Characterizing local-scale heterogeneity of malaria risk: a case study in Bunkpurugu-Yunyoo district in northern Ghana

Punam Amratia[1,2]* , Paul Psychas[2], Benjamin Abuaku[3], Collins Ahorlu[3], Justin Millar[1,2], Samuel Oppong[4], Kwadwo Koram[3] and Denis Valle[1,2]

Abstract

Background: Bayesian methods have been used to generate country-level and global maps of malaria prevalence. With increasing availability of detailed malaria surveillance data, these methodologies can also be used to identify fine-scale heterogeneity of malaria parasitaemia for operational prevention and control of malaria.

Methods: In this article, a Bayesian geostatistical model was applied to six malaria parasitaemia surveys conducted during rainy and dry seasons between November 2010 and 2013 to characterize the micro-scale spatial heterogeneity of malaria risk in northern Ghana.

Results: The geostatistical model showed substantial spatial heterogeneity, with malaria parasite prevalence varying between 19 and 90%, and revealing a northeast to southwest gradient of predicted risk. The spatial distribution of prevalence was heavily influenced by two modest urban centres, with a substantially lower prevalence in urban centres compared to rural areas. Although strong seasonal variations were observed, spatial malaria prevalence patterns did not change substantially from year to year. Furthermore, independent surveillance data suggested that the model had a relatively good predictive performance when extrapolated to a neighbouring district.

Conclusions: This high variability in malaria prevalence is striking, given that this small area (approximately 30 km × 40 km) was purportedly homogeneous based on country-level spatial analysis, suggesting that fine-scale parasitaemia data might be critical to guide district-level programmatic efforts to prevent and control malaria. Extrapolations results suggest that fine-scale parasitaemia data can be useful for spatial predictions in neighbouring unsampled districts and does not have to be collected every year to aid district-level operations, helping to alleviate concerns regarding the cost of fine-scale data collection.

Keywords: Malaria, Bayesian, Fine-scale, Geostatistical, Ghana

*Correspondence: punam.amratia@gmail.com
[1] School of Forest Resources and Conservation, University of Florida, Gainesville, USA
Full list of author information is available at the end of the article

Amratia *et al. Malar J*   (2019) 18:81

Page 2 of 14

Background

Over the past two decades, Ghana has made significant progress towards reducing malaria mortality [1]. This progress can be attributed to increasing coverage and improving access to rapid diagnostic tests and artemisinin-based combination therapy, implementing universal access to insecticide-treated bed nets, scaling-up indoor residual spraying (IRS) [2, 3] as well as climate change, urbanization patterns and infrastructural development [4, 5]. Despite these country-wide efforts for malaria control and prevention [6], and improved infrastructure, malaria morbidity remains relatively high [7]. The national early childhood (6 to 59 months old) malaria prevalence rate in Ghana has remained relatively stable (22–27%) from 2011 to 2016 [8]. However, it is important to consider that nationwide prevalence estimates mask significant spatial variability, as many parts of Ghana still experience intense seasonal malaria transmission, particularly in the northern regions [9, 10] where prevalence is greater than 40%. The 2016 Ghana Malaria Indicator Survey (MIS) revealed regional childhood prevalence estimates ranging between 5 and 31%, but these aggregated estimates may understate substantial heterogeneity within and between districts in each region [5].

Characterizing the spatial variation of disease prevalence by mapping exercises has proven to be useful for the strategic planning of malaria prevention and control activities at the national level [11–13]. In 2013, the Ghana National Malaria Control Programme (NMCP) collaborated with a team at Kenya Medical Research Institute-Wellcome Trust to develop a state-of-the-art map of malaria parasitaemia rates at the district level [5]. Their model estimates of *Plasmodium falciparum* in children aged 2 to 10 years in Northern Region ranged from 42% for the Tamale municipality to over 75% for most rural districts, presumably related to the residual malaria transmission in this area [14]. Linking data on climate and urbanization with locally available data yielded a more nuanced view of malaria distribution in the northern savannah compared to national averages, which allowed for the classification of districts into those that may be suitable/unsuitable for seasonal malaria chemoprophylaxis (SMC). Such work has paved the way for finer scale mapping efforts to describe the within-district local-scale heterogeneity in malaria parasite prevalence [15, 16], which is critical to increasing the effectiveness of current malaria control interventions implemented at the local level [17]. However, the ability to map at finer scale is contingent on the availability of rich surveillance data [18].

The purpose of this study was to characterize local spatial patterns in early childhood malaria prevalence in a single district in northern Ghana. A Bayesian hierarchical geostatistical model it was applied to a series of seasonal cross-sectional parasitaemia surveys conducted in the Bunkpurugu-Yunyoo district from 2010 to 2013, which were previously collected as part of an IRS intervention evaluation [3]. The high geographic resolution in these surveys provided a rich dataset for understanding the local scale malaria epidemiology in a residual transmission region [19]. Detailed surveys such as these can be prohibitively costly and would be untenable to scale to the entire Northern Region of Ghana, therefore exploratory work was conducted on how well the model could be extrapolated to unsampled neighbouring districts and how frequently these surveys would have to be implemented to achieve the same degree of resolution.

Methods

The outcome used throughout this paper was microscopy-based malaria status of individuals sampled by parasitaemia household surveys that were conducted between October 2010 and March 2013 in the Bunkpurugu-Yunyoo district (BYD), Northern Region, Ghana. This dataset arises from a collaboration between the University of Ghana and the President's Malaria Initiative (PMI) as part of an IRS evaluation project [20].

Study area

Bunkpurugu-Yunyoo is located in Northern Region, Ghana (10.6° N 0.0° W). The district exhibits a gradual slope from the rocky Gambaga escarpment in the north/northwest to riverine plains in the south/southeast, dropping from 518 to 128 m above sea level within a relatively small area (30 km by 40 km). The district lies in the Guinea Savannah zone, characterized by a unimodal rainy season from May to October, which peaks in August–September while the remainder of the year is typically dry. The mean annual rainfall is between 100 mm and 115 mm, with annual temperature ranging between 30 °C and 40 °C. The predominant malaria vector species are *Anopheles gambiae* sensu stricto and *Anopheles funestus* [21]. In 2010, the estimated total district population was 122,591, with an estimated under-five population of 21,373. The study district is predominantly rural, with the exception of two settlements that exceed a population threshold of 5000, and are thereby designated as 'urban' according to the Ghana Statistical Service, namely Nakpanduri (population 6179) and Bunkpurugu (population 11,106) [22]. Most of the population live in mud-walled compounds and engage in rain-fed, small-scale farming as well as small-scale trading.

Malaria transmission in BYD is characterized by strong seasonal variations that closely follow rainfall patterns, with a peak lasting up to 3–4 months between August and November [6]. The study area is defined as the set of

Amratia *et al. Malar J*   (2019) 18:81

Page 3 of 14

communities that were eligible for spraying by the PMI-funded IRS programme in 2011–2013 (Fig. 1). The study boundaries do not perfectly align with the official administrative district boundary because the sampling frame for the study consisted of the list of communities that fell under the public health jurisdiction of the BYD office of Ghana Health Service (GHS). As a result, this list leaves out communities handled by the East Mamprusi office of GHS along the extreme northeastern border and the southeastern corner of BYD.

The district had benefitted from two recent mass distributions of long-lasting insecticide-treated nets (LLINs). In 2010, the LLINs programme covered children under 5 years and pregnant women and in 2012 they covered all other members of the household, resulting in approximately > 75% coverage across BYD [2, 23]. The other major malaria prevention and control intervention in the region had been IRS, which began in 2011 using alpha-cypermethrin 0.4% WP (ICON®10CS, Syngenta, Basel, Switzerland) and shifted to organophosphates in 2013

due to declining susceptibility of *Anopheles spp.* to the pyrethroid insecticides [21].

Malaria survey data

Surveys were conducted twice annually at periods of expected peak (i.e., the end of the rainy season, October–November) and trough (i.e., the end of the dry season, March–April) levels of malaria parasitaemia. Each survey was collected over a 2-week period.

Children under 60 months of age and their caregivers were randomly selected for each survey, using a multi-stage randomized cluster sampling design. Probability proportional to population was used to randomly sample approximately 72 communities per survey across the study area from a GHS roster of 238 communities. Within each community, survey teams visited 15–17 households with children under 5 years old, selected randomly from an inventory of such households that had been conducted within 6 months prior to each survey. Teams were instructed to invite any child under
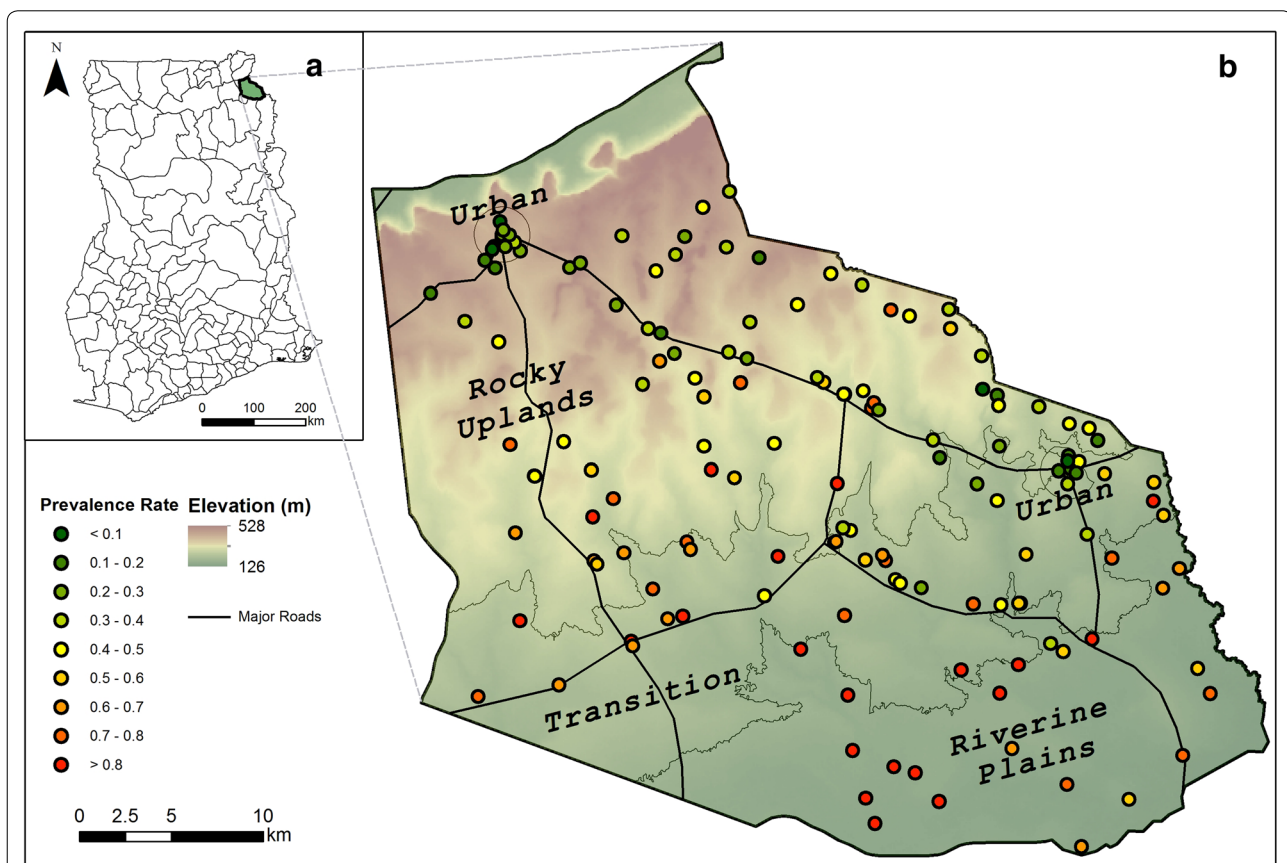

Fig. 1 Communities sampled (black dots) in Bunkpurugu-Yunyoo district across all 2011–2013 surveys. The observed prevalence rates are displayed from low (green) to high (red) on the elevation surface and major roads. Four ecological zones defined by study design are highlighted by the elevation as rocky uplands, transition vegetation, riverine plains and urban zones. Inset map shows all the districts within Ghana, highlighting the Bunkpurugu-Yunyoo district

Amratia *et al. Malar J*   (2019) 18:81

Page 4 of 14

5 years old who was available on the day of the survey and had slept in the household the night before for participation. The resulting sample (average 24 children per community) covered approximately 20% of the under-five population in each survey, based on the 2010 census data and community lists developed by the GHS [3]. Thin and thick blood films were prepared in the field and later read by two qualified microscopists at the Noguchi Memorial Institute for Medical Research laboratory in Accra to determine microscopic parasitaemia. In cases of discordant readings, a senior microscopist determined the final result. Prior to analysis, children aged under 6 months were removed from the dataset, based on previous studies suggesting that young infants have enhanced malaria protection due to maternal antibodies [24, 25].

Using a hand-held Etrex© GPS device (Garmin), field technicians obtained GPS coordinates for a readily identifiable central point in each community, such as a church, school, or chief's palace. GPS coordinates were not obtained at the household level because of feasibility and ethical considerations. The geo-coding of communities permitted the dataset to be enhanced with remotely sensed and geographic variables.

Variables employed to model malaria parasite prevalence

Remote sensed variables Malaria transmission has been shown to be strongly related to satellite-derived environmental and socio-demographic factors [26]. Weiss et al. [27] carried out a comprehensive assessment of spatially gridded covariates that are likely to be associated with the malaria transmission cycle. Based on their results a suite of satellite and geographically derived covariates were assembled for modelling purposes (Table 1). The selection of environmental covariates was partly based on availability of raster data that closely matched the survey times. Elevation is widely used in malaria mapping due to its association with precipitation and temperature [27] and it was extracted from the 90-m resolution Shuttle Radar Topography Mission Digital Elevation Model (SRTM-DEM). The normalized difference vegetation index (NDVI), a proxy for vegetation cover, was obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) products using the 16-day composite. Based on this product, the maximum NDVI was calculated within a 32-day period prior to the start date of each survey. Vegetation cover is a useful proxy for characterizing vector habitat for *Anopheles* spp. commonly found across Africa [28]. The land surface temperature (LST) for day and night was also obtained from MODIS (MOD11)

Table 1 List of spatial predictors, including their sources and spatial resolution (maps of spatial covariates can be found in Additional file 1)

Variable	Definition (units)	Spatial resolution	Source
Remote-sensed			
Elevation	Height above sea level (m)	90 m	CGIAR SRTM [49]
Normalized difference vegetation index (NDVI)	Index of vegetation conditions. Ranges from −1 (no vegetation) to 1 (complete vegetated)	250 m	NASA (Terra) MOD13A3 and (Aqua) MYD13A3 datasets [50]
Land surface temperature (LST)—day and night time	Kelvin (converted to °C)	1 km	NASA (Terra) MOD11A2 and (Aqua) MYD11A2 datasets [50]
Rainfall (seasonal 3 months cum.)	Actual cumulative 3-months rainfall (mm)	5 km	CHIRPS [51]
Long-term precipitation 1970–2000 (seasonal 3 months cum.)	Long-term cumulative 3 months rainfall based on average monthly rainfall (mm) data from 1970–2000	1 km	WorldClim [52]
Night-time lights	Average visible band digital number values ranging 1–63	1 km	NOAA [33]
Population density	Number of people per 100 sq m	100 m	WorldPop [36]
GIS-derived			
Distance to urban centre	Distance from urban centre (km)	1 km	Based on field data
Distance to health facility	Distance from active health facilities during study (km)	1 km	Based on field data
Distance to roads	Distance from established road-network (km)	1 km	CIESIN [46]
Distance to water bodies	Distance from permanent water bodies (km)	1 km	ESRI [45]
Accessibility to cities	Travel time to cities to assess inequalities in accessibility (travel time)	1 km	MAP [44]
Slope	Percentage rise in elevation	90 m	Derived from elevation product

Amratia *et al. Malar J*   (2019) 18:81

Page 5 of 14

products using the 8-day composites, which were then used to calculate a monthly average LST 32 days prior to survey start date. Temperature has been widely accepted to be an important component in malaria transmission, largely based on its influence on mosquito survival, development, breeding, and biting rates [8, 29, 30], whereas precipitation is a proxy for available stagnant water puddles that are ideal habitat for mosquito larvae. Previous studies have shown that a 9 to 12-week time lag exists between rainfall and onset of malaria transmission [31]. Therefore, the variable of cumulative rainfall for a 3-month period 30 days prior to survey start date was created, based on the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data while long-term precipitation patterns at finer resolution than CHIRPS (1 km) were extracted from WorldClim datasets [32]. Long-term precipitation was included because of its higher spatial resolution compared to cumulative rainfall. Night-time lights for 2011 were obtained from National Oceanic and Atmospheric Administration (NOAA) satellite products at 1-km resolution [33], being a useful proxy for poverty and infrastructure (i.e., electrified housing are more likely to be walled and be less suitable habitat for vector development) [34]. Because densely populated areas are often poor habitat for *Anopheles* to breed [35], Population density at 100-m resolution was extracted from WorldPop [36] and validated this product using the census data for the two main urban towns of Bunkpurugu and Nakpanduri. Finally, land use products by MODIS were expected to be an important factor [27] for vector habitat, however it was not included because there was little variation across the study district [37].

GIS derived variables
In relation to the geographically derived variables, Euclidean distances from roads, urban centres, permanent water bodies, and active health facilities were calculated using ArcGIS 10.3 [38]. A 5-km buffer was created around the study area to control for edge effects that might affect distance calculations. The inclusion of these distance variables was based on their biological importance to malaria transmission: urban centres are known to be associated with lower malaria prevalence due to their infrastructural development and access to resources such as medication [39]. Access to health facilities has been shown to be a key determinant in child mortality and primary usage of health clinics declines with increased travel time and distance [40]. Roads are a proxy for accessibility [41, 42] and water bodies are directly linked to mosquito breeding sites [42, 43]. Accessibility to cities, developed by the Malaria Atlas Project (MAP), was also included as a metric to account for distance to cities, transport infrastructure and distribution of resources [44]. A geo-coded

shape file with urban centres was created based on field work records. Urban centres were defined as villages with more than 5000 people in the 2010 census, resulting in two urban centres in the study region: Nakpanduri and Bunkpurugu. Field workers used GPS to record locations of active health facilities in each survey period. Shape files for water bodies and major roads were obtained from ESRI online [45] and the Center for International Earth Science [46], respectively. Finally, slope was derived from the SRTM elevation raster. A list of covariates and their details can be found in Table 1.

Adjusting the model for child age
The only individual-level variable used for modelling was child's age. Age is known to be strongly associated with individual-level parasitaemia and therefore was accounted for in the model to generate age-adjusted malaria prevalence predictions [47]. Diggle et al. explored a similar approach of including non-spatial covariates for malaria mapping such as age and bed-net use to explain the non-spatial variation in model-based geostatistical methods [48]. A detailed risk analysis involving other individual-level variables for this same dataset is provided elsewhere [4].

Selecting a suitable set of variables for prediction
The use of too many covariates can lead to over-fitting of the model, as well as multicollinearity [53, 54]. To address this issue, a two-step procedure, commonly used in mapping exercises [9, 55], was used to select the best predictors from Table 1. First, correlation between covariates was calculated, and in cases of high correlation (i.e., $R > 0.7$), a single representative covariate was selected. The choice of which covariate to retain was based on the strength of association with malaria prevalence, spatial resolution, and its relevance to malaria epidemiology [27]. Second, a bi-directional, step-wise, regression analysis was run on a full multivariate model using the remaining covariates. The covariates associated with the model with the lowest Akaike information criterion (AIC) were used for subsequent analysis.

Bayesian geostatistical model
A Bayesian hierarchical geostatistical model was fitted independently to data from each survey. Let the malaria status M_{ijt} for child i in village location j in survey t be a binary variable, equal to 1 for a positive microscopy result and 0 otherwise. Using a probit regression framework, assume that $M_{ijt} = 1$ if the latent variable z_{ijt} is greater than 0 and $M_{ijt} = 0$ otherwise. The latent variable z_{ijt} is modelled as:

$$z_{ijt} \sim N\left(\alpha_{jt} + \boldsymbol{x_{ijt}^T}\boldsymbol{\beta}_t, 1\right)$$

Amratia *et al. Malar J*   (2019) 18:81

Page 6 of 14

where $\boldsymbol{x}_{ijt}^T$ is the design vector containing the age of each child, $\boldsymbol{\beta}_t$ contains the corresponding regression coefficients, and α_{jt} is the village-survey level random-effect. Assume that:

$$\boldsymbol{\alpha}_t \sim N\left(\boldsymbol{W}_t \boldsymbol{\gamma}_t, \sigma_t^2 \boldsymbol{\Sigma}(\rho_t)\right)$$

where $\boldsymbol{W}_t$ is the design matrix and $\boldsymbol{\gamma}_t$ is a vector of the corresponding coefficients. Furthermore, $\boldsymbol{\Sigma}$ is the spatial correlation matrix, where the correlation between villages k and l is given by the exponential parametric function $exp\left(-\frac{d_{kl}}{\rho_t}\right)$. In this expression, d_{kl} is the Euclidean distance between villages k and l and ρ_t is the correlation decay parameter.

Finally, the priors are given by:

$$\boldsymbol{\beta}_t \sim N(0, \boldsymbol{I})$$

$$\rho_t, \sigma_t \sim Unif(0, 100)$$

$$\boldsymbol{\gamma}_t \sim \boldsymbol{N}(0, \mathrm{T})$$

where T is a diagonal matrix with diagonal elements equal to [10,1,...,1]. A customized Gibbs Sampler programmed in R [56] was used to fit this model separately for each survey.

Model validation

The geospatial model performance was assessed based on its out-of-sample predictive ability using a tenfold cross-validation approach for each survey. Villages in the data were randomly divided into 10 sub-sets. The model was trained on 9 of the 10 sub-sets, and then used the estimated parameters to predict the expected prevalence for the remaining withheld sub-set. This was repeated using each testing sub-set. The out-of-sample predictive skill of the geospatial Bayesian model was compared to that of a standard generalized linear regression model (GLM), which is a commonly used statistical framework for modelling this type of response variable. A temporal validation was also done for the Bayesian model, where the model was trained on one survey and used to predict prevalence for future surveys. These temporal predictions were stratified by season (i.e., rainy season surveys were used to predict future rainy season surveys but not

dry season surveys). Two statistical metrics were used for all validations: the log-likelihood and the mean absolute error (MAE) in relation to malaria prevalence in each village.

Extrapolation to neighbouring district

Once the geospatial model was fitted, the results were extrapolated to the neighbouring district: East Mamprusi. This district was chosen because it has similar topography and environment to Bunkpurugu–Yunyoo. Furthermore, a list of health facilities and urban centres for East Mamprusi was obtained from the GHS [57] and corroborated using expert information and the 2010 population census, respectively [22], enabling the determination of the distance to health facilities and urban centres. For an independent evaluation of the accuracy of these spatial extrapolation, the MAE was calculated between the predicted malaria prevalence rates in the 2011 rainy season and the 2011 Multiple Indicator Cluster Survey (MICS) prevalence estimates for locations that fell within this district [58].

Results
Malaria prevalence

The study included 10,518 children aged 6–59 months of age with a complete microscopy diagnostic test result for malaria parasitaemia collected in 438 communities across 3 rainy seasons and 3 dry seasons. The average age of participants was 31 months. The average community level prevalence was 44.46% (95% CI 43.83–45.74) during the entire study but prevalence varied considerably by survey, with a strong seasonal effect (Table 2).

Variable selection

As mentioned in Methods, an initial set of 14 spatial covariates were chosen based on the malaria mapping literature [27] (Table 1). NDVI, rainfall, long-term precipitation, and daytime LST were highly correlated (R > 0.7). Because NDVI had the strongest association with microscopy and this covariate had the finest spatial resolution, NDVI was retained and the other three highly correlated covariates were excluded from the step-wise selection. Because accessibility was highly correlated with distance to health facility (R = 0.71),

Table 2 Average prevalence across each survey

Survey year	2010	2011		2012		2013
Season	Rainy	Dry	Rainy	Dry	Rainy	Dry
Number of children	1547	1762	1794	1804	1803	1808
Mean prevalence (%)	60	39	56	36	51	28
95% CI	57–62	36–41	54–58	34–38	49–54	25–29

given that distance to health facility was based on ground-truth locations of active health facilities at the time of each survey, it was retained instead of accessibility. A step-wise, model-selection approach using AIC revealed that the most important covariates were distance to urban centre, distance to health facility, elevation, NDVI, distance to road, distance to water, LST at night, and night-time lights. These parameters were used for geostatistical predictions (parameter estimates can be found in Additional file 2).

Geospatial model results

The Bayesian geostatistical model revealed a strong rural–urban gradient with prevalence generally increasing northeast to southwest (top and middle panels in Fig. 2) ranging from 19 to 90%. These results were particularly surprising given that the malaria prevalence
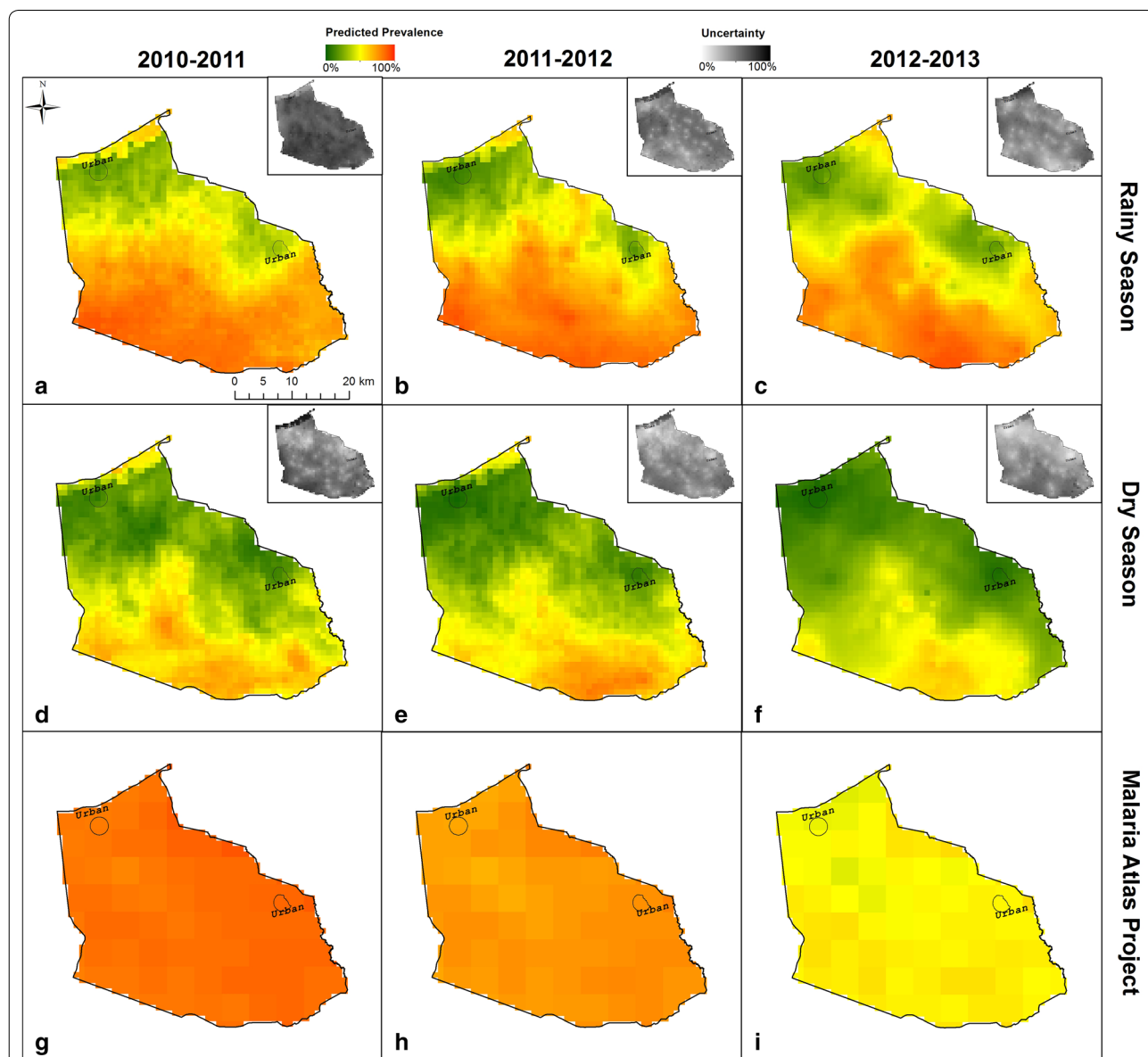
Fig. 2 Strong seasonal and fine-scale spatial variability in age-adjusted predicted malaria prevalence. Top and middle panels (**a–f**) display the age-adjusted predicted mean prevalence for the rainy and dry seasons, respectively. Age-adjusted malaria prevalence is based on children with average age (i.e., 31 months). The bottom panels (**g–i**) display the MAP parasite prevalence surface for *P. falciparum* between ages 2 and 10 years [11]. Insets show the uncertainty in prevalence predictions (top right in each panel), given by the width of the 95% CI. High uncertainty is represented by darker grey

maps for Ghana developed by the MAP suggest a much smaller range in prevalence (68–83%), without a clear geographical trend (bottom panels in Fig. 2). Malaria remained low in and around the urban centres throughout the study but this pattern was more distinct during rainy seasons versus dry seasons. The only notable temporal trend was a small reduction in prevalence in the southern portion of BYD in the final dry season.

The regression results indicate that increasing distance to urban centre was significantly associated with high malaria prevalence estimates during rainy seasons. Furthermore, higher elevation was linked to reduced parasite prevalence of malaria during the rainy seasons in 2010 and 2011, and the dry season in 2012. Finally, living further away from a health facility tended to increase the risk of high malaria prevalence, but this effect was only significant for dry season 2013. All other covariates showed no statistically significant association. Posterior estimates of the parameters used for predictive modelling are given in Additional file 2.

Model validation
The model validation results using the average log-likelihood across all surveys (Fig. 3a) revealed that the Bayesian model generally had better out-of-sample predictive performance when compared to the GLM model, particularly during the dry seasons. Similar results also arose in relation to the mean absolute errors (Fig. 3b).

The temporal validation results reveal an average MAE of approximately 12.7%, without much difference in error for temporal predictions of 1 versus 2 years (Table 3). These MAE values are similar to those obtained when performing the tenfold cross-validation, suggesting that temporal predictions within a 1- or 2-year time interval are as accurate as spatial predictions.

Extrapolation to neighbouring district
Predicted malaria prevalence in East Mamprusi was substantially higher during the rainy season when compared to the dry season (Fig. 4). As expected, there is a distinct trend of lower transmission close to urban areas and the 2012–2013 dry seasons showed a significantly

Table 3 Temporal validation for surveys with temporal prediction lag of one and two years

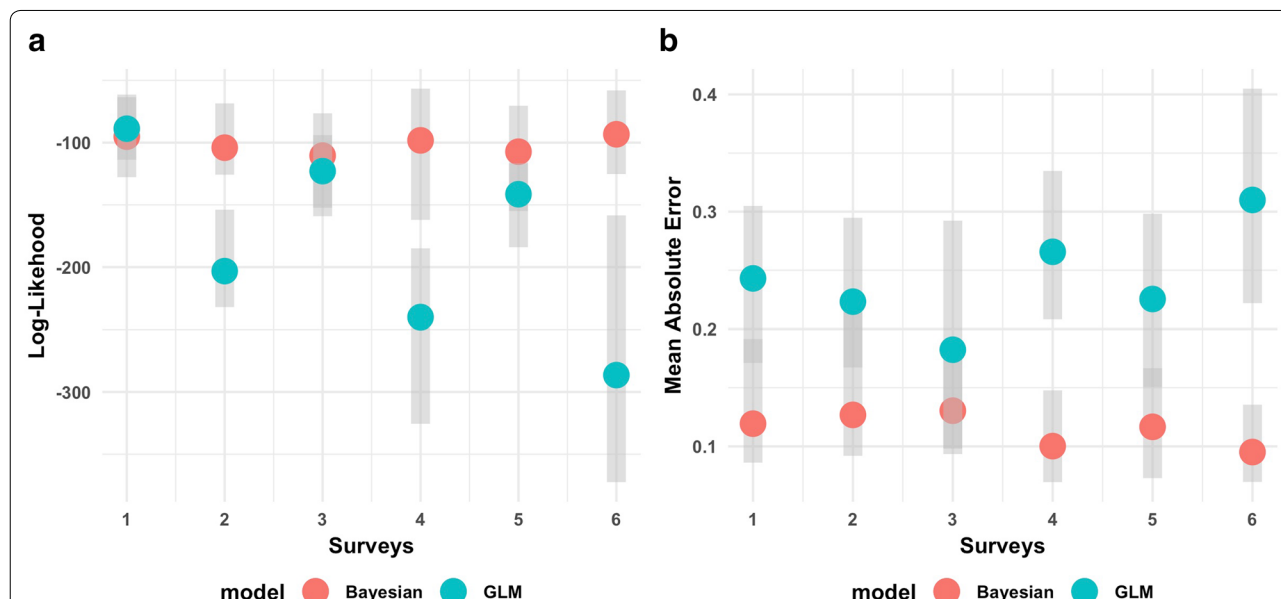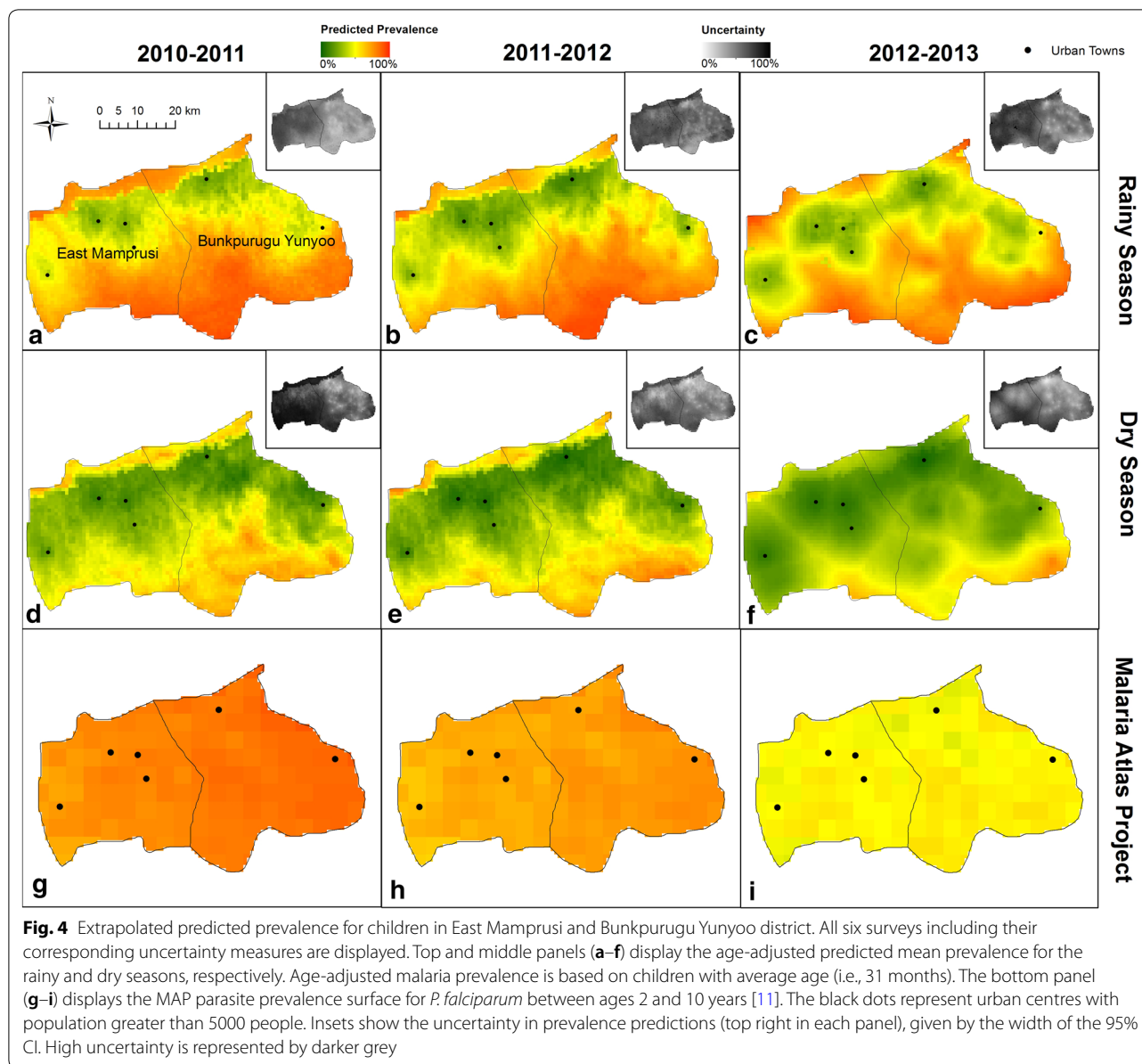| Season	Year of estimation	Year of prediction	Mean absolute error
Rainy	2010	2011	0.124
Rainy	2011	2012	0.126
Rainy	2010	2012	0.132
Dry	2011	2012	0.126
Dry	2012	2013	0.124
Dry	2011	2013	0.126

Fig. 3 Model validation comparisons between Bayesian geospatial model and generalized linear model (GLM). Predictions based on the Bayesian geospatial model generally outperform the GLM. Two out-of-sample metrics were calculated; the log-likelihood for each survey (**a**) and MAE (**b**). Grey vertical polygons depict the range of outcomes based on the 10 cross-validation folds. Higher values for the log-likelihood and lower values for the MAE indicate better out-of-sample predictive skill

Amratia *et al. Malar J*   (2019) 18:81

Page 9 of 14

Fig. 4 Extrapolated predicted prevalence for children in East Mamprusi and Bunkpurugu Yunyoo district. All six surveys including their corresponding uncertainty measures are displayed. Top and middle panels (**a**–**f**) display the age-adjusted predicted mean prevalence for the rainy and dry seasons, respectively. Age-adjusted malaria prevalence is based on children with average age (i.e., 31 months). The bottom panel (**g**–**i**) displays the MAP parasite prevalence surface for *P. falciparum* between ages 2 and 10 years [11]. The black dots represent urban centres with population greater than 5000 people. Insets show the uncertainty in prevalence predictions (top right in each panel), given by the width of the 95% CI. High uncertainty is represented by darker grey

lower predicted prevalence. Finally, as expected, there is distinctly higher uncertainty across East Mamprusi compared to Bunkpurugu-Yunyoo for all surveys. The results for survey three (rainy season 2011) and MICS 2011 survey data were used to determine the reliability of these spatial extrapolations. MICS 2011 contained 10 clusters that fell in Bunkpurugu-Yunyoo and East Mamprusi district. Of these, six were in Bunkpurugu-Yunyoo and four in East Mamprusi. No location in survey data matched the MICS locations; so unfortunately, direct ground-truthing was not possible. The MAE was equal to 19.5 and 11.2% for Bunkpurugu-Yunyoo and East Mamprusi, respectively. The overall MAE for all 10 locations in both districts was 16%. While this may seem large, it is important to note that this margin of error is not substantially larger than those obtained with the tenfold cross-validation and temporal prediction exercises in Bunkpurugu-Yunyoo. Furthermore, this margin of error is enough to adequately distinguish the main spatial prevalence trends in the region, something that current national prevalence maps are not able to detect. For comparison purposes, the MEA for these 10 locations with the MAP 2011 estimates was 37.6% (graphical representation can be found in Additional file 2: Figure S1).

Discussion

This study presents a Bayesian geostatistical analysis of the active surveillance data collected over a period covering high (rainy season) and low (dry season) transmission seasons in a region that has previously been characterized as uniformly high transmission. The application of high-resolution remote sensing data and geostatistics to develop malaria prevalence maps at a fine spatial resolution has revealed striking heterogeneity of malaria prevalence over a small geographic area, which were previously not expressed in country level and global mapping efforts [5, 11, 13]. For example, the malaria prevalence map for Ghana developed by the MAP for all survey years show predicted prevalence in BYD ranging from 68 to 83%, failing to encompass the range predicted by the model (19 to 90%) and to uncover the strong northeast to southwest trend that was detected. These differences can be explained in multiple ways. For instance, the prevalence surface from MAP is based on children 2–10 years old, focuses solely on *P. falciparum*, and relies on satellite imagery that was aggregated to 5 km and averaged temporally to suit their dataset's temporal scale. In contrast, the results in this paper rely on satellite imagery that matches the dates of the surveys, and is based on children between 6 months and 5 years of age. Comparisons of maps based on different age ranges may lead to spurious conclusions, therefore a sensitivity analysis was conducted by reconstructing the prevalence data using the MAP age-adjustment model [59] to attain prevalence estimates for age 2 to 10 years and then re-created the prevalence maps per survey. A tenfold cross-validation estimates MAE being 36.9% and 12.5% on average across all surveys for MAP and the Bayesian model, respectively (Additional file 3). These results suggest that the main reason for these strikingly different prevalence maps is because the model took advantage of fine-scale malaria surveillance data, whereas MAP had to rely on spatially sparse data from national surveys (i.e., Demographic Health Surveys (DHS), MICS, MIS), in which the spatial coordinates are displaced slightly. Ultimately, these findings highlight the importance of fine-scale data to enable strategic allocation of resources and malaria prevention and control interventions.

The resulting maps predicted elevated risk during the rainy season, particularly in low-lying areas near the riverine plains. The prevalence across rainy seasons showed very little change implying that temporal change between years is limited, despite substantial within-year seasonal differences. Importantly, regardless of season and year, malaria prevalence was higher in the southern regions of the study area. Furthermore, the temporal validation also finds that 2-year ahead predictions of the spatial distribution of malaria prevalence were relatively accurate.

Taken together, these results suggest that, in the absence of large-scale changes in malaria interventions, active data collection to spatially guide malaria prevention and control interventions might not be required multiple times within a year nor every single year, an important consideration given the high costs associated with data collection.

The maps in Fig. 2 also reveal a strong urban–rural relationship, where the urban areas experience lower rates of malaria prevalence when compared to the rural areas across all seasons and years. Although the association between lower malaria prevalence and urban areas has long been acknowledged [60, 61], it was found that even relatively modest urban centres (e.g., population 6000–12,000) in rural districts have a strong protective effect. This could be because these centres have fewer mosquito breeding sites [62], better housing conditions that provide protection against mosquito-human contact [63], and/or urban residents are more likely to have better access to medications [64]. A detailed analyses on the non-linear relationship between distance to urban centre and malaria prevalence in BYD has been reported elsewhere [4]. Acknowledging this strong effect of modest urban centres may help malaria prevention and control programmes better allocate their interventions. For example, after presenting the results to the local IRS programme, it was suggested that IRS could be prioritized to peri-urban and rural areas because urban areas, while easier to access, are more difficult to spray as heads of households are less likely to be at home during the day, making it difficult to reach the targeted structures to be sprayed.

Several environmental factors previously identified as driving malaria transmission in Ghana at the national level (e.g., rainfall and temperature [5, 64]) were not strong determinants of fine-scale malaria prevalence heterogeneity and were removed prior to modelling efforts. Furthermore, when fitting the geospatial model, many commonly used covariates were not statistically significant, including NDVI, distance to roads and water bodies, LST for night, and night-time lights. Although these covariates are identified as significant in the multivariate model used in step-wise selection, the effect of these covariates was probably well captured by the spatial random effects in the Bayesian model. Alternative reasons for these environmental covariates not being significant might be because, at small spatial scales, there is little variation in them and perhaps part of the effect of these variables is already being captured by distance to urban centre. It is possible that socio-economic variables (e.g., wealth) play a much more important role at this spatial scale than these environmental covariates. Finally, individual level non-spatial covariate were restricted to age only to avoid interpretation issues regarding the predicted

prevalence surface. However, the model is designed such that more individual level covariates can be included. A more detailed study on the inference of socio-economic, demographic, environmental, and intervention at individual level has been conducted elsewhere [4].

Given the prohibitive cost associated with collecting fine-scale data on malaria prevalence over a large geographical region, an important aim of this paper was to determine the reliability of extrapolating model predictions to neighbouring districts. Using readily available survey data from MICS 2011, 10 clusters were obtained (six in Bunkpurugu Yunyoo and four in East Mamprusi) with observed malaria prevalence for October 2011, which is a small validation set. Nevertheless, these independent data suggest that the extrapolations results for both East Mamprusi and Bunkpurugu Yunyoo were still relatively accurate for the Bayesian model compared to published maps (16 and 37.6%, respectively) and therefore can still be useful for programmatic malaria prevention and control activities. Note that the extrapolation to a neighbouring district with was done using a similar environmental conditions and for which GIS data on urban city localities and active health facilities was readily available. Additional work will be needed to determine how extrapolation quality is impacted when environmental conditions differ substantially between districts, and to determine how generalizable these findings are.

The data presented here were based on the first six surveys during which a pyrethroid-based insecticide was applied across the district. Although the direct effects of the IRS intervention have not been accounted for, it was expected that IRS would result in declines of malaria prevalence across the district. Interestingly, the model results did not find much change over time with the exception of the sixth survey in the dry season of 2013. This suggests a potential lag in the impact of IRS across the district or the presence of substantial resistance to pyrethroids. Indeed, pyrethroids were subsequently switched to organophosphates and follow-on surveys not presented here revealed a significant drop in malaria prevalence [3]. Further work in this area would include modelling these follow-on surveys whilst adjusting for IRS intervention in the study area. The methodology developed in this paper would help identify how IRS spraying is impacting the spatial and temporal patterns of malaria in the region.

An important component of the model outputs are uncertainty estimates, defined here as the width of the 95% credible interval. The predicted uncertainty maps are indications of the precision around the mean estimated prevalence at a given location (i.e., at the pixel level), which helps users understand the robustness of the predicted mapped surface. Factors that contribute to uncertainty can include sparseness in the observed survey data (i.e., not enough observations in a given area), and/or inability of the model to explain the variability in the data [65, 66]. The results suggest that areas predicted to have elevated risk as well as higher uncertainty were mostly rural. These communities tend to be spread across larger geographic areas compared to the urban centres and the geographical coordinates, which were collected at easily recognizable landmarks in each community, may not be an accurate representation of the environment of a typical household in these rural communities. These results suggest that studies in many of these more rural areas may benefit from collection of GPS coordinates for each household or at least for clusters of households, instead of village-level coordinates, or sampling frameworks that are designed based on spatially stratified random sampling instead of solely on population proportional random sampling [67]. Finally, these uncertainty maps may help in the identification of areas in need of additional sampling [68].

Conclusion

This study demonstrated how the use of high-resolution survey data and a geostatistical model can reveal local-scale spatial heterogeneity in an area previously assumed to be relatively uniform in terms of malaria risk. Characterizing the heterogeneity in the spatial distribution of malaria in this small geographic area enabled the identification of areas of high risk for which malaria prevention and control efforts can be strategically allocated to reduce malaria transmission in Bunkpurugu-Yunyoo, Ghana. Spatial extrapolations to neighbouring districts revealed that it is possible to take advantage of the rich data in one area to gain insight on the spatial heterogeneity in another, and temporal extrapolation results suggest that 2-year predictions can be made with similar accuracy as spatial predictions. These extrapolation results, together with limited seasonal and between-year variability in the location of hotspots, suggest that fine-scale data collection on malaria prevalence can be conducted less frequently, which is an important consideration for the long-term financial sustainability of these efforts, while still strategically guiding malaria prevention and control interventions.

Additional files

Additional file 1. Spatial covariates maps for study region.

Additional file 2. Variable selection, posterior estimates and model comparisons.

Additional file 3. Sensitivity analysis to comparing age standardized local maps.

Amratia *et al. Malar J*   (2019) 18:81

Page 12 of 14

Abbreviations

AIC: Akaike information criterion; BYD: Bunkpurugu-Yunyoo district; CHIRPS: Climate Hazards Group InfraRed Precipitation with Station data; CI: credible intervals; CIESIN: Center for International Earth Science Information Network; DHS: Demographic Health Surveys; GLM: Generalized linear regression model; GHS: Ghana Health Service; GIS: Geographical Information System; GPS: Geographic positioning system; MIS: Ghana Malaria Indicator Survey; IRS: indoor residual spraying; LLIN: long lasting insecticide treated nets; LST: land surface temperature; MAP: Malaria Atlas Project; MAE: mean absolute error; MODIS: Moderate Resolution Imaging Spectroradiometer; MICS: Multiple Indicator Cluster Survey; NASA: National Aeronautics and Space Administration; NMCP: National Malaria Control Programme; NOAA: National Oceanic and Atmospheric Administration; NDVI: normalized difference vegetation index; PMI: Presidents Malaria Initiative; SMC: seasonal malaria chemoprophylaxis; SRTM-DEM: Shuttle Radar Topography Mission Digital Elevation Model.

Author details

[1] School of Forest Resources and Conservation, University of Florida, Gainesville, USA. [2] Emerging Pathogens Institute, University of Florida, Gainesville, USA. [3] Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Accra, Ghana. [4] National Malaria Control Programme, Accra, Ghana.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The data files used in this analysis are available at https://github.com/justinmillar/bma-manuscript. Further information and code is available upon request from the authors.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Ethical approval for the data collection was granted by the Institutional Review Board (IRB) of the Noguchi Institute for Medical Research at the University of Ghana (NMIMR IRB CPN#009-10-11 revd 2013, FWA 001,824/IRB 908). Approval for faculty and student involvement in the follow-on analysis of de-identified data was given by the University of Florida (IRB201500051).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. WHO. World malaria report 2018. Geneva: World Health Organization; 2018.
2. PMI, Ghana. Ghana Malaria Operational Plan FY 2014. 2015. http://www.pmi.gov/docs/default-source/default-document-library/malaria-operational-plans/fy-15/fy-2015-ghana-malaria-operational.
3. PMI, Ghana. President's Malaria Initiative Indoor Residual Spraying in Bunkpurugu Yunyoo District, Northern Region of Ghana. 2015. https://www.pmi.gov/news/stories-from-the-field/stories-from-the-field—detail/president-s-malaria-initiative-indoor-residual-spraying-in-bunkpurugu-yunyoo-distict-northern-region-of-ghana. Accessed Jan 1 2017.
4. Millar J, Psychas P, Abuaku B, Ahorlu C, Amratia P, Koram K, et al. Detecting local risk factors for residual malaria in northern Ghana using Bayesian model averaging. Malar J. 2018;17:343.
5. INFORM project, NMCP Ghana. An epidemiological profile of malaria and its control in Ghana. A report prepared for the Ministry of Health, Ghana, the Roll Back Malaria Partnership and the Department for International Development, UK. November, 2013. 2013.
6. Baird JK, Owusu Agyei S, Utz GC, Koram K, Barcus MJ, Jones TR, et al. Seasonal malaria attack rates in infants and young children in northern Ghana. Am J Trop Med Hyg. 2002;66:280–6.
7. Awine T, Malm K, Bart-Plange C, Silal SP. Towards malaria control and elimination in Ghana: challenges and decision making tools to guide planning. Glob Health Action. 2017;10:1381471.
8. Abiodun GJ, Maharaj R, Witbooi P, Okosun KO. Modelling the influence of temperature and rainfall on the population dynamics of *Anopheles arabiensis*. Malar J. 2016;15:364.
9. Assoum M, Ortu G, Basáñez MG, Lau C, Clements ACA, Halton K, et al. Spatiotemporal distribution and population at risk of soil-transmitted helminth infections following an eight-year school-based deworming programme in Burundi, 2007–2014. Parasit Vectors. 2017;10:583.
10. Bannister-Tyrrell M, Verdonck K, Hausmann-Muela S, Gryseels C, Muela Ribera J, Peeters Grietens K. Defining micro-epidemiology for malaria elimination: systematic review and meta-analysis. Malar J. 2017;16:164.
11. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. Nature. 2015;526:207–11.
12. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al. A world malaria map: *Plasmodium falciparum* endemicity in 2007. PLoS Med. 2009;6:e1000048.
13. Noor AM, Kinyoki DK, Mundia CW, Kabaria CW, Mutua JW, Alegana VA, et al. The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000–2010: a spatial and temporal analysis of transmission intensity. Lancet. 2014;383:1739–47.
14. Monroe A, Asamoah O, Lam Y, Koenker H, Psychas P, Lynch M, et al. Outdoor-sleeping and other night-time activities in northern Ghana: implications for residual transmission and malaria prevention. Malar J. 2015;14:35.
15. Baidjoe AY, Stevenson J, Knight P, Stone W, Stresman G, Osoti V, et al. Factors associated with high heterogeneity of malaria at fine spatial scale in the Western Kenyan highlands. Malar J. 2016;15:307.
16. Sturrock HJW, Cohen JM, Keil P, Tatem AJ, Le Menach A, Ntshalintshali NE, et al. Fine-scale malaria risk mapping from routine aggregated case data. Malar J. 2014;13:421.
17. Walker M, Winskill P, Basáñez M-G, Mwangangi JM, Mbogo C, Beier JC, et al. Temporal and micro-spatial heterogeneity in the distribution of Anopheles vectors of malaria along the Kenyan coast. Parasit Vectors. 2013;6:311.
18. Macharia PM, Giorgi E, Noor AM, Waqo E, Kiptui R, Okiro EA, et al. Spatio-temporal analysis of *Plasmodium falciparum* prevalence to understand the past and chart the future of malaria control in Kenya. Malar J. 2018;17:340.
19. Killeen GF. Characterizing, controlling and eliminating residual malaria transmission. Malar J. 2014;13:330.
20. Abuaku B, Ahorlu C, Psychas P, Ricks P, Oppong S, Mensah S, et al. Impact of indoor residual spraying on malaria parasitaemia in the

Amratia *et al. Malar J*   (2019) 18:81

Page 13 of 14

Bunkpurugu-Yunyoo district in northern Ghana. Parasit Vectors. 2018;11:555.

21. Coleman S, Dadzie SK, Seyoum A, Yihdego Y, Mumba P, Dengela D, et al. A reduction in malaria transmission intensity in Northern Ghana after 7 years of indoor residual spraying. Malar J. 2017;16:324.

22. Ghana Statistical Service. 2010 Population and Housing Census. National Analytical Report. 2013. p. 305–37. http://www.statsghana .gov.gh/docfiles/2010phc/National_Analytical_Report.pdf. Accessed 5 Feb 2017.

23. The Presidents Malaria Initative (PMI). Ghana Malaria Operational Plan FY 2014. 2014. https://www.pmi.gov/docs/default-source/defau lt-document-library/malaria-operational-plans/fy14/ghana_mop_fy14. pdf?sfvrsn=20. Accessed 1 Feb 2017.

24. Macdonald G. The analysis of malaria parasite rates in infants. Trop Dis Bull. 1950;47:915–38.

25. Riley EM, Wagner GE, Akanmori BD, Koram KA. Do maternally acquired antibodies protect infants from malaria infection? Parasite Immunol. 2001;23:51–9.

26. Hay SI. Remote sensing and disease control: past, present and future. Trans R Soc Trop Med Hyg. 1997;91:105–6.

27. Weiss DJ, Mappin B, Dalrymple U, Bhatt S, Cameron E, Hay SI, et al. Re-examining environmental correlates of *Plasmodium falciparum* malaria endemicity: a data-intensive variable selection approach. Malar J. 2015;14:68.

28. Guerra CA, Snow RW, Hay SI. A global assessment of closed for- ests, deforestation and malaria risk. Ann Trop Med Parasitol. 2006;100:189–204.

29. Paaijmans KP, Thomas MB. The influence of mosquito resting behaviour and associated microclimate for malaria risk. Malar J. 2011;10:183.

30. Pascual M, Ahumada JA, Chaves LF, Rodó X, Bouma M. Malaria resurgence in the East African highlands: temperature trends revisited. Proc Natl Acad Sci USA. 2006;103:5829–34.

31. Krefis AC, Schwarz NG, Krüger A, Fobil J, Nkrumah B, Acquah S, et al. Modeling the relationship between precipitation and malaria incidence in children from a holoendemic area in Ghana. Am J Trop Med Hyg. 2011;84:285–91.

32. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. The WorldClim interpolated global terrestrial climate surfaces. 2005. http://www.world clim.org. Accessed 1 Aug 2016.

33. NOAA. Version 4 DMSP-OLS Nighttime Lights Time Series. Natl. Centers Environ. Inf. 2014. https://ngdc.noaa.gov/eog/dmsp/downloadV4compo sites.html. Accessed 1 Aug 2016.

34. Noor AM, Alegana VA, Gething PW, Tatem AJ, Snow RW. Using remotely sensed night-time light as a proxy for poverty in Africa. Popul Health Metr. 2008;6:5.

35. Tatem AJ, Guerra CA, Kabaria CW, Noor AM, Hay SI. Human population, urban settlement patterns and their impact on *Plasmodium falciparum* malaria endemicity. Malar J. 2008;7:218.

36. Stevens FR, Gaughan AE, Linard C, Tatem AJ, Sorichetta A, Hornby GM, et al. WorldPop-RF, Version 2b.1.1. figshare. 2015. http://figshare.com/artic les/WorldPop_RF/1491490.

37. Smith C, Whittaker M. Beyond mobile populations: a critical review of the literature on malaria and population mobility and suggestions for future directions. Malar J. 2014;13:307.

38. ESRI. ArcGIS Desktop: Release 10.3. Redlands CA. 2014. p. Environmental Systems Research Institute.

39. Siri JG, Lindblade KA, Rosen DH, Onyango B, Vulule J, Slutsker L, et al. Quantitative urban classification for malaria epidemiology in sub-Saharan Africa. Malar J. 2008;7:34.

40. Kadobera D, Sartorius B, Masanja H, Mathew A, Waiswa P. The effect of distance to formal health facility on childhood mortality in rural Tanzania, 2005–2007. Glob Health Action. 2012;5:1–9.

41. Pinchoff J, Chaponda M, Shields T, Lupiya J, Kobayashi T, Mulenga M, et al. Predictive malaria risk and uncertainty mapping in Nchelenge district, Zambia: evidence of widespread, persistent risk and implications for targeted interventions. Am J Trop Med Hyg. 2015;93:1260–7.

42. Ferrao JL, Niquisse S, Mendes JM, Painho M. Mapping and modelling malaria risk areas using climate, socio-demographic and clinical variables in Chimoio, Mozambique. Int J Environ Res Public Health. 2018;15:e795.

43. Zhou SS, Zhang SS, Wang JJ, Zheng X, Huang F, Li WD, et al. Spatial cor- relation between malaria cases and water-bodies in *Anopheles sinensis* dominated areas of Huang-Huai plain. China. Parasit Vectors. 2012;5:106.

44. Weiss DJ, Nelson A, Gibson HS, Temperley W, Peedell S, Lieber A, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature. 2018;553:333–6.

45. ESRI. Digital Chart of the World (DCW). Inland waters bodies. http://www. diva-gis.org/gdata. Accessed 5 Apr 2016.

46. CIESIN. Global Roads Open Access Data Set, Version 1 (gROADSv1). Pali- sades, NY: NASA Socioeconomic Data and Applications Center (SEDAC); 2013. http://dx.doi.org/10.7927/H4VD6WCT. Accessed 4 Apr 2016.

47. Griffin JT, Ferguson NM, Ghani AC. Estimates of the changing age-burden of *Plasmodium falciparum* malaria disease in sub-Saharan Africa. Nat Commun. 2014;5:3136.

48. Diggle P, Moyeed R, Rowlingson B, Thomson M. Childhood malaria in the Gambia: a case-study in model-based geostatistics. J R Stat Soc Ser C Appl Stat. 2002;51:493–506.

49. Jarvis A, Reuter HI, Nelson A, Guevara E. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90 m Database. Consort Spat Inf. 2016;1–9. http://www.cgiar-csi.org/data/srtm-90m-digital-eleva tion-database-v4-1.

50. NASA LP DAAC. MODIS Level 1 Land surface temperatures registered at-sensor radiance. Version 5. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. 2015. https://lpdaac.usgs.gov/dataset_discovery/modis/ modis_products_table/mod13c1. Accessed 15 Jan 2016.

51. Funk C, Verdin A, Michaelsen J, Peterson P, Pedreros D, Husak G. A global satellite-assisted precipitation climatology. Earth Syst Sci Data. 2015;7:275–87.

52. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37:4302–15.

53. Murtaugh PA. Performance of several variable-selection methods applied to real ecological data. Ecol Lett. 2009;12:1061–8.

54. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med. 2004;66:411–21.

55. Slater H, Michael E. Mapping, Bayesian geostatistical analysis and spatial prediction of lymphatic filariasis prevalence in Africa. PLoS ONE. 2013;8:e71574.

56. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. http:// www.r-project.org/.

57. Ministry of Health Ghana. Health Facilities. 2016. Available from: http:// data.gov.gh/dataset/health-facilities. Accessed 30 June 2016.

58. Ghana Statistical Service. Ghana multiple indicator cluster surveys for 2006. 2008. http://www.statsghana.gov.gh/nada/index.php/catalog/89. Accessed 5 Feb 2017.

59. Smith DL, Guerra CA, Snow RW, Hay SI. Standardizing estimates of the *Plasmodium falciparum* parasite rate. Malar J. 2007;6:131.

60. Pond BS. Malaria indicator surveys demonstrate a markedly lower preva- lence of malaria in large cities of sub-Saharan Africa. Malar J. 2013;12:313.

61. Tatem AJ, Gething PW, Smith DL, Hay SI. Urbanization and the global malaria recession. Malar J. 2013;12:133.

62. Cano J, Descalzo MA, Moreno M, Chen Z, Nzambo S, Bobuakasi L, et al. Spatial variability in the density, distribution and vectorial capacity of anopheline species in a high transmission village (Equatorial Guinea). Malar J. 2006;5:21.

63. Baragatti M, Fournet F, Henry M-C, Assi S, Ouedraogo H, Rogier C, et al. Social and environmental malaria risk factors in urban areas of Ouaga- dougou, Burkina Faso. Malar J. 2009;8:13.

64. De Silva PM, Marshall JM. Factors contributing to urban malaria transmission in sub-saharan Africa: a systematic review. J Trop Med. 2012;2012:819563.

65. Diggle PJ, Ribeiro PJ. Model-Based Geostatistics. Biometrics. 2008;64:653–4.

Amratia *et al. Malar J*   (2019) 18:81

Page 14 of 14

66. Gething P, Tatem A, Bird tom, Burgert-Brucker CR. Creating spatial interpolation surfaces with DHS data. DHS Spat Anal Reports No 11. 2015. http://dhsprogram.com/pubs/pdf/SAR11/SAR11.pdf. Accessed 5 Feb 2017.

67. Delmelle EM. Spatial sampling. In: Fischer MM, Nijkamp P, editors. Handbook of regional science. Berlin, Heidelberg: Springer; 2014. p. 1385–1399.

68. Grist EPM, Flegg JA, Humphreys G, Mas IS, Anderson TJC, Ashley EA, et al. Optimal health and disease management using spatial uncertainty: a geographic characterization of emergent artemisinin-resistant *Plasmodium falciparum* distributions in Southeast Asia. Int J Health Geogr. 2016;15:37.