



## An ecohydrological model of malaria outbreaks

E. Montosi<sup>1</sup>, S. Manzoni<sup>2,3</sup>, A. Porporato<sup>2,3</sup>, and A. Montanari<sup>1</sup>

<sup>1</sup>Faculty of Engineering, Bologna University, Bologna, Italy

<sup>2</sup>Nicholas School of the Environment, Duke University, Durham NC, USA

<sup>3</sup>Department of Civil and Environmental Engineering, Duke University, Durham NC, USA

Correspondence to: E. Montosi (elena.montosi@unibo.it)

Received: 22 February 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 5 March 2012

Revised: 2 July 2012 – Accepted: 13 July 2012 – Published: 16 August 2012

**Abstract.** Malaria is a geographically widespread infectious disease that is well known to be affected by climate variability at both seasonal and interannual timescales. In an effort to identify climatic factors that impact malaria dynamics, there has been considerable research focused on the development of appropriate disease models for malaria transmission driven by climatic time series. These analyses have focused largely on variation in temperature and rainfall as direct climatic drivers of malaria dynamics. Here, we further these efforts by considering additionally the role that soil water content may play in driving malaria incidence. Specifically, we hypothesize that hydro-climatic variability should be an important factor in controlling the availability of mosquito habitats, thereby governing mosquito growth rates. To test this hypothesis, we reduce a nonlinear ecohydrological model to a simple linear model through a series of consecutive assumptions and apply this model to malaria incidence data from three South African provinces. Despite the assumptions made in the reduction of the model, we show that soil water content can account for a significant portion of malaria's case variability beyond its seasonal patterns, whereas neither temperature nor rainfall alone can do so. Future work should therefore consider soil water content as a simple and computable variable for incorporation into climate-driven disease models of malaria and other vector-borne infectious diseases.

### 1 Introduction

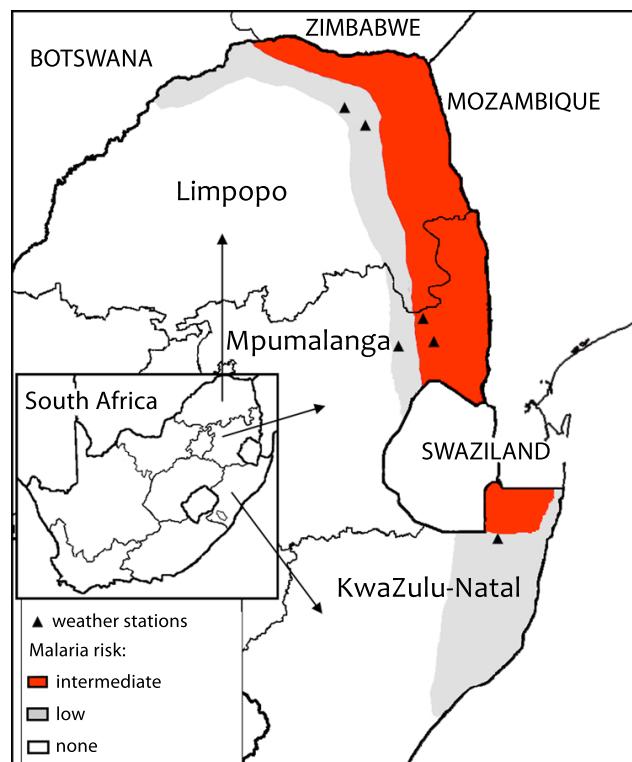
The World Health Organization estimates that 250 million clinical episodes of malaria occur annually, resulting in at least one million disease-associated deaths (World Health Organization, 2008). Malaria incidence is especially high in

developing countries, where it is a leading cause of morbidity and mortality, in particular among children and pregnant women. Since the pioneering work by Ross (1910) and MacDonald (1957), progress in understanding malaria dynamics has been made through the development of mathematical models and their statistical inference with incidence data (e.g. Bailey, 1982; Hay et al., 2002; Depinay et al., 2004; Zhou et al., 2004; Pascual et al., 2008; Chaves et al., 2012; Bhadra, 2011; Laneri et al., 2010, among others). A subset of these models has considered the role that external forcing plays in generating patterns of seasonal and interannual case variability. Despite these advances, early warning systems for malaria have still only limited ability (and thereby efficacy) to predict outbreaks, and the factors contributing to malaria case variability still require more thorough investigation (Pascual et al., 2008; Craig et al., 2004a,b).

Two climatic variables that have long been known to influence malaria's seasonal and interannual dynamics are temperature and rainfall. Temperature is known to affect the development time of mosquito larvae, the probability of mosquito survival, and the development time of the malaria parasite *Plasmodium falciparum* in infected mosquitoes (Bayoh and Lindsay, 2003; Hoshen and Morse, 2004). Rainfall is hypothesized to affect malaria incidence through the creation of mosquito breeding sites during wet periods and the reduction of their availability during droughts (Patz et al., 1998). Although a moderate level of rainfall appears to have a positive effect on mosquito recruitment, intense rainfall events may destroy mosquito habitats and thereby reduce malaria incidence shortly following their occurrence (Briet et al., 2008).

Due to the multiple effects of temperature and rainfall on the malaria parasite and its mosquito vector, previous work linking malaria incidence to climate forcing has frequently focused on the direct impacts of temperature and rainfall on the risk of malaria (Craig et al., 2004a; Lindsay et al., 2000). Several studies have found a correlation between malaria incidence and either minimum, mean or maximum temperatures (Hay et al., 2002; Zhou et al., 2004; Devi and Jauhar, 2007; Gomez-Elipe et al., 2007), while others have instead considered the effects of diurnal temperature fluctuations (Paamajans et al., 2009) and indices of the upper atmospheric circulation combined with sea-surface temperatures (Jury and Kanemba, 2007) on the disease. At regional spatial scales, malaria incidence has been correlated with rainfall amounts (Pascual et al., 2008; Craig et al., 1999; Hay et al., 2001), interestingly with time lags of several months, consistent with the presence of a physical “buffering” mechanism that dampens rainfall fluctuations. These latter findings hint at the fundamental role that soil water content dynamics may play in malaria outbreaks. Recent work has moved in this direction, with the role of surface hydrology being more fully recognized and explicitly included in mechanistic models. Some of these models have focused on water level fluctuations in small reservoirs and the effect of these fluctuations on malaria prevalence (Porphyre et al., 2005; Shaman et al., 2002). In an effort to predict the number of malaria cases, several statistical models have also considered vegetation density (which depends on soil water availability) and distances between water bodies and human populations (Gomez-Elipe et al., 2007; Kleinschmidt et al., 2001). Finally, detailed models linking hydrology with entomology have recently been proposed and tested for a semi-arid region (Bombali et al., 2008; Yamana et al., 2011). The role of hydrologic processes in the dynamics of other vector-borne infectious diseases is also starting to become more fully recognized (Bertuzzo et al., 2008).

Motivated by these previous efforts, here we test the hypothesis that soil water content is an important driver of malaria dynamic. To test this hypothesis, we develop a minimalist ecohydrological model of coupled surface water malaria dynamics and use it to describe malaria incidence in three South African provinces (Fig. 1). Although the ecohydrological model we derive makes a series of simplifying assumptions, our results show that variability in soil water content is significantly correlated with variability in malaria incidence, whereas neither rainfall nor temperature alone shows this correlation (beyond their seasonal associations). Future work should therefore focus on soil water content as a simple and computable environmental variable that can be incorporated into more mechanistic models of malaria transmission that include internal feedbacks.



**Fig. 1.** Map of the investigated provinces in South Africa, colored by level of malaria risk. The locations of the weather stations for each province are marked with black triangles (see: [http://www.malaria.org.za/Malaria\\_Risk/Risk\\_Maps/risk\\_maps.html](http://www.malaria.org.za/Malaria_Risk/Risk_Maps/risk_maps.html)).

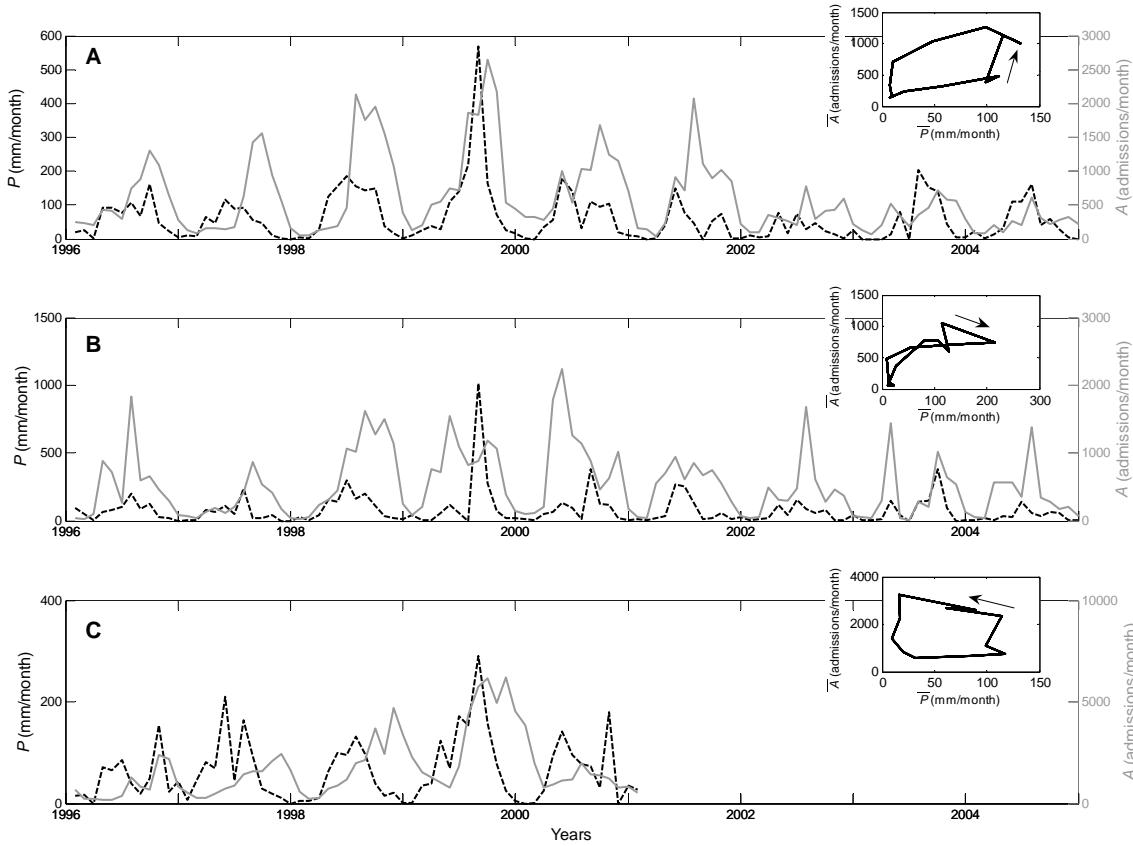
## 2 Materials and methods

### 2.1 Malaria data

Monthly malaria incidence data over the period July 1996–March 2007 were obtained from the South African Department of Health (<http://www.health-e.org.za/resources/statistics.php>, last access: February 2012) for three South African provinces: Limpopo, Mpumalanga and KwaZulu-Natal (Figs. 2–3). Implementation of malaria control strategies, mainly due to the Lumbobo Spatial Development Initiative, resulted in a steady decrease in the number of reported malaria cases in the provinces of Mpumalanga and Limpopo starting in 2005 and in the province of KwaZulu-Natal starting in 2002. To neglect this transient and consider a time series driven only by climate variability, we therefore limited our analysis to incidence data before June 2005 for Mpumalanga and Limpopo and before July 2001 for KwaZulu-Natal (Figs. 2–3).

### 2.2 Climate data

The meteorological data consisted of daily rainfall and daily minimum and maximum temperature records collected from weather stations managed by the South African Weather



**Fig. 2.** Monthly malaria cases (solid grey line) and rainfall levels (dashed black line) for (A) Mpumalanga, (B) Limpopo, and (C) KwaZulu-Natal provinces. Insets describe seasonal averages of malaria cases and rainfall rates (time proceeds according to the arrows).

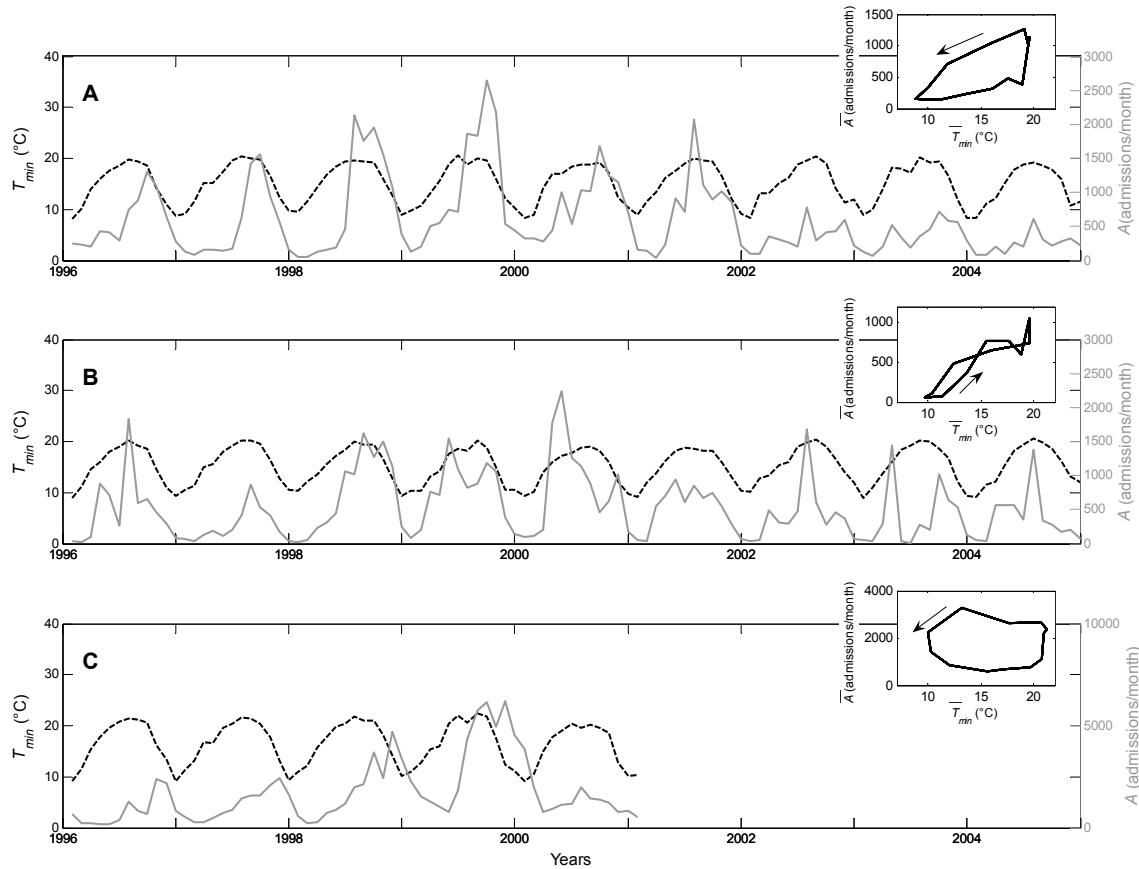
Service. We restricted our analysis to datasets with less than 20 % missing meteorological data that were obtained from stations in areas with high population densities and intermediate to high malaria risk (Fig. 1). For each province, spatially integrated daily time series were first obtained by averaging across the selected weather stations. Finally, to have the same temporal resolution for the malaria and climate datasets, we computed monthly meteorological data from these spatially integrated daily time series (Figs. 2–3). A preliminary inspection of monthly malaria data alongside the meteorological data highlighted different responses to temperature and rainfall. We observed that months characterized by one or more days with daily maximum temperatures above 39 °C tended to be followed by a decrease in the number of malaria cases, most likely resulting from the effect of heat-stress on mosquitoes (Craig et al., 1999). Periods of moderate to high precipitation had a positive and delayed effect on malaria incidence, presumably because of the increased availability of mosquito breeding sites (Wyse et al., 2007). The few instances of anomalously high daily precipitation were followed by a rapid decrease in malaria incidence, presumably a result of mosquito habitat destruction (Briet et al., 2008). Moreover, Limpopo province shows

a shorter delay in malaria occurrence with respect to the provinces of Mpumalanga and KwaZulu-Natal (insets of Figs. 2–3).

### 2.3 Model description

We assessed the association between climatic drivers and malaria cases in three ways: (i) standard linear regression between climate anomalies and malaria case anomalies, at time lags of zero and one month; (ii) an ecohydrological model of malaria dynamics; and (iii) a transfer function model accounting for delayed climate effects on malaria dynamics. The first approach (i) provided a baseline to quantify the improvement of models (ii) and (iii). We chose not to consider the monthly maximum temperature data as an explanatory variable for malaria cases, because it did not have significant explanatory power in preliminary analysis. Instead, we focused our analysis on monthly data of both minimum temperature and precipitation anomalies.

The full ecohydrological model explicitly coupled a model of malaria transmission dynamics with a hydrologic model describing soil water content. Through a series of assumptions, detailed below, this full ecohydrological model was reduced to a minimal, linear model that describes the expected



**Fig. 3.** Monthly malaria cases (solid gray line) and minimum temperature (dashed black line) for (A) Mpumalanga, (B) Limpopo, and (C) KwaZulu-Natal provinces. Insets describe seasonal averages of malaria cases and minimum temperature (time proceeds according to the arrows).

relationship between precipitation levels in previous months and current malaria incidence.

The full model of malaria transmission is given by the following equations, where  $M$  is the population size of the *Anopheles* mosquito vector,  $M_I$  is the population size of infected mosquitoes,  $H_S$  is the population size of susceptible individuals, and  $H_I$  is the population size of infected individuals (Ross, 1910; Porphyre et al., 2005; Kermack and McKendrick, 1927; Smith and McKenzie, 2004; McCallum et al., 2001):

$$\frac{dM(t)}{dt} = \Psi_0[w(t), T(t)]M(t - \tau_M) - \delta M(t) \quad (1)$$

$$\frac{dM_I(t)}{dt} = \alpha H_I(t - \tau_I)M(t - \tau_I) - \delta M_I(t) \quad (2)$$

$$\frac{dH_S(t)}{dt} = \mu H_{TOT} + \gamma [H_{TOT} - H_S(t) - H_I(t)] + \eta_0 M_I(t) \frac{H_S(t)}{H_{TOT}} - \mu H_S(t) \quad (3)$$

$$\frac{dH_I(t)}{dt} = \eta_0 M_I(t) \frac{H_S(t)}{H_{TOT}} - (\nu + \mu)H_I(t) \quad (4)$$

The number of individuals recovered and immune to malaria infection at time  $t$  is not explicitly modeled, but given by  $H_{TOT} - H_S - H_I$  where  $H_{TOT}$  is the constant total host population size.

The dynamics of the mosquito population, given by Eq. (1), are modeled by a delay differential equation where  $\Psi_0$  is the mosquito growth rate (a function of soil moisture,  $w$ , and temperature,  $T$ ),  $\frac{1}{\delta}$  ( $\sim 10$  days) is the average lifespan of a mosquito, and  $\tau_M$  ( $\sim 10$  days) is the time delay between oviposition and mosquito emergence (Chitnis et al., 2008). The number of infected mosquitoes, governed by Eq. (2), increases through feeding on infected hosts at a rate  $\alpha$  (whose unit is  $(\text{months} \cdot \text{number of infected individuals})^{-1}$ ) as we assumed a temporal resolution of 1 month), with  $\tau_I$  being a time delay (in months) representing the incubation period for malaria parasites, and decreasing with background mortality at a rate  $\delta$ . The number of susceptible hosts increases with births ( $\mu H_{TOT}$ ) and loss of immunity ( $\gamma (H_{TOT} - H_S - H_I)$ , where  $\gamma$  is dimensionless) and decreases with background mortality ( $\mu H_S$ ) and malaria transmission via infected mosquitoes ( $\frac{\eta_0 M_I H_S}{H_{TOT}}$ , Eq. 3). Finally, the number of infected hosts increases with transmission and

decreases with background mortality and recovery from infection ( $vH_I$ ).  $\mu$ ,  $\eta_0$  and  $v$  are all quantities expressed in terms of (months) $^{-1}$ .

Soil water content,  $w$ , is described by a soil-moisture and surface-water balance equation (Wyse et al., 2007; Porporato et al., 2004):

$$\frac{dw(t)}{dt} = P(t) - mw(t) \quad (5)$$

where  $P(t)$  is rainfall at time  $t$ , and  $mw(t)$  is a linearized soil water loss function accounting for plant transpiration, surface evaporation, and deep infiltration, where  $m$  is a lumped soil water loss rate (with units of months $^{-1}$ ). The use of this function is justified by the large spatial scale (Fig. 1) and relatively low temporal resolution ( $\sim 1$  month) we are considering. In contrast to classical models of soil water content, we chose a form for Eq. (5) that does not saturate, so as to account for both soil and surface water storage. By surface water storage, we mean primarily soil moisture, but also seasonal ponds and rivers that might be breeding sites for mosquitoes. In the following, the assumptions employed to reduce the full ecohydrological model (i.e., Eqs. 1–4) for malaria dynamics and Eq. (5) for surface hydrology to the linear model are described.

- Assumption 1: Because the number of infected is much smaller than the total population in the considered areas, as a first approximation we can assume that the susceptible fraction of the host population,  $\frac{H_S}{H_{TOT}}$ , is relatively stable through time. With this assumption, the dynamics in the number of infected individuals are simplified to

$$\frac{dH_I(t)}{dt} = \eta M_I(t) - (v + \mu) H_I(t), \quad (6)$$

where  $\eta = \eta_0 \frac{H_S}{H_{TOT}}$  is expressed in (months) $^{-1}$ .

- Assumption 2: Consistent with the available malaria data, we focus on the monthly time scale and we assume that both the time delay between mosquito oviposition and emergence, and the incubation period, can be neglected. These assumptions yield

$$\frac{dM(t)}{dt} = \Psi_0[w(t), T(t)]M(t) - \delta M(t), \quad (7)$$

$$\frac{dM_I(t)}{dt} = \alpha H_I(t)M(t) - \delta M_I(t). \quad (8)$$

- Assumption 3a: when climate is the only limiting factor to mosquito emergence, the effect of  $M(t)$  on the growth rate can be neglected, effectively approximating the exponential growth resulting from Eq. (7) with a linear growth with moisture- and temperature-controlled rate, which can be expressed as

$$\frac{dM(t)}{dt} = \Psi[w(t), T(t)] - \delta M(t). \quad (9)$$

- Assumption 3b: alternatively, in the case that the mosquito population exhibits logistic growth, with climate determining its carrying capacity, we can re-write Eq. (7) as

$$\frac{dM(t)}{dt} = r \left[ 1 - \frac{M(t)}{K[w(t), T(t)]} \right] M(t). \quad (10)$$

where  $r$  is the maximum growth rate and  $K$  the carrying capacity of the mosquito population.

- Assumption 4: we assume that the total mosquito density  $M$  is approximately in equilibrium at the monthly time scale  $\frac{dM(t)}{dt} = 0$ , on the grounds that both the soil water content (with a mean transit time of 1–3 months) and the size of the infected population (transit time  $\sim 10$  months) fluctuate at longer time scales than mosquito density (transit time  $\ll 1$  month) (Chitnis et al., 2008). With Assumption 3a in place, the equilibrium mosquito population,  $\widehat{M}$ , is found as

$$\widehat{M}(t) = \frac{\Psi[w(t), T(t)]}{\delta}. \quad (11)$$

Alternatively, with Assumption 3b in place,

$$\widehat{M}(t) = K[w(t), T(t)]. \quad (12)$$

- Assumption 5: since we are considering a large geographic area where spatial heterogeneity likely weakens the nonlinear nature of the interaction between infected humans and mosquitoes (as shown for other systems; see e.g. Katul et al., 2007b), we also assume that the density of infected mosquitoes is proportional to the total mosquito density:  $\widehat{M}_I(t) \cong \gamma \widehat{M}(t)$ . With these assumptions,  $\widehat{M}_I(t)$  is given by

$$\widehat{M}_I(t) = \gamma \frac{\Psi[w(t), T(t)]}{\delta}, \quad (13)$$

and

$$\widehat{M}_I(t) = \gamma K[w(t), T(t)] \quad (14)$$

from Eqs. (11) and (12), respectively. Equations (13) and (14), together with Eq. (5) and (6), lead to a system of two coupled linear equations driven by rainfall  $P$ , water content  $w$ , and temperature  $T$ :

$$\frac{dw(t)}{dt} = P(t) - mw(t) \quad (15)$$

$$\frac{dH_I(t)}{dt} = \eta \gamma \varepsilon[w(t), T(t)] - (v + \mu) H_I(t) \quad (16)$$

where  $\varepsilon[w(t), T(t)]$  is given by either  $\Psi[w(t), T(t)]$  or  $K[w(t), T(t)]$ . Equation (16) predicts a rapid growth in the number of infected individuals when climate is favorable for mosquito emergence (large  $\varepsilon$ ) or exponential decay in the number of infected individuals when

climate is unfavorable ( $\varepsilon \approx 0$ ). At the time scale of a month, malaria incidence  $A$  is provided by the first term of Eq. (16),  $A(t) = \eta\gamma\varepsilon[w(t), T(t)]$ .

- Assumption 6: the effects of environmental conditions on mosquito growth rate are assumed to be linear, i.e.,  $A(t) = a + bw(t) + cT(t)$ .

Under these assumptions, the minimalist model given by Eqs. (15) and (16) is linear in both the state variables and the climatic factors, allowing us to remove the seasonal component of these dynamics, leaving the dynamics themselves in terms of anomalies from the seasonal averages. This solution is appropriate because the disease and climate time series are both strongly seasonal, therefore hindering the identification of how climate variability drives disease dynamics on an interannual time scale (Hay et al., 2000; Briet et al., 2008). Seasonal averages for malaria case incidence and for precipitation levels are shown in the insets of Figs. 2–3 for each of the three provinces.

Indicating the monthly anomalies with prime signs and the seasonal monthly averages with an overbar, we have  $P(t) = \overline{P} + P'(t)$ ,  $A(t) = \overline{A} + A'(t)$  and  $H(t) = \overline{H} + H'(t)$ . Because the monthly temperature anomalies are negligible in comparison to the anomalies in the hydrologic variables (the maximum standard deviation of the anomalies of temperature is  $0.84^{\circ}\text{C}$ , recorded for the province of Limpopo, and the maximum standard deviation of the temperature seasonal average is  $4.19^{\circ}\text{C}$  in the province of KwaZulu-Natal), the effect of temperature anomalies on  $A$  can be neglected. Accordingly, Eqs. (15) and (16) can be re-written as

$$\frac{dw'(t)}{dt} = P'(t) - mw'(t) \quad (17)$$

$$\frac{dH'_I(t)}{dt} = A'(t) - (\nu + \mu)H'_I(t) \quad (18)$$

with

$$A'(t) = bw'(t) \quad (19)$$

Given monthly malaria case incidence anomalies  $A'$  and monthly precipitation anomalies  $P'$ , we sought to fit parameters  $m$  and  $b$ . To this end, Eq. (17) was integrated numerically with a given value of  $m$  through a finite difference approach to yield soil water content levels over time,  $w'(t)$ , and parameter  $b$  was then estimated through linear regression (Eq. 19). Our expectation is that  $b$  assumes positive values (anomalously high water soil content should have a positive effect on mosquito growth rate) and that the rate of soil water loss  $m$  is on the order of  $0.3\text{--}0.5\text{ months}^{-1}$  (Katul et al., 2007a). To yield predictions of monthly malaria cases (Fig. 4), malaria case anomalies  $A'$  were first estimated (given the best fit parameter values  $b$  and  $m$ ) and then added to the seasonal averages in malaria case numbers  $\overline{A}$ .

By looking at the monthly series of malaria and precipitation, we realized that there is a high seasonality both in

the data per se and in their standard deviations. Therefore, we developed the aforementioned ecohydrological model to describe these high seasonal dynamics mainly looking at the deseasonalized data by means of a physically based approach. However, it is interesting to assess the impact of the seasonality of the standard deviation of the data as well. To this aim, we developed a stochastic model, our third approach, to study the association between climatic drivers and malaria cases. Specifically, we used a transfer function model, i.e. a linear filter of the climatic variables to predict malaria cases. The model was conceived by mimicking the relationship suggested by the ecohydrological model (Eqs. 17 and 19 above).

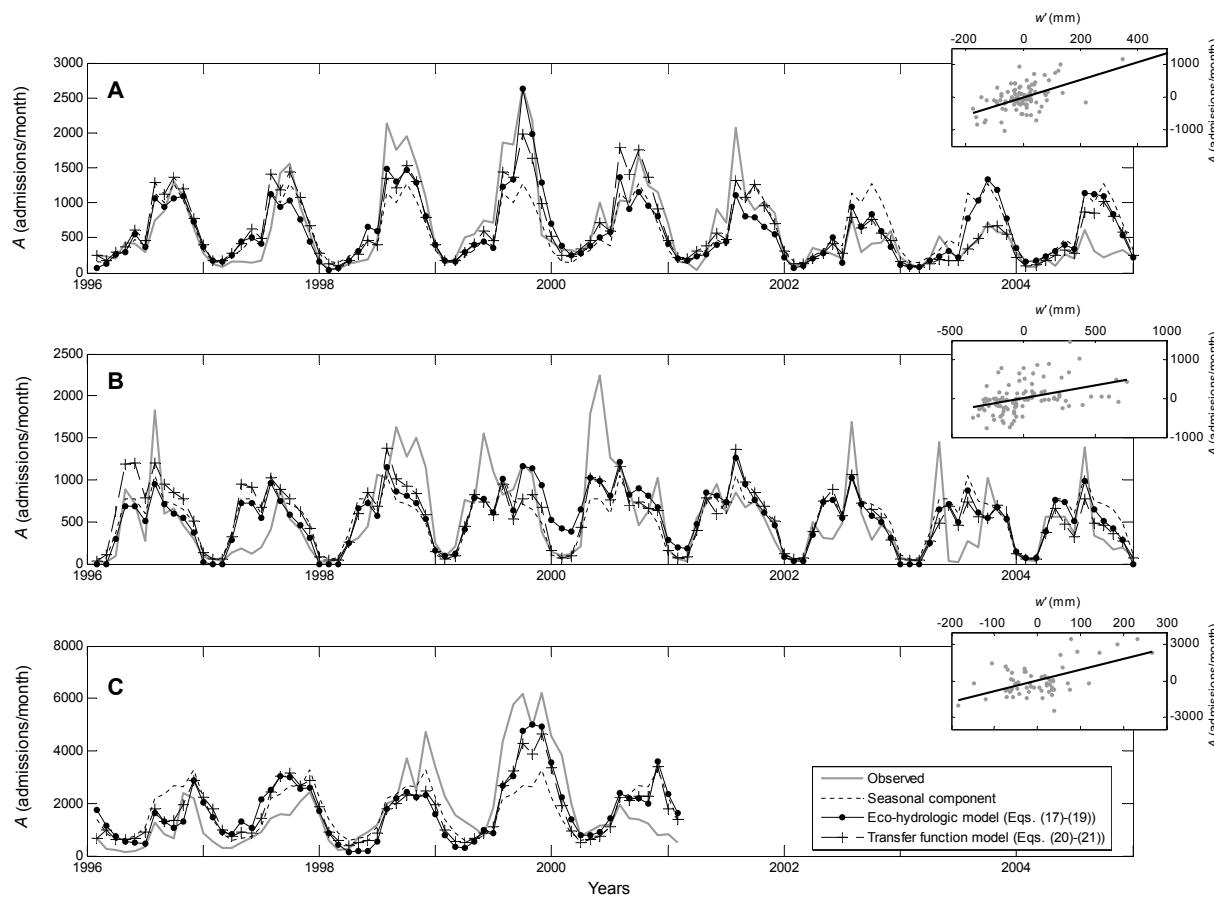
Given that the time series of both monthly malaria cases ( $A$ ) and monthly rainfall rates ( $P$ ) showed a seasonal pattern in their standard deviations, making the time series heteroscedastic and limiting robust parameter estimation, we adopt a transfer function model of the following form:

$$w''_{i+1} = P''_i + nw''_i \quad (20)$$

$$A''_i = dw''_i \quad (21)$$

where double prime signs indicate that the variables  $P'$ ,  $w'$ , and  $A'$  are normalized with respect to their seasonal standard deviation. Equation (20) describes soil moisture anomaly at month  $i$  as rainfall anomaly plus a portion of the previous month anomaly, while Eq. (21) is analogous to Eq. (19). Parameters  $d$  and  $n$  are conceptually similar to  $b$  and  $m$ . Again, we expect that  $d$  is positive. To obtain predictions of monthly malaria cases using this transfer function model (Fig. 4), malaria case anomalies  $A''$  were estimated (based on the best fit parameter values  $d$  and  $n$ ), de-normalized with respect to standard deviations and, as before, added to the seasonal averages in malaria case numbers  $\overline{A}$ . To check statistical consistency, transfer function model residuals were tested against Gaussianity, homoscedasticity, and independence, using the Kolmogorov-Smirnov, Bartlett and Portmanteau test, respectively.

Calibration for both models was performed by using the differential evolution adaptive Metropolis (DREAM) algorithm (Vrugt et al., 2008), yielding best fit values for the parameters along with their posterior probability density functions. DREAM algorithm convergence is monitored through the R-statistic of Gelman and Rubin (1992) and with a visual inspection of the likelihood behavior of each individual chain. To approximate the posterior probability density functions of the model parameters, a total of 80 000 model evaluations are performed with DREAM; the first 40 % of the samples of each chain are discarded and considered as burn-in, in order to evaluate the marginal densities of the parameters from the samples whose likelihood is close to convergence. Figure 5 shows the posterior probability distributions for the parameters of both models applied to the Mpumalanga data. It can be seen that calibration converged to well-defined optimal values.



**Fig. 4.** Observed and modeled monthly malaria incidence for (A) Mpumalanga, (B) Limpopo and (C) KwaZulu-Natal provinces. Observed incidence is shown alongside seasonal averages, case estimates from the ecohydrological model, and case estimates from the transfer function model. Insets are scatter plots of soil water content anomalies against malaria case anomalies for each province (dots, observations; solid lines, modeled relationships using Eqs. (17) and (19)). Best fit parameters are reported in Table 2.

### 3 Results and discussion

The seasonal component of malaria dynamics explained on average 41 % of the variability in cases (Table 1). None of the linear regressions between deseasonalized climate forcing and deseasonalized malaria cases explained significant variability in malaria case anomalies, resulting in at most a Nash efficiency (Nash et al., 1970) of 14 % for the rainfall-malaria correlation with a zero month time lag in the province of KwaZulu-Natal (Table 1). Introducing time lags improved the results only for the case of Mpumalanga province when one considers the rainfall-malaria correlation with one month time lag. These results indicate that a direct and linear association between climatic variables and malaria dynamics is not statistically supported, beyond seasonal effects.

We therefore considered whether soil water content dynamics may be associated with malaria case anomalies. The ecohydrological model was able to explain a significant fraction of the variation in malaria cases in all three provinces (Fig. 4 and Table 1). Furthermore, the parameter estimates of

the ecohydrological model, including the slope of the regression between soil water content anomalies and malaria case anomalies, are biologically and physically interpretable (Table 2). Estimates of the parameter  $m$ , the rate at which surface water is lost, yielded values between 0.1 and 0.44 months $^{-1}$ , although the related 95 % confidence limits, also shown in Table 2, show the presence of significant parameter uncertainty. The optimal values of parameter  $m$  translate into soil water transit times of 2–10 months, which highlight the persistence of perturbations induced in malaria cases by higher than usual precipitation. 95 % confidence limits of the transit times do not allow one rejecting the hypothesis that they are identical in all the three locations. Values of the soil water loss rate  $m$  between 0.3 and 0.5 months $^{-1}$  are comparable with estimates from other systems (Katul et al., 2007a), while the smaller estimate of this parameter for Limpopo suggests the possible presence of a biological delay in the disease dynamics not explicitly considered here (Pascual et al., 2008; Hay et al., 2000) and might be further imputable to Limpopo's different climatic conditions. As a whole, from

**Table 1.** Nash efficiencies for the seasonal component of malaria dynamics, the preliminary analysis, the ecohydrological model, and the transfer function model for the three South African provinces of Mpumalanga, Limpopo and KwaZulu-Natal. In the preliminary analysis, we examined the direct link between monthly data of malaria and rainfall, as well as malaria and minimum temperature applying linear regressions to the data before the removal of the seasonality denoted by “Non-deseasonalized” in the table and after the deseasonalization, considering only the “Anomalies” and taking into account the seasonal dynamics as well (Seasonal + anomalies).

	Equation	Mpumalanga	Limpopo	KwaZulu-Natal
Seasonal cycle		0.496	0.388	0.337
Rainfall-malaria linear regression				
Anomalies (0 month time lag)		0.117	0.027	0.140
Seasonal + anomalies (0 month time lag)		0.555	0.405	0.431
Non-deseasonalized (0 month time lag)		0.212	0.117	0.038
Anomalies (1 month time lag)		0.268	0.033	0.125
Seasonal + anomalies (1 month time lag)		0.631	0.408	0.420
Non-deseasonalized (1 month time lag)		0.434	0.111	0.148
Minimum temperature-malaria linear regression				
Anomalies (0 month time lag)		0.016	0.005	0.026
Seasonal + anomalies (0 month time lag)		0.505	0.391	0.351
Non-deseasonalized (0 month time lag)		0.231	0.001	0.001
Anomalies (1 month time lag)		0.046	0.001	0.069
Seasonal + anomalies (1 month time lag)		0.519	0.389	0.383
Non-deseasonalized (1 month time lag)		0.414	0.241	0.114
Ecohydrological model	(17)–(19)	0.351	0.160	0.313
Seasonal + ecohydrological		0.674	0.491	0.545
Transfer function model	(20)–(21)	0.351	0.190	0.156
Seasonal + transfer function		0.752	0.454	0.569

a physical perspective, these results suggest that unusually wet periods result in anomalously high surface and soil water storage, and thereby higher mosquito densities, ultimately yielding anomalously high malaria case numbers.

In our ecohydrological model, all the delays between favorable environmental conditions and malaria outbreaks are assumed to be induced by soil water dynamics. This is consistent with observed biological delays (in both the mosquito population and in disease dynamics) that are shorter than a month (e.g. Chitnis et al., 2008), while the memory induced by surface water storage is generally longer (e.g. Katul et al., 2007a). The proposed simplified model neglects compound delays and nonlinear interactions that might result in long delays between climatic forcing and malaria cases (Pascual et al., 2008; Hay et al., 2000). Our results are therefore likely to underestimate surface water loss rates  $m$  to some degree, since we attributed all physical and biological delays to soil water content dynamics alone.

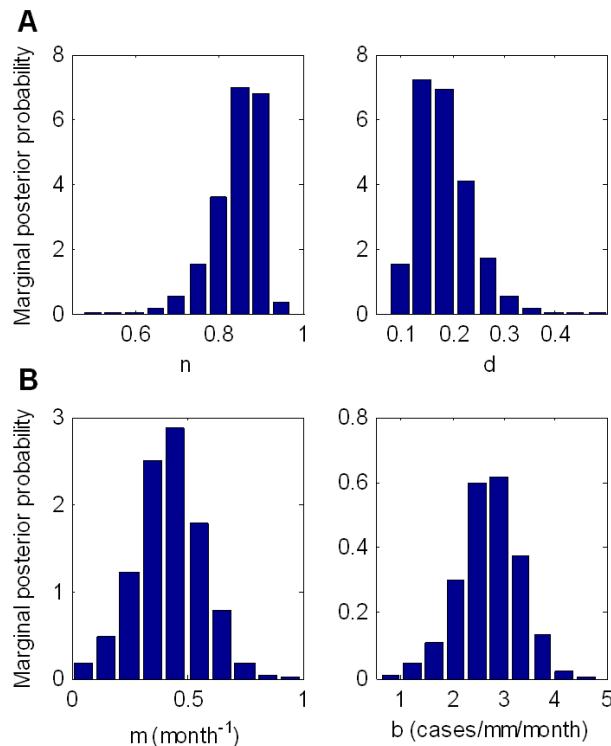
In the transfer function statistical model, we expect different parameter values, because deseasonalization of standard deviation was carried out. However, the resulting transit times are consistent with those obtained with the ecohydrological model. The results show that the hypotheses of Gaussianity and homoscedasticity of the residuals cannot be re-

jected at the 95 % confidence level in the three time series, therefore providing support for the assumption of model linearity. On the other hand, all the residuals exhibited a slight but still statistically significant autocorrelation (at the 95 % confidence level). This outcome is due to the inability of the model and the selected climatic determinants to fully account for the persistence of the malaria case anomalies. Such persistence could presumably be accounted for by other factors; however, this would introduce more parameters and thus uncertainties in their estimation. Alternatively, the residual autocorrelation could be eliminated introducing a further autoregressive component in the regression model to account for previous malaria cases, which would summarize the effects of the additional inputs above (see, for instance, the approach adopted by Zhou et al., 2004). However, this solution was not used here because it induces equifinality, therefore hindering the efficient identification of the causal relationship between rainfall/soil water content and malaria. Looking at the statistical model performances (Table 1), one can see that they are comparable with those of the ecohydrological model, with some differences depending on location.

It is not straightforward to provide a physical interpretation of the spatial variability for model parameter values. Additional information on the topography and climate

**Table 2.** Best set of model parameters (and 95 % confidence level) of the ecohydrological and transfer function models for the three provinces illustrated in Fig. 1.

Model	Parameter	Units	Mpumalanga	Limpopo	KwaZulu-Natal
Eco-hydrological model	$m$	$\text{month}^{-1}$	0.442 (0.164 ÷ 0.780)	0.107 (0.063 ÷ 0.359)	0.267 (0.180 ÷ 0.551)
	$b$	$\text{cases mm}^{-1} \text{month}^{-1}$	2.697 (1.394 ÷ 3.686)	0.673 (0.338 ÷ 1.135)	8.855 (5.457 ÷ 13.393)
Transfer function model	$n$	-	0.916 (0.726 ÷ 0.945)	0.871 (0.698 ÷ 0.931)	0.745 (0.103 ÷ 0.864)
	$d$	-	0.145 (0.108 ÷ 0.311)	0.242 (0.131 ÷ 0.362)	0.214 (0.074 ÷ 0.449)



**Fig. 5.** Posterior parameter distributions for (A) ecohydrological and (B) transfer function models at Mpumalanga. 95 % confidence bands for the parameters are given in Table 2.

of the study area can be obtained from the maps provided by the South African Department of Environmental Affairs and Tourism (see: <http://www.environment.gov.za/?q=content/maps-graphics>, last access: May 2012). The areas prone to malaria risk are located mainly in the low-lying parts of the three provinces where the climate is quite variable, as it is possible to assess from the maps of mean annual precipitation provided by the aforementioned department. It is moreover possible to see from the insets of Figs. 2–3 that the seasonal incidence of malaria in the province of KwaZulu-Natal is almost double with respect to the other two provinces. This high incidence of malaria cases is in line with the higher population counted in this province by the 2001 census (Statistics South Africa, 2003) and is also reflected by the results we obtained. In fact, parameter  $m$  is

relatively stable across the three provinces (Table 2), indicating relatively mild changes in soil and vegetation properties at the large scale we are considering. In contrast, the parameter  $b$  varies remarkably. In particular, a higher value is obtained for the KwaZulu-Natal province. Because  $b$  quantifies the sensitivity of malaria cases to rainfall, the higher values in KwaZulu-Natal province may be related to its coastal proximity, which implies the presence of a more humid climate. Moreover, wetlands and water bodies are characteristically more frequent in the land cover map of this province, if compared to what is depicted in Limpopo and Mpumalanga maps. A previous work by Kleinschmidt et al. (2001) carried out in two districts located in the northern side of KwaZulu-Natal province confirms the relevant role of wetlands and water bodies. In line with our results, they stressed the importance played by proximity to permanent water bodies, which are often surrounded by ephemeral mosquito habitats, as an additional risk factor for malaria incidence.

#### 4 Conclusions

Using time series of malaria cases from three South African provinces, we showed through two sets of analyses – an ecohydrological model and a transfer function model – that soil water content is an important driver of malaria dynamics. These analyses required a series of consecutive assumptions to be made in order to interface these models with available climate and malaria data. Nevertheless, we found a statistical association between modeled soil water content and malaria cases in all three provinces, with parameter estimates that were biologically and physically interpretable. Future work, with more extensive time series, should therefore focus on coupling soil water content dynamics to full epidemiological models (Eqs. 1–4), which require a larger number of parameters, but that maintain the non-linear feedbacks that are known to be important for malaria dynamics in endemic regions (Koelle and Pascual, 2004; Pascual et al., 2008; Laneri et al., 2010). Further developments should consider satellite soil moisture data as a possible alternative to direct estimates of soil moisture. Using such data would provide high resolution maps without invoking a water balance model dedicated to this purpose as Brugger et al. (2011) already proposed for the study of the seasonal cycle of *Culicoides* spp. abundance.

**Acknowledgements.** The authors wish to thank Charlotte McBride of the South Africa Weather Service who kindly provided us with the meteorological data for this study and are very grateful to Katia Koelle for the insightful discussions and comments on an earlier version of the manuscript. The authors also sincerely thank Luca Brocca, Alberto Viglione, Ignacio Andrés-Doménech and the anonymous reviewer whose comments helped to improve the manuscript quality. This research was partially supported by the US Department of Agriculture (USDA grant 2011-67003-30222) and by the United States Department of Energy (DOE) through the Office of Biological and Environmental Research (BER)Terrestrial Ecosystem Science (TES) program (DE-SC0006967). A. P. gratefully acknowledges the support of the Landolt & Cie visiting Chair “Innovative strategies for a sustainable future” at the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

Edited by: F. Laio

## References

- Bailey, N. T. J.: The biomathematics of malaria., Charles Griffin and Company Ltd., London, 1982.
- Bayoh, M. N. and Lindsay, S. W.: Effect of temperature on the development of the aquatic stages of *Anopheles gambiae* sensu stricto (Diptera:Culicidae), *B. Entomol. Res.*, 93, 375–381, 2003.
- Bertuzzo, E., Azale, S., Maritan, A., Gatto, M., Rodriguez-Iturbe, I., and Rinaldo, A.: On the space-time evolution of a cholera epidemic, *Water Resour. Res.*, 4, W01424, doi:10.1029/2007WR006211, 2008.
- Bhadra, A., Ionides, E. L., Laneri, K., Pascual, M., Bouma, M., and Dhaman, R. C.: Malaria in Northwest India: Data Analysis via Partially Observed Stochastic Differential Equation Models Driven by Lvy Noise, *J. Am. Stat. Assoc.*, 106, 440–451, doi:10.1198/jasa.2011.ap10323, 2011.
- Bombelles, A., Duchemin, J.-B., and Eltahir, E. A. B.: Hydrology of malaria: Model development and application to a Sahelian village, *Water Resour. Res.*, 44, W12445, doi:10.1029/2008WR006917, 2008.
- Briet, O. J. T., Vounatsou, P., and Amerasinghe, P. H.: Malaria seasonality and rainfall seasonality in Sri Lanka are correlated in space, *Geospatial Health*, 2, 183–190, 2008.
- Brugger, K., Rath, C., Stracke, D., Pathe, C., Wagner, W., and Rubel, F.: Statistical modelling of *Culicoides* spp. (Diptera: Ceratopogonidae) using ASCAT soil moisture data, EUMET-SAT/ESA Scatterometer Science Conference 2011, Darmstadt, Germany, 11–13 April, 2011.
- Chaves, L. F., Hashizume, M., Satake, A., and Minakawa, N.: Regime shifts and heterogeneous trends in malaria time series from Western Kenya Highlands, *Parasitology*, 139, 14–25, 2012.
- Chitnis, N., Hyman, J. M., and Cushing, J. M.: Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model, *B. Math. Biol.*, 70, 1272–1296, 2008.
- Craig, M. H., Snow, R. W., and Le Sueur, D.: A climate-based distribution model of malaria transmission in sub-Saharan Africa, *Parasitol. Today*, 15, 105–111, 1999.
- Craig, M. H., Kleinschmidt, I., Le Sueur, D., and Sharp, B. L.: Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: Part II. The impact of non-climatic factors, *Tropical Medicine and International Health*, 9, 1258–1266, 2004a.
- Craig, M. H., Kleinschmidt, I., Nawn, J. B., Le Sueur, D., and Sharp, B. L.: Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: Part I. The impact of climatic factors, *Tropical Medicine and International Health*, 9, 1247–1257, 2004b.
- Depinay, J., Mbogo, C. M., Killeen, G., Knols, B., Beier, J., Carlson, J., Dushoff, J., Billingsley, P., Mwambi, H., Githure, J., Toure, A. M., and McKenzie, F. E.: A simulation model of African Anopheles ecology and population dynamics for the analysis of malaria transmission, *Malaria J.*, 3, 29, doi:10.1186/1475-2875-3-29, 2004.
- Devi, N. P. and Jauhar, R. K.: Climatic variables and malaria incidence in Dehradun, Uttarakhand, India., *J. Vect. Borne Dis.*, 43, 21–28, 2007.
- Gelman, A. and Rubin, D. B.: Inference from iterative simulation using multiple sequences, *Sta. Sci.* 7, 457–472, doi:10.1214/ss/1177011136, 1992.
- Gomez-Elipe, A., Otero, A., Van Herp, M., and Aguirre-Jaime, A.: Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997–2003, *Malaria J.*, 6, 129, doi:10.1186/1475-2875-6-129, 2007.
- Hay, S. I., Myers, M. F., Burke, D. S., Vaughn, D. W., Endy, T., Ananda, N., Shanks, G. D., Snow, R. W., and Rogers, D. J.: Ethology of interepidemic periods of mosquito-borne disease, *Proc. Natl. Acad. Sci. USA*, 97, 9335–9339, 2000.
- Hay, S. I., Rogers, D. J., Shanks, G. D., Myers, M. F., and Snow, R. W.: Malaria early warning in Kenya, *Trends Parasitol.*, 17, 95–99, 2001.
- Hay, S. I., Cox, J., Rogers, D. J., Randolph, S. E., Stern, D. I., Shanks, G. D., Myers, M. F., and Snow, R. W.: Climate change and the resurgence of malaria in the East African highlands, *Nature*, 415, 905–909, 2002.
- Hoshen, M. B. and Morse, A. P.: A weather-driven model of malaria transmission, *Malaria J.*, 3, 32, doi:10.1186/1475-2875-3-32, 2004.
- Jury, M. R. and Kanemba, A. D.: A climate-based model for malaria prediction in south-eastern Africa, *S. Afr. J. Sci.*, 103, 57–62, 2007.
- Katul, G. G., Porporato, A., Daly, E., Oishi, A. C., Kim, H. S., Stoy, P. C., Juang, J. Y., and Siqueira, M. B.: On the spectrum of soil moisture from hourly to interannual scales, *Water Resour. Res.*, 43, W05428, doi:10.1029/2006WR005356, 2007a.
- Katul, G. G., Porporato, A., and Oren, R.: Stochastic dynamics of plant-water interactions, *Annu. Rev. Ecol. Evol. S.*, 38, 767–791, 2007b.
- Kermack, W. O. and McKendrick, A. G.: A Contribution to the Mathematical Theory of Epidemics, *Proc. R. Soc. Lond. A*, 115, 700–721, 1927.
- Kleinschmidt, I., Sharp, B. L., Clarke, G. P. Y., Curtis, B., and Fraser, C.: Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in KwaZulu Natal, South Africa, *Am. J. Epidemiol.*, 153, 1213–1221, 2001.
- Koelle, K. and Pascual, M.: Disentangling extrinsic from intrinsic factors in disease dynamics: A nonlinear time series approach with an application to cholera, *Am. Nat.*, 163, 901–913, 2004.
- Laneri, K., Bhadra, A., Ionides, E. L., Bouma, M., Dhaman, R. C., Yadav, R. S., and Pascual, M.: Forcing Versus Feedback:

- Epidemic Malaria and Monsoon Rains in Northwest India, PLoS Comput. Biol., 6, e1000898, doi:10.1371/journal.pcbi.1000898, 2010.
- Lindsay, S. W., Bodker, R., Malima, R., Msangeni, H. A., and Kisinza, W.: Effect of 1997–98 El Nino on highland malaria in Tanzania, Lancet, 355, 989–990, 2000.
- Macdonald, G.: The Epidemiology and Control of Malaria London, Oxford University Press, Oxford, 1957.
- McCallum, H., Barlow, N., and Hone, J.: How should pathogen transmission be modelled?, Trends Ecol. Evol., 16, 295–300, 2001.
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I: A discussion of principles, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.
- Paaijmans, K. P., Read, A. F., and Thomas, M. B.: Understanding the link between malaria risk and climate, Proc. Natl. Acad. Sci. USA, 106, 13844–13849, 2009.
- Pascual, M., Cazelles, B., Bouma, M. J., Chaves, L. F., and Koelle, K.: Shifting patterns: malaria dynamics and rainfall variability in an African highland, P. Roy. Soc. B-Biol. Sci., 275, 123–132, 2008.
- Patz, J. A., Strzepek, K., Lele, S., Hedden, M., Greene, S., Noden, B., Hay, S. I., Kalkstein, L., and Beier, J. C.: Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya, Tropical Medicine and International Health, 3, 818–827, 1998.
- Porphyre, T., Bicout, D. J., and Sabatier, P.: Modelling the abundance of mosquito vectors versus flooding dynamics, Ecol. Model., 183, 173–181, 2005.
- Porporato, A., Daly, E., and Rodriguez-Iturbe, I.: Soil water balance and ecosystem response to climate change, Am. Nat., 164, 625–632, 2004.
- Ross, R.: The prevention of malaria, Dutton, New York, 1910.
- Shaman, J., Stieglitz, M., Stark, C., Le Blancq, S., and Cane, M.: Using a dynamic hydrology model to predict mosquito abundances in flood and swamp water, Emerg. Infect. Dis., 8, 6–13, 2002.
- Smith, D. L. and McKenzie, F. E.: Statics and dynamics of malaria infection in Anopheles mosquitoes, Malaria J., 3, 13, doi:10.1186/1475-2875-3-13, 2004.
- Statistics South Africa, Census 2001: Census in brief, Report No: 03-02-03(2001), Pretoria, 111 pp., 2003
- Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 44, W00B09, doi:10.1029/2007WR006720, 2008.
- World Health Organization: World Malaria Report, WHO/HTM/GMP/2008.1, 2008.
- Wyse, A. P. P., Bevilacqua, L., and Rafikou, M.: Simulating malaria model for different treatment intensities in a variable environment, Ecol. Model., 206, 22–330, 2007.
- Yamana, T. K. and Eltahir, E. A. B.: On the use of satellite-based estimates of rainfall temporal distribution to simulate the potential for malaria transmission in rural Africa, Water Resour. Res., 47, W02540, doi:10.1029/2010WR009744, 2011.
- Zhou, G., Minakawa, N., Githeko, A. K., and Yan, G. Y.: Association between climate variability and malaria epidemics in the East African highlands, Proc. Natl. Acad. Sci. USA, 101, 2375–2380, 2004.