Simulations and Predictions of Mosquito Populations in Rural Africa Using Rainfall Inputs from Satellites and Forecasts

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
SEP 29 2010
LIBRARIES
ARCHIVES

by

Teresa K. Yamana

S. B. Environmental Engineering Science
Massachusetts Institute of Technology (2004)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of
Master of Science in Civil and Environmental Engineering

at the

Massachusetts Institute of Technology
September 2010

© 2010 Massachusetts Institute of Technology. All rights reserved.

Signature of Author: _____
Department of Civil and Environmental Engineering
July 9, 2010

Certified by: _____
Elfatih A. B. Eltahir
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by: _____
Daniele Veneziano
Professor of Civil and Environmental Engineering
Chairman, Departmental Committee for Graduate Students

Simulations and Predictions of Mosquito Populations in Rural Africa Using Rainfall Inputs from Satellites and Forecasts

by

Teresa K. Yamana

Submitted to the Department of Civil and Environmental Engineering
on July 9, 2010 in partial fulfillment of the requirements for the degree of Master of Science in Civil and Environmental Engineering

ABSTRACT

This thesis describes studies on the use of the Hydrology, Entomology and Malaria Transmission Simulator (HYDREMATS) developed and tested against field data by Bomblies *et al.* (2008) in simulating and predicting the potential for malaria transmission in rural Africa. The first study examined the temporal resolution of rainfall input required by HYDREMATS. Simulations conducted over Banizoumbou village in Niger showed that for reasonably accurate simulation of mosquito populations, the model requires rainfall data with at least 1 hour resolution. The second study investigated whether HYDREMATS could be effectively forced by satellite based estimates of rainfall instead of ground based observations. The CPC Morphing technique (CMORPH) (Joyce *et al.*, 2004) precipitation estimates distributed by NOAA are available at a 30-minute temporal resolution and 8 km spatial resolution. We compared mosquito populations simulated by HYDREMATS when the model is forced by adjusted CMORPH estimates and by ground observations. The results indicate that adjusted CMORPH rainfall estimates can be used with HYDREMATS to simulate the dynamics of mosquito populations and malaria transmission with accuracy similar to that obtained when using ground observations of rainfall. The third study tested the ability of HYDREMATS to make short term predictions about mosquito populations. A method was developed by which the rainfall forcing for HYDREMATS is constructed to suit a prediction mode. Observed rainfall is used up until the date of the prediction. The rainfall for the following two weeks (or four weeks) is assumed to be the seasonal mean for that period. HYDREMATS predictions using this method were not significantly different from simulations using observed data.

Thesis Supervisor: Elfatih A. B. Eltahir

Title: Professor of Civil and Environmental Engineering

Acknowledgements

Financial support for this project was provided by the U.S. National Science Foundation.

I would like to thank my advisor, Elfatih Eltahir, for his guidance and encouragement in all aspects of this project, from my decision to apply to graduate school, to the final revisions of this thesis.

I am grateful to Arne Bomblies for sharing his extensive knowledge and enthusiasm about malaria, mosquitoes, and HYDREMATS. I would also like to thank my colleagues in the Eltahir research group and Parson's Lab, especially Gayle Sherman.

I am thankful for the constant support I received from Kabir Mukaddam, and my friends and family. I am thankful to my parents and all my teachers for instilling in me a love of learning, and for helping me to achieve my goals.

Finally, I would like to acknowledge Ron Rivera, a dear friend and mentor, who passed away after acquiring malaria while setting up a water-filter factory in Nigeria. Ron, along with Susan Murcott and Kate Ricke, were instrumental in developing my desire to work towards public health in the developing world. Ron was a true inspiration, both personally and professionally. His untimely passing in the early months of my research added a very personal aspect to my motivation to conduct this work.

Contents

List of Figures

List of Tables

1 Introduction

1.1 *Background*

Malaria is responsible for nearly a million deaths each year, 90% of which occur in Africa (Aregawi *et al.*, 2008). The disease is caused by the plasmodium parasite, which is transmitted primarily by *Anopheles* mosquitoes. In arid areas such as the Sahel, malaria is closely linked to rainfall, as the mosquitoes that transmit the disease are limited by the availability of vector breeding habitat. In Niger, the primary malaria vector is *Anopheles gambiae,* which breeds in temporary rain-water fed pools on the order of tens of meters in diameter. The formation of these pools and their utilization by mosquitoes is mechanistically modeled by the Hydrology, Entomology and Malaria Transmission Simulator (HYDREMATS), developed by Bomblies *et al.* (2008). HYDREMATS has been used in a number of studies of malaria in Banizoumbou and Zindarou villages in Niger. (Bomblies *et al.*, 2008; Bomblies & Eltahir, 2009; R. Gianotti *et al.*, 2008; R. L. Gianotti *et al.*, 2008).

1.2 *Model Description*

The development of the Hydrology, Entomology and Malaria Transmission Simulator (HYDREMATS) is described in detail in Bomblies *et al.* (2008). The model was developed to simulate village-scale response of malaria transmission to interannual climate variability in semi-arid desert fringe environments such as the Sahel. The model provides explicit representation of the spatial determinants of malaria transmission. HYDREMATS can be

separated into two components: the hydrology component which explicitly represents pooled water available to anopheles mosquitoes as breeding sites, and the entomology component, which is an agent-based model of disease transmission.

The hydrology component of HYDREMATS is based on the land surface scheme LSX of Pollard and Thompson (1995). The model simulates momentum, energy and water fluxes within its vertical column of the atmosphere, six soil layers and two vegetation layers. Vegetation type and soil characteristics are required as model inputs, and strongly influence soil moisture and runoff in the model. Thicknesses and permeabilities of vertical soil layers are assigned to represent the soil structure observed in the Sahel, including the thin layer of low-permeability crust commonly observed in areas with sparse vegetation (Bomblies *et al.*, 2008).

Water at each grid cell is partitioned between runoff and infiltration, based on a Hortonian runoff process governed by hydraulic conductivity and porosity of the soil. Unsaturated zone hydraulic conductivity is calculated as a function of soil moisture following Campbell's equation. Infiltration through the unsaturated zone is calculated using an implicit Richard's equation solver. Uptake of soil water from evapotranspiration is calculated based on climactic variables (Bomblies *et al.*, 2008). Richard's equation is:

$$\frac{\partial \theta(z,t)}{\partial t} = \frac{\partial}{\partial z}\left[K_u(\theta)\frac{\partial \phi(\theta,z)}{\partial z} + K_u(\theta) \right]$$

where θ = soil moisture [$cm^3\ cm^{-3}$]

$K_u(\theta)$ = unsaturated hydraulic conductivity [$m\ sec^{-1}$]

$\phi(\theta,z)$ = head value [m]

z = elevation [m]

Overland flow is modeled using a finite difference solution of a diffusion wave approximation to the St. Venant equations following the formulation of Lal (1998). Flow velocity is represented by Manning's equation as a function of friction slope, flow depth, and the distributed roughness parameter n, which is derived from soil characteristics and vegetation type. The overland flow process is of critical importance for the modeling of water pool formation (Bomblies *et al.*, 2008).

The continuity equation for shallow flow is:

$$\frac{\partial h}{\partial t}+\frac{\partial(hu)}{\partial x}+\frac{\partial(hv)}{\partial y}-P+I+ET=0$$

where u and v are the flow velocities in the x and y directions, respectively, h is the water depth, P is precipitation, I is infiltration, and ET is evapotranspiration.

The momentum equations for the x and y directions are approximated as:

$$\frac{\partial H}{\partial x}=-S_{fx}$$

$$\frac{\partial H}{\partial y}=-S_{fy}$$

Where S_{fx} and S_{fy} are the friction slopes in the x and y directions, respectively, and H is the water depth (h) plus a reference elevation (z).

Manning's equation is:

$$u=\frac{1}{n}h^{\frac{2}{3}}S_{fx}^{\frac{1}{2}}$$

where *n* is the Manning's roughness coefficient which determines resistance to overland flow. These equations are solved using the alternate-direction implicit (ADI) method described by Lal (1998).

The meteorological inputs required by the model are temperature, humidity, wind speed and direction, incoming solar radiation, and rainfall. These variables are assumed to be uniform over the model domain in the simulations conducted in this thesis. Distributed rasters of vegetation, soil type, and topography are required at the grid resolution specified by the user. The hydrology component of HYDREMATS generates a grid of water depths and temperatures for each grid cell, for each timestep. These grids serve as the inputs for the entomology component of the model (Bomblies *et al.*, 2008).

The entomology component of HYDREMATS simulates individual mosquito and human agents. Human agents are immobile, and are assigned to village residences, as malaria transmission in this region occurs primarily at night when humans are indoors (Service, 1993). Mosquito agents have a probabilistic response to their environment based on a prescribed set of rules governing dispersal and discrete events including development of larval stages, feeding, egg-laying and death. The model tracks the location, infective status and reproductive status of each female mosquito through time. The malaria parasite is transmitted when a mosquito bites an infected human, and takes a second bloodmeal from an uninfected human (Bomblies *et al.*, 2008).

In addition to the water pool inputs supplied by the hydrology component of the model, the entomology component requires air temperature, humidity, wind speed and wind direction. Air temperature and relative humidity influence mosquito behavior and survival, while wind speed and direction influence mosquito flight, both by physical displacement by wind, and by attracting mosquitoes to upwind blood sources. The location of village residences is required in order to assign the location of human agents (Bomblies *et al.*, 2008).

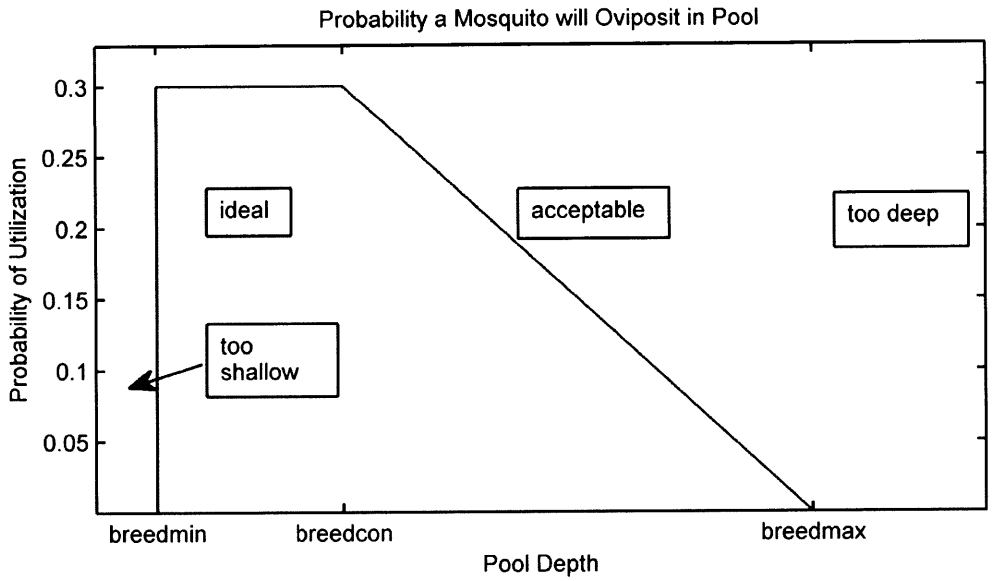

Figure 1.1 Probability of Utilization Given Pool Depth

Each simulated female mosquito ready to oviposit begins to seek water. At each time step, she finds herself a new grid point. The probability that she will deposit her eggs at that point is called the probability of utilization, and is a function of the depth of water at that point. The probability of utilization is calculated at every time step and each grid point based on the water depth at that point, and serves as the primary linkage between the hydrology and the entomology components of HYDREMATS. The water depth can fall into 4 categories: too shallow or dry, ideal, acceptable, or too deep. If a grid cell has less than a minimum depth of water, *breedmin*, the probability of utilization is zero. Ideal conditions for mosquitoes are shallow areas, often at the fringes of pools. Grid cells with water depth between the minimum acceptable value up to a certain point, *breedcon*, are considered to be ideal, and are assigned the maximum probability of utilization. Beyond this depth, the pool begins to have waves and be less appealing to mosquitoes and the probability of utilization decreases linearly to zero until

it reaches the maximum acceptable depth, *breedmax*. The relationship between pool depth and probability of utilization is shown in Figure 1.1. The probability of utilization does not take into account the probability that a mosquito will land on the grid point in question. The likelihood that a mosquito would arrive at given grid point would be a function of the pool's location relative to humans, wind speed and wind direction.

Mosquito eggs hatch and advance through four stages of larval development at rates dependent on water temperature, nutrient competition and predation given by Depinay et al. (2004). Surviving larvae pupate and emerge as adult mosquitoes. The duration of the aquatic stage of *Anopheles gambiae* mosquitoes ranges from 8 to 24 days (Depinay *et al.*, 2004). All aquatic stage mosquitoes in a pool that dries up are killed, emphasizing the importance of pool persistence for mosquito breeding(Bomblies *et al.*, 2008).

Adult female mosquitoes follow a cycle of seeking human bloodmeals, feeding, resting, and ovipositing for the duration of their lifespan. Mosquito flight velocity is assigned as a weighted random walk corrected for attraction to upwind CO_2 and wind influence. The effective flight velocity, which incorporates resting time and direction changes within the model time-step, is assumed to follow a normal distribution with mean 15 m/hr and variance 25 m/hr. Mortality of adult mosquitoes is a function of daily average temperature, with no survival above a daily average temperature of 41°C. The model outputs for each time step includes the number of live adult mosquitoes, their location and infective status, and the prevalence of malaria infections in humans (Bomblies *et al.*, 2008).

There are a number of differences between the parameterization of HYDREMATS in the simulations conducted in this thesis, and those conducted by Bomblies *et al.* (2008). These

parameters describe the behavior and characteristics of adult and larval stage mosquitoes, and their exact values are not known. They were changed in order to produce a better fit to observed mosquito data. All simulations in this thesis include the representation of bednets, or insecticide treated nets (ITNs), while Bomblies et al. do not represent ITNs in simulations conducted in 2005, as the nets were distributed in Banizoumbou in late 2005. Additionally, Bomblies' simulations of Banizoumbou in 2005 runs from July 1st through October 31st, while the simulations for 2005 conducted here cover the period between June 1st and October 31st. A summary of the differences between the two formulations are presented in Table 1-1.

Table 1-1 Differences in parameterization of HYDREMATS

Variable	Bomblies	Yamana	units
average mosquito flight velocity	15	25	m hr^{-1}
resting time	24	48	hr
weighting of random walk vs. straight line flight	0.2	0.7	
utilization probability of water	0.95	0.3	
probability that a mosquito takes a bloodmeal	0.7	0.6	
predation of aquatic stage mosquitoes	yes	no	
carrying capacity of pools	300	100	mg m^{-2}
Bed-nets represented	2006 only	all years	
Simulation time in 2005	Jul-Oct 2005	Jun-Oct 2005	

1.3 ***Study Location***

The simulations in this thesis, as with several of past studies of the Eltahir Research Group using HYDREMATS (Bomblies *et al.*, 2008; R. Gianotti *et al.*, 2008; R. L. Gianotti *et al.*, 2008), were conducted over the domain of our field site, Banizoumbou village in southwestern Niger shown in Figure 1.2. Banizoumbou is a typical Sahelian village in a semi-arid landscape,

with a population of roughly 1000. Land cover consists of tiger bush shrubland, millet fields, fallow and bare soil. Millet fields dominate near the village, while tiger bush is more common near surrounding plateau tops (Bomblies *et al.*, 2008).

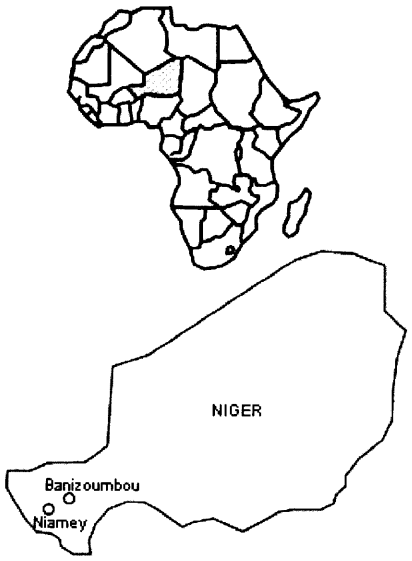

Figure 1.2 Location of Banizoumbou, Niger

The rainy season in Banizoumbou extends from May to October, with maximum rainfall occurring in August. During this time, water pools form in and around the village, providing ideal breeding habitat for *Anopheles gambiae* mosquitoes. Mosquito populations and malaria transmission in Niger increase dramatically during the rainy season, as seen in Figure 1.4, where monthly rainfall is shown by the solid blue line, and weekly malaria incidences in Niamey, Niger for 2001, 2002 and 2003 are shown by the dotted lines on the left axis (Bomblies *et al.*, 2008).

Malaria prevalence and mosquito density data are collected in Banizoumbou by our collaborators at the Centre de Recherche Medicale et Sanitaire, Niamey, Niger (CERMES).

Meteorological data for Banizoumbou is available through the African Monsoon Multidisciplinary Analyses (AMMA).

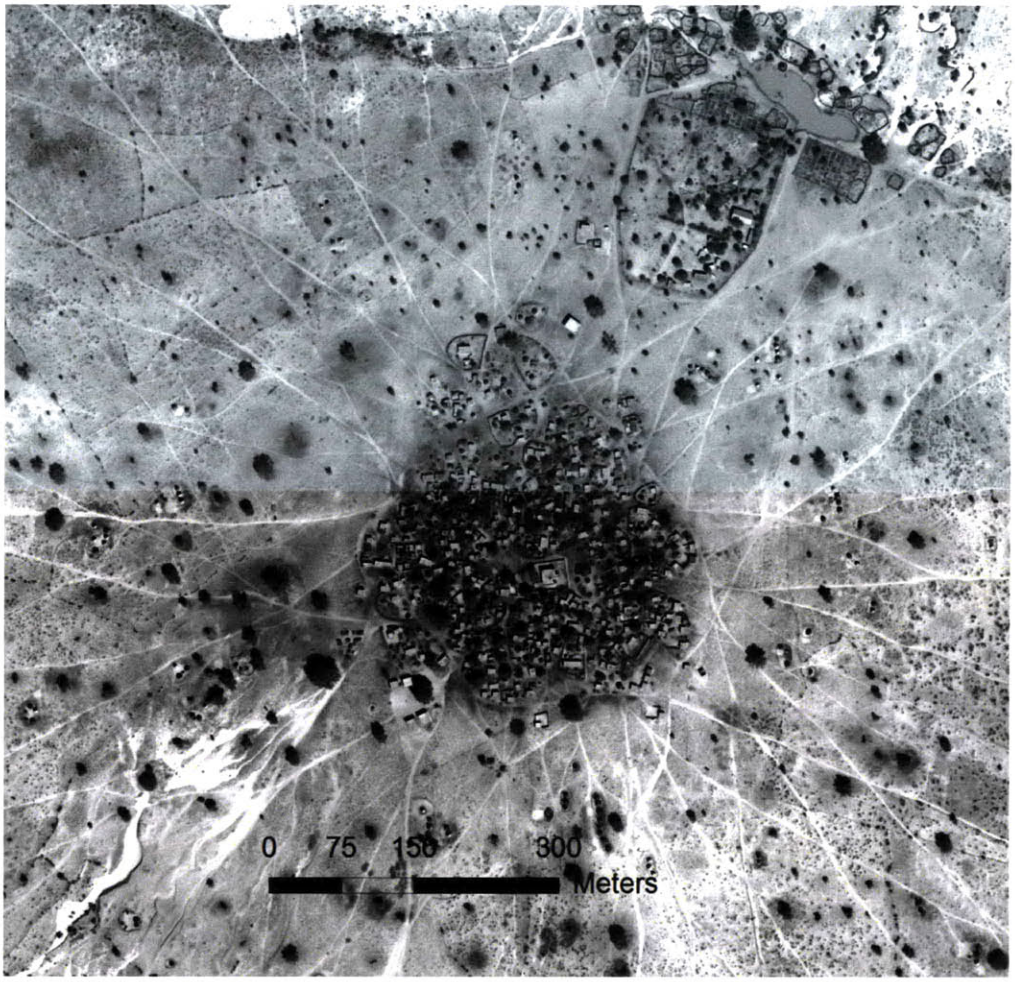

Figure 1.3 Quickbird image of Banizoumbou Village (copyright DigitalGlobe Incorporated, 2005)

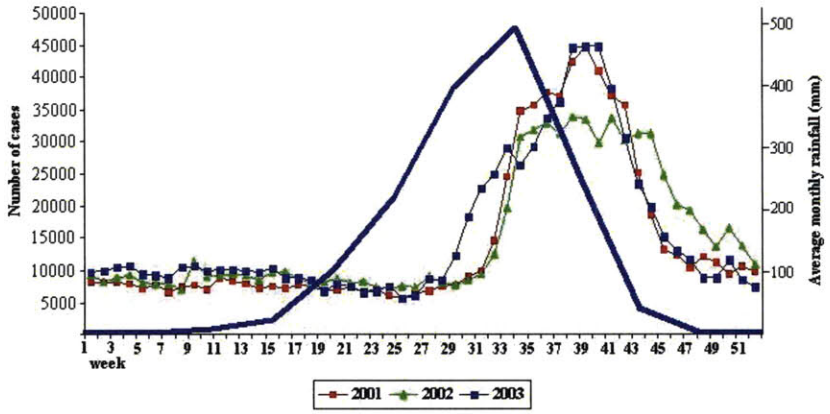

Figure 1.4 Relationship between malaria incidence and seasonal rainfall in Niger

1.4 ***Research Questions***

HYDREMATS is a powerful tool with great potential to assist public health officials in formulating malaria control programs in Western Africa. A main limitation of the model is the high amount of input data needed in order to properly simulate the dynamics of malaria transmission. The aim of this thesis was to investigate methods to make HYDREMATS more accessible for use in the field.

Rainfall data is a critical input variable for HYDREMATS. Because pool formation is modeled explicitly using overland flow and infiltration models, information about the distribution and intensity of rainfall events as important as the total amount of rainfall. For this reason, the model can be sensitive to the temporal resolution of rainfall. In Chapter 2, the response of HYDREMATS to different resolutions of rainfall was examined in order to determine the minimum requirements for rainfall data needed in order to accurately simulate the variability in mosquito populations.

Chapter 3 investigates whether satellite based estimates of rainfall can be used with HYDREMATS in areas where ground observations are not available. This would greatly expand

the range of applicability of the model, as many areas at risk of malaria transmission lack the ground observations that would be needed by HYDREMATS.

Chapter 4 investigates whether HYDREMATS can be used to make short term predictions of malaria transmission patterns. If the model could forecast sudden increases in malaria transmission, local officials could take precautions to reduce risk of infection.

In Chapter 5, we consider the possibility of running HYDREMATS using exclusively satellite and archived data sets so that the model could be used anywhere that malaria transmission is influenced by water availability.

2 The Effect of the Temporal Resolution of Precipitation on Simulations of Mosquito Habitat using HYDREMATS

2.1 *Introduction and Literature Review*

The connection between rainfall and malaria transmission has been observed for nearly a century (M. C. Thomson *et al.*, 1996). Mosquitoes of the *Anopheles gambiae* complex, the primary malaria vector in Africa, breed in small, temporary pools of standing water. Increased rainfall often means that there are increased breeding sites available to female mosquitoes, which leads to an increased mosquito population, and thus malaria transmission. Numerous studies, reviewed below, have demonstrated correlations between rainfall levels and malaria transmission at varying temporal and spatial scales. These studies used rainfall inputs with temporal resolutions of ten days to 3 months.

Several studies use statistical methods to relate rainfall to malaria risk. Craig *et al.* (1999) developed a climate-based tool for determining malaria risk on the continental scale, using monthly rainfall and temperature values. By comparing the monthly rainfall and temperature values to malaria transmission in 20 sites across Africa, they found threshold values of temperature and rainfall required for various levels of transmission. They reported that stable transmission occurs in areas where rainfall was at least 80 mm per month for more than five months, and mean temperature is above 22°C. Kilian et al. (1999) noted a correlation between mean monthly rainfall and malaria incidence with a two-month time lag between 1995 and 1998 in western Uganda. Thomson et al. (2005) used stepwise regression to relate seasonal rainfall totals between December and February and malaria incidence over a 22-year

period in Botswana. They found a strong correlation, with $r^2 > 0.66$. These correlation-based approaches do not directly model the relationship between rainfall and mosquito populations, and therefore do not require detailed information on the temporal distribution of rainfall.

Some studies use more sophisticated models to relate rainfall to mosquito populations. Patz *et al.* (1998) demonstrated the value of including hydrological information in predicting the human biting rate (HBR) of anopheles mosquitoes using a water balance model that took daily rainfall inputs, hourly temperature and dew point data, and output soil moisture. With a two-week time lag, the r^2 value between soil moisture and HBR was 0.45, compared with 0.08 when using raw rainfall. Hoshen and Morse (2004) compiled daily rainfall data into dekadal values, the sum of rainfall during the previous 10 days. Their weather-based model relates the dekadal rainfall to mosquito oviposition and larval survival rates.

While an increase in rainfall generally leads to an increase in mosquitoes, this is not always the case. Excess rainfall can flood the small pools and disrupt developing larvae. A decrease in rainfall can also increase breeding habitats by slowing flowing water, as observed when pools formed along a river bed in Sri Lanka during a prolonged drought (Wijesundera Mde, 1988). These processes can only be represented in a mechanistic model such as HYDREMATS, which explicitly represents the small pools in which mosquitoes breed, as well as the processes by which these sites are formed and emptied.

While the aforementioned studies relate rainfall to malaria transmission with a temporal resolution on the order of days, weeks, or even months, hydrology models require a much finer time scale. The main hydrological process of interest is surface runoff, as this is an important process affecting the mechanism by which water forms pools of water that can be

used by anopheles mosquitoes. Several studies have examined the effects of the temporal resolution of rainfall on the calculations of runoff models.

The following studies investigated the effect of averaging 5 or 6-minute resolution rainfall by up to an hour on the hydrologic response of distributed runoff models. Krajewski *et al* (1991) compared runoff generated a distributed hydrologic model using 5 minute resolution rainfall data to that using hourly rainfall rates constructed by averaging the fine resolution data. They also examined the runoff response to the spatial resolution of rain gauges, and determined that for their basin size of 7.5 km^2, the model was more sensitive to temporal resolution than to spatial resolution of rainfall, but that the hydrograph characteristics using the hourly rainfall did not differ significantly to that using the 5-minute data. Similarly, Pessoa *et al.* (1993) found that degrading 5-minute rainfall data to 15-minute and 30-minute averages had minimal effect on the hydrograph predicted by a distributed basin simulation model. However, Winchell *et al* (1998) saw significant differences in runoff when 6-minute rainfall data from NEXRAD was aggregated it into 24 minute, 42 minute and 60 minute data sets. They developed a rainfall-runoff model that could separately calculate infiltration-excess and saturation-excess runoff. They found that the coarsening of rainfall data did not affect saturation-excess runoff, but that it caused a decrease in infiltration-excess runoff. This decrease varied by individual storms, but could be as high as 58%.

The effect of averaging rainfall into timescales greater than 1-hour was investigated by Finnerty *et al.* (1997) compared the runoff simulated using 1-hour, 3-hour and 6-hour rainfall resolutions at a number of spatial scales. The model was most sensitive to temporal resolution at fine sub-basin scales. They found that with a basin size of 32 km x 32 km, the hourly rainfall

inputs cause 21% more surface runoff and 20% more interflow than the 6-hour inputs. They concluded that this reduction of runoff was the result of smoothing the short and high-intensity rainfall events that are important to runoff production into longer but lower-intensity events.

Until now, HYDREMATS has relied upon rain gauges for its rainfall inputs (Bomblies *et al.*, 2008; Bomblies *et al.*, 2009; Bomblies & Eltahir, 2009; R. Gianotti *et al.*, 2008; R. L. Gianotti *et al.*, 2008). However with emerging satellite data products with fine spatial and temporal resolutions, it may be possible to expand the range of HYDREMATS to areas with sparse ground-measurement data by using satellite data. Because a hydrology model considers the duration, intensity, and frequency of rainfall events in addition to total volume of rainfall in calculating pool availability and persistence, it is important to know what temporal resolution of rainfall is required to properly model the pool formation process. The aim of this study is to determine the minimum temporal resolution needed by HYDREMATS to model pool formation, mosquito populations and malaria transmission in West African terrain. This information will allow us to determine which, if any, satellite products have sufficient temporal resolution to be used with the model.

2.2 *Simulation Description*

The meteorological station in Banizoumbou is owned and operated by Institut de Recherche pour le Développement (IRD). Rainfall data are collected by a tipping bucket rain gauge at 5-minute intervals. To examine the model's sensitivity to the temporal resolution of rainfall, the 5-minute rainfall data during the rainy season of the year 2006 was aggregated to

construct rainfall data sets with 15-minute, 30-minute, 1-hour, 3-hour, 6-hour, 12-hour and 24-hour sampling resolutions. HYDREMATS simulations were conducted for each rainfall data set, operating at a 15-minute time step, from June 1, 2006 through October 31, 2006. The total rainfall measured during this time was 505.55 mm. The 15-minute resolution rainfall measurements were assumed to be the truth, and are shown in Figure 2.1, and the simulation using 15 minute rainfall resolution served as the control simulation. Simulations were not conducted using the original 5-minute rainfall resolution, due to the computational time requirements of using a 5-minute time step. All other meteorological inputs, temperature, humidity, wind speed and direction, and radiation, had a 1-hour resolution, and did not vary between simulations. The model domain was a 2.5 km x 2.5 km area centered over Banizoumbou village.

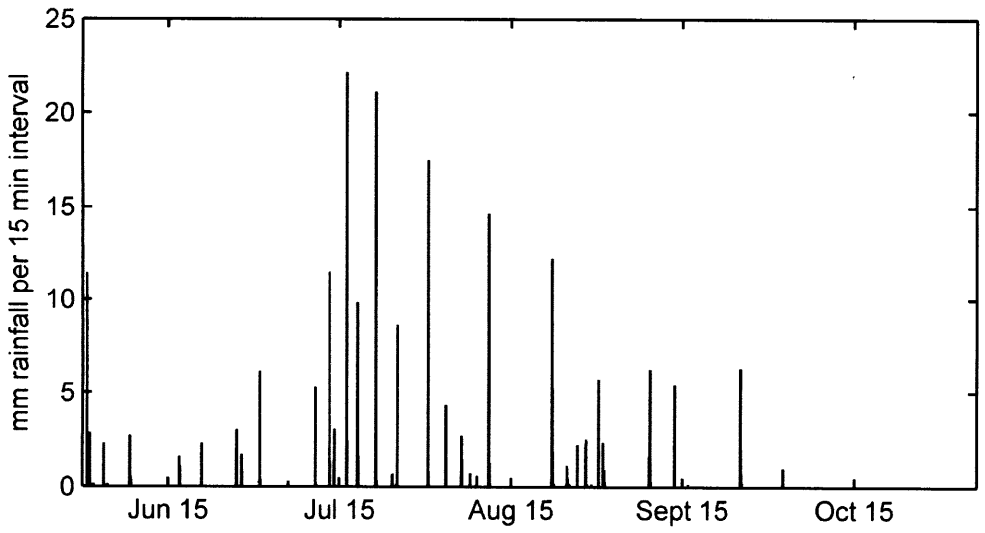

Figure 2.1 Rainfall in Banizoumbou, Niger, 2006

2.3 *Results*

2.3.1 Effect on Water Pools

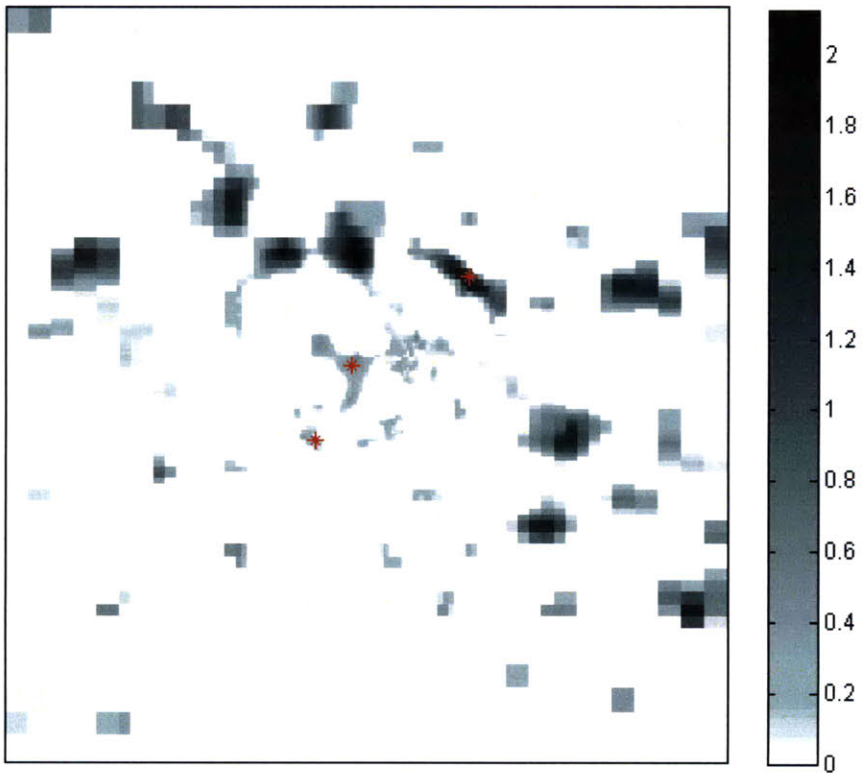

Figure 2.2 Pool depth output in meters for the 2.5 x 2.5 km Banizoumbou domain

HYDREMATS output rasters of pool water depths for each time step can be compared to assess the effect of reducing temporal resolution on the modeling of pool formations. Figure 2.2 shows a sample raster of pool depths for the 2.5 km by 2.5 km model domain. Each grid point has dimensions of 10 m by 10 m. The three red points correspond to three grid points for which a time series of pool depths is simulated by the model and examined. The top point corresponds to a large pool on the outskirts of the village, which we will call Pool 1. Mosquito larvae are not generally found in pools of this size, as they prefer shallower and calmer waters.

The middle point corresponds to a pool found in the center of Banizoumbou village, Pool 2, and is shown in Figure 2.3. This pool is a productive mosquito breeding site, which, combined with its proximity to households, makes it a significant public health concern. The third point represents Pool 3, another typical pool where mosquito larvae are found.

Figure 2.3 Pool 2 in the center of Banizoumbou Village

Figure 2.4 shows water depth Pool 1, Pool 2, and Pool 3 on the first day of each week throughout the rainy season. For each of the three pools, we see a good agreement between the water depths computed in the control 15 minute rainfall resolution simulation and the 30 minute and 1 hour rainfall resolution simulations. However, coarser time resolutions result in decreased water depths. A similar pattern is observed when we consider the total mass of water contained in pools found in the entire model domain, as summarized in Table 2-1. There was a very high correlation (correlation coefficient > 0.97) of the mass of pooled water

at each 15 minute timestep between the 30 minute, 1 hour and 3 hour rainfall resolution scenarios and the control scenario. However quantity of water stored in pools at each time step decreases as rainfall resolution decreases further. Despite the high correlation of mass of pooled water, the cumulative sum of pooled water over the entire rainy season is nearly 20% less than the control in the 1 hour simulation, and nearly 50% less than the control in the 3 hour simulation. Averaging rainfall over 24 hours results in a cumulative total of pooled water of only 0.3% of the amount in the control simulation, with a correlation coefficient of less than 0.005.

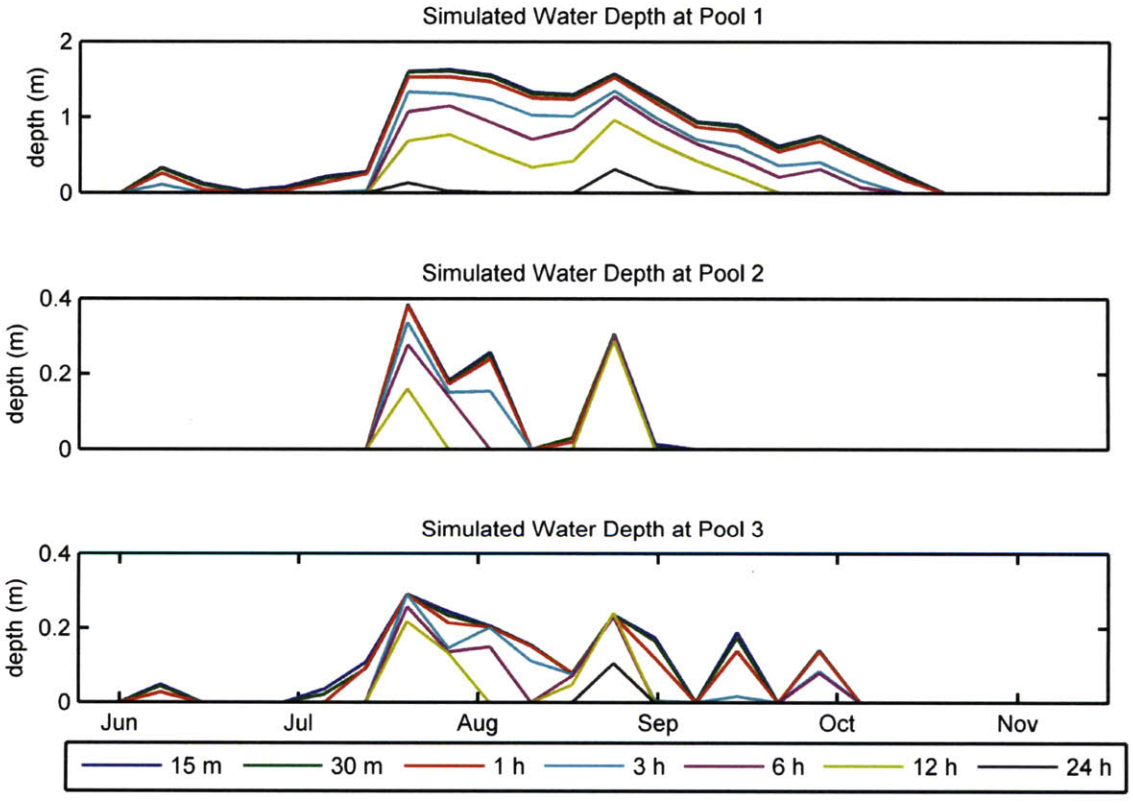

Figure 2.4 Weekly Water Depths at Pools 1-3

Another important characteristic of water pools is their surface area, as larvae are limited to the water surface. The total surface area of water pools over the model domain can be calculated by counting the number of grid points with water depth above a threshold value. The scatter plots in Figure 2.5 compare the proportion of the model domain covered in pools at the beginning of each day in the control simulation on the x-axis to that in the 30 minute, 1 hour, 3 hour and 6 hour simulations on the y-axis. The figure shows that daily outputs of surface area in the 30 minute and 1 hour scenarios correspond well with the control, while the 3-hour and 6-hour resolution scenarios underestimate the surface area of water pools. This is consistent with the decreasing volume of pooled water observed in these coarse-resolution scenarios.

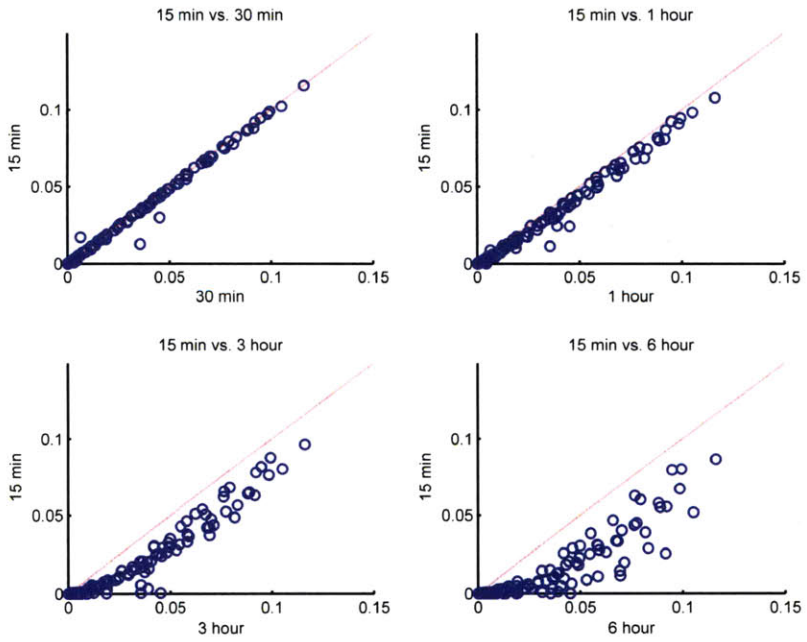

Figure 2.5 Daily Proportion of Pooled Area

Table 2.1 summarizes the comparison between the control simulation and the coarser resolution simulations. As with the mass of pooled water, we see that the surface area of water pools decreases as the temporal resolution of rainfall coarsens. The correlation coefficient between the simulations and control decreases at a faster rate for surface area than it does for water mass, decreasing to 0.91 for the 1-hour simulation and 0.80 for the 3-hour simulation.

Table 2-1 Volume and Surface Area of Water Pools

	15 min	30 min	1 hour	3 hour	6 hour	12 hour	24 hour
Mass of Pooled Water							
Cumulative Mass (10^5 kg)	1.36	1.31	1.11	0.70	0.43	0.16	0.00
Percent of 15 min	100	96	82	51	31	11	0.3
Correlation Coefficient to 15 min	1.00	1.00	1.00	0.97	0.89	0.72	0.34
Surface Area of Pooled Water							
Cumulative Surface Area (km^2)	2769	2698	2482	1823	1319	682	54
Percent of 15 min	100	97	90	66	48	25	2
Correlation Coefficient to 15 minute	1.00	0.97	0.91	0.80	0.69	0.53	0.27

The decrease in both the surface area and volume of pooled water observed when the model is run with coarsening rainfall resolution can be explained by the differences in rainfall intensity between the scenarios. In the control simulation, rainfall events occur as short-

duration, high-intensity events, with rain falling at a faster rate than can be absorbed by the soil. This leads to runoff and pool formation. When these rainfall events are averaged over one or more hours, the events begin to resemble long-duration, low-intensity rainfall, which is more easily infiltrated into the soil.

2.3.2 Effect on Infiltration and Evaporation

Rainfall over the Banizoumbou model domain dissipates by infiltration, evaporation, or runoff out of the model domain. Field studies in this region have shown that infiltration is the dominant mechanism for the draining of pools, while surface evaporation plays a much smaller role [*Desconnets et al.,* 1997]. The model results are consistent with these findings, with 86.8 to 99.8% of total rainfall being dissipated through infiltration. Although runoff out of the model domain was not explicitly measured, we can assume that any rainfall not accounted for by infiltration or evaporation was lost to runoff because no surface water remains at the end of each simulation. Table 2.2 shows the contribution of infiltration, evaporation and runoff to the dissipation of rainfall in each model simulation.

Table 2-2 Water Balance over Model Domain

	15 min	30 min	1 hour	3 hour	6 hour	12 hour	24 hour
Percent infiltrated	86.8	86.7	85.2	89.2	90.4	95.3	99.8
Percent evaporated	8.7	8.7	9	7.7	6.9	3.4	0.1
Percent lost	4.4	4.5	5.8	3.1	2.7	1.3	0.1

Evaporation and infiltration between the control and the lower rainfall resolution simulations are summarized in Table 2-3. We see an initial decrease in infiltrated water as the rainfall resolution decreases from 15 minutes to 30 minutes and 1 hour. However, simulations using rainfall resolutions of 3 hours or coarser experienced more infiltration than the control simulation.

Table 2-3 Evaporation and Infiltration over Model Domain

	15 min	30 min	1 hour	3 hour	6 hour	12 hour	24 hour
Infiltrated water (mm)	439.03	438.48	430.5	451.12	457.15	482.03	504.76
Percent of 15 min	100	100	98	103	104	110	115
Correlation to 15 min	1	0.98	0.92	0.73	0.64	0.54	0.41
Evaporated water (mm)	44.04	44.15	45.4	38.87	34.85	17	0.5
Percent of 15 min	100	100	103	88	79	39	1
Correlation to 15 min	1	0.99	0.98	0.91	0.82	0.73	0.61

Figure 2.6 shows the cumulative infiltration from soil and infiltration from pools for each simulation. At fine temporal resolutions, the high intensity rains descend at rates higher than the infiltration rate, thus running off and forming water pools or leaving the model domain. As the rainfall resolution is degraded, infiltration from soil increases because lower intensity rainfall is more easily infiltrated as it falls. Infiltration from pools decreases as rainfall resolution coarsens because there are fewer pools. This is why we see a slight decrease in total

infiltration from 15-minute to 30-minute to 1-hour, followed by a steady increase as infiltration from soil begins to dominate.

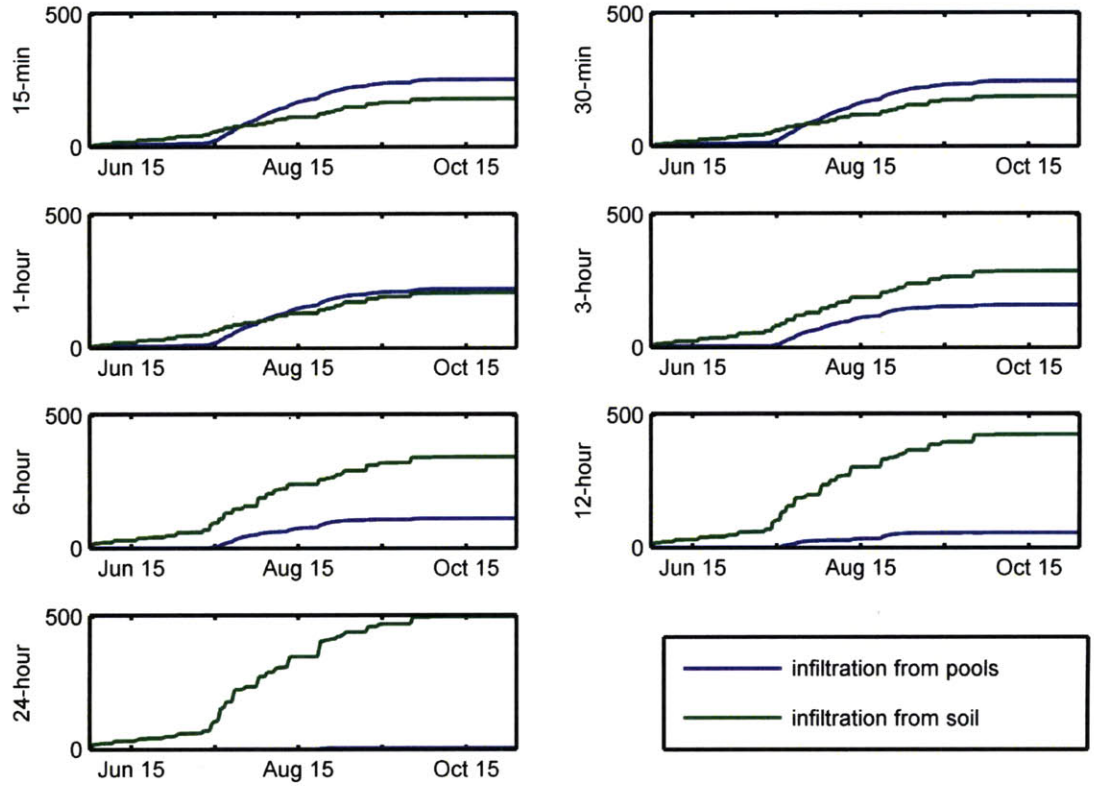

Figure 2.6 Cumulative Infiltration from Pools and Soil

2.3.3 Effect on Probability of Utilization

The probability of utilization, described in Section 1.3, is one of the main coupling factors between the hydrology and entomology components of the model, and is therefore crucial to this analysis. Figure 2.7 shows the cumulative number of cells that fall into each of the four categories shown in Figure 1.1. Although the 15-minute and 30-minute scenarios had

the greatest volume and surface area of pooled water, the 1 hour simulation has the greatest number of cells in the 'ideal' category. This was because the increased amount of pooled water caused a number of cells in the 15 minute and 30 minute scenarios to be too deep for mosquito larvae. Table 2-3 compares the probability of utilization between simulations, averaged over the entire model domain for each timestep. The cumulative probability of utilization was very similar in each of the first three scenarios, and significantly lower in the remaining scenarios. The time-series of mean probability of utilization were highly correlated for the control and the 30-minute simulations with a correlation coefficient 0.96, dropping to 0.87 for the 1-hour simulation, and 0.65 for the 3-hour simulation.

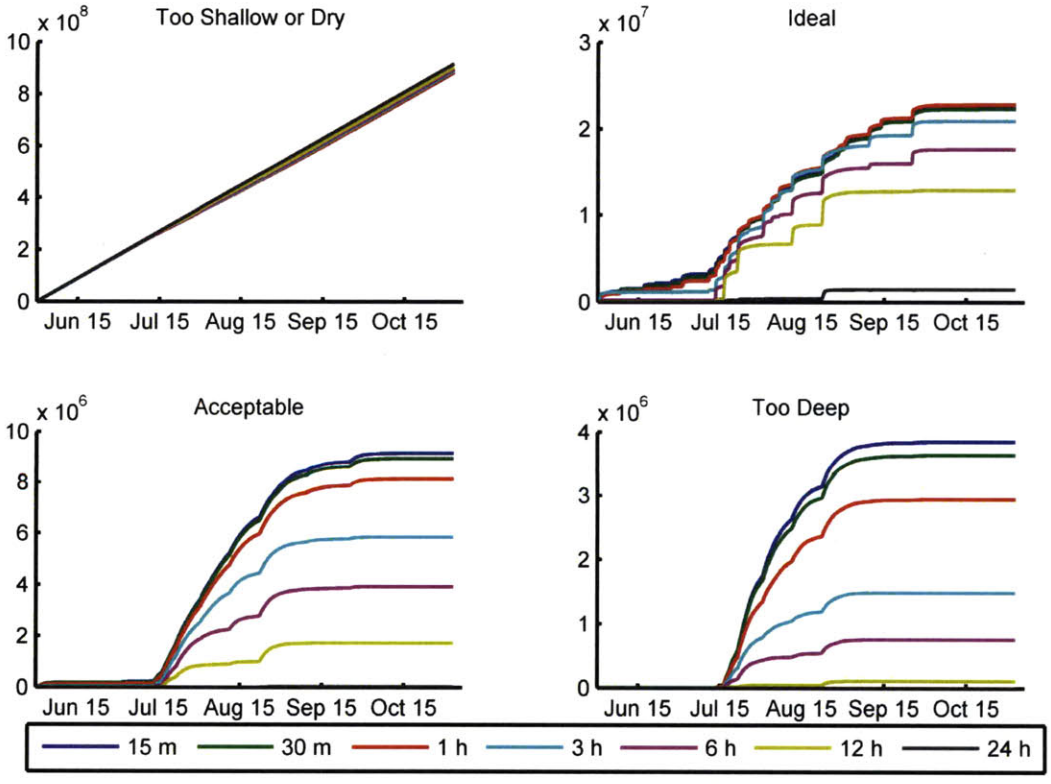

Figure 2.7 Cumulative Grid Points Falling in Each of the 4 Water Depth Categories

Table 2-4 Effect of Varying Rainfall Resolution on Probability of Utilization

	15 min	30 min	1 hour	3 hour	6 hour	12 hour	24 hour
Cumulative Probability of Utilization	133.63	132.35	132.43	117.51	96.52	67.37	7.03
Percent of 15 min	1.00	0.99	0.99	0.88	0.72	0.50	0.05
Correlation Coefficient to 15 minute	1.00	0.96	0.87	0.65	0.52	0.34	0.20

2.3.4 Effect on Mosquito Population

Figure 2.12 shows the number of female adult mosquitoes alive at each time step for the different simulations. We see that the number of mosquitoes between the 30-minute, 1-hour and 3-hour simulations is very highly correlated to the number in the control simulation, with the correlation coefficient of greater than or equal to 0.95. While the peaks and ebbs are closely matched between the various simulations, the coarser resolutions significantly underestimate the magnitude of mosquito populations. The number of adult mosquitoes at each timestep in the 30-minute and 1-hour simulations corresponds reasonably well with the control, underestimating the cumulative sum of mosquitoes by 12% and 15% respectively. The majority of this underestimation occurs in the first three peaks in the mosquito populations, occurring between late mid-July and late-August. When the rainfall resolution is degraded to 3-hour resolution, the model shows the cumulative sum of live mosquitoes as less than 50% that of the control. This dramatic drop in mosquito populations magnifies the decrease in surface area of pooled water presented in Table 2.1, where pooled water drops from 90% the amount seen in the control at 1-hour resolution, to 66% of the control at 3-hour resolution. When the

rainfall is averaged over 12 or 24 hours, the resulting simulations show only very minimal mosquito breeding, with the cumulative sum of mosquitoes being 13% and 7% respectively of that of the control.

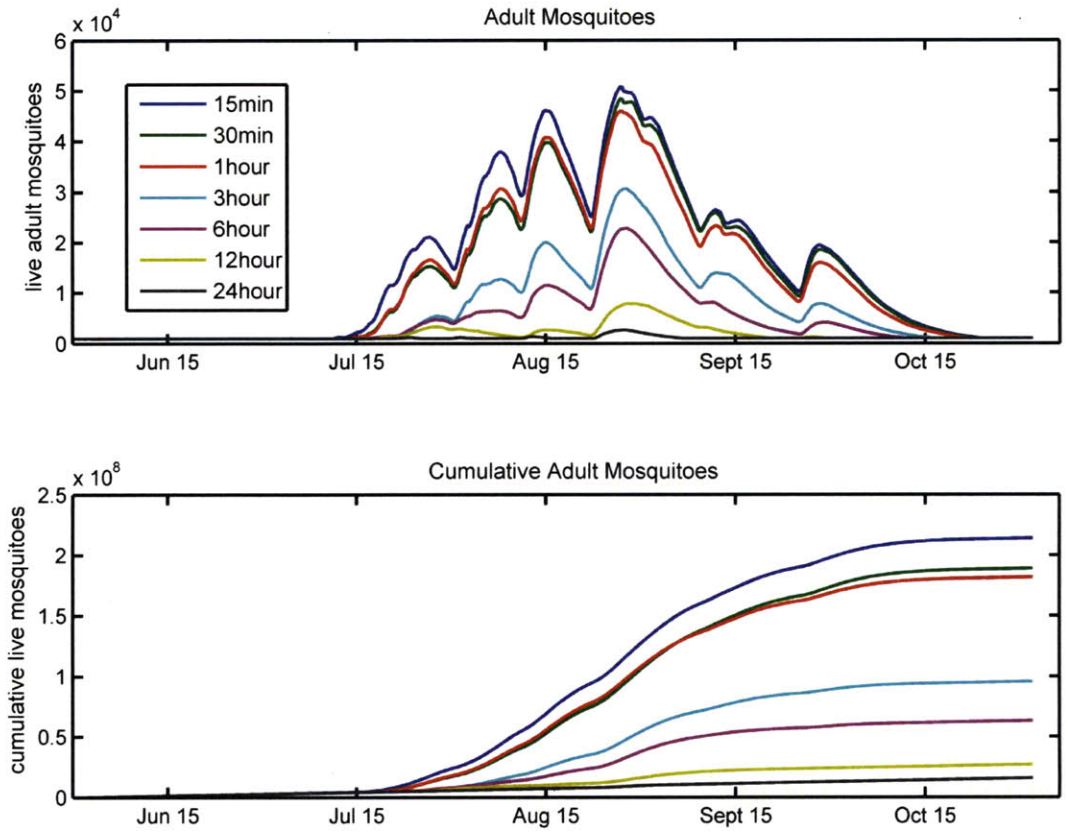

Figure 2.8 Adult Mosquitoes

Table 2-5 Effect of Rainfall Resolution on Mosquito Populations

	15 min	30 min	1 hour	3 hour	6 hour	12 hour	24 hour
Cumulative live adult mosquitoes (millions)	214.02	189.02	181.79	95.73	63.23	26.92	15.68
Percent of 15 min	1.00	0.88	0.85	0.45	0.30	0.13	0.07
Correlation Coefficient to 15 minute	1.00	0.99	1.00	0.95	0.91	0.78	0.62

2.4 *Conclusion and Discussion*

Hydrological modeling can add valuable information to malaria transmission models. However, in order to function properly, these models require a fine temporal resolution of rainfall inputs. These simulations show that while the availability and convenience of using a rainfall data set with low temporal resolution has certain advantages, HYDREMATS' skill in accurately modelling the hydrologic and entomologic systems decreases as the resolution of rainfall decreases. This is because low temporal resolution data sets average rainfall events into longer, lower intensity events that lead to more infiltration, and reduce the amount of pooling. The decreased amount of water in pools means that there is less breeding habitat available for mosquitoes, and thus the number of mosquitoes decreases.

This analysis also demonstrates that the individual storm hyetographs are crucial to determining the effect of rainfall on malaria transmission. Even though all of the simulations had the same total amount of rainfall, the difference in the distribution of rainfall led to drastically different outcomes. This means that it is not enough to look at the total rainfall of an area in order to make intelligent predictions about malaria risk. An area where rainfall occurs frequently but at low intensities is likely to have fewer pools suitable for mosquito larvae than an area with comparable terrain that experiences the same volume of rainfall, occurring in low frequency, high intensity rainfall events.

Given the results of these simulations, we conclude that a minimum rainfall resolution of 1 hour should be used with HYDREMATS. Although there is some loss in accuracy, the high levels of correlation of all examined variables between the 1-hour simulation and the control indicate that HYDREMATS can give a reasonable representation of its domain using rainfall

inputs of this resolution. Meanwhile, 1-hour resolution data is likely to be more available than rainfall data with finer resolution. While 3-hour and coarser resolution rainfall data sets are even more widely available, using these data sets would result in an unacceptable loss in accuracy in the model.

If hourly rainfall data is not available, the model's performance would be improved by applying a disaggregation algorithm in order to better simulate the distribution of rainfall. The aquatic stage of the mosquito's life cycle lasts for 7-10 days, and depends on the persistence of water pools; larvae die when pools dry up. Thus while hourly rainfall resolution is required to properly model the formation of pools, the fine-scale information about the actual sequence of rainfall events is less important to mosquito breeding, which operates on a coarser time scale. Daily accumulation totals can be useful for modeling pool formations if the rainfall is distributed in a way that allows the model to accurately represent the runoff, infiltration, and pool formation process, rather than averaging rainfall into equal amounts over 24 hours as was done in this investigation. Disaggregation algorithms use daily accumulation totals to infer the statistics of rainfall events on a time scale fine enough to be applied to hydrologic modeling. This allows rainfall inputs to have realistic distributions, even if the exact characteristics of each rainfall event are unknown. In one such algorithm, daily totals of rainfall were used to infer rainfall statistics on the 1-hour timescale, based on theoretical equations for the mean, variance, autocorrelation and probability of zero rainfall (Bo *et al.*, 1994).

3 Application of Satellite Estimates of Rainfall Distribution to Simulate the Potential for Malaria Transmission in Africa

3.1 *Introduction and Literature Review*

The Hydrology, Entomology and Malaria Transmission Simulator (HYDREMATS) is a mechanistic model developed to assess malaria risk in areas where transmission is limited by the availability of water pools suitable for mosquito breeding, and was developed based on extensive field work in Niger. This model relies on rainfall inputs as its primary forcing. Until now, HYDREMATS has used ground-based rainfall observations. However, rain gauge networks in the areas most affected by malaria are often sparse. The increasing availability of satellite based rainfall estimates could greatly extend the range of applicability of the model. The minimum temporal resolution of rainfall data was determined in Chapter 2 to be one hour. Rainfall estimates produced using the CPC Morphing technique (CMORPH) (Joyce *et al.,* 2004) and distributed by National Oceanic and Atmospheric Administration (NOAA) fit this criterion, as it provides 30-minute estimates at 8km resolution. CMORPH was chosen for this study because of this high temporal and spatial resolution, and because it performs better than rainfall estimates based only on passive microwave images (PMW) or PMW data blended with IR data (Joyce *et al.,* 2004).

3.1.1 Applications of Remote Sensing for Malaria Control

The application of remote sensing approaches to malaria control has been the subject of extensive research over the past 30 years, as has been reviewed in numerous articles (Ceccato

et al., 2005; S. Hay *et al.*, 1998; Kalluri *et al.*, 2007; D. J. Rogers *et al.*, 2002; M. C. Thomson *et al.*, 1996)). Early investigations were led by the National Aeronautics and Space Administration (NASA), through the Earth Observations Division (EOD), the Global Monitoring and Human Health program (GMHH), and the Center for Health Applications of Aerospace Related Technologies (CHAART) (Ceccato *et al.*, 2005). Now, remote sensing measurements are used to better understand the dynamics of malaria by many groups, including local malaria control programs (Grover-Kopec *et al.*, 2005).

3.1.1.1 Land Cover Mapping using Visible Data

High resolution images have been used in many settings to locate aquatic breeding habitats, and map areas at high-risk for malaria transmission. Beck et al. (1994)used a 30 meter resolution Landsat Thematic Mapper (TM) image to identify villages in Chiapas, Mexico that were at high risk of malaria transmission, based on abundance of adult *Anopheles albimanus* mosquitoes. When stepwise linear regression and stepwise discriminant analysis were used to determine relationships between landscape features visible on the processed satellite image and the relative abundance of adult mosquitoes captured in 40 villages, transitional swamp and unmanaged pasture were found to be strongly associated with mosquito abundance. Villages were then classified as having high or low vector density based on the proportion of these two landscape elements in the areas surrounding the villages. The classification had an overall accuracy of 90%. In a follow-up study, the transportability of the classification system was assessed by applying the equations generated in the first study to a group of 40 new villages and a new Landsat-TM image in an area 50 km away from the original study site. The villages

were classified as having high or low vector density with an overall accuracy of 70% (Beck *et al.*, 1997). A similar study was conducted in Chiapas by Pope et al. (1994). Landsat Enhanced Thematic Mapper (ETM+), also with 30 meter resolution, was used to identify larval habitats in rice fields in Mali (Diuk-Wasser *et al.*, 2004) and Korea (Masuoka *et al.*, 2003).

Information from Landsat-TM has also been used to estimate measures of malaria transmission. In the Gambia, West Africa, Bogh et al. (2007) related features from Landsat-TM to *Anopheles gambiae* breeding habitats found by a larval sampling survey, creating a map denoting areas of larval habitats. The authors then used linear regression to relate distance from larval habitats to the entomological inoculation rate (EIR), which is the number of infectious bites per person, per unit time. The calculated EIR values were highly correlated to EIR values measured in 12 villages (r=0.752). Castro et al. (2006) used land cover information from Landsat-TM along with the results of a field survey covering environmental, economic and behavioral risk factors to assess the risk of malaria on the Amazon frontier in Brazil. Landsat-TM has also been used to assess the risk of other vector-borne diseases including Lyme disease (Dister *et al.*, 1997), Trypanosomiasis (Kitron *et al.*, 1996), Sin Nombre (Glass *et al.*, 2002), and Rift Valley Fever (K. Pope *et al.*, 1992).

The strategy of mapping larval habitats have been also been used with High Resolution Visible (HRV) images from Satellite Pour l'Observation de la Terre (SPOT) (Rejmankova *et al.*, 1995; Thomas & Lindsay, 2000), IKONOS, and Indian Remote Sensing Satellite (IRS) 1-A and 1-B (Sharma *et al.*, 1996). Achee et al. (2006)investigated whether SPOT 10 m panchromatic and 20 m multispectral resolution images and Ikonos 1 m panchromatic and 4 m multispectral resolution images were suitable for identifying *Anopheles Darlingi* habitat in Belize, which

consist primarily of mats of debris floating in fresh-water rivers. No associations were found between positive larval habitats and the characteristics of the rivers or the surrounding land cover as identified by the satellite images. Mushinzimana et al. (2006) found that a supervised classification of an IKONOS image over the Kenyan highlands resulted in a 89% accurate classification of land cover, and the correct identification of 41% of aquatic larval habitats, while the same technique applied to a Landsat-TM image misclassified 39% of land cover, and failed to find any of 10 identified aquatic habitats.

3.1.1.2 Mapping of Mosquito Habitat and Malaria Transmission Based on NDVI

Lower resolution satellite images have also been used for habitat mapping. Normalized Difference Vegetation Index (NDVI) data from the National Ocean and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR), available at 1.1 km resolution, has been widely used as a proxy for rainfall data to map mosquito habitat suitability. The NDVI value gives information on the abundance of green vegetation, which can then be used to infer information about rainfall (M. C. Thomson *et al.*, 1996). NDVI ranges from -1 to 1, and is given by the formula *NDVI= (NIR-RED)/ (NIR+RED)*, where *NIR* is percentage reflectance in the near infrared channel and *RED* is the percentage reflectance in the red channel.

The following studies mapped larval habitats or malaria-risk areas using NDVI data. Hay et al. (1998) examined the relationship between the seasonality of pediatric severe malaria hospital admissions and remotely sensed data on NDVI, land surface temperature and middle infra-red radiance from AVHRR, as well as cold cloud duration data from the Meteosat-High

Resolution Radiometer (HRR) in three Kenyan communities. The data were subjected to maximum value compositing, temporal Fourier analysis and linear regression. The highest correlation was found between NDVI and malaria admissions (mean r^2=0.74), and months with greater than 5% of the annual malaria cases had a threshold NDVI value of between 0.3 and 0.4. This information was interpolated to create an 8 x 8 km map of Kenya denoting the number of months that were expected to have malaria hospital admissions, which corresponded well with an existing map, which was created manually from expert opinion. Similarly, Nihei et al. (2002) compared maps of malaria distribution on the Indochina peninsula with maps showing the number of months that NDVI was greater than or equal to 0.2, 0.3, 0.35 and 0.4, and found that the distribution of 0.3+ and 0.4+ monthly NDVI corresponded well with malaria cases. Eisele et al. (2003) used the NDVI approach with 5 x 5 m resolution data from the Multispectral Thermal Imager (MTI) aggregated to 270 x 270 m to map larval habitats in two urban areas of Kenya. After adjusting for household density, a correlation between NDVI and larval habitats was observed (r^2=0.56). Thomson et al. (1999) applied methodologies used by the Famine Early Warning Systems Network (FEWS) to find a significant association between age-related malaria prevalence rates in Gambian children and NDVI. Rahman et al. (2006) calculated vegetation health (VH) indices using NDVI and brightness temperature (BT) information from AVHRR and investigated their correlations with malaria prevalence in Bangladesh. The vegetation condition index (VCI), is a function of NDVI and is an estimate of moisture, and the temperature condition index (TCI) is a function of BT and estimated temperature. While neither index had high correlations in cooler months where mosquitoes were less active, during the warm and wet period between April and October, TCI reached a

maximum correlation of 0.7 and VCI had a maximum correlation of -0.66. NDVI from AVHRR data has been used to in conjunction with other diseases, including Rift Valley Fever (Linthicum *et al.*, 1999), trypanosomiasis (D. Rogers & Randolph, 1991), schistosomiasis (Malone *et al.*, 1994), and leishmaniasis (Cross *et al.*, 1996)

3.1.1.3 RADAR

Measurements made by RADAR microwave sensors have the advantage of being able to penetrate clouds. Ritchie (1993) demonstrated that land-based S-band radar estimates of rainfall could be used to locate larval broods of *Aedes taeniorhynchus* mosquitoes in Florida. Pope et al., (1992) combined Landsat-TM data with synthetic aperture radar (SAR) information to detect habitats of the Rift Valley Fever vector, *Aedes* mosquitoes. *Aedes* eggs are deposited in topographical low points known as dambos, and remain viable for up to a year until flooding induces hatching. While the Landsat-TM data was able to locate dambos, it could not distinguish between flooded and non-flooded statuses. The flooding status of dambos was determined using 1.6 x 2.4 m resolution data from an airborne SAR L-band horizontal transmit, horizontal receive (LHH) sensor. Kaya et al. (2007) used land classification of RADARSAT-1 data to identify wetland areas and flooded vegetation in coastal Kenya, which are associated with mosquito larval habitats.

3.1.1.4 Environmental Variables

While most of the previous studies used remotely sensed variables such as NDVI and land cover as proxies for rainfall and humidity, some authors have investigated satellite based measurements of environmental variables directly. Thomson et al. (2005) investigated the

relationship between malaria incidence in Botswana and estimates of rainfall and sea surface temperature (SST) between 1982 and 2003. Total rainfall between December and February of each year was calculated using 2.5^{o} x 2.5^{o} resolution Climate Prediction Center Merged Analysis of Precipitation (CMAP) data, which incorporates gauge observation, satellite estimates, and reanalysis data and is available at the daily time scale. The SST used was the Niño3.4 index, provided by NOAA's Climate Prediction Center (CPC) at the monthly timescale, and averaged for December through February by the authors. After adjusting for non-climatic trends and policy interventions, seasonal log malaria incidences were strongly correlated with total CMAP rainfall December through February (r^2=0.85 in quadratic model). The averaged SSTs were associated with both seasonal rainfall (r=-0.55) and standardized malaria incidence anomalies (r=-0.49). Since the peak malaria transmission occurs in March and April, the authors suggest that rainfall and SST from December through February could be used to give an early warning for high transmission years. Omumbo et al. (2002) examined the relationships between historical intensity of malaria transmission and NDVI, mid-infrared reflectance, land surface temperature and air temperature data obtained from AVHRR, as well as data on altitude and cold cloud duration in Tanzania, Kenya and Uganda. Altitude information was from GTOPO30, a 1 km resolution digital elevation model (DEM) provided by the United States Geological Survey (USGS). Cold cloud duration (CCD) was measured by the European Organisation for the Exploitation of Meteorological Satellites (EUMESAT) Meteosat series of satellites, and was used as a proxy for rainfall. Statistical analysis relating temporal Fourier processed satellite data to historical transmission intensity showed land surface temperature to be the best predictor of malaria transmission intensity, followed by CCD and NDVI.

Monitoring rainfall has been recognized as an essential component for the malaria early warning systems (MEWS) advocated for by the Roll Back Malaria initiative (Grover-Kopec *et al.*, 2005). In response to this need, dekadal (every 10 days) estimates rainfall anomalies are distributed by the Africa Data Dissemination Service (ADDS), a website supported by the U.S. Agency for International Development (USAID). The estimates incorporate rain gauge data with cloud top temperature from Meteosat 7, as well as estimates from the Special Sensor Microwave/Imager (SSM/I) on the Defense Meteorological Satellite Program satellites and the Advanced Microwave Sounding Unit (AMSU) on NOAA satellites (Ceccato *et al.*, 2006). Hay et al. (2003) retrospectively determined that this data would have provided a reliable warning to a major malaria epidemic that occurred in 2002 in Kenya. Monitoring of malaria cases detected the epidemic in June and July; rainfall data from May showed over 35% more rain than average. The Highland Malaria Project (HIMAL) was investigating the use of ADDS rainfall estimates for their MEWS in Uganda and Kenya, and malaria control programs in Uganda and Sudan reported monitoring the estimates on a regular basis (Grover-Kopec *et al.*, 2005).

Re-analysis data has also been used with malaria models. Hoshen & Morse (2004) used European Centre for Medium Range Weather Forecasting (ECMWF) second-generation global weather re-analysis data set ERA-40 with their mechanistic model of malaria transmission. ERA-40 is available for every 6 hours for the years 1960-2000 at $0.5^{\circ} \times 0.5^{\circ}$ resolution, and was converted by the authors to dekadal totals to act as a forcing for their model.

While all of these studies use remote sensing data of some form, none use satellite-based estimates of meteorological variables to force a mechanistic model of malaria transmission such as HYDREMATS. Rogers et al. (2002) and Kalluri et al. (2007) call for the use

of satellite data in mechanistic models of vector-borne pathogen transmission as an improvement over the statistical models currently being used and an important next step for malaria control.

3.1.2 CMORPH

The Climate Prediction Center morphing method (CMORPH) (Joyce *et al.*, 2004) provides global estimates of rainfall every 30 minutes at a 0.07277 degree (~8km) spatial resolution. CMORPH combines rainfall estimates from passive microwave (PMW) sensors with spatial propagation vectors derived from infrared (IR) data. PMW sensors can detect thermal emission and scattering patterns associated with rainfall. However, these sensors are only available on polar orbiting satellites, giving them limited spatial and temporal coverage. Infrared data from geostationary satellites are available globally at 30 minute resolution, and can be used to determine the movements of the precipitating systems sensed by PMW instruments.

CMORPH uses infrared measurements from the Geostationary Operational Environmental Satellite (GOES) 8 and 10, Meteosat-5 and 7, and Geostationary Meteorological Satellite-5 (GMS-5). PMW sensors used by CMORPH are aboard the NOAA polar-orbiting operational meteorological satellites, the U.S. Defense Meteorological Satellite Program (DMSP) satellites, and the Tropical Rainfall Measuring Mission (TRMM) satellite. Consecutive PMW images are propagated forward and backwards in time using motion vectors derived from the infrared images. The shape and intensity of the precipitating systems in the 30-minute intervals between PMW measurements are determined by using a time-weighted linear interpolation (Joyce *et al.*, 2004).

While CMORPH has been shown to perform significantly better at estimating rainfall than techniques that use only PMW images or PMW data blended with rainfall estimates derived from IR data (Joyce *et al.*, 2004), it does have some error. Tian et al. (2007) compared 3-hour $0.25^{\circ} \times 0.25^{\circ}$ resolution CMORPH estimates to ground measurements over the United Sates, and found CMORPH to have a substantial positive bias in the summer, and a minor negative bias in the winter. However the correlation to rain gauge data was higher in the summer than in the winter. There was a clear spatial pattern to correlation, with high latitudes, coastal regions and mountainous areas having the weakest correlation. They also observed that CMORPH overestimates the number of high-intensity rainfall events. Zeweldi and Gebremichael (2009) compared CMORPH to gauge-adjusted Next Generation Weather Radar (NEXRAD) data over the Little Washita watershed in Oklahoma. They noted a large variability in the hourly CMORPH bias, meaning that a correction factor based on the monthly bias would not be helpful in determining the hourly bias. They found that temporally averaging data is only effective in reducing error over timescales of 6 hours or greater. Laws et al. (2004) compared CMORPH estimates to data from rain gauges across Africa by finding the mean daily bias, calculated by subtracting CMORPH daily rainfall from daily totals of corresponding rain gauges. CMORPH was found to consistently overestimate rainfall, with the mean daily bias and standard deviation of the bias increasing as the daily rainfall recorded by CMORPH increases.

The wet bias of CMORPH data has been attributed to evaporation of rainfall below the cloud base (Tian & Peters-Lidard, 2007). This was observed by Rosenfeld and Mintz (1988) when they examined the evaporation of falling rainfall in a semi-arid region of South Africa. By taking radar measurements at the cloud base and then at a number of points between the

cloud and the surface, they developed a relationship between the fraction of rainfall evaporated and the fall distance, for varying levels of rainfall intensity. They found that rainfall with intensity of 80 mm/hr, 15% evaporated within 1 km of the cloud base. This fraction increased for lower intensity rainfall, as these were comprised of smaller raindrops that are more prone to evaporation.

CMORPH biases have accounted for over the United States by creating a new data set called RMORPH. RMORPH was created by scaling CMORPH values such that the daily rainfall estimate at each grid point is equal to the daily accumulation reported by the Climate Prediction Center gauge analysis (Janowiak *et al.*, 2007).

3.2 *Simulation Description*

This investigation was based on two HYDREMATS simulations conducted over the domain of Banizoumbou, Niger for the year 2007. The first simulation, the control, used rain gauge data as the rainfall input. The second simulation used adjusted CMORPH data as the rainfall forcing. The simulations ran using a 1-hour timestep.

In order to determine whether CMORPH could provide a reasonable estimate of rainfall over West Africa, CMORPH values were compared to rain gauge data at three West African locations for the year 2006: Zindarou village in Niger, Djougou in Benin, and Agoufou in Mali. Zindarou village is located approximately 20 km away from Banizoumbou. Djougou, Benin receives significantly more rainfall than Banizoumbou, with an average of over 1000mm per year, spread out over a longer rainy season. Agoufou, Mali, is notably drier than Banizoumbou,

with average annual rainfall between 200 and 400mm. The map depicted in Figure 3.1 shows the four locations. The dotted contour lines show average annual rainfall.

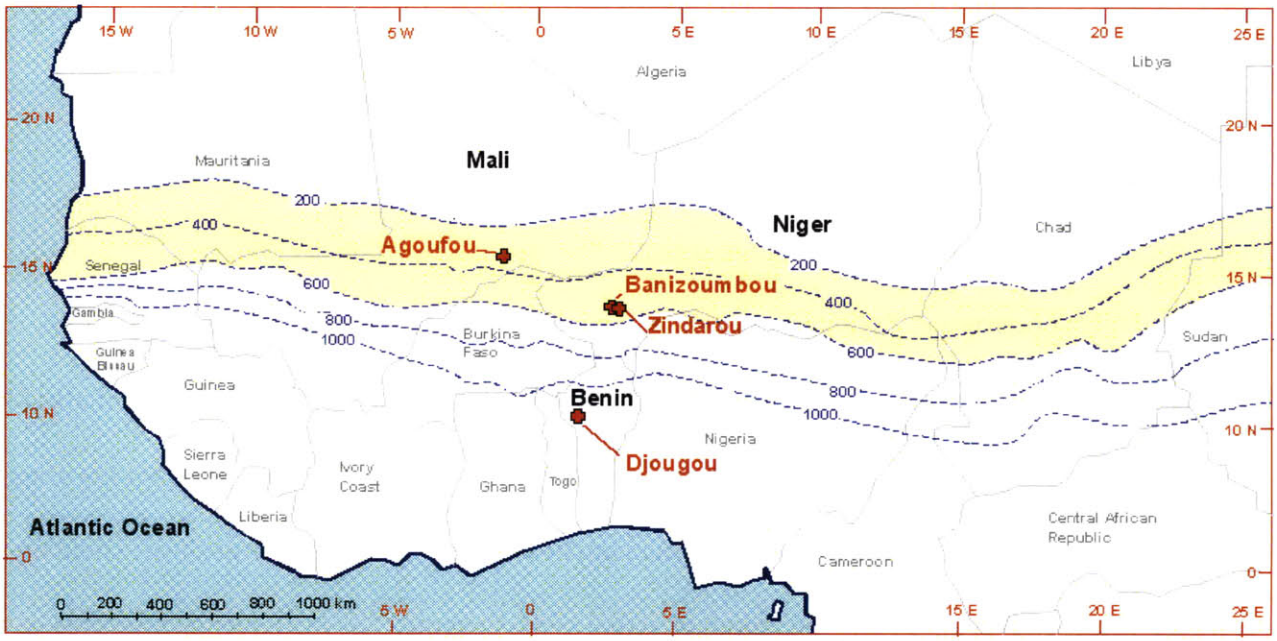

Figure 3.1 Locations of Banizoumbou, Zindarou, Agoufou and Djougou (adapted from Bomblies, 2008)

Rainfall in Zindarou was measured using a Texas Electronics TE525 rain gauge, installed by Bomblies (2008) and maintained by Centre de Recherche Médicale et Sanitaire (CERMES), our collaborators in Niger. Rain gauge data for the other three locations were obtained from the African Monsoon and Multidisciplinary Analyses (AMMA) database. The 30-minute CMORPH data and 5 to 20 minute rain gauge data for these four locations were averaged into hourly values of rainfall, which is the timestep determined by the previous investigation to be sufficient for accurate use of HYDREMATS.

An initial comparison of rainfall totals between the two types of data sets showed that CMORPH overestimated total annual rainfall by 45% to 68%. The frequency bias (FB), false

alarm ratio (FAR), and probability of detection (POD) were calculated on hourly and daily time scales. The FB refers to the number of hours or days where CMORPH estimates non-zero rainfall divided by the number of non-zero rainfall observations on the ground. FAR refers to the fraction of non-zero rainfall estimates in CMORPH that did not correspond to non-zero rainfall measurements on the ground. The POD is the fraction of non-zero rainfall measurements observed on the ground that were correctly detected by CMORPH. The POD, FB, FAR for hourly and daily data at each location are presented in Table 3-1. The overall FB was 1.81 for hourly data and 1.31 on for daily data. This means that CMORPH has estimates far more non-zero rainfall hours than is actually observed on the ground. The significant lowering of FB when rainfall is averaged over 24 hours implies that many of the false positives could be the result of overestimating the length of actual rainfall events. The POD over all four sites is 0.61 on the hourly scaly and 0.81 daily. The FAR is 0.67 hourly and 0.38 daily.

Examining the data, we observed that many of the false positives occurred when the CMORPH hourly rainfall estimates were less than 1mm. Such small amounts of rainfall are likely to evaporate before reaching the surface. It was found that the FB and FAR could be significantly reduced by setting all CMORPH data points less than or equal to 0.4 mm to 0 mm, with a tradeoff being a small decrease in the POD. On the hourly scale, this adjustment led to a 33% decrease in FB, a 15% decrease in the FAR, and a 15% decrease in the POD. On the daily time scale, the FB and the FAR decreased by 11%, while the POD decreased by 5%.

Table 3-1 CMORPH detection statistics

	Hourly			Daily		
	FB	PoD	FAR	FB	PoD	FAR
CMORPH						
Banizoumbou	2.24	0.65	0.71	1.27	0.75	0.41
Zindarou	2.11	0.61	0.71	1.34	0.83	0.38
Agoufou	2.01	0.68	0.66	1.21	0.83	0.31
Djougou	1.53	0.57	0.63	1.35	0.82	0.4
Overall	1.81	0.61	0.67	1.31	0.81	0.38
Adjusted CMORPH						
Banizoumbou	1.46	0.58	0.6	1.2	0.73	0.4
Zindarou	1.42	0.5	0.65	1.22	0.78	0.36
Agoufou	1.36	0.6	0.56	1.12	0.79	0.3
Djougou	1.04	0.48	0.53	1.15	0.77	0.33
Overall	1.22	0.52	0.57	1.17	0.77	0.34
Percent Change	33%	15%	15%	11%	5%	11%

After removing the low intensity rainfall estimates, CMORPH data was further adjusted by multiplying each value by the ratio of total rainfall measured by rain gauges to total rainfall estimated by CMORPH in Zindarou, Agoufou and Djougou, which was 0.73. Banizoumbou was not included in the calculation of the correction factor so that we could test the transportability of the adjustment method to other stations. This crude adjustment scaled down CMORPH estimates such that the annual totals were closer in magnitude to observed data. Figure 3.2 shows weekly and cumulative rainfall for the three locations for rain gauge data, raw CMORPH data, and adjusted CMORPH data. We see that the adjusted CMORPH data set agrees well with observed weekly and yearly rainfall totals.

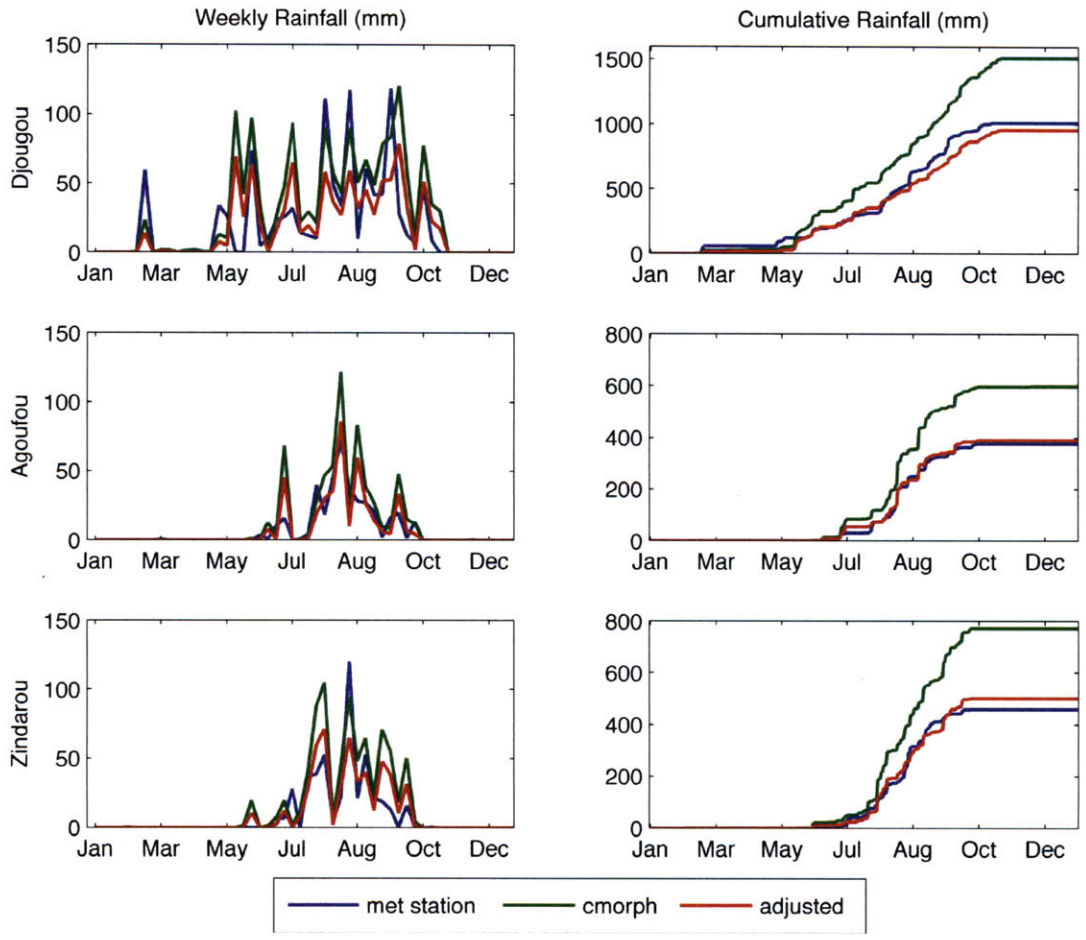

Figure 3.2 Weekly and Cumulative Rainfall in 2006

The adjustment method was applied to CMORPH data for Banizoumbou in 2007. The resulting rainfall series, shown in Figure 3.3, was the rainfall forcing for the CMORPH simulation in HYDREMATS. The correlation between the adjusted CMORPH rainfall and met station rainfall during the rainy season (June 1 – Oct 31) was 0.26, 0.74 and 0.77 for the hourly, daily and weekly time scales respectively.

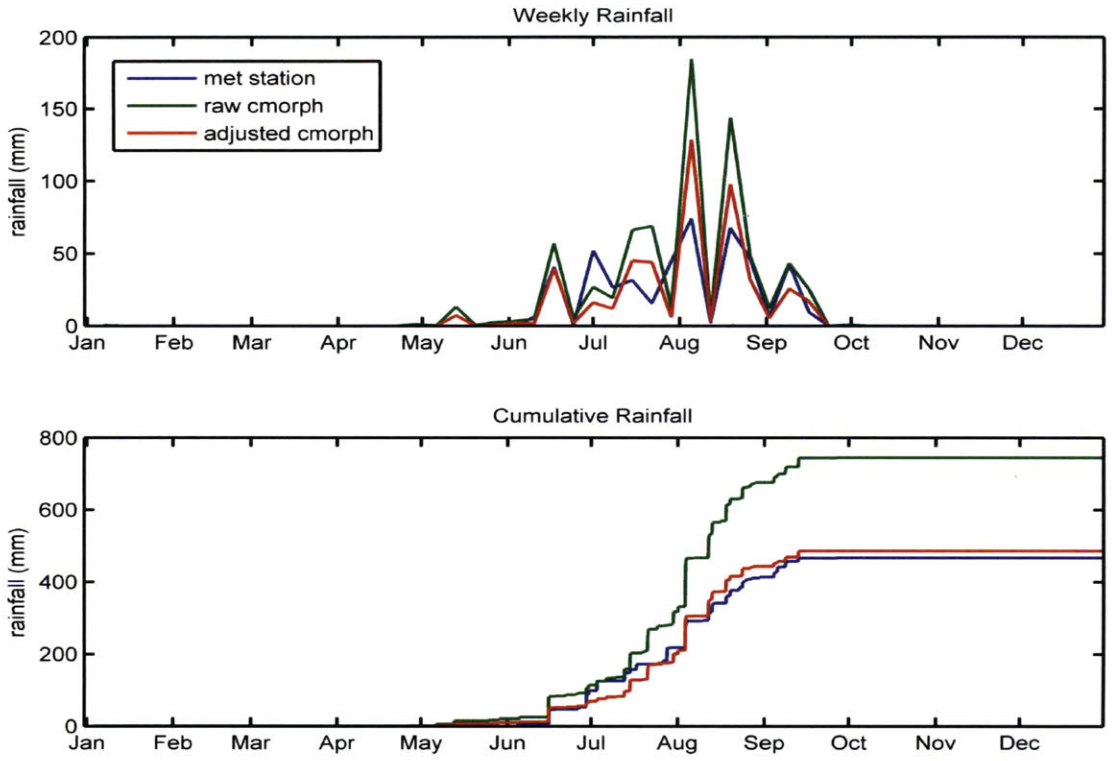

Figure 3.3 Banizoumbou Rainfall in 2007

3.3 ***Results***

The hourly depth of each of the three pools shown in Figure 2.2 is presented in Figure 3.4, with the blue line depicting pool depth output from the simulation using ground data, and the red line pool depth using adjusted CMORPH rainfall forcings. Pool 1, the deepest and largest of the pools, was consistently deeper in the ground data simulation than in the CMORPH simulation, despite the fact that the CMORPH simulation had slightly more total rainfall. The depths of Pools 2 and 3 are closer in magnitude between the two simulations, but their

correlations are lower. The correlation coefficients between the two simulations are 0.99, 0.78 and 0.83 for Pools 1, 2, and 3 respectively.

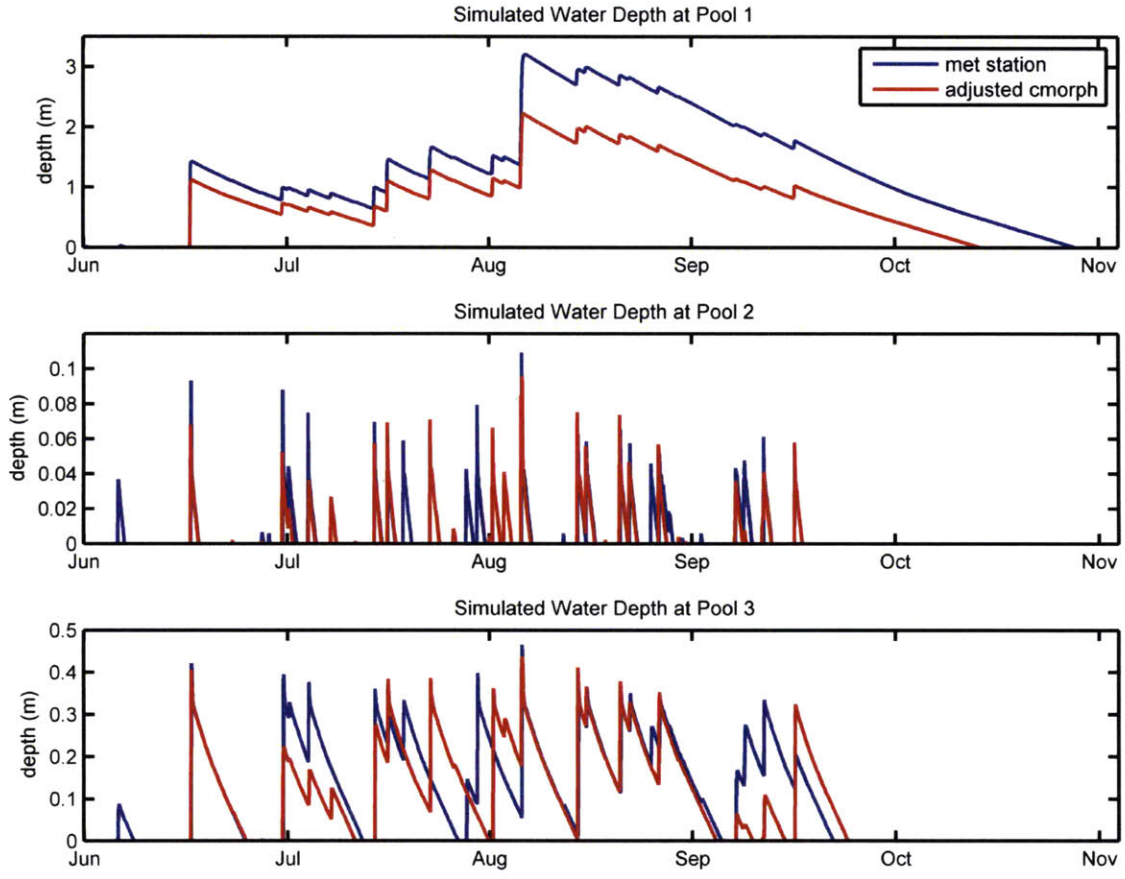

Figure 3.4 Pool depths

The initial disparity of depths between the two simulations observed in Pool 1, observed mid-June, is interesting because it accentuates the importance of rainfall distribution in the pool formation process. The rainfall event leading to the formation of Pool 1 occurred over a 9-hour period between June 15 and June 16, 2007. This event was recorded by both the rain gauge and CMORPH, with the rain gauge showing a total of 40.5 mm, while the adjusted CMORPH data set showed a total of 39.6 mm rainfall. Despite this very close agreement in

magnitude of the event, the gauge data recorded the rainfall over 3 hours with a maximum intensity of 34.6 mm/hr, while the adjusted CMORPH data showed the event as occurring over 7 hours with a maximum intensity of 16.1 mm/hr. The higher intensity rainfall in the rain gauge simulation caused higher rates of surface runoff leading to greater volumes of pooled water, thus Pool 1 was deeper in the rain gauge simulation than it was in the CMORPH simulation. The same reason explains the increased difference in depths of Pool 1 observed in early August. Pools 2 and 3 are much smaller and shallower, and are therefore less sensitive to the overall runoff patterns of the model domain. They quickly reach their maximum depth, after which excess water travels as runoff to a greater depression in the topography.

While differences in rainfall distribution accounted for the discrepancies seen in Pool 1 in mid-June and early-August, this is not representative of all rainfall events. Figure 3.5 shows the daily mean surface area of pools, cumulative surface area of pools, and cumulative rainfall. There is significant correlation between the daily mean surface area of water pools of the two simulations, with a correlation coefficient of 0.76. This correlation is comparable to that of rainfall at the daily timescale. Comparing cumulative surface area of pools to cumulative rainfall shows that the major discrepancies observed in the daily mean surface area of pools between the two simulations are primarily due to differences in the magnitude of rainfall events between the adjusted CMORPH and ground data rather than differences in their temporal distribution; the surface area of pools is closely correlated to rainfall, with a correlation coefficient of 0.76 in the met station simulation and 0.84 in the CMORPH simulation.

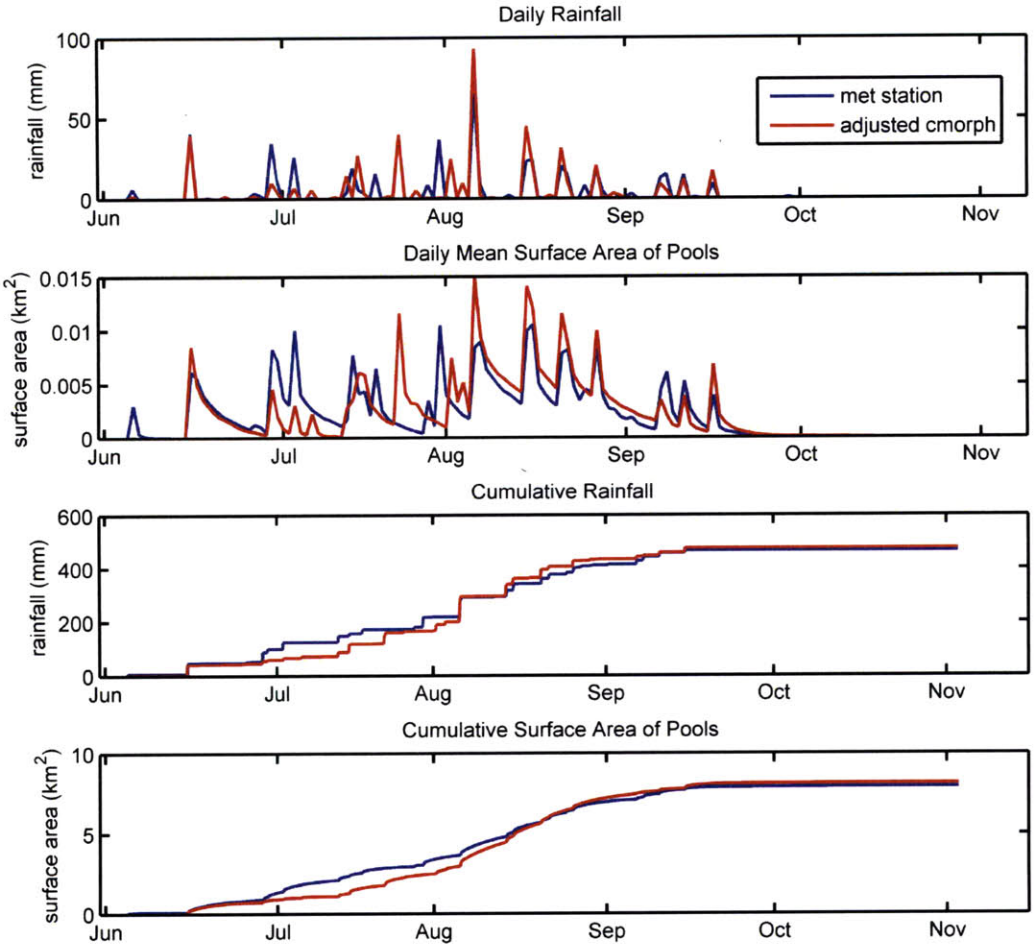

Figure 3.5 Surface Area of Pools and Cumulative Rainfall

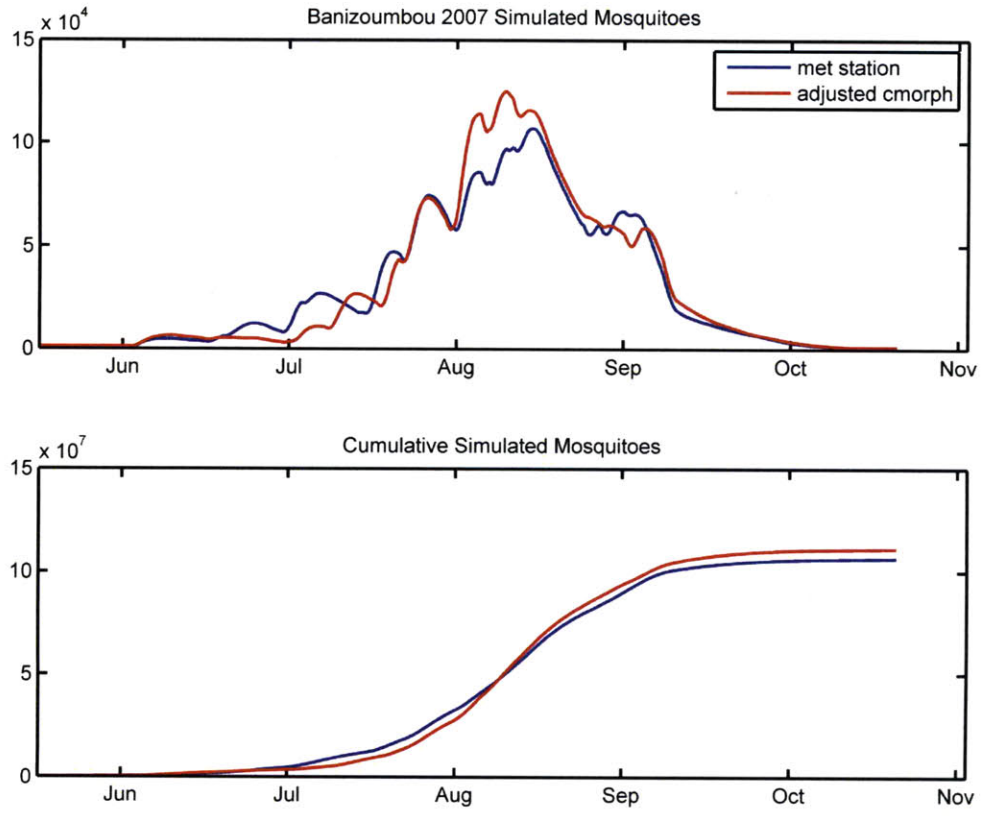

Figure 3.6 Simulated Mosquitoes

The number of simulated mosquitoes alive at each timestep for the two simulations is shown in Figure 3.6. The correlation coefficient for the two outputs is 0.98, and the root mean square error is 5.9×10^3. The CMORPH simulation shows lower numbers of mosquitoes than the rain gauge simulation in late June and July, and early September, and greater numbers in August. This is consistent with the relative amounts of pooled water available during these times.

3.4 *Conclusion and Discussion*

The results of this study show that satellite derived estimates of rainfall can be used in a mechanistic model to simulate mosquito populations and malaria transmission. The use of satellite data with such models has been described as the logical next step to existing studies using satellite data for malaria control (Rogers et al. 2002, Kalluri et al., 2007). While these studies that use statistical information to infer relationships between satellite based environmental observations and malaria transmission have great value in mapping malaria-risk areas on the country scale, they generally do not address village-scale variability. Since they most often do not explain causal pathways between remotely sensed data and malaria transmission, they have limited ability to predict the malaria response to a given environmental change.

Satellite technology is constantly improving, providing plentiful information which can be used in a model such as ours. In the future, we envision a system by which a large portion of the inputs to HYDREMATS could be obtained from satellite platforms and archived data sets. These inputs include temperature, humidity, wind speed and direction, topography, soil characteristics, and location of residences. If these data could all be applied to the model, the range of applicability of HYDREMATS could be extended to every village in Western Africa, as well as any other area where malaria transmission is limited by the availability of water. HYDREMATS could be a valuable tool for researchers and those working in malaria control programs. This would be especially useful for addressing changes in the environment such as droughts or floods, climate change and land-use change.

4 Early Warning of Malaria Transmission using HYDREMATS

4.1 *Introduction and Literature Review*

One important application for a model such as HYDREMATS is its potential to be used to provide early warnings of malaria epidemics. Epidemics occur most frequently in areas where malaria transmission is unstable, as the populations of these areas have lower levels of immunity than those in areas of constant transmission (WHO, 2001). Unstable transmission can occur in areas where disease transmission is limited by climatic variables such as temperature or rainfall. Other factors leading to unstable transmission include the movement of people between areas of high and low transmission, and interruptions of malaria control strategies (WHO, 2001). An early warning of an upcoming epidemic would give health officials the opportunity to prepare for or prevent the rapid onset of new malaria cases. As such, some African countries have set targets for early detection of epidemics. As part of the Roll Back Malaria in Africa (RBM) initiative, countries in the Southern Africa Development Community (SADC) have adopted targets stipulating that 60% of epidemics be detected within two weeks of onset, and that 60% of epidemics be responded to within two weeks of detection (DaSilva *et al.*, 2004).

The RBM initiative has published a framework for malaria early warning systems (MEWS) in Africa. These systems rely on indicators of vulnerability, transmission risk and early case detection in order to predict the onset and severity of malaria epidemics {{304 WHO 2001}}. Monitoring rainfall has been recognized as an essential component for MEWS (Grover-Kopec *et al.*, 2005). In response to this need, dekadal (every 10 days) estimates of rainfall anomalies are

distributed by the Africa Data Dissemination Service (ADDS), a website supported by the U.S. Agency for International Development (USAID). The estimates incorporate rain gauge data and satellite data (Ceccato *et al.*, 2006). Hay et al. (2003) retrospectively determined that this data would have provided a reliable warning to a major malaria epidemic that occurred in 2002 in Kenya. Monitoring of malaria cases did not detect the epidemic until June and July. Rainfall data from May, showing over 35% more rain than average, could have alerted authorities that an epidemic would be likely in the following months. The Highland Malaria Project (HIMAL) was investigating the use of ADDS rainfall estimates for their MEWS in Uganda and Kenya, and malaria control programs in Uganda and Sudan reported monitoring the estimates on a regular basis (Grover-Kopec *et al.*, 2005).

Thomson et al. (2005) investigated the relationship between malaria incidence in Botswana and estimates of rainfall between 1982 and 2003. Total rainfall between December and February of each year was calculated using Climate Prediction Center Merged Analysis of Precipitation (CMAP) data, which incorporates gauge observation, satellite estimates, and reanalysis data and is available at the daily time scale. After adjusting for non-climatic trends and policy interventions, seasonal logarithmic malaria incidences were strongly correlated with total CMAP rainfall December through February (r^2=0.85 in a quadratic regression based model). Since the peak malaria transmission occurs in March and April, the authors suggest that rainfall from December through February could be used to give an early warning for high transmission years.

While excess rainfall is often associated with increased malaria transmission, this is not always the case. For example, heavy rainfall associated with the 1997-98 El Nino event was

associated with decreased malaria transmission in the highlands of Tanzania, presumably by washing away larval breeding sites (Lindsay *et al.*, 2000). Similarly, decreases in rainfall have been observed to increase malaria transmission by creating breeding pools in areas where flowing water would normally wash larva away (Wijesundera Mde, 1988). With a hydrology driven model such as HYDREMATS, the relationships between anomalous levels of rainfall and malaria transmission can be explicitly represented, allowing the user to draw the correct conclusions from information regarding rainfall patterns. While HYDREMATS does not currently represent the flushing of breeding sites, excess rainfall can lead to decreased breeding site availability if pools exceed the maximum depth for ovipositing. In the case of breeding sites along a river, the model could be configured to prohibit ovipositing and survival of aquatic stage mosquitoes in cells connected by water to the river.

Another component of some MEWS includes seasonal forecasting of climate based on statistical or dynamic general circulation models linking climatic variation associated with the El Nino Southern Oscillation (ENSO) to malaria transmission. Seasonal climate forecasts have been used in some parts of Africa to make predictions for malaria transmission up to six months in advance. These long range climate forecasts typically give probabilistic data describing the likelihood of rainfall being below normal, normal, or above normal on the provincial or country scale (WHO, 2001). Hay *et al.* (2003) concluded that given the low temporal and spatial specificity and the probabilistic nature of these estimates, such long range forecasts were not useful for malaria control in the short-term. However Thomson et al. (2003) countered that best estimate area-average rainfall predictions from climate models for Kenya were highly correlated with observed rainfall (r=0.88), and could therefore be useful as a component of

MEWS. They stressed that closer dialogue between climate scientists and health officials was necessary in order to make the predictions useful, given their spatial limitations and probabilistic nature.

Some studies have shown that DEMETER multi-model seasonal climate predictions (Palmer *et al.*, 2004) can be used to give warnings of high malaria transmission months in advance. The DEMETER system consists of seven European global coupled ocean-atmosphere models, each of which is run using an ensemble of initial conditions. The resulting simulations provide a probability distribution of future climate (Palmer *et al.*, 2004). Morse et al. (2005) used rainfall and temperature output from the DEMETER multi-model seasonal climate predictions as the meteorological forcings for a mathematical model of malaria transmission developed by Hoshen and Morse (2004). Malaria prevalence outputs from these simulations were compared to outputs when ERA-40 reanalysis data were used to force the model. The results showed that DEMETER outputs combined with their malaria transmission model could make predictions on malaria transmission one month in advance. The model also made skilful predictions of above-median prevalences within a 4-6 month forecast window for the time period covering the seasonal malaria peak (Morse *et al.*, 2005).

The model used by Morse et al. (2005) differs from HYDREMATS in that it does not take into account topography and hydrology; rather, rainfall is aggregated over 10 day periods and multiplied by a factor as a proxy for pool availability. HYDREMATS is a mechanistic, spatially distributed model, while Morse et al.'s (2005) is a lumped model. Thomson et al. (2006) used seasonal forecasts of precipitation averaged over Botswana from DEMETER, and linked this to malaria incidence based on a regression relationship determined in an earlier study (M. C.

Thomson *et al.*, 2005). They found that using this statistical relationship linking climate predictions from DEMETER to malaria incidence led to skillful forecasts of probabilities of unusually high and low malaria incidence, issuing early warnings five months before the peak malaria season, compared to only one month's lead time when observing precipitation directly (M. Thomson *et al.*, 2006). Table 4-1 describes the main difference between HYDREMATS and the models used by Thomson et al. (2005, 2006) and Morse et al. (2005).

Table 4-1 Comparison of Malaria Models

Model	Model Type	Spatial Distribution
HYDREMATS	Mechanistic	Distributed
Morse et al. (2005)	Mechanistic	Lumped
Thomson et al. (2005, 2006)	Statistical regression	Lumped

In this study, we have developed a simple method to estimate rainfall two and four weeks in advance. We believe that combining the lead time gained by these forecasts with the lead time provided by the lag between rainfall and malaria transmission will allow us to use HYDREMATS to estimate malaria transmission as measured by vectorial capacity several weeks in advance. While this does not have the same advantages of a warning several months in advance, it has less uncertainty than longer range forecasts, and it could nonetheless be helpful as it would allow malaria control programs some time to redistribute drug supplies, prepare health clinics for an influx of cases, engage in vector control activities and raise public awareness.

4.2 *Predicting Mosquito Populations and Vectorial Capacity using HYDREMATS*

4.2.1 Predicting Rainfall

When attempting to predict rainfall with a lead time on the order of two weeks, the important factors to consider are the seasonality, the history, and the persistence of rainfall. The mean bi-weekly rainfall in Banizoumbou is taken from the Global Daily Merged Precipitation Analyses of the Global Precipitation Climatology Project (GPCP) (Huffman, 1997). GPCP values from 1997 to 2008 were averaged for each two-week period between June 1st and October 31st. Although a longer time-record would be desirable, GPCP with bi-weekly resolution was only available over these 11 years. The mean and standard deviation of biweekly precipitation totals over the 12 year period is shown in Figure 4.1. The graph shows a strong seasonality of rainfall, with a strong peak in August.

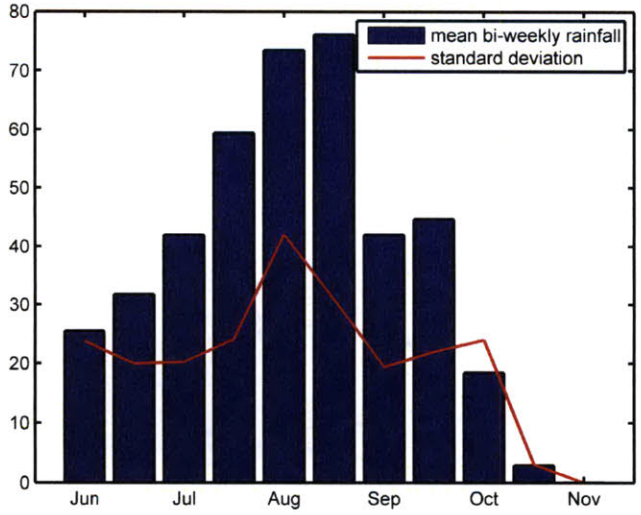

Figure 4.1 Mean and standard deviation of biweekly precipitation totals for Banizoumbou

To determine the extent of persistence of rainfall, the standardized anomaly of rainfall, which is defined as the departure from the seasonal mean divided by the standard deviation for that time period, during each two-week period in Banizoumbou between 1997 and 2008 was calculated, and compared to that of the following two-week period. This analysis, shown in Figure 4.2, showed that there was very little persistence of rainfall amounts between two-week periods, with a correlation coefficient of only 0.07. Since there was little persistence in rainfall patterns on the two-week time-scale, we relied on the history and seasonality of rainfall to make predictions of rainfall 14 days into the future.

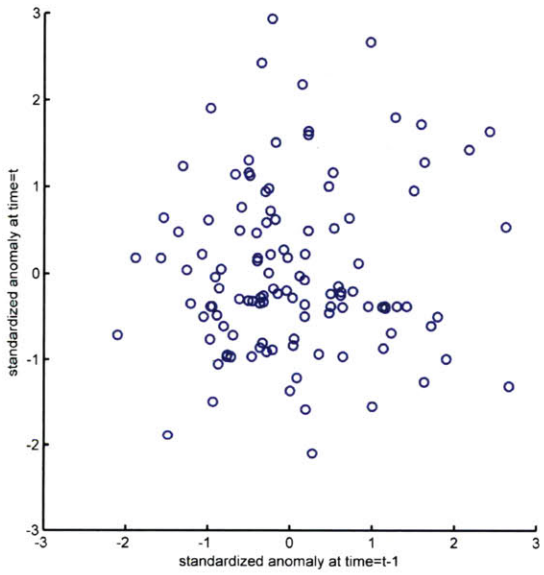

Figure 4.2 Standardized anomalies of adjacent biweekly precipitation totals

This investigation focused on Banizoumbou village during the rainy seasons of 2005, 2006 and 2007. Hourly rainfall measured by a rain gauge for this time period is shown in Figure 4.3.

HYDREMATS simulations using observed rainfall were conducted and taken as control simulations. The observed sequence of rainfall is referred to as pattern 1.

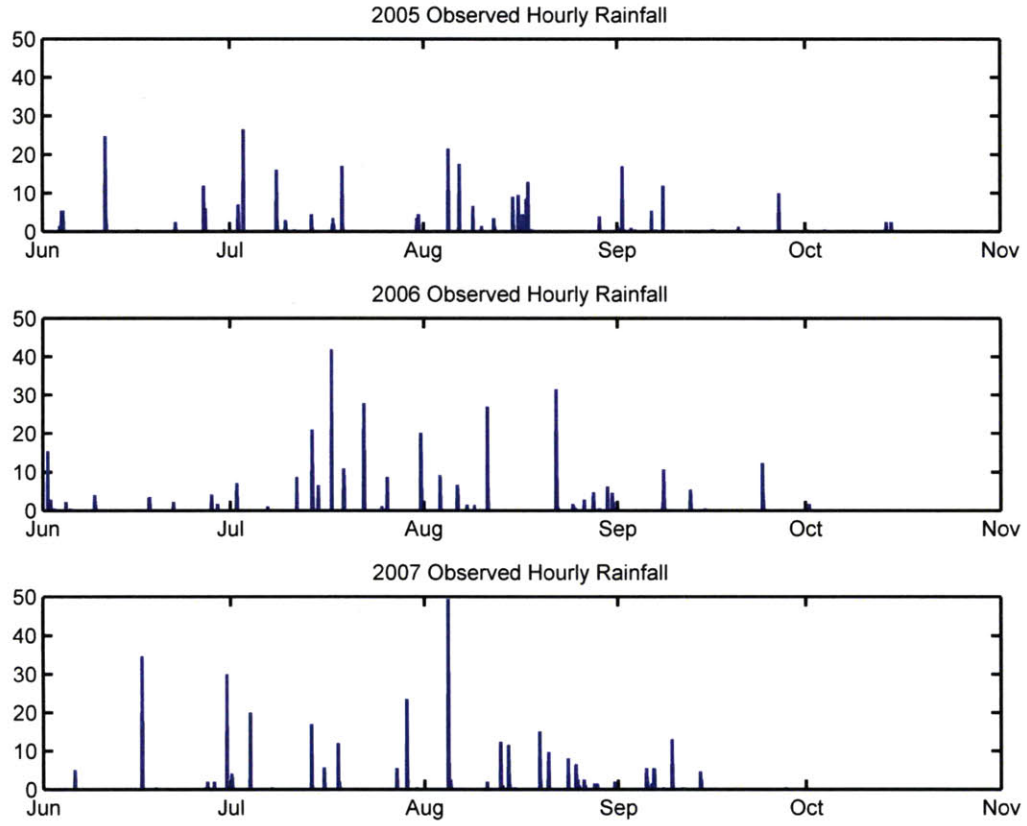

Figure 4.3 Observed Hourly Rainfall in Banizoumbou, Niger

The main objective in this chapter was to create a prediction for rainfall 2 weeks and 4 weeks into the future. The predicted rainfall data sets were created by taking the observed rainfall up until the time of the prediction, thus incorporating the history of the rainfall system. In order to forecast the rainfall sequence and magnitude for the future two weeks, the actual rainfall time series for the latest two weeks was scaled such that the total volume of rainfall was equal to the mean of historically observed rainfall corresponding to the future two-week

period based on the GCPC data. This method incorporates the information about seasonality, while distributing rainfall in a realistic manner. A rainfall time series was created in this manner starting on June 1st and continuing to each successive two-week period from June 28th until October 31st of 2005, 2006 and 2007. This pattern of rainfall is defined as pattern 2. Figure 4.4 shows rainfall inputs for simulations for the sixth bi-weekly period of the 2006 rainy season, which corresponds to late-August. The first plot shows observed rainfall. The second graph shows the predicted rainfall series. This series uses observed rainfall from the first five bi-weekly periods. Finally, the pattern of rainfall observed for the fifth period is repeated in the sixth period, scaled such that the total volume is equal to the mean volume for that period, which is 76 mm. A HYDREMATS simulation was conducted for each of these predicted rainfall series.

In order to isolate the effects on malaria forecasts based on rainfall inputs, real time measurements of the remaining environmental inputs to HYDREMATS, which are temperature, relative humidity, radiation, wind speed and wind direction, were used in all simulations. In a true prediction scenario, we would need to estimate future values of these variables as well. Possible prediction methods include extending the method used here for rainfall to include other environmental variables, or using predictions from medium range weather forecasting or seasonal forecasting systems.

While warnings of high malaria transmission two weeks in advance would give malaria control officials some time to prepare, warnings given four weeks in advance would be even more useful. To investigate whether this is feasible using our prediction method, the prediction simulations were repeated using four-week timesteps rather than bi-weekly timesteps.

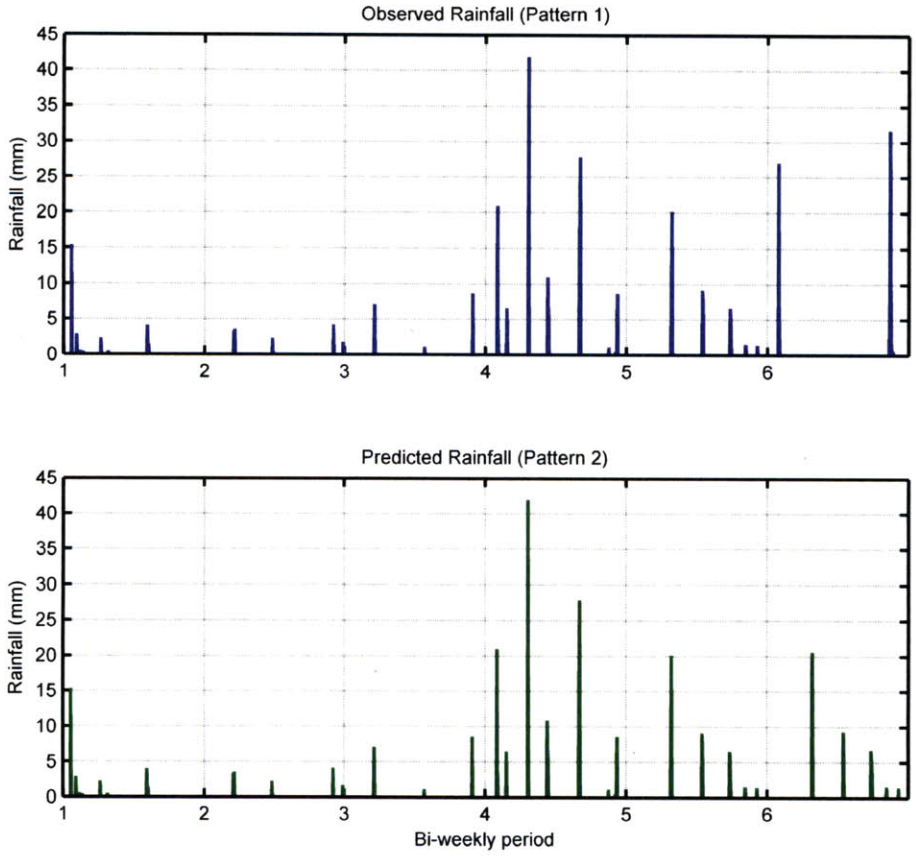

Figure 4.4 Rainfall Inputs for the 6th Bi-weekly Period in 2006

4.2.2 Results

4.2.2.1 Bi-weekly Simulations

Figure 4.5 shows the mean number of live female adult mosquitoes for each two-week period during the rainy seasons of 2005, 2006 and 2007 with the simulations forced by observed rainfall (pattern 1) in blue and 2-week prediction rainfall series in green. The number of mosquitoes is on the y-axis, and the bi-weekly timestep is on the x-axis, where timestep 1

corresponds to June 1, timestep 2 is June 14 and so on. The blue and green lines overlap at timesteps 10 and 11.

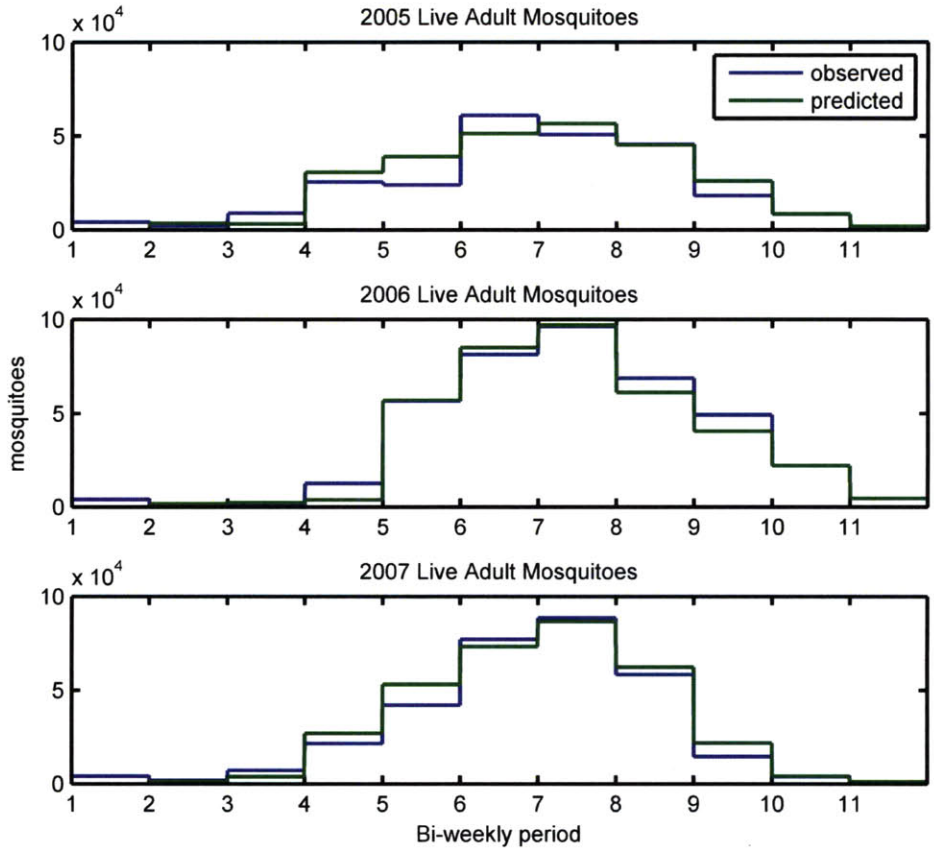

Figure 4.5 Bi-weekly estimates of adult mosquito population

The results shown in Figure 4.5 demonstrate that our simple method of predicting rainfall leads to a prediction of mosquito populations that is quite accurate. In all three years, the predicted number of mosquitoes follows the same seasonal cycle as the number of mosquitoes in the simulation using observed rainfall. This method correctly predicted that the peak number of mosquitoes in 2006 and 2007 would be significantly greater than the peak in 2005.

The timing of the peak was also correctly predicted in 2006 and 2007, but was predicted to be two weeks later than observed in 2005.

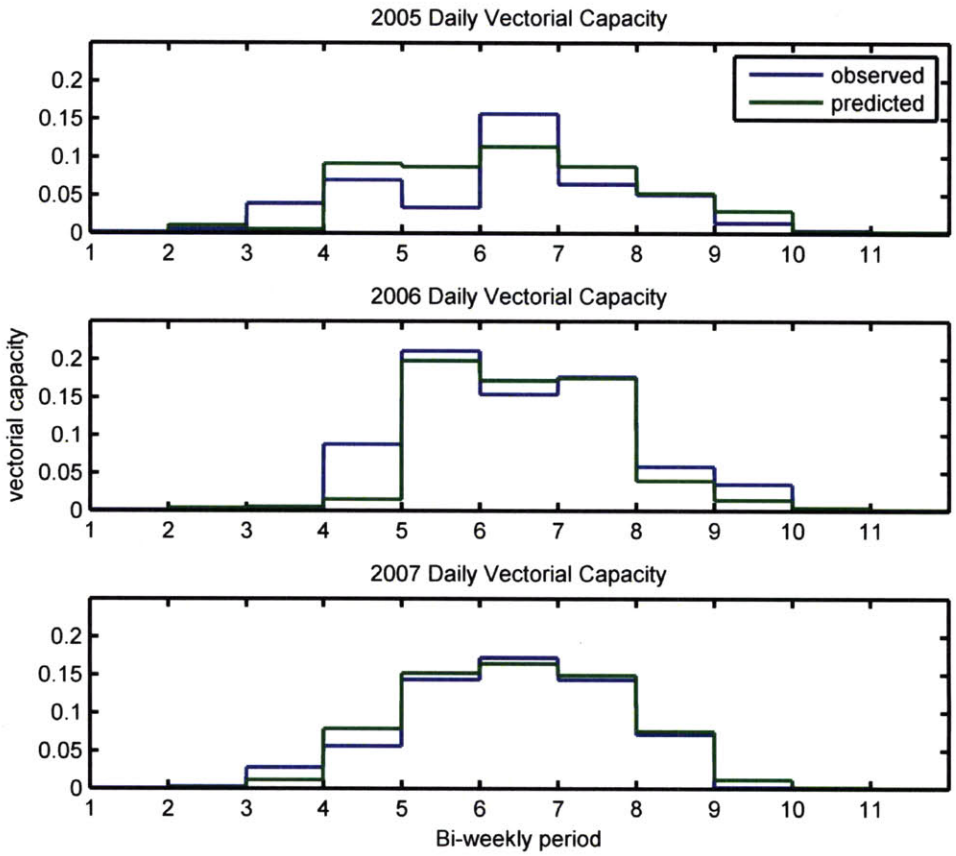

Figure 4.6 Bi-weekly estimates of vectorial capacity

Figure 4.6 shows the bi-weekly mean vectorial capacity outputs from the same simulations. Vectorial capacity is a measure of the mosquito's ability to transmit disease, and is defined as the average number of human inoculations of a parasite originating from a single case of malaria, if all vectors biting the original case were to become infected (Garrett-Jones & Grab, 1964). The equation for vectorial capacity, VC, is:

$$VC = \frac{ma^2p^n}{-\ln p}$$

Where

m is the number of female mosquitoes per human

a is the average number of bites each mosquito takes per timestep

p is the probability the mosquito survives the given timestep

n is the extrinsic incubation time of the parasite, which is the time the parasite resides in the vector before it can be transmitted to humans

In our simulations, *p* and *n* are the same between simulations over a given year, as they are primarily a function of air temperature, which does not change between simulations. The mosquito density, *m*, is directly proportional to the number of mosquitoes, as the number of humans in HYDREMATS simulations are held constant. The timestep used to calculate vectorial capacity is one day.

The accuracy of our prediction method can be assessed by comparing results from the simulations using observed and predicted rainfall, shown in the blue and green lines in Figure 4.5. In all three years, the period of peak transmission was correctly identified. Vectorial capacity was overestimated for the fourth and fifth bi-weekly periods of 2005, and underestimated during the sixth period, which was when peak transmission occurred. Vectorial capacity for the fourth bi-weekly period of 2006 was underestimated. Otherwise, predicted values of vectorial capacity were very close to the values obtained when the model was forced with observed rainfall.

4.2.2.2 4-Week Prediction Simulations

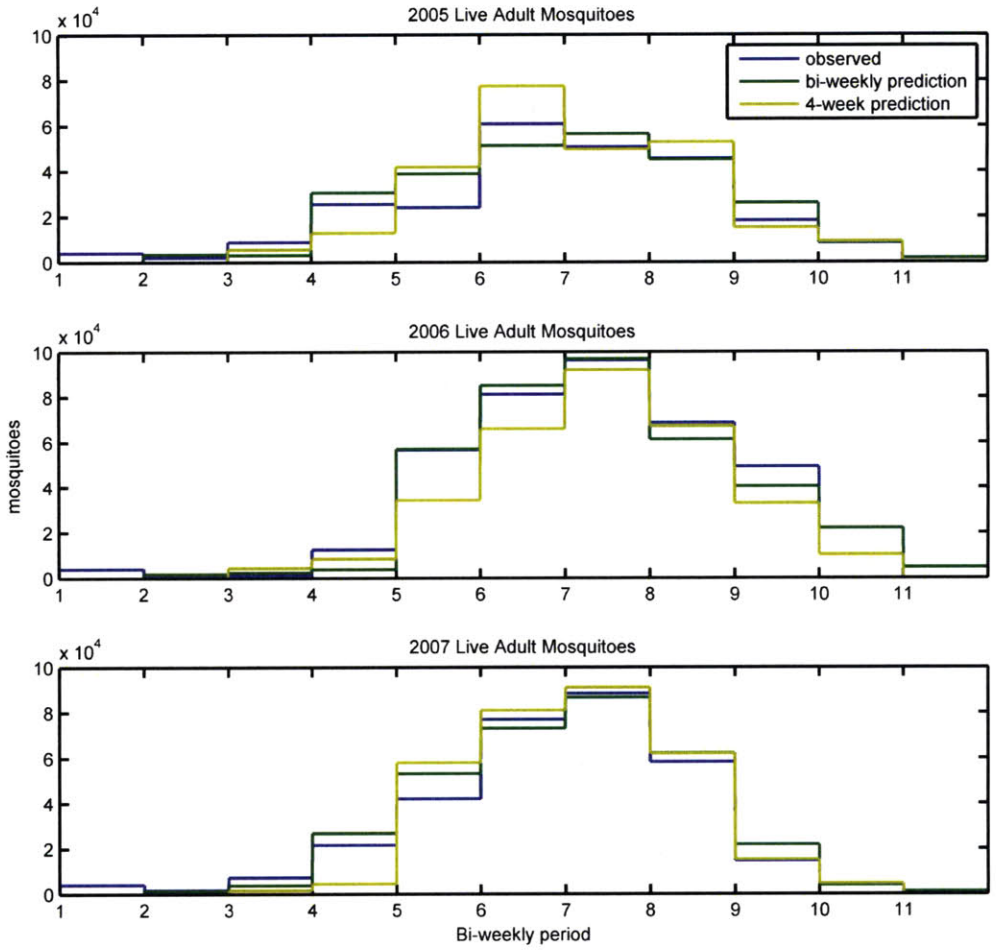

Figure 4.7 Bi-weekly and 4-week estimates of mosquito populations

The accuracy of the 4-week prediction can be assessed based on the results shown in Figure 4.7. This figure shows the bi-weekly mean mosquito populations for the observed rainfall simulation in blue, the 4-week prediction in yellow, and for comparison, the 2-week prediction in green. The numbers on the x-axis denote the times at which the predictions for the following four weeks were made, with 1 corresponding to June 1st, and each subsequent mark corresponding to an additional four weeks. Although the predictions were made every

four weeks, the results were calculated as two two-week estimates in order to maintain a two-week resolution for mosquito populations and vectorial capacity.

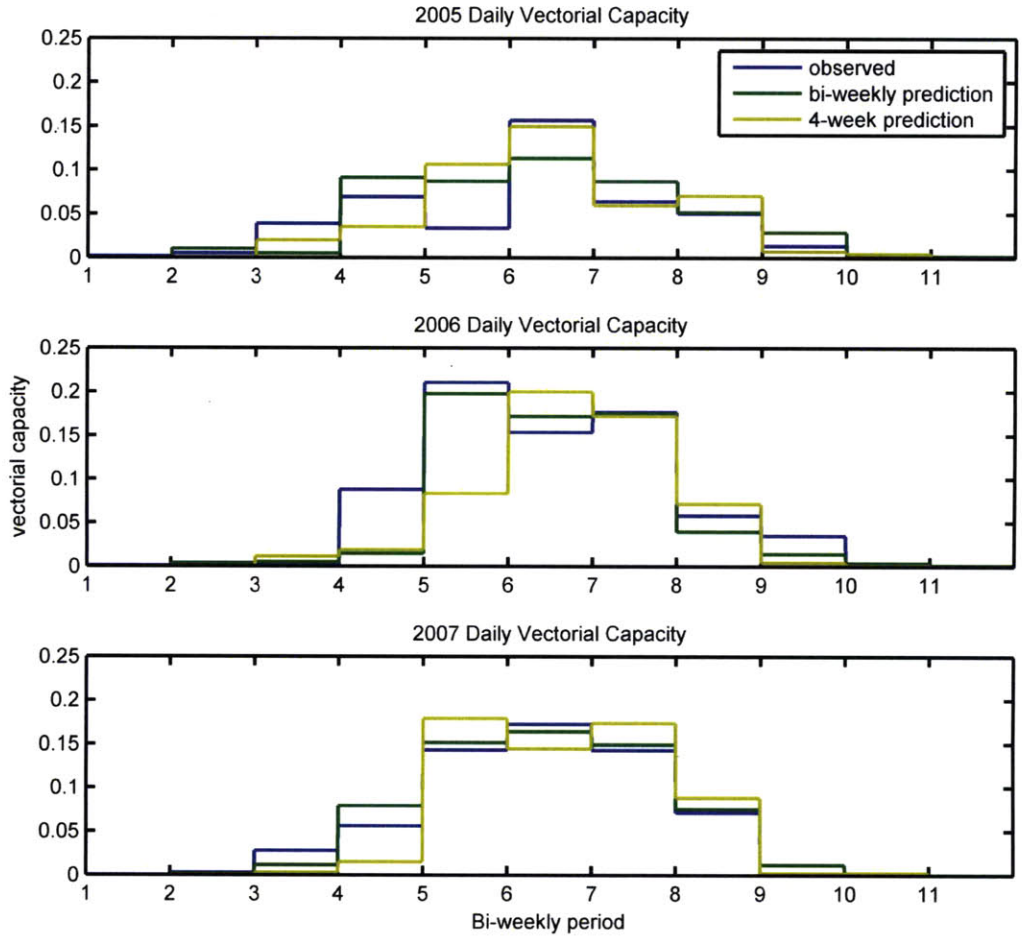

Figure 4.8 Bi-weekly estimates of vectorial capacity

Comparing mosquito populations from the observed rainfall simulations to the 4-week prediction simulations, we again see that the prediction simulations were able to forecast the seasonal pattern of mosquito populations. However, in most cases, estimates made four weeks in advance resulted in estimates further from the control than estimates made only two weeks

in advance. In general, the 4-week predictions made for the second half of the malaria season were very close to the control, while predictions for the early part of the season were less consistent. The predictions for vectorial capacity made four weeks in advance, shown in Figure 4.8, were also generally less accurate than estimates made two weeks in advance, with estimates for the second half of the season being significantly more accurate than estimates for the early months of the rainy season. These findings suggest that a reasonable prediction method would be to make predictions with two-week lead time for the early rainy season, transitioning to four-week predictions in the second half of the season.

4.3 *The Benefit of Using the History of Rainfall in a Prediction Scenario*

In order to examine the benefit of including the history of the magnitude of observed rainfall in the prediction, a new set of rainfall inputs was created by scaling the pattern 2 rainfall prediction series such that the total volume of rainfall in each two-week period was equal to the seasonal mean for that period. We call these rainfall inputs 'seasonal mean (pattern 2)'. Simulations forced by seasonal mean (pattern 2) rainfall are predictions for the following two weeks, made without the benefit of knowing the magnitude of past rainfall.

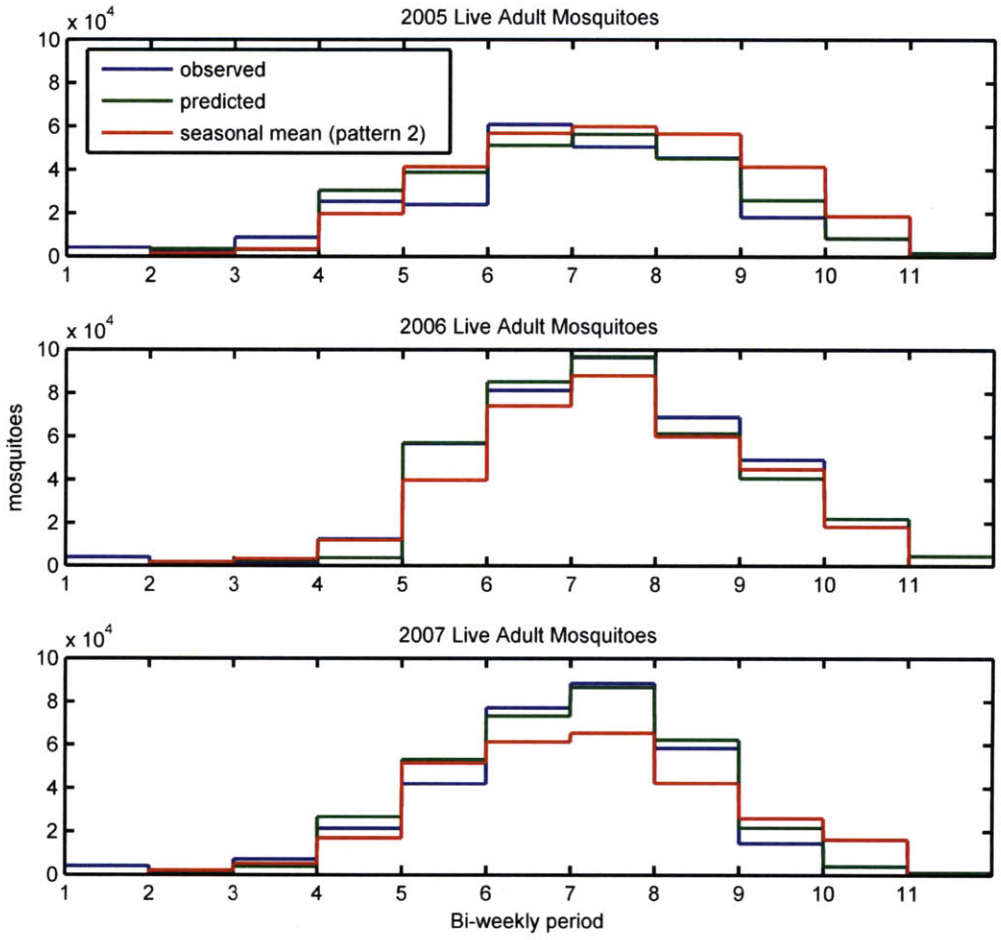

Figure 4.9 The benefit of rainfall history in predicting mosquito populations

Figure 4.9 shows the mean number of live female adult mosquitoes for each two-week period during the rainy seasons of 2005, 2006 and 2007 with the simulations forced by observed rainfall (pattern 1) in blue, 2-week prediction rainfall series (pattern 2) in green, and seasonal mean (pattern 2) rainfall series in red. The value of including the history of rainfall in the prediction can be assessed by comparing results from the prediction simulation and the seasonal mean (pattern 2) simulation.

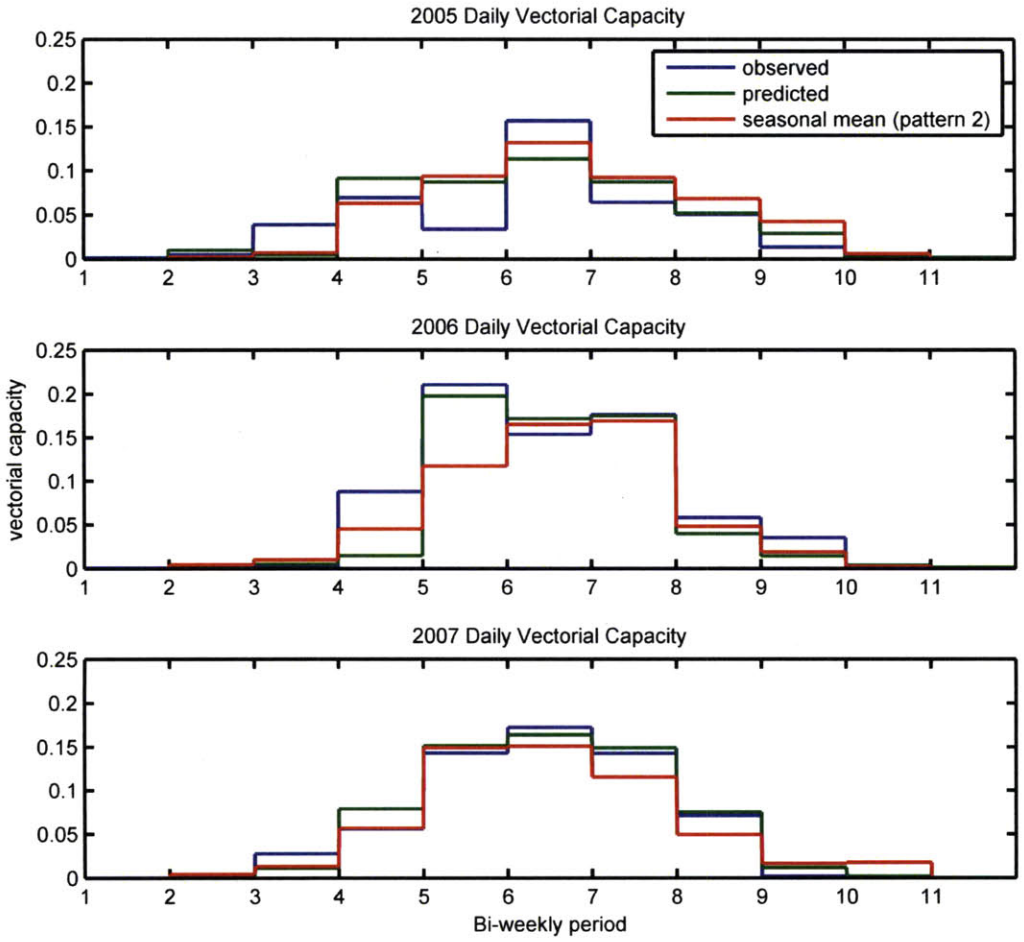

Figure 4.10 The benefit of using rainfall history in predicting vectorial capacity

We can see that for predicting mosquito populations, the history of rainfall is important, as the mean number of mosquitoes estimated by the prediction simulations were almost always closer to the observed rainfall simulations than the seasonal mean simulations were. This is most evident in the middle of the rainy season in 2007, where the seasonal mean simulation significantly underestimated the peak mosquito populations occurring in the sixth through ninth bi-weekly periods. Total rainfall in 2005 was close to the annual mean total, so

the difference between the prediction and the seasonal mean simulations is less apparent. However, we do observe a consistent over-estimation of mosquito populations in the later part of 2005 in the seasonal mean simulation, while the prediction simulation was very close to the observed rainfall simulation. The seasonal mean simulation underestimated mosquito populations in the fifth through ninth bi-weekly periods in 2006, which is the period of peak number of mosquitoes; estimates given by the prediction simulation during this period were closer to the values obtained in the observed rainfall simulation.

Interestingly, the advantage of using the history of rainfall magnitude is much less apparent in the vectorial capacity results, shown in Figure 4.10. The seasonal mean simulations produced estimates of vectorial capacity comparable to those produced by the prediction simulations. The seasonal mean simulation in 2006 did not predict the sharp increase in vectorial capacity in the fifth bi-weekly period, while this signal was successfully reproduced in the prediction simulation. Vectorial capacity was consistently underestimated in the sixth through eight bi-weekly periods of 2007 in the seasonal mean scenario, while the prediction simulation for this time period resulted in vectorial capacity estimates very close to the values obtained by the observed rainfall simulation.

4.4 *The Effect of Temporal Distribution of Rainfall*

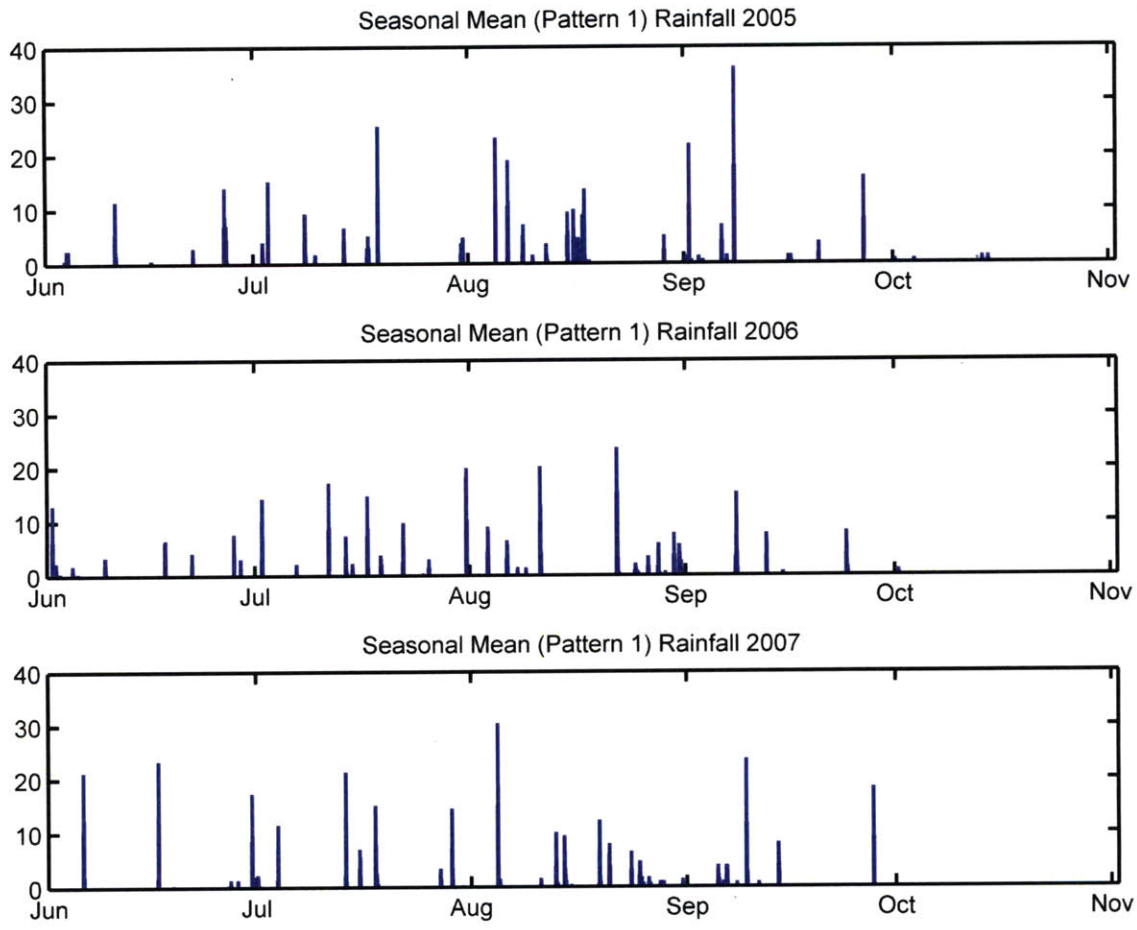

Figure 4.11 Seasonal Mean (Pattern 1) Rainfall

Finally, in order to examine the importance of the temporal distribution of rainfall, observed rainfall from 2005, 2006 and 2007 was modified such that the temporal distribution of rainfall follows pattern 1, but the total volume of rainfall for each two week period was equal to the seasonal mean volume for that period. Simulations were conducted using these three rainfall series, shown in Figure 4.11, and are referred to as the 'seasonal mean (pattern 1)' simulations. Since the purpose of this set of simulations was to compare differences resulting

from different rainfall patterns, 2006 data were used for all other environmental forcings for the 2005, 2006 and 2007 simulations so that any differences observed between the two simulations were solely the result of differences of temporal distribution of rainfall within two-week periods.

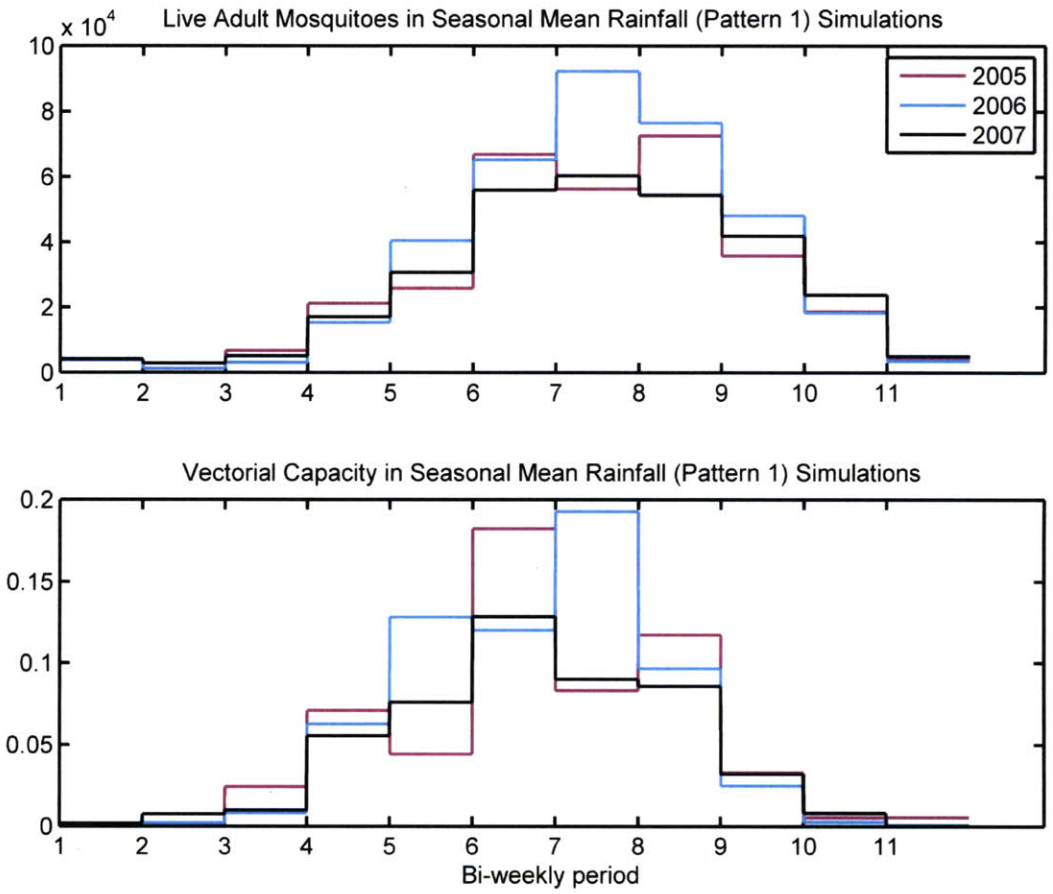

Figure 4.12 Comparison of Results with Differing Temporal Distribution of Rainfall

Results from the seasonal mean (pattern 1) simulations for 2005, 2006 and 2007 are shown in Figure 4.12. We see that the differences in temporal distribution of rainfall lead to significant differences in both mosquito populations and vectorial capacity. In the top graph,

showing live mosquitoes, we see that the pattern of rainfall observed in 2006 led to significantly more mosquitoes than the patterns observed in 2005 and 2007, with the estimated mosquitoes at the peak of the season reaching 1.5 times the number estimated for 2005 and 2007. The bottom graph in Figure 4.12 shows differences in vectorial capacity. Here, the differences between the two simulations are even more apparent, as even the seasonal distribution of vectorial capacity differs. In the 2005 simulation, the peak vectorial capacity occurred during the 6^{th} bi-weekly period; in the 2006 simulation, it occurred in the 7^{th} bi-weekly period, while in 2007, it arrived in the 5^{th} bi-weekly period. The magnitude of peak vectorial capacity in 2006 was 1.5 times as high as it was in 2007, but only 1.05 times as high as it was in 2005.

4.5 *Comparison with Field Observations*

The validity of our simulations and prediction can be assessed by comparing outputs to field observations. While it is impossible to quantify the absolute number of mosquitoes over the model domain, trends in mosquito abundance are reflected in trends in captured mosquitoes. During the rainy seasons of 2005 and 2006, six CDC miniature light traps were deployed in Banizoumbou, four inside residences, and two outdoors. Sampling occurred weekly, from 7 p.m. to 7 a.m. At the end of the sampling period, all mosquitoes were removed, and *Anopheles gambiae sensu lato* mosquitoes were identified and counted (Bomblies et al., 2008). The mosquito capture data had a pronounced dependence on lunar phase, as has been noted by numerous studies (e.g., (Horsfall, 1943),(Pratt, 1948)). This occurs because moonlight competes with bulbs from light traps as a mosquito attractant; the bulb attracts mosquitoes

from greater distances on moonless nights than on the brightly lit nights of a full moon. In order to correct for this bias, the number of captured mosquitoes at each sampling time were divided by the effective capture area of the trap for that night. The resulting mosquito densities are assumed to be independent of lunar phase ((Bomblies *et al.*, 2008)).

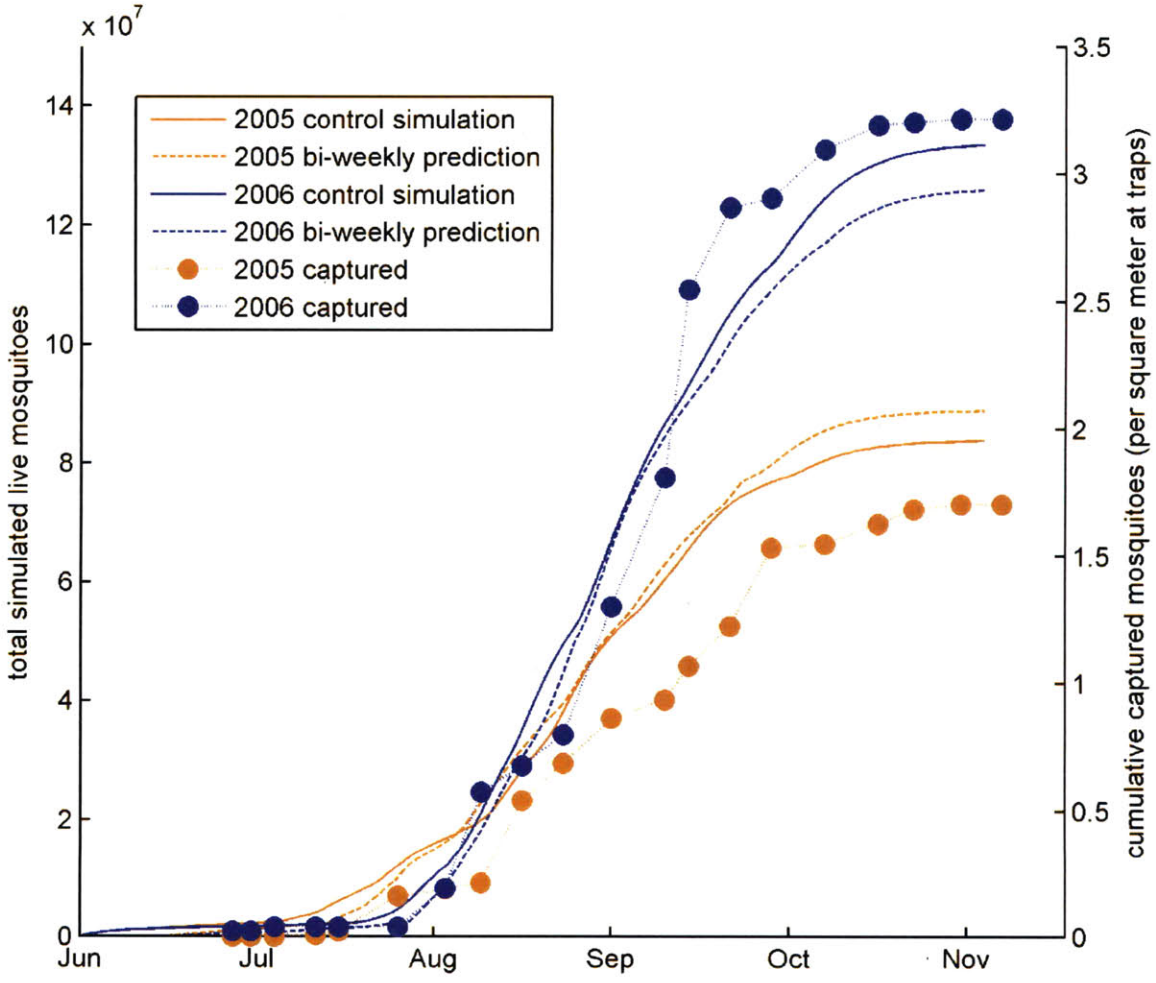

Figure 4.13 Simulated and Captured Mosquitoes

Cumulative simulated mosquitoes and cumulative density of captured mosquitoes are shown in Figure 4.13. In this figure, simulated mosquitoes are shown on the left axis, with the solid lines corresponding to the simulations using observed rainfall, and the dashed lines

corresponding to the bi-weekly prediction simulations. Mosquito densities are shown on the right axis. Data from 2005 are shown in orange, and data from 2006 are shown in blue. In this figure, we see that simulations using both observed and predicted rainfall resulted in roughly 1.5 times as many cumulative mosquitoes in 2006 than in 2005. This dramatic increase in simulated mosquito numbers is consistent with the observed mosquito densities. While the cumulative mosquitoes from the prediction simulations are further from the field observations than the control simulations for both, the predicted relative difference between 2005 and 2006 is clearly pronounced. This finding gives us confidence that HYDREMATS can serve as a reliable predictor of relative mosquito abundance and hence the potential for malaria transmission.

4.6 *Conclusion and Discussion*

This chapter explored a number of concepts. Section 4.2 proposed an effective method for forecasting mosquito populations and vectorial capacity two and four weeks in advance. Section 4.3 examined the importance of information regarding the history of rainfall in making predictions for future mosquito populations and vectorial capacity. Section 4.4 considered the effect of differences in temporal distribution of rainfall on HYDREMATS. Finally, section 4.5 compared HYDREMATS outputs from this chapter to field observations of mosquitoes.

The results from Section 4.2 showed that HYDREMATS can be effective at making short-term predictions of mosquito populations and vectorial capacity by using a simple method of predicting future rainfall, combined with known information on past rainfall. Specifically, HYDREMATS was able to replicate the seasonal cycle and interannual variations of both mosquito populations and vectorial capacities two and four weeks in advance. This could prove

extremely useful to malaria control programs by giving an advance warning of periods with higher than expected malaria transmission, allowing health officials to prevent or prepare for an oncoming epidemic. The two to four week warning could allow health officials to engage in activities such as redistributing limited medical staff and supplies, preparing health workers for prompt case detection and treatment, conducting vector control, and raising public awareness of heightened risk of malaria transmission.

In Section 4.3, we saw that using the history of observed rainfall was important in making an accurate prediction of mosquito populations, but somewhat less important in the prediction of vectorial capacity. This means that some of the information used by HYDREMATS in predicting mosquito populations is based on rainfall that is already known to have occurred, which increases the accuracy of the prediction, compared to a prediction based entirely on predicted future rainfall.

The investigation in Section 4.4 showed that equal volumes of rainfall with different temporal distributions lead to significantly different levels of malaria transmission. This reinforces the importance of considering the temporal distribution of rainfall as well as the magnitude of rainfall in order to assess malaria risk. It also emphasizes the value of hydrology based model such as HYDREMATS, since a statistical model which simply relates rainfall volume with malaria would have predicted identical results for the two rainfall series.

The cumulative number of simulated mosquitoes was significantly higher in 2006 than in 2005 for both the control and the prediction simulations. This relative difference in abundance agrees with field observations of adult mosquitoes. This finding indicates that HYDREMATS can serve as a reliable simulator and predictor of mosquito abundance.

5 Summary, Next Steps and Future Work

5.1 *Summary and Conclusions*

This thesis described the use of HYDREMATS in simulating mosquito populations and malaria transmission in West Africa. While past work using HYDREMATS was limited to simulations forced by meteorological data obtained by ground observations, this thesis sought to extend the use of HYDREMATS in space and time; in space by using satellite data instead of ground observations, and in time by using predictions for rainfall that has not yet occurred.

Chapter 1 provided a detailed description of HYDREMATS, and our study location in Banizoumbou, Niger. We suggested that HYDREMATS, being a mechanistic, hydrology based model, holds a significant advantage over other malaria transmission models in describing the details of village-scale malaria dynamics. HYDREMATS explicitly represents the causal pathways linking environmental variables to malaria transmission, which allows for better understanding of the malaria response to climate signals.

In Chapter 2, we determined that a minimum of one-hour temporal resolution of rainfall is required by HYDREMATS to accurately simulate surface runoff and pool formation, which is required for calculating the availability of mosquito breeding habitat and thus the potential for malaria transmission. When rainfall is averaged over a longer period of time, the resulting low-intensity rainfall is more readily infiltrated into the soil, and fails to produce water pools. The question of the temporal resolution of rainfall is relevant as we consider using a wider range of data products with HYDREMATS, as many such products are available only at 3-hour, daily, or even weekly temporal resolution.

In Chapter 3, we demonstrated that satellite-based estimates of rainfall can be used with HYDREMATS to simulate mosquito populations. One-hour resolution CMORPH data were corrected for a wet-bias and used to simulate mosquito transmission in Banizoumbou over two years. The resulting mosquito populations were very similar to outputs when the model was forced with ground observations of rainfall, showing that bias-corrected CMORPH can be used in areas where ground observations are not available.

In Chapter 4, HYDREMATS was used to make predictions of mosquito populations and vectorial capacity two and four weeks into the future. HYDREMATS was able to replicate the seasonal cycle and interannual variations of both mosquito populations and vectorial capacities two and four weeks in advance. We also demonstrate how past rainfall influences future mosquito populations, and how equal amounts of rainfall distributed differently over time can result in very different patterns of mosquito population and vectorial capacity.

5.2 *Next Steps and Future Work*

By demonstrating that HYDREMATS can be used with satellite-based rainfall data, we have made an important step towards extending the range of applicability of the model to areas where ground data is not available. However, HYDREMATS requires inputs in a number of other variables that would also have to be accounted for if we were to use HYDREMATS over new domains. The following section describes ways in which these variables could also be obtained from satellites or archived data sets. If this can be achieved, HYDREMATS could be used to model mosquito populations and malaria transmission dynamics anywhere in West

Africa, simply by collecting the relevant model inputs. This would prove to be a valuable tool to those seeking to understand the links between environmental variables and malaria transmission.

5.2.1 Archived and Satellite Data

In its current configuration, HYDREMATS requires distributed inputs of elevation, vegetation, soil type, and location of homes, as well as hourly space-invariant inputs of precipitation, temperature, relative humidity, radiation, wind speed and wind direction. In Chapter 3, we discussed how CMORPH estimates of precipitation can be used as the rainfall forcing for HYDREMATS. CMORPH combines the estimates from various passive microwave sensors and provides precipitation estimates every 30 minutes, which is a higher temporal resolution than can be obtained from individual satellites' microwave images. Currently, CMORPH incorporates data from the NOAA polar orbiting operational meteorological satellites, the U.S. Defense Meteorological Satellite Program (DMSP) satellites, and NASA's Tropical Rainfall Measuring Mission (TRMM) and Aqua satellites (Joyce *et al.*, 2010). Rainfall estimates from ERA-40 reanalysis data (S. Uppala *et al.*, 2005), which has been used in some of the studies reviewed in this thesis, is not recommended for use with HYDREMATS, as its temporal resolution of 6-hours would not allow for proper simulation of the hydrologic system. Here, we explore whether the remaining inputs could also be taken from satellite data or from archived datasets. Table 5-1 provides a summary of possible data sources for all inputs required by HYDREMATS.

Table 5-1 Data Sources for HYDREMATS Inputs

Variable	Type	Temporal requirement	Source	Spatial Resolution	Temporal Resolution	Reference
Vegetation	distributed	n/a	UMD Global Land Cover Classification	1km	n/a	Hansen et al., 2000
			IGBP-DIS DISCover	1km	n/a	Loveland et al., 2000
soil type	distributed	n/a	Global Soil Texture and Derived Water-Holding Capacities	1 degree	n/a	Webb et al, 2000
Topography	distributed	n/a	SRTM	90m	n/a	Farr et al., 2007
Households	distributed	n/a	Google Earth	various	n/a	http://earth.google.com
Precipitation	Lumped	hourly	CMORPH	8km	30 min	Joyce et al, 2004
Temperature	Lumped	hourly	ERA-40	1.125 deg	6 hr	Uppala et al., 2005
Humidity			ERA-Interim	0.703 deg	6 hr	Uppala et al., 2008
wind speed			Or			
wind direction			NCAR/NCEP reanalysis	2.5 deg	6 hr	Kalnay et al., 1996
surface radiation						

Vegetation inputs can be constructed from the supervised classification of multispectral images from satellites such as Landsat (Chander *et al.*, 2009). This process is similar to that described in Section 3.1.1.1 (McGraw & Tueller, 1983). Archived Landsat images are available globally at 30 meter spatial resolution, free of charge, at http://glovis.usgs.gov. The University of Maryland (UMD) Global Land Cover Classification (Hansen *et al.*, 2000) and the International Geosphere–Biosphere Programme Data and Information System (IGBP-DIS) DISCover (Loveland *et al.*, 2000) both provide global land cover data at 1km resolution, compiled from data from 1992–1993 from the Advanced Very High Resolution Radiometer (AVHRR). The two differ in

their classification method; the UMD used a supervised classification for the entire world, while DISCover used an unsupervised clustered classification scheme, calculating each continent separately (Hansen & Reed, 2000).

HYDREMATS is extremely sensitive to specified topography, as this determines the flow of water in the overland flow component of the model. 90 meter spatial resolution data are available between 60°N and 56°S from the Shuttle Radar Topography Mission (SRTM) at http://www2.jpl.nasa.gov/srtm/ (Farr *et al.*, 2007). Higher resolution digital elevation models can be constructed using interferometry on radar images obtained from commercial satellites. For example, a 2 to 5 meter resolution DEM made from stereo Ikonos images can be purchased from the Satellite Image Corporation.

The location of households can be determined by a high resolution satellite image. Currently, Landsat images do not have sufficient spatial resolution to detect small households (Zhang & Wang, 2002). However, the image shown in Figure 1.3 shows that households can be easily located using a Quickbird panchromatic image, with 60cm resolution. While this commercial satellite product may be prohibitively expensive, similar images can be obtained free of charge using Google Earth, as shown in Figure 5.1. While the resolution of Google Earth images varies by location, all land areas appear to be high enough resolution to see individual residences.

In areas where knowledge of local soil compositions is not available, soil characteristics can be taken from the Global Soil Texture and Derived Water-Holding Capacities data set (Webb et al., 2000). This data set provides information about soil thickness and types at a 1° by 1° spatial resolution, compiled from the Food and Agriculture Organization of the United

Nations/United Nations Educational, Scientific, and Cultural Organization (FAO/UNESCO) Soil Map of the World, Vols. 2–10 [1971–1981] (FAO-UNESCO, 1988) and combined with the World Soil Data File (Zobler, 1986).

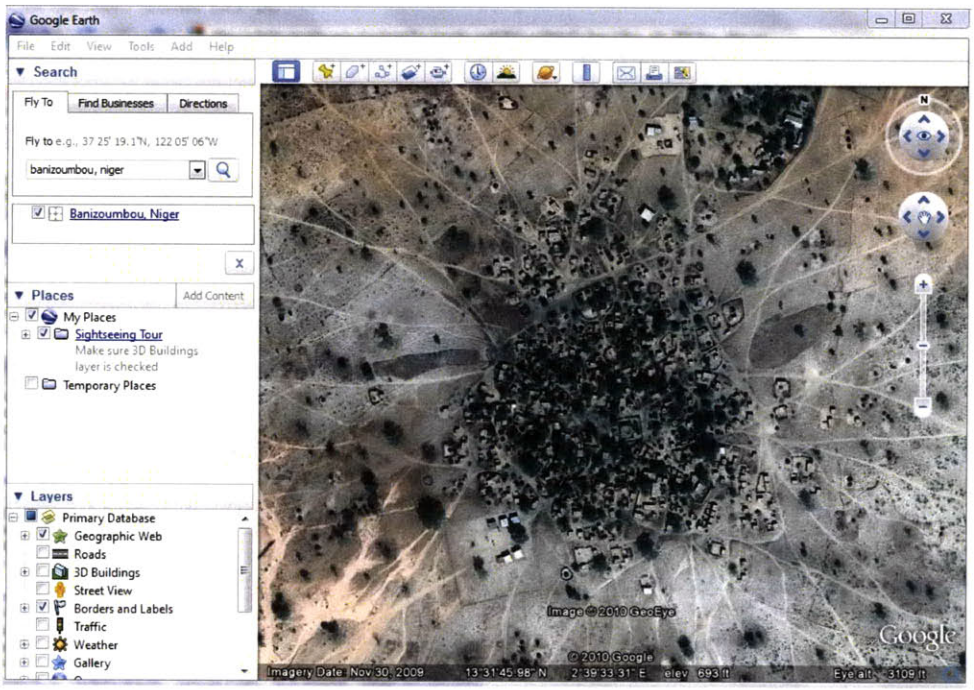

Figure 5.1 Image of Banizoumbou taken from Google Earth (Image copyright GeoEye, 2010)

Temperature, relative humidity, surface radiation, wind speed and wind direction data are required near real-time. These variables can be obtained from reanalysis data at 6 hour temporal resolution from ERA-40 (S. Uppala *et al.*, 2005) for the period 1957-2002 with a spatial resolution of roughly 1.125°, and ERA-Interim (S. Uppala *et al.*, 2008) for 1989-present, with a spatial resolution of roughly 0.703°. The National Centers for Environmental Protection (NCEP) and the National Center for Atmospheric Research (NCAR) also have a reanalysis product, available 1948-present for 6-hour periods at 2.5° resolution (Kalnay *et al.*, 1996). No studies

have been conducted to determine whether this spatial or temporal resolution is sufficient for use with HYDREMATS.

5.2.2 Forecasting Mosquito Populations and Vectorial Capacity using HYDREMATS

The investigation in Chapter 4 predicted malaria transmission two and four weeks into the future using a simple method of forecasting rainfall based on ground observations of past rainfall. An obvious extension of the work done in this thesis would be to combine the findings of Chapter 3 and 4 and use rainfall data from satellites to make the predictions for mosquito populations. The previous section proposed a system by which HYDREMATS could simulate mosquito populations and malaria transmission using only satellite data and archived data sets, so that it could be used throughout West Africa. By applying our methodology of predicting malaria transmission using HYDREMATS, we could not only simulate but also forecast Malaria in villages across West Africa.

Our study only attempted to forecast rainfall, but continued to use real time data for temperature, relative humidity, wind speed and direction and radiation. Future studies could find simple ways to make short term predictions for each of these variables, based on a combination of climatology and past observations within that season. Seasonal forecasts or medium range weather forecasts such as those provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) could be used for these variables.

While a warning of increased malaria transmission four-weeks in advance has great value, there would be advantages in even earlier warnings. We did not attempt to make forecasts longer than four weeks, as our simple method for forecasting rainfall is not likely to be

effective in making longer term forecasts. However, there is potential for medium-range weather forecasts to be used with HYDREMATS in order to make detailed predictions of malaria transmission months in advance. For example, the DEMETER multi-model ensemble system for seasonal forecasting provided daily values of radiation and precipitation, and 6-hourly values of temperature, humidity, wind speed and wind direction(Palmer *et al.*, 2004), and has been used with a malaria model(Morse *et al.*, 2005). While this resolution of precipitation is too coarse for HYDREMATS, further advances in medium range weather forecasting or downscaling of existing forecasts have great potential to be used with HYDREMATS to make seasonal predictions of malaria.

References

Achee, N. L., Grieco, J. P., Masuoka, P., Andre, R. G., Roberts, D. R., Thomas, J., et al. (2006). Use of remote sensing and geographic information systems to predict locations of anopheles darlingi-positive breeding sites within the Sibun River in Belize, Central America. *Journal of Medical Entomology, 43*(2), 382-392.

Aregawi, M., Cibulskis, R., & Otten, M. (2008). *World malaria report 2008* World Health Organization.

Beck, L. R., Rodriguez, M. H., Dister, S. W., Rodriguez, A. D., Rejmankova, E., Ulloa, A., et al. (1994). Remote sensing as a landscape epidemiologic tool to identify villages at high risk for malaria transmission. *The American Journal of Tropical Medicine and Hygiene, 51*(3), 271.

Beck, L. R., Rodriguez, M. H., Dister, S. W., Rodriguez, A. D., Washino, R. K., Roberts, D. R., et al. (1997). Assessment of a remote sensing-based model for predicting malaria transmission risk in villages of Chiapas, Mexico. *The American Journal of Tropical Medicine and Hygiene, 56*(1), 99.

Bo, Z., Islam, S., & Eltahir, E. (1994). *Water Resources Research, 30*(12), 3423-3435.

Bogh, C., Lindsay, S. W., Clarke, S. E., Dean, A., Jawara, M., Pinder, M., et al. (2007). High spatial resolution mapping of malaria transmission risk in the Gambia, West Africa, using LANDSAT TM satellite imagery. *The American Journal of Tropical Medicine and Hygiene, 76*(5), 875.

Bomblies, A., Duchemin, J. B., & Eltahir, E. A. B. (2008). Hydrology of malaria: Model development and application to a Sahelian village. *Water Resour.Res, 44*

Bomblies, A., Duchemin, J. B., & Eltahir, E. A. B. (2009). A mechanistic approach for accurate simulation of village-scale malaria transmission. *Malaria Journal, 8*(1), 223.

Bomblies, A., & Eltahir, E. A. B. (2009). Assessment of the impact of climate shifts on malaria transmission in the sahel. *EcoHealth,*6(3), 426-37.

Ceccato, P., Bell, M., Blumenthal, M., Connor, S., Dinku, T., Grover-Kopec, E., et al. (2006). Use of remote sensing for monitoring climate variability for integrated early warning systems: Applications for human diseases and desert locust management. *IEEE International Conference on Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006,* 270-274.

Ceccato, P., Connor, S., Jeanne, I., & Thomson, M. (2005). Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. *Parassitologia, 47*(1), 81-96.

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for landsat MSS, TM, ETM , and EO-1 ALI sensors. *Remote Sensing of Environment, 113*(5), 893-903.

Craig, M. H., Snow, R. W., & le Sueur, D. (1999). A climate-based distribution model of malaria transmission in sub-saharan africa. *Parasitology Today, 15*(3), 105-111.

Cross, E. R., Newcomb, W. W., & Tucker, C. J. (1996). Use of weather data and remote sensing to predict the geographic and seasonal distribution of phlebotomus papatasi in southwest asia. *The American Journal of Tropical Medicine and Hygiene, 54*(5), 530.

DaSilva, J., Garanganga, B., Teveredzi, V., Marx, S. M., Mason, S. J., & Connor, S. J. (2004). Improving epidemic malaria planning, preparedness and response in southern Africa. Report on the 1st southern African regional epidemic outlook forum, Harare, Zimbabwe, 26-29 September, 2004. *Malaria Journal, 3*, 37.

de Castro, M. C., Monte-Mór, R. L., Sawyer, D. O., & Singer, B. H. (2006). Malaria risk on the Amazon frontier. *Proceedings of the National Academy of Sciences, 103*(7), 2452.

Depinay, J. M., Mbogo, C. M., Killeen, G., Knols, B., Beier, J., Carlson, J., et al. (2004). A simulation model of African anopheles ecology and population dynamics for the analysis of malaria transmission. *Malaria Journal, 3*, 29.

Dister, S. W., Fish, D., Bros, S. M., Frank, D. H., & Wood, B. L. (1997). Landscape characterization of peridomestic risk for lyme disease using satellite imagery. *American Journal of Tropical Medicine and Hygiene, 57*(6), 687-692.

Diuk-Wasser, M., Bagayoko, M., Sogoba, N., Dolo, G., Touré, M., Traoré, S., et al. (2004). Mapping rice field anopheline breeding habitats in Mali, West Africa, using landsat ETM sensor data. *International Journal of Remote Sensing, 25*(2), 359.

Eisele, T. P., Keating, J., Swalm, C., Mbogo, C. M., Githeko, A. K., Regens, J. L., et al. (2003). Linking field-based ecological data with remotely sensed data using a geographic information system in two malaria endemic urban areas of kenya. *Malaria Journal, 2*(1), 44.

FAO-UNESCO. (1988). Soil map of the world. *World Soil Resources Rep., 60*, 119.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar topography mission. *Reviews of Geophysics,* 45(2).

Finnerty, B. D., Smith, M. B., Seo, D. J., Koren, V., & Moglen, G. E. (1997). Space-time scale sensitivity of the sacramento model to radar-gage precipitation inputs. *Journal of Hydrology, 203*(1-4), 21-38.

Garrett-Jones, C., & Grab, B. (1964). The assessment of insecticidal impact on the malaria mosquito's vectorial capacity, from data on the proportion of parous females. *Bull Wld Hlth Org, 31*, 71-86.

Gianotti, R., Bomblies, A., & Eltahir, E. (2008). Using hydrologic modeling to screen potential environmental management methods for malaria vector control in niger. *AGU Fall Meeting Abstracts,* 0494.

Gianotti, R. L., Bomblies, A., Dafalla, M., Issa-Arzika, I., Duchemin, J. B., & Eltahir, E. A. (2008). Efficacy of local neem extracts for sustainable malaria vector control in an african village. *Malaria Journal, 7*, 138.

Glass, G. E., Yates, T. L., Fine, J. B., Shields, T. M., Kendall, J. B., Hope, A. G., et al. (2002). Satellite imagery characterizes local animal reservoir populations of sin nombre virus in the southwestern united states. *Proceedings of the National Academy of Sciences, 99*(26), 16817.

Grover-Kopec, E., Kawano, M., Klaver, R. W., Blumenthal, B., Ceccato, P., & Connor, S. J. (2005). An online operational rainfall-monitoring resource for epidemic malaria early warning systems in africa. *Malaria Journal, 4*, 6.

Hansen, M., DeFries, R., Townshend, J., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing, 21*(6), 1331-1364.

Hansen, M., & Reed, B. (2000). A comparison of the IGBP DISCover and university of maryland 1 km global land cover products. *International Journal of Remote Sensing, 21*(6), 1365-1373.

Hay, S., Renshaw, M., Ochola, S. A., Noor, A. M., & Snow, R. W. (2003). Performance of forecasting, warning and detection of malaria epidemics in the highlands of western kenya. *Trends in Parasitology, 19*(9), 394-399.

Hay, S. I., Snow, R. W., & Rogers, D. J. (1998). Predicting malaria seasons in kenya using multitemporal meteorological satellite sensor data. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 92*(1), 12-20.

Hay, S. I., Were, E. C., Renshaw, M., Noor, A. M., Ochola, S. A., Olusanmi, I., et al. (2003). Forecasting, warning, and detection of malaria epidemics: A case study. *The Lancet, 361*(9370), 1705-1706.

Hay, S., Snow, R., & Rogers, D. (1998). From predicting mosquito habitat to malaria seasons using remotely sensed data: Practice, problems and perspectives. *Parasitology Today, 14*(8), 306-313.

Horsfall, W. R. (1943). Some responses of the malaria mosquito to light. *Annals of the Entomological Society of America, 36*(1), 41-45.

Hoshen, M., & Morse, A. (2004). A weather-driven model of malaria transmission. *Malaria Journal, 3*(1), 32.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology, 5*(3), 487-503.

Joyce, R. J., Xie, P., Yarosh, Y., Janowiak, J. E., & Arkin, P. A. (2010). CMORPH: A "Morphing" approach for high resolution precipitation product generation. *Satellite Rainfall Applications for Surface Hydrology,* 23-37.

Kalluri, S., Gilruth, P., Rogers, D., & Szczur, M. (2007). Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: A review. *PLoS Pathog, 3*, 1361-1371.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society, 77*(3), 437-471.

Kaya, S., Pultz, T., Mbogo, C., Beier, J., & Mushinzimana, E. (2007). The use of radar remote sensing for identifying environmental factors associated with malaria risk in coastal Kenya. *Canada Centre for Remote Sensing.Consultado, 11.*

Kilian, A. H. D., Langi, P., Talisuna, A., & Kabagambe, G. (1999). Rainfall pattern, El Niño and malaria in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 93*(1), 22-23.

Kitron, U., Otieno, L. H., Hungerford, L. L., Odulaja, A., Brigham, W. U., Okello, O. O., et al. (1996). Spatial analysis of the distribution of tsetse flies in the Lambwe valley, Kenya, using Landsat TM satellite imagery and GIS. *Journal of Animal Ecology, 65*(3), 371-380.

Krajewski, W. F., Lakshmi, V., Georgakakos, K. P., & Jain, S. C. (1991). A Monte-Carlo study of rainfall sampling effect on a distributed catchment model. *Water Resour Res* 1991;27:119–28.

Lal, A. M. W. (1998). Performance comparison of overland flow algorithms. *Journal of Hydraulic Engineering, 124*(4), 342-349.

Laws KB, Janowiak JE, Huffman G. 2004. Verification of rainfall estimates over Africa using RFE, NASA MAPRT-RT and CMORPH. *Preprints of 84th AMS Annual Meeting, Paper P2.2 in 18th Conf. on Hydrology, 11–15 January 2004, Seattle, WA.* 4 pp.

Lindsay, S. W., Bødker, R., Malima, R., Msangeni, H. A., & Kisinza, W. (2000). Effect of 1997–98 El Niño on highland malaria in Tanzania. *The Lancet, 355*(9208), 989-990.

Linthicum, K. J., Anyamba, A., Tucker, C. J., Kelley, P. W., Myers, M. F., & Peters, C. J. (1999). Climate and satellite indicators to forecast rift valley fever epidemics in Kenya. *Science, 285*(5426), 397.

Loveland, T., Reed, B., Brown, J., Ohlen, D., Zhu, Z., Yang, L., et al. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing, 21*(6), 1303-1330.

Malone, J. B., Huh, O. K., Fehler, D. P., Wilson, P. A., Wilensky, D. E., Holmes, R. A., et al. (1994). Temperature data from satellite imagery and the distribution of schistosomiasis in Egypt. *The American Journal of Tropical Medicine and Hygiene, 50*(6), 714.

Masuoka, P. M., Claborn, D. M., Andre, R. G., Nigro, J., Gordon, S. W., Klein, T. A., et al. (2003). Use of IKONOS and Landsat for malaria control in the Republic of Korea. *Remote Sensing of Environment, 88*(1), 187-194.

Matthews, E. (1983). Global vegetation and land use: New high-resolution data bases for climate studies. *Journal of Applied Meteorology, 22*(3), 474-487.

McGraw, J. F., & Tueller, P. T. (1983). Landsat computer-aided analysis techniques for range vegetation mapping. *Journal of Range Management, 36*(5), 627-631.

Morse, A. P., Doblas-Reyes, F. J., Hoshen, M. B., Hagedorn, R., & PALMER, T. (2005). A forecast quality assessment of an end-to-end probabilistic multi-model seasonal forecast system using a malaria model. *Tellus A, 57*(3), 464-475.

Mushinzimana, E., Munga, S., Minakawa, N., Li, L., Feng, C. C., Bian, L., et al. (2006). Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands. *Malaria Journal, 5*, 13.

Nihei, N., Hashida, Y., Kobayashi, M., & Ishii, A. (2002). Analysis of malaria endemic areas on the Indochina peninsula using remote sensing. *Jpn J Infect Dis, 55*(5), 160-166.

Omumbo, J., Hay, S., Goetz, S., Snow, R., & Rogers, D. (2002). Updating historical maps of malaria transmission intensity in east africa using remote sensing. *Photogrammetric Eng Remote Sens, 68*, 161-166.

Palmer, T., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délécluse, P., et al. (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society, 85*(6), 853-872.

Patz, J. A., Strzepek, K., Lele, S., Hedden, M., Greene, S., Noden, B., et al. (1998). Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya. *Tropical Medicine & International Health : TM & IH, 3*(10), 818-827.

Pessoa, M. L., Bras, R. L., & Williams, E. R. (1993). Use of weather radar for flood forecasting in the sieve river basin: A sensitivity analysis. *Journal of Applied Meteorology, 32*(3), 462-475.

Pollard, D., & Thompson, S. L. (1995). Use of a land-surface-transfer scheme (LSX) in a global climate model: The response to doubling stomatal resistance. *Global and Planetary Change, 10*(1-4), 129-161.

Pope, K. O., Rejmankova, E., Savage, H. M., Arredondo-Jimenez, J. I., Rodriguez, M. H., & Roberts, D. R. (1994). Remote sensing of tropical wetlands for malaria control in Chiapas, Mexico. *Ecological Applications, 4*(1), 81-90.

Pope, K., Sheffner, E., Linthicum, K., Bailey, C., Logan, T., Kasischke, E., et al. (1992). Identification of central Kenyan rift valley fever virus vector habitats with Landsat TM and evaluation of their flooding status with airborne imaging radar. *Remote Sensing of Environment (USA),*

Pratt, H. D. (1948). Influence of the moon on light trap collections of anopheles albimanus in Puerto Rico. *Journal.National Malaria Society, 7*(3), 212-220.

Rahman, A., Kogan, F., & Roytman, L. (2006). Analysis of malaria cases in Bangladesh with remote sensing data. *The American Journal of Tropical Medicine and Hygiene, 74*(1), 17.

Rejmankova, E., Roberts, D. R., Pawley, A., Manguin, S., & Polanco, J. (1995). Predictions of adult anopheles albimanus densities in villages based on distances to remotely sensed larval habitats. *The American Journal of Tropical Medicine and Hygiene, 53*(5), 482.

Ritchie, S. A. (1993). Application of radar rainfall estimates for surveillance of aedes taeniorhynchus larvae. *Journal of the American Mosquito Control Association, 9*(2), 228-231.

Rogers, D. J., Randolph, S. E., Snow, R. W., & Hay, S. I. (2002). Satellite imagery in the study and forecast of malaria. *Nature, 415*(6872), 710-715.

Rogers, D., & Randolph, S. (1991). Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature, 351,* 739-741.

Rosenfeld, D., & Mintz, Y. (1988). Evaporation of rain falling from convective clouds as derived from radar measurements. *Journal of Applied Meteorology, 27*(3), 209-215.

Service, M. (1993). Mosquito ecology: Field sampling methods. 2nd edn. Elsevier Applied Sciences, London, UK.

Sharma, V. P., Nagpal, B. N., Srivastava, A., Adiga, S., & Manavalan, P. (1996). Estimation of larval production in Sanjay lake and its surrounding ponds in Delhi, India using remote sensing technology. *The Southeast Asian Journal of Tropical Medicine and Public Health, 27*(4), 834-840.

Thomas, C., & Lindsay, S. (2000). Local-scale variation in malaria infection amongst rural Gambian children estimated by satellite remote sensing. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 94*(2), 159-163.

Thomson, M., Indeje, M., Connor, S., Dilley, M., & Ward, N. (2003). Malaria early warning in Kenya and seasonal climate forecasts. *The Lancet, 362*(9383), 580-580.

Thomson, M. C., Mason, S. J., Phindela, T., & Connor, S. J. (2005). Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. *The American Journal of Tropical Medicine and Hygiene, 73*(1), 214.

Thomson, M., Connor, S., D'Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M., et al. (1999). Predicting malaria infection in gambian children from satellite data and bed net use surveys: The importance of spatial correlation in the interpretation of results. *The American Journal of Tropical Medicine and Hygiene, 61*(1), 2.

Thomson, M., Doblas-Reyes, F., Mason, S., Hagedorn, R., Connor, S., Phindela, T., et al. (2006). Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature, 439*(7076), 576-579.

Thomson, M. C., Connor, S. J., Milligan, P. J., & Flasse, S. P. (1996). The ecology of malaria--as seen from earth-observation satellites. *Annals of Tropical Medicine and Parasitology, 90*(3), 243-264.

Tian, Y., & Peters-Lidard, C. D. (2007). Systematic anomalies over inland water bodies in satellite-based precipitation estimates. *Geophysical Research Letters, 34*(14), L14403.

Uppala, S., Dee, D., Kobayashi, S., Berrisford, P., & Simmons, A. (2008). Towards a climate data assimilation system: Status update of ERA-interim. *ECMWF Newsletter, 115*, 12-18.

Uppala, S., Kallberg, P., Simmons, A., Andrae, U., da Costa Bechtold, V., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society, 131*(612), 2961-3012.

Wijesundera Mde, S. (1988). Malaria outbreaks in new foci in Sri Lanka. *Parasitology Today (Personal Ed.), 4*(5), 147-150.

Winchell, M., Gupta, H. V., & Sorooshian, S. (1998). On the simulation of infiltration-and saturation-excess runoff using radar-based rainfall estimates: Effects of algorithm uncertainty and pixel aggregation. *Water Resources Research, 34*(10)

Zeweldi, D. A., & Gebremichael, M. (2009). Evaluation of CMORPH precipitation products at fine Space–Time scales. *Journal of Hydrometeorology, 10*(1), 300-307.

Zhang, Q., & Wang, J. (2002). Detection of buildings from landsat-7 ETM and SPOT panchromatic data in beijing, china. *2002 IEEE International Geoscience and Remote Sensing Symposium, 2002. IGARSS'02, , 5*

Zobler, L. 1986. A World Soil File for Global Climate Modelling. NASA Technical Memorandum # 87802. NASA Goddard Institute for Space Studies, New York, New York, U.S.A