



Bayesian analysis of zero inflated spatiotemporal HIV/TB child mortality data through the INLA and SPDE approaches: Applied to data observed between 1992 and 2010 in rural North East South Africa

Eustasius Musenge^{a,b,*}, Tobias Freeman Chirwa^b, Kathleen Kahn^{a,d,e}, Penelope Vounatsou^c

^a MRC/Wits Rural Public Health & Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

^b Biostatistics and Epidemiology Division, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

^c Swiss Tropical and Public Health Institute, Basel, Switzerland

^d Centre for Global Health Research, Umeå University, Umeå, Sweden

^e INDEPTH Network, Accra, Ghana

ARTICLE INFO

Article history:

Received 7 July 2011

Accepted 2 April 2012

Keywords:

GMRF

Big "N"

Zero inflated

INLA SPDE

HIV/TB mortality

Spatiotemporal

Agincourt South Africa

ABSTRACT

Longitudinal mortality data with few deaths usually have problems of zero-inflation. This paper presents and applies two Bayesian models which cater for zero-inflation, spatial and temporal random effects. To reduce the computational burden experienced when a large number of geo-locations are treated as a Gaussian field (GF) we transformed the field to a Gaussian Markov Random Fields (GMRF) by triangulation. We then modelled the spatial random effects using the Stochastic Partial Differential Equations (SPDEs). Inference was done using a computationally efficient alternative to Markov chain Monte Carlo (MCMC) called Integrated Nested Laplace Approximation (INLA) suited for GMRF. The models were applied to data from 71,057 children aged 0 to under 10 years from rural north-east South Africa living in 15,703 households over the years 1992–2010. We found protective effects on HIV/TB mortality due to greater birth weight, older age and more antenatal clinic visits during pregnancy (adjusted RR (95% CI): 0.73(0.53;0.99), 0.18(0.14;0.22) and 0.96(0.94;0.97) respectively). Therefore childhood HIV/TB mortality could be reduced if mothers are better catered for during pregnancy as this can reduce mother-to-child transmissions and contribute to improved birth weights. The INLA and SPDE approaches are computationally good alternatives in modelling large multilevel spatiotemporal GMRF data structures.

© 2012 Elsevier B.V. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

1. Introduction

Public Health data on mortality have been growing increasingly rich as more accurate information on “who”, “where” and “when” becomes available. These form hierarchical (multilevel) data structures which are correlated such that person-level (“who”) information can be repeated, geo-statistical (“where”) data often has spatial correlation and temporal (“when”) data can be auto-correlated. Classical statistical techniques are usually based on independent observations, but when applied to multilevel data structures they often underestimate the standard errors. As a result of this the statistical significance is overestimated leading to erroneous results and subsequent inferences (Cressie, 1993).

* Corresponding author at: MRC/Wits Rural Public Health & Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, 7 York road, Parktown 2193, Johannesburg, South Africa.

E-mail address: Eustasius.Musenge@wits.ac.za (E. Musenge).

This defeats the main goal in epidemiological analysis, which is to identify and quantify correctly any exposures, behaviours and characteristics that may modify a population's or individuals risk and use these to implement more appropriate interventions (Rose, 2001).

In modelling hierarchical data we can take into account spatial and temporal correlations by introducing in the model spatiotemporal random effects. Several other hurdles have to be overcome when modelling hierarchical mortality data such as: zero inflation when there is a greater proportion of non-occurrence for an outcome, handling large data structures, repeated measures and estimating many parameters rapidly and accurately. Bayesian techniques with the aid of the Markov chain Monte Carlo (MCMC) simulation methods have successfully overcome these hurdles and fit spatiotemporal random effects for reasonably sized geo-locations of Gaussian fields (GF) (Berliner et al., 2000; Gilks et al., 1996; Casella and Robert, 1999; Wikle, 2003; Wikle et al., 1998). However as the number of geo-locations increases, MCMC computations of a dense GF $m \times m$ spatial correlation matrix become

infeasible or extremely slow in the order of power three ($O(m^3)$), this problem is popularly known as the “big m” or “big N” (Banerjee et al., 2004). Several approaches have been used to resolve the “big m”. Banerjee et al. (2004) give brief summaries of these: sub-sampling, spectral, lattice, dimension reduction and course fine coupling methods (Banerjee et al., 2004, 2008; Banerjee and Carlin, 2003; Kamman and Wand, 2001; Johnson et al., 1990; French et al., 2002). Generally these techniques attempt to reduce the dimension of the GF by selecting a “representative” sub-sample or fixing some parameters or changing the scale from continuous to discrete with the aim of reducing the computational burden in running MCMC simulations.

We addressed this problem using techniques proposed by Rue and Held (2005) who changed the continuous scale GF to a discrete scale Gaussian Markov Random Field (GMRF), for the Matérn family of covariance structures (Rue and Held, 2005). More recently Lindgren et al. (2011) provides the detail of how the GF and GMRF relate via Stochastic Partial Differential Equations (SPDE) using basis functions (Lindgren et al., 2011; Cameletti et al., 2012). Secondly we performed inference and prediction using Integrated Nested Laplace Approximation (INLA) well suited for GMRF as opposed to the commonly used MCMC (Rue et al., 2009). Hence we greatly reduced the computational burden and could do in hours what usually took days having reduced the computational operations for a spatiotemporal model from power 3 to power 1.5 ($O(m^3) \rightarrow O(m^{3/2})$).

The aim of this paper is to discuss and apply a Bayesian model that can handle large zero-inflated spatiotemporal observational data on mortality producing reliable estimates speedily. In Section 2 we explore the Bayesian methods and model fitting inference, prediction and goodness of fit. Section 3 we apply the discussed Bayesian spatiotemporal model to the data from Agincourt in rural South Africa which has 71,057 children aged 0–9 years living in 15,703 households over the years 1992–2010. In Section 4 we discuss the merits of our model and distil the Public Health implications of our results in interventional studies.

2. Methods

2.1. Spatiotemporal model structure

The outcome $y_i(s_j, t)$ was the observed HIV/TB related death of a child $i = 1, \dots, N$ from a given household $j = 1, \dots, m$ in a specific year $t = 1, \dots, T$ which is a realisation of the spatio-temporal process $y(\dots) \in Y(\dots)$. Assuming the outcomes distribution belongs to the exponential family of distributions, we can fit flexible structural additive models belonging to the generalized linear mixed models (GLMM) (Brezger and Lang, 2006; Fahrmeir and Lang, 2001). Our data may be represented by the equation:

$$y_i(s_j, t) = X(s_j, t)\boldsymbol{\beta} + \sum_{j=1}^{n_f} f^{(j)}(u_{ji}, t) + \varepsilon_i(s_j, t) \quad (1)$$

where $X(s_j, t)$ is the design matrix with fixed p covariates, $\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)'$ is the regression coefficients vector, $f(\cdot)$ which is one of the $f^{(j)}$ used to relax the linear relationship or introduce random effects or both and $\varepsilon_i(s_j, t) \sim N(0, \sigma_\varepsilon^2)$ are the error terms which are neither temporally nor spatially correlated (Cameletti et al., 2012). As our data are spatially and temporally correlated we can introduce random effects $f(\cdot) = f(s_j, t)$ a Gaussian random field with a first order autoregressive temporal effect $\xi(s_j, t - 1)$ and coefficient ϕ and zero mean multivariate normal (temporally independent) spatial effects $\omega(s_j, t) \sim MVN(0, \Sigma = \sigma_\omega^2 C(\|s_j - s_k\| = h); j \neq k)$ resulting in the equation:

$$f(s_j, t) = \phi \xi(s_j, t - 1) + \omega(s_j, t) \quad (2)$$

where $|\phi| < 1$ in case of stationarity, $\xi(s_j, 1) \sim N(0, \sigma_\omega^2 / (1 - \phi^2) = 1 / \tau_\omega (1 - \phi^2))$ and the spatial effect is second order stationary and isotropic. When the spatial correlation follows a Matérn covariance structure we obtain $C(h) = (1 / (\Gamma(\nu) 2^{\nu-1})) (\kappa h)^\nu K_\nu(\kappa h)$ for the Euclidean distance lags h . The parameter ν measures the degree of smoothness and also the order of the modified Bessel function (when $\nu > 0$) of second kind K_ν , and finally $\kappa > 0$ is the scaling parameter with a range $\rho = (\sqrt{8\nu} / \kappa)$ where the spatial correlation is close to 0.1 for each ν (Lindgren et al., 2011; Rue et al., 2009).

2.2. Zero inflated Poisson and Binomial spatiotemporal models

Observational binary outcome data are commonly analysed using the logistic regression model, which has a logit linear predictor in the GLMMs canonical link structure. However this model has problems of instability especially with spatial random effects, which would be exacerbated due to zero inflation (Agarwal et al., 2002). In epidemiological cohort studies a relative risk is more preferred than an odds ratio as this caters for temporality and also a more intuitive measure of burden of morbidity or mortality (Barros and Hirakata, 2003; Fekedulegn et al., 2010). In light of this and in an endeavour to have better fitting models, two models that can handle zero inflation were employed. The two conditionally independent models fit were the zero inflated Poisson and zero inflated Binomial with a log and a logit canonical link functions respectively: $\mathbf{y}_t | \boldsymbol{\eta}_t, \theta_t \sim ZIP(\boldsymbol{\mu}_t, \theta_t)$ and $\mathbf{y}_t | \boldsymbol{\eta}_t, p_t, \theta_t \sim ZIBin(n_t, p_t, \theta_t)$.

The mortality outcome (count/binary) data $\mathbf{y}_t(s_j)$ observed at the households in the Agincourt area are zero inflated and assumed to follow either a Poisson (count) or Binomial distribution (binary). We will occasionally drop the s_j for notational convenience in the rest of the article. We therefore resorted to the zero inflated models to cater for the imbalance due to many zeros. The model that takes care of zero inflation (θ_t) can then be represented as:

$$\pi(\mathbf{y}_t | \boldsymbol{\eta}_t, \theta_t) = \begin{cases} \theta_t + (1 - \theta_t)\pi(0 | \boldsymbol{\eta}_t, \theta_t) & \text{if } \mathbf{y}_t = 0 \\ (1 - \theta_t)\pi(1 | \boldsymbol{\eta}_t, \theta_t) & \text{if } \mathbf{y}_t \neq 0 \end{cases} \quad (3)$$

With canonical links of the expected means: $g(E(y_t(\cdot))) = \log(\boldsymbol{\mu}_t) = \boldsymbol{\eta}_t$ and $g(E(y_t(\cdot))) = \text{logit}(p_t) = \boldsymbol{\eta}_t$ with means are $\boldsymbol{\mu}_t = \exp(\boldsymbol{\eta}_t)$ and $p_t = (\exp(\boldsymbol{\eta}_t) / (1 + \exp(\boldsymbol{\eta}_t)))$ for the Poisson and Binomial distributions, respectively. A spatiotemporal canonical link (linear predictor) model can be expressed as

$$\boldsymbol{\eta}_t(s_j) = X_t(s_j)\boldsymbol{\beta} + f_t(s_j) + \boldsymbol{\varepsilon}_t(s_j) \quad (4)$$

$$f_t(s_j) = \phi \xi_{t-1}(s_j) + \omega_t(s_j)$$

where $\boldsymbol{\varepsilon}_t(s_j) \sim N(\mathbf{0}, \sigma_\varepsilon^2 I_m)$ with identity matrix I_m of dimension $m \times m$, $\omega_t \sim N(\mathbf{0}, \Sigma = \sigma_\omega^2 \tilde{\Sigma})$ with a stationary AR(1) process $\boldsymbol{\xi}_1 \sim N(\mathbf{0}, \Sigma / (1 - \phi^2))$ (Cameletti et al., 2012). The Σ is a dense GF m by m dimensional matrix from a Matérn distribution with scale and smoothness parameters κ and ν (which is fixed in all our computations) respectively. As the size of m increases computations become increasingly more difficult due to the “big m” as previously highlighted.

2.3. Solving the “big m” using the SPDEs to estimate the spatial random effects

To resolve the computational burden associated with the GF Matérn covariance function we used a technique that changes this to a GMRF proposed by Rue and Held (2005). Briefly the locations are converted into areal triangulations firstly by making them the initial triangle vertices before adding more vertices for proper triangulation which extends the grid and very useful for prediction. Fig. 1 shows how we employed triangulation to our

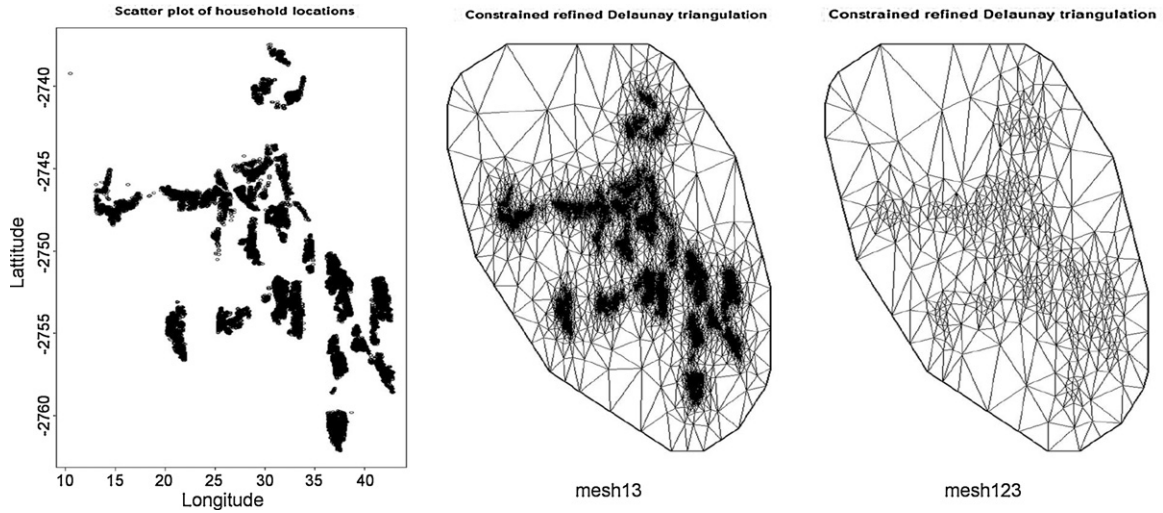


Fig. 1. Agincourt original household locations (left), triangulation of all household (centre) and triangulation of households within 500 m (right).

data, the diagram on the right was made assuming households within 500 m were similar we fit 488 vertices and 938 triangles.

The SPDE technique redefines the Matérn field as a finite “representative” linear combination of basis functions on a triangulation of the locations (Cameletti et al., 2012). Hence for our spatiotemporal random effects in GMRF representation we get $\hat{\omega}_t \sim N(\mathbf{0}, \Sigma = Q_s^{-1})$ and $\xi_1 \sim N(\mathbf{0}, (Q_s^{-1}/(1 - \phi^2)) = Q_T^{-1})$ where Q_s is a sparse time-invariant precision matrix with dimension m^* vertices from the triangulations. A joint spatiotemporal GMRF $\tilde{f}_t \sim N(\mathbf{0}, Q^{-1} = (Q_T \otimes Q_s)^{-1})$ whose precision matrix is the Kronecker product of the temporal and spatial precision matrices; is such that $\tilde{f}_t \approx \mathbf{B}f_t$ where the basis \mathbf{B} is a sparse matrix with unit elements for matching triangle vertices and zero's elsewhere (Lindgren et al., 2011). Therefore Eq. (4) becomes $\eta_t = X_t\beta + \mathbf{B}f_t + \varepsilon_t$, where we let $\mathbf{x}_t = \{\beta, \mathbf{B}, \theta_t\}$ be the vector of latent Gaussian fields and $\varphi_t = \{\sigma_\omega^2, \phi, \kappa, \sigma_\varepsilon^2\}$ being a vector of unknown parameters. We can thus express our model into a hierarchical Gaussian latent variable fashion as follows, stage 1 – observational equation, stage 2 – latent Gaussian field and stage 3 – parameter model (Simpson et al., 2011):

$$\begin{aligned} \text{stage 1: } & \mathbf{y}_t | \mathbf{x}_t, \varphi_t \sim N(X_t\beta + \mathbf{B}f_t; \sigma_\varepsilon^2 I_m) = N(\mu_{y_t}(\varphi_t) = A\mathbf{x}_t, Q_{y_t}^{-1}(\varphi_t)) \\ \text{stage 2: } & x_t | \varphi_t \sim N(\mu_{x_t}(\varphi_t); Q_{x_t}^{-1}(\varphi_t)) \\ \text{stage 3: } & \varphi_t \sim \pi(\varphi_t) \end{aligned} \quad (5)$$

where the precision matrices $Q^{-1}(\cdot)$ are either small enough (for easier multiple factorisation) or sparse (Simpson et al., 2011b). These models cover a wide range of models and are easily estimable using INLA as shown in the next subsection.

2.4. Bayesian inference using INLA

In accordance with the Bayesian paradigm we aim to find the posterior distribution of the processes and parameters updated by data (Wikle, 2003). This could be expressed as:

$$\begin{aligned} \text{Probability (process, parameters|data)} & \propto \text{Likelihood} \\ & \times (\text{data|process, parameters}) \times \text{Probability (process|parameters)} \\ & \times \text{Probability (parameters)} \end{aligned}$$

Applying this expression to our model, letting $\Theta = \{\beta, \phi, \theta_t, \sigma_\omega^2, \kappa, \sigma_\varepsilon^2\}$ denote the vector of all parameters and dropping the subscripts to present in vector form, $\xi = \{\xi_t\}$ and

data $\mathbf{y} = \{y_t\}$ (Cameletti et al., 2012; Rue et al., 2009), their joint posterior distribution is thus:

$$\pi(\Theta, \xi, \eta | \mathbf{y}) \propto \pi(\mathbf{y} | \Theta, \xi, \eta) \pi(\eta | \xi, \Theta) \pi(\xi, \Theta) \quad (6)$$

Fig. 2 gives a simplified pictographical view, where level 1 are the data and assumed distributions, level 2 is a process, level 3 and 4 are parameters and level 5 gives the default hyper-parameters used in the INLA package.

The posterior marginals are required, standard (i), nested approximation (ii) and numerical integrations (iii) for latent fields $\mathbf{x} = \{\beta, f, \theta\}$ and hyper-parameters $\varphi = \{\phi, \sigma_\omega^2, \kappa, \sigma_\varepsilon^2\}$ respectively (Rue and Held, 2005; Rue et al., 2009):

$$\begin{aligned} \text{i) } \pi(\mathbf{x}, \varphi | \mathbf{y}) &= \int \pi(\mathbf{x}_i | \varphi, \mathbf{y}) \pi(\varphi | \mathbf{y}) d\varphi & \pi(\varphi_j | \mathbf{y}) &= \int \pi(\varphi | \mathbf{y}) d\varphi_j \\ \text{ii) } \tilde{\pi}(\mathbf{x}, \varphi | \mathbf{y}) &= \int \tilde{\pi}(\mathbf{x}_i | \varphi, \mathbf{y}) \tilde{\pi}(\varphi | \mathbf{y}) d\varphi & \tilde{\pi}(\varphi_j | \mathbf{y}) &= \int \tilde{\pi}(\varphi | \mathbf{y}) d\varphi_j \\ \text{iii) } \tilde{\pi}(\mathbf{x}, \varphi | \mathbf{y}) &= \sum_k \tilde{\pi}(\mathbf{x}_i | \varphi_k, \mathbf{y}) \tilde{\pi}(\varphi_k | \mathbf{y}) \Delta_k & \tilde{\pi}(\varphi_j | \mathbf{y}) &\propto \int \frac{\pi(\mathbf{x}, \varphi | \mathbf{y})}{\tilde{\pi}_G(\mathbf{x} | \varphi, \mathbf{y})} \Big|_{\mathbf{x}=\mathbf{x}^*(\varphi)} d\varphi_j \end{aligned} \quad (7)$$

Using this technique we aim to initially get the terms “nested” inside the integrand in Eq. (7ii) left hand side without integration (Simpson et al., 2011b). To do so we firstly estimate the marginal $\tilde{\pi}_G(\mathbf{x} | \varphi, \mathbf{y})$ which is a Gaussian approximation of \mathbf{x} with mode $\mathbf{x}^*(\varphi)$, for a given φ . Secondly we estimate $\tilde{\pi}(x_i | \varphi, \mathbf{y}) = N(x_i; \mu_i(\varphi), \sigma_i^2(\varphi))$ using either Gaussian or Laplace or a simplified Laplace approximations (Rue et al., 2009). We computed these marginals using the R package Integrated Nested Laplace Approximation (INLA) which uses the simplified Laplace approximations (Rue and Held, 2005; Rue et al., 2009; R-cran, 2010). The INLA procedure also enables easier spatial prediction since it computes posterior conditionals for the spatial random effects on all triangulation vertices including the extensions as shown in Fig. 1.

2.5. Model goodness of fit and convergence diagnostics

We assessed the accuracy of $\tilde{\pi}(\varphi | \mathbf{y})$ using the effective number of parameters, which can be approximated as the difference between the dimension of the normalised integral $\{\tilde{\pi}(\varphi | \mathbf{y})\}$ n and the trace of the product of the prior precision matrix and the posterior covariance (Spiegelhalter et al., 2002):

$$p_D(\varphi) \approx n - \text{tr}\{Q(\varphi)Q^*(\varphi)^{-1}\} \quad (8)$$

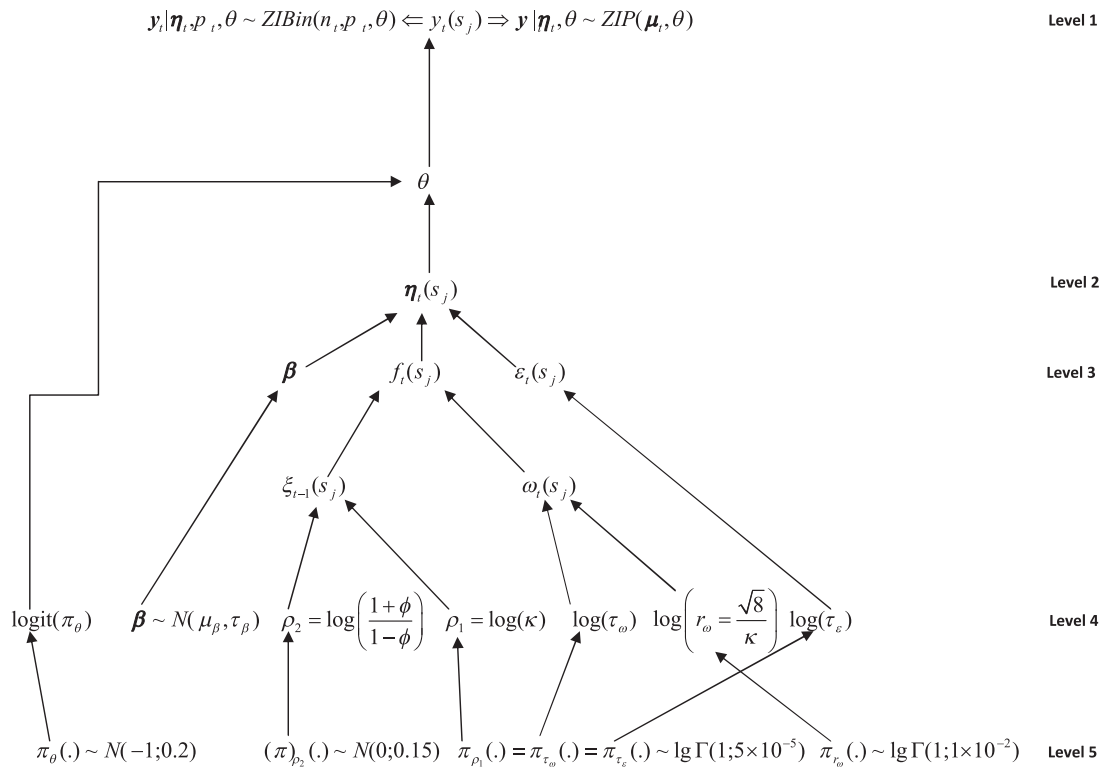


Fig. 2. Hierarchical structure of a zero inflated spatiotemporal model fit using INLA.

The deviance information criteria (DIC) was also used which is defined as the difference between twice the mean of the deviance and the deviance of the mean according to Spiegelhalter et al. (2002) and expressed as:

$$D(x, \varphi) = -2 \sum_i \log(\tilde{\pi}(y_i | x_i, \varphi)) + \text{constant} \quad (9)$$

Interpretations of these is quite straight forward the smaller the effective number of parameters the more parsimonious the model and the smaller the DIC the better the model fit, more-so most parsimonious is not always the best model.

3. Application

3.1. Rural South Africa Agincourt HDSS data, study design and ethics

The Agincourt health and demographic surveillance system (HDSS) site was set up in the Agincourt sub-district in 1992 due to its remote location, availability of several clinics and presence of Mozambican in-migrants (Tollman et al., 1999). By 2010, the Agincourt HDSS had a population of over 84,000 persons living in approximately 17,000 households scattered throughout 27 neighbouring villages. Cause of death data were obtained through verbal autopsies conducted on every recorded death (Clark et al., 2007). Interviews were conducted by trained field worker. This was done

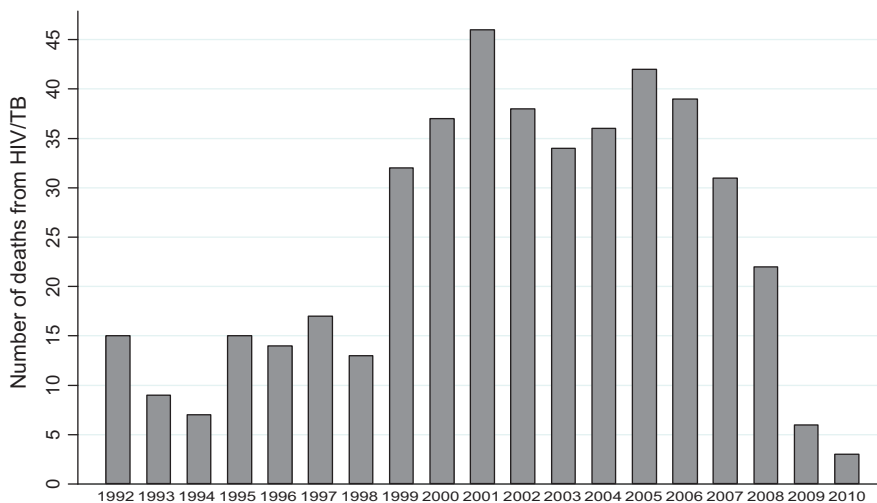


Fig. 3. Year specific child deaths due to HIV/TB from 1992 to 2010.

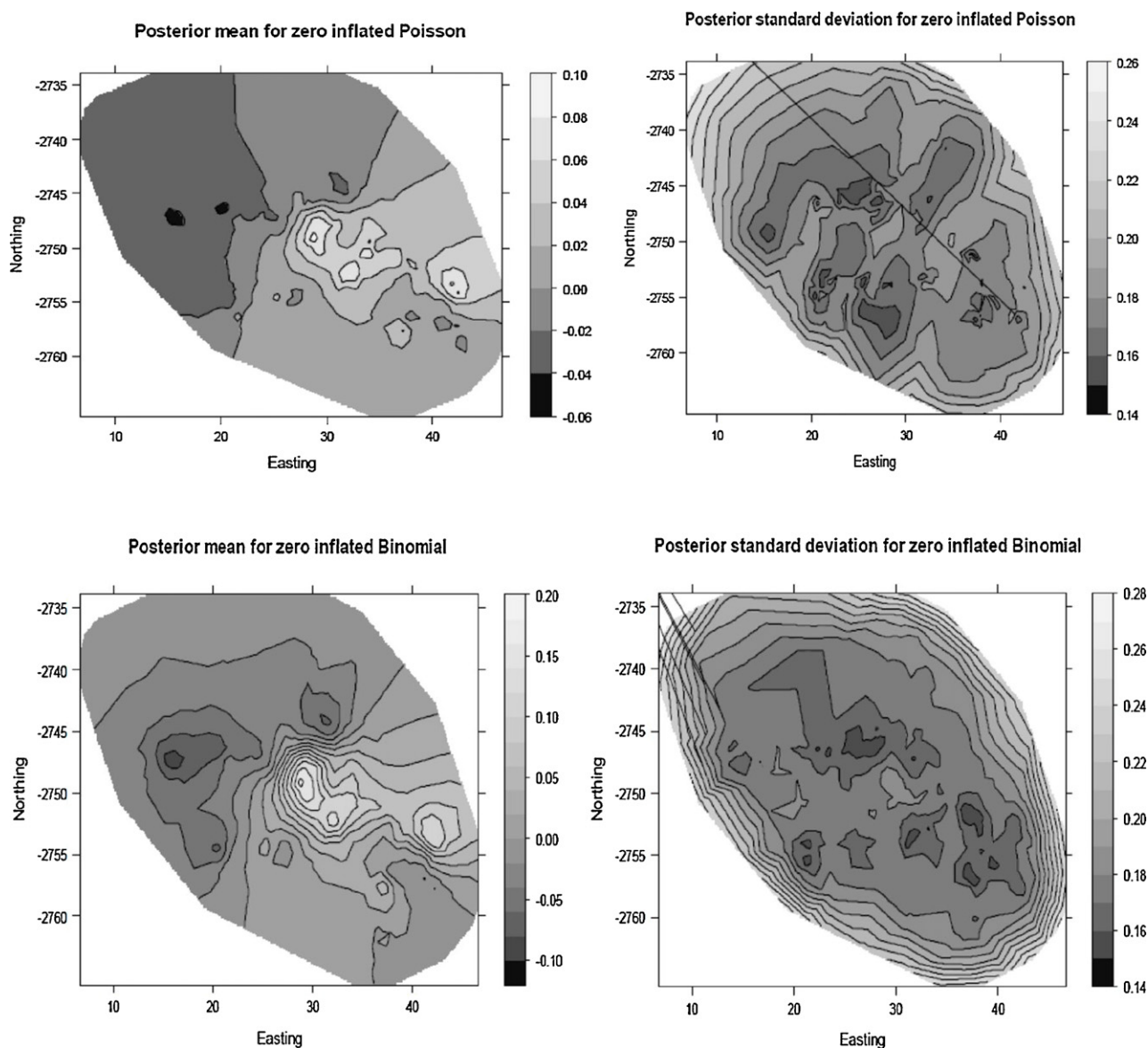


Fig. 4. Posterior point estimates for ZIP (top) and ZIB (bottom) models.

within 1 year after a death, with the closest caregiver of the deceased in their mother tongue. Cause of death was independently determined by two medical practitioners with the third as a tie breaker. Their consensus cause of death was classified according to the World Health Organization's International Classification of Diseases (ICD10) (Kahn et al., 2000). HIV/TB mortality in children was ascertained by the reported signs and symptoms, and in some instances this was verified through the mother's cause of death (Kahn et al., 2000). Over 90% of the HDSS households were geo-coded by 1992 and by 2010 all the households were geo-coded, thus enabling spatial analyses at household as well as village level. The study design was a retrospective cohort study covering households observed from the onset of the site to December 2010. The Agincourt HDSS site was granted ethical clearance by the University of the Witwatersrand's Committee for Research on Human Subjects (No. 960720). This work was also granted ethical clearance by the University of the Witwatersrand's Committee for Research on Human Subjects (M081145). Verbal informed consent was obtained when the census rounds were conducted and also when verbal autopsy data were collected from a close relative of the deceased.

3.2. Dependent and independent variables

The persons included in the study were all the children aged between 0 and under 10 years who lived or had lived in the Agincourt HDSS between January 1992 and December 2010. The independent variables used were: child's gender, birth-weight category, age category, slum (availability of water, electricity and toilet) and number of mother's antenatal clinic visits during pregnancy and year of observation. The household's latitude and longitude were used to construct the latent variables for the spatial random effects and year the AR(1) temporal effects. The dependent variable was death due to HIV and or tuberculosis (TB) determined by the WHO's ICD10 verbal autopsy codes A16–A19¹ for HIV and B20–B24² for TB. These data were extracted using Structured Query

¹ A16=respiratory tuberculosis; not confirmed bacteriologically or histologically; A17=tuberculosis of nervous system; A18=tuberculosis of other organs; A19=miliary tuberculosis.

² B20=human immunodeficiency virus (HIV) disease resulting in infectious and parasitic diseases; B21=human immunodeficiency virus (HIV) disease resulting in

Table 1
Univariate and multiple regression results models using zero inflated Poisson adjusting for spatiotemporal random effects.

| Variable | Summary, <i>n</i> (%) or mean ± SD | Univariate results, RR (95% credible interval) | Non-spatial multiple variable model, adjusted RR (95% credible interval) | Temporal multiple variable model, adjusted RR (95% credible interval) | Spatial multiple variable model, adjusted RR (95% credible interval) | Spatiotemporal multiple variable model, adjusted RR (95% credible interval) |
|---|------------------------------------|--|--|---|--|---|
| Sex | | | | | | |
| Male | 35,317(49.70) | 1.00 | | | | |
| Female | 35,740(50.30) | 0.89(0.74;1.07) | | | | |
| Birth weight category | | | | | | |
| Low weight | 6,320(8.89) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Moderate weight | 15,241(21.45) | 0.81(0.61;1.07) | 0.70 [*] (0.53;0.93) | 0.85(0.63;1.16) | 0.71 [*] (0.53;0.93) | 0.85(0.63;1.16) |
| High weight | 49,496(69.66) | 0.39 [*] (0.30;0.51) | 0.57 [*] (0.42;0.76) | 0.73 [*] (0.53;0.99) | 0.58 [*] (0.43;0.77) | 0.73 [*] (0.53;0.99) |
| Age category | | | | | | |
| 0–1 years | 8,580(12.07) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1–5 years | 19,619(27.61) | 0.21 [*] (0.17;0.26) | 0.20 [*] (0.16;0.25) | 0.18 [*] (0.23;0.64) | 0.20 [*] (0.16;0.25) | 0.18 [*] (0.14;0.22) |
| 5–9 years | 42,858(60.31) | 0.05 [*] (0.04;0.07) | 0.04 [*] (0.03;0.05) | 0.04 [*] (0.05;0.54) | 0.04 [*] (0.03;0.05) | 0.04 [*] (0.03;0.05) |
| Slum (electricity, water and toilet) | | | | | | |
| None of the three | 10,051(14.14) | 1.00 | 1.00 | | 1.00 | |
| At least one | 27,329(38.46) | 0.78(0.60;1.01) | 0.79(0.61;1.02) | | 0.77(0.60;1.00) | |
| At least two | 21,871(30.78) | 0.74 [*] (0.56;0.97) | 0.73 [*] (0.56;0.96) | | 0.71 [*] (0.54;0.93) | |
| All three | 5,748(8.09) | 0.57 [*] (0.37;0.86) | 0.55 [*] (0.35;0.83) | | 0.53 [*] (0.34;0.81) | |
| Antenatal clinic visits | | | | | | |
| Zero inflation parameter | 6.00(±1.50) | 0.98 [*] (0.97;0.99) | 0.98 [*] (0.97;0.99) | 0.96 [*] (0.94;0.97) | 0.98 [*] (0.96;0.99) | 0.96 [*] (0.94;0.97) |
| Precision for year | | | 0.06 [*] (0.002;0.25) | 0.05 [*] (0.003;0.18) | 0.07 [*] (0.003;0.28) | 0.04 [*] (0.002;0.15) |
| Rho for year | | | | 0.91 [*] (0.23;2.19) | | 0.86 [*] (0.16;2.32) |
| Kappa | | | | 0.88 [*] (0.72;0.97) | | 0.89 [*] (0.70;0.98) |
| Tau | | | | | 1.98 [*] (1.09;3.66) | 1.55 [*] (1.41;1.63) |
| Moran's Indexes: | | | | | 1.37 [*] (0.18;2.600) | 2.31 [*] (1.19;9.08) |
| Observed(Expected ± standard deviation) | | | | | | 0.32 [*] (2.1 ± 3.33 × 10 ⁻³) |
| Effective number of parameters | | | 12.05(0.05) | 25.61(1.10) | 23.18(5.04) | 29.16(6.33) |
| DIC | | | 4855.31 | 4532.69 | 4846.13 | 4532.31 |

^{*} Statistical significance at the 5% level.

Language (SQL). Preliminary data analyses and data management were done using STATA 10.1 (StataCorp, 2007). The Bayesian analysis was done using an R software package called INLA (Rue et al., 2009; R-cran, 2010).

3.3. Descriptive statistics

There were similar proportions of boys and girls (50.3%) with an average age of 6.14 years (standard deviation of 3.38). Almost 9% of the births were low-weight, the majority (60.3%) of the children were only observed after their 5th birthday. Only 8.09% had access to all three: clean tap water, flush toilet inside house and electricity, which is expected in typical rural areas. Health check-ups during pregnancy averaged about 6 visits, with the majority (73%) still not going at all. A total of 456 deaths were HIV/AIDS (including HIV/tuberculosis) which was a small proportion of the population indicative of the presence of zero inflation. Mortality was moderate in the early to mid 1990s before gradually growing from 1999 and reaching a peak in 2001, remaining high for several years before gradually declining from 2007 as shown in Fig. 3.

3.4. Zero inflated Poisson and Binomial spatiotemporal modelling results

We performed univariate and multiple variable analyses using both the ZIP and ZIB models. Multiple variable models were fit systematically starting with one without random effects, spatial only, temporal only and finally spatiotemporal random effects. The R-codes using INLA are given in Appendix 1. From our results shown

in Tables 1 and 2, our discussion will be centred on the ZIP spatiotemporal model which was the best fitting model since this had the lowest DIC (4532.31) and catered for spatiotemporal random effects, as later indicated in Section 3.5.

Two variables consistently showed significant associations with child HIV/TB mortality, firstly the age category of the child, which showed a decrease in mortality with increase in age (Chi-Square *p*-value < 0.001). Those aged 1 to under 5 (127/19,619) were 82% {0.18(0.14; 0.22)} less likely to experience a death due to HIV/TB compared to those under 1(262/8580) and those over 5 and under10 (67/42,858) were 96%{0.04(0.03;0.05)} less likely to die

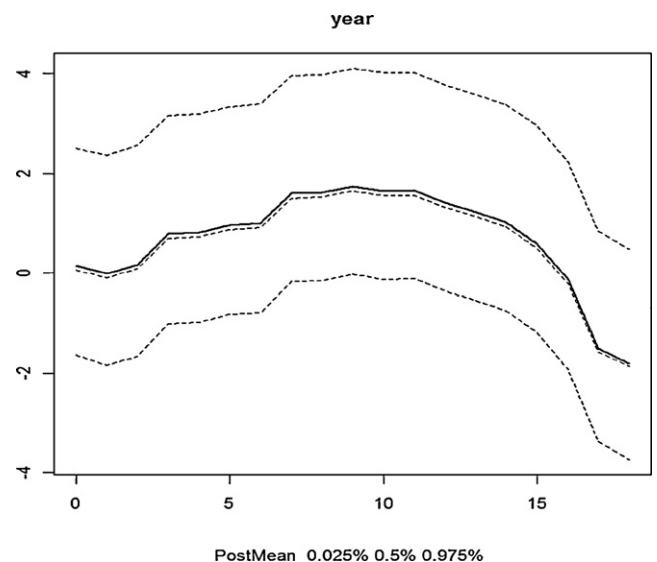


Fig. 5. Posterior means (bold), medians (0.5% = 50th percentiles, middle), lower credible limits (0.025% = 2.5th percentiles) and upper credible limits (0.975% = 97.5th percentiles).

malignant neoplasms; B22 = human immunodeficiency virus (HIV) disease resulting in other specified diseases; B23 = human immunodeficiency virus (HIV) disease resulting in other conditions; B24 = unspecified human immunodeficiency virus (HIV) disease.

Table 2
Multiple regression results of four models using zero inflated Binomial adjusting for spatiotemporal random effects.

| Variable | Non-spatial multiple variable model, adjusted OR (95% credible interval) | Temporal multiple variable model, adjusted OR (95% credible interval) | Spatial multiple variable mode, adjusted OR (95% credible interval) | Spatiotemporal multiple variable model, adjusted OR (95% credible interval) |
|---|--|---|---|---|
| Sex | | | | |
| Male | | | | |
| Female | | | | |
| Birth weight category | | | | |
| Low weight | 1.00 | 1.00 | 1.00 | 1.00 |
| Moderate weight | 0.70 [*] (0.52;0.92) | 0.86(0.62;1.20) | 0.70 [*] (0.52;0.93) | 0.005 [*] (0.003;0.07) |
| High weight | 0.62 [*] (0.46;0.83) | 0.67 [*] (0.48;0.95) | 0.63 [*] (0.47;0.85) | 0.05 [*] (0.007;0.04) |
| Age category | | | | |
| 0–1 years | 1.00 | 1.00 | 1.00 | 1.00 |
| 1–5 years | 0.07 [*] (0.05;0.09) | 0.05 [*] (0.04;0.07) | 0.07 [*] (0.06;0.09) | 0.18 [*] (0.14;0.22) |
| 5–9 years | 0.006 [*] (0.004;0.008) | 0.005 [*] (0.003;0.95) | 0.006 [*] (0.004;0.008) | 0.04 [*] (0.03;0.05) |
| Slum (electricity, water and toilet) | | | | |
| None of the three | 1.00 | | 1.00 | |
| At least one | 0.77(0.60;1.01) | | 0.76 [*] (0.58;0.99) | |
| At least two | 0.71 [*] (0.54;0.95) | | 0.69 [*] (0.52;0.93) | |
| All three | 0.52 [*] (0.34;0.81) | | 0.51 [*] (0.32;0.78) | |
| Antenatal clinic visits | | | | |
| Zero inflation parameter | 0.22 [*] (0.042;0.66) | 0.43 [*] (0.19;0.79) | 0.15 [*] (0.031;0.45) | 0.42 [*] (0.18;0.67) |
| Precision for year | | 0.78 [*] (0.17;1.96) | | 0.80 [*] (0.20;1.93) |
| Rho for year | | 0.87 [*] (0.67;0.97) | | 0.87 [*] (0.69;0.97) |
| Kappa | | | 1.98 [*] (1.14;3.55) | 1.54 [*] (1.41;1.62) |
| Tau | | | 1.37 [*] (0.27;2.54) | 2.23 [*] (1.22;6.91) |
| Moran's Indexes: | | | | |
| Observed(Expected ± standard deviation) | | | | 0.30 [*] (2.1 ± 3.33 × 10 ⁻³) |
| Effective number of parameters | 11.10(0.080) | 24.58(1.03) | 22.97(5.09) | 31.12(5.18) |
| DIC | 5179.43 | 4754.60 | 5180.50 | 4803.93 |

^{*} Statistical significance at the 5% level.

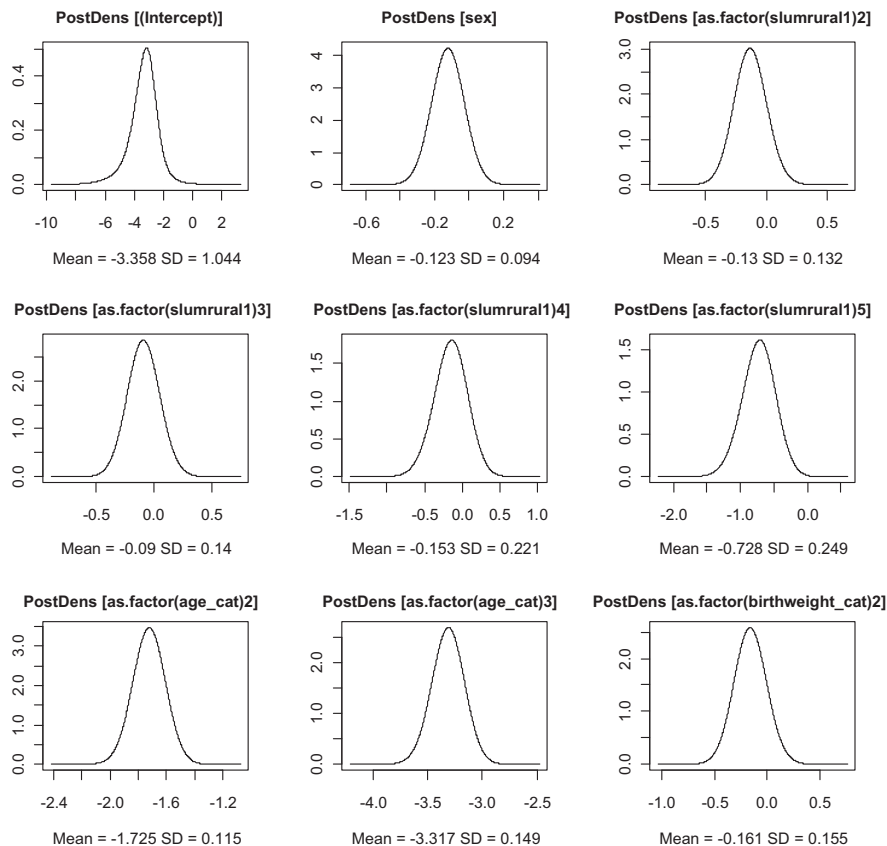


Fig. 6. Posterior density plots of log-means for fixed effects of the ZIP spatiotemporal model.

from HIV/TB relative to the under 1 s keeping all other variables constant. Secondly the category of the birth weight was significantly associated with deaths due to HIV/TB, showing a decreasing trend with greater birth weight (Chi-Square p -value < 0.001). Those who had a birth weight of 3.5 kg and above (232/49,496) were 27% {0.73(0.53;0.99)} less likely to die due to HIV/TB relative to those who were born with a low birth weight of less than 2.5 kg (76/6329) independent of other risk factors. The only other significant protective factor was the number of antenatal clinic visits made by the mother; for every additional visit the mother attended the risk of losing the child from HIV/TB decreased by 4% {0.96(0.94;0.97)} keeping other variables constant as shown by the spatial only and no-random effects model. Those who had access to all three electricity, water and flush toilet (who in this area are the more affluent) experienced fewer deaths due to the HIV/TB compared to those who had none. We show the posterior marginal point estimates from the ZIP and ZIB models in Fig. 4 top and bottom respectively.

As shown in Fig. 4, the region with the greatest risk of child HIV/TB mortality is the central region with a trend towards the south-westerly. In the context of our findings, this is the region whose households have lower birth weights, greater under 1 mortality and having lower visits to the health facility compared to households in other regions. The lowest risk region is the top north-easterly with the ZIP model showing few clusters than the ZIB model. The standard error maps indicate the greatest errors on the peripherals which are also the regions where the extensions of the triangulation shown in Fig. 1. Also consistent with our results from

Tables 1 and 2 there is a greater margin of error with the ZIB model as compared with the ZIP model.

3.5. Model assumptions, goodness of fit and convergence diagnostics

Using the final spatiotemporal models in both ZIB and ZIP we investigated the presence of zero inflation, temporal and spatial correlation. The null hypothesis for zero inflation was that the inflation parameter θ_t is zero (no zero inflation) our results in Tables 1 and 2 indicate significant presence of zero inflation of 0.04 and 0.42 for ZIP and ZIB respectively. Investigation of temporality was done by testing the ϕ coefficient of the first order autocorrelation model with null hypothesis that the process was not stationary $|\phi| > 1$. Both our models indicated significance presence of stationarity that is 0.86 and 0.87 for ZIP and ZIB respectively. Fig. 5 shows the posterior means plus 95% credible bands, time series plots from the modelling which resembles the observed data shown in Fig. 3. Lastly on checking model assumptions we investigated the presence of spatial effects by inspecting the significance of the components (κ and τ) of the Matérn covariance (see Tables 1 and 2). Also this was investigated with the classical Moran's indices on the spatial residuals testing the null hypothesis that there is zero spatial autocorrelation which from both tests yielded significant results, hence rejecting the null.

We compared the two models the ZIB and ZIP and found that the latter fitted better in all the variants from the no-random effects

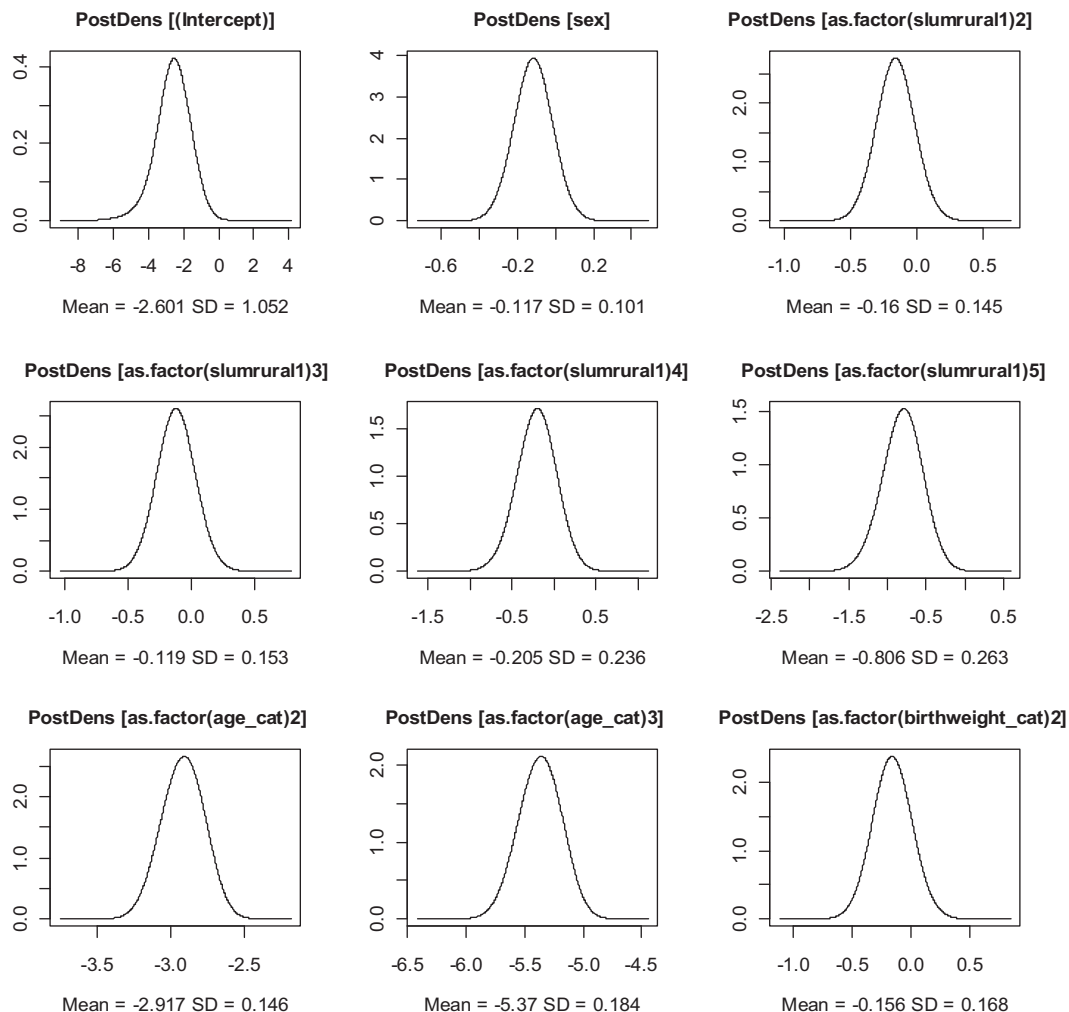


Fig. 7. Posterior density plots of log-means for fixed effects of the ZIB spatiotemporal model.

model to the spatiotemporal as this yielded smaller DICs and model parsimony. Some of the parameter estimates were similar, except for the zero inflated parameter, age categories and birth weights. The ZIP model also showed more stable estimates and consistent results in the adjusted relative risks (adjusted RRs) compared to the ZIB and narrower credible intervals. As a way of showing the convergence pictographically for some parameter estimates we show the fixed effects logarithmic form posterior means density plots in Figs. 6 and 7.

The plots in Figs. 6 and 7 show that the parameters fitted well as GMRF since they are symmetrical about thus the mean equals the mean and mode.

4. Conclusions

The Public Health rationale of our results is centred on lack of health seeking behaviour in rural areas. Interventions could target two groups of mothers (also caregivers) those not yet infected and those already infected. For both groups maternal attendance of health facility can minimize the risk of under one mortality and possibly giving birth to low-weight children. For those mothers' already infected several options are there; before child birth (prevention of mother to child) and after birth (anti-retro-virals – ARV uptake) and also early treatment on ARVs for the parent to increase their longevity. A significant decline in mortality can already be observed in our cohort post 2007 the year when ARVs were started in the area. Poverty however still remains the greatest challenge in the developing countries and mostly interventions when available are very scarce. As such our results from spatiotemporal Bayesian modelling coupled with maps can be very handy in allocation of the limited resources of aid.

Modelling that controls for potential confounding is an addition to epidemiological studies and gives strength to the estimates derived from such modelling procedures. Our modelling approach caters for potential spatial and temporal confounders. This approach was able also to cater for “large” zero inflated spatiotemporal data. The years 1992–2010 cover a span of 19 years (76%) of the United Nation's Millennium Development Goals (MDG) period of 25 years. Goal 4.1 aims at reducing child (under five) mortality by two thirds over that period (Black et al., 2003). Our results demonstrate such a significant decline in mortality due to HIV/TB.

A major limitation in developing nations is lack of reliable and user friendly analyses software, which was addressed by using INLA

a package available on the public domain. Zero inflated Poisson (ZIP) and zero inflated Negative Binomial (ZINB) models were used for log link function linear predictors and also catered for spatial random effects (Lambart, 1992; Ridout et al., 2001). It has also been noted from other studies that the log link function linear predictor models were more stable than logit link function models when they contained a spatial component. This finding was also confirmed in our study as we observed that models from the exponential family with a log-linear predictor were more stable and able to converge better than the logit-link function model. This motivated us to treat the outcome as a discrete measure as opposed to binary mostly used for logistic regression (Agarwal et al., 2002). We demonstrated that the “big m” can be resolved with great ease for Public Health mortality hierarchical structures with the aid of SPDEs and INLA.

Contributions

EM drafted the manuscript and did the statistical analysis. KK provided critical intellectual content and TFC and PV the statistical content. KK was also responsible for the design and quality of the Agincourt verbal autopsy data. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgements

This work was made possible by a doctoral fellowship support from World Health Organisation's, Tropical Disease Research (WHO/TDR) from 2008 to 2010 the Swiss South African Joint Research Program (SSAJP), project number JRP IZLSZ3122926, from 2009 to 2010. The Agincourt HDSS is funded by the Wellcome Trust UK (Grant number 058893/Z/99/A, 069683/Z/02/Z and 085477/Z/08/Z) and the University of the Witwatersrand and Medical Research Council, South Africa, and the Andrew Mellon and Hewlett Foundations, U.S.A. Special thanks goes to Benn Sartorius for data extraction, community leaders, field workers, team supervisors and the Agincourt communities. A very special thanks to Finn Lindgren for answering all my questions promptly whilst I developed models using INLA.

Appendix A. Zero inflated Poisson R-code using INLA adapted from <http://www.r-inla.org/> (Lindgren et al., 2011)

```

#1.#IMPORTING LIBRARIES
library(foreign)
library(INLA)
library(coda)
library(fields)

#2.# Read data:
decdata = read.dta("c:/jagnovredo/childleveldec.dta")
attach(decdata)
names(decdata)

#3.# Build triangular mesh of locations within 0.5 km see figure 1
mesh123 = (inla.mesh.create(cbind(decdata$coordlongx/1000,
decdata$coordlaty/1000),refine=TRUE,cutoff=0.5))
plot(mesh123)

#4.# Store the data-->vertex mapping:
## ( mesh$idx$loc[m] is the mesh vertex for data location number m )
decdata$hhid = mesh123$idx$loc

#5.# Create the SPDE/GMRF model,  $(\kappa^2 - \Delta)(\tau x) = W$ :
spde123 = inla.spde.create(mesh123, model="matern")

#6.# Functions of four models multivariate models fit
#6a.No random effects model with household and individual random effects
norand=hivtb2~sex+as.factor(slumrural1)+as.factor(age_cat)+as.factor(birthwei
ght_cat)+antenatalvisits+f(hhid, model="iid")+f(idno, model="iid")

#6b.Temporal random effects model with household and individual random
effects
temponly=hivtb2~sex+as.factor(slumrural1)+as.factor(age_cat)+as.factor(birthw
eight_cat)+antenatalvisits+f(year, model="ar1")+f(hhid, model="iid")+f(idno,
model="iid")

#6c.Spatial random effects model with individual random effects
spatonly=hivtb2~sex+as.factor(slumrural1)+as.factor(age_cat)+as.factor(birthw
eight_cat)+antenatalvisits+f(idno,model="iid")+f(hhid,model=spde123)

#6d.Spatiotemporal random effects model with individual random effects
spatiotemp=hivtb2~sex+as.factor(slumrural1)+as.factor(age_cat)+as.factor(birt
hweight_cat)+antenatalvisits+f(year, model="ar1")+f(idno,
model="iid")+f(hhid,model=spde123)

#7.# Running INLA models on the four stated in step 6
norand = inla(norand, family="zeroinflated.poisson1",data =
decdata,control.results=list(return.marginals.random=TRUE,return.marginals.pr
edictor=TRUE),control.compute=list(dic=TRUE,cpo=TRUE))
summary(norand)

spatonly = inla(spatonly, family="zeroinflated.poisson1",data =
decdata,control.results=list(return.marginals.random=TRUE,return.marginals.pr
edictor=TRUE),control.compute=list(dic=TRUE,cpo=TRUE))
summary(spatonly)

temponly = inla(temponly, family="zeroinflated.poisson1",data =
decdata,control.results=list(return.marginals.random=TRUE,return.marginals.pr
edictor=TRUE),control.compute=list(dic=TRUE,cpo=TRUE))
summary(temponly)

```

```
spatiotemp = inla(spatiotemp, family="zeroinflated.poisson1", data =
decdata, control.results=list(return.marginals.random=TRUE, return.marginals.pr
edictor=TRUE), control.compute=list(dic=TRUE, cpo=TRUE))
summary(spatiotemp)
```

```
#Saving output for model to store posterior estimates and other
save(spatiotemp, file="c:/resultsdec/spatiotemp.Rdata")
```

```
#-----#
```

```
#8.#Loading data from previous analyses in new session
spatiotemp<-load(file="c:/jagnovredo/spatiotemp.Rdata")
names(spatiotemp)
```

```
#9.#Assessing for posteriors Spatial auto-correlation (Moran's I)
library(ape)
D.dists <- as.matrix(dist(cbind(mesh123$loc[,1], mesh123$loc[,2])))
D.dists.inv <- 1/D.dists
diag(D.dists.inv) <- 0
Moran.I(spatiotemp$summary.random$hhid[, "mean"], D.dists.inv)
```

```
#10.# Diagnostic plots see figures 6 or 7
```

```
plot.inla(spatiotemp,
plot.fixed.effects = TRUE, constant=FALSE,
plot.lincomb = TRUE,
plot.random.effects = TRUE,
plot.hyperparameters = TRUE,
plot.predictor = TRUE,
plot.q = TRUE,
plot.cpo = TRUE)
```

```
#11.# Extract the SPDE parameters, point and interval estimates
```

```
tauconf = exp(spatiotemp$summary.hyperpar[2,])
kappaconf = exp(spatiotemp$summary.hyperpar[3,]/2)
```

```
tau = exp(spatiotemp$summary.hyperpar[2, "mean"])
kappa = exp(spatiotemp$summary.hyperpar[3, "mean"]/2)
```

```
#12a.# Get the spatial precision matrix (Qs) see equation section 2.1 and 2.3
```

```
Q = inla.spde.query(spde123, precision=list(tau=tau,
kappa2=kappa^2))$precision
```

```
#12b.# Reference point for covariance/correlation comparisons:
```

```
ref.s = (which.min((mesh123$loc[,1]-mean(range(mesh123$loc[,1])))^2 +
(mesh123$loc[,2]-mean(range(mesh123$loc[,2])))^2))
```

```
#12c.# Calculate covariances (S) and correlations (SS):
```

```
S = solve(Q)
SS = diag(1/sqrt(diag(S))) %*% S %*% diag(1/sqrt(diag(S)))
D = as.matrix(dist(mesh123$loc))
```

```
#12d.# Theoretical Matérn correlations and covariances:
```

```
dd = (0:1000)/1000
SS.theory = (dd*kappa)*besselK(dd*kappa, 1)
SS.theory[1] = 1
S.theory = SS.theory/(4*pi*kappa^2)/tau^2
```

```
#-----#
```

```

#13.# Prediction, plotting posterior prediction maps figure 4

## Calculate mapping between triangulation vertices and grid points:
## Resolution for gridded output was dims=c(300,300)
proj = inla.mesh.projector(mesh123, dims=c(300,300))

## Construct greyscale palette function:
my.grey.palette = function(n,...) { return (grey.colors(n,0.05,0.95,...))}
## Use it:
my.palette = my.grey.palette

## Construct map data appropriate for easy plotting:
mm = calc.map(map)

## Plot results:

par(mfrow=c(1,2))
## Map resulting posterior mean field to a grid:
plotdata = inla.mesh.project(proj, spatiotemp$summary.random$hhid[, "mean"])
## Plot PM contours:

postmap = (levelplot(row.values=proj$x, column.values=proj$y, x=plotdata,
                    mm=mm, panel=levelplotmap,
                    col.regions=my.palette,
                    xlim=range(proj$x), ylim=range(proj$y), aspect="iso",
                    contour=TRUE, cuts=11, labels=FALSE, pretty=TRUE,
                    xlab="Easting", ylab="Northing", main="Posterior mean for
zero inflated Poisson"))
plot(postmap)

## Map resulting posterior standard deviation field to a grid:
plotdata = inla.mesh.project(proj, spatiotemp$summary.random$hhid[, "sd"])
## Plot std.dev. contours:
#dev.new()
postmap2 = (levelplot(row.values=proj$x, column.values=proj$y, x=plotdata,
                    mm=mm, panel=levelplotmap,
                    col.regions=my.palette,
                    xlim=range(proj$x), ylim=range(proj$y), aspect="iso",
                    contour=TRUE, cuts=11, labels=FALSE, pretty=TRUE,
                    xlab="Easting", ylab="Northing", main="Posterior standard
deviation for zero inflated Poisson"))
plot(postmap2)

```

References

- Agarwal, D.K., Gelfand, A.E., Citron-Pousty, S., 2002. Zero-inflated models with application to spatial count data. *Environmental and Ecological Statistics* 9, 341–355.
- Banerjee, S., Carlin, B., 2003. Semiparametric spatio-temporal frailty modeling. *Environmetrics* 14, 523–535.
- Banerjee, S., Carlin, B.P., Gelfand, A.E., 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall/CRC.
- Banerjee, S., Gelfand, A.E., Finley, A.O., Sang, H., 2008. Gaussian predictive process models for large spatial datasets. *Journal of the Royal Statistical Society, Series B*, 70 (Part 4), 825–848.
- Barros, A.J.D., Hirakata, V.N., 2003. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Medical Research Methodology* 3, 21.
- Berliner, L.M., Wikle, C.K., Cressie, N., 2000. Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *Journal of Climate* 13, 3953–3968.
- Black, R.E., Morris, S.S., Bryce, J., 2003. Where and why are 10 million children dying every year? *Lancet* 361, 2226–2234.
- Brezger, A., Lang, S., 2006. Generalized structured additive regression based on Bayesian P-splines. *Computational Statistics and Data Analysis* 50, 967–991.
- Cameletti, M., Lindgren, F., Simpson, D., Rue, H., 2012. Spatio-temporal modelling of particulate matter concentration through the SPDE approach. *ASTA Advances in Statistical Analysis*, accepted.
- Casella, G., Robert, C.P., 1999. *Monte Carlo Statistical Methods*. Springer-Verlag, New York.
- Clark, S.J., Collinson, M.A., Kahn, K., Drullinger, K., Tollman, S.M., 2007. Returning home to die: circular labour migration and mortality in South Africa. *Scandinavian Journal of Public Health Supplement* 69, 35–44.
- Cressie, N.A.C., 1993. *Statistics for Spatial Data*. Wiley, New York.
- Fahrmeir, L., Lang, S., 2001. Bayesian inference for generalized additive mixed models based on Markov Random Field priors. *Journal of the Royal Statistical Society Series C-Applied Statistics* 50, 201–220.
- Fekedulegn, D., Andrew, M., Violanti, J., Hartley, T., Charles, L., Burchfiel, C., 2010. Comparison of statistical approaches to evaluate factors associated with metabolic syndrome. *The Journal of Clinical Hypertension* 12, 365–373.
- French, J.L., Kammann, E.E., Wand, M.P., 2002. Comment on paper by Ke and Wang. *Journal of American Statistical Association* 96, 1285–1288.
- Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. *Markov chain Monte Carlo in Practice*. Chapman & Hall/CRC.
- Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin designs. *Journal of Statistical Planning and Inference* 26, 131–148.
- Kahn, K., Tollman, S.M., Garenne, M., Gear, J.S., 2000. Validation and application of verbal autopsies in a rural area of South Africa. *Tropical Medicine and International Health* 5, 824–831.
- Kamman, E., Wand, M., 2001. *Geoadditive Models*. Department of Biostatistics, School of Public Health, Harvard University.
- Lambert, D., 1992. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34, 1–14.
- Lindgren, F., Rue, H., Lindstrom, J., 2011. An explicit link between Gaussian fields and Gaussian Markov Random Fields: the Stochastic Partial Differential Equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73, 423–498.
- R-cran, 2010. R: A Language and Environment for Statistical Computing. R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria.
- Ridout, M., Hinde, J., Demetrio, C., 2001. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative Binomial alternatives. *Biometrics* 57, 219–223.
- Rose, G., 2001. Sick individuals and sick populations. *International Journal of Epidemiology* 30, 427.
- Rue, H., Held, L., 2005. *Gaussian Markov Random Fields: Theory and Applications*. Chapman & Hall.
- Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71, 319–392.
- Simpson, D., Lindgren, F., Rue, H., 2011. Fast Approximate Inference with INLA: The Past, The Present and The Future. Arxiv preprint arXiv:1105.2982.
- Simpson, D., Lindgren, F., Rue, H., 2011b. Think Continuous: Markovian Gaussian Models in Spatial Statistics. Arxiv preprint arXiv:1110.6796.
- Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 64, 583–639.

StataCorp, 2007. Stata Statistical Software Release 10. StataCorp LP, College Station, TX, USA.

Tollman, S.M., Herbst, K., Garenne, M., Gear, J.S., Kahn, K., 1999. The Agincourt Demographic and Health Study: site description, baseline findings and implications. *South African Medical Journal* 89, 858–864.

Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* 84, 1382–1394.

Wikle, C.K., Berliner, L.M., Cressie, N., 1998. Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics* 5, 117–154.