

# Chapter 13

## Bayesian Spatial-Temporal Disease Modeling with Application to Malaria



Ropo Ebenezer Ogunsakin and (Din) Ding-Geng Chen

**Abstract** *Background:* Malaria remains a major public health challenge in Nigeria. Considerable effort has been made to reduce the prevalence and impact of the disease. The National Malaria Control Programme conducted a nationally representative Malaria Indicator Survey (MIS) within the malaria peak transmission season in 2008, 2010, 2013 and 2015 which comprises of all the six region of Nigeria. In this study, the spatial and temporal modeling of malaria risk within each region of Nigeria were studied using the MIS survey data. *Methods:* This study used data obtained from the Nigeria demographic health survey (NDHS) database to assess models; data were collected in 37 states between 2008, 2010, 2013 and 2015. We examine associations between malaria risk and socio-demographic factors using 16 Bayesian Poisson spatial-temporal models that incorporate spatial and temporal autocorrelations. The optimum model selected according to the deviance information criterion and effective number of parameters in the Bayesian paradigm. The models were implemented in R-INLA package. *Results:* The model included spatially uncorrelated heterogeneity, temporally correlated random-walk autocorrelation, and spatial temporal interaction model had small deviance information criteria. This model was the best in examining the association between malaria risk and socio-demographic factors using NDHS. The relationship between malaria risk and socio-demographic factor is statistically significant. *Conclusion:* The spatial-temporal interaction was statistically meaningful and the prevalence of malaria was influenced by the time and space interaction effect. Wealth index and place of residence have influence on malaria. To further reduce malaria burden, current tools should be supplemented by socio-demographic development.

---

The original version of this chapter was revised: An appendix has been added at the end of this chapter and page numbers in the subsequent chapters were corrected. The correction to this chapter is available at [https://doi.org/10.1007/978-3-030-35260-8\\_17](https://doi.org/10.1007/978-3-030-35260-8_17)

R. E. Ogunsakin (✉)  
University of Pretoria, Pretoria, South Africa

(Din) D.-G. Chen  
School of Social Work, University of North Carolina, Chapel Hill, NC, USA  
e-mail: [dinchen@email.unc.edu](mailto:dinchen@email.unc.edu)

© Springer Nature Switzerland AG 2020, corrected publication 2020  
X. Chen, (Din) D.-G. Chen (eds.), *Statistical Methods for Global Health and Epidemiology*, ICSA Book Series in Statistics,  
[https://doi.org/10.1007/978-3-030-35260-8\\_13](https://doi.org/10.1007/978-3-030-35260-8_13)

## 13.1 Introduction

Malaria is endemic in Nigeria and remains a major public health burdens affecting the world despite the remarkable accomplishment made towards its control and prevention. Most of the burden of malaria is concentrated in Sub-Saharan Africa (SSA) (Israel et al. 2018). Estimates in 2016 affirmed that 90% and 92% of the global proportion of malaria cases and death were recorded in this region (Awuah et al. 2018; Israel et al. 2018; Odugbemi et al. 2018; World Health Organization 2015) and Nigeria accounts for about 29% of this burden. Malaria is the third leading cause of death among under five children globally and accounts for almost one out of every five deaths in under five children (Abah & Temple 2015; Israel et al. 2018; Singh, Musa, Singh, & Ebere 2014). In Nigeria, it is estimated that about 110 million clinically diagnosed cases of malaria and nearly 300,000 malaria-related childhood deaths occur each year (Israel et al. 2018; Kyu, Georgiades, Shannon, & Boyle 2013). Evidence shows that the disease contributes to about 60% of all outpatients visits, 30% of hospitalizations and 11% of maternal mortality in the country (Bennett et al. 2017; Kassegne et al. 2017).

Considerable effort has been made to reduce the prevalence and impact of the disease, however, the last decade of malaria control has witnessed increased support by government and its partners in the areas of insecticide-treated nets (ITNs), intermittent preventive treatment (IPT), indoor residual spraying (IRS), integrated programme (IVM) and environmental management (EM), long-lasting insecticidal net (LLIN) campaigns, replacement campaigns, intermittent preventive treatment (IPT), and a massive scale up in malaria case management. The National Malaria Control Programme (NMCP) in collaboration with Roll Back Malaria (RBM) also keying into these global strategies plan (2009–2013) (Kilian, Boulay, Koenker, & Lynch 2010). In 2010, more than 24 million long lasting impregnated net (LLIN) were distributed across 14 states of Nigeria through a campaign supported by the partners (Adigun, Gajere, Oresanya, & Vounatsou 2015). Preceding this time, one of the state in South-South Nigeria have received more than 600,000 LLINs between 2008 and part of 2009 through the help of United State Agency for International Development (USAID) (Kyu et al. 2013) for children under the age of five. These efforts resulted into about 425 of households having at least one ITN (Adigun et al. 2015).

In addition, more than 70 million rapid diagnostic tests (RDTs) were distributed among all the health facilities in the country between 2008 and 2010 which could be freely used in malaria diagnosis and to provide immediate treatment based on the results (World Health Organization 2015). It was further reported that 5% of malaria cases were screened with RDTs in 2008. But in 2010, the number of pregnant women who received preventive therapy during their routine antenatal care reached 13% which is an indication of low turnout for health care seeking behavior. In view of the aforementioned, the effective malaria control strategies suggest a better and comprehensive map of the spatial distribution of malaria prevalence. This can help in efficient resource allocation for planning and intervention implementation as well as the evaluation of their impact (Gemperli et al. 2006; Giardina et al. 2012;

Gosoni, Msengwa, Lengeler, & Vounatsou 2012; Hay & Snow 2006; Riedel et al. 2010). It is essential to identify the association between malaria risk and socio-demographic factors. Such a study of the identification of the socio-demographic risk factors is helpful in identifying region who have a critical need for intervention. In Nigeria, previous studies have concluded that malaria risk are associated with environmental and climatic factors (Adigun et al. 2015; World Health Organization 2017). In particular it was noted that intervention appear not to have important effect on malaria risk. Nevertheless the spatial distribution of malaria was not investigated (Adigun et al. 2015). However, the modeling of malaria risk in each of the region in Nigeria has to be explored. Meanwhile, the spatial pattern of malaria risk is known to vary, its temporal evolution has yet to be evaluated. Therefore, the objective of this study was to determine the spatial-temporal modeling of malaria risk in Nigeria taking into consideration socio-demographic factors.

In this research, we introduce Bayesian spatial-temporal modeling that incorporate spatial information in such a way that not only reflect the influences of space and time but also reflect the interaction of space time on the preferred variable of interest. In doing so, we use 16 Bayesian Poisson spatial-temporal techniques in estimating model parameters.

## 13.2 Spatial-Temporal Data in Nigeria

### 13.2.1 Study Area

Nigeria is the most populous country in the continent of Africa, which is located in the west sub region of Africa. The country is divided into 37 states grouped into six (13.6) regions and covers an area of about 923,768 km<sup>2</sup>. Nigeria has the largest population in Africa and the seventh largest in the world. The current population is estimated at 177.1 million based on an annual growth rate of 3.2% (National Population Commission [NPopC] 2016). Nigeria's population is young, with persons age 0–24 accounting for more than 62% of the country's residents (National Population Commission 2010). According to the World Bank's definition, Nigeria is a lower middle income country. The country has tropical climate with two rainfall seasons in a year (wet and dry season) which is accompanied with the movement of two dominant winds: the rain bearing south westerly winds, and the cold, dry and dusty north easterly wind generally referred to as the Harmattan. The wet season occurs from April to September, and the dry season from October to March. The annual rainfall ranges between 550 mm in some part of the north mainly in the fringes of Sahara desert to 4000 mm in the coastal region around Niger delta area in the south. The temperature in Nigeria ranges between 25 and 40 °C. The geographic location of Nigeria makes suitable climate for malaria transmission throughout the country and it is all year round in most part of the country (Adigun et al. 2015). *Plasmodium falciparum* is the most prevalent malaria parasite species in Nigeria (Mouzin et al. 2012; National Population Commission 2012). Malaria transmission intensity, and

seasonality vary among the country's five ecological strata that extend from south to north (National Population Commission 2012). Considering population density and distribution of risk areas, an estimated 3%, 67% and 30% live in very low to low, moderate, and high to very high transmission intensities area, respectively (Mouzin et al. 2012). The transmission season increases from north to south in terms of duration, in the space of 3 months in the north area bordering Chad to perennial in the most southern part (Mouzin et al. 2012).

### ***13.2.2 Country Profile***

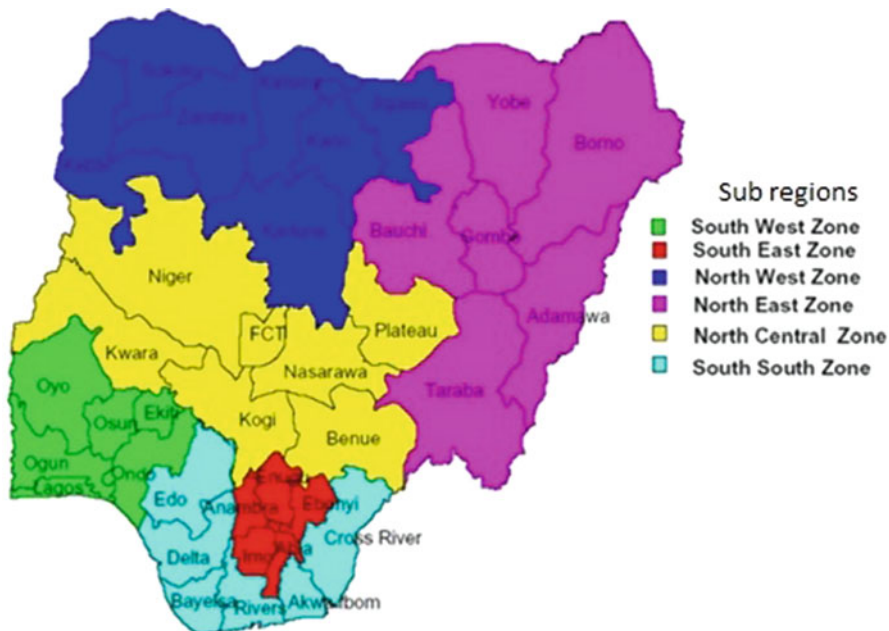
The data were collected using the standard malaria indicator questionnaires developed by the RBM and the demographic health surveillance programme. The dataset consists of information such as, demographic characteristics and socio-economic status which is on a nationally representative sample of around 6000 households from about 240 clusters. Detail description of the sampling strategies is reported in the final report of NMIS 2010 (National Population Commission 2012). The blood samples were taken from 239 clusters due to some security challenges in one of the clusters in northern part of Nigeria (National Population Commission 2012). The prevalence from two diagnostic methods: RDT and microscopy was recorded in the data (Wongsrichanalai, Barcus, Muth, Sutamihardja, & Wernsdorfer 2007). In 2015, malaria testing was done through both rapid diagnostic testing (RDT) as well as blood smear microscopy. Of the 6316 eligible children, 95% provided blood for RDT and 91% for malaria microscopy. The 2015 NMIS shows a malaria prevalence of 45% by RDT and 27% by microscopy. The geographical representation of the clusters involved and observed prevalence in the NMIS is displayed in Fig. 13.1. Figure 13.1 shows the map of Nigeria divided into various regions.

### ***13.2.3 Ethical Approval***

This study was based on the analysis of existing survey data-sets in the public domain that are available free online. The first author obtained permission for the download and usage of the NDHS dataset from [http://www.dhsprogram.com/data/dataset\\_admin/login\\_main.cfm](http://www.dhsprogram.com/data/dataset_admin/login_main.cfm).

### ***13.2.4 Predictor Variables***

The transmission of malaria is known to be influenced by several factors such as socioeconomic, demographic factors and environmental/climatic. Demographic



**Fig. 13.1** Map of Nigeria showing the 37 states with six Geo-political region

variables were captured on survey tools, which include area type of the household, age, and mother's educational level. Information on socioeconomic status was measured by a wealth index. It was calculated as a weighted sum of household assets using principal component analysis.

## 13.3 Statistical Methodology

### 13.3.1 Malaria Spatial-Temporal Modeling

Spatial-temporal disease mapping has become an important tool in passive surveillance of diseases. Understanding how disease risks and prevalence and/or incidence vary over time may provide information that may be of great epidemiological significance. Spatial-temporal models are extensions of the basic spatial models by simply including a linear or a non-parametric trend in time, time space, time covariate and time-space-covariate interactions. When using spatial-temporal data to study occurrences such as diseases, researchers are often interested in both the spatial and temporal aspects of these data. For instance, researchers might want to investigate disease location and time of diagnosis along with the disease counts. This goal could be achieved by modeling the disease counts as a Poisson process while

concurrently incorporating the space and time data with all other risk covariates. Because of the spatial-temporal autocorrelations, spatial-temporal disease data are typically modeled as multivariate with correlated observations of Poisson disease counts at a fixed spatial location that evolves over time.

In this study, our focus is on malaria data collected over a 4-year period (2–4) from 37 States in Nigeria. Suppose we let  $i$  represent the spatial location  $i = 1, \dots, K$ ,  $K (=37)$  states and  $t = 1, \dots, T (=4)$  years, the number of malaria cases,  $y_{it}$ , is modeled as a Poisson spatial-temporal model with the expected incidence rates  $E_{it}$ , and the associated risk  $\theta_{it}$ . The standard Besag-York-Mollie spatial analytic model is represented as follows:

**Data Distribution**

$$y_{it} \sim Pois(E_{it} \times \theta_{it}) \tag{13.1}$$

where  $y_i$  counts in area  $i$  are independently identically Poisson distributed and have an expectation in area  $i$  of  $E_i$ , the expected count, times  $\theta_i$ , the risk for area  $i$

**Spatial-Temporal Mixed-Effects Regression Model**

$$\log(\theta_{it}) = \beta_o + \beta_1 x_{1it} + \dots + \beta_j x_{jit} + S_i + T_t + ST_{it} \tag{13.2}$$

where  $S$  represent the random spatial term,  $T$  is the random temporal term, and  $ST$  is the random space-time interaction. Meanwhile, the fixed-effects component is  $\beta_o + \beta_1 x_{1it} + \dots + \beta_j x_{jit}$  where  $x_{1it}, \dots, x_{jit}$  are the risk factors to be modeled with the disease risk  $\theta_{it}$ . In the present study, the two covariates included is wealth index (WI), and area type (AT). Hence, the model (13.2) is simplified as

$$\log(\theta_{it}) = \beta_o + \beta_1 WI_{it} + \beta_2 AT_{it} + S_i + T_t + ST_{it} \tag{13.3}$$

From model (13.1),  $E_{it}$  represent the expected incidence rates and its values can be estimated by several approaches. The simplest overall average for the expected counts is given by:

$$E_{it} = p_{it} \times \frac{\sum_{i=1}^K \sum_{t=1}^T y_{it}}{\sum_{i=1}^k \sum_{t=1}^T p_{it}} \tag{13.4}$$

where  $p_{it}$  is the population at  $i$ th location (i.e., state) and  $t$ th time point (i.e., year) in this malaria data.

Specifically, eight models were constructed by considering the spatial effect, and the interaction between time and space (see Table 13.2). To evaluate the regional effects, the spatial-temporal model in expression (13.3) is built to include the six Nigeria region (Region) as:

$$\log(\theta_{it}) = \beta_o + \beta_1 W I_{1it} + \beta_2 A T_{1it} + \beta_4 R e g i o n_{it} + S_i + T_t + S T_{it} \quad (13.5)$$

Hence, additional of eight spatial-temporal models are included, yielding a total of 16 fitted spatial-temporal models. These eight models comprises of different combinations of spatial random effect (UH), spatially structured heterogeneity, linear time trend, identically independent distributed time variable, random walk as well as spatial-temporal interactions (see Table 13.2 for the description of those models).

### 13.4 Bayesian Spatial-Temporal Models with INLA

In this section, we introduce how Bayesian spatial-temporal model can be implemented using R-INLA. Spatio-temporal disease mapping models are a well-known tool to explain the pattern of disease counts. Model of this kind is usually formulated within a Bayesian framework (Banerjee, Carlin, & Gelfand 2004) and computationally expensive Markov Chain Monte Carlo (MCMC) are needed to obtain the respective parameter estimates. Also, in order to get a reliable estimate for a complex spatial and spatio-temporal models, a specific block-sampling algorithms have to be applied. Furthermore, Bayesian spatial-temporal disease mapping via MCMC methods involve computationally and time intensive simulations to obtain the posterior distribution for the parameters. An approximate technique for parameter estimation in latent Gaussian models was proposed by Banerjee et al. (2004). This technique uses Integrated Nested Laplace Approximation (INLA). The advantage of INLA method is that it does not use iterative computation techniques like MCMC and it returns precise parameter estimates. The posterior approximation is achieved by applying numerical integrations for fixed effects and Laplace integral approximation to the random effects (Chen, Wakefield, & Lumely 2014). Primarily, INLA is designed for latent Gaussian models, a very wide and flexible class of models like spatial and spatio-temporal models, making INLA to be used widely in a great variety of applications (Spiegelhalter, Best, Carlin, & van der Linde 2003). In addition, the deviance information criterion (DIC) is provided by INLA for Bayesian model choice. For our analysis, INLA was implemented in the R package “INLA” (R-INLA). We used R for data management and R package , maptools for reading the shapefile.

### 13.4.1 Goodness of Fit Statistics

Modeling was done in R using the R-INLA package. The model were compared using the Deviance Information Criterion (DIC) as recommended by Khana, Rossen, Hedegaard, and Warner (2018) and Spiegelhalter, Best, Carlin, and Van Der Linde (2002). The ability to fit complex multilevel models using Markov Chain Monte Carlo (MCMC) techniques presents a need for methods to compare alternative models. The standard model comparison techniques such as AIC and BIC require the specification of the number of parameters in each model. For multilevel models which contain random effects, the number of parameters is not generally obvious and as such an alternative technique of comparison is demanded. The most widely used of such alternative technique is the Deviance information Criteria (DIC) as suggested by Spiegelhalter et al. (2002). The DIC statistic is a generalization of the AIC, and is based on the posterior mean of the deviance, which is also a measure of model complexity and fit. The deviance is defined as

$$D(\theta) = -2 \log f(y|\theta).$$

since DIC is a measure of model complexity, it considers a measure of the effective number of parameters in a model, and is defined by

$$pD = \bar{D}(\theta) - (\check{\theta}).$$

where  $\bar{D}(\theta)$  is the posterior expectation of the deviance, given by

$$\bar{D}(\theta) = -2E \left[ \log f(y|\theta) | y \right].$$

and  $(\check{\theta})$  is the deviance evaluated at some estimate  $\check{\theta}$  of  $\theta$ . Therefore, we now define the deviance information criteria (DIC) by

$$DIC = \bar{D}(\theta) + pD = 2\bar{D}(\theta) - \hat{\theta}. \quad (13.6)$$

where  $\bar{D}$  is the posterior mean of the deviance that measures the goodness of fit, and  $pD$  represent the effective number of parameters in the model. In the case of the Bayesian and bootstrapping models, low values of  $\bar{D}$  imply a better fit, while small values of  $pD$  imply a parsimonious model.  $pD$  is higher for a more complex model, and DIC appears to select the correct model. The best fitting model is one with the smallest DIC, as suggested by Lesaffre and Lawson (2012) and Spiegelhalter et al. (2002). When comparing different models, how big the difference between the DIC value of the models need to be revealed so as to declare that one model is better than the other. Previous studies have shown that a difference of 3 in DIC between two models cannot be distinguished while a difference of between 3 and 7 can be weakly differentiated (Kazembe, Chirwa, Simbeye, & Namangale 2008; Spiegelhalter, Best, Carlin, & Linde 2014). For context, a DIC difference 3 to 5 is considered significant.



## 13.5 Results

To illustrate how the 16 Bayesian Poisson spatial-temporal models can be applied to real life data, we used the data on malaria risk from 37 states of Nigeria (see Fig. 13.1). In Table 13.1, we grouped the 37 states into six regions (i.e., North central, North east, North west, South east, South south and South west) to investigate regional differences. We extracted the malaria prevalence rate for the 4-year period. To obtain the malaria incidence rates, we merged these data to calculate the associated malaria incidences rates  $y_{it}$  and the expected incidence rates  $E_{it}$  to be used in Eqs. (13.1) and (13.4). As mentioned previously, malaria can be related to many risk factors. From the epidemiological perspective, malaria risk factors includes environmental/ climatic, socioeconomic status and socio-demographic and so on. The DHS database consists an extensive list of risk covariates that could be used to model the predictability of these risk factors to malaria prevalence rates; meanwhile, most of the covariates have a higher percentage of missing data (>90%). Hence, for demonstration purposes the current study utilizes wealth index, and area type as a possible covariates.

Meanwhile all the spatial-temporal data from 37 states collected for 4 years period were incorporated for a unified Bayesian spatial-temporal modeling. Table 13.2 presents series of spatial-temporal models fitted with the R-INLA package. Comparison results among different models affirmed that the DIC values of the two models with only spatial heterogeneity effect were: 1358.55 and 1358.36 respectively while the DIC values for models incorporating temporal heterogeneity were: 1338.26, 1336.49, 1339.86, and 1336.59, respectively. The last sets of two models considered assesses the spatial-temporal interaction. DIC values of the two UH random effect and convolution model with interaction term were: 1126.44 and 1125.92 respectively. Among the two interaction models, model taking the spatially temporally uncorrelated heterogeneity + UH, temporally correlated random walk autocorrelation, and spatial temporal interaction effect into consideration was the best fitting one with a smallest DIC as well as pD value. It should be acknowledged that the DIC values from the models 1–8 space-time interaction do not exhibit extreme differences. This can be attributed to all models taking the form shown in Eq. (13.3). Between the eight models fitted, model 7 and 8 has larger pD values which indicate that the two models are more complex, apparently because it incorporates a spatio-temporal interaction effect that is not part of model 1–6. Although, model 7 and 8 is weakly indistinguishable because of the differences between the DIC value is lesser than 3. Therefore, the higher complexity was beneficial as it led to lower DIC values in model 8 which indicates a better fit model to the data. Therefore, the best fitting DICs are seen with the interaction models.

With model 8 as the best fitting model, the estimated coefficients of place of residence (rural) and wealth index (poorer), (middle), (richer) and (richest) were: 0.04525, 0.01380, 0.11115, 0.000180 and 0.05793 respectively. Moreover, the estimated  $\beta$  for these socio-demographic variables were 1.04629(95% BCI: 0.905–1.209), 1.01389 (95% BCI: 0.887–1.158), 1.11797 (95% BCI: 0.966–1.292), 1.0001

**Table 13.1** List of 37 Nigeria states and associated state code

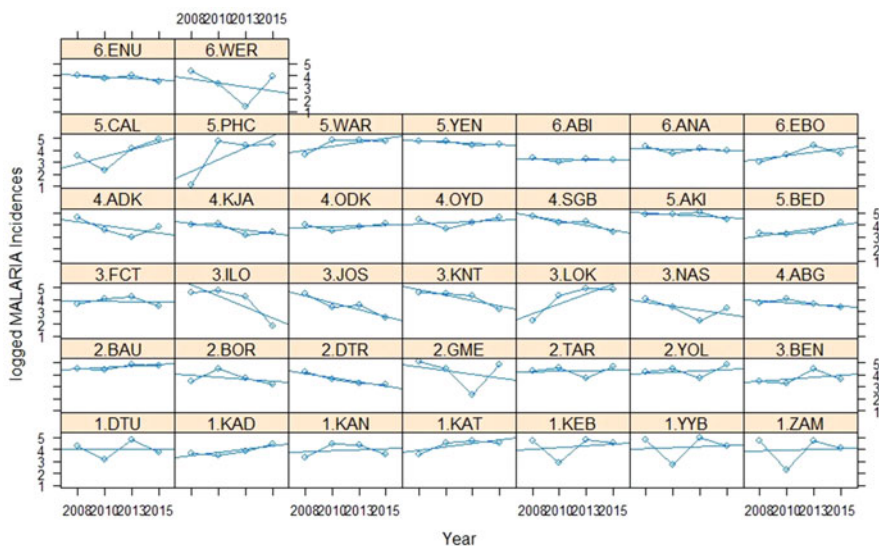
RegionNum	RegionName	StateName	StateCode	RegionCode
1	North West	Sokoto	YYB	1.YYB
1	North West	Zamfara	ZAM	1.ZAM
1	North West	Katsina	KAT	1.KAT
1	North West	Jigawa	DTI	1.DTI
1	North West	Kano	KAN	1.KAN
1	North West	Kaduna	KAD	1.KAD
1	North West	Kebbi	KEB	1.KEB
2	North East	Yobe	DTR	2.DTR
2	North East	Borno	BOR	2.BOR
2	North East	Adamawa	YOL	2.YOL
2	North East	Gombe	GME	2.GME
2	North East	Bauchi	BAU	2.BAU
2	North East	Taraba	TAR	2.TAR
3	North Central	Niger	KNT	3.KNT
3	North Central	Abuja	FCT	3.FCT
3	North Central	Nasarawa	NAS	3.NAS
3	North Central	Plateau	JOS	3.JOS
3	North Central	Benue	BEN	3.BEN
3	North Central	Kogi	LOK	3.LOK
3	North Central	Kwara	ILO	3.ILO
4	South West	Oyo	OYD	4.OYD
4	South West	Osun	SGB	4.SGB
4	South West	Ekiti	ADK	4.ADK
4	South West	Ondo	ODK	4.ODK
4	South West	Lagos	KJA	4.KJA
4	South West	Ogun	ABG	4.ABG
5	South South	Edo	BED	5.BED
5	South South	Cross River	CAL	5.CAL
5	South South	Akwa Ibom	AKI	5.AKI
5	South South	Rivers	PHC	5.PHC
5	South South	Bayelsa	YEN	5.YEN
5	South South	Delta	WAR	5.WAR
6	South East	Anambra	ANA	6.ANA
6	South East	Enugu	ENU	6.ENU
6	South East	Ebonyi	EBO	6.EBO
6	South East	Abia	ABI	6.ABI
6	South East	Imo	WER	6.WER

Note: These 37 states are grouped into 6 regions (“RegionNum”) under the region names (“RegionName”). For ease of representation in Figs. 13.1 and 13.2, we created the RegionCode abbreviation that combines the RegionNum and the StateCode

**Table 13.2** Specific spatial-temporal models and associated fit statistics

Model	Details	DIC	n.eff
1	Spatial Only (UH)	1358.55	33.633
2	Spatial Only (UH+CH)	1358.36	33.664
3	Spatial(UH+CH) +Temporal trend	1338.26	34.364
4	Spatial(UH+CH) +Temporal(UH)	1336.49	36.358
5	Spatial(UH+CH)+Temporal(CH)	1339.86	42.938
6	Spatial(UH+CH)+Temporal(UH+CH)	1336.59	36.056
7	Spatial(UH)+Temporal(CH)+ST	1126.44	109.213
8	Spatial(UH+CH)+Temporal(CH)+ST	1125.92	108.740

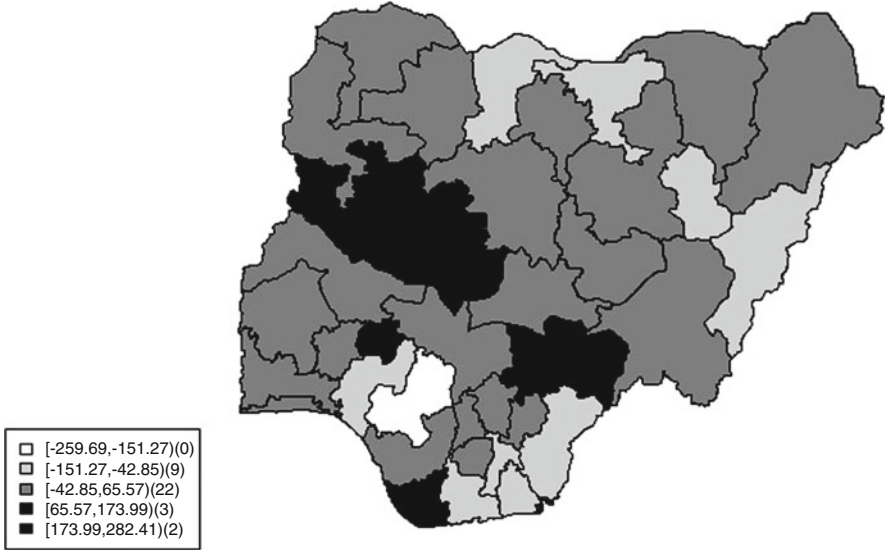
Abbreviations: Spatial—UH: uncorrelated effect model; CH: correlated effect model. Temporal—UH: uncorrelated heterogeneity, CH: random walk; spatial-temporal interaction. ST: spatial-temporal. DIC: deviance information criterion, n.eff: effective number of parameters



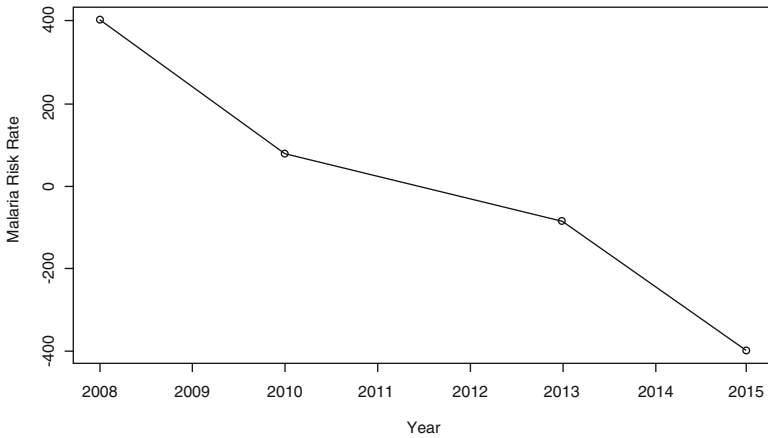
**Fig. 13.2** Temporal trends for malaria incidence rates (logged) for 37 States of Nigeria from six regions included in the analyses

(95% BCI: 0.860–1.162) and 1.0596(95% BCI: 0.897–1.250). Both the place of residence and wealth index had a positive influence on the prevalence of malaria risk. Moreover, the malaria rates as depicted in Figs. 13.1, 13.3, and 13.5 reveal some signs of spatial trends despite the fact that there are no statistically significant spatial patterns. As shown in Fig. 13.2, reported cases of malaria prevalence in Nigeria declined year by year across the 37 states over the 4-year period.

Also, the map depicted in Fig. 13.3 shows the estimated overall pattern in the spatial random-residual effects revealing spatial autocorrelation as represented by  $S_i$  in expression (13.3). The implication of the map is that all the six regions in the country have had a mix of high and low malaria prevalence over time, which is

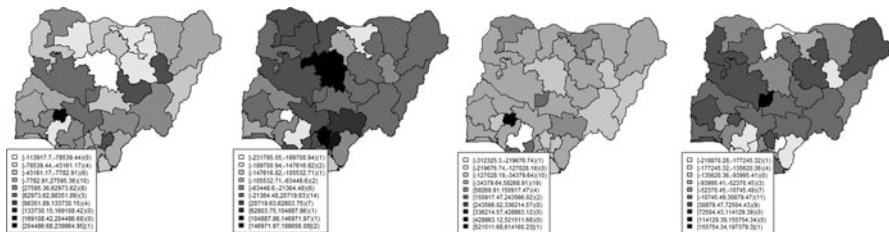


**Fig. 13.3** Spatial random-residual effects showing spatial autocorrelation as indicated by  $S_i$  in expression (13.3)



**Fig. 13.4** Temporal random-residual effects showing temporal autocorrelation as indicated by  $T_t$  in expression (13.3)

indicated by the random effects  $S_i$  and fixed effects presented in expression (13.3). Figure 13.4 depicts only the overall temporal pattern of the malaria risk prevalence as reported by  $T_t$  in expression (13.3); the map indicates that Nigeria experienced uneven risk of malaria infection without giving much knowledge about differences across the geopolitical zone of the country. Hence, both the spatial-only patterns in Fig. 13.3 and temporal-only trends in Fig. 13.4 should be interpreted simultaneously.



**Fig. 13.5** Spatial and temporal random-residual effects showing spatial and temporal autocorrelation as indicated by  $ST_{it}$  in expression (13.3)

**Table 13.3** Bayesian spatial-temporal models with regional effect

Model	Details	DIC	n.eff
9	Spatial Only (UH)	1359.45	35.221
10	Spatial Only (UH+CH)	1359.31	35.162
11	Spatial(UH+CH) +Temporal trend	1339.11	35.778
12	Spatial(UH+CH) +Temporal(UH)	1337.05	37.694
13	Spatial(UH+CH)+Temporal(CH)	1339.39	42.581
14	Spatial((UH+CH)+Temporal(UH+CH))	1337.21	37.427
15	Spatial(UH)+Temporal(CH)+ST	1126.34	112.183
16	Spatial(UH+CH)+Temporal(CH)+ST	1125.58	109.873

In addition, the interaction of spatial and temporal factors during 4 year period suggests the presence of convoluted spatial and temporal autocorrelation as indicated by  $ST_{it}$  in expression (13.3). Figure 13.5 also indicated some considerable differences in the relative risk of malaria across the six regions of Nigeria.

Moreover, in order to account for the remaining eight (8) models of our 16 Bayesian spatial-temporal model, we fitted a model accounting for the regional effects. The findings affirmed that the regional effects were statistically significant and the result is presented in Table 13.3. The results presented in Figs. 13.2, 13.3, 13.4 and 13.5 are for the first eight (8) models without the regional effect.

### 13.6 Conclusion and Summary of Findings

In this study, different models were compared for modeling and mapping of malaria risk in Nigeria. In particular, we considered series of Bayesian spatial-temporal models to examine the association or effects of socio-demographic on the malaria risk across the 37 states of Nigeria. This relationship is important to enable an effective policies as well as tools to tackle the menace of malaria transmission in Nigeria. These models were fitted to NDHS malaria prevalence data for 4-year period. Among the different spatial-temporal models examined, the model with spatial-temporal interaction fit the data well but model 8 appears better than

model 7. The findings indicate that model with spatially uncorrelated heterogeneity, temporally correlated random-walk autocorrelation, and spatial temporal interaction was the best model for goodness of fit for modeling malaria risk. The finding is similar with previous studies (Abellan, Richardson, & Best 2008; Popoff 2014) which found potential use of spatial and temporal terms in the model. Although our study did not use the Bayesian hierarchical but previous studies that used Bayesian hierarchical framework for diseases mapping as well as ecological studies of health environment association (Ehlers & Zevallos 2015; Popoff 2014) affirmed that if data are collected over space and time, spatial and temporal terms in the model becomes necessary. The reason for this could be due to the complex dependence patterns over space and over time of the occurrence of malaria deaths. The study findings indicated an estimated positive association between socio-demographic factors and malaria risk. This finding confirms previous results that showed that malaria risk is positively associated with socio-economic status (Adigun et al. 2015; Giardina et al. 2012; Gosoniu et al. 2012; Gosoniu, Veta, & Vounatsou 2010). Moreover, we observed that the overall malaria risk among the 37 states was spatially uncorrelated when viewed from a historical point for the 4 years period. Estimations from model 8 affirmed that wealth index could be an influential factor on the prevalence of malaria. Specifically, with one unit increase of wealth index (poorer), the risk of new malaria case increased by 1.0139 times. This finding is similar to what the previous findings on Bayesian geostatistical modeling of malaria from Nigeria (Adigun et al. 2015). Therefore, the results of this study provide evidence on the spatial-temporal distribution of socio-demographic risk factors in the occurrence of malaria. Hence, the utilization of socio-demographic data on malaria rapid diagnosis test (RDT), clarifies the association of these factors. From the study it was affirmed that those people living in the North Central region were found to be more at risk of malaria compared to those living in the South West.

Meanwhile, the malaria map produced in this study affirms considerable shrinkage in malaria burden in comparison to results from the first MIS survey of 2010 that showed a high burden of malaria in the entire country. There are some limitations to consider when interpreting the findings of this study. Foremost, the current study relied on malaria test results from RDT. Secondly, one can think of the limitation of the current study in line with the data used which may contain spatially correlated malaria prevalence trends across the local government or towns that are not noticeable at the state level. Hence, for future study it is advisable to perform a sensitivity analysis in case of a study utilizing Bayesian spatial temporal modeling to check whether the results vary at different geographical scales. Thus, this will help the researchers to discuss the research policy in case the results differ. Moreover, in obtaining the incidence rates when using Bayesian spatial-temporal approaches, decisions as to whether to calculate the incidence rate as population-based, geographical area-based, or combination of both should be put into consideration. In case of the current study, the incidence rate in Eq. (13.4) is obtained using the population-based. As pointed out by Lesaffre (Lawson 2013; Lesaffre and Lawson 2012), that is the most commonly used incidence rate in spatial-temporal disease mapping. An important aspect that needs to be highlighted

regarding this study is that the regional effects is statistically significant which requires the attention of the health policies makers in controlling the prevalence of malaria. Therefore, adequate effort should be targeted further in order to uncover factors that is responsible for the transmission so as to allow for the development of better malaria control measures.

## A.1 Appendix 1: R Program Codes for Analysis.

```

StateList = rep(State,each=T)
CodeList = rep(Code, each=T)
region<-rep(1:m,each=T)
region2<-region
ind2<-rep(1:(m*T))
data<-data.frame(Mal,E,year,Year,region,region2,CodeList,
  StateList,ind2,Res, Weat)
data
####1. Spatial-temporal data are pre-processed and load them in

source("dataMALARIA.R")
# check the data
print(data)
summary(data)
subregion = read.csv("subregionMALARIApaper.csv", header=T)
data = merge(data,subregion)

#calculation of the population
ff = mean(data$Mal/data$E)
data$pop = data$Mal/ff

#create a new variable for plotting since the State names are
to long
data$NewRegion = paste(data$subregion1,".",data$CodeList, sep
  = "")
# order the data by the new var
data= data[order(data$NewRegion),]
head(data)

#make table 1
d1 = data.frame(StateList=unique(data$StateList),Code=unique(
  data$CodeList),Region= unique(data$NewRegion))
d2 =merge(d1, subregion)
d3 = data.frame(RegionNum = d2$subregion1, RegionName =
  d2$subregion, RegionCode=d2$Region,StateName= d2$StateList,
  StateCode = d2$Code)
d3

```

```

###2. pre-analysis for to identify the relationship for
      modelling

library(lattice)
# incidence time series
xyplot(log(Mal)~Year|as.factor(NewRegion),type=c("b","r"), data
      , xlab="Year", ylab="logged MALARIA Incidences")

xyplot(log(Mal)~Year|as.factor(NewRegion),layout = c(7, 6),type
      =c("b","r"), data, xlab="Year", ylab="logged MALARIA
      Incidences")

### 3. Now Spatial-temporal modelling
#
# R libraries
# load necessary packages and download the map shapefile and
      therefater, read it into R  \\

library(maptools)
# get INLA
library(INLA)
inla.setOption(scale.model.default=FALSE)
require(splancs)
require(sp)
require(fields)
require(maptools)
require(abind)
library(rgdal)

## Meaning###
## The R tools maptools::readShapePoly() will read shapefiles
      into R, and spdep::poly2nb() followed by INLA::nb2INLA()
      are used to create the adjacency matrix neighbor structures
      for use with a CAR model. To map results, you can use sp::
      splot()

source("Malaria3.R")

# model 1: spatial only UH
modell = Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(
      region,model="iid")
result1 = inla(modell,family="poisson",data=data,E=E,control.
      compute=list(dic=TRUE,cpo=TRUE))
UH<-result1$summary.random$region[,2]
summary(modell)

## where:

# Mal is the disease count or outcome from your dataset.
## 1 forces an intercept onto the model.
## f() specify the spatial region and how it should be modeled.

```



```

## In the case of our study, spatial region is modeled as "iid"
  which is a random effects term. One of the advantage of
  this function is that it is useful when invoking any
  spatial model, especially the geographically weighted
  regressions.
## Also, f() functions can be added to each other in order to
  build up models.
## model1 refers to and invokes the previously defined model.
## family = specifies the likelihood.
## data = specifies the data.

## control.compute specifies options like DICand DCO.
## Note: CPO is the conditional predictive ordinate, a cross
  validation tool that predicts an area value using all the
  data except that area and compare that value to the actual
  value.
## E = specifies the offset variable required for a Poisson
  likelihood

# model 2: UH and CH effects
model2<-Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(region
  ,model="iid")+f(region2,model="bym",graph="
  nga_admbnda_adml_osgof_20161215.graph")
result2<-inla(model2,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE))
summary(result2)

# model 3: spatial + time trend (model 1a)
model3<-Mal~1+as.factor(Res)+as.factor(Weat)+year+subregion+f(
  region,model="iid")+f(region2,model="bym",graph="
  nga_admbnda_adml_osgof_20161215.graph")
result3 = inla(model3,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE,cpo=TRUE))
summary(result3)

# model 4: UH + CH + year IID
model4<-Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(region
  ,model="iid")+f(region2,model="bym",graph="
  nga_admbnda_adml_osgof_20161215.graph")+f(year,model="iid")
result4<-inla(model4,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE,cpo=TRUE))
summary(result4)

# model 5: UH + CH + year RW1 (model 1b)
model5<-Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(region
  ,model="iid",param=c(2,1))+f(region2,model="bym",graph="
  nga_admbnda_adml_osgof_20161215.graph")+f(year,model="rw1",
  param=c(1,0.01))
result5<-inla(model5,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE,cpo=TRUE))
summary(result5)

```

```

UH<-result4$summary.random\$region[,2]
yearR<-result4$summary.random\$year[,2]

# model 6: UH +CH +year UH +CH      (model 2)
year2<-year
model6<-Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(region
  ,model="iid")+f(region2,model="bym",graph="
  nga_admbnda_adm1_osgof_20161215.graph")+f(year,model="rw1")
  +f(year2,model="iid")
result6<-inla(model6,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE,cpo=TRUE))

# modle 7: UH+ year RW1 +INT  IID
model7<-Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(region
  ,model="iid")+f(year,model="rw1")+f(ind2,model="iid")
result7<-inla(model7,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE,cpo=TRUE))

# model 8: UH +CH + year RW1 + INT IID  (model 3)
model8<-Mal~1+as.factor(Res)+as.factor(Weat)+subregion+f(region
  ,model="iid")+f(region2,model="bym",graph="
  nga_admbnda_adm1_osgof_20161215.graph")+f(year,model="rw1")
  +f(ind2,model="iid")
result8<-inla(model8,family="poisson",data=data,E=E,control.
  compute=list(dic=TRUE,cpo=TRUE))

result1$dic$dic;result1$dic$sp.eff
result2$dic$dic;result2$dic$sp.eff
result3$dic$dic;result3$dic$sp.eff
result4$dic$dic;result4$dic$sp.eff
result5$dic$dic;result5$dic$sp.eff
result6$dic$dic;result6$dic$sp.eff
result7$dic$dic;result7$dic$sp.eff
result8$dic$dic;result8$dic$sp.eff

##The best model is model 8
#
# summary of the model 8
summary(result8)
# fixed effects
betas = result8$summary.fixed
betas

exp(betas)

## results for model 8

# get the shape file
library(maptools)

```

```

cities <- readOGR(dsn=dsn,layer="
  nga_admbnda_adml_osgof_20161215")
plot(cities)
names(cities)
# extract the data
UH = result7$summary.random$region[,2]*100000
yearR<-result7$summary.random$year[,2]*100000
STint<-result7$summary.random$ind2[,2]*100000

# plot risk rate by state
fillmap(cities,"Spatial Pattern for Nigeria Malaria Prevalence
  Risk ",UH,n.col=10)
fillmap(cities,"",UH,n.col=5)

plot(cities)
fillmap(cities,"",UH,n.col=5)

# plot risk rate by year
time<- c("2008", "2010","2013","2015")
plot(time,yearR, xlab="Year", ylab = " Risk Rates",main="
  Temporal Pattern for Malaria Risk Rates")
plot(time,yearR, xlab="Year", ylab = " Malaria Risk Rate")
lines(time,yearR)

# the S-T interaction
STest<-matrix(STint,ncol=4, byrow=T)

  ST1<-STest[,1]
  ST2<-STest[,2]

par(mfrow=c(1,2), mai=c(0,0,0.3,0),mar=c(2,1,1,1))
for(i in 1:4){
  #x11()
  fillmap(cities,paste("Spatial-Temporal in Year",2008+i,sep="
    "),STest[,i]*5,n.col=10)
}
x11()
for(i in 3:4){
  fillmap(cities,paste("Spatial-Temporal in Year",2008+i,sep="
    "),STest[,i]*10,n.col=10)
}

STest<-matrix(0,nrow = 88, ncol=10)

for(i in 1:4){i=ceiling(i/10) j=i-10*(k-1) STest[i,j]<-STint[i]
}

```

## References

- Abah, A., & Temple, B. (2015). Prevalence of malaria parasite among asymptomatic primary school children in Angiama Community, Bayelsa State, Nigeria. *Tropical Medicine & Surgery*, 4, 203–207.
- Abellan, J. J., Richardson, S., & Best, N. (2008). Use of space–time models to investigate the stability of patterns of disease. *Environmental Health Perspectives*, 116(8), 1111.
- Adigun, A. B., Gajere, E. N., Oresanya, O., & Vounatsou, P. (2015). Malaria risk in Nigeria: Bayesian geostatistical modelling of 2010 malaria indicator survey data. *Malaria Journal*, 14(1), 156.
- Awuah, R. B., Asante, P. Y., Sakyi, L., Biney, A. A., Kushitor, M. K., Agyei, F., & Aikins, A. d.-G. (2018). Factors associated with treatment-seeking for malaria in urban poor communities in Accra, Ghana. *Malaria Journal*, 17(1), 168.
- Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2004). *Hierarchical modeling and analysis for spatial data*. New York, NY: Chapman and Hall/CRC.
- Bennett, A., Bisanzio, D., Yukich, J. O., Mappin, B., Fergus, C. A., Lynch, M., . . . , Eisele, T. P. (2017). Population coverage of Artemisinin-based combination treatment in children younger than 5 years with fever and Plasmodium falciparum infection in Africa, 2003–2015: A modelling study using data from national surveys. *The Lancet Global Health*, 5(4), e418–e427.
- Chen, C., Wakefield, J., & Lumely, T. (2014). The use of sampling weights in Bayesian hierarchical models for small area estimation. *Spatial and Spatio-Temporal Epidemiology*, 11, 33–43.
- Ehlers, R., & Zevallos, M. (2015). Bayesian estimation and prediction of stochastic volatility models via INLA. *Communications in Statistics-Simulation and Computation*, 44(3), 683–693.
- Gemperli, A., Sogoba, N., Fondjo, E., Mabaso, M., Bagayoko, M., Briët, O. J., . . . , Vounatsou, P. (2006). Mapping malaria transmission in West and Central Africa. *Tropical Medicine & International Health*, 11(7), 1032–1046.
- Giardina, F., Gosoni, L., Konate, L., Diouf, M. B., Perry, R., Gaye, O., . . . , Vounatsou, P. (2012). Estimating the burden of malaria in Senegal: Bayesian zero-inflated binomial geostatistical modeling of the MIS 2008 data. *PLoS One*, 7(3), e32625.
- Gosoni, L., Msengwa, A., Lengeler, C., & Vounatsou, P. (2012). Spatially explicit burden estimates of malaria in Tanzania: Bayesian geostatistical modeling of the malaria indicator survey data. *PLoS One*, 7(5), e23966.
- Gosoni, L., Veta, A. M., & Vounatsou, P. (2010). Bayesian geostatistical modeling of malaria indicator survey data in Angola. *PLoS One*, 5(3), e9322.
- Hay, S. I., & Snow, R. W. (2006). The malaria atlas project: Developing global maps of malaria risk. *PLoS Medicine*, 3(12), e473.
- Israel, O. K., Fawole, O. I., Adebowale, A. S., Ajayi, I. O., Yusuf, O. B., Oladimeji, A., & Ajumobi, O. (2018). Caregivers’ knowledge and utilization of long-lasting insecticidal nets among under-five children in Osun State, Southwest, Nigeria. *Malaria Journal*, 17(1), 231.
- Kassegne, K., Zhang, T., Chen, S.-B., Xu, B., Dang, Z.-S., Deng, W.-P., . . . , Zhou, X.-N. (2017). Study roadmap for high-throughput development of easy to use and affordable biomarkers as diagnostics for tropical diseases: A focus on malaria and schistosomiasis. *Infectious Diseases of Poverty*, 6(1), 130.
- Kazembe, L. N., Chirwa, T. F., Simbeye, J. S., & Namangale, J. J. (2008). Applications of bayesian approach in modelling risk of malaria-related hospital mortality. *BMC Medical Research Methodology*, 8(1), 6.
- Khana, D., Rossen, L. M., Hedegaard, H., & Warner, M. (2018). A Bayesian spatial and temporal modeling approach to mapping geographic variation in mortality rates for subnational areas with R-INLA. *Journal of Data Science*, 16(1), 147.
- Kilian, A., Boulay, M., Koenker, H., & Lynch, M. (2010). How many mosquito nets are needed to achieve universal coverage? recommendations for the quantification and allocation of long-lasting insecticidal nets for mass campaigns. *Malaria Journal*, 9(1), 330.

- Kyu, H. H., Georgiades, K., Shannon, H. S., & Boyle, M. H. (2013). Evaluation of the association between long-lasting insecticidal nets mass distribution campaigns and child malaria in Nigeria. *Malaria Journal*, 12(1), 14.
- Lawson, A. B. (2013). *Bayesian disease mapping: Hierarchical modeling in spatial epidemiology*. Boca Raton, FL: Chapman and Hall/CRC.
- Lesaffre, E., & Lawson, A. B. (2012). *Bayesian biostatistics*. New York, NY: Wiley.
- Mouzin, E. (2012). *Focus on Nigeria*. [https://apps.who.int/iris/bitstream/handle/10665/87100/9789241503310\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/87100/9789241503310_eng.pdf). Accessed 10 Jan 2020.
- National Population Commission. (2010). Population and housing census of the Federal Republic of Nigeria 2006.
- National Population Commission. (2012). Nigeria malaria indicator survey 2010.
- National Population Commission. (2016). *Nigeria population projections by age and sex from 2006 to 2017*. Abuja: National Population Commission.
- Odugbemi, B., Ezeudu, C., Ekanem, A., Kolawole, M., Akanmu, I., Olawole, A., . . . , Babatunde, S. (2018). Private sector malaria RDT initiative in Nigeria: Lessons from an end-of-project stakeholder engagement meeting. *Malaria Journal*, 17, 70.
- Popoff, E. (2014). *An approximate spatio-temporal Bayesian model for Alberta wheat yield*. PhD thesis, University of British Columbia.
- Riedel, N., Vounatsou, P., Miller, J. M., Gosoni, L., Chizema-Kawesha, E., Mukonka, V., & Steketee, R. W. (2010). Geographical patterns and predictors of malaria risk in Zambia: Bayesian geostatistical modelling of the 2006 Zambia National Malaria Indicator Survey (ZMIS). *Malaria Journal*, 9(1), 37.
- Singh, R., Musa, J., Singh, S., & Ebere, U. V. (2014). Knowledge, attitude and practices on malaria among the rural communities in Aliero, Northern Nigeria. *Journal of Family Medicine and Primary Care*, 3(1), 39.
- Spiegelhalter, D., Best, N. G., Carlin, B. P., & van der Linde, A. (2003). Bayesian measures of model complexity and fit. *Quality Control and Applied Statistics*, 48(4), 431–432.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3), 485–493.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583–639.
- Wongsrichanalai, C., Barcus, M. J., Muth, S., Sutamihardja, A., & Wernsdorfer, W. H. (2007). A review of malaria diagnostic tools: Microscopy and rapid diagnostic test (RDT). *The American Journal of Tropical Medicine and Hygiene*, 77(6\_Suppl.), 119–127.
- World Health Organization. (2015). *World malaria report 2014*. World Health Organization.
- World Health Organization. (2017). *Global hepatitis report 2017*. World Health Organization.