Estimating malaria burden in Nigeria: a geostatistical modelling approach

Nnadozie Onyiri

Swiss Tropical and Public Health Institute, Basel, Switzerland

Abstract

This study has produced a map of malaria prevalence in Nigeria based on available data from the Mapping Malaria Risk in Africa (MARA) database, including all malaria prevalence surveys in Nigeria that could be geolocated, as well as data collected during fieldwork in Nigeria between March and June 2007. Logistic regression was fitted to malaria prevalence to identify significant demographic (age) and environmental covariates in STATA. The following environmental covariates were included in the spatial model: the normalized difference vegetation index, the enhanced vegetation index, the leaf area index, the land surface temperature for day and night, land use/land-cover (LULC), distance to water bodies, and rainfall. The spatial model created suggests that the two main environmental covariates correlating with malaria presence were land surface temperature for day and rainfall. It was also found that malaria prevalence increased with distance to water bodies up to 4 km. The malaria risk map estimated from the spatial model shows that malaria prevalence in Nigeria varies from 20% in certain areas to 70% in others. The highest prevalence rates were found in the Niger Delta states of Rivers and Bayelsa, the areas surrounding the confluence of the rivers Niger and Benue, and also isolated parts of the north-eastern and north-western parts of the country. Isolated patches of low malaria prevalence were found to be scattered around the country with northern Nigeria having more such areas than the rest of the country. Nigeria's belt of middle regions generally has malaria prevalence of 40% and above.

Introduction

In spite of being entirely preventable, malaria has a high level of mortality and is the world's most prevalent parasitic disease. It is caused by infection with single-celled parasites of the genus *Plasmodium*, which are transmitted by the bite of *Anopheles* mosquitoes. Apart from the endemic tropical and sub-tropical regions, malaria was once widespread in North America and other temperate countries. Today the disease occurs mostly in sub-Saharan Africa and Southeast Asia. In 2013, 97 countries had ongoing malaria transmission (WHO, 2013). According to the World Health Organization (WHO, 2014), malaria is the second leading cause of death from infectious diseases in Africa after HIV/AIDS, with Nigeria and the Democratic Republic of Congo (DRC) accounting for 40% of the global malaria deaths. Almost 20% of all deaths of children under-five in Africa are due to malaria (WHO, 2014).

Each year, over 500 million people suffer clinical malaria episodes caused by *P. falciparum* infection alone, resulting in a conservative estimate of 1 million deaths annually (Guinovart *et al*., 2006; Vaughan *et al*., 2008). Despite concerted efforts, of which the Roll Back Malaria programme (http://www.rollbackmalaria.org/) is an example, malaria remains a major health challenge. Between 2000 and 2012, the scale-up of interventions helped reduce malaria incidence rates by 25% globally, while reaching as high as 31% in WHO's African region (WHO, 2013). The global malaria mortality rate was reduced by 42% during the same period, while the decrease in African was 49% (WHO, 2013). During the same period, an estimated 3.3 million lives were saved by the scaled-up malaria interventions, 3 million of which (90%) concerned the under-five age group in sub-Saharan Africa (WHO, 2013). In Nigeria, statistics shows that malaria accounts for 25% of the under-five mortality, 30% of childhood mortality and 11% of maternal mortality (Okonko *et al*., 2009). All Nigerians are at risk of malaria and the problem is compounded by the increasing resistance of malaria to hitherto cost-effective drugs (Okonko *et al*., 2009).

Describing spatial and temporal variation in transmission and disease risk is fundamental to epidemiological understanding and control

of malaria. Risk maps are, by definition, outcomes of models of disease transmission based on spatial and temporal data. These models incorporate, by varying degrees, epidemiological, entomological, climate and environmental information (Kitron, 2000). Decades of experience confirm that successful malaria control depends on accurate identification and geographical reconnaissance of high-risk areas (Wijeyaratne, 1999; Carter *et al*., 2000). In the past, malaria risk maps at different geographical levels were largely based on expert opinion based on limited data, crude climate isolines with no clear and reproducible numerical definition (Craig *et al*., 1999). In recent years, the availability of new data sources such as remote sensing (RS) and mapping tools, such as computerized geographic information systems (GIS) for quantitative analysis of spatial data, have provided an unprecedented amount of information and increased capability to describe, predict and communicate risk and outcome of interventions (Hay *et al*., 2000; Kitron, 2000; Thomson and Connor, 2000; Berquist, 2001).

Measures that might be mapped include categories of endemicity (*e.g.* unstable, mesoendemic or holoendemic), vector density and capacity, entomological inoculation rate (EIR) and incidence of disease. However, although malaria endemicity can vary widely over only short distances, most of these measures have only been studied in a few, widely separated localities. In general, results from different sites differ. The most broadly available measure is point prevalence assessed by microscopy. Estimates of malaria prevalence at unsampled locations can be made by incorporating information from environmental covariates (Hay *et al*., 2000). The precision of such estimates can be further improved by using spatial smoothing or geostatistical methods (Ribeiro *et al.,* 1996; Kleinschmidt *et al*., 2000, 2001a, 2001b; Diggle *et al*., 2002), while Bayesian geostatistical methods have demonstrated their value for mapping childhood malaria risk in the Gambia (Diggle *et al*., 2002) and generally for relating infant mortality to malaria risk (Gemperli and Vounatsou, 2004). Spatial statistical models have also been used to produce malaria maps of Mali (Kleinschmidt *et al*., 2000) as well as for the whole of West Africa (Kleinschmidt *et al*., 2001a). All these analyses modelled the prevalence data directly without taking into account age-dependence of the malaria risk. Malaria prevalence data are usually reported by age, but with different age-groupings used in different surveys. Direct mapping of age-prevalence data therefore involves choosing a target age-group (with some flexibility in the choice of age-category boundaries), and discarding data for other age-groups and for sites where data for the age-group are not available. A number of malaria distribution maps are available for Africa based on climatic and other environmental predictors of malaria transmission (Craig *et al*., 1999; Snow *et al*., 1997; Kleinschmidt *et al*., 2000; Rogers *et al*., 2002; Gemperli, 2003; Gosoniu *et al.,* 2006; Gething *et al*., 2011). However, they make little or no use of the data from field surveys of malaria prevalence, which form the largest body of relevant information.

The Mapping Malaria Risk in Africa (MARA) project is a collaborative network of key African scientists and institutions with the aim of providing empirical risk maps of malaria in Africa (Snow *et al*., 1999). Initially, this involved the development of continent-wide climate-based theoretical models of climatic suitability (Craig *et al*., 1999) and the collection of parasite prevalence data to validate and/or improve these models. The MARA database was established in 1996 using published and unpublished malaria survey data compiled by a collaborative network of African scientists and institutions with the aim of providing an atlas of malaria for evidence-based and targeted malaria control in Africa. This is the most comprehensive database on malaria data in Africa so far, but information with respect to Nigeria, comprising 2581 age-specific surveys carried out at 126 distinct locations between 1930 until 2007, is comparatively scarce. As malaria is an environmentally-driven disease, the Swiss Tropical and Public Health Institute (Swiss TPH), in collaboration with its partners at the Malaria Research and Training Centre (MRTC) in Bamako, Mali initiated an updating the MARA database for Nigeria. This is important, not only because recent climate changes, but also given the absence of recent data from large parts of northern Nigeria. We aimed to estimate malaria prevalence in areas without survey data using rigorous statistical methods with environmental parameters as predictors. In this study, we used Bayesian geostatistical approaches to assess the malaria-environmental relationship for the purpose of malaria risk mapping.

Table 1. Summary of fieldwork in Nigeria.

Location	Number of surveys (years)	Type of survey examined	Number of individuals
University of Port-Harcourt	3 (2000 to 2006)	Community-based surveys	50 to 500
University of Calabar	1 (2004)	Published community-based	594
University of Benin	4 (1998 to 2005)	Published surveys	120 to 500
Imo State University	15 (1999 to 2006)	Unpublished theses (5) surveys, published surveys (10)	134 to 1200
University of Nigeria	4 (1998 to 2006)	Unpublished thesis work	80 to 500
University of Lagos	10 (1999 to 2004)	NA	150 to 1563
Nigerian Institute for Medical Research	2 (2000 to 2006)	Published surveys	165 to 350
University of Ibadan	3 (1996 and 2003)	Published surveys	100 to 405
Ahmadu Bello University	8 (2000 to 2005)	Unpublished theses	150 to 220
University of Abuja	1 (2000)	Published survey	200
University of Jos/Federal University of Technology Yola	5 (2000 to 2005)	Published works (two of the surveys were malaria drug efficacy trials)	120 to 250

NA, not available.

Materials and Methods

The data collected over the years by MARA up to 1999 form part of the data used for this study. As data for Nigeria were sparse or relatively old, more data had to be collected from the year 2000 and onwards and these included both published and unpublished survey data.

Malaria prevalence data

The fieldwork was carried out from March 1 to June 30 2007 and was aimed at obtaining unpublished malaria prevalence data from various Nigerian sources. The data collection was done in phases according to Nigeria's six geopolitical zones. For the south-south zone, I visited the Ministries of Health in Calabar (Cross River State), Port-Harcourt (Rivers State) and Benin (Edo State). However, these Ministries had no survey data in their archives except recent malaria incidence data. I also visited the local universities in these states obtaining unpublished malaria survey data from all three states with the University of Port-Harcourt providing the greatest number.

In the southeastern zone, University of Nigeria campuses at Nsukka and Enugu city (Enugu State) were visited. Imo State University provided a much larger number of malaria surveys than any other institution in the entire country. From this location, I got data not only from Imo State and the rest of the southeast zone, but also from just about every other part of the country. The visit to the southeastern zone was followed by a visit to the Capital of Nigeria, Abuja, to get a national perspective on the malaria situation. From the Roll Back Malaria Office (a division of the Federal Ministry of Health), I got only malaria incidence data, which were ultimately not used for this study. However I was able to get some malaria prevalence data from the University of Abuja.

Data from the southwestern zone was collected from Ibadan (Oyo State) and Lagos city. At the University of Ibadan, our data came from the College of Medicine and the Department of Zoology. In Lagos, the data was sourced from the Nigerian Institute of Medical Research and the Department of Zoology of the University of Lagos. These two institutions are both strong in malaria research.

Finally, data from the North (the north-eastern, north-central and north-western zones) were obtained from the University of Abuja (as mentioned earlier), the University of Jos, Ahmadu Bello University Zaria, and the Federal University of Technology in Yola. The bulk of the data came from Ahmadu Bello University, which provided much of the data from the northeastern and north-central zones. A summary of the data collection in Nigeria is shown in Table 1.

Additional survey data were obtained through online searches. Longitude and latitude co-ordinates were determined for each parasitological survey using the *Geonames geographical database* (http://www.geonames.org/). Each location was described as a set of either a first, second or third order administrative region in Nigeria.

Environmental data

The following factors were considered: the normalized difference vegetation index (NDVI), the enhanced vegetation index (EVI), the leaf area index (LAI), the amount of rainfall, the land surface temperature for day and night (LST_{day} and LST_{night}), respectively, land use/land-cover (LULC), elevation and distance to the nearest water body. The databases from which the environmental data emanate are given in Table 2.

Environmental predictors were extracted from RS sources at spatial and temporal resolutions and shown in Table 2. To take into account of the elapsing time between the climatic suitability for malaria transmission and parasitaemia, the climatic data were gathered for different periods (up to one year) prior to the survey starting from July 2006. The variables LST, NDVI, EVI, LAI and land-cover were downloaded from the Moderate Resolution Imaging Spectroradiometer (MODIS) (http://modis.gsfc.nasa.gov/), from the U.S. Geological Survey (USGS) Land Processes Distributed Active Archive Center (LP DAAC). LST data were extracted as averages over 8-day periods at 1-km spatial resolution. The NDVI was obtained as a 16-day average at 0.25-km spatial resolution. EVI data, available from MODIS for the year 2006, represents an improvement on the NDVI because it corrects some distortions in the reflected light caused by particles in the air as well as ground cover below the vegetation (Weier and Herring, 2000). The EVI data product also does not become saturated as easily as NDVI when viewing rainforests and other areas of the earth with large amounts of chlorophyll (Weier and Herring, 2000). The LAI data for 2006 were used. They define the number of equivalent layers of leaves relative to a unit of the ground and are computed daily at 1-km resulution from MODIS spectral reflectances for all vegetated land surface globally (USGS, 2014). The MODIS data, obtained at a spatial resolution of 1 km, not only provides land-cover (characterising five global land cover classification systems) and land-cover type assessment, but also includes a quality control mechanism (USGS, 2014). Altitude data were extracted from an interpolated USGS digital elevation model (DEM) available at a spatial resolution of 1 km.Digital maps for three different kinds of water bodies in Nigeria (lakes, rivers and wetlands) were acquired from the

Table 2. Sources and resolution of remote sensing data.

Data	Source	Spatial resolution (km^2)	Temporal resolution
LST_{day}	MODIS	1×1	8 days
LST_{night}	MODIS	1×1	8 days
LAI (vegetation)	MODIS	1×1	8 days
NDVI (vegetation)	MODIS	0.25×0.25	16 days
EVI (vegetation)	MODIS	1×1	16 days
Land cover	MODIS	1×1	1 year
Rainfall estimate	ADDS	8×8	Daily
Evapotranspiration	ADDS	8×8	Daily
Elevation	USGS	1×1	-
Water bodies (rivers, lakes, wetlands)	HealthMapper	1×1	-

LST, land surface temperature; LAI, leaf area index; NDVI, normalized difference vegetation index; EVI, enhanced vegetation index.

HealthMapper database (WHO, 2010). The distance from each location to the nearest water body source was calculated in IDRISI 32, which is an integrated geographic information system (GIS) and RS software developed by Clark Labs (http://www.clarklabs.org/). Rainfall estimates were obtained daily at a spatial resolution of 8 km from the Aviation Digital Data Service (ADDS) (http://www.aviationweather. gov/adds). The MODIS reprojection tool (USGS) was used to convert the RS data to geo-referenced maps and ArcMap, v. 9.1 (http://www.esri.com/) was used as mapping tool. Additional data processing was performed in STATA/SE 9.2 (Stata Corporation, College Station, TX, USA).

Statistical modelling

Logistic regression was fitted to malaria prevalence to identify significant demographic (age) and environmental covariates using STATA v. 9.0. Covariates with a significance level below 0.15 were fitted into Bayesian geostatistical logistic models using WinBUGS v. 1.4 (Imperial College and Medical Research Council, London, UK) in Fortran code. To take into account spatial heterogeneity, location-specific random effects were integrated in the logistic models, assuming that they were distributed according to a multivariate normal distribution with variance-covariance matrix parameterizing the correlation structure in the data as an exponential function of distance.

Let N_i be the number of children tested at location s_i, $i = 1,…,n$, Y_i the number of those found with malaria parasites in a blood sample and $X_i = (X_{i1}, X_{i2},…,X_{ip})T$ the vector of p associated environmental predictors observed at location s_i. We assumed that Y_i arised from a binomial distribution $Y_i \sim Bin(N_i, p_i)$ with parameter p_i measuring the malaria risk at location s_i and modelled the relation between the malaria risk and environmental covariates X_i via the logistic regression $logit(p_i) = X_i T ß$, where $ß = (ß_1, ß_2,…,ß_p)T$ are the regression coefficients. This model assumes independence between the surveys. However, the geographical location introduces correlation since the malaria risk at nearby locations is influenced by similar environmental factors and it is therefore expected that the closer the location, the more similar the way malaria risk varies. To account for the spatial variation in the data we introduced an error term (random effect) Φ_i at each location s_i, $logit(p_i) = X_i T ß + \Phi_i$ and modelled the spatial correlation on the i parameters, i.e. the i's are not independent but derive from a distribution which models the correlation, or equivalently the covariance between every pair of random effects. We adopted the multivariate normal distribution, for the Φ_i's since they represent error terms and are therefore defined on a continuous scale, i.e. $\Phi_i = (\Phi_1, \Phi_2,…,\Phi_n)T \sim N(0, \sum)$. $\sum$ is a matrix with elements $\sum_{ij}$ quantifying the covariance $Cov(\Phi_i, \Phi_j)$ between every pair (Φ_i, Φ_j) at locations s_i and s_j, respectively. The distribution of random effect Φ defines the Gaussian spatial process.

To complete the Bayesian model formulation of the geostatistical models mentioned above, we needed to specify prior distributions for their parameters. For the regression coefficients we adopted a non-informative uniform prior distribution with bounds -∞ and ∞. For the spatial parameters σ^2, σ^2_k, ρ, and ρ_k we adopted inverse gamma and gamma prior distributions, respectively, with parameters chosen to have means equal to 1 and variances equal to 100. We estimated the parameters of the model using Markov chain Monte Carlo simulation with Gibbs sampling (Gelfand and Smith, 1990). Starting with initial values about the parameters, the algorithm iteratively updates the parameters by simulating from their full conditional distributions, i.e. the posterior distribution of each parameter is conditional on the remaining parameters. The full conditional distributions of σ_2 and σ^2_k, $k=1,…, K$ are inverse gamma distributions and simulation from them

Table 3. Bivariate associations between malaria prevalence and environmental indicators arising from non-spatial and spatial logistic models.

Environmental indicators	Non-spatial models		Spatial models	
	PE	95% CI	PE	95% CI
LST$_{day}$				
>300 K	0.59	0.55, 0.63	0.18	0.06, 0.53
>305 K	0.64	0.60, 0.68	0.14	0.03, 0.57
LST$_{night}$	0.95	0.94, 0.96	1.00	0.99, 1.01
Altitude	1.001	0.999, 1.002		
Land use				
Forest	1.45	1.33, 1.58	0.12	0.02, 0.77
Build	0.70	0.68, 0.73	0.81	0.31, 2.00
Crop	10.65	8.84, 12.83	5.44	0.47, 86.53
Distance to water bodies (km)				
1-2	1.04	0.97, 1.11	0.71	0.29, 1.78
2-3	1.25	1.18, 1.33	1.50	0.59, 3.71
3-4	2.11	1.97, 2.27	1.16	0.36, 2.98
>4	1.65	1.58, 1.72	1.37	0.63, 3.15
Vegetation				
NDVI >0.35	1.53	1.48, 1.59	1.23	0.61, 2.64
EVI >0.20	1.21	1.17, 1.26		
LAI	1.013	1.011, 1.014		
Rainfall	1.00	1.00, 1.00	0.78	0.39, 1.28
Spatial correlation parameters				
Φ			2303.0	136.8, 4486.0
σ^2			0.88	0.56, 1.5

PE, posterior estimates; CI, confidence interval; LST, land surface temperature; NDVI, normalized difference vegetation index; EVI, enhanced vegetation index; LAI, leaf area index. Regression coefficients representing odds ratios estimated by the median of the posterior estimates of the corresponding coefficients.

is straightforward. The rest of the parameters do not have full conditional distributions of known forms. We simulated from the non-standard distributions by employing a random walk Metropolis Algorithm (Tierney, 1994) having a normal proposal density with a mean equal to the estimate of the corresponding parameter from previous Gibbs iteration and a variance equal to a fixed number, iteratively adapted to optimize the acceptance rates. We ran five chains with a burn-in of 5000 iterations. Convergence was assessed by inspection of ergodic averages of the selection model parameters. The analysis was implemented using Fortan 95 (Compaq Visual Fortran Professional 6.6.0) and standard numerical libraries supported by the Numerical Algorithms Group (NAG) Ltd (http://www.nag.co.uk/).

The model was used to predict malaria risk throughout Nigeria. This approach treats the malaria risk at a new location as random and calculates its predictive posterior distribution, which not only provides a single risk estimate, but also gives a whole range of likely values together with their probabilities to be the true values at a specific location. This makes it possible to estimate the prediction error, a substantial advantage over the classical Kriging methods. We estimated the predictive posterior distributions at prediction locations by simulation.

Results

The univariate non-spatial analysis indicated that the environmental factors, NDVI, EVI, LAI, distance to water bodies, rainfall, LST for day and night and land-use, were related to malaria prevalence. Malaria prevalence in Nigeria, estimated at 46 locations, had a median of about 45% with values ranging from 0 to over 70%. The results of the bivariate associations between malaria prevalence and environmental indicators arising from non-spatial and spatial logistic models are shown in Table 2.

Estimates of the odds ratios (OR) indicate that in the non-spatial analysis, all the environmental covariates were related to estimated malaria prevalence after adjusting for other risk factors. All covariates significant at the 5% significance level and having low area under the curve (AIC) values were included in the spatial analysis. The relation between malaria risk and rainfall was linear. The logarithmic transformation of NDVI described its relation with malaria risk the best. Second order polynomial terms for minimum temperature, maximum temperature and distance to water bodies gave the best associations with the malaria prevalence. The spatial model suggests that only LST_{day} and land use (forest) were related to malaria. Surprisingly, the models estimated a positive relation with the distance to water bodies, implying that the risk of malaria increased with the distance from permanent water bodies until about 4 km and then decreased. The function 3/rho*100 gives the minimum distance in km with a spatial correlation less than 5%.

Figure 1 shows the malaria prevalence at the various survey locations in Nigeria. There were relatively fewer surveys from northern Nigeria compared to the South. The southern part of the country also generally had more areas with malaria prevalence above 70%. This agrees with the generally held view considering that the southern part has more water bodies, is heavily forested and has more rain than the North.

The malaria risk maps (Figures 2 and 3) estimated from the spatial model shows that malaria prevalence in Nigeria varied from less than 20% in certain areas to over 70% in others. The highest prevalence rates of 70% and above were found in the Niger Delta states of Rivers and Bayelsa, the areas surrounding the confluence of the Rivers Niger

and Benue and in isolated parts of the northeastern and northwestern parts of the country. Isolated patches of low malaria prevalence were scattered around the country with northern Nigeria having more such areas. The middle belt regions had general malaria prevalence of 40% and above. This pattern of malaria distribution is shown in Figures 2 and 3 with the standard deviation (SD) of the median malaria prevalence ranging from 0.1 to 0.32.
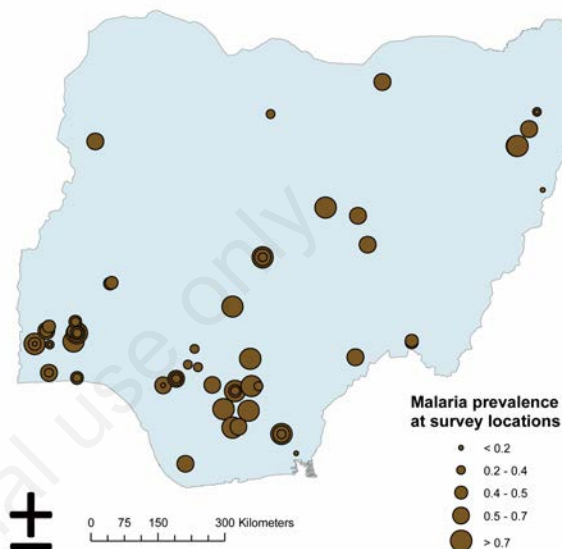

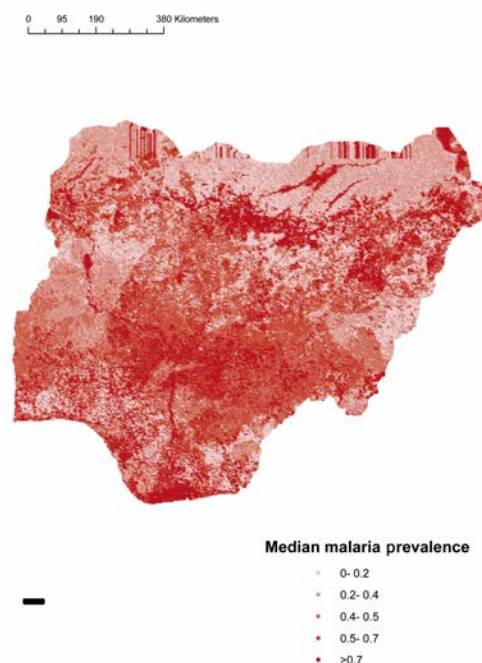**Figure 1. Malaria prevalence at the various survey locations in Nigeria.**

Figure 2. Median of the malaria prevalence across Nigeria.

Discussion

Malaria is an environmental disease and environmental factors are good predictors of transmission, but the relationship between environmental factors, mosquito abundance and malaria prevalence is not linear. This relationship can be established only by means of adequate, spatial statistical models, which can be used for improving predictions of malaria transmission not only in space (for risk mapping) but also in time (for developing early warning systems for malaria epidemics). The Bayesian model that was used has the advantage that it is age-adjusted and makes use of all the survey data available, so I did not have to discard any surveys because of inappropriate age groups. The Bayesian approach also allows flexible model fitting and estimation and mapping of the prediction error. The prevalence map produced (Figures 2 and 3) broadly corresponds to the known distribution of malaria in Nigeria and in particular indicate high transmission in the areas around the main rivers. The result of the survey showed that NDVI, EVI, LAI, distance to water bodies, rainfall, LST_{day}, and land-use were related to malaria prevalence with rainfall having a linear relationship with malaria risk. The models, however, estimated a positive relation with the distance to water bodies up to 4 km before decreasing. This could be due to the fact that even if people like to live near permanent water bodies, they do not usually live on the banks but rather about 500 m away. This implies that water collected in cans, open jars, pits and drainage canals – which serve as vector breeding grounds – will mostly be found at distances at least 500 m away from the banks. Thomas *et al.* (2013) found that 95% of female *Anopheles gambiae* moved within 1.67 km from the nearest breeding site, which means that from about 500 meters from permanent water bodies, malaria prevalence will increase for as far as the flight range of the female anopheles mosquitoes before decreasing. Thomson *et al.* (1997); Kleinschmidt *et al.* (2000); Carter *et al.* (2000); Gosoniu *et al.* (2006); Gemperli *et al.* (2006) also found this relationship. Indeed, Kleinschmidt *et al.* (2000), in their work on the analysis of malaria prevalence in Mali, found that malaria risk was estimated to be reduced to a level lower than that measured close to water only for distances of more than 4 km away from the nearest permanent water body. While vector abundance is supposed to be high closer to the breeding sites, Carter *et al.* (2000) found a negative association between malaria infection and vector abundance. They attributed this to the propensity of people to use bednets (Thomson *et al.*, 1997) or the stimulated development of immunity during early childhood in high-risk areas (Thomas and Lindsay, 2000). Most malaria surveys include people from areas of several km^2, so surveys close to water bodies may include some people from the riverbank and some further away. It is therefore not obvious what relationship to expect with respect to distance to water. The exact relationship between proximity to rivers and malaria appears to be very sensitive to which data points are included and to the details of the model, especially when there are very few data points in the critical areas. It may also be that the lack of adjustment for age in earlier models biased some of the covariate effects.

Empirical maps of malaria risk are important tools in malaria control as they can guide interventions and help assess their effectiveness. Maps are useful in malaria control by identifying areas according to their degrees of risk, thereby guiding resource allocation. They are easy to understand and can therefore be appreciated by people of different educational levels. With accurate maps the effectiveness of malaria control measures can be evaluated after a few years, by comparing the baseline map with subsequent maps. By identifying survey points it was possible to show areas of the country, for which there are

hardly any data and for which future surveys should be done. These maps rely on predictions of risk at locations without observed prevalence data. The prevalence estimates from both the non-spatial and spatial approaches have broad agreement, though the spatial estimates tend to have larger standard errors. This may be due to the fact that in the non spatial analysis there is the assumption of independence which was not upheld. There is still much to be done in the area of malaria control and elimination in Nigeria. The data collection for this work was difficult, primarily due to the fact that malaria prevalence data are few and old. The data that were used to produce the prevalence maps were collected from 1983 to 2007 with about 60% of these surveys done before 2000, which means that these maps may not accurately capture the current malaria prevalence picture of Nigeria.

The maps shown here (Figures 2 and 3) belong to a series of maps that have been produced showing malaria prevalence in Nigeria. Previous malaria maps (Kleinschmidt *et al.*, 2000; Gemperli *et al.*, 2003; Gosoniu *et al.*, 2006) concern all of West Africa and that of Gething *et al.* (2011) for the whole world. A comparison of the estimated malaria prevalence shown in the maps in this paper with those produced by Kleinschmidt *et al.* (2001a) for West Africa reveals similar patterns, but the predicted prevalence in the present maps show more regions with prevalence above 70% and below 30%. There is, however, agreement between the maps, with both showing high risk for malaria in the region of central Nigeria. The maps shown here, however, are in disagreement with the map produced by Gemperli *et al.* (2006b). Their map shows only two regions of Nigeria with prevalence of malaria above 70%. These are in the far north in the area of Kano State and in the southwestern tip of the country in the area of Lago state. These differences may be due to the use of different dataset and also the fact
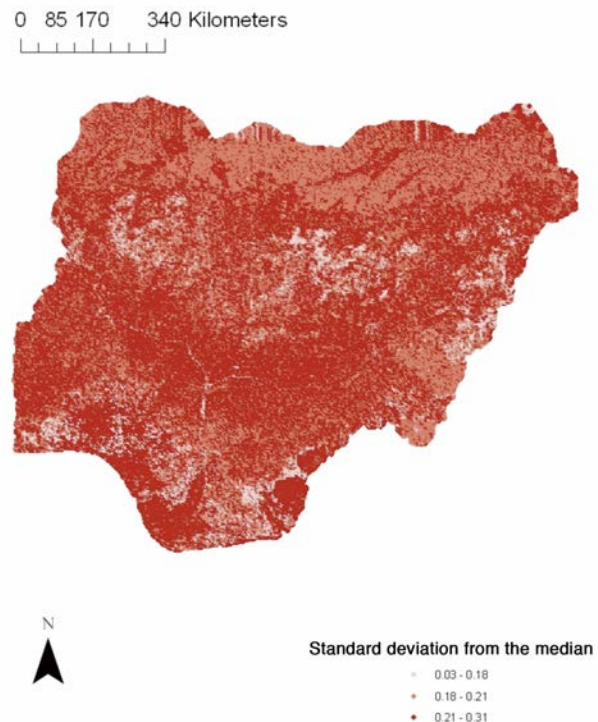
0 85 170   340 Kilometers

N

Standard deviation from the median

. 0.03 - 0.18
. 0.18 - 0.21
. 0.21 - 0.31

Figure 3. The standard deviation from the median of malaria prevalence in Nigeria.

that their map was for malaria prevalence in children ten years old and below. The Gemperli map also has 80% of the Nigerian population living in places with malaria prevalence of 40% and below. The map produced here, on the other hand, shows that 50% of the Nigerian population live in places with malaria prevalence of 50% and above. The Gethin map, however, generally agrees with the map shown here as it shows that virtually every region of Nigeria with a *P. falciparum* parasite rate of at least 50% with the southern tip of the country having a rate of 70% and above as shown here.

Conclusions

A national risk map for Nigeria will allow planners to identify malaria high-endemicity areas, where long term use of insecticide treated nets (ITNs) would be useful. Two distinct areas that can be identified with the aid of these maps are epidemic-prone areas and malaria-endemic areas. These areas require distinct intervention measures. For the epidemic areas, surveillance, indoor spraying of insecticides would be helpful in controlling malaria. In the endemic areas, widespread use of ITNs, behavioural changes, such as avoiding storing water in open cans outdoors and clearing of bushes around dwelling places would help. Planners can also assess the possible health impacts of measures aimed at improving food security through the promotion of large-scale irrigation and wetland management projects. Road construction companies should be requested to fill up construction ponds (burror pits) once they are no longer needed. Finally, the maps constructed will also guide public health researchers in identifying appropriate study environments for intervention trials as well as assist with identification of populations potentially benefiting from new interventions.

Though the data used for this study came from an imperfectly sampled population of Nigeria, this study nevertheless is the first known attempt to produce a malaria risk map for Nigeria based entirely on malariometric data. It is anticipated that it will provide useful additional guidance to control programme managers, and that it can be refined once sufficient additional data become available.

References

Berquist NR, 2001. Vector-borne diseases: new trends in data collection and risk assessment. Acta Tropica 79:13-20.

Carter R, Mendis KN, Roberts D, 2000. Spatial targeting of interventions against malaria. B World Health Organ 78:1401-11.

Craig HM, Snow RW, Le Sueur D, 1999. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitol Today 15:105-11.

Diggle PJ, Moyeed RA, Rowlinson B, Thomson M, 2002. Childhood malaria in the Gambia: a case study in model-based geostatistics. J R Stat Soc 51:493-506.

Gelfand AE, Smith AFM, 1990. Sampling-based approach to calculating marginal densities. J Am Stat Assoc 85:398-409.

Gemperli A, 2003. Development of spatial statistical methods for modeling point-referenced spatial data in malaria epidemiology. PhD Diss. University of Basel, Basel, Switzerland.

Gemperli A, Sogoba N, Fonjo E, Mabaso M, Bagayoko M, Briet O, Anderegg D, Liebe J, Smith T, Vounatsou P, 2006a. Mapping malaria transission in West and Central Africa. Trop Med Int Health 11:1032-46.

Gemperli A, Vounatsou P, 2004. Fitting generalized linear mixed models for point-referenced data. J Mod Appl Stat Methods 2:497-511.

Gemperli A, Vounatsou P, Sogoba N, Smith T, 2006b. Malaria mapping using transmission models: An application to survey data in Mali. Am J Epidemiol 163:289-97.

Gething PW, Patil AP, Smoth DL, Guerra CA, Elyazar IRF, Johnston GL, Tatem AJ, Hay SI, 2011. A new world malaria map: *Plasmodium Falciparum* endemicity in 2010. Malaria J 10:378.

Gosoniu L, Vounatsou P, Sogoba N, Smith T, 2006. Bayesian modeling of geostatistical malaria risk data. Geosp Health 1:127-39.

Guinovart C, Navia MM, Tanner M, Alonso PL, 2006. Malaria: burden of disease. Curr Mol Med 6:137-40.

Hay SI, Rogers DJ. Toomer JF, Snow RW, 2000. Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, internet access and review. T Roy Soc Trop Med H 94:113-27.

Kitron U, 2000. Risk maps: transmission and byrden of vector-borne diseases. Parasitol Today 16:324-5.

Kleinschmidt I, Bagayoko M, Clarke GPY, Craig M, LeSueur DA, 2000. A spatial statistical approach to malaria mapping. Int J Epidemiol 29:355-61.

Kleinschmidt I, Omumbo J, Bret O, Van de Giesen N, Sogoba N, Mensah NK, Windmeijar P, Moussa M, Teuscher T, 2001a. An emperical malaria distribution map for west Africa. Trop Med Int Health 6:779-86.

Kleinschmidt I, Sharp BL, Clarke GPY, Curtis B, Fraser C, 2001b. Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in KwaZulu, Natal, South Africa. Am J Epidemiol 153:1213-21.

Okonko IO, Soleye FA, Amusan TA, Ogun AA, Udeze AO, Nkang AO, Ejembi J, Faleye TOC, 2009. Prevalence of malaria *Plasmodium* in Abeokuta, Nigeria. Malays J Microbiol 5:113-8.

Rogers DJ, Randolf SE, Snow RW, Hay SI, 2002. Satellite imagery in the study and forecast of malaria. Nature 415:710-5.

Ribeiro JMC, Seulu F, Abose T, Kidane G, Teklehaimanot A, 1996. Temporal and spatial distribution of anopheline mosquito in an Ethiopian village: implications for malaria control strategies. B World Health Organ 74:299-305.

Snow EW, Omumbo JA, Lowe B, Molineaux CS, Obiero JO, Palmer A, Weber MW, Pinder M, Nahlen B, Obonyo C, Newbold C, Gupta S, Marsh K, 1997. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. Lancet 349:1650-4.

Thomas CJ, Cross DE, Bøgh C, 2013. Landscape movements of *Anopheles gambiae* Malaria vector mosquitoes in rural Gambia. PLoS One 87:e68679.

Thomas CJ, Lindsay SW, 2000. Mapping and estimating the population at risk fromlymphatic filariasis in Africa. T Roy Soc Trop Med H 94:37-45.

Thomson MC, Connor SJ, 2000. Environmental information systems for the control of arthropod vectors of disease. Med Vet Entomol 14:227-44.

Thomson MC, Connor SJ, Milligan P, Flasse SP, 1997. Mapping malaria risk in Africa: what can satellite data contribute. Parasitol Today 13:313.

Tierney L, 1994. Markov chains for exploring posterior distributions. Ann Stat 22:1701-28.

USGS, 2014. Land processess distributed archive center (LP DAAC) modis products. Available from: https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd15a3

Vaughan AM, Aly ASI, Kappe SHI, 2008. Malaria parasite pre-erythrocytic stage infection: gliding and hiding. Cell Host Microbe 4:209-18.

Weier J, Herring D, 2000. Measuring vegetation (NDVI & EVI) Nasa earth observatory. Available from: http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php

Wijeyaratne P, 1999. Malaria prevention: lessons learned. Environmental Health Project Publ., Washington, DC, USA.

WHO, 2010. The HealthMapper database. Available from: https://health-map.wordpress.com/2010/04/01/world-health-organizations-health-mapper/

WHO, 2013. Expert Committee Report on Malaria. World Health Organization, Geneva, Switzerland.

WHO, 2014. Expert Committee Report on Malaria. World Health Organization, Geneva, Switzerland.

World Resources Institute, 1995. African data sampler. World Resources Institute, Washington, DC, USA.