



RESEARCH ARTICLE

REVISION **Geostatistical analysis of Malawi's changing malaria transmission from 2010 to 2017** [version 2; peer review: 2 approved, 1 approved with reservations]

Michael Give Chipeta ¹, Emanuele Giorgi ², Donnie Mategula ¹, Peter M. Macharia ³, Chimwemwe Ligomba ¹, Alinane Munyenembe ¹, James Chirombo¹, Austin Gumbo⁴, Dianne J. Terlouw ^{1,5}, Robert W. Snow ^{3,6}, Michael Kayange⁴

¹Malaria Epidemiology Group, Malawi-Liverpool Wellcome Trust Research Programme, Blantyre, Malawi

²Lancaster Medical School, Lancaster University, Lancaster, LA1 4YW, UK

³Population Health Unit, Kenya Medical Research Institute - Wellcome Trust Research Programme, Nairobi, Kenya

⁴National Malaria Control Programme, Malawi Ministry of Health, Lilongwe, Malawi

⁵Liverpool School of Tropical Medicine, Liverpool, L3 5QA, UK

⁶Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, OX1 2JD, UK

v2 First published: 27 Mar 2019, 4:57 (<https://doi.org/10.12688/wellcomeopenres.15193.1>)

Latest published: 04 Jul 2019, 4:57 (<https://doi.org/10.12688/wellcomeopenres.15193.2>)

Abstract



Background: The prevalence of malaria infection in time and space provides important information on the likely sub-national epidemiology of malaria burdens and how this has changed following intervention. Model-based geostatistics (MBG) allow national malaria control programmes to leverage multiple data sources to provide predictions of malaria prevalence by district over time. These methods are used to explore the possible changes in malaria prevalence in Malawi from 2010 to 2017.

Methods: *Plasmodium falciparum* parasite prevalence (*PfPR*) surveys undertaken in Malawi between 2000 and 2017 were assembled. A spatio-temporal geostatistical model was fitted to predict annual malaria risk for children aged 2–10 years (*PfPR*_{2–10}) at 1×1 km spatial resolutions. Parameter estimation was carried out using the Monte Carlo maximum likelihood methods. Population-adjusted prevalence and populations at risk by district were calculated for 2010 and 2017 to inform malaria control program priority setting.

Results: 2,237 surveys at 1,834 communities undertaken between 2000 and 2017 were identified, geo-coded and used within the MBG framework to predict district malaria prevalence properties for 2010 and 2017. Nationally, there was a 47.2% reduction in the mean modelled *PfPR*_{2–10} from 29.4% (95% confidence interval (CI) 26.6 to 32.3%) in 2010 to 15.2% (95% CI 13.3 to 18.0%) in 2017. Declining prevalence was not equal across the country, 25 of 27 districts showed a substantial decline ranging from a 3.3% reduction to 79% reduction. By 2017, 16% of Malawi's population still lived in areas that support *PfPR*_{2–10} ≥ 25%.

Open Peer Review

Reviewer Status

	Invited Reviewers		
	1	2	3
<div>REVISED</div> <div>version 2</div> <div>published 04 Jul 2019</div>	<div>✓</div> <div>report</div>		<div>✓</div> <div>report</div>
<div>version 1</div> <div>published 27 Mar 2019</div>	<div>↑</div> <div>?</div> <div>report</div>	<div>?</div> <div>report</div>	<div>↑</div> <div>?</div> <div>report</div>
<div>1 Rocco Panciera, United Nation Children's Fund (UNICEF), New York, USA</div> <div>2 Tobias Chirwa , University of the Witwatersrand, Johannesburg, South Africa</div> <div>3 Theresa Smith , University of Bath, Bath, UK</div>			
Any reports and responses or comments on the article can be found at the end of the article.			

Conclusions: Malawi has made substantial progress in reducing the prevalence of malaria over the last seven years. However, Malawi remains in *meso*-endemic malaria transmission risk. To sustain the gains made and continue reducing the transmission further, universal control interventions need to be maintained at a national level.

Keywords

Model-based geostatistics, malaria, Malawi, Plasmodium falciparum



This article is included in the [Malawi-Liverpool Wellcome Trust Clinical Research Programme gateway](#).

Corresponding author: Michael Give Chipeta (mgchipeta@mlw.mw)

Author roles: **Chipeta MG:** Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Giorgi E:** Data Curation, Formal Analysis, Investigation, Methodology, Software, Writing – Review & Editing; **Mategula D:** Writing – Original Draft Preparation, Writing – Review & Editing; **Macharia PM:** Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ligomba C:** Writing – Review & Editing; **Munyenyerembe A:** Writing – Review & Editing; **Chirombo J:** Writing – Review & Editing; **Gumbo A:** Writing – Review & Editing; **Terlouw DJ:** Investigation, Project Administration, Resources, Supervision, Validation, Writing – Review & Editing; **Snow RW:** Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Kayange M:** Resources, Supervision, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was funded by the UK's Department for International Development under a project entitled Strengthening the Use of Data for Malaria Decision Making in Africa (DFID Programme Code 203155). MGC was supported by Wellcome Trust Career Bridging Fellowship (101113); PMM acknowledges support of the IDeALs Project (107769); RWS is funded as a Principal Wellcome Fellow (103602). DJT acknowledges the support of the Wellcome Trust to the Malawi Major Overseas Programme (206545). PMM and RWS acknowledge the support of the Wellcome Trust to the Kenya Major Overseas Programme (203077).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Chipeta MG *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Chipeta MG, Giorgi E, Mategula D *et al.* **Geostatistical analysis of Malawi's changing malaria transmission from 2010 to 2017 [version 2; peer review: 2 approved, 1 approved with reservations]** Wellcome Open Research 2019, 4:57 (<https://doi.org/10.12688/wellcomeopenres.15193.2>)

First published: 27 Mar 2019, 4:57 (<https://doi.org/10.12688/wellcomeopenres.15193.1>)

REVISED Amendments from Version 1

We would like to thank the reviewers for their helpful feedback. The manuscript has been modified to address the comments made by the three reviewers and improve on clarity. The main modifications in this version are related to the methods and results sections. We have changed statistical “significance” to “substantial” district level prevalence changes in the results. We have clarified the number of unique and repeat survey locations. We also include maps showing exceedance and non-exceedance probabilities to quantify decrease in prevalence between 2010 and 2017. A new reference to climate changes over the study period has been added. We have uploaded a dataset to figshare repository which now includes minimum and maximum ages of subjects.

See referee reports

Introduction

Malaria-endemic countries are increasingly encouraged to define their sub-national epidemiology to understand the likely rational allocation of intervention mixes and the changing malaria landscape [WHO, 2010; WHO, 2015]. Within a country, variations in vector ecology, environment, and intervention coverage all determine the patterns of malaria risk and incipient disease burden. Consequently, maps of malaria risk are required by national malaria control programmes to guide decision making in heterogeneous settings.

National cartographies of malaria risk were common during the 1950s and 1960s [Snow & Noor, 2015]. Often these maps were based on accepted definitions of malaria endemicity [Lysenko & Semashko, 1968; Metselaar & van Thiel, 1959], derived from community-based surveys of malaria infection prevalence and climate ecologies [Snow & Noor, 2015]. The importance of malaria risk maps began to re-emerge during the 1990s, coincident with the pan-African malaria resurgent epidemic [Snow *et al.*, 1996], resulting in a new generation of national malaria strategic plans that articulated the sub-national disparities in malaria ecology, transmission and disease burden [Omumbo *et al.*, 2013].

In Malawi, as part of the 2001–2005 national malaria strategic plan [NMCP, 2001], the epidemiology of malaria was described based on a World Health Organization (WHO) mission to the country 30 years earlier [Cheyabejara *et al.*, 1974] “In 1973, the WHO mission determined malaria to be meso- to hyper-endemic in Malawi, except in isolated higher altitudes mountainous regions” [NMCP, 2001]. However, during the second national malaria strategic plan 2005–2010 [NMCP, 2005], no references to the epidemiological patterns of transmission, dominant vector species or disease burden were provided.

It was not until 2006 that empirical data on malaria infection prevalence was used with model-based geostatistical (MBG) methods to provide predictive quantities of risk across Malawi [Kazembe *et al.*, 2006]. This work used data on malaria prevalence from 73 survey locations, where children aged 1–10 years that had been sampled between 1970 and 2001. Temperature, rainfall, potential evapotranspiration, and elevation were all used as covariates to help predict infection prevalence at un-sampled

locations using information and correlates with sampled locations. This map was used in the malaria programme review in 2010 [NMCP, 2010] and the national strategic plan 2011–2015 to highlight the hyper-endemic nature of malaria transmission in the country, with variations in higher altitude areas [NMCP, 2011].

In 2013, the application of MBG methods was extended to include 1057 surveys of malaria infection undertaken between 2000 to 2010, employing covariates related to urbanization and temperature suitability for malaria transmission to provide a 1×1 km posterior prediction of malaria risk in 2000, 2005 and 2010 [Bennet *et al.*, 2013]. These analyses showed significant sub-national variations in malaria prevalence; however, there was little change across the prediction time-periods.

Since 2010, there have been significant additional community-based malaria surveys undertaken, sampled both at district and national levels. This paper describes the re-assembly of malaria infection data in Malawi and their use within a MBG framework to understand changing sub-national malaria endemicity between 2010 and 2017 at a time of increased malaria intervention coverage throughout the national strategic plan 2011–2015 [NMCP, 2011].

Methods

Geography

Malawi is a landlocked country located in the South Eastern region of Africa along the Great Rift Valley, bordered by Tanzania, Mozambique and Zambia (Figure 1). The country has an estimated population of 17.6 million people in 2017 [NSO, 2018], and is one of the poorest countries in Africa with a per capita GDP of US\$350 [World Bank, 2018]. Altitude varies considerably from 37 metres above mean sea level (MASL) to 3003 MASL. Both proximity to Lake Malawi and altitude define the variable climate patterns [Vincent *et al.*, 2014]. The country is divided into three regions, namely Northern, Central, and Southern regions, which encompass 28 districts (Figure 1). This includes the small islands of Likoma and Chizumulu that form Likoma district, with *circa* 9,000 people, in Mozambique waters (Figure 1). For the purposes of the present analysis Likoma district is excluded as it is situated more than 69 km from the mainland.

Malaria control 20 years on

Malawi launched a national malaria strategic plan in 2001 in line with recommendations made by the Roll Back Malaria initiative [NMCP, 2001]. With rapidly emerging sulphadoxine-pyrimethamine (SP) resistance, Malawi proposed a change in its first-line treatment policy from SP to Artemether-Lumefantrine (AL) in 2004, but the policy was not implemented until 2007–2008 [Malenga *et al.*, 2009]. A programme of socially marketed, subsidized insecticide-treated net (ITN) distributions was launched in 1998 [Mathanga *et al.*, 2012; MoH, 2005], and was largely the sole source of ITN access nationwide until 2007 when the free delivery of ITN through public health facilities to children attending immunization and pregnant mothers was launched [Mathanga *et al.*, 2012]. Between 2007 and 2010, approximately 4 million nets were distributed free of charge to vulnerable populations, including 1.1 million nets provided during the first nationwide

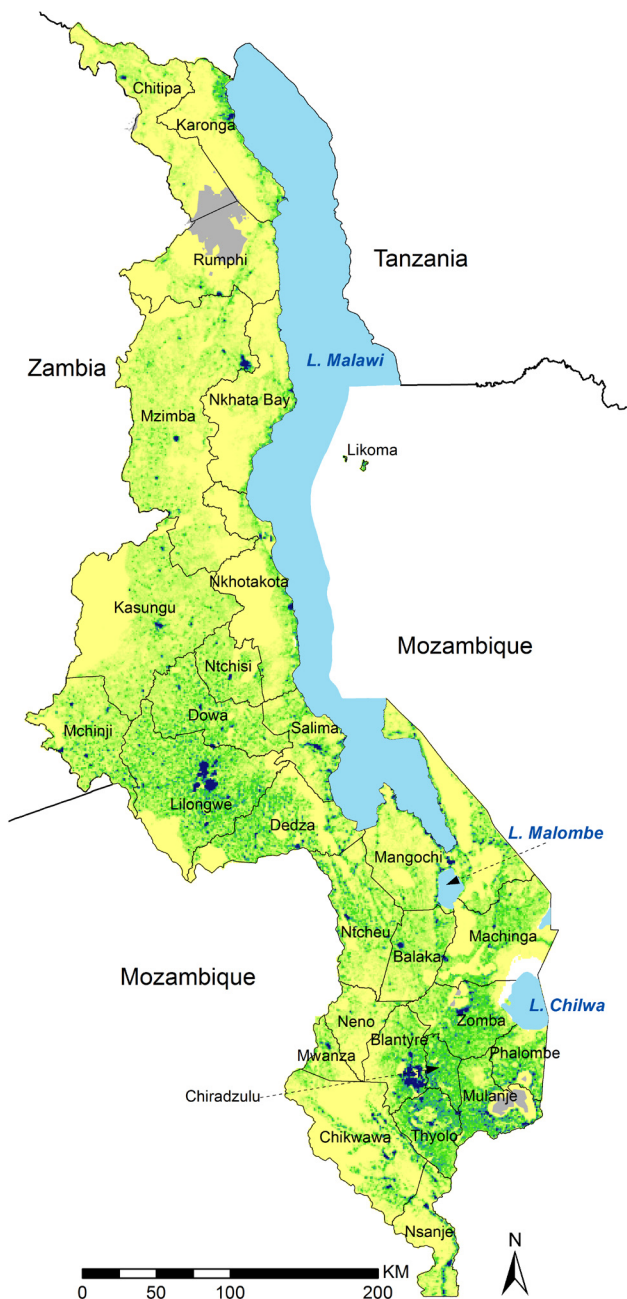


Figure 1. The Geography, population density, districts and unsuitability for malaria transmission in Malawi. Population density ranges from zero (yellow) to 37,332 person per 1 km grid (dark blue). Grey areas represent a temperature suitability index (TSI) of zero which indicates a temperature range that cannot support malaria parasite development cycles in the mosquito [Gething *et al.*, 2011], and all correspond to unpopulated areas in the Nyika Plateau in the north and Mulanje Massif range in the south.

mass-campaign in 2008. Before 2010, only pilot indoor-residual house spraying (IRS) using pyrethroids were undertaken in small community studies in Ntchisi, Mzimba, Kayelekera uranium mine and the Nchalo and Dwangwa Estates of the Illovo sugar

company [Chunga & Kumwenda, 2014; MoH, 2011]. Between 2000 and 2010, partial increases in vector control coverage [Chanda *et al.*, 2016] and the delayed introduction of effective therapeutics [Malenga *et al.*, 2009] was reflected in limited reductions in national community-based malaria prevalence, 2000 (36.4%) and 2010 (36.3%) [Bennett *et al.*, 2013] and increases over the same period in malaria hospitalizations [Okiro *et al.*, 2013; Roca-Feltrer *et al.*, 2012a].

In 2011, the National Malaria Strategy 2011–2015 was launched with a vision that all people in Malawi are free from the burden of malaria and an ambition to halve the malaria disease and mortality burden by 2016 [NMCP, 2011]. Substantial increases in long-lasting insecticide-treated net (LLIN) distribution were achieved from 2010 with over 1 million nets being delivered through public health facilities each year since 2011, and 5.4, 7.0 and 8.6 million nets distributed during a mass campaigns in 2012, 2014 and 2016, respectively. Between July 2010 and 2011, IRS using pyrethroids was undertaken in seven districts (Nkhosakota, Salima, Karonga, Nkhata Bay, Mangochi, Chikwakwa and Nsanje) protecting approximately 3 million people [NMCP, 2012]. Fuel shortages, funding interruptions and increasing Lambda-cyhalothrin and carbamate resistance [Mzilahowa *et al.*, 2016; Wondji *et al.*, 2012] led to a significant reduction in IRS activities, with Salima district only under IRS in 2013 where after IRS was suspended nationwide [Chanda *et al.*, 2015]. In 2011, the introduction of rapid diagnostic tests (RDTs) was launched nationwide with supplies to peripheral health facilities and training of health workers in 2012 and 2013 [PMI, 2016], which were included for health workers in hard to reach areas as part of integrated community case management (iCCM) in 2015 [Phiri *et al.*, 2016].

Malaria prevalence survey data assembly

The process of identifying community-based malaria survey data is described in detail elsewhere [Bennett *et al.*, 2013; Snow *et al.*, 2017]. Of importance have been national household sample surveys that have included malaria infection prevalence among children in national micro-nutrient surveys conducted in 2001, 2006, 2009 and 2016; national demographic and/or malaria indicator surveys undertaken in 2010, 2012, 2014 and 2017 [NMCP, 2017]; sub-national surveys undertaken at district levels by the College of Medicine Malaria Alert Centre between 2005–2009 [Mathanga *et al.*, 2010], repeat surveys in three districts between 2012 and 2014 [Walldorf *et al.*, 2015], and district-wide surveys in Chikwawa 2010–2016 [Buchwald *et al.*, 2016; Kabaghe *et al.*, 2017; Kabaghe *et al.*, 2018a; McCann *et al.*, 2017; Roca-Feltrer *et al.*, 2012b]. Other data were derived from published sources and the generous help with unpublished data from national malaria scientists and collaborators, listed at the end of the paper. Data were restricted to surveys undertaken between January 2000 and December 2017.

For each of the identified data sources, information on the month and year of the survey, age range of the surveyed population, the numbers tested versus numbers identified as harbouring *Plasmodium falciparum* and the methods of parasite detection were all extracted. The location (longitude and latitude) of each surveyed village, school or enumeration cluster was checked

using national statistical office high-resolution global positioning system (GPS) databases and other publicly available, online digital gazetteers. The raw information is provided as underlying data [Chipeta *et al.*, 2019; Snow, 2017].

Geostatistical spatio-temporal analysis

MBG [Diggle *et al.*, 1998; Diggle & Giorgi, 2019] is a likelihood-based approach that allows prediction of a health outcome of interest using sparsely sampled data. This modelling framework has also been extended to interpolate both the spatial and temporal variation of disease prevalence through the analysis of repeated cross-sectional data [Giorgi *et al.*, 2018]. MBG has become a well-established tool in statistics for modelling the spatio-temporal correlation induced by unmeasured risk factors to predict prevalence at any desired place and time.

To model changes in $PfPR_{2-10}$ by borrowing strength of information across time and space, an MBG model was used. Unlike previous MBG approaches [Bennett *et al.*, 2013] a decision was made not to include human settlement, climate or other environmental covariates during the modelling exercise. The inclusion of covariates (climate, land use, social economic status and intervention), when used to assist predictions at locations without data, presume a clearly defined *a priori* biological relationship with prevalence and are only valuable when predictions must be made without large volumes of input empirical prevalence data, which themselves represent the product of all the possible covariate influences [Macharia *et al.*, 2018].

The model is described as follows. Let x be the location of a surveyed community in year t . Define a spatio-temporal Gaussian process, $S(x, t)$, and unstructured random effects, $Z(x, t)$, to account for the unexplained variation between and within communities, respectively. Conditionally on $S(x, t)$ and $Z(x, t)$, the counts of positive tests for *P. falciparum* were assumed to follow mutually independent binomial distributions with number of trials N , corresponding to number of sampled individuals, and probability of a positive outcome $p(x, t)$ at location x (n =surveyed locations) and year t (2000–2017) given by

$$\log \left\{ \frac{p(x, t)}{1 - p(x, t)} \right\} = \alpha + \beta mA + \gamma MA + S(x, t) + Z(x, t). \quad (1)$$

where mA and MA are the minimum and maximum age among the sampled individuals at a location x and time t . In carrying the spatio-temporal predictions, mA and MA were set to 2 and 10 respectively to standardise to the age group 2–10 years. A stationary and isotropic Gaussian process for the spatio-temporal random effects is assumed $S(x, t)$, with an exponential correlation function given as

$$\text{cor}\{S(x, t), S(x', t')\} = e^{-\|u\|/\varphi} e^{-|v|/\psi} \quad (2)$$

where φ and ψ are scale parameters which regulate the rate of decay of the spatial and temporal correlation for the increasing distance and time separation, respectively; $u = \|x - x'\|$ is the distance in space between the location of any two communities, one at x and the other at x' ; $v = |t - t'|$ is the time separation in years between any two surveys.

The model parameters were estimated via Monte Carlo maximum likelihood in the R statistical software environment [R Core Team, 2017] using the [PrevMap](#) package version 1.4.2 [Giorgi & Diggle, 2017]. The targets for predictions were $PfPR_{2-10}$ over the 1×1 km regular grid covering the whole of mainland Malawi. Maps of malaria risk were generated for the two reference years 2010 and 2017 using ArcMap 10.4 (ESRI Inc., Redlands, CA, USA).

Model validation

The model was validated using two methods. First by testing evidence against the residual spatio-temporal correlation in the data through the following variogram-based validation algorithm [Giorgi *et al.*, 2018]: 1) Generate a point estimate $Z(x_i, t_i)$ i.e. $\tilde{Z}(x_i, t_i)$ from a non-spatio-temporal model, for each observed location x_i and time t_i ; 2) Permute the order of the data, including $\tilde{Z}(x_i, t_i)$, while holding (x_i, t_i) fixed; 3) Compute the empirical semi-variogram for $\tilde{Z}(x_i, t_i)$; 4) Repeat steps (1) and (2) a large number of times, say B ; 5) Using the resulting B empirical variograms to generate 95% confidence intervals at each of the pre-defined distance bins. To conclude that there is no evidence against the adopted spatio-temporal model correlation the empirical semi-variogram from the original data must fall within the generated 95% confidence intervals. Second, validation statistics based on a 10% hold-out dataset or correlation against observed and predicted estimates of $PfPR_{2-10}$, bias and mean absolute error was done.

Population-adjusted risk

Neither human settlement nor malaria risk are evenly distributed, and therefore to ensure that malaria risk maps converge with human population density, similar gridded surfaces of population are required. Dasymetric modelling techniques for the reallocation of populations within census units have been developed to overcome the difficulties caused by input census data of coarse spatial resolution [Linard *et al.*, 2012; Stevens *et al.*, 2015]. In brief, the 2008 Malawi national census, organized by 12,666 enumeration areas (lowest available level of aggregation), was reallocated using a Random Forest model in combination with land cover specific weightings, protected areas, night-time lights, roads, rivers, altitude and settlement data to adjust and re-allocate population densities within each enumeration area ([available here](#)). United Nations rural and urban growth rates were used to project population's forward to 2010 and 2017 ([available here](#)) to provide a gridded dataset of population distribution (counts) at 0.1×0.1 km resolution. The population maps were resampled to 1×1 km grids ([Figure 1](#)) to match the malaria risk mapped outputs.

Results

Description of survey data

A total of 2,276 independent survey data points at 1,874 unique locations were identified during the data assembly process. Two survey locations could not be geo-coded and 37 surveys were undertaken on the islands of Likoma district and were excluded. Of the remaining 2,237 surveys, 403 (18%) were repeat surveys, taken at the same geo-location but in different years and 1,834 were unique locations. The data covered the entire period 2000–2017

and represented the examination of 59,920 individuals. The majority of surveys employed microscopy (83%), rather than RDTs (17%) for parasite detection. Most survey locations (86%) were positioned using local GPS sources. The location of the age-corrected survey data is shown in [Figure 2](#).

Spatio-temporal variation in malaria risk

The assembled data were used in the spatio-temporal model Equation 1 to generate the 1×1 km grids of mean predictions of $PfPR_{2-10}$ in 2010 and 2017 ([Figure 3](#)). The validity of the spatio-temporal model indicated that the empirical semi-variogram

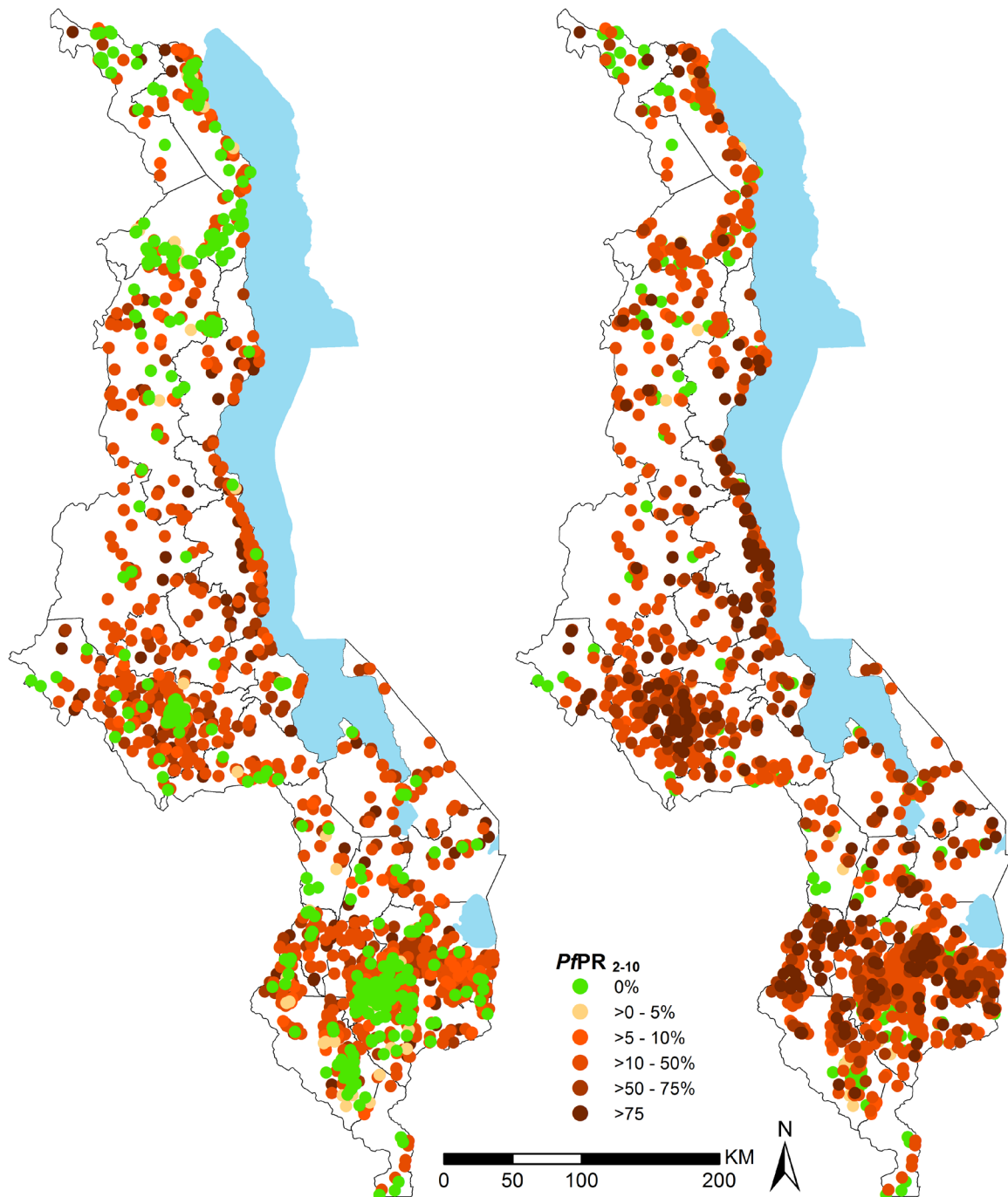


Figure 2. Spatial distribution of $PfPR_{2-10}$ surveys in Malawi between 2000 and 2017. Data assembled from 2,237 surveys at 1,834 unique locations of community parasite prevalence showing the lowest values of $PfPR_{2-10}$ on top (left panel) and highest values of $PfPR_{2-10}$ on top (right panel) to reflect locations sampled more than once during the period.

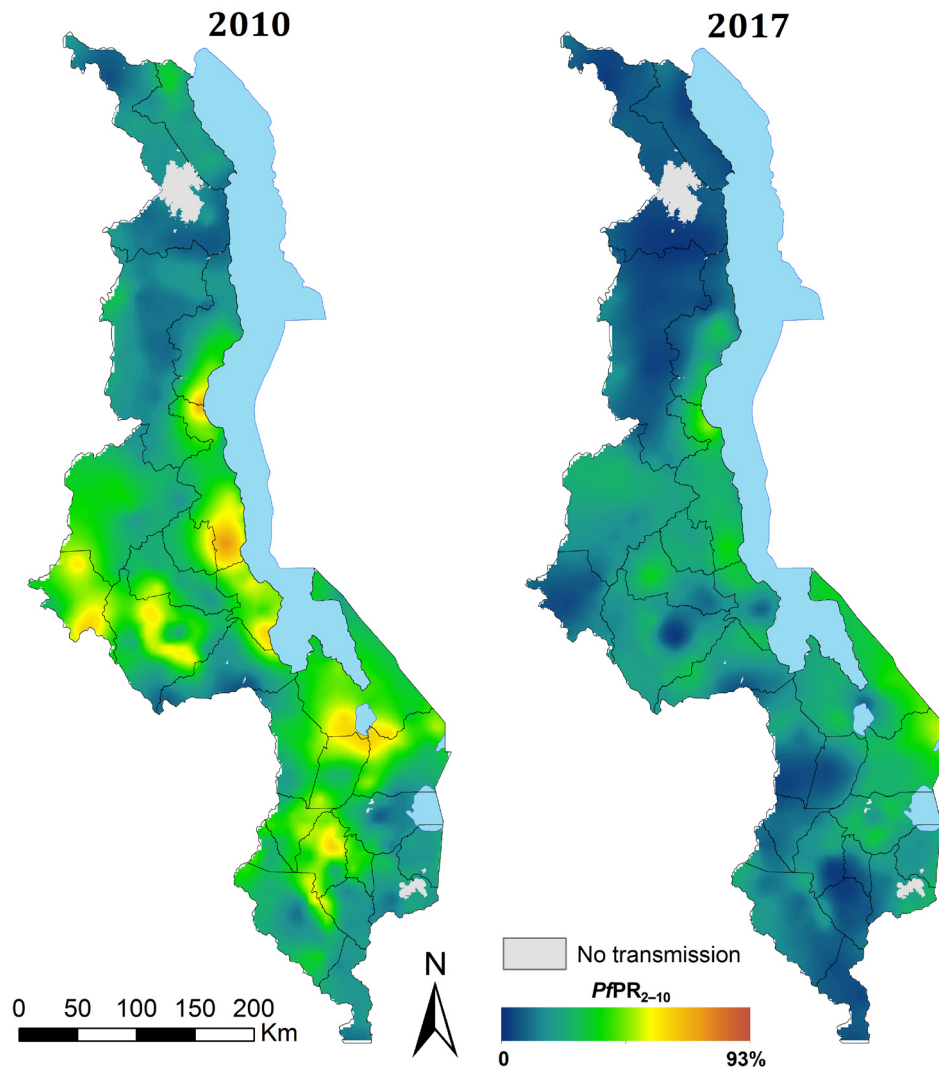


Figure 3. Mean standardized *Plasmodium falciparum* parasite rate ($PfPR_{2-10}$) for 2010 (left) and 2017 (right). The predicted posterior mean community $PfPR_{2-10}$ is presented at 1×1 km ranging from zero (dark blue) to 93% (dark red) in Malawi. Grey areas represent TSI values of zero, unable to support transmission.

falls within the 95% confidence intervals (Figure 4), indicating that there is no evidence against the adopted spatio-temporal model. The predictive performance of the model using cross validation, holding 10% of the total sample (224 surveys), indicated a high correlation between observed and predicted $PfPR_{2-10}$ with $\rho = 0.72$, a bias of 0.28% and mean absolute error of 15%.

The national mean predicted $PfPR_{2-10}$ in 2010 was 29.4% (95% confidence interval (CI) 26.6–32.3%) compared to 15.6% (95% CI 13.3–18.0%) in 2017. When combined with population density in each year this corresponds to a reduction of 94.7% in the numbers of people living in areas where $PfPR_{2-10}$ is greater than 40% and a corresponding 216.3% increase in populations living under areas of $PfPR_{2-10} < 20\%$ (Figure 5). As shown in Figure 3 and Figure 6, declines in $PfPR_{2-10}$ were witnessed nationwide. However, the largest declines ($\geq 60\%$ reductions)

in the population-weighted mean $PfPR_{2-10}$ by 2017 using 2010 as the baseline were observed in Karonga, Rumphi, Mchinji, Lilongwe, Balaka, Blantyre, Chiradzulu, Neno and Thyolo districts (Table 1; Figure 7). Conversely, those witnessing the lowest reductions ($< 10\%$) by 2017 compared to 2010 were observed in Dedza and Mulanje districts and two districts showing a rise during the interval, Zomba and Phalombe districts (Table 1; Figure 7). However, the confidence intervals for prevalence changes in Machinga, Mulanje and Phalombe districts contain a zero. To emphasise on prevalence decline, we categorised prevalence into two thresholds. In Figure 6, we show areas with low $PfPR_{2-10}$ prevalence where prevalence lies below 20% (non-exceedance probability-NEP, 80 or 90% sure) and areas with very high prevalence where $PfPR_{2-10}$ prevalence is above 30% (exceedance probability-EP, 80 or 90% sure), i.e. areas where intensive and sustained control is required.

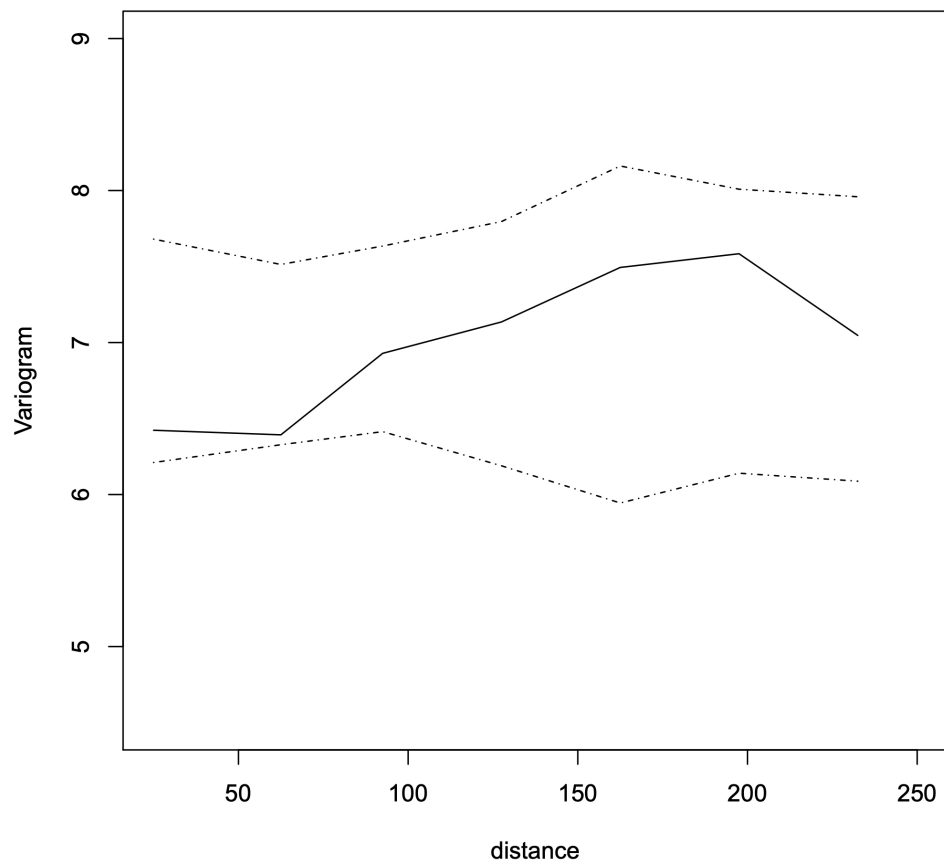


Figure 4. Validity of the assumed covariance model for the spatial correlation. The empirical semi-variogram (solid line) falls within the 95% tolerance intervals (dashed lines), indicating that the adopted covariance model was compatible with the data.

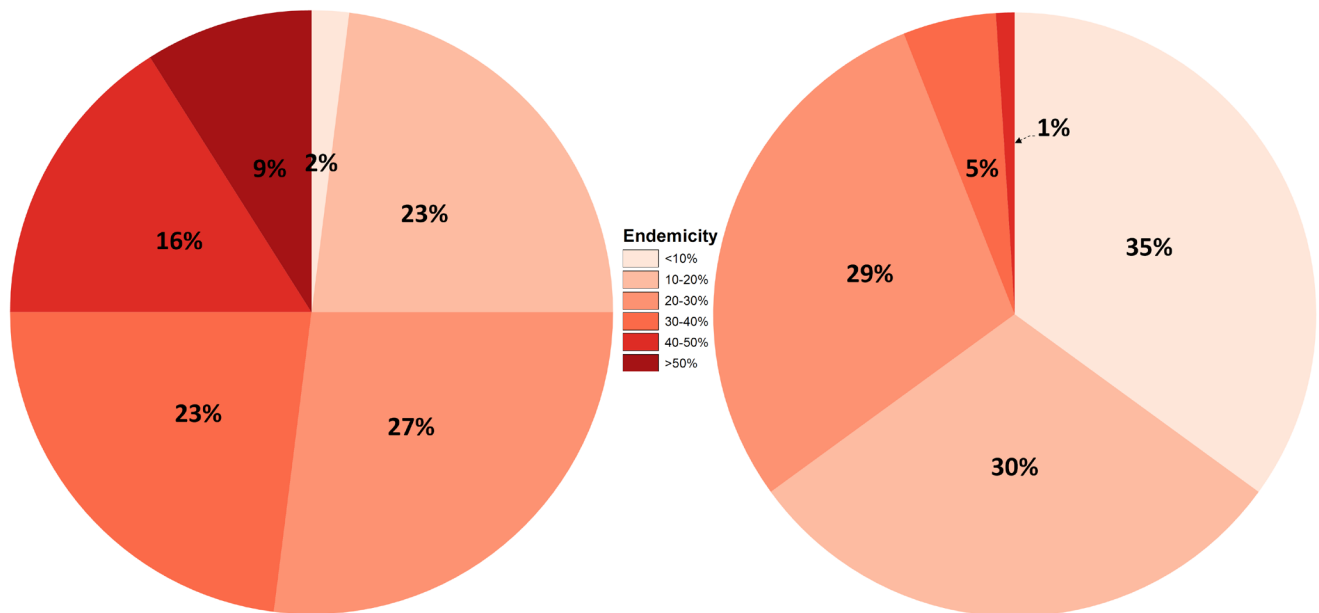


Figure 5. Percentage population of people living under different endemicity classes in 2010 (left) and 2017 (right). The mean community *Plasmodium falciparum* parasite rate ($PfPR_{2-10}$) have been grouped into six classes ranging from light red (<10%) to dark red (>50%).

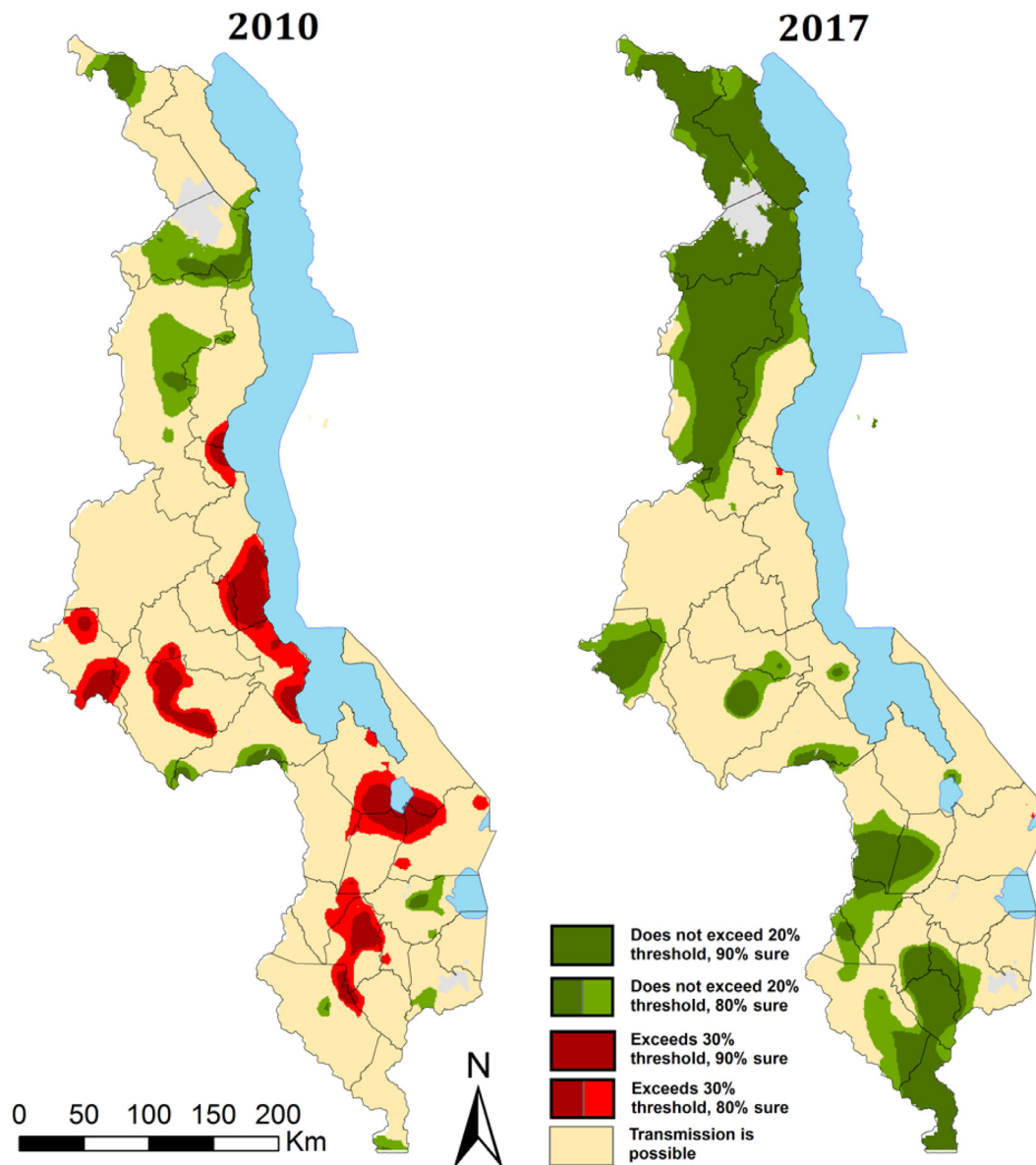


Figure 6. Non - exceedance and exceedance probabilities map. Showing areas where predicted $PfPR_{2-10}$ is less (non-exceedance probability) than 20% which were > 80% confidently predicted (light green and dark green) or > 90% confidently predicted (dark green); and areas where $PfPR_{2-10}$ is greater (exceedance probability) than 30% which were > 80% confidently predicted (light red and dark red) or > 90% confidently predicted (dark red). Areas which do not support malaria transmission are shown in grey (see Figure 1); all other areas where transmission can occur are shown in yellow.

Discussion

Earlier investigations of changing malaria prevalence [Bennett *et al.*, 2013] and malaria hospitalisation [Okiro *et al.*, 2013; Roca-Feltre *et al.*, 2012a] suggested that between 2000 and 2010 there was little evidence in support of a reduction in malaria transmission or disease burden. The launch of the 2011, five-year national malaria strategic plan, over US\$150 million in donor support has been provided for malaria control between 2010–2015 [WMR, 2018], resulting in a significant increase in vector

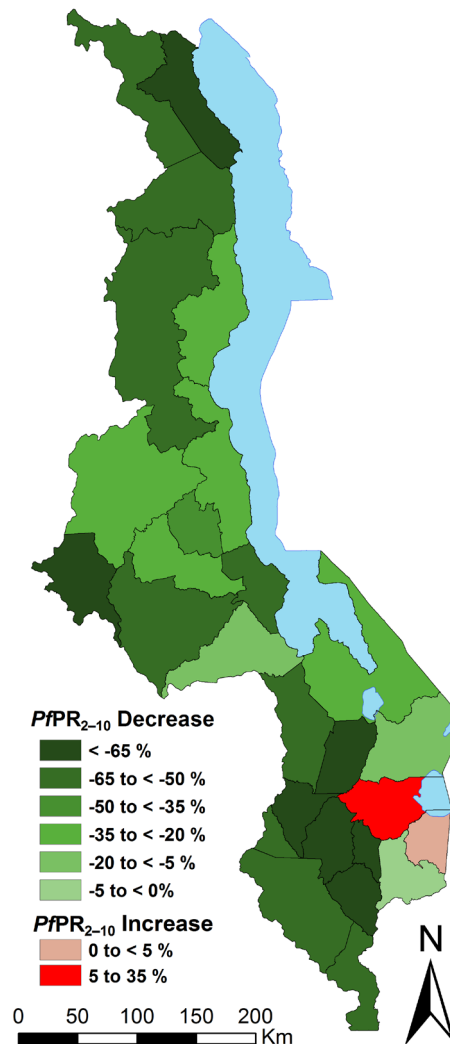
control coverage and improved malaria case-management, including expanded community-based care. This has corresponded with a dramatic decline in infection prevalence nationwide. Compared to 2010, the national mean $PfPR_{2-10}$ has declined by 47.2% (Figure 3, Figure 6 and Figure 7; Table 1). In 2017, 6% of the population lived under conditions of intense malaria transmission ($PfPR_{2-10} > 40\%$) compared to 25% in 2010 (Figure 5). It is not possible to directly attribute the reduction in malaria transmission to any specific intervention or combination of interventions.

Table 1. Predicted average $PfPR_{2-10}$ population adjusted and relative change in 27 districts between 2010 and 2017.

District	<i>PfPR</i> ₂₋₁₀ Population adjusted estimates, %			Number of surveys
	2010	2017	Per cent change (95% CI)	
Northern region*				
Chitipa	13.2	6.3	-52.7 (-59.7, -46.6)	29
Karonga	19.1	5.3	-72.1 (-74.9, -68.2)	94
Mzimba	15.1	7.1	-53.1 (-53.5, -50.8)	109
Nkhata Bay	30.7	20.3	-33.9 (-21.1, -41.3)	28
Rumphi	9.5	3.5	-63.1 (-59.8, -64.0)	176
Central region				
Dedza	20.0	18.2	-8.4(-2.5, -9.2)	45
Dowa	30.6	20.5	-33.2(-38.1, -27.5)	26
Kasungu	27.0	19.0	-29.6 (-40.6, -20.0)	35
Lilongwe	36.2	13.8	-61.9 (-69.7, -54.6)	266
Mchinji	49.4	10.4	-79.0 (-86.5, -68.8)	19
Nkhotakota	48.2	31.7	-34.3 (-48.3, -20.9)	87
Ntcheu	31.9	14.1	-55.8 (-63.2, -48.4)	22
Ntchisi	36.6	23.7	-35.2 (-47.5, -26.8)	11
Salima	48.0	19.5	-59.5 (-70.9, -48.5)	31
Southern region				
Balaka	39.9	12.8	-68.0 (-76.3, -58.5)	24
Blantyre	36.0	7.9	-78.2 (-85.2, -70.9)	277
Chikwawa	26.9	12.2	-54.8 (-71.0, -38.7)	205
Chiradzulu	33.3	11.4	-65.9 (-72.8, -59.1)	150
Machinga	35.2	31.4	-10.6 (-34.5, 7.9)	32
Mangochi	40.8	27.0	-33.8 (-43.4, -26.1)	52
Mulanje	18.2	17.6	-3.3 (-16.0, 4.9)	27
Mwanza	29.0	11.7	-59.8 (-70.0, -51.7)	51
Neno	39.9	11.3	-71.6 (-83.8, -55.6)	72
Nsanje	15.6	6.7	-56.8 (-68.7, -48.0)	15
Phalombe	19.5	19.6	0.26 (-12.2, 8.0)	158
Thyolo	19.5	6.8	-65.2 (-77.8, -52.7)	69
Zomba	18.3	24.7	34.9 (20.7, 47.2)	90
Total	29.4	15.6	-47.2 (-50.2, -44.4)	

*Predictions do not include Likoma Island, which is in northern Lake Malawi.

However, in the absence of any significant climate anomalies during this interval that might have lowered malaria risks [Future Climate for Africa, 2017], it seems plausible that the reductions seen were a result of direct intervention.

**Figure 7.** Percentage change in predicted mean $PfPR_{2-10}$ by district, between 2010 and 2017. The percentage change in mean $PfPR_{2-10}$ is shown in shades of green for decreasing malaria risk and shades of red for increasing risk.

It should, however, be highlighted that Malawi remains a highly malaria endemic country. Over 16.4% of the population live in areas where at least 1 in 4 children aged 2–10 years harbour malaria infection; the populations living in districts along the shore of Lake Malawi and central region remain under highest risk. Notably, there are few areas where extremely low, pre-elimination conditions exist ($PfPR_{2-10} < 1\%$) [Cohen *et al.*, 2010]. Consequently, the National Malaria Control Programme (NMCP) needs to sustain efforts at universal infection prevention and disease management. Unlike neighbouring countries, which exhibit a substantial sub-national diversity of malaria transmission warranting a sub-national tailoring of interventions, Malawi should

be considered as a country that should, at present, maintain a single, national strategic approach to malaria control.

The present analysis focuses on malaria infection prevalence, a frequently used malaria metric and used for over 100 years across Africa [Snow *et al.*, 2017]. Prevalence data for Malawi have been assembled from multiple sources including district-level research platforms, school surveys and nutritional surveys. The leveraging of multiple national survey data from diverse research and health constituents improves the precision of predictions over sparse data collected during single cross-sectional national malaria or health surveys powered to provide information on variables other than prevalence at low spatial resolutions. Malawi has several sentinel research districts which provide platforms to investigate specific intervention access, attribution and impact questions [Buchwald *et al.*, 2016; Escamilla *et al.*, 2017; Kabaghe *et al.*, 2018a; Kabaghe *et al.*, 2018b; Roca-Feltrer *et al.*, 2012a]; these serial, repeat observational data significantly contribute to informing the changing national profile of malaria risk in the MBG models. A key function of the NMCP remains curating and updating these data from all partners in-country. However, the analysis of sub-national variations in risk and epidemiological transitions should be triangulated with additional routine data from health information systems and malaria hospitalization. Through a process of data triangulation, a more granular understanding of the epidemiological transition is possible. The use of MBG methods to interpolate information on malaria prevalence at community levels, is less perfect when compared to complete, reliable routine data on the monthly presentation of parasitologically diagnosed fevers to health facilities. However, in the absence of complete routine data, both routine and survey data provide opportunities to understand the impact of scaled intervention on the malaria burden sub-nationally.

Conclusion

Malawi has made substantial progress in reducing the prevalence of malaria over the last seven years. It seems plausible that this transition has been a direct result of substantial investment in improving the scale and range of intervention coverage. More detailed interrogation, and triangulation, of intervention and routine data, is required to understand the sub-national impact of control. Malawi remains a high burden country. To accelerate future progress will require further prioritization of existing interventions and increasing their reach for

several years before sub-national targeting of resources becomes a priority.

Data availability

Figshare: Geostatistical analysis of Malawi's changing malaria transmission from 2010 to 2017. <https://doi.org/10.6084/m9.figshare.7856990.v2> (Chipeta *et al.*, 2019).

This project contains survey data from 2000–2017.

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

Grant information

This work was funded by the UK's Department for International Development under a project entitled *Strengthening the Use of Data for Malaria Decision Making in Africa* (DFID Programme Code 203155). MGC was supported by Wellcome Trust Career Bridging Fellowship (101113); PMM acknowledges support of the IDeALs Project (107769); RWS is funded as a Principal Wellcome Fellow (103602). DJT acknowledges the support of the Wellcome Trust to the Malawi Major Overseas Programme (206545). PMM and RWS acknowledge the support of the Wellcome Trust to the Kenya Major Overseas Programme (203077).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The following individuals were instrumental in providing assistance in identifying, and sharing unpublished survey data or have provided assistance in geo-coding of the assembled survey data: Adam Bennett, Bernard Brabin, Simon Brooker, Rachael Bronson, Marian Bruce, Job Calis, Ben Chilima, Michael G Chipeta, Isaac Chirwa, Michael Coleman, Arantxa Roca-Feltrer, Prinsen Geerligs, Timothy Holtz, Gertrude Kalanda, Lawrence Kazembe, David Laloo, Miriam Laufer, Don Mathanga, Robert McCann, Kelias Msyamboza, Themba Mzilahowa, Kamija Phiri, Natalie Roschnik, Bertha Simwaka, Rick Steketee, Terrie Taylor, Dianne J Terlouw, Lindsay Townes, Mark Wilson. The authors are grateful to the following people who supported the LINK project in Malawi: Lauren Hashiguchi, Emelda Okiro, David Kyalo, Joe Maina, Tessa Lennemann and Carrie Lynch.

References

- Bennett A, Kazembe L, Mathanga DP, *et al.*: **Mapping malaria transmission intensity in Malawi, 2000–2010.** *Am J Trop Med Hyg.* 2013; **89**(5): 840–849. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Buchwald AG, Walldorf JA, Cohee LM, *et al.*: **Bed net use among school-aged children after a universal bed net campaign in Malawi.** *Malar J.* 2016; **15**: 127. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chanda E, Mzilahowa T, Chipwanya J, *et al.*: **Preventing malaria transmission by indoor residual spraying in Malawi: grappling with the challenge of uncertain sustainability.** *Malar J.* 2015; **14**: 254. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chanda E, Mzilahowa T, Chipwanya J, *et al.*: **Scale-up of integrated malaria vector control: lessons from Malawi.** *Bull World Health Organ.* 2016; **94**(6): 475–480. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cheyabekara S, Sobti SK, Payne D: **Investigations of malaria situations in Malawi. Report on a mission 10th October 1973 to 10th December 1973.** World Health Organization unpublished document, AFR/MAL/137 15th March 1974. 1974.

- Chipeta MG, Giorgi E, Mategula D, *et al.*: **Geostatistical analysis of Malawi's changing malaria transmission from 2010 to 2017.** *figshare*. Dataset. 2019. <http://www.doi.org/10.6084/m9.figshare.7856990.v2>
- Chunga C, Kumwenda S: **Community satisfaction with indoor residual spraying for malaria control in Karonga, Northern Malawi.** *Malawi Med J*. 2014; 26(3): 71–77.
[Reference Source](#)
- Cohen JM, Moonen B, Snow RW, *et al.*: **How absolute is zero? An evaluation of historical and current definitions of malaria elimination.** *Malar J*. 2010; 9: 213.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Diggle PJ, Giorgi E: **Model-based Geostatistics for Global Public Health: Methods and Applications.** CRC/Chapman & Hall. 2019.
[Reference Source](#)
- Diggle PJ, Tawn JA, Moyeed RA: **Model-based geostatistics.** *J Roy Stat Soc C App Stat*. 1998; 47(3): 299–350.
[Publisher Full Text](#)
- Escamilla V, Alker A, Dandolo L, *et al.*: **Effects of community-level bed net coverage on malaria morbidity in Lilongwe, Malawi.** *Malar J*. 2017; 16(1): 142.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Future Climate for Africa: **Future climate projections for Malawi.** 2017; Accessed May 14th, 2019.
[Reference Source](#)
- Gething PW, Van Boeckel TP, Smith DL, *et al.*: **Modelling the global constraints of temperature on transmission of *Plasmodium falciparum* and *P. vivax*.** *Parasit Vectors*. 2011; 4: 92.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Giorgi E, Diggle PJ: **PrevMap: an R package for prevalence mapping.** *J Stat Softw*. 2017; 78(8): 29.
[Publisher Full Text](#)
- Giorgi E, Diggle P, Snow RW, *et al.*: **Geostatistical methods for disease mapping and visualisation using data from spatio-temporally referenced prevalence surveys.** *Int Stat Rev*. 2018; 86(3): 571–597.
[Publisher Full Text](#)
- Kabaghe AN, Chipeta MG, Gowelo S, *et al.*: **Fine-scale spatial and temporal variation of clinical malaria incidence and associated factors in children in rural Malawi: a longitudinal study.** *Parasit Vectors*. 2018a; 11(1): 129.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kabaghe AN, Chipeta MG, McCann RS, *et al.*: **Adaptive geostatistical sampling enables efficient identification of malaria hotspots in repeated cross-sectional surveys in rural Malawi.** *PLoS One*. 2017; 12(2): e0172266.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kabaghe AN, Chipeta MG, McCann RS, *et al.*: **Access and adequate utilization of malaria control interventions in rural Malawi: a descriptive quantitative study.** *Malar J*. 2018b; 17(1): 104.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kazembe LN, Kleinschmidt I, Holtz TH, *et al.*: **Spatial analysis and mapping of malaria risk in Malawi using point-referenced prevalence of infection data.** *Int J Health Geogr*. 2006; 5: 41.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Linard C, Gilbert M, Snow RW, *et al.*: **Population distribution, settlement patterns and accessibility across Africa in 2010.** *PLoS One*. 2012; 7(2): e31743.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lysenko AJ, Semashko IN: **Geography of malaria. A medico-geographic profile of an ancient disease.** [in Russian]; Lebedev A, editor. Moscow: Academy of Sciences USSR. 1968.
- Macharia PM, Giorgi E, Noor AM, *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(1): 340.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Malenga G, Wirima J, Kazembe P, *et al.*: **Developing national treatment policy for falciparum malaria in Africa: Malawi experience.** *Trans R Soc Trop Med Hyg*. 2009; 103 Suppl 1: S15–S18.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Mathanga DP, Campbell CH Jr, Vanden Eng J, *et al.*: **Comparison of anaemia and parasitaemia as indicators of malaria control in household and EPI-health facility surveys in Malawi.** *Malar J*. 2010; 9: 107.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mathanga DP, Walker ED, Wilson ML, *et al.*: **Malaria control in Malawi: current status and directions for the future.** *Acta Trop*. 2012; 121(3): 212–217.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- McCann RS, van den Berg H, Diggle PJ, *et al.*: **Assessment of the effect of larval source management and house improvement on malaria transmission when added to standard malaria control strategies in southern Malawi: study protocol for a cluster-randomised controlled trial.** *BMC Infect Dis*. 2017; 17(1): 639.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Metselaar D, van Thiel PH: **Classification of malaria.** *Trop Geogr Med*. 1959; 11: 157–161.
[Reference Source](#)
- Ministry of Health: **Malaria Strategic Plan 2005-2010.** Lilongwe, Malawi: Malawi Ministry of Health. 2005.
[Reference Source](#)
- Ministry of Health: **Malaria Strategic Plan: 2011—2015: Towards universal access.** National Malaria Control Programme, Ministry of Health, Government of Malawi, Lilongwe, Malawi. 2011.
[Reference Source](#)
- Mzilahowa T, Chiumia M, Mbewe RB, *et al.*: **Increasing insecticide resistance in *Anopheles funestus* and *Anopheles arabiensis* in Malawi, 2011-2015.** *Malar J*. 2016; 15(1): 563.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- National Malaria Control Programme: **Malawi Roll Back Malaria five-year strategic plan 2001-2005.** National Malaria Control Programme, Ministry of Health & Population, Lilongwe, 2001.
- National Malaria Control Programme: **Malaria Strategic Plan 2005 - 2010: Scaling up malaria control interventions.** National Malaria Control Programme, Ministry of Health & Population, Lilongwe. 2005.
[Reference Source](#)
- National Malaria Control Programme: **Malawi Malaria Program Performance Review.** National Malaria Control Programme, Ministry of Health & Population, Lilongwe, 2010.
[Reference Source](#)
- National Malaria Control Programme: **Malaria Strategic Plan 2011 - 2015: Towards universal access.** National Malaria Control Programme, Ministry of Health & Population, Lilongwe. 2011.
[Reference Source](#)
- National Malaria Control Programme: **Malawi National Malaria Indicator Survey 2017.** National Malaria Control Programme Ministry of Health, Lilongwe. 2017.
[Reference Source](#)
- National Malaria Control Programme: **Report on the implementation of Indoor Residual Spraying campaign in Malawi.** Ministry of Health, Government of Malawi, Lilongwe, Malawi. 2012.
- National Statistical Office: **2018 Malawi Population and Housing Census. Preliminary report.** Zomba. 2018.
[Reference Source](#)
- Okiro EA, Kazembe LN, Kabaria CW, *et al.*: **Childhood malaria admission rates to four hospitals in Malawi between 2000 and 2010.** *PLoS One*. 2013; 8(4): e62214.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Omumbo JA, Noor AM, Fall IS, *et al.*: **How well are malaria maps used to design and finance malaria control in Africa?** *PLoS One*. 2013; 8(1): e53198.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Phiri TB, Kaunda-Khangamwa BN, Bauleni A, *et al.*: **Feasibility, acceptability and impact of integrating malaria rapid diagnostic tests and pre-referral rectal artesunate into the integrated community case management programme. A pilot study in Mchinji district, Malawi.** *Malar J*. 2016; 15: 177.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- President's Malaria Initiative: **Evaluation of the impact of Malaria control interventions on all-cause mortality in children under five years of age in Malawi.** Malawi Malaria Impact Evaluation Group. Ministry of Health, Lilongwe, Malawi. 2016.
[Reference Source](#)
- R Core Team: **R: A Language and Environment for Statistical Computing.** Vienna, Austria. 2017; Accessed February, 15th 2019.
[Reference Source](#)
- Roca-Feltre A, Kwizombe CJ, Sanjaquin MA, *et al.*: **Lack of decline in childhood malaria, Malawi, 2001-2010.** *Emerg Infect Dis*. 2012a; 18(2): 272–278.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Roca-Feltre A, Lalloo DG, Phiri K, *et al.*: **Rolling Malaria Indicator Surveys (rMIS): a potential district-level malaria monitoring and evaluation (M&E) tool for program managers.** *Am J Trop Med Hyg*. 2012b; 86(1): 96–98.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Snow RW: **The prevalence of *Plasmodium falciparum* in sub Saharan Africa since 1900.** Harvard Dataverse, V1, UNF:6:HTeB0mwkXpFKpfEUtXM0tg== [fileUNF]. 2017.
<http://www.doi.org/10.7910/DVN/Z29FR0>
- Snow RW, Marsh K, le Sueur D: **The need for maps of transmission intensity to guide malaria control in Africa.** *Parasitol Today*. 1996; 12(12): 455–457.
[Publisher Full Text](#)
- Snow RW, Noor AM: **Malaria risk mapping in Africa: The historical context to the Information for Malaria (INFORM) project.** Working Paper in support of the INFORM Project funded by the Department for International Development and the Wellcome Trust, Nairobi, Kenya June 2015. 2015.
[Reference Source](#)
- Snow RW, Sartorius B, Kyalo D, *et al.*: **The prevalence of *Plasmodium falciparum* in sub-Saharan Africa since 1900.** *Nature*. 2017; 550(7677): 515–518.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Stevens FR, Gaughan AE, Linard C, *et al.*: **Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data.** *PLoS One*. 2015; 10(2): e0107042.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vincent K, Dougill AJ, Mkwambisi DD, *et al.*: **Analysis of Existing Weather and Climate Information for Malawi.** Lilongwe. 2014.
[Reference Source](#)
- Walldorf JA, Cohee LM, Coalson JE, *et al.*: **School-Age Children Are a Reservoir of Malaria Infection in Malawi.** *PLoS One*. 2015; 10(7): e0134061.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Wondji CS, Coleman M, Kleinschmidt I, *et al.*: **Impact of pyrethroid resistance on operational malaria control in Malawi**. *Proc Natl Acad Sci U S A*. 2012; **109**(47): 19063–19070.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

World Bank: **The World Bank Data Bank: Malawi**. 2018; Accessed January 14th, 2019.
[Reference Source](#)

World Health Organization: **Malaria programme reviews: a manual for reviewing the performance of malaria control and elimination programmes**. Geneva,

World Health Organization 2010. 2010; Accessed February 15th, 2019.

[Reference Source](#)

World Health Organization (WHO): **Global technical strategy for malaria 2016-2030**. 2015; Accessed February 15th, 2019.

[Reference Source](#)

World Health Organization (WHO): **World malaria report 2018**. 2018; Accessed February 15th, 2019.

[Reference Source](#)

Open Peer Review

Current Peer Review Status:   

Version 2

Reviewer Report 29 July 2019

<https://doi.org/10.21956/wellcomeopenres.16757.r35914>

© 2019 Smith T. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Theresa Smith 

Department of Mathematical Sciences, University of Bath, Bath, UK

The authors have addressed my concerns.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 26 July 2019

<https://doi.org/10.21956/wellcomeopenres.16757.r35913>

© 2019 Panciera R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Rocco Panciera

United Nation Children's Fund (UNICEF), New York, NY, USA

The authors have addressed my concerns satisfactorily.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Geographic Information Systems for Health.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 June 2019

<https://doi.org/10.21956/wellcomeopenres.16576.r35579>

© 2019 Smith T. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Theresa Smith**

Department of Mathematical Sciences, University of Bath, Bath, UK

In this paper, Chipeta and co-authors use model based geostatistics to analyse malaria prevalence data collected in a large number of surveys throughout Malawi and across several years. There is a particular emphasis on estimating changes in prevalence from 2010 to 2017 during which several malaria control strategies were implemented. Using a spatio-temporal model taking into account geographic and time correlations between surveys as well as unstructured heterogeneity, the authors produce estimates of *Plasmodium falciparum* prevalence in children aged 2-10 (PfPR₂₋₁₀) at a spatial resolution of 1 km x 1 km in 2010 to 2017 along with regional estimates in the change in prevalence for 27 districts.

Overall, this is a well-written, technically sound paper. I have some small questions and suggestions:

Are sufficient details of methods and analysis provided to allow replication by others?/**Are all the source data underlying the results available to ensure full reproducibility?**

I could not find the source data in the links the authors provided. For example, the figshare reference (Chipeta *et al.*, 2019) does not contain the minimum and maximum age or the month of the survey.

Could the authors provide R code for their models and data preparation? In particular, the population re-adjustment and age standardisation are not described in enough details to allow replication by others.

Are the conclusions drawn adequately supported by the results?

Some of the conclusions in the discussion do not seem fully supported by the statistical results presented. The authors state, on page 9, 'the populations living in districts along the shore of Lake Malawi and central region remain under highest risk.' It seems to me (a statistician with limited expertise in the application) that Phalombe and Zomba would be at highest risk because these are the two regions with potentially increasing prevalence (especially Zomba). Perhaps the authors could elaborate on how they determined 'highest risk'?

On the same page, the authors recommend, 'Malawi should be considered as a country that should, at present, maintain a single, national strategic approach to malaria control'. How is this recommendation supported by the modelling?

Some smaller questions

Are the underlying survey data based on random samples from each location? Are there potential design factors or biases that should be taken into account (as in Giorgi *et al.*, 2014¹)?

Is month used in the temporal part of the model or year of the survey used? Only year is in the shared data.

Fig 3: How much uncertainty is there in these estimates?

Fig 5: How much uncertainty is there in the classification? A different figure type (perhaps a stacked bar chart) would help emphasize the decrease in prevalence between the two years.

Table 1: Could the authors add the number of surveys per district as a column in this table?

References

1. Giorgi E, Sesay S, Terlouw D, Diggle P: Combining data from multiple spatially referenced prevalence surveys using generalized linear geostatistical models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2015; **178** (2): 445-464 [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Jun 2019

Michael Give Chipeta, Malawi-Liverpool Wellcome Trust Research Programme, Blantyre, Malawi

We thank Dr Theresa Smith for thoughtful and helpful comments made on the first version. We have improved the clarity of the paper by responding to the points raised as follows:

Summary

In this paper, Chipeta and co-authors use model based geostatistics to analyse malaria prevalence data collected in a large number of surveys throughout Malawi and across several years. There is a particular emphasis on estimating changes in prevalence from 2010 to 2017 during which several malaria control strategies were implemented. Using a spatio-temporal model taking into account geographic and time correlations between surveys as well as unstructured heterogeneity, the authors produce estimates of *Plasmodium falciparum* prevalence in children aged 2-10 (PfPR₂₋₁₀) at a spatial resolution of 1km x 1km in 2010 to 2017 along with regional estimates in the change in prevalence for 27 districts.

Overall, this is a well-written, technically sound paper. I have some small questions and suggestions:

1. I could not find the source data in the links the authors provided. For example, the figshare reference (Chipeta *et al.*, 2019) does not contain the minimum and maximum age or the month of the survey.

Response: The data contains the year of the survey only as described in the data section of the manuscript. We have now updated the dataset on figshare repository ([here](#)), that contains minimum and maximum ages. We have also added this link in the updated manuscript.

2. Could the authors provide R code for their models and data preparation? In particular, the population re-adjustment and age standardisation are not described in enough details to allow replication by others.

Response: The assembled surveys contained a variety of sampled age groups and it is necessary to have a single age-group in comparing data in time and space. The theory behind changing age profiles of infection prevalence depending on the intensity of transmission in a given location is provided in the reference by Smith *et al.* (2007). To standardise age to a single age range of 2 – 10 years (PfPR₂₋₁₀), we incorporate this process in the model by setting minimum (*mA*) and maximum (*MA*) ages to 2 and 10 respectively, in Equation (1) in the manuscript. Analysis was carried out in PreVMap package in R, the code is available upon reasonable request to the first author (MGC) (this is being prepared for online hosting). Population adjustment (to account for where people live) was done in ArcMap 10.4, using WorldPop data. We provide the details in the manuscript on the links available [here](#) and [here](#).

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

3. Some of the conclusions in the discussion do not seem fully supported by the statistical results presented. The authors state, on page 9, 'the populations living in districts along the shore of Lake Malawi and central region remain under highest risk.' It seems to me (a statistician with limited expertise in the application) that Phalombe and Zomba would be at highest risk because these are the two regions with potentially increasing prevalence (especially Zomba). Perhaps the authors could elaborate on how they determined 'highest risk'?

Response: Highest risk is a factor of the change between the two timepoints and the current risk (in 2017) We agree with the reviewer, Phalombe and Zomba districts have the high prevalence, in addition, the transmission risk along the lake shore and central region remains high because of the enabling environment for the vectors. At the same time, the districts along the lake shore and central region have experienced least changes despite the interventions and control efforts.

4. On the same page, the authors recommend, 'Malawi should be considered as a country that

should, at present, maintain a single, national strategic approach to malaria control'. How is this recommendation supported by the modelling?

Response: This is based on WHO recommendations on subnational stratification for malaria control. Malawi remains a meso-endemic country, with very few with extremely low pre-elimination settings (< 1% of the population). This in our view, still calls for universal coverage control effort. Targeted interventions should be applied in pre-elimination and highly heterogeneous areas with substantial areas covering high and others covering low risk.

5. Are the underlying survey data based on random samples from each location? Are there potential design factors or biases that should to be taken into account (as in Giorgi *et al.*, 2014¹)?

Response: The underlying surveys are based on a collection of multiple random surveys. For national household surveys, these are provided in the links to national household survey designs, mostly two-stage random surveys. For other published works, these are often provided in detail in the peer reviewed publications. Overall, our only requirement was that a community or a school was sampled as a single cross-section and that information was available on the date, number examined and numbers positive. It is beyond the scope of these large data assembly approaches, leveraging all possible national survey data, to build a detailed sampling strategy per data point. This is possible when ONLY using multi-stage, household sample surveys. And some investigators only use these for mapping work, building in survey weights. However, it is our view that for NMCPs, these are often limited in the information they can provide on malaria prevalence and ignore the vast amounts of other unpublished works that exist within a country. Here we have levered all possible data from all possible sources to be able to provide the most complete source of data on malaria infection prevalence in Malawi.

6. Is month used in the temporal part of the model or year of the survey used? Only year is in the shared data.

Response: Only year of the survey is used in the modelling of the data.

7. Fig 3: How much uncertainty is there in these estimates?

Response: Uncertainty in the estimates was quantified using the standard errors. The standard errors ranged from 0.0005 to 0.3 across space and time and were largely influenced with availability of data across space and time. Areas with more data were seen to have low standard errors and narrow Confidence Intervals. Standard error maps are available upon request to the first author (MGC).

8. Fig 5: How much uncertainty is there in the classification? A different figure type (perhaps a stacked bar chart) would help emphasize the decrease in prevalence between the two years.

Response: To emphasise the decrease in prevalence between 2010 and 2017, we now provide a new Figure for exceedance and non-exceedance probabilities maps, showing areas where predicted PfPR₂₋₁₀ is less (non-exceedance probability) than 20% which were > 80% confidently predicted or > 90% confidently predicted; and areas where PfPR₂₋₁₀ is greater (exceedance probability) than 30% which were > 80% confidently predicted or > 90% confidently predicted, for both 2010 and 2017.

9. Table 1: Could the authors add the number of surveys per district as a column in this table?

Response: We have added this information in Table 1 as suggested.

Once again, we thank the reviewer and offer the improved manuscript based on this and two other reviews.

Competing Interests: No competing interests were disclosed.

Reviewer Report 21 May 2019

<https://doi.org/10.21956/wellcomeopenres.16576.r35159>

© 2019 Chirwa T. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Tobias Chirwa

Division of Epidemiology and Biostatistics, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

Summary

I have read the manuscript report and, subject to addressing my comments, should be considered for indexing. The paper looks at prevalence of malaria in space and time, taking into account multiple data sources to counter sparse data. The setting is Malawi and the authors utilize various data sources to estimate prevalence in 2010 and 2017. They argue that this change can be attributed to interventions implemented during the period. Generally, it is an important paper showing decline in prevalence over time.

General comment:

The authors need to justify why they only considered changes or trends based on two time points (i.e. 2010 and 2017). They had an opportunity to optimally use the different multiple sources over time. One would have assumed the estimation would be enriched with more time points considering surveys were also conducted in the intermediate period. Can the method provide estimates for several time points? For the two time points, I therefore feel the authors should be cautious of use of statistical significance.

Specific comments:

Abstract:

Under results, it is not clear whether the 1834 communities are repeat communities and whether this was taken into account during analysis i.e. weighting.

On declining prevalence, it is not clear why the authors say the decline was significant, and how this was determined. Was this by district, 2017 vs 2010?

Introduction:

In paragraph 5, the authors should clarify whether the 1057 surveys are in addition to the 73 mentioned in the previous paragraph.

Method:

The authors should give an exclusion and inclusion criteria. At the moment, it is not clear why Likoma District was excluded, although distance is mentioned.

On page 4 – the authors show that there were different interventions in selected districts. Did these overlap and how were they handled in the analysis? Were the mass campaigns a national effort? And

how do we explain the differences in prevalence decline?

Under section on malaria prevalence the dates are from 2000 to 2017 while the title states 2010 to 2017 – this has been repeated elsewhere. Can the authors also explain how the age standardization was done?

Under the geostatistical spatio-temporal analysis, can the authors explain the model and how this fits in with sparse data scenario considering national surveys were used? What would also be useful is to be able to compute inter-cross sectional survey estimates besides 2010 and 2017 since we have direct estimates at the time of cross-sectional surveys. I am looking for added value of the model than classical analysis.

On Page 5 (population-adjusted risk), can the authors explain whether they adjusted for repeat surveys?

Results:

There are varying case detection methods from survey to survey. How were these used to estimate prevalence. Sensitivity, specificity issues? Under Table 1, I wish there were estimates for intermediate years. I think the authors have not optimally utilized the data. A trend analysis would be useful as well. Can the authors explain the small changes in some districts; and also why there is such a wide variation in the changes by district?

Discussion:

Can the authors discuss the design variation for the different multiple sources of data; One of the references has same title as the manuscript.

Do we have any reference to climatic changes within the period of analysis?

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Biostatistics and Epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Jun 2019

Michael Give Chipeta, Malawi-Liverpool Wellcome Trust Research Programme, Blantyre, Malawi

We thank Prof Tobias Chirwa for thoughtful and helpful comments made on the first version. We have improved the clarity of the paper by responding to the points raised as follows:

Summary

I have read the manuscript report and, subject to addressing my comments, should be considered for indexing. The paper looks at prevalence of malaria in space and time, taking into account multiple data sources to counter sparse data. The setting is Malawi and the authors utilize various data sources to estimate prevalence in 2010 and 2017. They argue that this change can be attributed to interventions implemented during the period. Generally, it is an important paper showing decline in prevalence over time.

A) General comment:

The authors need to justify why they only considered changes or trends based on two-time points (i.e. 2010 and 2017). They had an opportunity to optimally use the different multiple sources over time. One would have assumed the estimation would be enriched with more time points considering surveys were also conducted in the intermediate period. Can the method provide estimates for several time points? For the two time points, I therefore feel the authors should be cautious of use of statistical significance.

Response: We used all the available malaria survey data in Malawi, which included national surveys, subnational surveys, published and unpublished sources between 2000 and 2017 to make predictions to two-time points 2010 and 2017. Changes in prevalence between 2000 and 2010, using data assembled up to 2010, have been presented elsewhere by Bennet *et al.* (2013). The current work updated data through to 2017. Our aim was to explore changes in malaria prevalence in 2017 compared to 2010. It is always tempting to estimate malaria every year, but we argue that predictions should be made to years of maximal data, and years that are important for national malaria programme planning. We used data pre-2010 simply to improve the 2010 prediction.

We have taken note and revised accordingly to replace the word “significance” with “substantial” to better present our findings on the changes between estimates in 2010 and 2017. Additionally, we have added a map of exceedance and non-exceedance probabilities to quantify with 80% or 90% certainty that a location is above or below a given prevalence threshold.

Bennett A, Kazembe L, Mathanga DP, Kinyoki D, Ali D, Snow RW, Noor AM (2013). Mapping malaria transmission intensity in Malawi, 2000-2010. *American Journal of Tropical Medicine & Hygiene*, 89: 840–849

B) Specific comments:Abstract:

1. Under results, it is not clear whether the 1834 communities are repeat communities and whether this was taken into account during analysis i.e. weighting.

Response: The 1834 are unique communities, among the 2237 communities. That's 403 (18%) were repeat survey communities, taken at the same geo-location but in different years.

Response: Weighting was directly taken into account by the spatio-temporal correlation structure as defined in equation (2). More specifically, locations that are repeatedly sampled will have a spatial correlation of 1 (first factor on the right-hand side of equation (2)) but the model will still allow for variation over time as resulting from the temporal correlation (second factor on the right-hand side of the equation (2)) and from the unstructured random effects in equation (1).

2. On declining prevalence, it is not clear why the authors say the decline was significant, and how this was determined. Was this by district, 2017 vs 2010?

Response: The prevalence decline was determined by the percentage change per district, from Table 1. We have reworded "significance" to "substantial", indicating the change between prevalence in 2010 and 2017.

Introduction:

3. In paragraph 5, the authors should clarify whether the 1057 surveys are in addition to the 73 mentioned in the previous paragraph.

Response: The 1057 surveys include the 73 surveys mentioned in the previous paragraph.

4. The authors should give an exclusion and inclusion criteria. At the moment, it is not clear why Likoma District was excluded, although distance is mentioned.

Response: The model-based geostatistical methodology used for analysis in the current paper models prevalence by borrowing strength of information across space and time. Considering that the islands are at a distance of more than 60 kilometres from the mainland, it was deemed necessary to exclude them from the analysis. Spatial dependence of information tends to diminish with increasing separation distance and this island setting may have a different ecology, by its island status, to the contiguous areas of mainland Malawi.

5. On page 4 – the authors show that there were different interventions in selected districts. Did these overlap and how were they handled in the analysis? Were the mass campaigns a national effort? And how do we explain the differences in prevalence decline?

Response: Indeed, the interventions were overlapping, including insecticide treated-bed nets and indoor residual spraying. Net distribution campaigns were a nation-wide effort for years 2008, 2011, 2012, 2014 and 2016. IRS, however, was applied in a select number of districts as outlined in the paper. We intentionally did not include interventions as a variable in the analysis. The aim was to describe parasite prevalence without adjusting for, or testing for, interventions or other variables that might influence the prevalence of infection. We feel that climate, intervention, ecological changes would be reflected in the parasite rate in a given year. Therefore, the spatial predictions in 2010 and 2017 reflect the effects of intervention at these times. We have equally avoided specific statistical testing of intervention. Rather we have taken a plausibility approach [Habbict *et al.*, 1999] and used in previous works looking at changing malaria infection prevalence in Africa [Snow *et al.*, 2017; Macharia *et al.*, 2018; Giorgi *et al.*, 2018]. Whether the interventions were overlapping or not, their cumulative effect is reflected in the prevalence and changes in prevalence.

Giorgi E, Osman AA, Hassan AH, Ali AA, Ibrahim F, Amran JGH, et al. Using non-exceedance probabilities of—relevant malaria prevalence thresholds to identify areas of low transmission in Somalia. *Malar J.* 2018;17:88.

Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol.* 1999;28:10–8.

Macharia PM, Giorgi E, Noor AM, Waqo E, Kiptui R, Okiro EA, Snow RW (2018). Spatio-temporal analysis of *Plasmodium falciparum* prevalence to understand the past and chart the future of malaria control in Kenya. *Malaria Journal*, 17: 340.

Snow RW, Sartorius B, Kyalo D, Maina J, Amratia P, Mundia CW, Bejon B, Noor AM (2017). The prevalence of *Plasmodium falciparum* in sub Saharan Africa since 1900. *Nature*, 550: 515-518.

6. Under section on malaria prevalence the dates are from 2000 to 2017 while the title states 2010 to 2017 – this has been repeated elsewhere. Can the authors also explain how the age standardization was done?

Response: The estimated prevalence was at two-time points (2010 and 2017). However, we used data pre-2010 (between 2000 and 2010) to provide stable estimates in 2010 by leveraging on the temporal autocorrelation.

Response: The assembled surveys contained a variety of sampled age groups and it is necessary to have a single age-group in comparing data in time and space. The theory behind changing age profiles of infection prevalence depending on the intensity of transmission in a given location is provided in the reference by Smith *et al.* (2007). To standardise age to a single age range of 2 – 10 years (PPR_{2-10}), we incorporate this process in the model by setting minimum (mA) and maximum (MA) ages to 2 and 10 respectively, in Equation (1).

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

7. Under the geostatistical spatio-temporal analysis, can the authors explain the model and how this fits in with sparse data scenario considering national surveys were used? What would also be useful is to be able to compute inter-cross sectional survey estimates besides 2010 and 2017 since we have direct estimates at the time of cross-sectional surveys. I am looking for added value of the model than classical analysis.

Response: The model implemented is explained under this said section. A binomial model was implemented to model probability of having a positive outcome (infected) (Equation 1). The model had a stationary and isotropic Gaussian process, to account for the spatio-temporal random effects modelled with an exponential correlation function (equation 2 on the manuscript). Is the Gaussian noise/unexplained variation within communities or the unstructured random effects. The parameters were estimated using Monte Carlo maximum likelihood. All the data were considered as an individual survey point surveyed at a certain location and specific time irrespective of whether it was a cross section national survey or several data points from a village. Our rationale for selecting two time-periods is discussed earlier.

8. On Page 5 (population-adjusted risk), can the authors explain whether they adjusted for repeat surveys?

Response: Population adjusted risk in this case means adjusting for uneven distribution of population when calculating malaria risk. The use of repeat surveys in the same location is discussed above.

Results:

9. There are varying case detection methods from survey to survey. How were these used to estimate prevalence. Sensitivity, specificity issues?

Response: As detailed in the description of survey data: Most surveys employed microscopy (83%), rather than RDTs (17%) for parasite detection. The only gold standard is the PCR, however, they are too few to model and compare. Standardizing between microscopy and RDTs is fraught with difficulties because there are often very few details per survey, where both used, to describe the quality of microscopy (e.g. quality of staining, slide storage, how many high-power fields examined, and magnification of microscopy) nor any reliable algorithm to standardize between different RDTs. Where surveys have compared high quality slide reading and RDTs, for example in Kenyan school children, there has always been a close correlation between methods [Gitonga *et al.*, 2012]. During previous work by the Malaria Atlas Project, supplementary information to data use described results obtained by RDTs and Microscopy as equivalent for the purposes of risk mapping [Gething *et al.*, 2011]. For these reasons we did not attempt to standardize between detection methods. Future work, might provide novel ways of standardising between diagnostic methods, however at present these reliable algorithms are not available.

Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, et al. A new world malaria map: *Plasmodium falciparum* endemicity in 2010. *Malaria journal*. 2011;10:378.

Gitonga CW, Kihara JH, Njenga SM, Awundo K, Noor AM, Snow RW et al. Use of rapid diagnostic tests in malaria school surveys in Kenya: does under-performance matter for planning malaria control? *Am J Trop Med Hyg*. 2012; 87:1004-11.

10. Under Table 1, I wish there were estimates for intermediate years. I think the authors have not optimally utilized the data. A trend analysis would be useful as well.

Response: The trend between 2010 and 2017 is unlikely to have been linear and would also be hampered by large confidence intervals due to few data points by year. Notwithstanding that, we aimed to provide only the overall difference between 2010 and 2017 in the current work. The predictions are available between the principle prediction years and the data can be obtained for those interested from the primary author (MGC).

11. Can the authors explain the small changes in some districts; and also, why there is such a wide variation in the changes by district?

Response: The changes observed by district are a function of covariates such as interventions, urbanization, land use etc. We did not assemble all possible covariates by district to allow direct comparison by district in a plausibility framework due to data limitations. The wide variation is likely due to sparse data and prevalence within a district spanning different endemicity levels (heterogeneity within districts) and this will be reflected in the Confidence Bounds of each prediction.

Discussion:

12. Can the authors discuss the design variation for the different multiple sources of data;

Response: For national household surveys, these are provided in the links to national household survey designs. For other published works, these are often provided in detail in the peer reviewed publications. This often lead to contacts with the authors for finer spatial and temporal level data not provided in the publication. Other data has been provided without the context of a publication by malariologists working in Malawi, and all acknowledged in the paper. Overall, our only

requirement was that a community or a school was sampled as a single cross-section and that information was available on the date, number examined and numbers positive. It is beyond the scope of these large data assembly approaches, leveraging all possible national survey data, to build a detailed sampling strategy per data point. This is possible when ONLY using multi-stage, household sample surveys. And some investigators only use these for mapping work, building in survey weights. However, it is our view that for NMCPs, these are often limited in the information they can provide on malaria prevalence and ignore the vast amounts of other unpublished works that exist within a country. Here we have levered all possible data from all possible sources to be able to provide the most complete source of data on malaria infection prevalence in Malawi.

13. One of the references has same title as the manuscript.

Response: The reference similar to the title of the manuscript references the dataset used in the analysis for future citations where the data provided will be used elsewhere.

14. Do we have any reference to climatic changes within the period of analysis?

Response: We have added a reference on the climatic changes within the study period.

Once again, we thank the reviewer and offer the improved manuscript based on this and two other reviews.

Competing Interests: No competing interests were disclosed.

Reviewer Report 02 May 2019

<https://doi.org/10.21956/wellcomeopenres.16576.r35161>

© 2019 Panciera R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Rocco Panciera

United Nation Children's Fund (UNICEF), New York, NY, USA

The article presents a geostatistical analysis of the temporal and spatial variation of malaria incidence estimated by assembling a national and multi-temporal dataset of malaria prevalence surveys in Malawi between 2010 and 2017. Geostatistical spatio-temporal models are applied to provide granular estimates of malaria prevalence standardized to age 2-10, which are then weighted by population density to estimate changes in malaria prevalence within the period and for all districts in Malawi. Such changes are attributed to Malaria-specific interventions undertaken in the same period under the National Malaria Strategy 2011–2015.

The paper is well written, scientifically sound and presents conclusions of relevance to inform the success of intervention and their variability across the country. Some comments are provided below in relation to some weaknesses in specific areas.

Are sufficient details of methods and analysis provided to allow replication by others?

- Authors should clarify whether the 10% cross-validation exercise was repeated for multiple extractions.

If applicable, is the statistical analysis and its interpretation appropriate?

- Could the author clarify whether the model produces an estimate uncertainty surface as usual for geostatistical models, and whether this was seen to vary significantly across space (and time), and how would this might affect the conclusion drawn on sub-national variation in rates of decrease in prevalence for particular districts.
- Is it possible that model performance as measured based on 10% hold-out dataset through might vary throughout the temporal, window of analysis 2010-2017?
- When discussing district-level temporal changes in prevalence, authors might want to indicate when the CI includes the null (i.e., in the case of Phalombe district, a decrease of incidence is also compatible with the data (i.e., lower CI of -12%)).
- Authors should comment on how the MAE of 15% might impact the district level changes observed in Table 1.

Are the conclusions drawn adequately supported by the results?

- The Authors claim that no significant climate anomalies were observed during the analysis window (2010-2017). However the work they cite to sustain this claim seems to refer to the previous decade (2000 to 2010). Could a more recent reference relevant to the period at hand be found?
- Other factors, such as socio-economic factors, education, access to improved water and sanitation, and proximate determinants including access to health services are known to affect malaria-related morbidity. The authors should comment on whether such factors might be associated to prevalence rate, and to the extent possible provide argument against their potential impact on the observed reduction

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Geographic Information Systems for Health.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 19 Jun 2019

Michael Give Chipeta, Malawi-Liverpool Wellcome Trust Research Programme, Blantyre, Malawi

We thank Dr Rocco Panciera for thoughtful and helpful comments made on the first version. We have improved the clarity of the paper by responding to the points raised as follows:

1. Authors should clarify whether the 10% cross-validation exercise was repeated for multiple extractions.

Response: The 10% cross-validation exercise was repeated for multiple extraction simulations to ensure that Monte-Carlo error was reduced as much as possible.

2. Could the author clarify whether the model produces an estimate uncertainty surface as usual for geostatistical models, and whether this was seen to vary significantly across space (and time), and how would this might affect the conclusion drawn on sub-national variation in rates of decrease in prevalence for particular districts.

Response: The model produces standard error maps, and the 2.5 and 97.5 quintiles. The standard error ranged from 0.0005 to 0.3 across space and time and was largely influenced with availability of data across space and time. Areas with more data were seen to have low standard errors and narrow CIs. Thus, for any policy relevant work, the managers would be confident in areas with more data and add more data collection initiatives where data were sparse. Additionally, we have added a map of exceedance and non-exceedance probabilities to quantify with 80% or 90% certainty that a location is above or below a given prevalence threshold.

3. Is it possible that model performance as measured based on 10% hold-out dataset through might vary throughout the temporal, window of analysis 2010-2017?

Response: While temporal variation would be expected, the amount of data by year is not enough for the exercise.

4. When discussing district-level temporal changes in prevalence, authors might want to indicate when the CI includes the null (i.e., in the case of Phalombe district, a decrease of incidence is also compatible with the data (i.e., lower CI of -12%)).

Response: We have now indicated explicitly in the discussion that changes in the CIs in Machinga, Mulanje and Phalombe districts included the null value.

5. Authors should comment on how the MAE of 15% might impact the district level changes observed in Table 1.

Response: The mean absolute error (MAE) represents the average magnitude of the errors (the absolute differences between the predictions and actual the observations). As this is an average measure, some data points with larger values in MAE may have contributed to the value obtained. However, if the interpretation is made at district level, changes ranging between 0 and 15 % would be assumed to be within the MAE bounds and thus null. The correlation was 0.72 indicating a good correspondence between observed and predicted values.

6. The Authors claim that no significant climate anomalies were observed during the analysis window (2010-2017). However, the work they cite to sustain this claim seems to refer to the previous decade (2000 to 2010). Could a more recent reference relevant to the period at hand be found?

Response: We have now updated the reference to a more recent one.

7. Other factors, such as socio-economic factors, education, access to improved water and sanitation, and proximate determinants including access to health services are known to affect malaria-related morbidity. The authors should comment on whether such factors might be associated to prevalence rate, and to the extent possible provide argument against their potential impact on the observed reduction.

Response: Our modelling framework did not use covariates (such as socio-economic, climate, land use, ecology and interventions) during the spatial-temporal modelling. This is so because we regard parasite prevalence observed at each location and time reflects the effects of all climate, land-use, urbanization and interventions at the time a survey is undertaken. In addition, if these covariates were to be used, we do not share the view that the right data exist, or the right directionality is presumed of covariate selection over-time. The inclusion of intervention covariates introduces a circularity that often leads to an underestimate of predicted PfPR, for example PfPR in 2017 in a given location is a product of the use of LLIN in that location, to include a covariate that adjusts this value according to LLIN coverage data from other sources seems inappropriate. For these combined reasons we have increasingly defaulted to using the empirical data on parasite rate without covariates, enabling us to more reliably understand the likely impacts of time-varying environmental and intervention effects at a continental (Snow *et al.*, 2017), national (Macharia *et al.*, 2018; Giorgi *et al.*, 2018) and sub-national scale (Snow *et al.*, 2015) using a qualitative plausibility framework.

Giorgi E, Osman AA, Hassan AH, Ali AA, Ibrahim F, Amran JGH, et al (2018). Using non-exceedance probabilities of—relevant malaria prevalence thresholds to identify areas of low transmission in Somalia. *Malar J.*, 17:88

Macharia PM, Giorgi E, Noor AM, Waqo E, Kiptui R, Okiro EA, Snow RW (2018). Spatio-temporal analysis of *Plasmodium falciparum* prevalence to understand the past and chart the future of malaria control in Kenya. *Malaria J.*, 17: 340

Snow RW, Sartorius B, Kyalo D, Maina J, Amratia P, Mundia CW, Bejon B, Noor AM (2017). The prevalence of *Plasmodium falciparum* in sub Saharan Africa since 1900. *Nature*, 550: 515-518

Once again, we thank the reviewer and offer the improved manuscript based on this and two other reviews.

Competing Interests: No competing interests were disclosed.