



A breeding site model for regional, dynamical malaria simulations evaluated using *in situ* temporary ponds observations

Ernest O. Asare,^{1,2} Adrian M. Tompkins,² Leonard K. Amekudzi,¹ Volker Ermert³

¹Department of Physics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana; ²Abdus Salam International Centre for Theoretical Physics, Trieste, Italy; ³Institute of Geophysics and Meteorology, University of Cologne, Cologne, Germany

Abstract

Daily observations of potential mosquito developmental habitats in a suburb of Kumasi in central Ghana reveal a strong variability in their water persistence times, which ranged between 11 and 81 days. The persistence of the ponds was strongly tied with rainfall, location and size of the puddles. A simple power-law relationship is found to fit the relationship between the average pond depth and area well. A prognostic water balance model is derived that describes the temporal evolution of the pond area and depth, incorporating the power-law geometrical relation. Pond area increases in response to rainfall, while evaporation and infiltration act as sink terms. Based on a range of evaluation metrics, the prognostic model is judged to provide a good representation of the pond coverage evolution at most sites. Finally, we demonstrate that the prognostic equation can be generalised and equally applied to a grid-cell to derive a fractional pond coverage, and thus can be implemented in spatially distributed models for relevant vector-borne diseases such as malaria.

Correspondence: Ernest O. Asare, Department of Physics, Kwame Nkrumah University of Science and Technology, Accra Road, Kumasi, Ghana.
Tel: +233 24 3425243.
E-mail: eoheneasare@gmail.com

Key words: Pond water area; Pond water depth; Area-depth relation; Water fraction.

Acknowledgements: EOA was funded by two International Centre of Theoretical Physics (ICTP) programme, namely the Italian government's funds-in-trust programme and the ICTP PhD Sandwich Training and Educational Programme (STEP). The study was also partly funded by two European Union Seventh Framework Programmes projects; HEALTHY FUTURES under the grant agreement number 266327 and QWeCI (Quantifying Weather and Climate Impacts on health in developing countries) under the grant agreement number 243964. The authors thank three anonymous reviewers who greatly improved the original submission.

Received for publication: 17 June 2015.
Accepted for publication: 4 February 2016.

©Copyright E.O. Asare *et al.*, 2016
Licensee PAGEPress, Italy
Geospatial Health 2016; 11(s1):390
doi:10.4081/gh.2016.390

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Introduction

Surface hydrology and water body temperature are two key factors that control the aquatic stage life cycle of mosquitoes and thus adult abundance by influencing the stability of habitat and larvae growth rates, respectively. Mosquitoes may exploit any available water for oviposition, natural or man-made (Imbahale *et al.*, 2011; Fillinger *et al.*, 2004), permanent or temporary (Fillinger *et al.*, 2004), clean or polluted (Sattler *et al.*, 2005; Awolola *et al.*, 2007; Chinery, 1984) and of various sizes from hoof-prints of animals to the edges of large water bodies (Sattler *et al.*, 2005; Mutuku *et al.*, 2006b; Imbahale *et al.*, 2011), although individual species have preferences of habitat type. For example, *Anopheles gambiae* complex mosquitoes, the principal malaria vector in Sub-Sahara Africa prefer temporary, sunlit water bodies for their breeding, which become abundant during the rainy season (Mutuku *et al.*, 2006a; Minakawa *et al.*, 2004), although their larvae have also been found in polluted waters (Imbahale *et al.*, 2011; Awolola *et al.*, 2007; Sattler *et al.*, 2005).

Small O(metre)-sized breeding habitats have many advantages over larger permanent breeding sites that increase the developmental rate or survival probability of the aquatic stage. Firstly, these habitats contain small amounts of water and therefore their temperatures are high relative to deeper water bodies which shortens the larval-pupal development time (Munga *et al.*, 2005; Bayoh and Lindsay, 2004; Ndenga *et al.*, 2011), although high temperature increases larval mortality (Bayoh and Lindsay, 2003). Several studies have found these micro habitats to be productive and therefore their contribution to mosquito vector abundance, especially during the rainy season, should not be ignored in dynamical models (Bomblies, 2012; Imbahale *et al.*, 2011; Sattler *et al.*, 2005). For example, in western Kenya, Minakawa *et al.* (2004) found more than 80% of *Anopheles gambiae* s.s. in isolated pools with water surface areas lower than 0.1 m². Secondly, such pools are temporary and therefore contain fewer or no competitors and predators decreasing the larvae mortality rate (Koenraad *et al.*, 2004; Sunahara *et al.*, 2002). Thirdly, human activities contribute to the creation of these habitats which are found near human settlements and thus time spent by the gravid mosquito to locate surface water for oviposition is reduced (Mutuku *et al.*, 2006b; Minakawa *et al.*, 1999). Another critical factor determining mosquito breeding habitat productivity is its stability because only habitats that persist long enough can produce adult mosquitoes (Gianotti *et al.*, 2009).

These mechanisms emphasise the importance of small-sized breeding habitats during the rainy season and how they can account for the seasonal variation in malaria vector abundance. Therefore, critical factors influencing the stability and productivity of these transient habitats need to be examined. Low frequency and intensity of rainfall can lead to desiccation of habitats before adult emergence (Himeidan *et al.*, 2009; Srivastava *et al.*, 2001). Small bodies on the sub-metre scale

can produce adult vectors if they are contained in larger scale depressions that collect water (Bomblies *et al.*, 2008), although often the most productive pond size in terms of number and duration are on the 1 to 10 metre scale (Gianotti *et al.*, 2009). Other important factors that can control the stability of these small-sized habitats are hydrological parameters of the area which include soil type, soil moisture content and water table depth (Bomblies *et al.*, 2009; Montosi *et al.*, 2012).

In order to model malaria effectively using a dynamical modeling approach, a realistic representation of the surface hydrology is required. However, most dynamical malaria models have no representation of surface hydrology. One impediment to the development of more realistic treatment of the availability of ponds is their small spatial scale, which limits the use of remote sensing techniques to make parametrisation assumptions, a situation not aided by the lack of *in situ* observations. The aim of this paper is therefore to derive a water balance model for the surface area of ponds, driven by the processes of precipitation and surface runoff, with water lost through overflow, infiltration and evaporation. The model will be calibrated and evaluated using *in situ* pond measurements. The final goal is then to generalise the model to represent the aggregate total pond availability of order of kilometre sized grid-box such that the model can be applied to regional-scale vector-borne disease modeling.

Hayashi and Van der Kamp (2000) introduced a power function area-depth (A-h) relationship that requires two independent measurements of pond area and depth to determine scaling and shape constants. This model (hereafter referred to as Hayashi model) is a diagnostic model of water bodies that relates the volume, area and depth. This Hayashi model has been used extensively to study permanent and semi-permanent pond dynamics in Senegal relevant for vectors of rift valley fever that have spatial scales of tens of metres (Soti *et al.*, 2010). In addition, the model has also been evaluated for regional terrain (Brooks and Hayashi, 2002; Minke *et al.*, 2010). However, no study has examined whether such a

model can also be successfully applied to small temporary sub-metre scale water bodies. We therefore report results of *in situ* measurements of a range of such developmental habitats in a peri-urban area of Kumasi in Ghana, and then evaluate whether the Hayashi model can describe the relationship between area and depth of these small breeding sites.

There is considerable variability in the availability and size of small-scale breeding sites over the course of a rainy season. Thus to model the evolution of these pools a prognostic treatment is required. A water balance model is derived that represents some of the key sources and sinks of ponds, namely rainfall, surface runoff, pond overflow, evaporation and infiltration. The model incorporates the geometrical model of Hayashi to convert water volume to pond coverage and thus breeding site availability, which is then evaluated using the *in situ* pond data.

Finally, to apply the surface hydrology model to regional or national-scale distributed disease transmission models, the equation is generalised to predict the fractional water coverage constituted by all the ponds present in a grid-cell that may range in size from a kilometre to hundreds of kilometres. The aggregate statistics of all the monitored ponds are used to evaluate the potential of this approach for application in regional-scale vector-borne disease models.

Materials and Methods

Study area and data

The study was conducted in the Kumasi (Ghana) metropolis comprising parts of the Kwame Nkrumah University of Science and Technology (KNUST) campus and peri-urban areas of Ayeduase, a town sharing a boundary with the university (Figure 1). For 81 days between 6 June (day 157 in 2011) and 25 August (day 237 in 2011) daily obser-

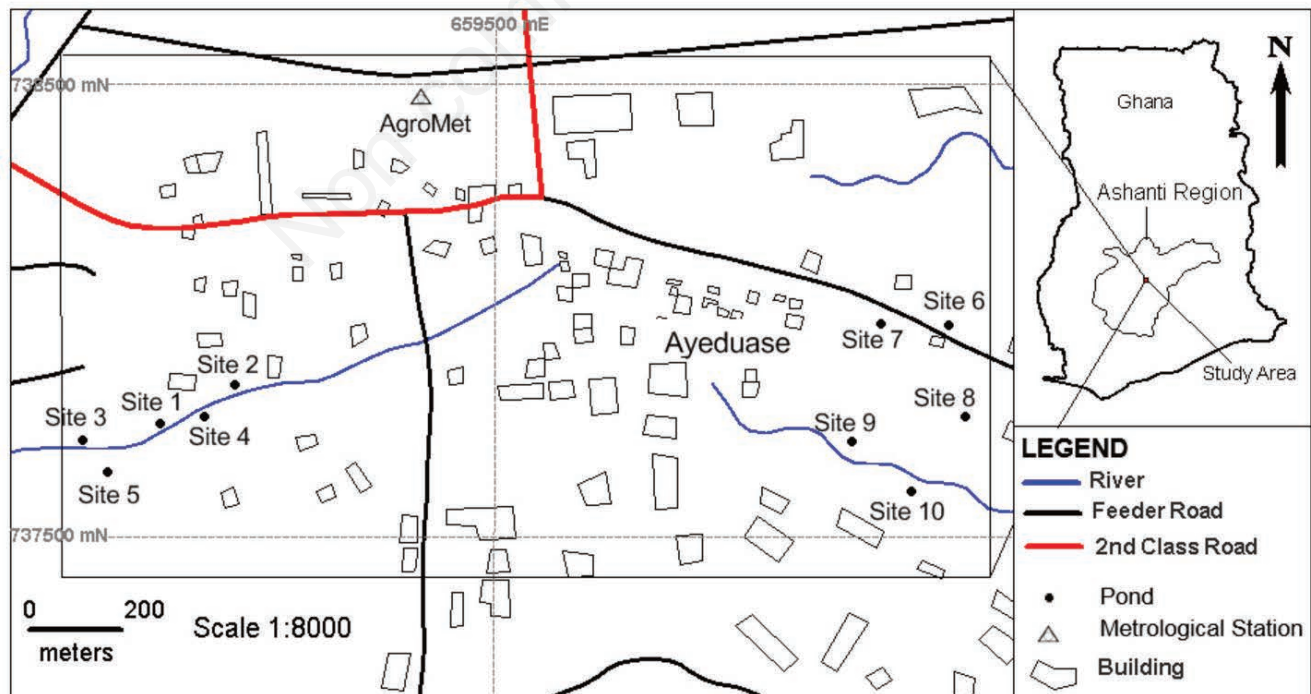


Figure 1. Location of the study area: Kwame Nkrumah University of Science and Technology campus and peri-urban areas of the Ayeduase. Also indicated are the 10 temporary sites together with the meteorological station AgroMet.

vations of small-scale potential mosquito breeding habitats were undertaken. Ten sites with continuous water presence for at least 10 days are considered for this study. The monitored ponds were generated during the course of the monsoonal rains of 2011.

The potential mosquito developmental habitats consisted of a truck tyre truck, small surface depressions, puddles and ditches which collect water during the wet season (Figure 2). At each site, parameters measured include water temperature, water depth and the major and minor dimensions of water. However, irregular puddles were divided into segments before measuring major and minor dimensions. The daily depth of water in each pond was calculated as average of three readings taken at three marked points within the water coverage area of the pond using a tape measure with an estimated 1 mm measurement accuracy.

Although study of mosquito larvae was not undertaken, mosquito larvae were observed (by visual inspection without larvae identification) in all habitats at some point during the study period. Pond measurements were not performed every day during the study period resulting in some data gaps existing in the time series. The number of potential water bodies decreased over the period since various ponds dried out and those that persisted decreased in size and depth in August 2011 due to the reduction of rainfall during that month (Owusu and Waylen, 2009; Manzanos *et al.*, 2014). While some of the habitats dried out within a few days of rainfall events, others persisted throughout the study period of 81 days. The geographical-coordinates of the sites were recorded using a simple hand held global position system receiver (GARMIN eTrex series). Daily rainfall and evaporation data were obtained from a minute temporal resolution automated rain gauge installed at the KNUST AgroMet station (Figure 1).

The precipitation and evaporation data used to drive the models developed in this work were obtained from the KNUST AgroMet weather station which was located very close to the pond sites (see location AgroMet in Figure 1). In this study, evaporation from the water surface was simply equated to the ambient air evaporation measured by the weather station.

Diagnostic pond geometry model

Hayashi and Van der Kamp (2000) developed an area-depth (A-h) relation for water bodies expressed as $A \propto h^{2/p}$, with the pond water area estimated from eq. 1:

$$A = A_{ref} \left(\frac{h}{h_{ref}} \right)^{2/p} \quad (\text{eq. 1})$$

where A is the pond water surface area, h is the pond water depth, p represents the pond geometry and h_{ref} and A_{ref} are reference pond water depth and surface area measurements, respectively.

The key assumption of the model is that, averaged radially, variations in the water body geometry average out, so that the relationship between the water body depth and its areal coverage can be given by a simple power law. The relationship is specified by the exponent factor p, which is a constant representing the geometry of the habitat, describing how the depth relates to the area. Since p is assumed a constant, only two, coincident representative values of pond depth and area are required in order to close the equation, h_{ref} and A_{ref} . Several options exist to set these two parameters. For example, they could be set using the maximum values of the pond, an *ad hoc* pair of measurements taken on a random day, or the average values over an entire measurement campaign. The latter approach is employed here to reduce sensitivity to measurement error.

Once h_{ref} and A_{ref} are known, the key p parameter is estimated using a least-squared fit of the power function of the form of eq. 1 to all the data points. Again, by using all measurements, the sensitivity to measurement error should be reduced relative to the method of Minke *et al.* (2010), who computed p based on two separate measurements of water surface area and depth. The p parameter is therefore equal to 2 divided by the index of the best-fit regression line.

Prognostic pond area model

A prognostic water balance model for small-sized transient breeding habitats is developed (Figure 3). The model assumes that ponds are filled either by precipitation directly falling on their surface area A, or by surface runoff that occurs in small scale catchments of scale A_{max} which represents the maximum possible dimension of the pond. Rainfall that falls outside the catchment drains to the larger scale river network and is not collected in ponds. Ponds lose water through the processes of evaporation, infiltration or overflow losses. The daily volumetric water balance of a puddle is approximated by eq. 2:



Figure 2. Two typical monitored potential mosquito breeding habitats within the study area. A) site 10; B) site 9.

$$\frac{dV}{dt} = [PA + R(A_{max} - A)] \left(1 - \frac{A}{A_{max}}\right) - A(E + I) \quad (\text{eq. 2})$$

where $\frac{dV}{dt}$ represents the rate of change in puddle water volume, A_{max}

is the pond catchment area approximated as maximum measured water surface area of the pond, A is the daily pond water surface area, P , R , E and I are the daily amount of precipitation, runoff, evaporation and infiltration, respectively. The first term on the right-hand side of eq. 2 represents the direct contribution of precipitation falling on the pond area A , the second term $(A_{max} - A)$ represents the input surface runoff from the remainder of the catchment area, while the last term represents loss of water from the puddle through evaporation and infil-

tration. The $\left(1 - \frac{A}{A_{max}}\right)$ term represents the losses due to overflow, which is approximated by a simple linear function, such that overflow increases with the pond volume. The term is zero when the area of the puddle reaches its maximum (*i.e.*, A_{max}) in which case additional water from rainfall and runoff is balanced by outflow.

The runoff R term was estimated using the Soil Conservation Service curve number (SCS-CN) method developed by the United States Department of Agriculture (USDA) USDA (1972):

$$R = \frac{(P - 0.2S)^2}{P + 0.8S} \quad (\text{eq. 3b})$$

$$S = \frac{25400}{CN} - 254 \quad (\text{eq. 3b})$$

where S is potential maximum retention (mm) and CN (range between 0 and 100) is the curve number, a dimensionless parameter indicating the land surface and soil type characteristics. At the lower CN value, all rainfall infiltrates without generating runoff while all rainfall becomes runoff at the upper CN threshold. USDA (1972) provides CN values for various land cover and soil types. Note that CN is also a function of soil

moisture, which is neglected, in the present model.

The loss of pond water due to infiltration increases with increasing water surface area: with maximum infiltration occurring after rainfall events, while minimum infiltration occurs when the water level reduces to reach the clogged region of the pond (Porphyre *et al.*, 2005; Martin-Rosales and Leduc, 2003). This is represented using a linear relation with the daily pond water surface area and depth as given by eqs. 4a and 4b for area and depth, respectively:

$$I = I_{max} \left(\frac{A}{A_{max}}\right) \quad (\text{eq. 4a})$$

$$I = I_{max} \left(\frac{h}{h_{max}}\right)^{2/p} \quad (\text{eq. 4b})$$

where I_{max} is the estimated daily maximum infiltration, h_{max} is the measured maximum water depth and h is the daily water depth.

In order to translate the volumetric pond equation (eq. 2) to prognostic equations for pond area and depth, we introduce the volume, height, area relationship of the diagnostic Hayashi model. For generic depression shape, Hayashi and Van der Kamp (2000) further provided a relation that links the water depth and surface area to its volume

$$V = \frac{A_{ref}}{(1+2/p)} \frac{h^{(1+2/p)}}{h_{ref}^{2/p}}, \text{ and substituting this into the Hayashi}$$

model (eq. 1) gives the relationship between the derivative of pond area and pond volume:

$$\frac{dA}{dt} = \frac{2}{ph_{ref}} \left(\frac{A_{ref}}{A}\right)^{\frac{p}{2}} \frac{dV}{dt} \quad (\text{eq. 5a})$$

$$\frac{dh}{dt} = \frac{1}{A_{ref}} \left(\frac{h_{ref}}{h}\right)^{\frac{2}{p}} \frac{dV}{dt} \quad (\text{eq. 5b})$$

It is then straightforward to substitute eq. 2 into eqs. 5a and 5b to

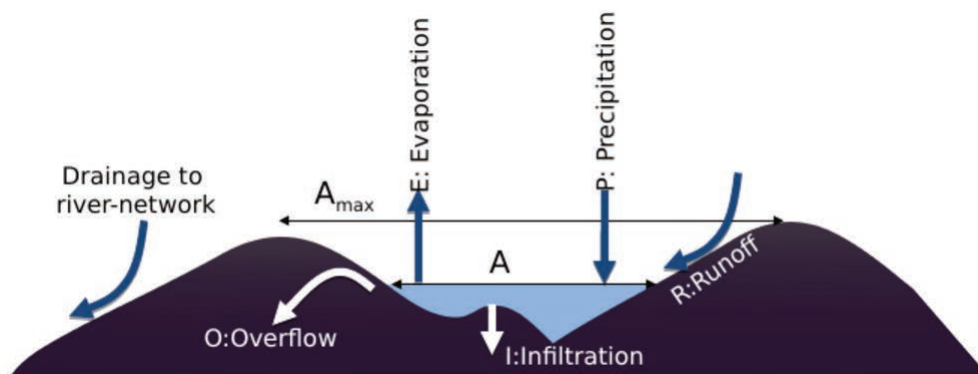


Figure 3. Scheme illustrating the various processes in the pond model.



derive the prognostic pond area and depth model:

$$\frac{dA}{dt} = \frac{2}{ph_{ref}} \left(\frac{A_{ref}}{A}\right)^{\frac{p}{2}} \left([PA + R(A_{max} - A)] \left[1 - \frac{A}{A_{max}}\right] - A(E + I) \right) \quad (\text{eq. 6a})$$

$$\frac{dh}{dt} = \frac{1}{A_{ref}} \left(\frac{h_{ref}}{h}\right)^{\frac{2}{p}} \left([PA + R(A_{max} - A)] \left[1 - \frac{A}{A_{max}}\right] - A(E + I) \right) \quad (\text{eq. 6b})$$

A simple semi-implicit, backward in time numerical technique is used to integrate the equation forward in time stably using a daily timestep, treating linear terms at time level $t + 1$ for stability. However, solving for the power terms $A^{-p/2}$ and $h^{-2/p}$ on the right hand side of eq. 6 at future time levels makes the solution intractable, while solving for these terms at time level t would imply that the puddle would not be refilled once it dries out (*i.e.*, $A(t) = 0 \Rightarrow \frac{dA}{dt} = 0$). A simple two-step solver approach is thus introduced to address this issue. The $A^{-p/2}$ and $h^{-2/p}$ terms are approximated with reference values, $A_{ref}^{-p/2}$ and $h_{ref}^{-2/p}$, respectively to reduce eq. 6 to a simple linear equation. This is then solved implicitly to provide first guess puddle area and depth values at the first step:

$$A_{fg,t} = \frac{A_t + \frac{2R}{ph_{ref}}(A_{max} - A_{ref})\Delta t}{1 + \frac{2}{ph_{ref}} \left[E + R - P + \frac{A_{ref}}{A_{max}}(I_{max} + P - R) \right] \Delta t} \quad (\text{eq. 7a})$$

$$h_{fg,t} = \frac{h_t + \frac{A_{max}R}{A_{ref}} \left[1 - \left(\frac{h_{ref}}{h_{max}}\right)^{2/p} \right] \Delta t}{1 + \frac{h_t^{2/p-1}}{h_{ref}^{2/p}} \left[E + R - P + \left(\frac{h_{ref}}{h_{max}}\right)^{2/p} (I_{max} + P - R) \right] \Delta t} \quad (\text{eq. 7b})$$

where $A_{fg,t}$ and $h_{fg,t}$ provides first guess puddle area and depth values. For the second step solution, eq. 6 is solved again implicitly but replacing the power terms $A^{-p/2}$ and $h^{-2/p}$ terms with $A_{ref}^{-p/2}$ and $h_{ref}^{-2/p}$ respectively to provide the final model equation:

$$A_{t+1} = \frac{A_t + \frac{2R}{ph_{ref}} \left(\frac{A_{ref}}{A_{fg,t}}\right)^{p/2} (A_{max} - A_{fg,t})\Delta t}{1 + \frac{2}{ph_{ref}} \left(\frac{A_{ref}}{A_{fg,t}}\right)^{p/2} \left[E + R - P + \frac{A_{fg,t}}{A_{max}}(I_{max} + P - R) \right] \Delta t} \quad (\text{eq. 8a})$$

$$h_{t+1} = \frac{h_t + \frac{A_{max}R}{A_{ref}} \left(\frac{h_{ref}}{h_{fg,t}}\right)^{2/p} \left[1 - \left(\frac{h_{fg,t}}{h_{max}}\right)^{2/p} \right] \Delta t}{1 + \frac{h_t^{2/p-1}}{h_{fg,t}^{2/p}} \left[E + R - P + \left(\frac{h_{fg,t}}{h_{max}}\right)^{2/p} (I_{max} + P - R) \right] \Delta t} \quad (\text{eq. 8b})$$

Aggregate pond fraction model for regional simulations

The pond area model is generalised to represent the total coverage of ponds over grid-cells that may range from several kilometres to over 100 km in dimension, such that it can be used in large-scale, distributed model for vectors or vector-borne disease transmission (*e.g.*

Hoshen and Morse, 2004; Tompkins and Ermert, 2013).

If we denote the distributed model grid-cell size as A_{grid} , and the area of pond i to be A_i , then the fractional coverage F_i of that single pond is simply $F_i = A_i/A_{grid}$. Thus the rate of change of the fractional coverage of this single pond that has a shape factor p_i is straightforward to derive which is deduced from eq. 6a:

$$\frac{dF_i}{dt} = \frac{2}{ph_{ref}} \left(\frac{F_{ref}}{F_i}\right)^{\frac{p_i}{2}} \left([PF_i + (F_{i,max} - F_i)R] \left[1 - \frac{F_i}{F_{i,max}} \right] - (E + I)F_i \right) \quad (\text{eq. 9})$$

The aggregate total fractional coverage is the summation of the n individual ponds in a grid cell:

$$\frac{dF}{dt} = \sum_i^n \frac{dF_i}{dt} = \sum_i^n \frac{2}{ph_{ref}} \left(\frac{F_{ref}}{F_i}\right)^{\frac{p_i}{2}} \left([PF_i + (F_{i,max} - F_i)R] \left[1 - \frac{F_i}{F_{i,max}} \right] - (E + I)F_i \right) \quad (\text{eq. 10})$$

This summation would be straightforward, except for the fact that each individual pond has its own value for the shape factor p_i . If we assume, however, that an appropriate value of p can be adopted that adequately describes the mean shape of the aggregate pond (this is not simply $\frac{1}{n} \sum_i^n p_i$ due to the nonlinearity), which will be referred to as \bar{p} then the aggregate fractional pond coverage is

$$\frac{dF}{dt} = \sum_i^n \frac{dF_i}{dt} = \sum_i^n \frac{2}{ph_{ref}} \left(\frac{F_{ref}}{F_i}\right)^{\frac{p_i}{2}} \left([PF_i + (F_{i,max} - F_i)R] \left[1 - \frac{F_i}{F_{i,max}} \right] - (E + I)F_i \right) \quad (\text{eq. 11})$$

Performance diagnostics

To assess the performance of the models, Nash-Sutcliffe efficiency [NSE; Nash and Sutcliffe (1970)], the coefficient of determination (R^2) and the relative mean absolute error (RMAE) defined by eq. 12 are used:

$$NSE = 1 - \frac{\sum_{i=1}^N (O_i - S_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}$$

$$R^2 = \left(\frac{\sum_{i=1}^N (O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^N (O_i - \bar{O})^2 \sum_{i=1}^N (S_i - \bar{S})^2}} \right)^2$$

$$RMAE = \frac{1}{\bar{O}} \left(\frac{1}{N} \sum_{i=1}^N |O_i - S_i| \right) \quad (\text{eq. 12})$$

where S_i refers to the i_{th} simulated value, O_i is the i_{th} observation, \bar{O} is the mean of observed data, \bar{S} is the mean of simulated data and N is the total number of observations. NSE ranges from $-\infty$ to 1 where NSE=1 indicates a perfect fit to the data (*i.e.* a scatter plot of model versus observations produces a perfect 1:1 line). A value of NSE=0 implies the

model performance is equivalent to using climatology, that is, simply using the mean of the observations. Zero or negative NSE values indicates that model has no skill, and its performance is worse than the mean of the observations.

Results

Pond measurements

The depth and dimension of the ponds were strongly influenced by rainfall amounts and its frequency as well as the local hydrology of the pond location. The average pond depths of the ten monitored temporary small-sized potential mosquito breeding habitats varied between 5.9 and 14.5 cm with the average areas ranging between 1.4 and 9.1 m² and decreasing significantly towards the end of the study period (Figure 4). Note that at the end of the observational period only about half of the potential breeding habitats were available. Sites 2, 5, 6, 7 and 8 dried out, site 3 was destroyed, while sites 1, 4, 9 and 10 remained with water at the end of observational period. This observation also reveals the link between pond stability and local hydrology of the pond location.

The average pond water depth and surface area, maximum pond water depth and surface area as well as elevation of these breeding habitats are shown in Table 1. The total rainfall recorded throughout the study period was 541.1 mm with July recording the highest rainfall amount (Figure 4). During the months of June and parts of July (up to 21 July, day 202), the maximum dry period (number of days between rainy days) was 5 days with 49.9 mm being the maximum daily rainfall recorded. After 1 July 2011 a dryer period started. The maximum dry period increased to 9 days and the maximum daily rainfall amounted only to 15.3 mm. Most of the ponds remained stable containing water from the start of the experiment until day 207 (26 July 2011) with the exception of sites 6 and 7, which are located far from the stream and also at a higher elevation (Figure 1 and Table 1). After this date, about 5 of the ponds desiccated and also significant reduction in both average area and depth occurred at the end of the study period (Figure 4). For the ten sites studied, variability in the stability of temporary mosquito developmental habitats ranged from 11 to 81 days and was strongly linked with rainfall, local hydrology of habitat location and size of the habitat (Table 1). For instance, due to the short dry spell occurring at the Guinea Coast in August between the major and minor rainfall sea-

sons (Owusu and Waylen, 2009; Manzanos *et al.*, 2014), only about half of the monitored ponds contained water towards the end of August. However, those ponds with persisted water at the end of observational period were located close to permanent streams and waterlogged areas (Figure 1 and Table 1). This is in agreement with the study of Himeidan *et al.* (2009), who also found that longer-lived developmental habitats were located near streams and water sources in the eastern African highlands. In the same district, Mushinzimana *et al.* (2006) made similar observations with more than 60% of positive habitats found within 50 metres from streams in both dry and rainy seasons. Similar observations were made by Bomblies *et al.* (2009); a different

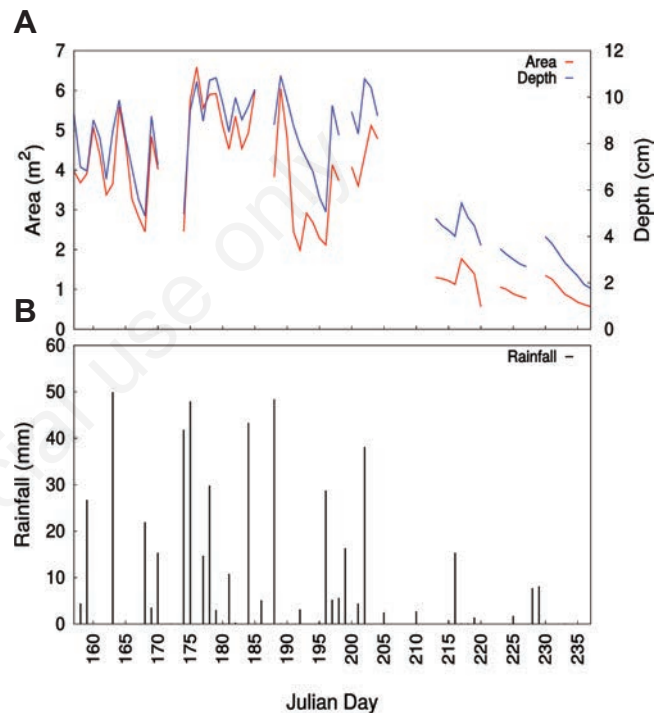


Figure 4. A) Average area (left axis, red line) and average depth (right axis, blue line) of the monitored breeding habitats. B) Daily precipitation amounts of the automated rain gauge from the Kwame Nkrumah University of Science and Technology campus.

Table 1. Average area and depth, maximum area and depth, elevation, total number of days when pond contained water and maximum number of days of continuous water presence of the ten temporary breeding habitats.

Name	Elevation (m)	Pond water (days)	Average area (m ²)	Maximum area (m ²)	Average depth (cm)	Maximum depth (cm)	Maximum water (days)
Site 1	259	81/81	1.9	3.8	5.5	8.8	81
Site 2	258	65/81	2.4	4.6	5.8	9.0	46
Site 3	257	<i>31/31</i>	5.0	8.0	7.7	10.3	31
Site 4	256	81/81	4.7	8.8	10.7	16.7	81
Site 5	258	79/81	2.5	4.9	6.8	11.0	79
Site 6	283	45/79	3.6	5.0	6.7	9.7	11
Site 7	282	51/79	7.5	12.5	7.7	10.2	13
Site 8	263	56/80	5.9	8.8	7.5	9.8	56
Site 9	262	<i>60/60</i>	1.3	1.8	8.0	11.3	60
Site 10	260	81/81	5.0	7.4	12.9	16.8	81

Italics represent sites that did not endure the entire experiment: site 3 was destroyed and site 9 measurement started later on 26 June 2011.

water table was detected for two Nigerian villages, which led to a strong difference in the persistence of pools. In addition, these puddles are potential mosquito breeding habitats since their stability far exceeds the time required between the eggs laying and adult emergence mosquitoes (Depinay *et al.*, 2004). However, the productivity of the longer-lived ponds might be affected by predators due to their long water persistence times (Chase and Knight, 2003).

Minakawa *et al.* (2005) observed that pond stability positively correlated with habitat size and location with the former having a higher correlation coefficient. Their results are consistent with our findings, where both size and location influence pond stability. However, our results point to habitat location having a more pronounced impact. For instance, site 6 (average area: 3.6 m²; average depth: 7.0 cm) and site 7 (average area: 8.5 m²; average depth: 7.7 cm) are located within the same area. In this case, site 6 with a smaller water volume had a shorter water persistence time (Table 1). Conversely, site 1 (average area: 1.9 m²; depth: 5.5 cm) reveals a greater stability than sites 6 and 7, due to its close location to a permanent stream. This site contained water throughout the study period despite its smaller dimension. This confirms that both habitat size and local hydrological conditions influence pond stability and within the same area, size might be the dominant factor (Minakawa *et al.*, 2005), however, over a wider area, habitats location might be the key factor controlling its variability in stability. For instance, Himeidan *et al.* (2009) found that the percentage of stable habitats was 48.76% and 80.79% for habitats located on top of the hills and near streams, respectively.

Pond geometry

Using the site data, the best-fit p parameter for the Hayashi model was derived for each site. Figure 5 shows example for some of the sites and the p values for all the sites are given in Table 2. The p values range between 1.1 and 2.0 with an average of 1.6. This range of values lies within the expected range reported for temporary pools and ponds (Brooks and Hayashi, 2002; Hayashi and van der Kamp, 2007). For instance, Brooks and Hayashi (2002) reported p values ranging between 0.6 and 2.24 for 34 vernal pools. The R^2 of the power function fit range between 0.54 and 0.94. These results indicate that the Hayashi diagnostic relationship can also describe reasonably well the geometry of small ponds and forms a good basis on which to build the prognostic model for pond coverage.

Evaluating the prognostic pond model

The performance of the area-depth (A-h) relation for the 10 micro-habitats was validated against field observations. The results (Figure 6) demonstrate the model potential to simulate the daily variability of the pond water surface area and depth. The water area was used to simulate the daily pond water depth as shown in Figure 6. To assess the performance of the model, RMAE, R^2 and NSE differences between observation and model output were computed and the results are summarised in Table 2. The model results are subdivided in relation to ponds location.

Sites 1 to 5 are located close to permanent stream (Figure 1), the model captured the observed variability of ponding due to rainfall events (Figure 6A-D). During the later period of the study, the infiltration rate reduces when the pond area is small due to the clogging effect of clay, the nonlinear representation of infiltration in the model was able to account for this effect to some extent. The R^2 and NSE range between 0.78 and 0.93 and 0.64 and 0.85, respectively, indicating that the model was able to reproduce the observed surface area and depth (Table 2). In addition, the RMAE ranges from 0.10 to 0.23 (see Table 2 for the range) values were obtained for these sites for both the area and depth.

The model for habitats located in areas characterised by high infiltration and far from permanent sources of water underestimates the generation of ponds (Figure 6E-G). The model fails to simulate the intermittent drying of ponds as observed and also overestimates the area during the latter and fairly dry part of the study period for these sites. Regarding the dry period, the simulation developed ponding on day 216 due to a 15.3 mm rainfall event. However, no ponding was observed. A high infiltration rate seems to characterise these sites. This indicates that the model could be improved by incorporating a treatment of soil moisture, which would increase or decrease infiltration or runoff respectively, in the drier periods as the soil dries out.

Regarding sites 6 to 8, the R^2 and NSE values range between 0.77 and 0.95 and 0.51 and 0.86, respectively (Table 2). The RMAE range between 0.23 and 0.52 for both depth and area simulations.

The water table of the waterlogged areas penetrated the surface at the peak of the rainy season, an effect which is not simulated by the model. In this case, the infiltration becomes zero and water loss is

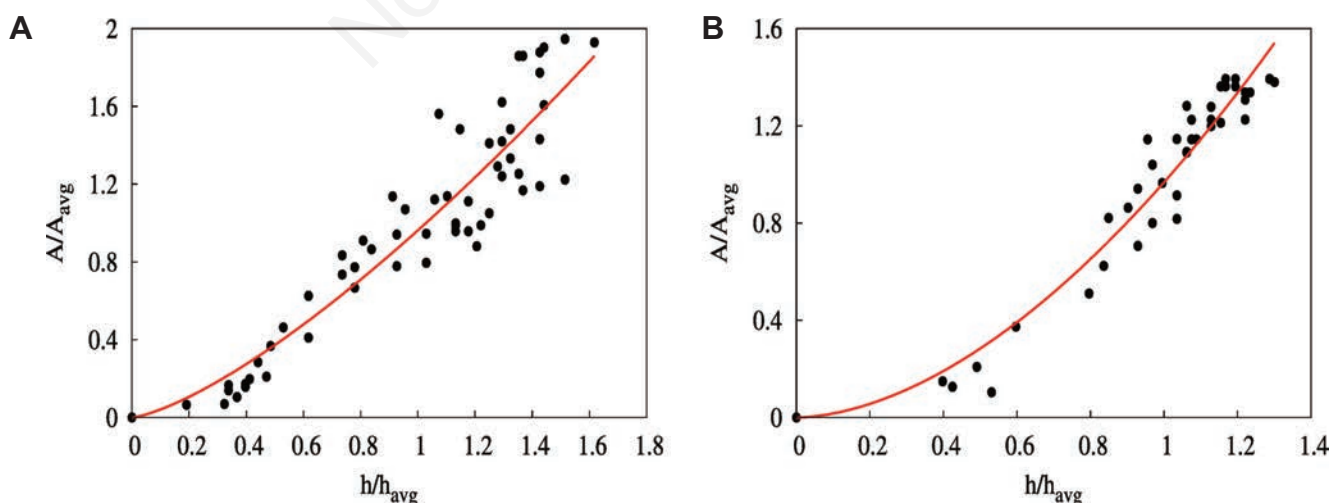


Figure 5. Example of power function fit for sites 5 (A) and 8 (B) using the area-depth relation. The estimated p values for all the sites are listed in Table 2.

governed only by evaporation. For these reasons, the model underestimates the surface area and depth of the water bodies at sites 9 and 10 (Figure 6 H and I) for the latter part of the study period. The R^2 and NSE values range between 0.87 and 0.95 and 0.57 and 0.79, respectively (Table 2). The RMAE values range between 0.08 and 0.15 (Table 2) for both area and depth simulations. Overall, the results from the presented prognostic geometrical model demonstrate the potential of the model to simulate daily and intraseasonal variability in surface water area and depth of individual ponds under different local hydrological conditions, except in water-logged situations where the loss terms are overestimated. These could be addressed by including a representation of soil moisture in the model, but at the expense of a considerable more complex model system.

Evaluation of the fraction model

It has been demonstrated that the prognostic water balance model for pond area presented in this work gives a good representation when validated with individual ponds. To evaluate the generalised pond fraction model, a value for the aggregate pond geometry \bar{p} is required. The value of \bar{p} is set using the best fit between daily average area and depth (Figure 7A), which gives $\bar{p}=1.4$, which is close to the mean value $\bar{p}=1.6$ or the 8 sites used to estimate the fractional water coverage.

Figure 7B shows the time series of simulated and estimated daily water fraction. The observed value is calculated by simply summing the pond fractions and assuming a value of A_{grid} equal to $100 A_{\text{max}}$. Instead the simulated value integrates eq. 11 forward in time implicitly, assuming $\bar{p}=1.4$. The fact that the fit is close to that achieved for the individ-

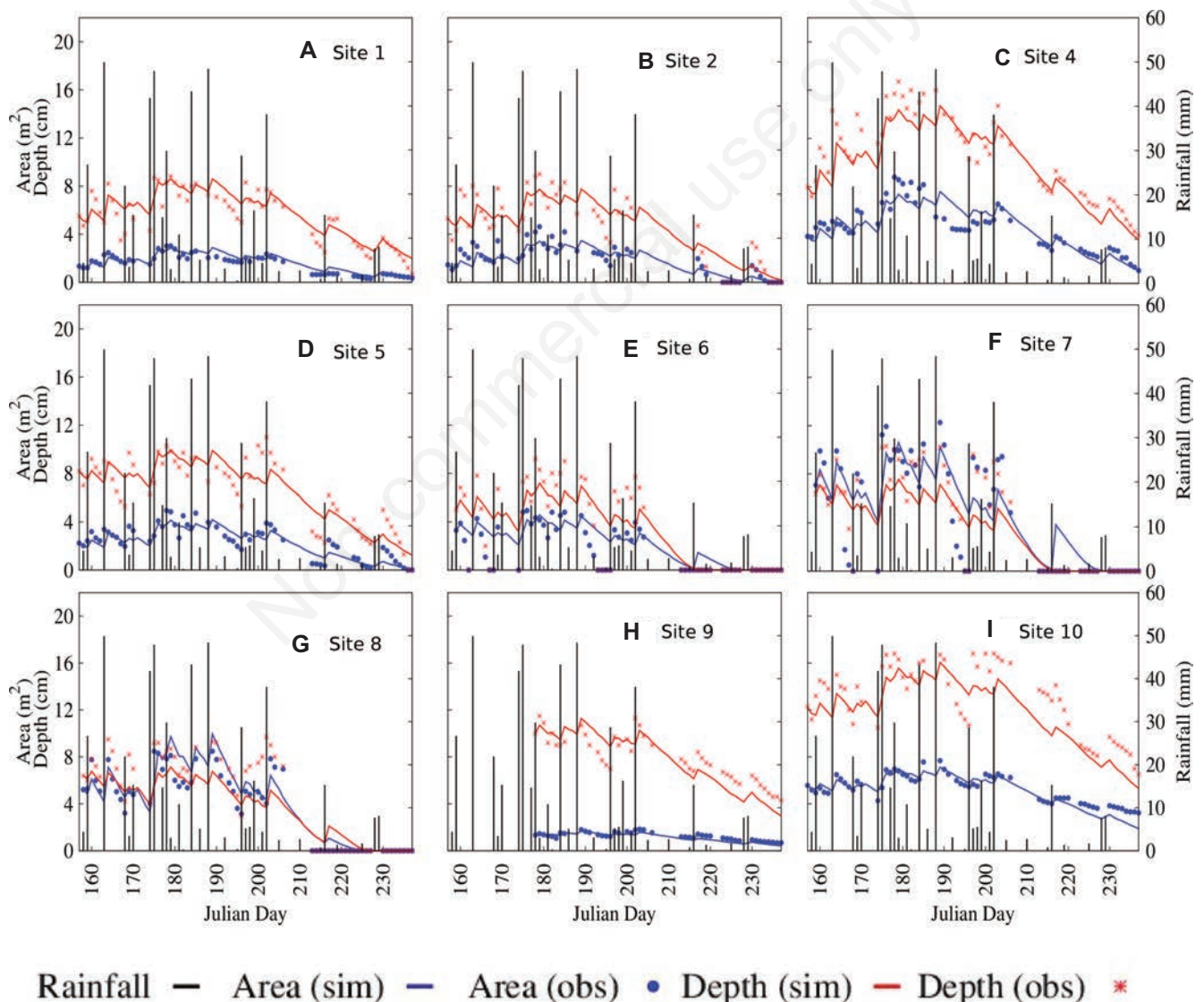


Figure 6. Comparison of daily observed and simulated pond water surface areas and water depth of temporary water bodies. Solid lines are simulated values of area (blue) and depth (red). Dots and crosses represent the observed values of the surface area and water depth, respectively. Sites names are shown under each plot.



ual ponds indicates that the approximation of assuming a generalised shape parameter across all individual ponds is a reasonable one. This demonstrates the useful application of the presented geometrical model for gridded malaria vector borne disease models to improve their representation of surface hydrology using a simple model.

Discussion

Surface hydrology is one key factor that controls the life-cycle of mosquito larvae and thus a treatment is required for dynamical disease models. A prognostic water balance model was developed using the geometrical assumptions of Hayashi to predict pond coverage and depth, with sources and sinks due to precipitation, runoff, evaporation and infiltration. The infiltration increased with pond coverage to represent the clogged inner regions of puddles. Based on a range of evaluation metrics, the prognostic model is judged to provide a good representation of the pond depth and coverage evolution under different hydrolog-

ical conditions. The single pond coverage model was then generalised to simulate the total fractional coverage of ponds in an area that may measure tens of kilometres, to be able to apply the model to gridded regional predictions. If an appropriate aggregate pond geometrical factor was selected, the model was also able to simulate the aggregate fractional coverage of the ponds. Thus the model could be used to improve the representation of surface hydrology in regional vector borne disease models. The model represented by eq. 11 has been implemented into the Vector-borne disease community model of the International Centre for Theoretical Physics, Trieste (VECTRI) malaria model and is available from v1.3.1 onwards.

Conclusions

During the later dry phase of the study period, the model tends to overestimate the pond area and depth for the sites located far away from permanent streams (sites 6 to 8), while underestimating within

Table 2. Computed p values, relative mean absolute error, coefficient of determination and Nash-Sutcliffe efficiency between observation and model output for various sites for both area and depth simulations.

Values	Sites									
	1	2	3	4	5	6	7	8	9	10
p	1.5	1.7	2.0	1.6	1.4	1.7	1.5	1.1	1.9	2.0
Area										
RMAE	0.15	0.23	0.12	0.15	0.21	0.52	0.40	0.23	0.15	0.08
R ²	0.93	0.90	0.90	0.89	0.78	0.77	0.83	0.95	0.90	0.95
NSE	0.85	0.78	0.71	0.78	0.79	0.51	0.68	0.86	0.57	0.79
Depth										
RMAE	0.13	0.23	0.12	0.10	0.16	0.44	0.34	0.31	0.13	0.13
R ²	0.91	0.89	0.86	0.93	0.89	0.80	0.87	0.93	0.92	0.87
NSE	0.81	0.79	0.64	0.80	0.78	0.64	0.70	0.74	0.63	0.59

R², the coefficient of determination; RMAE, relative mean absolute error; NSE, Nash-Sutcliffe efficiency.

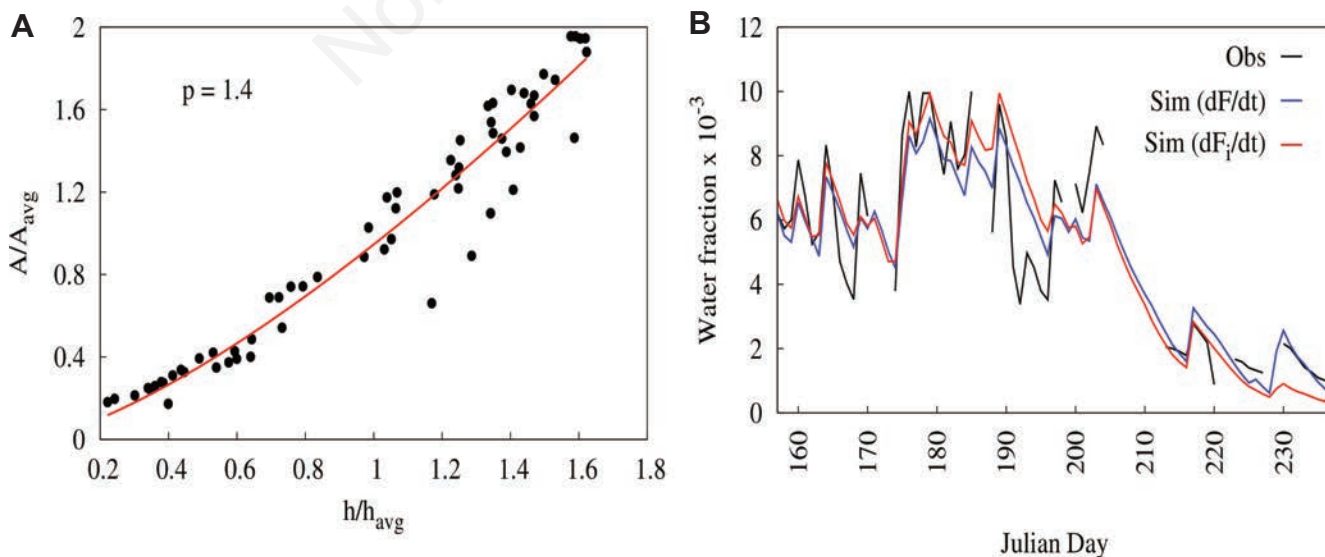


Figure 7. A) Power function fit for average of 8 sites (excluding sites 3 and 9), and B) comparison of simulated and estimated water fraction.

waterlogged areas (sites 9 and 10). This highlights that although the model was successful in representing the broad sub-seasonal evolution of ponds and thus represents the zero-order processes that determine pond fraction, further improvements might be achievable by including further secondary processes that are presently neglected. One process would be to incorporate a treatment of soil moisture, which would increase infiltration in the drier periods prior to the monsoon onset or as the soil dries out post-monsoon. This would also allow the model to distinguish between lower lying water-logged areas and areas with higher infiltration rates. However, accurate treatment of soil moisture is a challenge and would substantially increase the model complexity. For example, the agricultural model GLAM (Challinor *et al.*, 2004; Li *et al.*, 2016) represents the soil column with 25 soil layers, thus requiring an additional 25 prognostic equations to be solved. The specification of the soil properties is challenging due to sparsity of data. Moreover, topography and slope details would be beneficial, in particular to determine the pond catchments. However, there will always be processes that cannot be represented in the model. For example, shading and productivity of ponds is a function of the vegetation present, which would be impossible to determine for all ponds on a regional scale. Thus, it is clear that many of these secondary processes will have to be accounted for implicitly using calibration of the key pond model parameters. A generalised calibration technique to be applied in each grid-cell is the subject of current work and will be reported in a future paper.

References

- Awolola T, Oduola A, Obansa J, Chukwura N, Unyimadu J, 2007. *Anopheles gambiae* ss breeding in polluted water bodies in urban Lagos, southwestern Nigeria. J Vector Dis 44:241.
- Bayoh M, Lindsay S, 2003. Effect of temperature on the development of the aquatic stages of *Anopheles gambiae sensu stricto* (Diptera: Culicidae). B Entomol Res 93:375-82.
- Bayoh M, Lindsay S, 2004. Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito *Anopheles gambiae* in the laboratory. Med Vet Entomol 18:174-9.
- Bombliès A, 2012. Modeling the role of rainfall patterns in seasonal malaria transmission. Climatic Change 112:673-85.
- Bombliès A, Duchemin JB, Eltahir EAB, 2008. Hydrology of malaria: model development and application to a Sahelian village. Water Resour Res 44:W12445.
- Bombliès A, Duchemin JB, Eltahir EAB, 2009. A mechanistic approach for accurate simulation of village scale malaria transmission. Malaria J 8:223.
- Brooks RT, Hayashi M, 2002. Depth-area-volume and hydroperiod relationships of ephemeral (vernal) forest pools in southern New England. Wetlands 22:247-55.
- Challinor AJ, Wheeler TR, Craufurd PQ, Slingo JM, Grimes DIF, 2004. Design and optimisation of a large-area process-based model for annual crops. Agr Forest Meteorol 124:99-120.
- Chase JM, Knight TM, 2003. Drought-induced mosquito outbreaks in wetlands. Ecol Lett 6:1017-24.
- Chinery W, 1984. Effects of ecological changes on the malaria vectors *Anopheles funestus* and the *Anopheles gambiae* complex of mosquitoes in Accra, Ghana. Trop Med Int Health 87:75-81.
- Depinay JMO, Mbogo CM, Killeen G, Knols B, Beier J, Carlson J, Dushoff J, Billingsley P, Mwambi H, Githure J, Toure AM, McKenzie FE, 2004. A simulation model of African *Anopheles* ecology and population dynamics for the analysis of malaria transmission. Malaria J 3:29.
- Fillinger U, Sonye G, Killeen GF, Knols B, Becker N, 2004. The practical importance of permanent and semipermanent habitats for controlling aquatic stages of *Anopheles gambiae sensu lato* mosquitoes: operational observations from a rural town in western Kenya. Trop Med Int Health 9:1274-89.
- Gianotti RL, Bombliès A, Eltahir EA, 2009. Hydrologic modeling to screen potential environmental management methods for malaria vector control in Niger. Water Resour Res 45:W08438.
- Hayashi M, Van der Kamp G, 2000. Simple equations to represent the volume-area-depth relations of shallow wetlands in small topographic depressions. J Hydrol 237:74-85.
- Hayashi M, Van der Kamp G, 2007. Water level changes in ponds and lakes: the hydrological processes. In: Johnson EA, Miyanishi K, eds. Plant disturbance ecology: the process and the response. Elsevier Academic Press, Amsterdam, The Netherlands, pp 311-39.
- Himeidan YE, Zhou G, Yakob L, Afrane Y, Munga S, Atieli H, El-Rayah EA, Githeko AK, Yan G, 2009. Habitat stability and occurrences of malaria vector larvae in western Kenya highlands. Malaria J 8:234.
- Hoshen MB, Morse AP, 2004. A weather-driven model of malaria transmission. Malaria J 3:32.
- Imbahale SS, Paaajmans KP, Mukabana WR, van Lammeren R, Githeko AK, Takken W, 2011. A longitudinal study on *Anopheles* mosquito larval abundance in distinct geographical and environmental settings in western Kenya. Malaria J 10:81.
- Koenraadt C, Githeko A, Takken W, 2004. The effects of rainfall and evapotranspiration on the temporal dynamics of *Anopheles gambiae* ss and *Anopheles arabiensis* in a Kenyan village. Acta Tropica 90:141-53.
- Li S, Tompkins AM, Lin E, Ju H, 2016. Simulating the impact of flooding on wheat yield. Case study in East China. Agr Forest Meteorol 216:221-31.
- Manzanas R, Amekudzi L, Preko K, Herrera S, Gutierrez JM, 2014. Precipitation variability and trends in Ghana: an intercomparison of observational and reanalysis products. Climatic Change 124:805-19.
- Martin-Rosales W, Leduc C, 2003. Dynamiques de vidange d'une mare temporaire au Sahel: l'exemple de Banizoumbou (Sud-Ouest du Niger). C R Geosci 335:461-8.
- Minakawa N, Mutero CM, Githure JI, Beier JC, Yan G, 1999. Spatial distribution and habitat characterization of anopheline mosquito larvae in Western Kenya. Am J Trop Med Hyg 61:1010-6.
- Minakawa N, Sonye G, Mogi M, Yan G, 2004. Habitat characteristics of *Anopheles gambiae* ss larvae in a Kenyan highland. Med Vet Entomol 18:301-5.
- Minakawa N, Sonye G, Yan G, 2005. Relationships between occurrence of *Anopheles gambiae* sl (Diptera: Culicidae) and size and stability of larval habitats, 2005. J Med Entomol 42:295-300.
- Minke AG, Westbrook CJ, van der Kamp G, 2010. Simplified volume-area-depth method for estimating water storage of prairie pot-holes. Wetlands 30:541-51.
- Montosi E, Manzoni S, Porporato A, Montanari A, 2012. An eco-hydrologic model of malaria outbreaks. Hydrol Earth Syst Sc 9:2831-54.
- Munga S, Minakawa N, Zhou G, Barrack O, Githeko AK, Yan G, 2005. Oviposition site preference and egg hatchability of *Anopheles gambiae*: effects of land cover types. J Med Entomol 42:993-7.
- Mushinzimana E, Munga S, Minakawa N, Li L, Feng C, Bian L, Kitron U, Schmidt C, Beck L, Zhou G, Githeko AK, Yan G, 2006. Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands. Malaria J 5:13.
- Mutuku FM, Alaii JA, Bayoh MN, Gimnig JE, Vulule JM, Walker ED, Kabiru E, Hawley WA, 2006a. Distribution, description, and local



- knowledge of larval habitats of *Anopheles gambiae* sl in a village in western Kenya. *Am J Trop Med Hyg* 74:44-53.
- Mutuku FM, Bayoh MN, Gimnig JE, Vulule JM, Kamau L, Walker ED, Kabiru E, Hawley WA, 2006b. Pupal habitat productivity of *Anopheles gambiae* complex mosquitoes in a rural village in western Kenya. *Am J Trop Med Hyg* 74:54-61.
- Nash J, Sutcliffe J, 1970. River flow forecasting through conceptual models. Part I. A discussion of principles. *J Hydrol* 10:2 82-90.
- Ndenga BA, Simbauni JA, Mbugi JP, Githeko AK, Fillinger U, 2011. Productivity of malaria vectors from different habitat types in the western Kenya highlands. *PLoS One* 6:e19473.
- Owusu K, Waylen P, 2009. Trends in spatio-temporal variability in annual rainfall in Ghana (1951- 2000). *Weather* 64:115-20.
- Porphyre T, Bicout D, Sabatier P, 2005. Modelling the abundance of mosquito vectors versus flooding dynamics. *Ecol Model* 183:173-81.
- Sattler MA, Mtasiwa D, Kiama M, Premji Z, Tanner M, Killeen GF, Lengeler C, 2005. Habitat characterization and spatial distribution of *Anopheles* sp. mosquito larvae in Dar es Salaam (Tanzania) during an extended dry period. *Malaria J* 4:4.
- Soti V, Puech C, Seen DL, Bertran A, Vignolles C, Mondet B, Dessay N, Tran A, 2010. The potential for remote sensing and hydrologic modelling to assess the spatio-temporal dynamics of ponds in the Ferlo Region (Senegal). *Hydrol Earth Syst Sc* 14:1449-64.
- Srivastava A, Nagpal B, Subbarao RSS, 2001. Predictive habitat modelling for forest malaria vector species *An. dirus* in India. A GIS-based approach. *Curr Sci* 80:1129-34.
- Sunahara T, Ishizaka K, Mogi M, 2002. Habitat size: a factor determining the opportunity for encounters between mosquito larvae and aquatic predators. *J Vector Ecol* 27:8-20.
- Tompkins AM, Ermert V, 2013. A regional-scale, high resolution dynamical malaria model that accounts for population density, climate and surface hydrology. *Malaria J* 12:65.

Non commercial use only