Investigating the spatial patterns of cough, fever and diarrhea among young children in West African countries

Ezra Gayawan
Department of Statistics, Federal University of Technology, Akure
egayawan@futa.edu.ng

O. B. Aladeniyi
Department of Statistics, Federal University of Technology, Akure
obaladeniyi@futa.edu.ng

Samson B. Adebayo
Planning Research and Statistics Directorate, National Agency for Food and Drug Administration and Control, Abuja, Nigeria
samson.adebayo@nafdac.gov.ng

Abstract

Children in developing countries have continued to suffer morbidity and mortality arising from a few number of illnesses. In this study, we adopt a spatial modeling techniques and pooled data from six neighboring West African countries, obtained from Demographic and Health Survey, to elucidate the geographical pattern of cough, fever and diarrhea among young children in a manner that transcend boundary. Analysis was through the Bayesian framework with appropriate priors assigned to the different parameters. Results show similar significant spatial distributions for the three illnesses, and that demonstrate that children in Benin Republic and Mali are less likely to suffer from these illnesses while higher likelihood were found in Cote d'Ivoire, Burkina Faso, Togo and some part of Ghana. The nonlinear effects of child's age show that children around age 8-10 months have higher risks of suffering from these diseases while, as the mothers advance in age, their children have lower risks. Breastfeeding, woman working status and education are among the significant factors that either enhance or inhibit these illnesses.

Keywords: Spatial analysis; childhood illnesses; fever; West Africa.

1. Introduction

The world has been accelerating progress in reducing the under-5 mortality rate. Sub-Africa countries also registered remarkable decline in under-5 mortality since year 2000, but this progress was insufficient to reach the Millennium Development Goal (MDG 4) target of a two-third reduction in the under-5 mortality rate by year 2015, and is unlikely, in the present state, to attain the Sustainable Development Goal (SDG) target of a neonatal mortality rate of 12 deaths per 1000 live birth by 2030 (WHO factsheet, 2016). West African countries have been identified as the major contributors to child mortality in Sub-Saharan Africa, recording more than double the case recorded in the southern or northern Africa (Balk et al. 2004; Gayawan et al. 2016). Apart from malnutrition, leading causes of death in children under-5 years in Africa, like in most developing countries have been identified to include cough (pneumonia), diarrhea and malaria/fever (WHO 2008).

It is evident that most diseases are spatially structured and therefore, policy makers are often interested in the distributions of diseases, not just the prevalence. Though most survey reports provide the disease prevalence across different geographical locations, estimates from these reports may not be too reliable as they depend entirely on sample sizes taken from different locations and, the effects of other contributory factors have not been accounted for. Model based estimation of the distributions across space offers more reliable results as the procedures involve mechanisms that borrow strength from neighboring locations in the estimation process thereby yielding more reliable estimates.

There have been extensive discussions in the literature on the determinants and spatial distributions of childhood illnesses including fever, pneunomia, diarrhea, malaria and anaemia in most sub-Saharan African countries (Kazembe and Namangale 2007; Gayawan et al. 2014; Kandala et al. 2006; Kandala et al. 2007) and all the studies confirm the earlier suggestion of Kalipeni (1993) that the geographic location which a child lives sets the context for child morbidity and mortality. All the studies cited above have been based on country specific analysis. Although the distribution of the risk of anaemia in pre-school children after adjusting for nutritional status, parasitic infections and other individual variables have been considered in West Africa countries, the work was however based on three countries: Burkina Faso, Mali and Ghana (Magalhaes and Clements 2011). The present study takes advantage of the geo-reference data provided by the Demographic and Health Surveys (DHS) to map the distribution of childhood illnesses in multiple West African countries. Specifically, we investigate, in a manner that transcends geographical boundary, the spatial distributions and determinants of fever, cough and diarrhea in six West African countries namely: Benin, Ghana, Togo, Cote' de I'vore, Burkina Faso and Mali. We adopt structured additive regression analysis, approach that allows to considered nonlinear effects of continuous variables, structured and unstructured random effects and the usual linear effects in a single framework (A Brezger and S Lang 2006). With this methodology, we were able to discern not only the geographical distributions of the illnesses but also the detailed functional relationship between the diseases and continuous variables including child's and mother's age.

2. Data

The study utilizes data from Demographic and Health Surveys (DHS) conducted in the six West African countries between 2010 and 2014. The DHS programme has, over the years, provided technical assistance to surveys in several developing countries, thereby advancing global understanding of health and population trends. DHS has developed standard procedures, methodologies and manuals to guide the survey processes and to ensure that the data, not only properly reflect the situations they are intended to describe, but are also comparable across countries and over time. The primary sampling units are usually defined on the basis of the enumeration areas (EAs) from the Census frames of the country to be surveyed. DHS samples are usually selected using two- or three-stage stratified design. Information are collected from all women aged 15‐49 years present at the selected households as well as on all children younger than five years of age.

West African countries included in the study and the year of survey are as follows: Benin (2011–2012), Burkina Faso (2010), Cote d'Ivoire (2011–2012), Ghana (2014), Mali (2012–2013), and Togo (2013-2014). In the case of Mali, three regions, namely Tombouctou, Gao and Kidal, were not covered during the surveys and were therefore excluded. Demographic variables that were considered include mother's age, child's age, type of place of residence, child's birth order, sex of the child mother's educational attainment, household wealth index, mother's working status, electricity, toilet facility, water source, exposure to media (newspaper, radio and television; whether or not the mother was exposed to each of these at least once a week) and whether or not he is being breastfed. The regions, being the spatial unit of analysis, were geo-referenced.

3. Statistical analysis

Consider observations (y_i, x_i, s_i, v_i), i = 1, …, n, where y_i is a binary response variable:

$$y_i \sim Bin(n_i, p_i)$$

where p_i is the proportion of the children that suffer from the diseases (cough, fever or diarrhea) breastfed, x=(x_1, …, x_p)′ is a vector of metrical covariates e.g. mother's age and child's age, s_i=(1,…, s_q) the region where child i lived during the survey; and vector v=(v_1, …, v_q)′ of categorical covariates. Usually intention is to model the dependence of y_i on metrical, spatial and categorical covariates within the context of generalized additive model (Hastie and Tibshirani 1990). The logistic model with structured additive predictors is defined as:

$$\log\left(\frac{p_i}{1-p_i}\right) = \sum_{j=1}^{p} f_j(x_{ij}) + f_{spat}(s_i) + v_i'\beta \qquad (1)$$

where f_1,… ,fp are nonlinear (unknown) smooth functions of the metrical covariates, f_{spat} is the nonlinear effect of spatial covariates and β_i = (β_1 ,…, β_L)′ is a vector of fixed effect parameters for the categorical covariates. One may further split up the spatial effects f_{spat} into spatially correlated (structured) and uncorrelated (unstructured) effects as

$$f_{spat}(s_i) = f_{str}(s_i) + f_{unstr}(s_i)\,.$$

A rationale behind this is that a spatial effect is a surrogate of many unobserved influential factors, some of which may be a strong spatial structure and others may only be present locally. The modeling approach was through the binary logistic model within a Bayesian perspective that jointly accounts for nonlinear, spatial and random effects. Bayesian structured additive regression of logit model is preferred because of ease of interpretation and the possibility of computing posterior odds ratio (which is a measure of strength of association in logistic regression) directly from the Markov chain Monte Carlo (MCMC) output. Fitting Bayesian models via MCMC entails treating all parameters as randomly distributed according to some prior distribution. The posterior distribution is intractable, so MCMC algorithms are used to generate samples from this prior distribution which allow estimation and inference for the parameters (Diggle et al. 1998).

Within a Bayesian context, all parameters and functions (say f for nonlinear effects) are considered as random variables upon which appropriate priors are assumed. Independent diffuse priors are assumed on the parameters of fixed effects. For the non-linear effects, Bayesian P-splines prior based on Lang and Brezger (2004) and A. Brezger and S. Lang (2006) was assumed. The P-splines prior allows for nonparametric estimation of *f* as a linear combination of basis function (B-spline)

$$p(z) = \sum_{j=1}^{J} \beta_j B_j(z)$$

where $B_j(z)$ are B-splines basis function and the coefficients β_j corresponds to the vector of unknown regression coefficients. Smoothness of function *f* is achieved by penalizing differences of coefficients of adjacent B-splines (Eilers and Marx 1996) or in a Bayesian approach as in this case, β_j are further defined to follow a first or second order Gaussian random walk smoothness priors

$$\beta_t = \beta_{j-1} + u_t \qquad \beta_t = 2\beta_{j-1} - \beta_{j-2} + u_t\,,$$

with *i.i.d.* errors $u_t \sim N\left(0, \tau^2\right)$. The variance τ^2 controls the smoothness of *f*. Assigning a weakly informative inverse Gamma prior $\tau^2 \sim \mathrm{IG}(\varepsilon, \varepsilon)$, ε small, it is estimated jointly with the basis function coefficients. For further clarifications and understanding of this concept, readers are referred to Eilers and Marx (1996); Lang and Brezger (2004) and A. Brezger and S. Lang (2006).

Random effects $\left(z_{gi}\right)$ were modeled by assuming exchangeable normal priors, $z_{ij} \sim N\left(0, \tau_b^2\right)$, where τ_b^2 is a variance component that incorporates over-dispersion and heterogeneity. For the spatial effects $f_{spat}\left(s\right)$, we chose a Gaussian Markov random field prior, which is common in spatial statistics, see (Besag et al. 1991). It is given as

$$[f_{str}(s) \mid f_{str}(t); t \neq s, \tau^2] \sim N\left(\sum_{t \in \partial_s} \frac{f_{str}(t)}{N_s}, \frac{\tau^2}{N_s}\right)$$

where N_s is the number of adjacent sites and $t \in \partial_s$ denotes that site *t* is a neighbor of site *s*. Thus the (conditional) mean of $f_{spat}\left(s\right)$ is an average of function evaluations $f_{spat}\left(t\right)$ of neighboring sites *t*. Again τ^2 controls the amount of spatial smoothness.

In order to be able to estimate the smoothing parameters for non-linear and spatial effects simultaneously, highly dispersed but proper hyper-priors are assigned to them. Hence for all variance components, an inverse gamma distribution with hyperparameters *a* and *b* is chosen, e.g. $\tau^2 \sim IG(a,b)$. Standard choices of hyperparameters are a=1 and b=0.005 or a=b=0.001. The latter option was used in this case study.

4. Results and Discussion

Results for the spatial effects for cough, fever and diarrhea are presented in Figures 1-3 respectively. Part (a) of the figures present the maps of posterior means while the (b) part show the maps of 95% credible interval, which are used in deciding the significance of the posterior mean estimates. From the maps of credible intervals, regions in black (white) shading are places where the estimates are significantly lower (higher) whereas, estimates for regions shaded in grey colour are not significant. The results show some interesting spatial pattern for the three diseases under consideration. As for cough, results show that children are less likely to have suffered from cough throughout Benin Republic and Mali as well as in some regions of Burkina Faso and these are significant. However, in Togo, where all the regions share boundary with Benin Republic, findings show that children in that country have higher likelihood of suffering for cough. Estimates also show that but for three regions of Cote D'Ivoire, children in all other regions of the country have higher likelihood of suffering from cough while, on the other hand, significantly higher likelihood exists in only the Eastern region of Ghana.

Results for fever show similar pattern as those of cough. Again, children in Benin Republic, those residing in all but Sikasso regions of Mali as well as in Central region of Ghana faced significantly less likelihood of suffering from fever. Places with higher likelihood are obtained are: all regions of Cote d'Ivoire except Zanzan, Vallee du Bandama, Sud-Comoe and N'zi-Comoe; Cascade, Sud-Ouest, Hauts-Bassins, Centre_Ouest, Boucle du Mouhoun, Nord, Centre-Sud and Plateau-Centralregions of Burkina Faso and Plateaux and Savanes regions of Togo. As for diarrhea, the results are again similar to those of cough and fever. Children from Benin Republic and Mali have less likelihood of suffering from the illness while, those from northern Ghana, Eastern Cote d'Ivoire, Western Burkina Faso and Northern Togo have higher likelihood.

Estimates for the nonlinear effects of child's age and mother's age on cough, fever and diarrhea are presented in Figure 2(a-f). The results show similar pattern for all the three illnesses under consideration. The results show an inverse 'U' relationship between child's age and the three illnesses, signifying that the likelihood that a child would suffer from the illnesses increase as the child grows older up to around age 10 months from where it drops drastically with age. This finding conforms to other studies in African and point to the influence of breastfeeding (Kandala et al. 2007; Kandala et al. 2006; Kazembe and Namangale 2007). As for the mother's age, there is evidence of linear relationship that shows that as the mother advance in age, the child is less likely to suffer from the diseases.

Findings for the fixed effects are presented in Table 1. The Table shows the posterior means and 95% credible intervals. Children who were being breastfed at the time of the survey were having lower odds of having coughing whereas, the results show that children whose mothers watch television at least once a week, those belonging to the richest households and children of working women have higher odds of coughing. As for fever, results show that children whose parents attained higher education have lower odds of having the illness when compared with those whose mothers had no education, those form households with improved toilet facilities and those being breastfed were also less likely to have fever compared with their other counterparts. The categories of children with higher odds of having cough are also found to have fever. As for diarrhea, children whose mothers listen to radio at least once a week, female children and those receiving breastfeeding possess the lower odds of having the disease against their other counterparts. On the other hands the odds are significantly higher for children whose mothers attained primary education, working mothers and those of fourth or higher birth order. All other estimates are not significant.

5. Conclusions

This study adopts structured additive regression modeling techniques to elucidate the spatial pattern and determinants of common childhood illnesses pooling data from six West African countries. The results show that the illnesses: cough, fever and diarrhea are spatially structured displaying similar patterns across the countries under consideration. The results have however pinpointed reions of West African countries with high and low risks of the illnesses and this would enhance intervention strategies of policy makers and donors in the sub region. Also worthy of note is the findinds that show high risk of the illnesses among children under age 10 months. More emphasis would need to be given to ensuring the that the WHO recommended 6 months exclusive breastfeeding is adhered to more so that the fixed effects show that childrren receiving breast milk have lower risks of suffering from any of the illnesses.

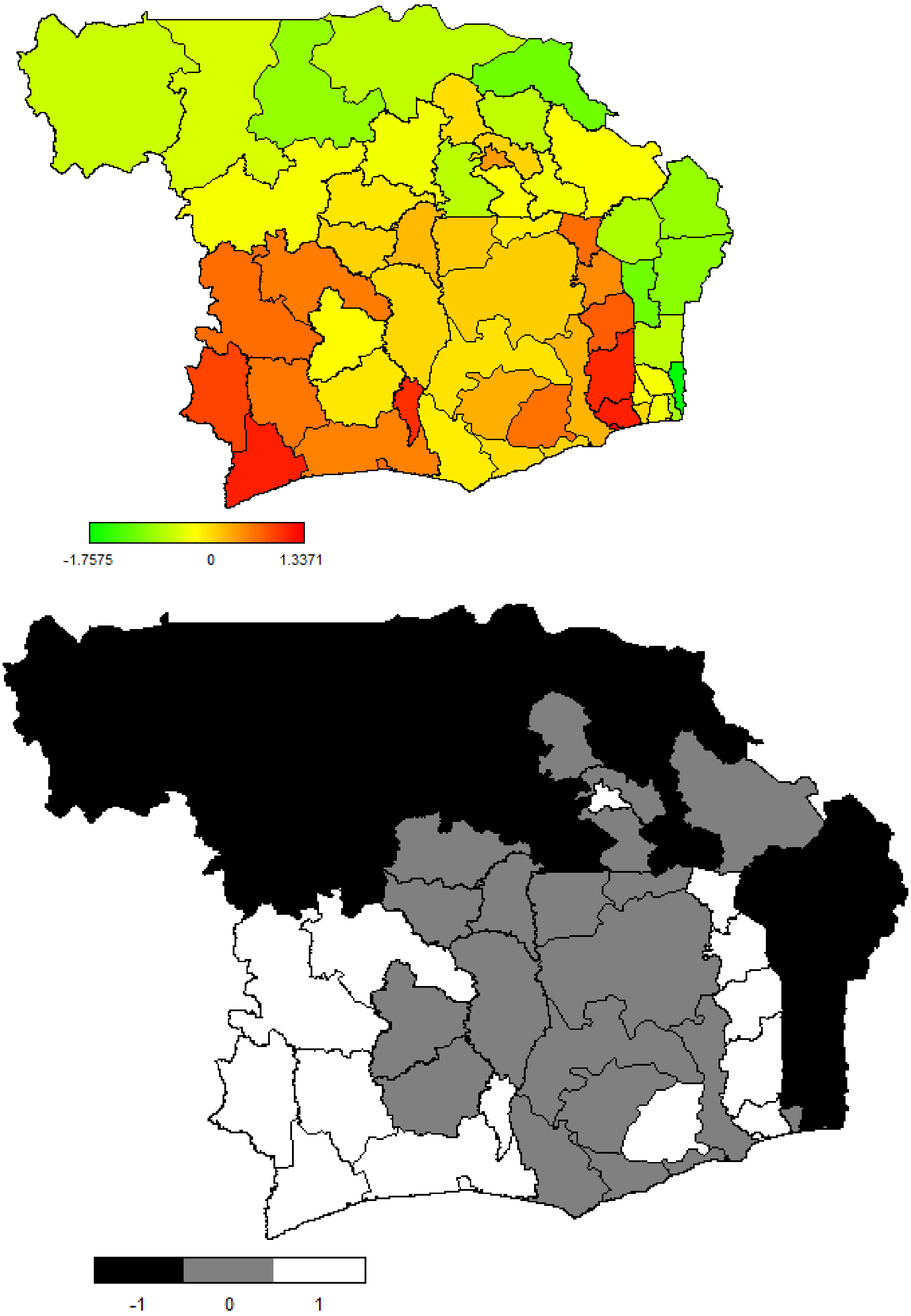

Figure 1: Estimated posterior mean spatial effects of cough (a) and the 95% credible interval (b)

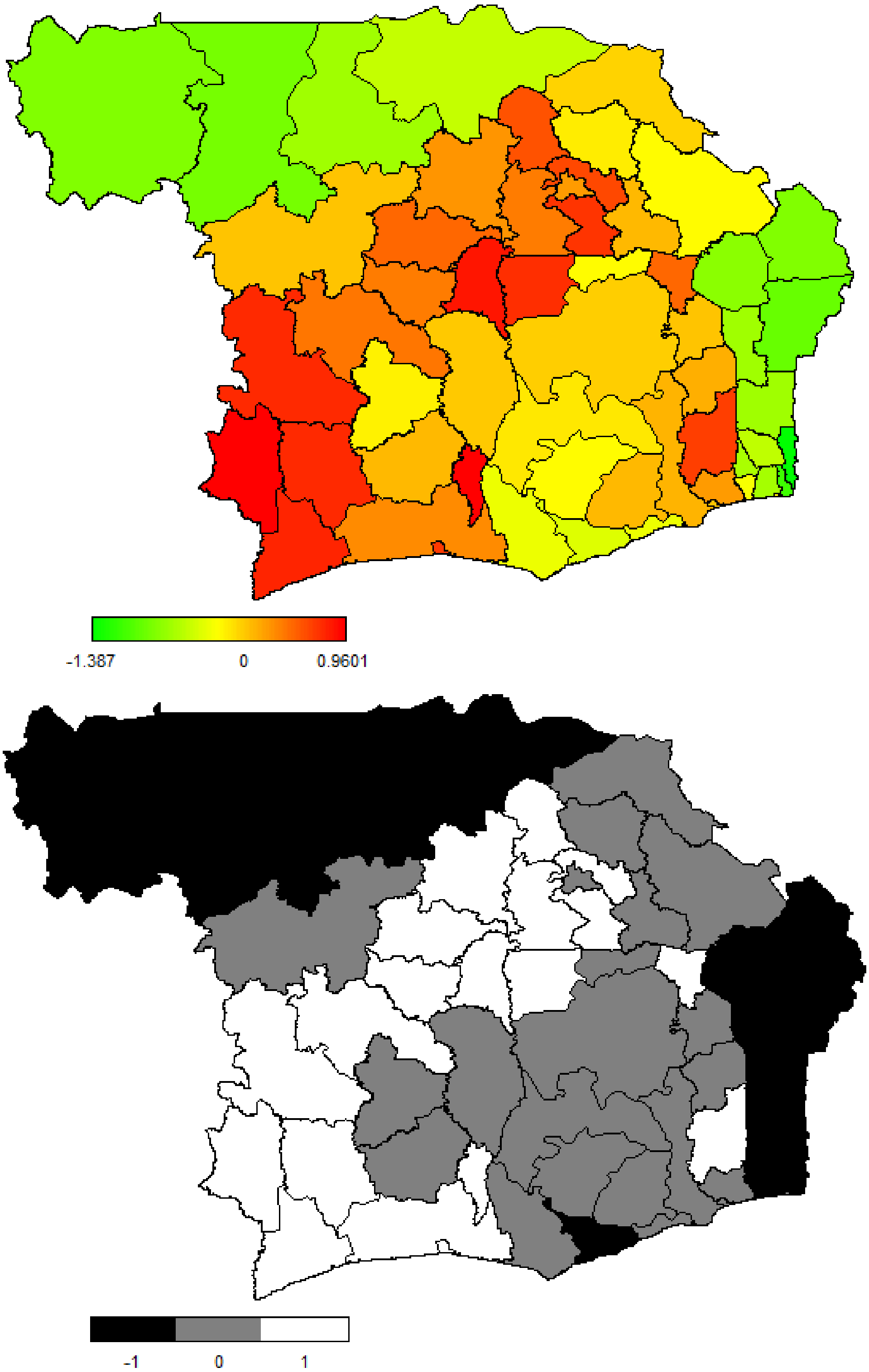

Figure 2: Estimated posterior mean spatial effects of fever (a) and the 95% credible interval (b)

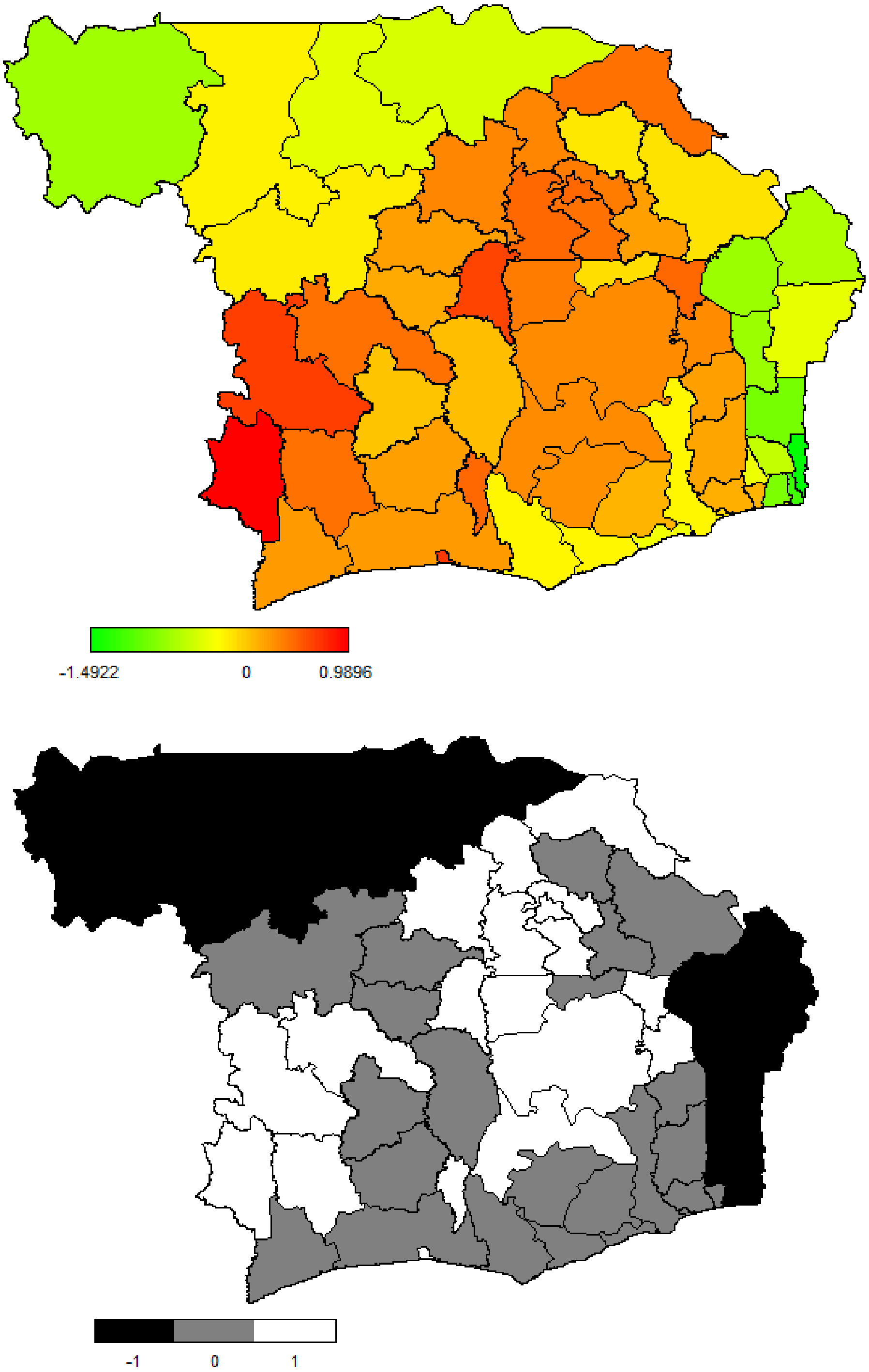

Figure 3: Estimated posterior mean spatial effects of diarrhea (a) and the 95% credible interval (b)

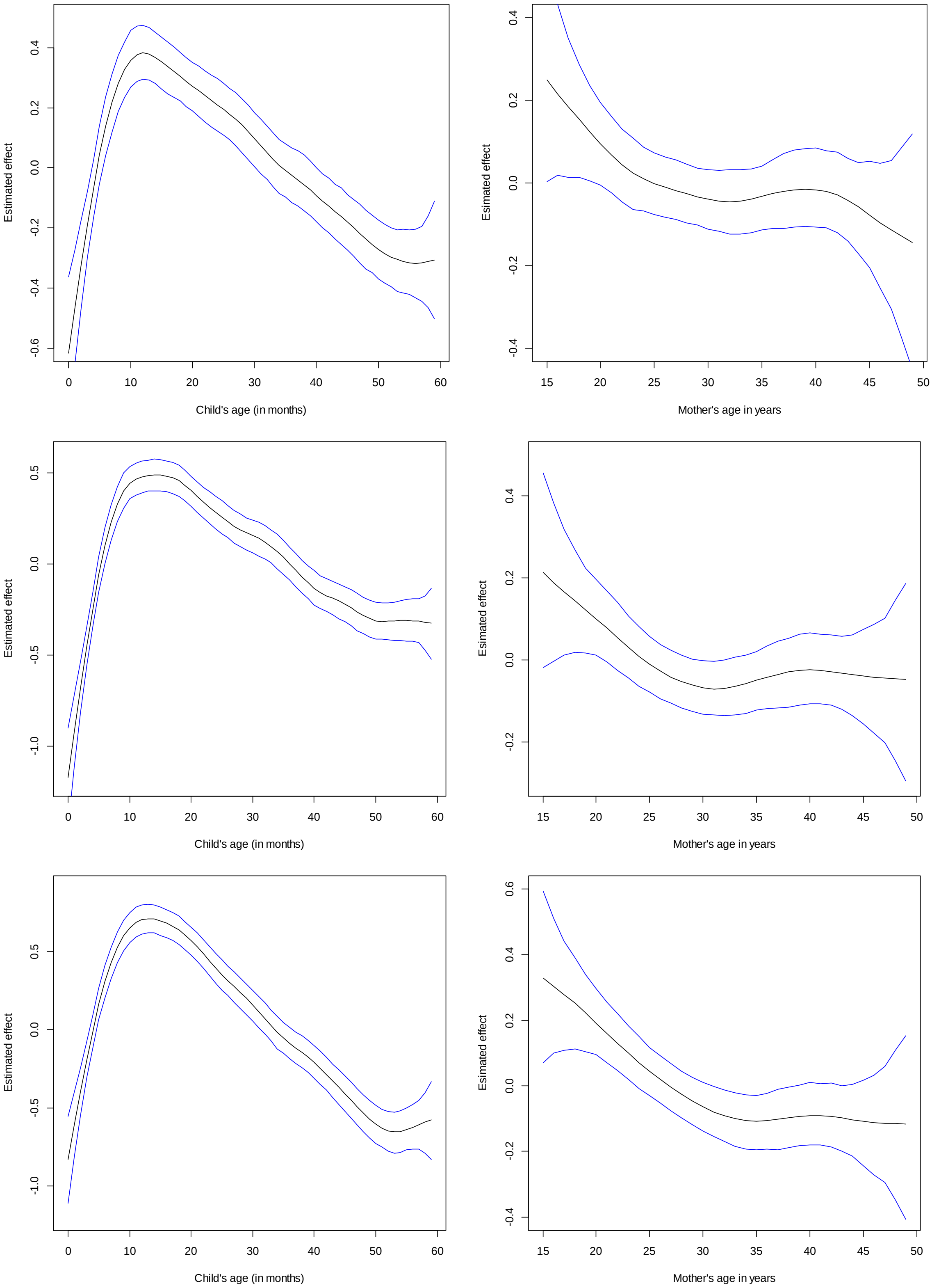

Figure 4: Nonlinear effects of child's age and mother's age for cough (a & b), fever (c & d) and diarrhea (e & f)

Table 1: Posterior means and 95% credible intervals for the fixed effects

Variable	Cough		Fever		Diarrhea	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
Place of residence						
Rural	1		1		1	
Urban	1.009	0.910, 1.122	0.996	0.905, 1.097	0.984	0.889, 1.098
Level of education						
No Education	1		1		1	
Primary	1.112	1.008, 1.224	1.088	0.992, 1.195	1.214	1.091, 1.341
Secondary	1.062	0.931, 1.215	0.986	0.868, 1.122	1.097	0.961, 1.257
High School	0.877	0.618, 1.226	0.582	0.392, 0.866	0.632	0.389, 1.019
Water source						
Not protected	1		1		1	
Protected	0.966	0.887, 1.055	0.984	0.910, 1.062	0.965	0.886, 1.054
Toilet facility						
Non-improved	1		1		1	
Improved	1.001	0.904, 1.107	0.878	0.795, 0.962	0.952	0.849, 1.054
Electricity						
No	1		1		1	
Yes	0.961	0.858, 1.076	0.911	0.819, 1.013	0.900	0.799, 1.009
Access to mass media						
Newspaper						
No	1		1		1	
Yes	1.088	0.939, 1.255	1.039	0.897, 1.191	0.911	0.763, 1.067
Radio						
No	1		1		1	
Yes	0.953	0.882, 1.037	1.042	0.972, 1.122	0.842	0.777, 0.913
Television						
No	1		1		1	
Yes	1.151	1.041, 1.267	1.100	1.005, 1.208	1.065	0.965, 1.176
Wealth Index						
Poorest	1		1		1	
Poorer	0.934	0.837, 1.048	0.967	0.884, 1.066	1.022	0.921, 1.138
Middle	0.988	0.875, 1.104	1.054	0.949, 1.167	1.057	0.943, 1.182
Richer	1.037	0.897, 1.192	1.122	1.001, 1.270	1.119	0.976, 1.299
Richest	1.236	1.014, 1.475	1.149	0.963, 1.361	1.106	0.906, 1.319
Mother's Working Status						
Not working	1		1		1	
Working	1.272	1.168, 1.377	1.358	1.257, 1.461	1.095	1.006, 1.191
Birth order						
1st birth	1		1		1	
2nd or 3rd birth	0.940	0.846, 1.045	1.010	0.916, 1.118	1.065	0.955, 1.190
4th or higher birth	0.971	0.854, 1.104	1.119	0.996, 1.260	1.212	1.063, 1.387
Sex						
Male	1		1		1	
Female	1.022	0.954, 1.096	0.950	0.891, 1.013	0.897	0.835, 0.963
Breastfeeding						
No	1		1		1	
Yes	0.889	0.818, 0.963	0.824	0.763, 0.889	0.858	0.785, 0.934

References

Balk, D., Pullum, T., Storeygard, A., Greenwell, F., & Neuman, M. (2004). A spatial analysis of childhood mortality in West Africa. *Population, Space and Place, 10*, 175-216.

Besag, J., York, Y., & Mollie, A. (1991). Bayesian Image Restoration with two Applications in Spatial Statistics (with discussion). *Annals of the Institute of Statistical Mathematics, 43*, 1-59.

Brezger, A., & Lang, S. (2006). Generalized structured additive regression based on Bayesian P-splines. *Computational Statistics and Data Analysis, 50*, 967-991.

Brezger, A., & Lang, S. (2006). Generalized Structured Additive Regression based on Bayesian P splines. *Computational Statistics and Data Analysis, 50*, 967-991.

Diggle, P. J., Tawn, J. A., & Moyeed, R. A. (1998). Model-based geostatistics (with discussion). *Applied Statistics, 47*, 299-350.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science, 11*(2), 89-121.

Gayawan, E., Adarabioyo, M. I., Okewole, D. M., Fashoto, S. G., & Ukaegbu, J. C. (2016). Geographical variations in infant and child mortality in West Africa: a geo-additive discrete-time survival modelling. *Genus, 72*(5), doi:10.1186/s41118-016-0009-8.

Gayawan, E., Arogundade, E. D., & Adebayo, S. B. (2014). A Bayesian multinomial modeling of spatial pattern of co-morbidity of malaria and non-malarial febrile illness among young children in Nigeria. *Transactions of the Royal Society for Tropical Medicine and Hygiene, 108*, 415-424.

Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*: Chapman and Hall, London.

Kalipeni, E. (1993). Determinants of infant mortality in Malawi: a spatial perspective. *Social Science and Medicine, 37*(2), 183-198.

Kandala, N. B., Ji, C., Stallard, N., Stranges, S., & Cappuccio, F. P. (2007). Spatial analysis of risk factors for childhood morbidity in Nigeria. *American Journal of Tropical Medicine and Hygiene, 77*(4), 770-778.

Kandala, N. B., Magadi, M. A., & Madise, N. J. (2006). An investigation of district spatial variations of childhood diarrhoea and fever morbidity in Malawi. *Social Science and Medicine, 62*, 1138-1152.

Kazembe, L. N., & Namangale, J. J. (2007). A Bayesian multinomial model to analyse spatial patterns of childhood comorbidity in Malawi. *European Journal of Epidemiology, 22*, 545-556.

Lang, S., & Brezger, A. (2004). Bayesian P-Splines. *Journal of Computational and Graphical Statistics, 13*, 183-212.

Magalhaes, R. J. S., & Clements, A. C. A. (2011). Mapping the risk of anaemia in preschool-age children: the contribution of malnutrition, malaria, and helminth infections in West Africa. *PLoS Medicine, 8*(6), e1000438.

WHO (2008). World Malaria Report 2008. Geneva: WHO.