

RESEARCH

Open Access

Geo-additive modelling of malaria in Burundi

Hermenegilde Nkurunziza^{1*}, Albrecht Gebhardt² and Jürgen Pilz²

Abstract

Background: Malaria is a major public health issue in Burundi in terms of both morbidity and mortality, with around 2.5 million clinical cases and more than 15,000 deaths each year. It is still the single main cause of mortality in pregnant women and children below five years of age. Because of the severe health and economic burden of malaria, there is still a growing need for methods that will help to understand the influencing factors. Several studies/researches have been done on the subject yielding different results as which factors are most responsible for the increase in malaria transmission. This paper considers the modelling of the dependence of malaria cases on spatial determinants and climatic covariates including rainfall, temperature and humidity in Burundi.

Methods: The analysis carried out in this work exploits real monthly data collected in the area of Burundi over 12 years (1996-2007). Semi-parametric regression models are used. The spatial analysis is based on a geo-additive model using provinces as the geographic units of study. The spatial effect is split into structured (correlated) and unstructured (uncorrelated) components. Inference is fully Bayesian and uses Markov chain Monte Carlo techniques. The effects of the continuous covariates are modelled by cubic p-splines with 20 equidistant knots and second order random walk penalty. For the spatially correlated effect, Markov random field prior is chosen. The spatially uncorrelated effects are assumed to be i.i.d. Gaussian. The effects of climatic covariates and the effects of other spatial determinants are estimated simultaneously in a unified regression framework.

Results: The results obtained from the proposed model suggest that although malaria incidence in a given month is strongly positively associated with the minimum temperature of the previous months, regional patterns of malaria that are related to factors other than climatic variables have been identified, without being able to explain them.

Conclusions: In this paper, semiparametric models are used to model the effects of both climatic covariates and spatial effects on malaria distribution in Burundi. The results obtained from the proposed models suggest a strong positive association between malaria incidence in a given month and the minimum temperature of the previous month. From the spatial effects, important spatial patterns of malaria that are related to factors other than climatic variables are identified. Potential explanations (factors) could be related to socio-economic conditions, food shortage, limited access to health care service, precarious housing, promiscuity, poor hygienic conditions, limited access to drinking water, land use (rice paddies for example), displacement of the population (due to armed conflicts).

Background

In Burundi, malaria is a major public health issue in terms of both morbidity and mortality with around 2.5 million clinical cases and more than 15,000 deaths each year. In 2001, Burundi was the world's most affected country by malaria [1]. Malaria is the main cause of mortality among pregnant women and children under

five years of age, accounting for more than 50% of all cases.

Many studies have been undertaken to understand factors that are associated with malaria in many countries. Most of them found a strong association between malaria and climate [2-5]. For example, the results in [2] suggest that the variability of the climate played an important role in initiating epidemics of malaria in the highlands of East Africa. A significant positive correlation between the number of malaria cases and temperature and rainfall has been identified. Pemola and Jauhari

* Correspondence: nkuhermes@yahoo.fr

¹Department of Mathematics, Institute of Applied Pedagogy, University of Burundi, Burundi

Full list of author information is available at the end of the article

[3] found higher positive correlation between monthly malaria parasite incidence and climatic variables (temperature, rainfall and humidity) in Dehradun, India. Gallup and Sachs [4] suggested that the location and severity of malaria are mostly determined by climate and ecology. Bouma *et al* [5] concluded that rainfall and humidity were able to predict malaria rates fairly well in Pakistan.

However, other studies on the same topic suggested that factors other than climate may explain the distribution of malaria [6-11]. For example, Cox *et al* [6] noted that the relatively high rates of malaria morbidity in Africa could result from poor access to health services, inadequate case management, overwhelmed health services, poor immunological competence because of malnutrition, a general disruption to livelihoods because of often-associated flooding, or a combination of these factors. Patz and Lindsay [7] suggested the existence of many variables affecting malaria transmission beside the climatic changes, such as environmental factors, the population growth, a limited access to health care systems, and lack of or unsuccessful malaria control measures. Kigbafori *et al* [8] concluded that risk factors for malaria infection include age, socioeconomic factors, not sleeping under a bed net, lack of health care facilities and various environmental features, such as vegetation, rainfall and distance to rivers. Tren [9] suggested that though climate can affect the incidence of malaria, man's economic activities and malaria control policy play a very important role in the incidence of the disease. Hay *et al* [11] suggested that the claimed association between local malaria resurgence and regional changes in climate, in Eastern Africa, is overly simplistic. They suggest that economic, social and political factors explain recent resurgence in malaria and other mosquito-born diseases with no need to invoke climate change.

In this study, a geo-additive model is proposed to understand the dependence of malaria cases on spatial effects and climatic covariates including rainfall, maximum and minimum temperature, maximum and minimum humidity in Burundi.

Methods

Study area

Burundi is located in East-central Africa, between 2°20 and 4°27 of latitude south and between 28°50 and 30°53 of longitude east; the altitude varies between 775 metres (Lake Tanganyika) and 2,670 metres (Crest Congo - Nil). Burundi has in general a tropical highland climate with a significant daily temperature variation in many areas [12]. Temperature also varies significantly from one region to another mainly due to differences in altitude. The area in the central plateau is cool, with temperature averaging 20°C. The area near Lake Tanganyika is warmer, averaging

23°C; the areas in the highest mountains are cooler with temperature averaging 16°C. Rain is irregular and falls most heavily in the northwest region [12]. Dry season varies in length with sometimes longer periods of drought. Most parts of Burundi receive rainfall between 130 cm and 160 cm per year [12]. Bounded on the north by Rwanda, in south-east by Tanzania and in west by the Democratic Republic of Congo, Burundi covers an area of 27,834 km² (of which 2,634 km² are occupied by Tanganyika Lake) and has a population estimated at about 8 million. In terms of habitat, it remains essentially rural, with 91.6% of the population living in rural area. The urban population is 8.4% with an annual growth rate of 5.7%. The Burundi population is young: 46.1% are under 15 years of age, while people aged 60 and above represent only 5.4%. With an average density of 266 inhabitants per km², a population growth rate of 3.44% and a total fertility rate of 6 children per woman, Burundi is one of Africa's most densely populated countries [13]. Burundi is structured in 17 provinces. The epidemiological profile can be summarized as follows. The health system suffers from a shortage of qualified personnel with 1 doctor per 34,750 inhabitants and 1 nurse for 3,500 inhabitants [13]. 17.4% of patients do not have access to health care, while 81.5% of patients are forced to go into debt or sell property to pay the health costs. There is a big disparity between the capital Bujumbura and the remainder of the country as 80% of doctors and more than 50% of nurses are engaged in Bujumbura. Responsible for more than 50% of hospital deaths in children under five years of age and more than 40% of all consultations in health centres, malaria is undoubtedly the main public health problem, the main cause of mortality and morbidity in Burundi [13].

Data description

The goal in our study is to understand the dependence of malaria cases on factors such as climatic variables and spatial (correlated and uncorrelated) effects in Burundi. Monthly data on malaria morbidity in Burundi over 12 years (from 1996 to 2007) were collected from EPI-STAT (Epidemiology and Statistics in Burundi) [14], a department of the Burundi Ministry of health in charge of collecting and storing data on epidemiology all over the country. The well-known nearest neighbour method was used to fill the missing data (~5%). The estimated population for each province, for the study period, was obtained from the Institute of Statistics and Economic Studies in Burundi (ISTEEBU)[[15] Malaria incidence in a given province was computed by dividing the number of malaria cases by the total population of the province, assuming that the whole population is susceptible. Monthly data on cumulative precipitation, monthly average of daily maximum temperature, minimum temperature, maximum humidity and minimum humidity for

1996-2007 was obtained from the Geographic Institute of Burundi (IGEBU) [16]. The record of these variables from 1996 through 2007 has remained uniform, with the same calibration and the same precision. The missing data (2% - 3%) were filled by the same method as in Malaria data (nearest neighbour and cross-validation). Data for three provinces (Bubanza, Bujumbura rural and Cibitoke) were not available for the study period; they were estimated using ordinary kriging [17]. The data are available on different scales and units (malaria incidence and humidity are unit free, rainfall is measured in centimetre (cm), temperature in degree centigrade (°C)). They were then standardized to avoid the effect of scale in the modelling.

Model formulation

In a previous study [18], assuming that climatic covariates have a nonlinear effect on malaria incidence and based on the Akaike information criterion (AIC) using the algorithm described in [23], the following generalized additive mixed model (GAMM) [24] was proposed to assess the dependence of malaria cases on climatic variables.

$$\eta_{it} = \alpha_0 + f_1(R_{nit}) + f_2(H_{xit}) + f_3(T_{xp_{it}}) + f_4(H_{xp_{it}}) + \alpha_1 T_{xit} + \alpha_2 T_{nit} + \alpha_3 T_{np_{it}} + \varepsilon_{it} \quad (1)$$

Here η_{it} is the predictor of malaria incidence assumed to have a gamma distribution, R_{nit} is the rainfall, H_{xit} is the maximum humidity, T_{xit} is the maximum temperature and T_{nit} is the minimum temperature, of the province i in month t . T_{xp} , T_{np} , H_{xp} are the same variables for the previous month. f_1, \dots, f_4 are unknown nonlinear smooth functions of the covariates. The α_i ($i = 1, \dots, 3$) are the regression coefficient of the linear effects. α_0 is the intercept (accounting for unmeasured covariates). ε_{it} is the error.

The aim here was to assess the climatic factors that are highly associated with monthly malaria incidence in Burundi; hence spatial effect was not included in the model. The results have shown that malaria incidence in a given month is positively associated with the minimum temperature in the previous month. In this study, the GAMM in (1) is replaced by a geo-additive model by incorporating the spatial effects as follows [25-32].

$$\eta_{it} = \alpha_0 + f_1(R_{nit}) + f_2(H_{xit}) + f_3(T_{xp_{it}}) + f_4(H_{xp_{it}}) + \alpha_1 T_{xit} + \alpha_2 T_{nit} + \alpha_3 T_{np_{it}} + f_{spat}(p_i) + \varepsilon_{it} \quad (2)$$

Here, as above, f_1, \dots, f_4 are nonlinear smooth functions of the metrical continuous covariates and f_{spat} is the effect of the spatial covariate p_i , ($i = 1, \dots, 17$) representing province i . The spatial effect f_{spat} is then split up into correlated (structured) and uncorrelated (unstructured or random) effects as follows [30,31].

$$f_{spat}(p_i) = f_{str}(p_i) + f_{unstr}(p_i) \quad (3)$$

The logic behind this is that a spatial effect is usually a combination of many unobserved influences, some of them obeying a strong spatial structure and others being present only locally [26-31,33]. Eq. (2) is then written as

$$\eta_{it} = \alpha_0 + f_1(R_{nit}) + f_2(H_{xit}) + f_3(T_{xp_{it}}) + f_4(H_{xp_{it}}) + \alpha_1 T_{xit} + \alpha_2 T_{nit} + \alpha_3 T_{np_{it}} + f_{str}(p_i) + f_{unstr}(p_i) + \varepsilon_{it} \quad (4)$$

This geo-additive model assumes that the nonlinear effects f_1, \dots, f_4 are the same for all provinces.

Prior assumptions and inference

For Bayesian inference, the unknown functions f_1, \dots, f_4 in predictor (4), the vector of the linear effects parameter $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, are considered as random variables and are supplemented by prior assumptions. In the absence of any prior knowledge, diffuse priors are the appropriate choice for fixed effects parameters, i.e. $p(\alpha_i) \propto \text{const}$ [32,34,35]. Another common choice are highly dispersed Gaussian priors [31].

For the continuous (smooth) functions f_1, \dots, f_4 , a second order random walk prior is considered for f defined as follows. Consider the case of a metrical covariate x with equally spaced observations x_i , $i = 1, \dots, m$, $m \leq n$ (n is the number of observations). Suppose that $x_{(1)} < \dots < x_{(t)} < \dots < x_{(m)}$ is an ordered sequence of distinct values for a covariate and define $f(t) = f(x_{(t)})$. The second order random walk is then defined by

$$f(t) = 2f(t-1) - f(t-2) + u(t) \quad (5)$$

with Gaussian errors $u(t) \sim N(0, \tau^2)$ and diffuse priors $f(1) \propto C^{st}$ and $f(2) \propto C^{st}$, for initial values. A second order random walk penalizes deviations from the linear trend $2f(t-1) - f(t-2)$ [33,36,37]. For the spatially correlated effect f_{str} , Markov random field prior is chosen [32,38]. This prior indicates spatial neighborhood relationship. For geographical data, a common assumption is that two sites or regions r_1 and r_2 are neighbors if they have a common boundary [25-32]. Thus, a spatial extension of the random walk model leads to the following conditional spatially autoregressive specification [25-32]

$$f_{str}(p)/f_{str}(p'), p' \neq p, \tau_{str}^2 \sim N\left(\frac{1}{N_s} \sum_{p' \in \partial p} f_{str}(p'), \frac{\tau_{str}^2}{N_s}\right) \quad (6)$$

Here N_s is the number of adjacent provinces and $p' \in p$ denotes that province p' is a neighbour of province p . The prior is called a Markov random field (MRF) [31,32,38]. We define provinces as neighbours if they share the same boundary and assume that the effect of a province p is conditionally Gaussian with expectation equals to the mean of the effects of neighbouring

provinces and a variance that is inversely proportional to the number of its neighbours N_s [26,31]. The conditional mean of $f_{str}(p)$ is an unweighted average of function evaluations of neighbouring provinces. For the spatially uncorrelated (unstructured) effect, f_{unstr} are assumed to be i.i.d. Gaussian (this is a common assumption) [26-31]:

$$f_{unstr}(s)/\tau_{unstr}^2 \sim N(0, \tau_{unstr}^2) \quad (7)$$

The variance parameters $\tau_j^2, j = 1, \dots, 4, str, unstr$ control the trade-off between flexibility and smoothness [36,37]. They are also considered as unknown and estimated simultaneously with corresponding unknown functions f_j . Weakly informative inverse Gamma hyperprior $\tau_j^2 \sim IG(a_j, b_j)$ are assigned to τ_j^2 . The corresponding probability density function is given by [39].

$$p(\tau_j^2) \propto (\tau_j^2)^{-a_j-1} \exp\left(-\frac{b_j}{\tau_j^2}\right) \quad (8)$$

Using proper priors for τ_j^2 ($a_j > 0$ and $b_j > 0$) ensures propriety of the joint posterior [39].

Bayesian inference is based on the posterior of the model and is carried out using MCMC simulation techniques. For the predictor (4), let γ denotes the vector of all unknown parameters in the model. Then, under conditional independence assumptions, the posterior of the model is given by [26-31].

$$p(\gamma/\gamma) \propto \prod_{i=1}^n L_i(y_i, \eta_i) \prod_{j=1}^4 \left\{ p(f_j/\tau_j^2) p(\tau_j^2) \right\} \prod_{i=0}^3 p(\alpha_i) p(f_{str}/\tau_{str}^2) p(f_{unstr}/\tau_{unstr}^2) \quad (9)$$

The full conditionals for the parameter vectors $f_j, j = 1, \dots, 4$ as well as the full conditionals for f_{str}, f_{unstr} are multivariate Gaussian. The MCMC simulation is used for successive draw of $f_1, \dots, f_4, f_{str}, f_{unstr}, \tau_j^2, j = 1, \dots, 4$ from the full conditionals [26-31]. The model is implemented in BayesX, a public domain software for Bayesian inference in structured Additive Regression Models [40]. Only the main effects are modelled. The effects of two-factor interactions are assumed to be smaller and are omitted. The main reason is that we wish to preserve the simplicity and easy interpretation of the effects, which are often lost by including interactions [24]. The effects of the continuous covariates are modelled by cubic p-splines [41,42] with 20 equidistant knots and second order random walk penalty [36,43]. Positive hyperparameters $a = 0.0001$ and $b = 0.0005$ have been chosen for τ^2 to ensure the propriety of the posterior [39]. 12,000 iterations of the MCMC were run with a burn-in phase of 2,000 iterations. Thinning was applied to the Markov

Chain to reduce autocorrelations, by requiring the programme to store only every 10th sampled parameter. Single block updating scheme is adopted, with inverse weighted least square (IWLS) proposal [35,37]. Sensitivity of the results with respect to changes in the hyperparameters a and b was checked. The model was then re-estimated with different choices for the hyperparameters a and b for each effect in the model by $(a = 1, b = 0.005)$; $(1 = 0.001, b = 0.001)$; $(a = 0.001, b = 0.005)$; $(a = 0.001, b = 0.005)$ $(a = 0.0001, b = 0.0001)$; $(a = 0.001, b = 0.0005)$ to assess the dependence of results on minor changes in the model assumptions. The results showed any significant change.

Results and discussion

The aim in this study is to analyse the dependence of malaria cases on factors, such as climatic variables and spatial (correlated and uncorrelated) effects in Burundi. Table 1 presents the estimate of the linear effects parameters.

In Table 1. α_0, α_2 and α_3 have a positive mean. α_0 and α_3 have a positive credible interval (CI). α_1 has a negative mean with a negative 95% credible interval (CI). These results suggest that malaria incidence in a given month is positively associated with the minimum temperature of the same month and more strongly with the minimum temperature of the previous month. In contrast, the results suggest that malaria incidence in a given month is negatively associated with maximum temperature of the same month. α_0 (the intercept) has the largest value, suggesting that unmeasured covariates have larger effect on malaria incidence. Figure 1 presents the nonlinear effects in model (4), with 95% credible interval. The upper-left plot of Figure 1 suggests that malaria incidence in a given month is negatively associated with rainfall of the same month. The above results may be explained as follows. Minimum temperature is the most influential factor of malaria incidence as it is observed at night and mosquitoes are active only at night; by day time they hide themselves in houses or vegetation. Moreover, when the night temperature is high, people do not cover themselves, increasing the risk of being bitten by Mosquitoes. Furthermore, due to the development cycle of the parasite into mosquitoes and the incubation period, those who became ill in a given month were bitten by mosquitoes in the previous

Table 1 Estimate of the linear effects parameters of model (4)

Parameter	Mean	Std. Dev.	Median	95% Credible Interval(CI)
α_0	0.8470	0.0551	0.8482	[0.7410, 0.9586]
α_1	-0.0303	0.0134	-0.0300	[-0.0563, -0.0019]
α_2	0.0144	0.0156	0.0140	[-0.0152, 0.0463]
α_3	0.0595	0.0142	0.0591	[0.0323, 0.0873]

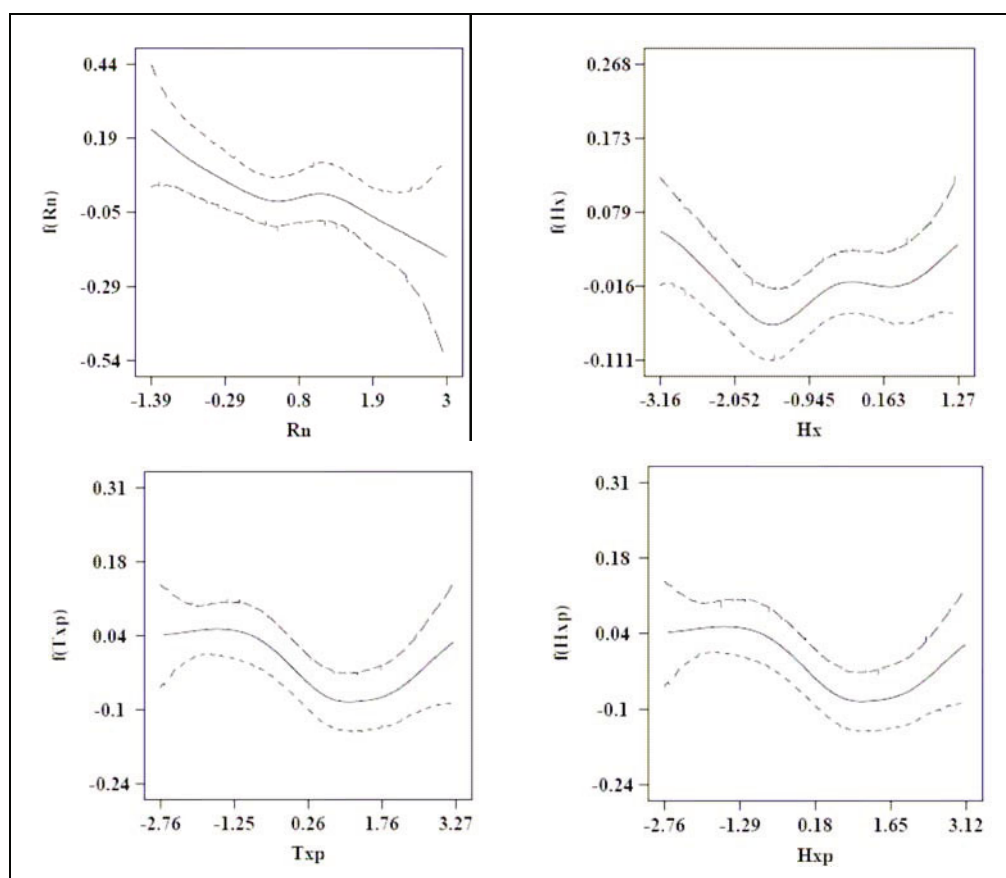


Figure 1 Nonlinear effect of the continuous covariates, with 95% credible interval.

month. This explains why malaria incidence in a given month is strongly associated with the minimum temperature of the previous month.

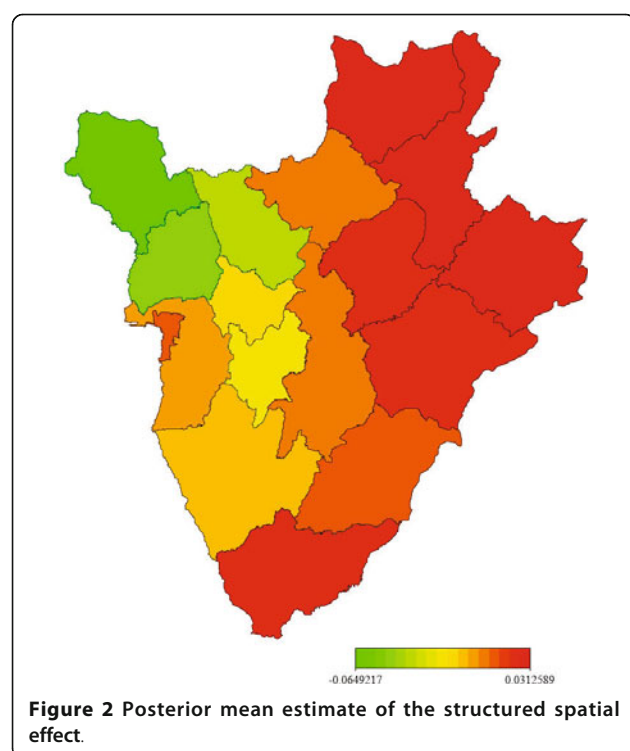
In contrast, the maximum temperature has a negative effect because mosquito's development is interrupted at higher temperature [44]. Too much rainfall may flush away the breeding larvae, decreasing the number of mosquitoes. Figures 2 and 3 show distinct spatial patterns that point to the influence of variables other than climate on malaria.

Figure 2 presents the posterior mean estimates of the structured smooth spatial component f_{str} . The map shows two main patterns: the western part, less affected by structured effect and the eastern part displaying a high risk of structured spatial effect. Figure 3 displays the posterior mean estimates of the unstructured (random) component f_{unstr} . The map shows similar trend as in Figure 2, but two provinces (Bujumbura Rural and Gitega) seem to present higher risk than others. This is probably because those provinces have a high population density, but more explanations are needed to understand the clear difference among provinces. The

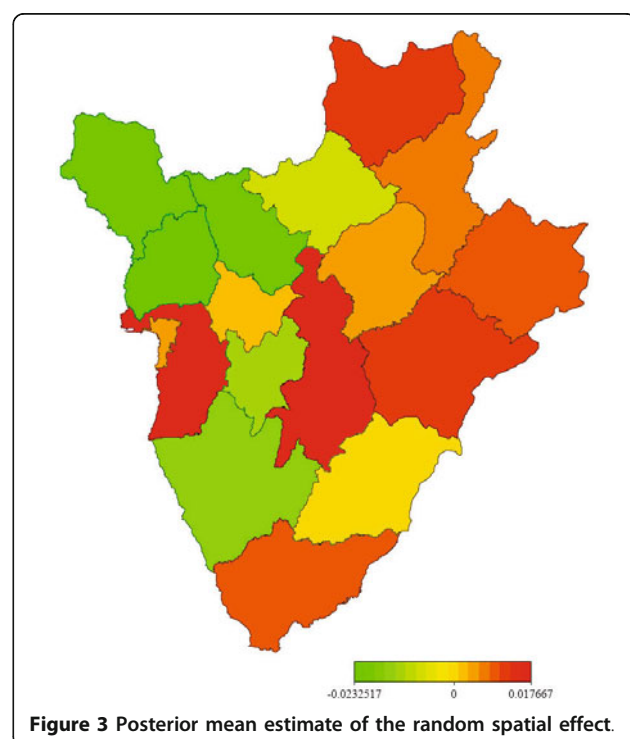
generated maps in this study could be used for targeting provinces of high risk of malaria in view to initiate control policy.

Conclusion

In this paper, semiparametric models were used to model the effects of both climatic covariates and spatial effects on malaria distribution in Burundi. The spatial analysis was based on a geo-additive model in which the province is the geographic unit of analysis. The spatial effect was split into smooth structured and unstructured (random) components. Inference was fully Bayesian and was based on Markov chain Monte Carlo techniques. The effects of climatic covariates and the effects of other spatial determinants were estimated simultaneously, in a unified regression framework. The obtained results suggest that malaria incidence in a given month is positively associated with the minimum temperature of the same and the previous months. In contrast, it is found that malaria incidence is negatively associated with rainfall and maximum temperature of the same month. From the spatial effects, important spatial patterns of malaria that are



related to factors other than climatic variables were identified without being able to explain them. Potential explanations (factors) could be related to socio-economic conditions, food shortage, limited access to health care



service, precarious housing, promiscuity, poor hygienic conditions, limited access to drinking water, land use (rice paddies for example), displaced population camps (due to armed conflicts) [6,10]. Unfortunately most of these factors are difficult to quantify in the context of poor countries like Burundi, where the record of such features is rare or nonexistent.

Acknowledgements

This study was carried out in the frame of the doctorate studies of the first author (HN). He is extremely grateful to the Austrian Agency for International Cooperation in Education and Research (ÖAD) for offering him the opportunity to study in Austria through providing him with the necessary financial support. HN would also like to acknowledge the valuable cooperation and support from the workers in Epistat and IGBU in Burundi by providing him with the data used in this study.

Author details

¹Department of Mathematics, Institute of Applied Pedagogy, University of Burundi, Burundi. ²Department of Statistics, University of Klagenfurt, Austria.

Authors' contributions

HN collected the data, performed the statistical analysis and drafted the manuscript. AG and JP reviewed and improved the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 20 January 2011 Accepted: 11 August 2011

Published: 11 August 2011

References

1. Ndayiragije A, Niyungeko D, Karenzo J, Niyungeko E, Barutwanayo M, Ciza A, Bosman A, Moyou-Somo R, Nahimana A, Nyarushatsi JP, Barihuta T, Mizerio L, Ndaruhutse J, Delacollette C, Ringwald P, Kamana J: *Efficacité de combinaisons thérapeutiques avec des dérivés de l'artémisinine dans le traitement de l'accès palustre non-complicqué au Burundi*. *Trop Med Int Health* 2004, **9**:673-679.
2. Zhou G, Minakawa N, Githeko A, Yan G: *Association between climate variability and malaria epidemics in the East African highlands*. *Proc Natl Acad Sci USA* 2004, **101**:2375-2380.
3. Pemola ND, Jauhari RK: *Climatic variables and malaria incidence in Dehradun, Uttaranchal, India*. *J Vector Borne Dis* 2006, **43**:21-28.
4. Gallup JL, Sachs JD: *The Economic burden of malaria*. *Am J Trop Med Hyg* 2001, **64**:85-96.
5. Bouma MJ, Dye C, Van der Kaay HJ: *Falciparum malaria and climate change in the north west frontier province of Pakistan*. *Am J Trop Med Hyg* 1996, **55**:131-137.
6. Cox J, Hay SI, Tarekegn AA, Checchi F, Snow RW: *The uncertain burden of Plasmodium falciparum epidemics in Africa*. *Trends Parasitol* 2007, **23**:142-148.
7. Patz JA, Lindsay SW: *New challenges, new tools: the impact of climate change on infectious diseases*. *Current Opinion in Microbiology* 1999, **2**:445-451.
8. Kigbafori DS, Raso G, Yapi A, Vounatsou P, Tanner M, N'Goran EK, Utzinger J: *Spatially-explicit risk profiling of Plasmodium falciparum infections at a small scale: a geostatistical modelling approach*. *Malar J* 2008, **7**:111.
9. Tren R: *Malaria and climate change* Working papers series, Julian Simon Centre for Policy Research; 2002.
10. Gomez-Elipe A, Otero A, Herp VM, Aguirre-Jaime A: *Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997-2003*. *Malar J* 2007, **6**:129.
11. Hay SI, Cox J, Rogers DJ, Sarah ER, Stern DI, Shanks DG, Myers MF, Snow RW: *Climate change and the resurgence of malaria in the East African highlands*. *Nature* 2002, **415**:905-909.

12. Encyclopedia of the Nations. [http://www.nationsencyclopedia.com/Africa/Burundi-CLIMATE.html].
13. WHO: *Stratégie de coopération de l'OMS avec les pays* République du Burundi 2005-2009;.
14. Ministry of Public Health in Burundi, EPISTAT. .
15. Ministry of Economy, Finance and development co-operation, ISTEEDU. .
16. Ministry of Planning and Environment in Burundi, IGEPU. .
17. Matheron G: Principles of geostatistics. *Economic and Geology* 1963, **58**:1246-1267.
18. Nkurunziza H, Gebhardt A, Pilz J: Bayesian modelling of the effect of climate on malaria in Burundi. *Malar J* 2010, **9**:114.
19. Kleinschmidt I: *Spatial statistical analysis, modelling and mapping of malaria in Africa. PhD dissertation, South Africa* 2001.
20. Hastie TJ, Tibshirani RJ: *Generalized additive models*, Chapman & Hall 1997..
21. Lang S, Brezger A: Bayesian P-splines. *J Comput Graph Stat* 2004, **13**:183-212.
22. Fahrmeir L, Kneib T, Lang S: Penalized structured additive regression for space- time data: A Bayesian perspective. *Statistica* 2004, **14**:731-761.
23. Belitz C, Lang S: Simultaneous selection of variables and smoothing parameters in structured additive regression models. *Computational Statistics and Data Analysis* 2008, **53**:61-81.
24. Fahrmeir L, Tutz G: *Multivariate Statistical Modelling Based on Generalized Linear Models*, Springer-Verlag 2001.
25. Kammann EE, Wand MP: Geoadditive models. *Applied Statistics* 2003, **52**:1-18.
26. Kneib T, Fahrmeir L: A mixed model approach for geoadditive hazard regression. *Scandinavian Journal of Statistics* 2007, **34**:207-228.
27. Kneib T: Mixed model-based inference in geoadditive hazard regression for interval- censored survival times. *Computational Statistics and Data Analysis* 2006, **5**:777-792.
28. Adebayo SB, Fahrmeir L, Klasen S: Analyzing infant mortality with geoadditive categorical regression models: a case study for Nigeria. *Economics and Human Biology* 2004, **2**:229-244.
29. Hennerfeind A, Brezger A, Fahrmeir L: Geoadditive survival models. *Journal of the American Statistical Association* 2006, **101**:1065-1075.
30. Kandala NB, Lang S, Klasen S, Fahrmeir L: Semiparametric analysis of the socio- demographic and spatial determinants of undernutrition in two African countries. *Collaborative Research Center* 2001, 386, paper 245, University of Munich.
31. Kandala NB, Fahrmeir L, Klasen S, Priebe J: Geo-additive models of childhood undernutrition in three Sub-Saharan African countries. *International Journal of Population Geography* 2008, **15**:461-473.
32. Fahrmeir L, Lang S: Bayesian inference for generalized additive mixed models based on Markov random fields. *Applied statistics* 2001, **50**:201-220.
33. Fahrmeir L, Lang S: Bayesian semiparametric regression analysis of multicategorical time-space data. *Annals of the Institute of Statistical Mathematics* 2001, **53**:11-30.
34. Brezger A, Kneib S, Lang S: BayesX: Analysing Bayesian structured additive regression models. *J Stat Software* 2005, **14**:1-22.
35. Echavarria LEO: *Semiparametric Bayesian count data models. Dissertation Ludwig-Maximilians-University of Munich*; 2004.
36. Fahrmeir L, Kneib T, Lang S: Penalized structured additive regression for space- time data: A Bayesian perspective. *Statistica Sinica* 2004, **14**:731-761.
37. Brezger A, Lang S: Generalized structured additive regression based on Bayesian P- splines. *Computational Statistics & Data Analysis* 2006, **50**:967-991.
38. Rue H, Held L: *Gaussian Markov Random Fields: Theory and Applications* Chapman and Hall/CRC; 2005.
39. Fahrmeir L, Kneib T: Propriety of posteriors in structured additive regression models: Theory and empirical evidence. *Journal of Statistical Planning and Inference* 2009, **139**:843-859.
40. [http://www.stat.uni-muenchen.de/~bayesx/bayesx.html].
41. Brezger A: *Bayesian P-splines in structured additive regression models. PhD Dissertation University of Munich*; 2004.
42. Kneib T: *Mixed model based inference in structured additive regression. PhD Dissertation University of Munich*; 2005.
43. Kneib T, Fahrmeir L: Structured additive regression for categorical space-time data: A mixed model approach. *Biometrics* 2006, **62**:109-118.
44. Sachs J, Malaney P: The economic and social burden of malaria. *Nature* 2002, **415**:680-685.

doi:10.1186/1475-2875-10-234

Cite this article as: Nkurunziza et al.: Geo-additive modelling of malaria in Burundi. *Malaria Journal* 2011 **10**:234.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

