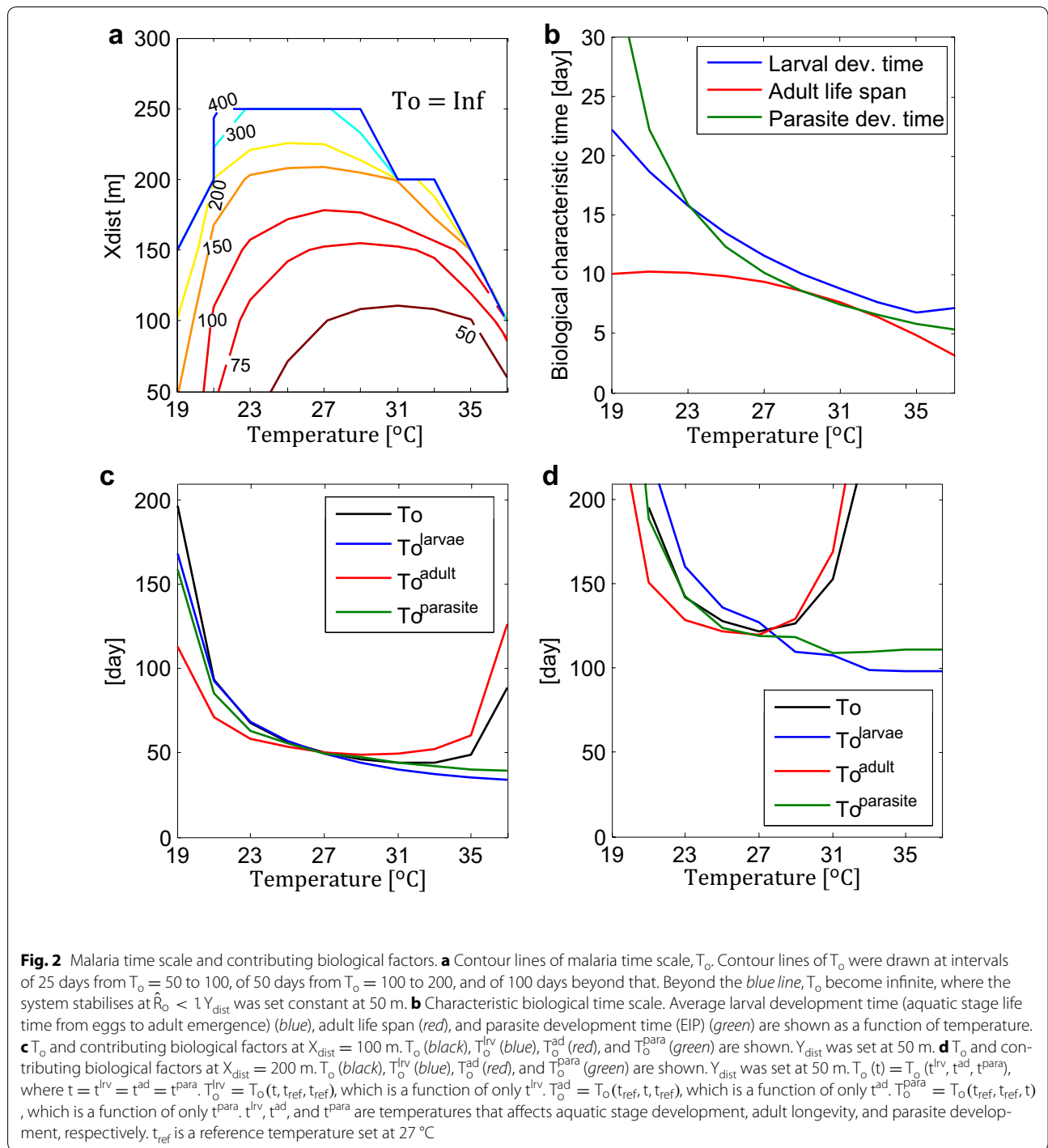
## Results

### Timescale for stable malaria transmission

A time scale for malaria transmission, T<sub>o</sub>, was introduced. T<sub>o</sub> was defined as the time that T<sub>wet</sub> is required to last to reach  $\hat{R}_0 = 1$  under the hydrologically-saturated condition. In other words, if a static water body exists for T<sub>o</sub> days of a year, annual R<sub>o</sub> would reach the critical value of one, and malaria transmission is stable. T<sub>o</sub> is thus independent of hydrological factors, T<sub>wet</sub>, T<sub>on</sub> and T<sub>int</sub>, but is dependent on temperature and spatial factors, X<sub>dist</sub> and Y<sub>dist</sub>. T<sub>o</sub> was obtained for each environmental condition, i.e. for different combinations of temperature, X<sub>dist</sub>, and Y<sub>dist</sub>, by running simulations under the hydrologically-saturated condition.

The contour lines of T<sub>o</sub> are shown in Fig. 2a on the plane of temperature and X<sub>dist</sub> fixing Y<sub>dist</sub> at 50 m (results for different Y<sub>dist</sub> can be found in Additional file 5). Smaller T<sub>o</sub> indicates higher malaria potential, which is always the case for smaller X<sub>dist</sub> values. Above critical values of X<sub>dist</sub>,  $\hat{R}_0$  never reaches one, and hence T<sub>o</sub> can be assumed as infinite. The condition with T<sub>o</sub> = ∞ means that malaria cannot be sustained under this environmental setup even with the existence of a permanent water body.

Responses of T<sub>o</sub> to X<sub>dist</sub> and temperature are non-linear, whose behaviours can be understood by looking at governing biological factors. Figure 2b illustrates the three important biological characteristic times: larval development time (blue), adult life span (red), and parasite development time in mosquito midgut (a.k.a. extrinsic incubation period, EIP) (green). Larval development time spans mosquitoes’ aquatic life from eggs, to larvae, to pupae, and to adult emergence. In this model, the larval development time follows the Depinay’s equation [16], and the adult life span follows Martens’ equation



[17]. The EIP is calculated using Detinova’s equation [18]. Because these biological time scales do not translate linearly to malaria transmission potential, the behaviour of  $T_o$  differs from that of these timescales. In addition, the importance of these biological factors varies not just with temperature, as often asserted in the literature, but also with spatial collocation of pools and houses (Fig. 2c, d).

The contributions of larval development time, adult life span, and EIP were expressed in line with  $T_o$ , as  $T_o^{larv}$ ,  $T_o^{ad}$ , and  $T_o^{para}$ , respectively (for definition, see caption of Fig. 2 and “Methods” section). At small  $X_{dist}$ , malaria transmission is limited by EIP (Fig. 2c); even though infected mosquitoes find hosts, they are not often infectious yet. This effect is stronger at colder temperatures,

when EIP is large. At large  $X_{dist}$ , mosquito life span is the main limiting factor for malaria transmission (Fig. 2d). Mosquitoes have to fly at least  $X_{dist}$  between houses and pools, three times within their life time to transmit malaria. This limitation is especially strong when mosquitoes' life span is short at high temperature.

**Universality in malaria endemic conditions in dimension-less space**

The conditions for  $\hat{R}_0 = 1$  were mapped in a dimension-less space constructed by  $D_1 = T_{wet}/T_o$  and  $D_2 = T_o/T_{int}$ .

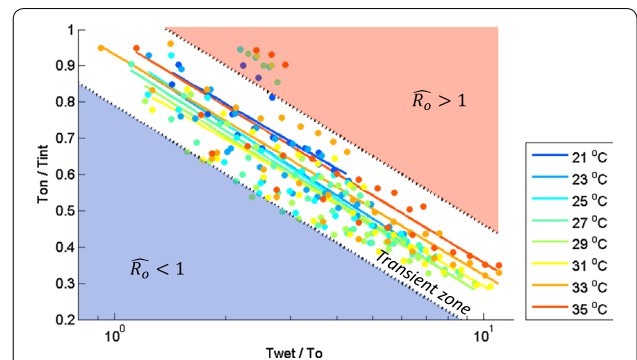
Simulations were conducted with various spatial settings and hydrological conditions at different temperatures. In Fig. 3, the conditions that led to  $\hat{R}_0 = 1$  were presented in the dimension-less space, for all the listed values of  $X_{dist}$ ,  $T_{wet}$ ,  $T_{on}$ , and temperature. The values of  $Y_{dist}$  and  $T_{int}$  were fixed in Fig. 3, but the results for different values of  $Y_{dist}$  and  $T_{int}$  can be found in Additional file 6. The observed points for  $\hat{R}_0 = 1$  (circles) were fitted with natural logarithmic functions (solid lines) for each temperature, which serves as contour lines for  $\hat{R}_0 = 1$ . A logarithmic function fitted with all the presented conditions across temperature values reads  $D_2 = -0.27 \log D_1 + 0.94$  ( $\equiv f_o$ ); the area bounded by the black dotted lines in Fig. 3 shows  $\pm 1.5$  in y-direction from the fitted line,  $f_o$ . The conditions above the contour line for a given temperature have high probability of having malaria endemics, where  $\hat{R}_0 > 1$ , and the conditions that fall below the lines are likely to be malaria-free, where  $\hat{R}_0 < 1$ .

The figure demonstrates that the conditions for  $\hat{R}_0 = 1$  fall in a narrow area (*transient zone*) bounded by  $D_2 = f_o + 1.5$  and  $D_2 = f_o - 1.5$  in the  $D_1$ - $D_2$  space, independent of temperature. Note that Fig. 3 aggregates the results from not only all the temperatures but also all the values of  $X_{dist}$ ,  $T_{wet}$ , and  $T_{on}$  tested. The results for different  $Y_{dist}$  and  $T_{int}$  revealed the same characteristics (Additional file 6). On the other hand, when plotted in the  $T_{wet}$ - $T_{on}$  plane (Additional file 7), the complex and non-linear contributions of  $X_{dist}$  and temperature were apparent. With the introduction of  $D_1$  and  $D_2$ , a universality of the conditions for  $\hat{R}_0 = 1$  was found in this theoretical study.

**Malaria transmission SCORE**

The estimates of malaria transmission potential are presented as *SCORE*, based on the universal relationship found between  $D_1$  and  $D_2$ . A malaria transmission score,  $s(D_1, D_2)$ , was first calculated for a given year in the following way:

$$s(D_1, D_2) = \begin{cases} 1 & \text{if } D_2 > f_o(D_1) + 1.5, \\ 0 & \text{if } D_2 \leq f_o(D_1) + 1.5 \text{ and } D_2 \geq f_o(D_1) - 1.5, \\ -1 & \text{if } D_2 < f_o(D_1) - 1.5. \end{cases}$$



**Fig. 3**  $\hat{R}_0 = 1$  contour lines for different temperatures. Observed points for  $\hat{R}_0 = 1$  (circles) were fitted with natural logarithmic functions (solid lines) on the non-dimensional  $D_1$ - $D_2$  space for each temperature, where  $D_1 = T_{wet}/T_o$  and  $D_2 = T_{on}/T_{int}$ . The x-axis in this figure is shown on a logarithmic scale. Each color represents a different temperature, which applies both to the circles and the lines. No line was drawn for 19 °C because almost no combination of conditions lead to  $\hat{R}_0 > 1$ . A natural logarithmic function fitted to all the observed points of  $\hat{R}_0 = 1$  reads as  $D_2 = -0.27 \log D_1 + 0.94$  ( $\equiv f_o$ ), and the area  $\pm 1.5$  in y-direction from it is denoted as transient zone. Note that the figure contains the results from all the values of  $T_{wet}$ ,  $T_{on}$  and  $X_{dist}$  tested, as well as temperatures

Subsequently, 10-year average of  $s(D_1, D_2)$  from 2001 to 2010 was obtained as *SCORE*.

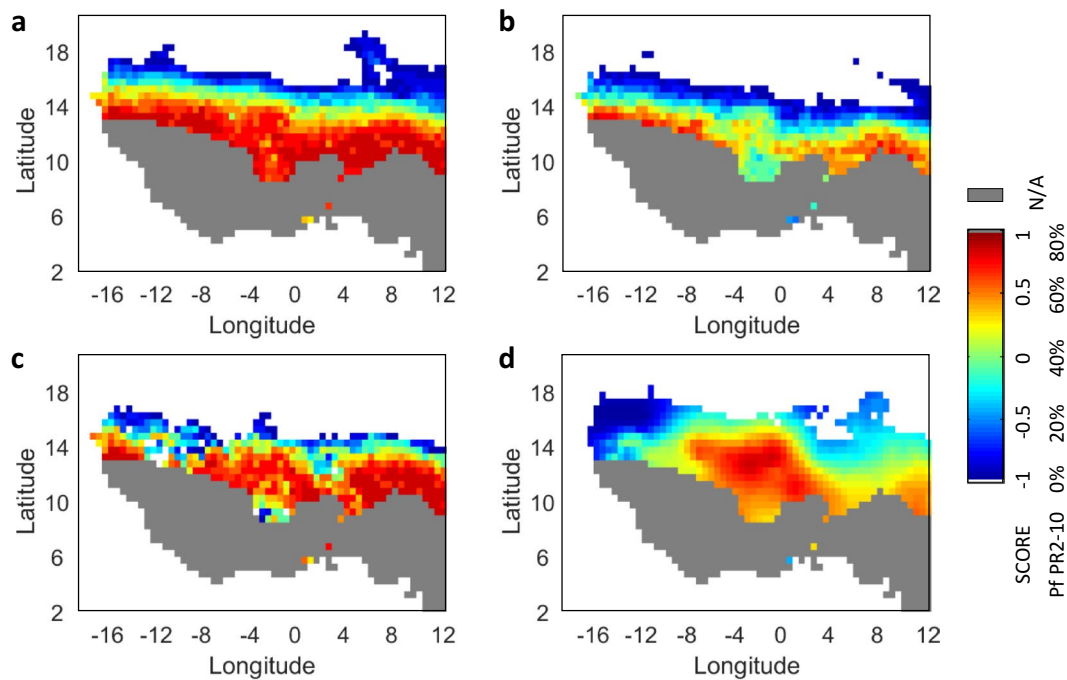
The estimates were made both with and without *PD* information. The estimate without the information of *PD*, non-adjusted-*SCORE*, was obtained by applying a fixed value of  $X_{dist}$ . The estimate adjusted for *PD*, adjusted-*SCORE*, was obtained by adopting  $X_{dist}$  values inferred from population density. Note that any values of  $X_{dist} \geq 300$  results in  $s(D_1, D_2) = -1$ . In both estimates,  $Y_{dist}$  was fixed at 50 m.

**Predictive theory of malaria transmission**

The two dimension-less variables introduced are not only intuitive, but also useful in analysing malaria transmission potential because they reduce the number of parameters involved. Based on the universality found of the conditions for  $\hat{R}_0 = 1$ , the *predictive theory* of malaria transmission is proposed: endemic malaria when  $D_2 > f_o + 1.5$ ; critical condition to sustain malaria with potential epidemics when  $D_2 \leq f_o + 1.5$  and  $D_2 \geq f_o - 1.5$ , in the *transient zone*; malaria phaseout when  $D_2 < f_o - 1.5$ .

**Testing the theory for West Africa**

Based on the predictive theory and observed climatological data, malaria transmission potential for West Africa was estimated as *SCORE*, and compared with observational malaria transmission data (Fig. 4). *SCORE* was calculated both with and without population density information as adjusted- and non-adjusted-*SCORE*,



**Fig. 4** Comparisons of estimated and observed malaria transmission intensity. **a** Non-adjusted-SCORE at  $X_{\text{dist}} = 100$  m. **b** Non-adjusted-SCORE at  $X_{\text{dist}} = 200$  m. **c** Adjusted-SCORE with  $X_{\text{dist}}$  inferred from PD. The estimates of malaria transmission potential, SCORE, are shown in color with minimum and maximum values being  $-1$  and  $1$ . Areas with annual rainfall more than  $1100$  mm are not suitable for the analysis and are shown in gray. **d** Observed malaria intensity. *Plasmodium falciparum* parasite rate in 2- to 10-year-old children estimated for the year 2010 is shown, with minimum and maximum values being  $0$  and  $80\%$ . The figure was adapted from the Malaria Atlas Project (MAP) [19]. Areas with annual rainfall more than  $1100$  mm are also masked in gray

respectively. In the adjusted-SCORE, population density data were used to extract spatial information; higher population density was associated with smaller  $X_{\text{dist}}$ . In the non-adjusted-SCORE, fixed values of  $X_{\text{dist}}$  were applied.

The non-adjusted-SCORE for  $X_{\text{dist}} = 100$  and  $200$  m are shown in Fig. 4a, b, respectively. Figure 4c presents the adjusted-SCORE. The estimates were compared with the observational malaria transmission intensity data in Fig. 4d, which presents the *Plasmodium falciparum* parasite rate in 2- to 10-year-old children obtained from the Malaria Atlas Project (MAP) [19]. Because HYDRE-MATS has been calibrated only for semi-arid parts of West Africa, no estimate was made for wet regions where annual rainfall exceeds  $1100$  mm (shown in gray).

The non-adjusted-SCORE reflects climate suitability of malaria transmission. The non-adjusted-SCOREs (Fig. 4a, b) captured the broad spatial limits for malaria transmission; however, they missed details of regional transmission intensity differences. The adjusted-SCORE, which also includes the spatial information, improved the predictability of malaria transmission potential as compared to the non-adjusted-SCORE, which was estimated using climate factors alone. For example, lower transmission potential is expected around an area  $10^{\circ}\text{E}$  and  $2^{\circ}\text{E}$  than

the non-adjusted-SCORE with  $X_{\text{dist}} = 100$  m (Fig. 4a); higher potential is expected over the area between  $-6^{\circ}\text{E}$  to  $0^{\circ}\text{E}$  and  $12^{\circ}\text{N}$  to  $14^{\circ}\text{N}$  than the non-adjusted-SCORE with  $X_{\text{dist}} = 200$  m (Fig. 4b). These details were captured in the adjusted-SCORE (Fig. 4c, d). The adjusted-SCORE, however, overestimated the current malaria transmission patterns west to  $-12^{\circ}\text{E}$  and east to  $4^{\circ}\text{E}$ .

## Discussion

The malaria transmission potential estimate compared well with contemporary observational data, when  $X_{\text{dist}}$  (hypothetical distance between pools and human habitats) was adjusted for population density (Fig. 4c). Even without adjustment for population density (Fig. 4a), the estimate performed as well as other well-known estimates, such as those of the MARA project [12].

One of the strengths of the estimates is the detailed parametrization of hydrology. Instead of the commonly used parameter of monthly or daily rainfall amount [12, 20], three hydrological variables were employed in this study:  $T_{\text{wet}}$ ,  $T_{\text{on}}$ , and  $T_{\text{int}}$ . The same amount of rainfall, but with different temporal rainfall patterns, results in different malaria transmission potentials [21]. Having a large amount of rainfall at a time followed by subsequent dry



days is different from having continuous but low intensity rainfall for many days. Pool persistence,  $T_{on}$ , and storm inter-arrival time,  $T_{int}$ , account for both rainfall amount and temporal patterns. The information of  $T_{on}$  is rarely available from either ground observations or satellite data sets; however, a previous study [9] made it possible to estimate  $T_{on}/T_{int}$  through the association of the annual rainfall and  $T_{wet}$ . In reality, pools persist over a range of period.  $T_{on}$  represents the average behaviour of pools for a given hydrological condition. Some pools disappear quicker, decreasing the total area of breeding pools. The effect of such heterogeneous pool persistence was implicitly modeled through imposing declining probability of mosquitoes laying eggs and persistence-dependent larval mortality.

The length of the rainy season,  $T_{wet}$ , is also an important determinant of malaria transmission, because it takes a certain time for mosquitoes to increase their populations. The effect of the rainy season length, however, is often demonstrated only implicitly, both in statistical models and in computational models. In the MARA project, Craig et al. [12] argued that malaria transmission requires 3 and 5 months of rain at relatively warm and cold regions of Africa, respectively. The MARA project, however, did not use the rainy season length itself as a model parameter, but used the 3 or 5 months as given “time windows” for calculation. The Malaria Early Warning System (MEWS) is an improved version of MARA [22], and it outputs the number of months in a year during which climate conditions are suitable for malaria transmission. This notion is similar to  $T_{wet}$ ; yet the importance of  $T_{wet}$  is better understood in comparison to  $T_o$ . The close description of hydrological conditions provides the predictive theory with an advantage over other malaria transmission models.

The consideration of the spatial effect made the estimate even more accurate. This study employed the hypothetical spatial setup, and the conceptual relationship between the house-to-pool distance and population density. The hypothetical setting is relevant to villages near water-resources reservoirs, having a long shoreline as a vector breeding pool. In reality, the distribution of houses and pools varies significantly from place to place, and hence the hypothetical setup is not valid for many locations. However, the relationship between the house-to-pool distance and population density is based on an implied relationship and should represent the characteristic relationship of regions on a large scale.

The testing of the predictive theory produced a good malaria estimate over West Africa; however, it also has noteworthy shortcomings. First of all, although the inclusion of population density data improved the quality of the estimation of malaria transmission potential, the

effect of population density on malaria transmission is not straightforward. This analysis assumed that having a large population density is associated with a denser house network, and so a small distance between houses and pools, resulting in large transmission potential. A large population density, on the other hand, may be linked to economic development and to enhanced malaria controls, especially when an analysis is conducted at a village scale [23, 24]. The former effect appears to be stronger than the latter at the scale of the analysis, given that the spatial patterns of malaria transmission intensity were more closely reproduced by positively associating the malaria risk and population density. However, the latter effect, the negative association between the malaria risk and population density, could be dominating in some parts of the region.

Second, the estimate did not account for past and current levels of malaria interventions and human immunity, which may explain the overestimated malaria transmission levels west to  $-12^\circ\text{E}$  and east to  $4^\circ\text{E}$  (Fig. 4c, d) in this study. In Senegal, located around  $-16^\circ\text{E}$  to  $-12^\circ\text{E}$ , malaria parasite rate dropped significantly (by around 30%) between 2000 and 2015 due to extensive malaria control programmes [25, 26]. Around  $4^\circ\text{E}$  to  $12^\circ\text{E}$  and  $12^\circ\text{N}$  to  $14^\circ\text{N}$  is Southern Niger, where Hausa-Fulani people live. The overestimation around this region may be accounted for by the unique low-susceptibility to malaria of the Hausa-Fulani people [27]. In this sense, this estimate addressed the climatic and environmental suitability for malaria transmission in the absence of human controls. The estimate of actual malaria transmission intensity is useful; however, the climatic and environmental suitability for malaria transmission is also vital information for malaria control programmes.

$T_o$  is a biological time scale introduced in this study, which is associated with the time needed to reach  $\hat{R}_o = 1$  under the assumption of the hydrologically-saturated condition. The calculation of  $T_o$  depends on temperature and spatial settings of houses and pools but is independent of hydrological conditions, i.e.,  $T_{wet}$ ,  $T_{on}$ , and  $T_{int}$ .

Malaria transmission is complex, requiring many factors to be considered. The introduction of  $T_o$ , together with the two dimension-less variables,  $D_1 = T_{wet}/T_o$  and  $D_2 = T_{on}/T_{int}$  disclosed a universality in the conditions for stable malaria transmission.  $D_1$  and  $D_2$  not only reduce the number of malaria transmission determinants, but are also intuitive. Readers should, however, be aware that the values of  $T_o$  were obtained for relatively dry regions in West Africa. For other regions with more or less competent vectors, or with other limiting factors in play, the values of  $T_o$  should be evaluated accordingly.

The application of  $T_o$  and the predictive theory is not limited to the estimation of malaria transmission

potential; it also offers guidelines regarding the minimum distances that houses and pools should be apart to prevent malaria, as shown in the following examples. Firstly,  $T_o$  helps in defining the house-to-pool distances needed to prevent malaria transmissions around permanent water bodies, such as water-resources reservoirs. To prevent malaria near a permanent water body, where the condition compares with the hydrologically-saturated condition,  $T_o$  should be infinite.  $T_o$  can be rendered infinite by locating houses away from pools by a certain distance, whose value depends on temperature (Fig. 2). Secondly, malaria-preventive distances between houses and pools can also be inferred for other general conditions from the predictive theory and the  $T_o$ - $X_{\text{dist}}$  relationship [from climatological data,  $T_{\text{on}}/T_{\text{int}}$  can be estimated as described in this paper's "Methods" section. By using the  $T_{\text{on}}/T_{\text{int}}$  with the predictive theory (Fig. 3),  $T_{\text{wet}}/T_o$  associated with  $R_o = 1$  can be found. As  $T_{\text{wet}}$  can be obtained from climatological data,  $T_o$  can be calculated. Finally, applying the  $T_o$ - $X_{\text{dist}}$  relationship (Fig. 2a) for a given temperature, an  $X_{\text{dist}}$  value leading to the  $R_o = 1$  condition can be found]. Malaria-resistant villages would require that houses should be located further than the critical distances, or that pools should be removed within the critical distances from houses.

The critical distances to prevent malaria found in this study also provide guidance for resource allocation. When relocation of settlements or removal of pools is not feasible, conventional interventions, such as distribution of bednets, remain primary malaria control strategy. The critical distances presented in this study shows the distance from pools within which such intervention should be prioritized. This study provides a useful guidance for estimating malaria transmission potential, and for designing malaria resistant villages by controlling house-to-pool distances.  $R_o$  values and critical distances suggested in this study should not be taken as absolute values that fit every condition, but they should be adjusted depending on the details of the local environment. In particular, malaria transmission potential and critical malaria-preventive distances depend on the distribution of houses and the initial number of mosquitoes in the beginning of the rainy seasons. The impact of different house distribution could be a future research topic. The sensitivity to the initial mosquito population is not the focus of this study; how and how many mosquitoes emerge after a long dry period are the questions yet to be clarified. At the limit of no mosquitoes surviving after a long dry season, the potential for malaria transmission would not exist even under the wettest hydrologic condition. The appearance of mosquito population after a long dry period can be explained either by aestivation or migration [28], and imposing the minimum mosquito

population in HYDREMARS conceptually agree with the aestivation hypothesis.

## Conclusion

In fighting malaria, the use of medicine and bed-nets are effective approaches that can be deployed in malaria-endemic regions. Here, this study emphasizes alternative preventive measures and suggested that effective environmental planning may offer a more sustainable approach that addresses the root cause of the problem. In order to reduce the dimension of malaria transmission determinants, and to aid environmental manipulation for malaria prevention, the predictive theory was developed and tested against observations.

The predictive theory, constructed with the two dimensionless parameters, is useful in measuring malaria transmission potential. The two dimensionless parameters contain information of many hydroclimatological and environmental factors. The comparison between the malaria estimate from the predictive theory and the observational data illustrated the importance of the spatial collocation of vector breeding pools and human habitats.

The predictive theory can provide guidelines on how to plan the layout of human habitats in order to prevent endemic malaria. Malaria-resistant villages can be designed by locating houses further than critical distances away from breeding pools or by removing pools within a critical distance from houses; the critical distance is described in the context of local climatology and hydrology.

## Additional files

**Additional file 1.** Malaria intensity and hydroclimatological conditions in West Africa.

**Additional file 2.** Cumulative rainfall and  $T_{\text{wet}}$  for twelve sites in West Africa.

**Additional file 3.** Population density and inferred  $X_{\text{dist}}$  in West Africa.

**Additional file 4.** Comparison between Tontot  $T_{\text{wet}}$  and annual-rainfall/ $T_{\text{wet}}$ .

**Additional file 5.** Contour lines of  $T_o$  at different  $Y_{\text{dist}}$  values.

**Additional file 6.** Conditions for  $\hat{R}_o = 1$  and the sensitivity to  $X_{\text{dist}}$ ,  $Y_{\text{dist}}$  and  $T_{\text{int}}$ .

**Additional file 7.**  $\hat{R}_o = 1$  contour lines on the plane of  $T_{\text{on}}$  and  $T_{\text{wet}}$ .

## Abbreviations

HYDREMATS: hydrology, entomology, and malaria transmission simulator;  $R_o$ : basic reproduction rate;  $\hat{R}_o$ : long-term average of  $R_o$  calculated in HYDREMATS.  $\hat{R}_o = \left( \sum_{i=1}^{N_{\text{inf}}} R_o^i \right) / N_{\text{inf}}$ ;  $P_o$ : the number of infections generated from the  $i$ -th infectious person, assuming that the population is fully susceptible;  $N_{\text{inf}}$ : the number of infectious people simulated over the simulation period of four years;  $X_{\text{dist}}$ : distance between vector breeding pools and human habitats (horizontal distance) (m);  $Y_{\text{dist}}$ : distance between human habitats (longitudinal distance) (m);  $T_{\text{wet}}$ : length of the rainy season (days);  $T_{\text{int}}$ : storm inter-arrival time (days);  $T_{\text{on}}$ : pool persistence (days);  $T_o$ : malaria time scale. The length of the rainy season required to reach  $\hat{R}_o = 1$  under the hydrologically-saturated condition at a given temperature and a given spatial condition.

$T_o(t) = T_o(t^{lv}, t^{ad}, t^{para})$ , where this study assumes  $t = t^{lv} = t^{ad} = t^{para}$  (days);  $t$ : temperature ( $^{\circ}\text{C}$ );  $t^{lv}$ : temperature that influences aquatic-stage development ( $^{\circ}\text{C}$ );  $t^{ad}$ : temperature that influences adult life span ( $^{\circ}\text{C}$ );  $t^{para}$ : temperature that influences parasite development ( $^{\circ}\text{C}$ );  $t_{ref}$ : reference temperature set at  $27^{\circ}\text{C}$ ;  $T_o^{lv}(t) = T_o(t, t_{ref}, t_{ref})$ ;  $T_o^{ad}(t) = T_o(t_{ref}, t, t_{ref})$ ;  $T_o^{para}(t) = T_o(t_{ref}, t_{ref}, t)$ ;  $D_1 = T_{wet}/T_o$ ;  $D_2$ : dimensionless parameter.  $D_2 = T_{on}/T_{ini}$ ;  $PD$ : population density (per  $\text{km}^2$ );  $PD_{imp}$ : implied population density (per  $\text{km}^2$ ).  $PD_{imp} = 5 \times X_{dist}^{-2} \times \alpha^{-1} \times 10^{-6}$ , with  $\alpha$  being set at 15;  $s(D_1, D_2)$ : malaria transmission score calculated for a year;  $SCORE$ : malaria transmission estimate. Ten-year average of  $s(D_1, D_2)$  from 2001 to 2010;  $EIP$ : extrinsic incubation period.

#### Authors' contributions

NE and EABE conceived and designed the study. EABE supervised the research. NE performed the computational experiments, and wrote the manuscript. NE and EABE edited and approved the final manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

This work was supported by the U.S. National Science Foundation and by the Cooperative Agreement between the Masdar Institute of Science and Technology (Masdar Institute), Abu Dhabi, UAE and the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA.

#### Competing interests

The authors declare that they have no competing interests.

#### Availability of data and materials

All the data used in this study are available upon request.

#### Funding

This work was funded by the U.S. National Science Foundation grant EAR-0946280 and by the Cooperative Agreement between the Masdar Institute of Science and Technology (Masdar Institute), Abu Dhabi, UAE and the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA Reference 02/MI/MI/CP/11/07633/GEN/G/00.

Received: 6 September 2016 Accepted: 23 November 2016

Published online: 01 December 2016

#### References

- Gu W, Regens JL, Beier JC, Novak RJ. Source reduction of mosquito larval habitats has unexpected consequences on malaria transmission. *Proc Natl Acad Sci USA*. 2006;103:17560–3.
- Bombliès A. Agent-based modeling of malaria vectors: the importance of spatial simulation. *Parasit Vectors*. 2014;7:308.
- Minakawa N, Seda P, Yan G. Influence of host and larval habitat distribution on the abundance of African malaria vectors in Western Kenya. *Am J Trop Med Hyg*. 2002;67:32–8.
- van der Hoek W, Konradsen F, Amerasinghe PH, Perera D, Piyaratne M, Amerasinghe FP. Towards a risk map of malaria for Sri Lanka: the importance of house location relative to vector breeding sites. *Int J Epidemiol*. 2003;32:280–5.
- Cano J, Descalzo MA, Moreno M, Chen Z, Nzambo S, Bobuakasi L, et al. Spatial variability in the density, distribution and vectorial capacity of anopheline species in a high transmission village (Equatorial Guinea). *Malar J*. 2006;5:21.
- Zhou G, Munga S, Minakawa N, Githeko AK, Yan G. Spatial relationship between adult malaria vector abundance and environmental factors in Western Kenya highlands. *Am J Trop Med Hyg*. 2007;77:29–35.
- Bombliès A, Duchemin JB, Eltahir EAB. Hydrology of malaria: model development and application to a Sahelian village. *Water Resour Res*. 2008;44:W12445. doi:10.1029/2008WR006917.
- Bombliès A, Duchemin JB, Eltahir EAB. A mechanistic approach for accurate simulation of village scale malaria transmission. *Malar J*. 2009;8:223.
- Yamana TK, Bombliès A, Eltahir EAB. Climate change unlikely to increase malaria burden in West Africa. *Nat Clim Chang*. 2016;6:1009–13.
- Macdonald G. The measurement of malaria transmission. *Proc R Soc Med*. 1955;48:295–302.
- WHO. World malaria report 2015. Geneva: World Health Organization; 2015.
- Craig M, Le Sueur D, Snow B. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today*. 1999;15:105–11.
- Kleinschmidt I, Omumbo J, Briet O, Van De Giesen N, Sogoba N, Mensah NK, et al. An empirical malaria distribution map for West Africa. *Trop Med Int Health*. 2001;6:779–86.
- Mitchell TD, Jones PD. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol*. 2005;25:693–712.
- Columbia University. Gridded Population of the World Version 3 (GPWv3). 2005.
- Depinay JMO, Mbogo CM, Killeen G, Knols B, Beier J, Carlson J, et al. A simulation model of African Anopheles ecology and population dynamics for the analysis of malaria transmission. *Malar J*. 2004;3:29.
- Martens P, Gellert GA. Health and climate change: modelling the impacts of global warming and ozone depletion. *Nature*. 1998;393:176.
- Detinova TS. Age-grouping methods in Diptera of medical importance with special reference to some vectors of malaria. *Monograph Ser World Health Organ*. 1962;47:13–191.
- Hay SI, Guerra CA, Gething PW, Patil A, Tatem AJ, Noor AM, et al. A world malaria map: *Plasmodium falciparum* endemicity in 2007. *PLoS Med*. 2009;6:0286–302.
- Erment 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.
- Bombliès A. Modeling the role of rainfall patterns in seasonal malaria transmission. *Clim Change*. 2012;112:673–85.
- Grover-Kopec EK, Blumenthal MB, Ceccato P, Dinku T, Omumbo JA, Connor SJ. Web-based climate information resources for malaria control in Africa. *Malar J*. 2006;5:38.
- Hay SI, Guerra CA, Tatem AJ, Atkinson PM, Snow RW. Urbanization, malaria transmission and disease burden in Africa. *Nat Rev Microbiol*. 2005;3:81–90.
- Tatem AJ, Guerra CA, Kabaria CW, Noor AM, Hay SI. Human population, urban settlement patterns and their impact on *Plasmodium falciparum* malaria endemicity. *Malar J*. 2008;7:218.
- Roll back malaria: focus on senegal. Progress & impact series, Number 4. Technical report, Geneva: World Health Organization; 2010.
- Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature*. 2015;526:207–11.
- Torcia MG, Santarasci V, Cosmi L, Clemente A, Maggi L, Mangano VD, et al. Functional deficit of T regulatory cells in Fulani, an ethnic group with low susceptibility to *Plasmodium falciparum* malaria. *Proc Natl Acad Sci USA*. 2008;105:646–51.
- Lehmann T, Dao A, Yaro AS, Adamou A, Kassogue Y, Diallo M, et al. Aestivation of the African malaria mosquito, *Anopheles gambiae* in the Sahel. *Am J Trop Med Hyg*. 2010;83:601–6.