Spatial Effects on Childhood Mortality in Nigeria: An application of Geo-Additive Bayesian Discrete-time Survival Model*

Gebrenegus Ghilagaber† Ngianga-Bakwin Kandala‡ Diddy Antai§

Abstract

Childhood mortality reflects the overall health and development in a country. Mortality rates in developing countries are not only influenced by socio-economic, demographic and health variables; they also vary significantly across regions and districts. This study analyzes under-five mortality in Nigeria using flexible geo-additive Bayesian survival model, which enables the measurement of small-area district-specific spatial effects simultaneously with possibly nonlinear or time-varying effects of other predictors. Inference is fully Bayesian and is based on Markov Chain Monte Carlo (MCMC) simulation. Data for the study come from the 2003 Nigeria Demographic and Health Survey (NDHS) and includes 6029 children born between 1999 and 2003. Results indicate that district-level socio-economic characteristics are important determinants of under-five mortality. More importantly, we find district clustering of under-five mortality, which indicates the importance of spatial effects. The presentation of this clustering through maps facilitates visuality and highlights differentials across geographical areas that would, otherwise, be overlooked in traditional data-analytic methods.

---

*Contribution to Session 68 (*Understanding health and population dynamics through longitudinal demographic surveillance systems*) at the 5^{th} **African Population Conference (*Emerging Issues on Population and Development in Africa*),** Arusha, Tanzania, 10-14 December 2007.

†Department of Statistics, Stockholm University, Sweden. e-mail: Gebre@stat.su.se

‡Clinical Sciences Research Inst., Warwick Medical School, University of Warwick, UK

§Div. of Social Medicine, Dept. of Pub. Health, Karolinska Inst., Stockholm, Sweden.

1 Introduction

Childhood mortality is an important indicator of overall health and development in a country. It is the result of a complex interplay of determinants at many levels. As such some studies have documented the association of maternal factors (Caldwell, 1979; Cleland & van Ginneken, 1988), socio-economic factors (Castro-Leal et al. 1999; Wagstaff, 2001), as well as environmental factors (Wolfe & Behrman, 1982; Lee et al. 1997) with childhood mortality. However, only a few studies have incorporated environmental factors that are spatial in nature and derived from geographic databases, such as distances from households or communities (Watson et al. 1997).

While the commonly used approaches, such as correlation coefficients and regression analysis may produce statistical outcomes and measures of association in a particular location, these relationships cannot be readily generalized for other locations within a country. Spatial analysis could be defined as a quantitative data analysis which focuses on the role of space and relies explicitly on spatial variables in order to explain or predict the phenomenon under investigation (Cressie, 1993; Chou, 1997). It tests theories which stress that the location of an individual influences social attitudes and behaviour, and that observed social phenomena are not distributed in a spatially random fashion (Weeks, 2004).

Studies of childhood mortality in developing countries using aggregated data and methodologies that ignore spatial dimensions run the risk of explaining very little of the variations in mortality rates as well as masking spatial variations. For instance, results of the 2003 Nigeria Demographic and Health Survey (NDHS), disaggregated by geopolitical zones, shows that the under-five mortality rate for the period 1999-2003 at the national level was 218 deaths per 1000 live-births, while the corresponding figures for the 6 geopolitical zones was 172 (North Central), 270 (North East), 264 (North West), 92 (South East), 187 (South South), and 101 (South West). Further strafication of the under-five mortality rates by districts (states) as displayed in Table 1, reveals wide variations between districts within the same geopolitical region. Such variations would be "hidden" in the overall picture of crude mortality rate for that region or states and, thus, prompt the importance of incorporating a spatial diemension in the analyis of under-five mortality.

This study is intended to account simultaneously for spatial and time-varying effects on childhood mortality by employing a geo-additive Bayesian model with dynamic and spatial extensions of discrete-time survival model.

Temporal and spatial variation in the determinants of childhood mortality, as well as any associations between risk factors and childhood mortality in the presence of spatial correlation. According to Weeks (2004), ignoring this correlation would lead to an underestimation of the variance of the effects of risk factors. The impact of some determinant factors of child survival is allowed to vary over time, as well as allowing for non-linear effects of some covariates on child survival. The model introduces appropriate smoothness priors for spatial and non-linear effects, and uses Markov Chain Monte Carlo simulation techniques (Gelfand and Smith, 1990; Smith and Roberts, 1993) to estimate the model parameters. The models are subsequently used to examine spatial variation in childhood mortality rates in Nigeria, and explore district-level clustering of mortality rates across both space and time. Figure 1 (a) shows current geopolitical districts of Nigeria. Due to lack of spatial data including for the five new states, however, this study will be based on the older 31 states (i.e. states created before 1996).

(A paragraph on the organization of the paper may be included here)

2 Study Area and Study Population

Nigeria, with a 2006 population of 140 million people, is the most populous country in Africa (Onuah, 2007). It is also the tenth largest country by population in the World. The country lies on the west coast of Africa between 4 and 14° North latitude and 2 and 15° East longitude, and is bordered by Benin, Niger, Chad, Cameroon, and the Gulf of Guinea. It has a landmass extending over 923,768 square kilometres and is located on the eastern terminus of the bulge of West Africa (Population Resource Centre, 2000). An average density of approximately 124 persons per square kilometer (Ali-Akpajiak and Pyke, 2003) makes Nigeria one of the most densely populated countries in the World. The spatial distribution of the population is uneven, with some areas of the country sparsely inhabited while other areas densely populated. With the exception of Lagos, which has the highest population density in the country, the South East of Nigeria has the highest densities. Sixty four percent of the population is concentrated in the rural areas (Ali-Akpajiak and Pyke, 2003). Nigeria is made up of 36 states (districts) and a Federal Capital Territory at Abuja. The 36 states are grouped into six geopolitical zones (regions). The mean temperature ranges between 25 and 40 °C, and rainfall ranges between 2650 mm in the Southeast and less than 600 mm in some parts of northern Nigeria that lie mainly in the Sahara desert. These climatic differences give rise to both vegetational differences ranging from mangrove swamp forest in the Niger delta and Sahel grassland in the North, and different soil conditions. This results in a variation in agricultural products and natural resources in the different parts of Nigeria. A map of Nigeria indicating the geographical location of the states (districts) is given in Figure 1.

(A paragraph that connects this section well to the next section is needed here)

3 Geo-Additive Bayesian Discrete-Time Survival Model

3.1 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:

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

Survival information is recorded by (t_i, δ_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 of the child at time of death (when $\delta_i = 1$), or the age of the child at date of interview (when $\delta_i = 0$).

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 define a binary event indicator y_{it} $\{i \in R_t, t = 1, \ldots, t_i\}$:

$$y_{it} = \begin{cases} 1 & \text{if } t = t_i \text{ and } \delta_i = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

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.

3.2 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; 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

$$\phi(\eta_{it}) = pr(y_{it} = 1 | x_{it}^{\star}). \tag{3}$$

The usual form of the predictor is

$$\eta_{it} = f_0(t) + X_{it}^{*}\beta \tag{4}$$

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 (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

$$\eta_{it} = f_0(t) + f_1(X) + f(t)X_{it} + f_{spat}(s_i) + X_{it}^{*}\beta. \tag{5}$$

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 (4) and (5) are the basis of our analysis and will be referred, henceforth, as constant fixed effects model (M_1) and geo-additive model (M_2), respectively.

3.3 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

$$f \mid \tau^2 \propto exp\left(-\frac{1}{2\tau^2} f\prime K f\right), \tag{6}$$

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

$$f(t) = f(t-1) + u(t), \quad \text{and} \quad f(t) = 2f(t-1) - f(t-2) + u(t), \tag{7}$$

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

$$f_{str}(s) \mid f_{str}(r), r \neq s, \tau^2_{str} \sim N\left(\frac{\sum_{r \in \partial_s} f_{str}(r)}{N_s}, \frac{\tau^2_{str}}{N_s}\right), \tag{8}$$

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:

$$f_{unstr}(s) \mid {\tau^2}_{unstr} \sim N(0, \tau^2_{unstr}). \tag{9}$$

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 $p(\tau^2_j) \sim IG(a_j, b_j)$ 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

$$DIC(M) = \overline{D(M)} + pD. \tag{10}$$

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.

3.4 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). Since Demographic and Health Survey data is based on a random sample of districts, they 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.

4 Illustration: Spatial Modelling of Under-five Mortality in Nigeria

4.1 Data set

Data from the 2003 Nigeria Demographic and Health Survey (NDHS) was used in this study. The survey included 7620 women aged 15-49 years, and all men aged 15-59 in a sub-sample of one-third (i.e. 2346) of the households. The data in the present study contains 6029 children born within 5 years prior to the survey (ca 1999-2003) coming from 3725 mothers who contributed between 1 child and 6 children. Technical details of the survey have been reported in the official 2003 NDHS report (NPC, 2003). Of the 6029 children 843 children (14%) died before their fifth birthday while the rest were still alive (censored) by the date of interview.

Each live birth and each subsequent child health outcome contains information on the household and each parent, thereby constituting the basic analytic sample.

4.2 Specification and measurement of variables

The indicator variable used in this study is:

$$y_{it} = \begin{cases} 1: & \text{if child } i \text{ dies in month } t \\ 0: & \text{if child } i \text{ survives beyond time } t, \end{cases} \tag{11}$$

and is linked to covariates as shown in the equations (3) - (5).

On the basis of previous studies, a selection of theoretically relevant variables was chosen as covariates of childhood mortality. These include:

- **mab**: mother's age at birth of the child (in years) – assumed to have nonlinear effect;
- **dobt**: duration of breastfeeding - assumed to have time-varying effect;
- **dist**: district (state) in Nigeria - **spatial covariate**;
- **X**: a vector of categorical covariates including:
  - *sex* of the child (male or female),
  - *asset index* (1^{st} quntile, 2^{nd} quitile, 3^{rd} quintile, 4^{th} quintile),
  - *place of residence* (urban or rural),
  - *mother's educational level* (no education, at least primary),
  - *partner's educational level* (no education, at least primary),
  - *place of delivery* (hospital, home/other),
  - *preceding birth interval* ($<$ 24 months, $\geq$ 24 months),
  - *antenatal visits during pregnancy* (at least one visit , no visit),
  - *marital status of mother* (single, married)
  - *household size* (small, medium, large)

The last levels of each covariate were selected as reference or baseline levels. Descriptive statistics of covariates used in the analysis are shown in Table 2.

4.3 Statistical method

An analysis and comparison of simpler parametric probit models, and probit models with dynamic effects, $pr(y_{it} = 1|x^{\star}_{it}) = \phi(\eta_{it})$, was made for the probability of dying in month t. In other words, the conditional probability of a child dying at time t (given the child's age in months, the district where the child lived before death, and covariates in X above), is modeled with the following predictors:

$$M_1 \ : \ \eta_{it} = f_0(t) + X^*_{it}\beta$$
$$M_2 \ : \ \eta_{it} = f_0(t) + f_1(mab) + f(t)X_{it} + f_{unstr}(dist) + f_{str}(dist) + X^*_{it}\beta$$

The fixed effects in model M_1 include all fixed covariates described above. Further, mother's age at birth (mab) was split into two categories as shown in Table 2, and duration of breastfeeding was included as dichotomous (0, 1) variable in Model 1. Model M_2 which uses mother's age at birth (mab) as non-linear and duration of breastfeeding as time-varying covariate, will be superior to model M_1 because Model M_2 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 Markov random field priors for $f_{str}(s)$. The analysis was carried out using BayesX-version 0.9 (Brezger et al. 2002), a software for Bayesian inference based on Markov Chain Monte Carlo simulation techniques. The sensitivity of the effects to choice of different priors for the non-linear effects ($p-splines$) and the choice of the hyperparameter values a and b are investigated.

Previous studies, for example, Berger et al. (2002), have shown that breastfeeding is an important factor. In order to assess its effect, a time-varying indicator variable (see Kandala, 2002), that takes the value 1 in the months a child is breastfed, and 0 otherwise, is generated. In addition, temporal and spatial variations in the determinants of child mortality are also assessed. Common choices for discrete survival models are the grouped Cox model and probit or logit models. For this study, probit model for discrete survival data is used because binary response models (see equation 3) can be written equivalently in terms of latent Gaussian utilities, which lead to very efficient estimation algorithms. In addition, since 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. A constant hazard within each month is assumed.

At the exploratory stage, a probit model with constant covariate effects (M_1) for the effects of breastfeeding and mother's age are fitted with a view to compare them to the dynamic probit models (M_2).

4.4 Results

4.4.1 Fixed effects

The estimates of posterior odds ratio of the fixed effect parameters for under-five mortality in Nigeria (Model 2) together with their standard errors and quantiles are presented in Table 3. Results indicate that children living in urban areas at lower risk of dying than children living in rural areas (posterior odds ratio 0.54 with 2.5%- and 97.5% quantiles (0.38 and 0.83, respectively) below 1 - indicating that the effect is statistically significant. The results also show that a short birth interval significantly reduces a child's chances of survival, as children with longer preceding birth interval were at lower risk of dying (posterior odds ratio 0.71). Children whose mothers had at least one antenatal visit were at lower risk of dying (posterior odds ration 0.57) and the effect being statistically significant. Lack of at least primary-level education raises mortality risk significantly (posterior odds 1.51). The other fixed variables didn't show any significant association with the risk of under-five mortality.

The children of single mothers were at higher risk of dying (posterior odds ratio 1.27) compared to children whose mothers were married; both quantiles were positive, and therefore the relationship was significant. Remarkably, the larger the household size, the lower the risk of the children dying. Children living in medium-size households (posterior odds ratio 0.99), and those living in large-size households (posterior odds ratio 0.96), were at lower risk of dying compared to children living in small-size households; both relationships had positive quantiles and were therefore significant.

4.4.2 Baseline effects

The estimated nonlinear effect of child's age (baseline time) obtained from the Bayesian *p-splines* are shown in Figure 2(a). The posterior means of log-odds are presented within 80% & 95% credible intervals, and show that starting from a comparably high level in the first month, the baseline effect remains more or less high until about three years (36 months) and declines thereafter. The peaks at months 24 and 36 may reflect a "heaping" effect from the large number of deaths being reported at these times.

4.4.3 Time-varying effects

Figure 2b) displays the time-varying effect of breastfeeding, and indicates that breastfeeding is on average associated with lower risk of mortality within the first 2 years (24 months). Given the wide range of the 80% & 95% credible region at the end of the observation period (most likely due to fewer numbers of cases), the results beyond 24 months should be interpreted with caution.

4.4.4 Nonlinear effects

Figure 2c) shows the non-linear time-varying effect of mother's age at birth of the child. Mortality risk increases with age of mother and more strongly so after age 35.

4.4.5 Spatial effects

Posterior means of the estimated residual spatial states-effects on under-five mortality are presented in Figure 1c). This map shows a strong spatial pattern, which suggests that survival chances of children under-five years of age are highest within the North Western (Sokoto, and Kebbi) and South Western (Lagos) regions compared to the other regions. On the other hand, the survival chances of children under-five years are lowest among children from Jigawa, Taraba, Delta, Rivers and Adamawa states compared to the children from the rest of the states. A comparison between the under-five mortality rates (Table 1 and Figure 1b)) and the estimated odds ratio (Figure 1c)) reveals the emergence of a clear spatial pattern of under-five mortality risk. These spatial effects could therefore be interpreted as representing the cumulative effect of unidentified or unmeasured additional covariates that may reflect impacts of environmental and socio-cultural factors. Thus, failure to take into consideration the posterior uncertainty in the spatial location (states or districts) would invariably lead to an overestimation of the precision in predicting childhood mortality risks in unsampled districts.

5 Discussion and Conclusion

Available statistics suggest that child mortality levels in Nigeria exhibit wide geographic disparities (NPC, 2000; NPC, 2004), with the northern regions and rural areas generally having higher childhood mortality rates compared to the southern regions and urban areas respectively. While the focus of previous studies in Nigeria have mainly been on effect of individual and household factors in explaining childhood mortality differences in the country, they have largely neglected the impact of small area variations and community-level variables (see Iyun, 1992; Adetunji, 1994; Folasade, 2000; NPC, 2004).

The aim of the present study was to highlight the regional- and district-level variations in under-five mortality in Nigeria, while improving current knowledge of district-level socio-economic and demographic determinants (thereby warranting the inclusion of a geographic location [districts] covariate).

After controlling for the spatial dependence in the data, most of the covariates associated with under-five mortality in the fixed part of the model were found to have effects in the expected directions.

The time-varying effects of breastfeeding emphasize the importance of breastfeeding, which is widely believed to be the most beneficial source of infant nutrition for the attainment of health and well-being of the infant (Weimer, 2001). Results of this study show a lowered risk of mortality associated with breastfeeding within the first 2 years. The results for the period after 2 years do not provide reliable information on the dynamic effect of breastfeeding due, mainly, to few cases. Results of the nonlinear effect of mother's age at the birth of the child are in the expected direction, emphasizing the risk associated with younger motherhood (also seen in Alam, 200) and childbirth at older ages (see Hobcraft et al., 1985).

The estimated residual spatial effects for under-five mortality in Figure 1c) show clear differences between the significantly better survival chances of children in the North West (Sokoto, and Kebbi) and South West (Lagos) regions compared to the North East (Adamawa, Taraba, Yobe, Borno), South South (Delta, Rivers, Akwa Ibom) and South East (Enugu) regions. These state patterns are similar to analysis of poverty in Nigeria in which the Northeast zone had the highest poverty incidence with 67.3 per cent, followed by the Northwest with 63.9 per cent; the South South zone had the highest poverty rates (55 percent) among the southern states, while the lowest poverty rates were recorded in the South East at 34.2 per cent, followed by Southwest with 43.0 per cent (National Bureau of Statistics, 2005).

While some of these effects have been shown using traditional parametric methods, using Bayesian geo-additive models uniquely shows subtle differences when examining small-area spatial effects. Though the spatial effects do not show causality, careful interpretation could identify latent and unobserved factors that directly influence mortality rates. This geographic semiparametric approach therefore appears to be able to discern subtle influences of the determinants, and identifies district-level clustering of under-five mortality.

The variation in the probability of childhood survival in Nigeria is spatially structured. This implies that adjusted mortality risks are similar among neighbouring states or districts, which may partly be explained by general health care practices, similar prevalence of common childhood diseases, and the residual spatial variation induced by variation in unmeasured district-specific characteristics.

It is also our hope that the results presented here assist policy makers in evaluating and designing programme strategies needed to improve child health services, and reduce childhood mortality levels in Nigeria.

References

[1] Adetunji JA. Infant mortality in Nigeria: effects of place of birth, mother's education and region of residence. J Biosoc Sci. 1994; 26(4): 469-77.

[2] Alam N. teenage motherhood and infant mortality in Bangladesh: Maternal age-dependent effect of party one. Journal of biosocial Science, 2000, 32: 229–236.

[3] Ali-Akpajiak SCA, Pyke T. Measuring Poverty in Nigeria. Oxfam Working Papers, Oxfam, 2003.

[4] Berger U, Fahrmeir L, Klasen S. (2002): Dynamic Modelling of Child Mortality in Developing Countries: Application for Zambia. Sonderforschungsbereich 386, Discussion Paper 299, University of Munich, Germany, 2002.

[5] Besag J, York Y, Mollie A. Bayesian image restoration with two applications in spatial statistics. Annual Institute of Statistical Mathematics 1991; 43: 1–59.

[6] Bolstad WM, Manda SO. Investigating child mortality in Malawi using family and community random effects: A Bayesian analysis Journal of the American Statistical Association 2001; 96(453): 12-19.

[7] Borgoni R, Billardi FC. Bayesian spatial analysis of demographic survey data: An application to contraceptive use at first sexual intercourse. Demographic Research 2003; 83: 61–92.

[8] Brezger A, Kneib T, Lang S. BayesX-Software for Bayesian Inference Based on Markov Chain Monte Carlo Simulation Techniques, 2002. Available at: http://www.stat.uni-muenchen.de/∼lang/.

[9] Caldwell JC. Education as a Factor in Mortality Decline an Examination of Nigerian Data. Population Studies 1979; 33(3): 395-413.

[10] Castro-Leal F, Dayton J, Demery L, Mehra K. Public social spending in Africa: Do the poor benefit? World Bank Res Obs. 1999; 14:49-72.

[11] Chou Y-H. Exploring Spatial Analysis in Geographic Information Systems. Santa Fe, NM: OnWard Press, 1997.

[12] Cleland JG, van Ginneken JK. Maternal education and child survival in developing countries: the search for pathways of influence. Soc Sci Med. 1988; 27(12):1357-1368.

[13] Cressie NAC. Statistics for Spatial Data: Revised Edition. New York: John Wiley & Sons, 1993.

[14] Fahrmeir L, Lang, S. Bayesian Inference for generalized additive mixed models based on Markov random field priors. Applied Statistics 2001a (JRSS C) 50: 201–220.

[15] Fahrmeir L, Lang S. Bayesian semi-parametric regression analysis of multi-categorical time-space data. Annual Institute of Statistical Mathematics 2001b; 53: 11–30.

[16] Folasade IB. Environmental factors, situation of women and child mortality in southwestern Nigeria. Soc Sci Med. 2000; 51(10): 1473-89.

[17] Gamerman D. Efficient sampling from the posterior distribution in generalized linear mixed models. Statistics Computing 1997; 7: 57–68.

[18] Gelfand AE, Smith AFR. Sampling-based approaches to calculating marginal densities. J Am Stat Assoc. 1990; 85: 398-409.

[19] Goldstein H. Multilevel statistical models. London: Institute of Education, Multilevel Models Project, April, 1999.

[20] Hobcraft JN, McDonald JW, Rutstein SO. Demographic Determinants of Infant and Early Child Mortality: A Comparative Analysis. Population Studies, 1985; 39(3): 363-385.

[21] Hobart J, Casella G. The effect of improper priors on gibbs sampling in hierarchical linear mixed models. Journal of the American Statistical Association 1996; 91: 1461-1473.

[22] Iyun BF. Women's status and childhood mortality in two contrasting areas in south-western Nigeria: a preliminary analysis. GeoJournal. 1992; 26(1): 43-52.

[23] Kandala N-B. Spatial modelling of socio-economic and demographic determinants of childhood undernutrition and mortality in Africa. Ph.D Thesis, University of Munich, Shaker Verlag, 2002.

[24] Kandala N-B., and Ghilagaber, G (2006), A Geo-additive Bayesian discrete-time survival model and its application to spatial analysis of childhood mortality in Malawi. *Quality and Quantity* **40**, 935-957.

[25] Knorr-Held L. (1999). Conditional prior proposals in dynamic models. Scandinavian Journal of Statistics 1999; 26: 129–144.

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

[27] Kuate Defo B. Areal and socioeconomic differentials in infant and child mortality in Cameroon. Social Science and Medicine, 1996; 42: 399-420.

[28] Lagakos SW. General right censoring and its impact on the analysis of survival data. Biometrics 1979; 35-139-156.

[29] Langford IH, Leyland AH, Rabash J, Goldstein H. Multilevel modeling of the geographical distributions of diseases. Journal of Royal Statistical Society, Series A (Applied´ Statistics) 1999; 48: 253–268.

[30] Lee L-F, Rosenzweig M, Pitt M. The effects of improved nutrition, sanitation and water quality on child health in high-mortality populations. J of Econometrics 1997; 77:209-35.

[31] Madise NJ, Matthews Z, Margetts B. Heterogeneity of child nutritional status between households: A comparison of six sub-Saharan African countries. Population Studies. 1999; 53: 331-343.

[32] Madise NJ, Diamond I. Determinants of infant mortality in Malawi: an analysis to control for death clustering within families. Journal of biosocial Science, 1995; 27: 95-106.

[33] National Population Commission (NPC) [Nigeria]. Nigeria Demographic and Health Survey 2003. Calverton, Maryland: National Population Commission and ORC/Macro, 2004.

[34] Onuah F. Nigeria Gives Census Result, Avoids Risky Details. Reuters, Dec. 30, 2006, accessed online at http://za.today.reuters.com, on March 30, 2007.

[35] National Bureau of Statistics (NBS). Poverty Profile for Nigeria Report (1980-1996). Federal Republic of Nigeria, 2005

[36] National Population Commission (NPC) [Nigeria]. Nigeria Demographic and Health Survey 1999. Calverton, Maryland: National Population Commission and ORC/Macro, 2000.

[37] Population Resource Centre. Nigeria Demographic Profile (2000). Accessed at: http://www.prcdc.org/summaries/nigeria/nigeria.html

[38] Rabe-Heskesth S, Everitt B. A Handbook of Statistical Analysis Using Stata, 2nd ed. Chapman and Hall/CRC, 2000.

[39] Sastry N. What explains rural-urban differentials in child mortality in Brazil? Social Science and Medicine, 1997; 44: 989-1002.

[40] Smith AFM, Roberts GO. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. J R Stat Soc. (B). 1993; 55:3-23.

[41] Spiegelhalter D, Best N, Carlin B, Van der Line A. Bayesian measures of models complexity and fit. Journal of the Royal Statistical Society Series B 2002; 64: 1–34.

[42] Vella V, Tomkins A, Nidku J, Marshall T. Determinants of child mortality in South-West Uganda. Journal of biosocial Science, 1992; 24: 103-112.

[43] Wagstaff A. Poverty and health. Geneva: Commission on Macroeconomics and Health working paper series, 2001. Washington DC: The World Bank.

[44] Watson RT, Zinyowera MC, Moss RH. (eds). The regional impacts of climate change: An assessment of vulnerability. A special report of the Intergovernmental Panel on Climate Change, Working Group II. Cambridge: Cambridge UP, 1997.

[45] Weeks JR. The Role of Spatial Analysis in Demographic Research. In Goodchild MF, Janelle DG. (eds). Spatially Integrated Social Science. Oxford University Press, 2004.

[46] Weimer JP. The Economic Benefits of Breastfeeding: A Review and Analysis. Food and Rural Food Assistance and Nutrition Research Report No. 13, U.S. Department of Agriculture, 2001.

[47] Wolfe B, Behrman J. Determinants of child mortality, health and nutrition in a developing country. J Dev Economics 1982; 11:163-193.

Table 1: Under-five mortality rates (per 1000 live births) across districts, Nigeria (1999-2003) (* indicates imputed values)

North Central		**All**	**172**
	1	Plateau	65*
	2	Benue	112*
	3	Kogi	131
	4	Kwara	96*
	5	Niger	202
	6	Abuja (FCT)	123*
North East		**All**	**270**
	7	Taraba	132*
	8	Adamawa	270*
	9	Borno	262
	10	Bauchi	278*
	11	Yobe	299
North West		**All**	**264**
	12	Jigawa	263*
	13	Kano	266
	14	Kebbi	240
	15	Kaduna	221
	16	Katsina	222
	17	Sokoto	304*
South East		**All**	**92**
	18	Anambra	54*
	19	Enugu	192
	20	Abia	126
	21	Imo	98*
South South		**All**	**187**
	22	Cross River	136*
	23	Akwa Ibom	154*
	24	Rivers	242*
	25	Delta	117*
	26	Edo	134*
South West		**All**	**101**
	27	Lagos	101
	28	Oyo	52
	29	Osun	86*
	30	Ogun	124
	31	Ondo	118*
National			**218**

Table 2: Descriptive statistics of covariates used in the analysis, Nigeria DHS, 2003.

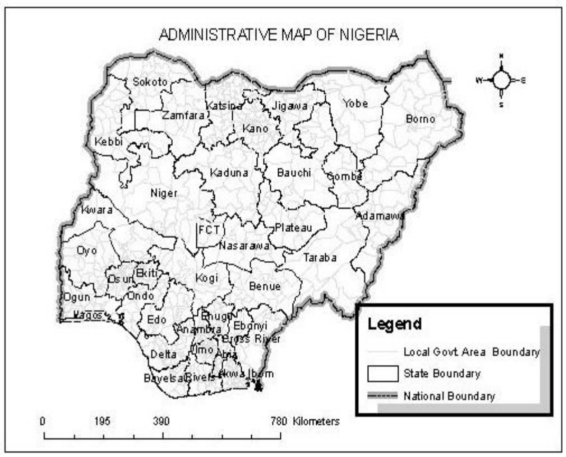
Variable	Level	Frequency (%)	Coding
Place of residence	Urban	2118 (35%)	1
	Rural	3911 (65%)	-1 (ref)
Sex	Male	3062 (51%)	1
	Female	2967 (49%)	-1 (ref)
Preceding Birth Interval	< 25 months	3266 (58%)	-1 (ref)
	25+ months	2326 (42%)	1
Mother's age at birth of child	< 20 years	264 (4%)	1
	20 - 35 years	5765 (96%)	-1 (ref)
Antenatal visits during pregnancy	At least one visit	2337 (64%)	1
	No visit	1339 (36%)	-1 (ref)
Place of delivery	Hospital	2094 (35%)	1
	Home/other	3878 (65%)	-1 (ref)
Asset Index	1st quantile	970 (16%)	1
	2nd quintile	2332 (39%)	2
	3rd quintile	1322 (22%)	3
	4th quintile	1405 (23%)	-1 (ref)
Mother's education	No educ	3033 (50%)	1
	At least primary	2966 (50%)	-1 (ref)
Partner's education	No educ	2343 (40%)	1
	At least primary	3501 (60%)	-1 (ref)
Marital status	Single	483 (8%)	1
	Married	5546 (92%)	-1 (ref)
Household size	Large size	1724 (29%)	1
	Medium size	2927 (48%)	2
	Small size	1378 (23%)	-1 (ref)

Table 3: Posterior Odds ratio of the fixed effect parameters for under-five mortality in Nigeria
Model 2:

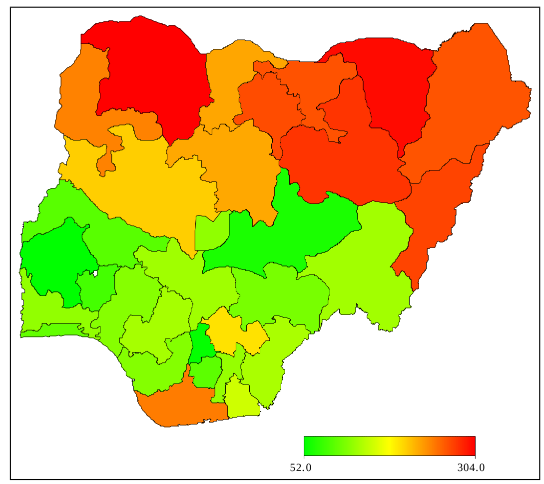
Variable	Level	**Odds Ratio**	**2.5% quantile**	**97.5% quantile**
Residence	**Urban**	**0.54**	**0.38**	**0.83**
Sex	Male	1.08	0.83	1.40
Preceding Birth Interval	**25+ months**	**0.71**	**0.55**	**0.94**
Antenatal visits during pregnancy	**At least one visit**	**0.57**	**0.40**	**0.77**
Place of delivery	Hospital	0.95	0.68	1.40
Asset Index	1st quintile	0.86	0.55	1.23
	2nd quintile	1.09	0.78	1.54
	3rd quintile	0.93	0.64	1.37
Mother's education	**None**	**1.51**	**1.06**	**2.25**
Partner's education	None	0.76	0.54	1.20
Marital status	Single	1.27	0.66	2.47
Household size	Medium	0.99	0.67	1.68
	Large	0.96	0.64	1.51

Figure 1: Maps of Nigeria with crude and estimated mortality indicators

a) Nigeria: location of the 36 states/districts



b) Raw under-five mortality by districts



c) Estimated posterior means (log-odds) of spatial effects

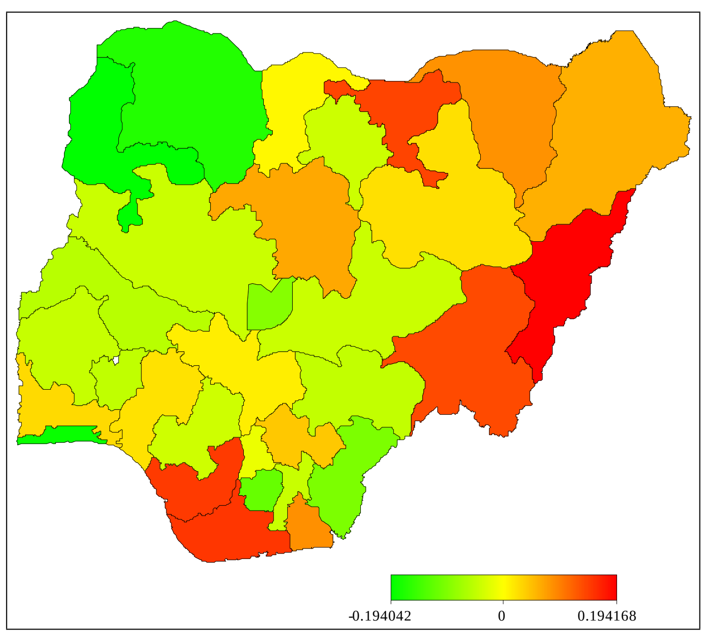
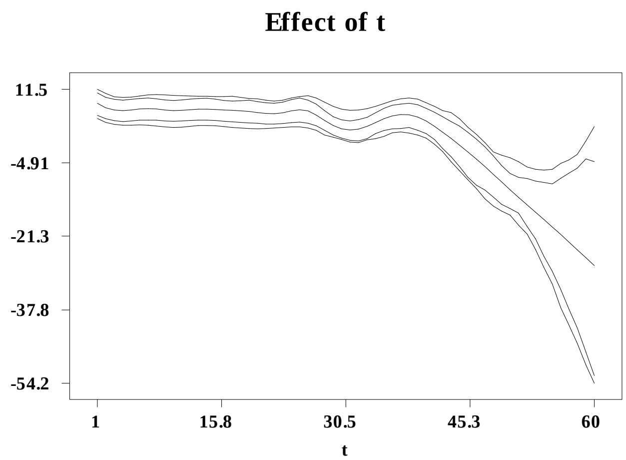
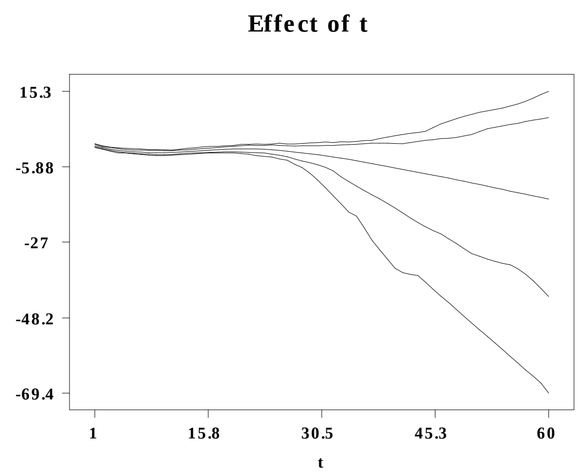


Figure 2: Estimated posterior mean log-odds of death risk with 80% and 95% credible intervals)

a) Nonlinear effect of baseline time



b) Time-varying effect of breasfeeding



c) Non-linear effect of mother's age at child's birth