Research Report

Department of Statistics

No. 2005:1

A Geo-additive Bayesian Discrete-time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi

Gebrenegus Ghilagaber
Ngianga Bakwin Kandala

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

A Geo-additive Bayesian Discrete-time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi

Gebrenegus Ghilagaber *and Ngianga-Bakwin Kandala †

Abstract

We describe a flexible geo-additive Bayesian survival model that controls, simultaneously, for spatial dependence and possible nonlinear or time-varying effects of other variables. Inference is fully Bayesian and is based on recently developed Markov Chain Monte Carlo techniques. In illustrating the model we introduce a spatial dimension in modelling under-five mortality among Malawian children using data from Malawi Demographic and Health Survey of 2000. The results show that district-level socioeconomic characteristics are important determinants of childhood mortality. More importantly, a separate spatial process produces district clustering of childhood mortality indicating the importance of spatial effects. The visual nature of the maps presented in this paper highlights relationships that would, otherwise, be overlooked in standard methods.

Keywords: Under-five mortality; Geo-additive models; Discrete-time survival models; Bayesian inference; Markov Chain Monte Carlo (MCMC); Time-varying effects; Spatial modelling.

*Author for correspondence: Gebrenegus Ghilagaber, Associate Professor, Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden. Fax. 46-8-167511, E-mail: Gebre@stat.su.se

†Department of HIV and GU Medicine, King's College London School of Medicine at Guy's, King's College and St Thomas' Hospitals. London SE5 9RJ, UK.

1 Introduction

Investigations on trends in, patterns of, and associations to childhood mortality rates are worthwhile efforts because mortality in childhood is a sensitive indicator of the quality of life in society (WHO, 1998). Since causes of childhood mortality are multifaceted and may operate in many complex ways, appropriate methodology that address these complexities are called for. In particular, socioeconomic and demographic patterns of child mortality vary a great deal from place to place and over time. Standard approaches such as correlation coefficients and regression analysis may produce summary statistics and measures of association at one particular site. But it cannot be assumed that these relationships hold everywhere within a country. As we expand our domain from biological factors to exogenous factors that work at household or community levels, there will be more and more variability in the phenomenon under investigation. This, in turn, prompts us to question the value of a single framework describing the pathways of determinants of child mortality that is universal for a given country.

Previous works on childhood mortality have been limited to examining socioeconomic, demographic and health-related determinants in specific contexts but have generally failed to incorporate spatial aspects. Mosley and Chen (1984), Boerma and Bicego (1992), Madise and Diamond (1995), Curtis and Steele (1996), Guilkey and Riphahn (1997), Defo and Khassoum (2002), and Berger, Fahmeir, and Klasen (2002), are few examples.

In these and other investigations that ignore the spatial dimension in the study of childhood mortality, population-level socioeconomic variables and health resources have explained very little of the variation in mortality rates. On the other hand, it is well documented that aggregate levels of mortality in many developing countries mask spatial variations. For instance, results from the Malawi Demographic and Health Survey show that the national level under-five mortality for the period 1996-2000 was 204 deaths per 1000 live-births, while the

corresponding figures for the Northern, Central, and Southern regions were 166, 204, and 212, respectively. A further stratification by districts, as shown in Table 1, reveals wide differences between districts within the same region.

Table 1 Under-five Mortality Rates by District during the period 1996-2000, that is 0 - 4 years prior to the survey, MDHS 2000[a]

Region	District no.	District name	Mortality Rate (per 1000)
North		All	166
	1	Chitipa	160
	2	Karonga	130
	3	Nkhata Bay	251
	4	Rumphi	196
	5	Mzimba	185
Central		All	204
	10	Kasungu	181
	11	Nkhota Kota	267
	12	Ntchisi	222
	13	Dowa	218
	14	Salima	255
	15	Lilongwe	186
	16	Mchinji	227
	17	Dedza	197
South		All	212
	18	Ntcheu	186
	20	Mangochi	195
	21	Machinga	195
	22	Zomba	181
	23	Chiradzulu	233
	24	Blantyre	205
	25	Mwanza	269
	26	Thyolo	237
	27	Mulanje	246
	28	Chikwawa	188
	29	Nsanje	368
	31	Balaka	186
	32	Phalombe	257
	Malawi	All districts	204

[a] The figures for Nkhata Bay, Rumphi, and Nsanje refer to 1986-90 (10-14 years before the survey)

Thus, the present paper's intended contribution to the literature is to account, simultaneously, for spatial and time-varying effects on childhood mortality and,

thereby, highlight relationships that would be overlooked in standard methods that fail to take due account of such effects. We achieve this task by introducing geo-additive Bayesian model, with dynamic and spatial effects, in discrete-time survival data in assessing temporal and spatial variation in the determinants of childhood mortality. The impact of some factors on child survival is allowed to vary over time. Our model also allows for nonlinear effects of some covariates on child survival. Appropriate smoothness priors for spatial and nonlinear effects are introduced and recently developed Markov chain Monte Carlo simulation techniques are utilized in the computation. The models are then used to examine spatial variation in under-five mortality rates in Malawi and explore district-level clustering of mortality rates across both space and time.

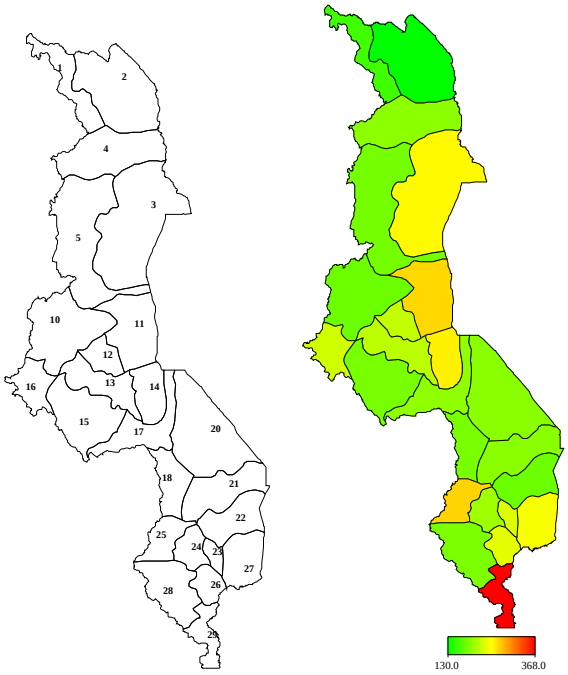

Figure 1: Physical map of Malawi showing districts (left) and spatial distribution of Under-five Mortality Rates by District during the period 1996-2000 as shown in Table 1 (right).

We present the Geo-additive Bayesian model for discrete-time - and highlight its advantages over conventional models in the next Section. In Section 3, the

model is illustrated using data obtained from the Malawi Demographic and Health Survey of 2000. The contents of the paper are tied up together in the last section by way of summary and concluding remarks, while some of our empirical results are presented in an Appendix at the end of the paper.

2 Geo-additive Bayesian Discrete-time Survival Model

2.1 The basic model

Let T denote a discrete survival time where $t \in \{1, \ldots, q+1\}$ represents the t^{th} month after birth and let $x_t^\star = (x_1, \ldots, x_t)$ denote the history of a covariate up to month t.

The discrete-time conditional probability of death at month t is then given by:

$$\lambda(t, x_t^\star) = pr(T = t | T \geq t, x_t^\star), t = 1, \ldots, q. \tag{2.1}$$

Survival information is recorded by $(t_i, \delta_i), i \in \{1, \ldots, N\}$, where $t_i \in \{1, \ldots, 60\}$ denotes the child's observed survival time in months, and δ_i is a censoring indicator with value 1 if child i died, and 0 if it is still alive. In other words, t_i represents either the age (in months) of the child at time of death (when $\delta_i = 1$), or (when $\delta_i = 0$) the current age of the child (in months) at date of interview.

We assume noninformative censoring in the sense of Lagakos (1979), so that the risk set R_t includes all individuals who are censored in interval ending in t.

Let us now defines a binary event indicator $y_{it} \quad \{i \in R_t, t = 1, \ldots, t_i\}$:

$$y_{it} = \begin{cases} 1 & \text{if } t = t_i \text{ and } \delta_i = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

The death process of individual i can, then, be considered as a sequence of binary 'outcomes' - dying at age t ($y_{it} = 1$) or surviving beyond age t ($y_{it} = 0$). Such formulation yields a sequence of 0s and 1s indicating survival histories of each child at the various time points.

2.2 Incorporating Fixed-, Time-varying and Spacial-Effects

Parallel with the sequence of 0s and 1s, we can also have records on values of relevant explanatory variables $x_{it}^{\star} = (x_{i1}, \ldots, x_{it}), i = 1, 2, \ldots$. These variables may be fixed over time such as sex, place of residence; or may vary over time, such as breast-feeding of a child at time t.

The indicator y_{it} can be linked to the covariates $x_{it}^{\star}$ by an appropriate link function for binary response model such as probit, logit or multinomial link function, and a predictor $\eta_{it}(x_{it})$. Assuming that y_{it} has a binomial distribution and using a probit link function for $i \in R_t$, the probability of death for a child i is given by

$$\phi(\eta_{it}) = pr(y_{it} = 1 | x_{it}^{\star}). \tag{2.3}$$

The usual form of the predictor is

$$\eta_{it} = f_0(t) + X_{it}^{*}\beta \tag{2.4}$$

where the baseline effect $f_0(t), t = 1, 2, \ldots$ is an unknown, usually non-linear, function of t to be estimated from data and β is the vector of fixed covariate effects. In parametric framework, the baseline hazard is often modelled by a few dummy variables dividing the time-axis into a number of relatively small segments or by some low-order polynomial. In practice, however, it is difficult to correctly specify such parametric functional forms for the baseline effects in advance. Nonparametric modelling based on some qualitative smoothness restrictions offers a more flexible framework to explore unknown patterns of the baseline.

Restriction to fixed effects alone might not be adequate because, in most cases, we have covariates whose value may vary over time. The predictor in (2.4) is, therefore, extended to a more flexible semiparametric model that can accommodate time-varying effects. If we further include another term representing spatial effects, this semiparametric predictor is given by

$$\eta_{it} = f_0(t) + f_1(X) + f(t)X_{it} + f_{spat}(s_i) + X_{it}^{*}\beta. \tag{2.5}$$

Here, $f_0(t)$ is the baseline function of time and f_1 is a nonlinear effect of metrical covariate X. The effects, $f(t)$, of the covariates in X_{it} are time-varying; while X^*_{it} comprises fixed covariates whose effect is represented by the parameter vector β; and f_{spat} is the nonlinear spatial component of, say, district s ($s = 1, \ldots, S$), where the child lives. The spatial effects $f_{spat}(s_i)$ may be split-up further into spatially correlated (structured) and uncorrelated (unstructured) effects of the form $f_{str}(s_i) + f_{unstr}(s_i)$. A rational behind this is that a spatial effect is a surrogate of many unobserved influential factors, some of which may obey a strong spatial structure while others may only be present locally.

Equations (2.4) and (2.5) are the basis of our analysis and will be referred, henceforth, as constant fixed effects model and geo-additive model, respectively.

2.3 The Estimation Process

Second-order random walk priors are used to smooth the functions f_0, f_1, and f using the MCMC techniques implemented in *BayesX* (see, for instance, Fahrmeir and Lang, 2001a; b; and Brezger, Kneib and Lang, 2002).

Let $f = \{f(1), \ldots, f(m), m \leq n\}$ be a vector of corresponding function evaluations at the observed values of x. Then, the general form of the prior for f is

$$f \mid \tau^2 \propto exp\left(-\frac{1}{2\tau^2} f\prime K f\right), \tag{2.6}$$

where K is a penalty matrix that penalizes too abrupt jumps between neighboring parameters. In most cases, K is rank deficient and, hence, the prior for f is improper.

In traditional approaches the smoothing parameter is equivalent to the variance parameter τ^2 which controls the trade off between flexibility and smoothness. A highly dispersed but proper hyperprior is assigned to τ^2 in order to estimate the smoothness parameter simultaneously with f. A proper prior for τ^2 is required in order to obtain a proper posterior for f (Hobert and Casella, 1996). If we choose

an Inverse Gamma distribution with hyperparameters a and b, ($\tau^2 \sim IG(a, b)$), then, a first- and second-order random walk priors for f are defined by

$$f(t) = f(t-1) + u(t), \quad \text{and} \quad f(t) = 2f(t-1) - f(t-2) + u(t), \tag{2.7}$$

respectively, with Gaussian errors $u(t) \sim N(0; \tau^2)$ and diffuse priors $f(1) \propto$ const, or $f(1)$ and $f(2) \propto$ const, as initial values.

A first order random walk penalizes abrupt jumps $f(t) - f(t-1)$ between successive states and a second order random walk penalizes deviations from the linear trend $2f(t-1) - f(t-2)$.

The trade off between flexibility and smoothness of f is controlled by the variance parameter τ^2. The goal in our approach is to estimate the variance parameter and the smoothing function simultaneously. This is achieved by introducing an additional hyperprior for τ^2 at a further stage of the hierarchy. We choose a highly dispersed but proper Inverse Gamma prior, $p(\tau^2) \sim IG(a; b)$, with $a = 1$ and $b = 0.005$. In analogy, we also define for the overall variance σ^2 a highly dispersed Inverse Gamma prior.

For the spatially correlated or structured effect, $f_{str}(s)$, $s = 1, \ldots, S$, we choose Markov random field priors common in spatial statistics (Besag, *et al.* 1991) of the form

$$f_{str}(s) \mid f_{str}(r), r \neq s, \tau^2_{str} \sim N\left(\sum_{r \in \partial_s} f_{str}(r)/N_s, \tau^2_{str}/N_s\right), \tag{2.8}$$

where N_s is the number of adjacent regions, and $r \in \partial_s$ indicates that region r is a neighbor of region s. Thus, the conditional mean of $f_{str}(s)$ is an unweighted average of function evaluations for neighboring regions. Again the variance parameter τ^2_{str} controls the degree of smoothness.

For a spatially uncorrelated (unstructured) effect, f_{unstr}, $s = 1, \ldots, S$, common assumptions are that the parameters $f_{unstr}(s)$, are i. i. d. Gaussian:

$$f_{unstr}(s) \mid \tau^2{}_{unstr} \sim N(0, \tau^2_{unstr}). \tag{2.9}$$

In a fully Bayesian analysis, variance or smoothness parameters $\tau^2_j, j = str$, $unstr$, are also considered as unknown and estimated simultaneously with the

corresponding unknown functions f_j. Therefore, hyperpriors are assigned to them in a second stage of the hierarchy by highly dispersed Inverse Gamma distributions $p(\tau_j^2) \sim IG(a_j, b_j)$ with known hyperparameters a_j and b_j.

Standard choices for the hyperparameters are $a = 1$ and $b = 0.005$ or $a = b = 0.001$. In our illustration, however, the results are not sensitive to the choice of a and b, and the later choice is close to Jeffrey's noninformative prior.

Fully Bayesian inference is based on the posterior distribution of model parameters whose form is not known. Therefore, MCMC sampling from full conditionals for nonlinear effects, spatial effects, fixed effects and smoothing parameters is used for posterior analysis. For the nonlinear and spatial effects, we apply the sampling scheme of Iterative Weighted Least Squares (IWLS) implemented in BayesX (see Brezger, Kneib and Lang, 2002). This is an alternative to the general Metropolis-Hastings algorithms based on conditional prior proposals that was first suggested by Knorr-Held (1999) in the context of state space models as an extension to Gamerman (1997). A more detailed related work is also given in Knorr-Held and Rue (2002).

An essential task in the model-building process is the comparison of a set of plausible models, for example rating the impact of covariates and assessing if their effects are time-varying or not; or comparing geo-additive models with simpler parametric alternatives. We adopt the measure of complexity and fit suggested by Spiegelhalter et. al. *(2002)* for comparison and select the model that takes all relevant structure into account while remaining parsimonious.

The *Deviance Information Criteria* (DIC) which may be used for model comparison is defined as

$$DIC(M) = \overline{D(M)} + pD. \tag{2.10}$$

Thus, the posterior mean of the deviance $\overline{D(M)}$ is penalized by the effective number of model parameters pD. Models can be validated by analyzing the DIC, which is smaller in models with covariates of high explanatory value.

2.4 Advantages of the Geo-additive Model

There are many potential advantages of the approach described above over more conventional approaches like discrete-time Cox models with time-varying covariates and fixed or random districts effects; or standard 2-level multilevel modelling with unstructured spatial effects (Goldstein, 1999). In the conventional models, it is assumed that the random components at the contextual level (district in our case) are mutually independent. In practice, these approaches specify correlated random residuals (see, for instance, Langford et *et al.*, 1999) which is contrary to the assumption. Further, Borgoni and Billari (2003) point out that the independence assumption has an inherent problem of inconsistency. They argue that if the location of the event matters, it makes sense to assume that areas close to each other are more similar than areas that are far apart. Moreover, treating groups (in our case districts) as independent is unrealistic and lead to poor estimates of the standard errors. As Rabe-Heskesth and Everitt (2000) pointed out, standard errors for between-district factors are likely to be underestimated because we are treating observations from the same districts as independent, and thus increasing the apparent sample size. On the contrary, standard errors for within district factors are likely to be overestimated (see also Bolstad and Monda, 2001). On the other hand, Demographic and Health Survey data is based on a random sample of districts which, in turn, introduces a structured component. Such component allows us to borrow strength from neighbors in order to cope with the posterior uncertainty of the district effect and obtain estimates for areas that may have inadequate sample sizes or are not represented in the sample.

In an attempt to highlight the advantages of our approach in a spatial context and examine the potential bias incurred when ignoring the dependence between aggregated spatial areas, we shall fit several models with and without the structured and random components in our illustration below.

3 Illustration: Spatial Modelling of Under-five Mortality in Malawi

3.1 Data Set and choice of variables

The data on which our illustration is based comes from the 2000 Malawi Demographic and Health Survey (MDHS 2000). For details on the data collection, sampling procedure and summary report, see National Statistics Office (2001).

Each record represents a child born within five years before the interview date and consists of survival information and a list of covariates. Individual data records were available for 10367 children who survived their first month of life. Of these, 1559 children had died before their fifth birthday.

Our indicator variable is

$$y_{it} = \begin{cases} 1: & \text{if child } i \text{ dies in month } t \\ 0: & \text{if child } i \text{ survives beyond time } t, \end{cases} \tag{3.1}$$

On the basis of previous work, we have selected the following array of theoretically relevant variables as covariates of childhood mortality.

- *mab* mother's age at birth of the child (in years), nonlinear
- *dobt* duration of breast-feeding, time-dependent
- *dist* district in Malawi, spatial covariate
- X vector of categorical covariates including:

  child's gender (male or female)

  asset index (low, middle or higher income household),

  residence (urban or rural),

  mother's educational attainment (up to primary or secondary and higher),

  place of delivery (hospital or other),

preceding birth interval (long birth interval, $\geq$ 24 months, or short interval, $<$ 24 months),

antenatal visit during pregnancy (at least one visit, or none),

marital status of mother (single or married),

district level mortality rate per 1000 (at least 20, or less than 20),

district level total fertility rates (at least 6 children or less than 6 children per woman).

The last levels of each covariate were selected as baseline (reference) levels. Summary statistics of these covariates is shown in Table 2.

Most of these are variables that have been found to be associated with childhood mortality in previous studies of childhood mortality in developing countries. The studies include Millard (1994), Curtis and Steele (1996), Desai and Alva (1998), Macassa et al. (2003), Da Vanzo *et al. (1983),* Woldemicael (1999), Brockerhoff (1990; 1993), Brockerhoff and Derose (1996), Madise and Diamond (1995), Kandala (2002), Cleland and Sathar (1984), Koenig *et al. (1990),* Whitworth and Stephenson (2002), Geronimus and Kerenman (1993), Bicego and Ahmad (1996), Manda (1998), Sastry (1997), Timaeus *et al. (1998),* and Claeson *et al. (2000*).

Census results from Malawi also indicate geographic variation in the rates of infant and under-five mortality with highest mortality rates in the Southern region and the lowest in the Northern region. The present study's main aim is to shed light on such regional- and district- variations and advance our knowledge of district-level socio-economic and demographic determinants of under-five mortality in the context of Malawi. We have, therefore, included an indicator of geographic location (district) among our covariates.

Table 2 Descriptive statistics of covariates used in the analysis

Variable	Frequency (%)	Coding
Place of residence		
Urban	2084(17.5)	1
Rural	9842(82.5)	-1 reference category
Sex of child		
Male	5951(49.9)	1
Female	5975(50.1)	-1 reference category
Preceding birth interval		
Greater than 24 months	10412(87.3)	1
Less than 24 months	1514(12.7)	-1 reference category
Mother's age at child birth		
Less than 20 years	2617(21.9)	category 1
20-35 years	7866 (66.0)	category2
Greater than 35 years	1443(12.1)	-1 reference category
Antenatal visit		
At least one antenatal visit	11629(97.5)	1
No antenatal visit	297(2.5)	-1 reference category
Place of delivery		
Hospital	6738(56.5)	1
Other	5148(43.2)	-1 reference category
Asset index:economic status of household		
low income household	4560(38.2)	category1
Middle income household	4724(39.6)	category2
Higher income household	2262(19.0)	-1 reference category
Mother's educational attainment		
up to primary	11060(92.7)	1
secondary and higher	866 (7.3)	-1 reference category
Marital status of mother		
Single mother	1366(11.5)	1
Married	10560 (88.5)	-1 reference category
District level crude mortality rates		
Greater than 20 deaths per 1000	5313(44.55)	1
Less than 20 deaths per 1000	6613 (55.45)	-1 reference category
District level total fertility rates (TFR)		
6 or more children per women	2410(20.21)	1
Less than 6 children per women	9516 (79.99)	-1 reference category

3.2 Statistical Method

We analyzed and compared simpler parametric probit models and probit models with dynamic effects, $pr(y_{it} = 1|x_{it}^{\star}) = \phi(\eta_{it})$, for the probability of dying in

month t. In other words, we model the conditional probability of a child dying, given child's age in months, the district where the child lived before she or he died, and covariates in x above, with the following predictors:

M1: $\eta_{it} = f_0(t) + X^*_{it}\beta$

M2: $\eta_{it} = f_0(t) + f_1(mab) + f(t)X_{it} + f_{unstr}(dist) + f_{str}(dist) + X^*_{it}\beta$

The fixed effects in model $M1$ include all covariates described above with constant fixed effects. Thus, mother's age at birth was split into three categories as indicated in Table 2, and duration of breast-feeding was included as dichotomous (0, 1) variable. Model $M2$ will be superior to model $M1$ not only in terms of the DIC (as will be seen Table 4) but because it also accounts for the unobserved heterogeneity that might exist in the data, all of which cannot be captured by the covariates (see, Madise *et al.*, 1999).

The effects of $f_0(t), f_1$ and $f(t)$ are estimated using second-order random walk prior, and Marked random field priors for $f_{str}(s)$. The analysis was carried out using BayesX-version 0.9 (Brezger, Kneib and Lang, 2002), a software for Bayesian inference based on Markov Chain Monte Carlo simulation techniques. We investigated the sensitivity of the effects to choice of different priors for the nonlinear effects (P-splines) and the choice of the hyperparameter values a and b.

As in Kandala (2002), we introduced a time-varying effect for breast-feeding. Duration of breast-feeding is an internal covariate that is observed only so long as the child survives and is uncensored. It carries survival information of the corresponding child as it can never exceed its survival time.

Instead of using duration of breast-feeding in months, we generated a binary covariate process, which is equal to 1 during the months the child was breast-fed and 0 otherwise. For instance, for a child that survived only 7 months and was breast-fed in all 7 months, the duration of breast-feeding is equal to 7 while the corresponding covariate-process is equal 1 for each of these seven months, but undefined thereafter. If, on the other hand, the child survived more than 7 months but was not breast-fed after seven months for some other reasons like

illness of the mother, then duration of breast-feeding is still equal to seven, but the covariate-process is equal to 1 for each of the first seven months and equal to 0 for every month thereafter until end of the observation.

Temporal and spatial variation in the determinants of child mortality are also assessed.

Common choices for binary response models are grouped Cox model and probit or logit models. We settled on a probit model because in this case the binary response model (2.3) can be written equivalently in terms of latent Gaussian utilities which lead to very efficient estimation algorithms. Further, because survival time in the DHS data set is recorded in months and the longest observation time for this study is limited to 60 months, the data naturally contain a high amount of tied events. Thus, a probit model for discrete survival data is a reasonable choice. A constant hazard within each month is assumed.

At the exploratory stage, we fitted a probit model with constant covariate effects (M1) for the effects of breast-feeding and mother's age with a view to compare them to the dynamic probit models (M2).

3.3 Results

3.3.1 Fixed effects

Table 3 contains estimates of posterior means of the fixed effects for categorical covariates together with their standard errors and quantiles. The results suggest that boys are at higher risk of dying than girls. This is indicated by the fact that the corresponding posterior mean, 0.012, is positive and the 10% and 90% quantiles are both positive - indicating that the effect is statistically significant.

Children of mothers with high education are at lower risk of dying than those of less-educated mothers (posterior mean of "up to primary education", 0.032, is positive and both quantiles are positive and, thus, significant).

Table 3 also indicates that survival chance is associated with economic wealth of the household. Children from poor families have higher risk compared to those

from rich families (posterior mean 0.034). On the other hand, the difference in mortality risks between children from rich- and the middle-income families is not statistically significant - this being indicated by the opposite signs of the posterior quantiles. This asset index variable captures the role of economic wealth of the household and communication infrastructure. For example, ownership of a radio facilitates acquisition of child care or nutrition information allowing a more effective allocation of resources to produce child health care.

After controlling for child, household and districts characteristics, children from urban areas seem to be better off compared to their counterparts in rural areas. This may be captured by the districts effect since the boundaries of the maps do not show urban and rural areas.

The results also show that children from single mothers are at higher risk of dying than those from married mothers. Further, children born in hospitals, and those from mothers seeking antenatal care are better off that their counterparts BUT the effects are not statistically significant (10% and 90% credible intervals are in opposite directions). We suspect that this is due to selection in the propensity to make use of health inputs, and a more appropriate modelling as in Ghilagaber (2004) may be needed before we can draw valid conclusion.

The results also show that a short birth interval significantly reduces a child's chance of survival. The district-levels factors (district's crude mortality rate and district's total fertility rate taken from census data) do not show statistically significant association with mortality at this stage of the analysis (in the fixed-effects model).

3.3.2 Baseline effects

The estimated baseline effects were almost similar in models $M1$ and $M2$. The most pronounced effects of the baseline time effects on child survival occur during the first month of life (Figure 2), although the excess risk persists throughout the first 6 months period.

Figure 2 also shows that the baseline effects peak at months 24, 36, and 48. These observed peaks are caused by the large number of deaths reported at these time points. It is, therefore, plausible to suspect that this is a "heaping" effect due to incorrect reporting of large number of deaths at these ages, which would appear to reflect digit preference in reporting deaths at 2, 3 and 4 years.

3.3.3 Time-varying effects

Figure 3 shows that there is a time-varying effect of breast-feeding in Malawi. Breast-feeding is associated with lower risk of mortality in the first 7-8 months using 80% credible region. At the mean value this could be the first 10 months and months 42-60 but given the wide range of the 80% credible region at the end of observation period (mainly, due to fewer number of cases), the results for 42-60 months should be interpreted cautiously.

3.3.4 Nonlinear effects

As shown in Figure 4, survival chance of children is associated with the mother's age at birth of the child. Children from younger mothers (less than 20 years) and older mothers (more than 45 years) are at higher risk of dying compared to children from middle age group (20-35 years). Figure 4 also shows that mortality to children of mothers aged 35-40 and even 40-45 is lower than those of mothers aged 20-25.

3.3.5 Spatial effects

Posterior means of the estimated residual spatial effects are shown in Figure 6. Both maps show a strong spatial pattern. This becomes even more clear in Figure 7 which shows "probabilities maps". On a nominal level of 80% the different colors indicate to regions with high mortality-risk (dark), moderate-level

mortality-risks (grey)", and low mortality-risk (white). Although the unstructured spatial effects are statistically insignificant in terms of posterior probability maps (Figure 7 right), the maps show interesting spatial pattern and confirm our initial thought of dependence between districts. The structured spatial effects (Figure 6 left) suggest that chances of survival are better in the North (Chitipa, Rumphi and Karonga) compared to the South and Central regions of the country. The unstructured spatial effects show that children from Karonga districts in the North; and Lilongwe, Dowa, Kasungu, and Dedza districts in the Central region; as well as Chikwawa, Machinga, Zombi, Ntcheu and Chiradzulu districts in the South have better chances of survival than those in other districts of the country (Figure 6 right). What emerges from the unstructured spatial effects is that the major centers are associated with lower risk of childhood mortality compared to the rural areas, probably because urban areas tend to enjoy better access to health services. This is the case for the largest city Lilongwe. Nsanje district has the highest under-five mortality risk. A comparison between the under-five mortality rates (Figure 1) and the estimated relative risks (in the right hand of Figure 5) indicates that, after controlling for fixed effects and other factors, a clear spatial pattern of under-five mortality risk has emerged with the residual effects.

The results in the fixed parts of the model are very similar to those obtained by including both spatial components, and are therefore not reported explicitly. Instead, the posterior effects and the maps of posterior probabilities in models with both components as shown in figures 6 and 7 clearly reflect spatial heterogeneity across the country and relative homogeneity among neighboring districts.

Failure to take due account of the posterior uncertainty in the spatial location (district) would, therefore, lead to an overestimation of the precision in predicting childhood mortality risks in unsampled districts. The general interpretation of the spatial effect is, therefore, that it represents the cumulative effect of unidentified or unmeasured additional covariates that may reflect impacts of

environmental, social and even cultural factors.

4 Summary and Concluding Remarks

4.1 Summary

After controlling for the spatial dependence in the data, most of the covariates were found to have effects in the expected directions. The factors that were associated with under-five mortality in the fixed part of the model include mother's age at child birth and her educational level and marital status; household economic status; residence; length of the preceding birth interval; and sex of the child.

An interesting, but not totally surprising, finding of the present study is that children of married mothers are at lower risk of dying than those living with a single mothers. Children living with two parents may benefit from extra care of both parents. Alternatively couples may benefit from economies of scales for child care and expenditure (Kandala, 2002).

We have also established that mortality, especially during the early months of life, is sensitive to low economic status and low levels of maternal education.

The time-varying effects of breast-feeding point to the importance of breast milk of the child after birth as recommended by WHO that a child should receive exclusively breast milk after birth until 6 months of age. Results at the end of the observation period do not, however, provide any reliable information on the dynamic effects of breast-feeding (due to few cases) and should therefore, be interpreted cautiously.

The most important finding of this paper is the sizeable district-specific geographical variation in the level of under-five mortality in Malawi which need to be scrutinized in further work. Over and above the impact of the fixed effects, there appear to be negative influences on child survival in the southern region

that are spread over and affect most of the districts there. The southern districts are at a lower altitude than other parts of the country. It is likely that climatic factors and associated diseases are responsible for this pronounced district patterns. Food insecurity associated with drought and flooding in the shire valley, which is a result of hazardous effect of climate variation are among possible explanations for these negative effects. Furthermore the southern districts are among high density population areas which can affect the child's physical environment and susceptibility to infections.

The structured effects on the left panel of Figure 6 show a sizeable difference between significantly worse child survival in the central and southern districts, and significantly better survival in the northern districts (in particular Chipita, Rumphi and Karonga districts). These district patterns are similar, but not identical to analysis of poverty and deprivation (World Bank, 2000). In terms of deprivation (based on a mean score of various services), the World Bank found the southern part of the country among the worst off. Considering the neighboring relationship, the right panel of Figure 6, surprisingly include the second largest city Blantyre among worse-off district because of the negative effects of neighboring "bad" districts. While we also found the south to be among the worst off in the country, our analysis shows a clear geographic pattern with the central-southern districts being worst off and the northern districts being well-off.

The unstructured random effects (Figure 6 right) suggest a fair amount of variation over and above the structured effects. This is particularly obvious for districts near the rift valley, but also visible for Nkhata-Bay in the north and Nkhotakola in the centre of the country. This may be related to the impact of drought that has affected the agriculture activities. The unstructured random effects also separate the positive effects in the capital city (Lilongwe) despite being surrounded by areas with negative effects. Living in the capital provides access to nutrition and health care that is superior in ways that have not been captured adequately with the fixed effects.

While some of these effects have been identified in univariate analysis, this study is the first of its kind to show that these subtle influences remain in a multivariate context, controlling for a range of fixed effects and using a flexible approach to modelling these influences. The spatial effects have no causal impact but careful interpretation can identify latent and unobserved factors which directly influence mortality rates. This gee-additive semi-parametric approach thus appears to be able to discern subtle influences on under-five mortality and identifies district-level clustering of under-five mortality. It could also be of value for a flexible modelling of other determinants of survival in developing countries.

4.2 Concluding Remarks

In this study we have shown that variation in childhood survival probability in Malawi is spatially structured. It implies that adjusted mortality risks are similar among neighboring districts, which may partly be explained by general health care practices and common childhood diseases prevalence. Another possible explanation is that of the residual spatial variation induced by variation in unmeasured districts-specific characteristics. In the light of this, a simple standard 2-level model with unstructured spatial effects which assume independence among districts is bound to yield estimates that lead to incorrect conclusions with regard to the phenomena under investigation.

From a methodological point of view, different types of covariates, such as categorical covariates with fixed effects, metrical covariates with nonlinear effects and spatial covariates are all treated within the same GLM framework by assigning an appropriate prior. For planning purposes, in constructing estimates of child mortality that include small scale spatial information we suggest a straightforward idea: maps could be used for targeting development efforts at a glance, or for exploring relationships between welfare indicators and other variables. For example, a mortality map could be overlaid with maps of other types of data, say on undernutrition, poverty, agro-climatic or other environmental characteristics. The visual nature of the maps may highlight unexpected

relationships that would be overlooked in a standard analysis.

References

[1] Berger U., Fahmeir, L., Klasen, S. (2002). *Dynamic modelling of child Mortality in Developing Countries: Application for Zambia. SFB Discussion Paper 299, University of Munich.*

[2] Besag, J., York, Y. and Mollie, A. (1991). *Bayesian Image Restoration with two Applications in Spatial Statistics. Ann. Inst. Statist. Math., **43**, 1-59.*

[3] Bicego, G. and Ahmad, (1996). *Infant and child mortality. Demographic and Health Surveys, Comparative Studies No.**20** Macro Int. Inc..*

[4] Boerma J.T. and Bicego, G. (1992). *Preceding Birth Intervals and Child Survival: Searching for Pathways of Influence. Studies in Family Planning **23(4)**:243-256.*

[5] Bolstad, W. M., and Manda, S. O. (2001), Investigating child mortality in Malawi using family and community random effects: A Bayesian analysis *Journal of the American Statistical Association* **96** (453), 12-19

[6] Borgoni, R. and Billardi, F.C. (2003). *Bayesian spatial analysis of demographic survey data: An application to contraceptive use at first sexual intercourse. Demographic Research,**8**3, 61-92. .*

[7] Brezger, A., Kneib, T. and Lang, S.(2002). *BayesX-Software for Bayesian Inference Based on Markov Chain Monte Carlo Simulation Techniques (`http://www.stat.uni-muenchen.de/~lang/`.).*

[8] Brockerhoff, M. (1990). *Rural-to-urban migration and child survival in Senegal. Demography **27(4):601-116**.*

[9] Brockerhoff, M. (1993). *Child survival in big cities: are the poor disadvantaged? Working Paper **58** New York: The pop. Council, Research Division.*

[10] Brockerhoff, M. and Derose, L.F.(1996). *Child survival in East Africa: the impact of preventive health care. World Development. **24(12**: 1841-1857.*

[11] Claeson, M., Bos, E.R., Mawji, T., and Pathmanthan, I., (2000). *Reducing child mortality in India in the new millenium. Bulletin of the Wolrd Health Organisation.* ***78(10)****: 1192-1199.*

[12] Cleland, J., and Sathar, Z.A., (1984). *The effect of birth spacing on childhood mortality in Pakistan. Population studies* ***38(3)****: 401-418.*

[13] Curtis, S.L., and Steele, F., (1996). *Variations in familial neaonatal mortality risks in four countries. Journal of Biosocial Science* ***28****: 141-159.*

[14] DaVanzo, J., Butz W.P., and Habicht, J.P., (198§). *How biological and behavioural influences on mortality in Malaysia vary during the first year of life. Population studies* ***37(3)****: 381-402.*

[15] Defo B.K. and Khassoum, D., (2002) *Geography of child mortality clustering within African families. Health and Place,* ***8****: 93-117.*

[16] Desai, S., and Alva, S., (1998) *Maternal education and child health: is there a strong causal relationship? Demography****35(1)****:71-81.*

[17] Fahrmeir, L. and Lang, S. (2001a). *Bayesian Inference for Generalized Additive Mixed Models Based on Markov Random Field Priors. Applied Statistics (JRSS C),* ***50****, 201-220.*

[18] Fahrmeir, L. and Lang, S. (2001b). *Bayesian Semiparametric Regression Analysis of Multicategorical Time-Space Data. Ann. Inst. Statist. Math.,* ***53****, 11-30.*

[19] Gamerman, D. (1997). *Efficient Sampling from the Posterior Distribution in Generalized Linear Mixed Models. Statistics Computing* ***7****, 57-68.*

[20] Geronimus, A.T., and Korenman, S.,(1993). *The socio-economic costs of teenage childbearing: evidence and interpretation. Demography* ***30(2):*** *281-290.*

[21] Ghilagaber, G. (2004), *Disentangling Selection and Causality in Assessing the Contribution of Health Inputs to Child Survival: Evidence from East*

Africa. ***Research Report 2004:7****, Department of Statistics, Stockholm University.*

[22] Goldstein, H. (1999). *Multilevel statistical models. First Internet Edition. (*`http://multilevel.ioe.ac.uk/index.html`*.)*

[23] Guilkey, D. and Riphahn, R (1997). *The Determinants of Child Mortality in the Philipines: Estimation of structural Model. Journal of Development Economics, 56(2):281-305.*

[24] Hobart, J. and G. Casella (1996). *The effect of Improper Priors on Gibbs Sampling in Hierachical Linear Mixed Models. J. Am. Statist. Asso. ,* ***91****, 1461-1473.*

[25] Kandala, N., -B. (2002). *Spatial modelling of Socio-Economic and Demographic Determinants of Childhood Undernutrition and Mortality in Africa, Ph.D Thesis, University of Munich, Shaker Verlag*

[26] Knorr-Held, L. (1999). *Conditional prior proposals in Dynamic Models. Scandinavian Journal of Statistics* ***26****, 129-144.*

[27] Knorr-Held, L.and Rue, H. (2002). *On block updating in Markov random field models for disease mapping. Scandinavian Journal of Statistics, 29, 597-614, 2002.*

[28] Koenig, M.A., Phillips, J.F., Campbell, O.M., and D'Souza, S., (1990). *Birth intervals and child hood mortality in rural Bangladesh. Demography* ***27(2)****: 251-265.*

[29] Lagakos, S. W. (1979), General right censoring and its impact on the analysis of survival data. *Biometrics* **35**, 139-156.

[30] Langford, I.H.; Leyland, A.H., Rabash, J., Goldstein, H.(1999). *Multilevel modeling of the geographical distributions of diseases. Journal of Royal Statistical Society, Series A(Applied Statistics) 48:253-268.*

[31] Macassa, M., Ghilagaber, G., Bernhardt, E., Diderichsen, F., & Burström, B. (2003), Inequalities in child mortality in Mozambique: differentials by parental socioeconomic position. *Social Science and Medicine* **57** (#12), 2255-2264.

[32] Madise,N.J. and Diamond, I. (1995). *Determinants of infant mortality in Malawi: an analyssis to control for death clustering within families. Journal of Biosocial Science, vol.**27(1)**: 95-106.*

[33] Madise,N.J., Z. Matthews, and B. Margetts (1999). *Heterogeneity of Child Nutritional Status between Households: A Comparison of six Sub-Saharan African Countries. Population Studies **53**: 331-343.*

[34] Manda, S.O.M., (1998). *Unobserved family and community effects on infant mortality in Malawi. Genus **LIV(1-2)**: 141-164.*

[35] Millard, A., (1994). *A Causal model of high rates of child mortality. Social Science and Medecine **38**: 253-268.*

[36] Mosley, W.H. and Chen, L.C. (1984). *Child Survival Strategies for Research. Population Development Review 10:25-45.*

[37] National Statistical Office [Malawi] and ORC Macro. (2001). *Malawi Demographic and Health Survey 2000. Zomba, Malawi and Calverton, Maryland, USA: National Statistical Office [Malawi] and ORC Macro*

[38] Rabe-Heskesth, S. and Everitt, B., (2000). *A Handbook of Statistical Analysis using Stata: second edition, Chapman and Hall/CRC.*

[39] Sastry, N., (1997). *Family-level clustering of childhood mortality risk in Notheast Brazil. Population Studies **51**: 245-261.*

[40] Spiegelhalter D., Best N., Carlin B., and Van der Line A. (2002). *Bayesian measures of models complexity and fit. Journal of the Royal Stat. Soc. **B(64)**:1-34.*

[41] Timaeus, I., K. Harris, and F., Fairbarn. (1998). *Can use of health care explain sex differentials in child mortality in the developing world? In United Nations (ed.), Too young to die: Genes or Gender? New york: United Nations Secretariat.*

[42] UNICEF in Malawi and Government of Malawi (1993). *Situation Analysis of Poverty in Malawi. Mantfort press, Limbe.*

[43] Whitworth A., and Stephenson R., (2002). *Birth spacing, sibling rivalry and child mortality in India. Social Science and Medecine* ***55****: 2107-2119.*

[44] WHO (1998). *WHO Global Database on Child Growth and Malnutrition. Geneva: WHO.*

[45] Woldemicael, G (1999), I*nfant and child mortality in Eritrea: Levels, trends and determinants*. Doctoral Dissertation, Demography Unit, Stockholm University.

[46] World Bank (2000). *Profile of poverty in Malawi, 1998. Poverty analysis of the Malawian Integrated Household Survey, 1997-98. Washington, DC: The World Bank.*

Appendix

Table 3: Estimates of Posterior means of the fixed-effect parameters in model M2

Variable	mean	std. error	10%	90%
constant	-3.41	0.13	-3.57	-3.22
Maternal education: secon. and higher (ref. category)				
Up to primary	0.032	0.014	0.014	0.050
Sex of child: female (ref. category)				
Boy	0.012	0.007	0.004	0.020
Marital status: married (ref. category)				
Single mothers	0.025	0.011	0.011	0.039
Antenatal visit: no antenatal visit (ref. category)				
At least one antenatal visit	-0.032	0.028	-0.068	0.006
Asset index: rich (ref. category)				
Poor	0.034	0.010	0.020	0.046
Middle	-0.006	0.009	-0.018	0.006
Place of delivery: other (ref. category)				
Hospital	-0.002	0.007	-0.010	0.007
Birth interval: short birth (ref. category)				
Long birth interval	-0.052	0.010	-0.064	-0.038
Place of residence: rural (ref. category)				
Urban	-0.030	0.011	-0.045	-0.016
District level mortality rates: $<$ 20 (ref. category)				
$>$ 20 deaths per 1000	-0.009	0.016	-0.031	0.013
District level total fertility rates: $<$ 6 (ref. category)				
6 or more children per women	-0.008	0.019	-0.034	0.017

Table 4 Summary of DIC for models M1 and M2

Model	Deviance	pD	DIC
M1	611.9	26.7436	6738.6
M2	6080.8	32.4146	6113.2

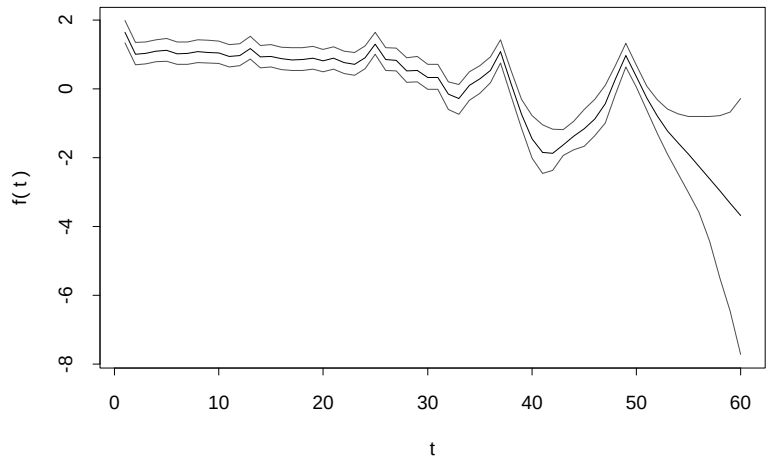

Figure 2 Estimated nonparametric effect of baseline time. Shown is the posterior mean within 80% credible regions.

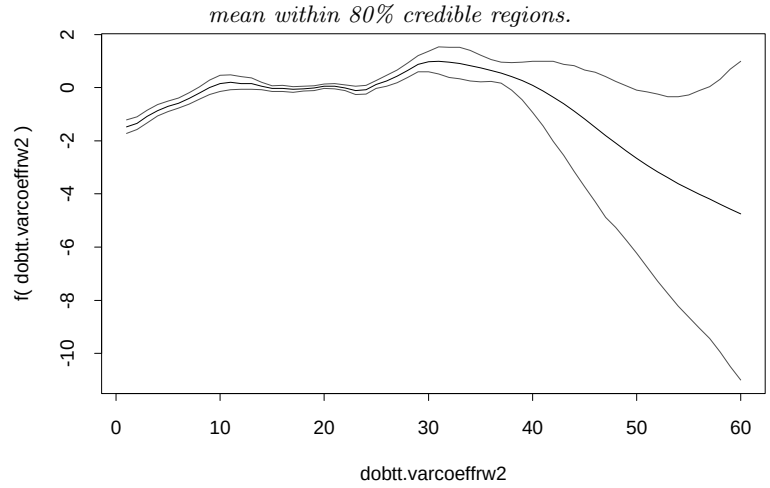

Figure 3 Estimated nonparametric effect of time-varying breastfeeding. Shown is the posterior mean within 80% credible regions.

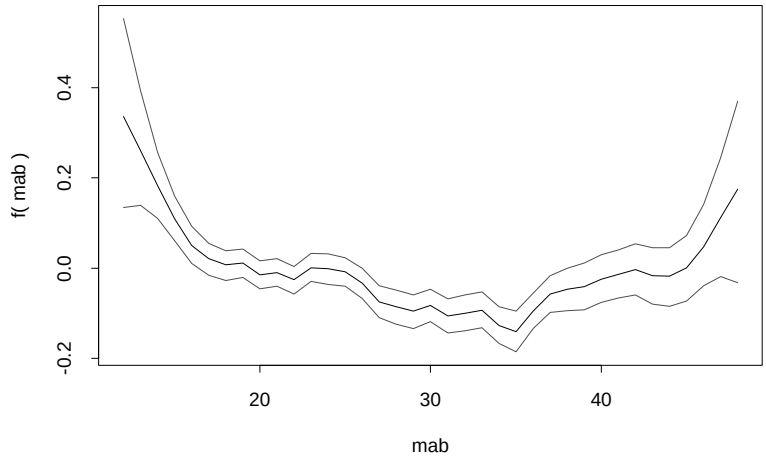

Figure 4 Estimated nonparametric effect of mother's age at child's birth. Shown is the posterior mean within 80% credible regions.

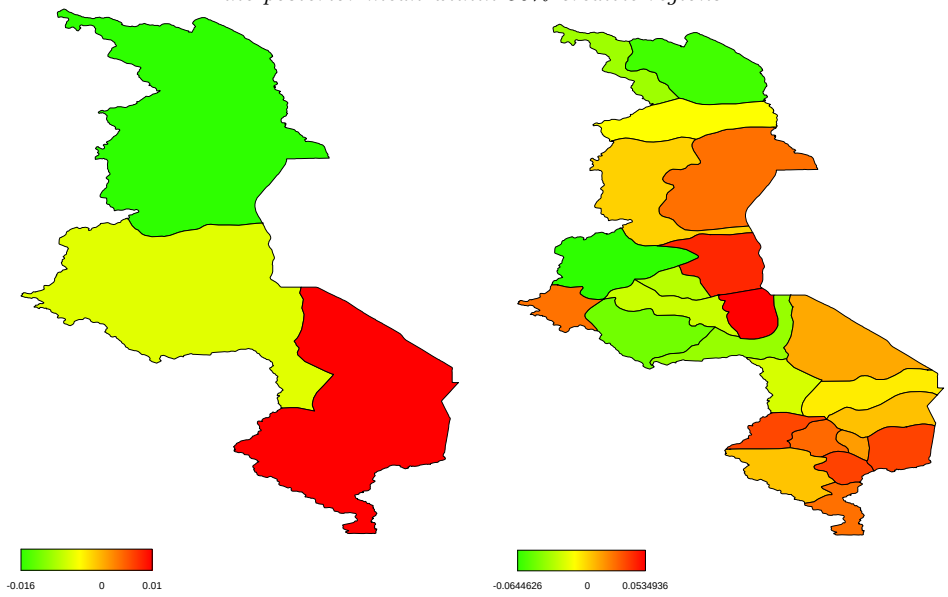

Figure 5 Regional fixed effects (left: Model M1) and total posterior mean of spatial effects (right: Model M2).

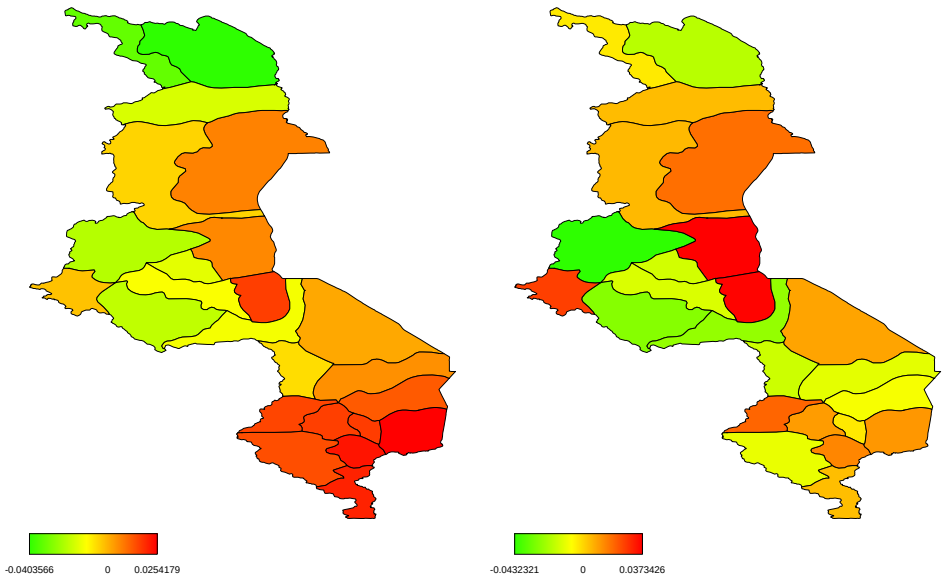

Figure 6 Posterior mean of structured (left) and unstructured (right) spatial effects (Model M2).

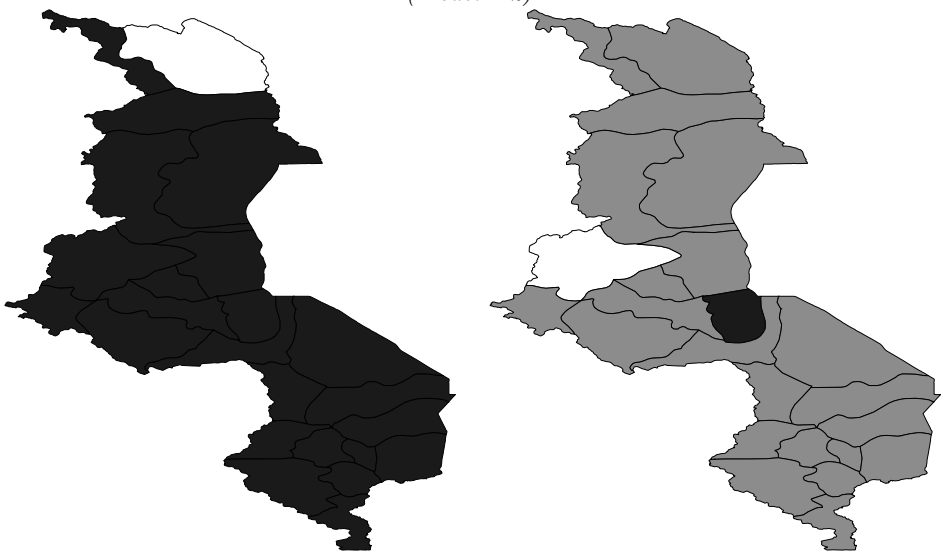

Figure 7 Maps of 80% posterior probabilities for the structured (left) and unstructured (right) spatial effects (Model M2).