A Zero-Inflated Spatial Gamma Process Model With Applications to Disease Mapping

L.E. NIETO-BARAJAS and D. BANDYOPADHYAY

In this paper, we introduce a novel discrete Gamma Markov random field (MRF) prior for modeling spatial relations among regions in geo-referenced health data. Our proposition is incorporated into a generalized linear mixed model zero-inflated (ZI) framework that accounts for excess zeroes not explained by usual parametric (Poisson or Negative Binomial) assumptions. The ZI framework categorizes subjects into low-risk and high-risk groups. Zeroes arising from the low-risk group contributes to structural zeroes, while the high-risk members contributes to random zeroes. We aim to identify explanatory covariates that might have significant effect on (i) the probability of subjects in low-risk group, and (ii) intensity of the high risk group, after controlling for spatial association and subject-specific heterogeneity. Model fitting and parameter estimation are carried out under a Bayesian paradigm through relevant Markov chain Monte Carlo (MCMC) schemes. Simulation studies and application to a real data on hypertensive disorder of pregnancy confirms that our model provides superior fit over the widely used conditionally auto-regressive proposition.

Key Words: Bayesian inference; Gamma Markov random field; Gibbs sampling; Latent variables; Mortality; Spatial.

1. INTRODUCTION

Epidemiological and public health studies often generate count data aggregated over a set of disjoint geographical regions with the underlying assumption that 'geographically close' counts would display residual spatial dependence (Wakefield 2007) not explained by relevant confounders/covariates. Analysis of these kinds of data are typically carried out using disease mapping (DM) models (Lawson 2009) which explains the spatial variations in disease risk/incidences, identify areas of unusually high risks and to provide a clean map for better resource allocation and risk assessments.

L.E. Nieto-Barajas (✉) is Professor, Department of Statistics, ITAM, 01080 Mexico City, Mexico (E-mail: *lnieto@itam.mx*). D. Bandyopadhyay is Associate Professor of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN 55455, USA (E-mail: *dbandyop@umn.edu*).

Journal of Agricultural, Biological, and Environmental Statistics, ***Accepted for publication***
DOI: 10.1007/s13253-013-0128-z

Published online: 07 February 2013

Our motivating example here is the study of mortality in pregnant women due to hypertensive disorder (also known as preeclampsia or eclampsia condition) in Mexico in the year 2009. Preeclampsia is a medical condition in which hypertension arises in pregnancy in association with significant amounts of protein in the urine. Eclampsia is an acute and life-threatening complication of pregnancy, characterized by the appearance of tonic-clonic seizures, usually in a patient who had developed preeclampsia. According to the National System on Health Information of Mexico (Annual Statistics Book 2009), 32 % of the deaths of pregnant women in 2009 were caused by a hypertensive disorder, being this the main cause of dead. The maternity mortality rate, defined as the number of maternity deaths for every 100 thousand births, is an indicator that reflects the status of the health services in a country. As reported in Women and Men in Mexico (2009), in 1990, this rate was 90.4 whereas in 2009 the indicator decreased to 55.8. We are interested in estimating the maternity mortality ratio due to hypertensive disorder, smoothed over spatially associated neighboring counties, after controlling for the effect of possible confounders of maternity mortality like the number of medical units, percentage of pregnant women with social security, etc.

To introduce a general framework of DM models, we consider a study region partitioned into n non-overlapping areas with the observed disease count/cases Y_i (for region i) naturally assumed to be Poisson (P) distributed with mean $\lambda_i \times E_i$, where E_i is the expected number of cases for area i under indirect standardization, and λ_i is the relative risk. The classical maximum likelihood (ML) estimate of λ_i is the standardized mortality ratio (SMR) (Marshall 1991) defined as $\hat{\lambda}_i = y_i/E_i$. Because $\text{Var}(\text{SMR}_i) = \lambda_i/E_i$, smaller values of E_i (mostly observed in areas with small expected counts) might inflate the SMR. If overdispersion is present in the count data, an alternative modeling for Y_i is to consider a Negative Binomial (NB) distribution with integer parameter E_i and probability of success $1/(1+\lambda_i)$. This implies that $\text{E}(Y_i) = \lambda_i E_i$, as in the Poisson model, but $\text{Var}(Y_i) = (1+\lambda_i)\lambda_i E_i$. In any case, it is assumed that the expected residual relative risks (not explained by confounders) in proximal areas are more similar than areas that are not 'geographically' close. This spatial variation, along with additional (random) non-spatial noise are taken to consideration while building DM models for providing reliable (efficient) estimates of λ_i.

Inferences for λ_i are usually considered within a hierarchical generalized linear mixed model (GLMM) framework assuming a (latent) spatial process of interest contaminated by observational noise (Best, Richardson, and Thomson 2005). Under a Bayesian framework (both empirical or fully Bayes), these hierarchical spatial (or spatio-temporal) formulations provide easy borrowing of information across the whole study region to provide efficient estimates of the overall spatial risk patterns as well as variance reduction through the use of shrinkage estimators. With the recent advances in Bayesian computation powered by Markov chain Monte Carlo (MCMC) techniques, these spatial dependencies are routinely incorporated into the covariance structure (implicitly or explicitly) through conditionally auto-regressive (CAR) specifications (Banerjee, Wall, and Carlin 2003) and can be implemented using conventional software like `OpenBUGS`. However, maternity mortality being a rare event is observed with considerable sparsity in the counts, and hence the data exhibit

excess of zeroes (or zero-inflation) that cannot be explained by typical P or NB assumptions. This leads to the zero-inflated Poisson (ZIP) or Negative Binomial (ZINB) formulations of our problem. Introduced by Lambert (1992), ZI models have been widely used in dental epidemiology (Böhning et al. 1999), pharmaceutical utilization (Street, Jones, and Furuta 1999), growth/developmental study (Cheung 2002), public health (Lam, Xue, and Cheung 2006), healthcare (Yau, Wang, and Lee 2003) etc. They consist of an underlying (binary) latent construct that dichotomizes the population into two-subpopulations, viz. the low-risk (disease free) group contributing to 'structural' zeroes and the high-risk group where the counts arises from a P/NB regression. The zeroes from the high-risk group are termed 'sample zeroes' and hence we have a mixture formulation where the excess of zeroes are resulting from two separate regimes. Adopting this ZI formulation, the primary objective of our study would be identifying confounders that would significantly predict (i) probability that a subject is from the low-risk (no mortality) group, and (ii) magnitude of the Poisson intensity, given that the subjects belong to high-risk (mortality) group.

The widely explored parametric CAR process assumes a strong (latent) spatial structure, determined by a single association parameter ρ, and models the 'log' of the relative risk λ_i through a multivariate normal density, avoiding direct modeling of λ_i. Despite alternative semi-parametric specifications (Denison and Holmes 2001; Green and Richardson 2002; Knorr-Held and Rasser 2000) introduced to control over-smoothing of relatively heterogeneous risk surfaces, there had been a dearth of considering multivariate specifications of λ_i, other than some Markov random field (MRF) models by Wolpert and Ickstadt (1998) and Ickstadt and Wolpert (1999). Motivated by the Markov gamma process of Nieto-Barajas and Walker (2002), Nieto-Barajas (2008) introduced dependencies between λ_i (under a Bayesian paradigm) not directly through its set of neighbors, but through the edges, by considering a discrete independent latent process in a 2-level hierarchical model, producing a gamma-type MRF with second order properties $\mathrm{E}(\lambda_i) = \alpha/\beta$ and $\mathrm{Var}(\lambda_i) = \alpha/\beta^2$. Our aim in this paper is to introduce a spatial gamma process (SGP) model with strict marginal gamma distributions to be used under a GLMM framework for modeling ZI counts. As in Nieto-Barajas (2008), dependencies in our model are also introduced through the edges via a latent process, but we now take a continuous dependent process, as opposed to a discrete independent process. The dependence in the latent edges is achieved via exchangeability for which we require a third level in the hierarchical structure. The new 3-level hierarchical model has the feature to be strictly stationary with gamma marginal distributions and turns out to be very flexible in capturing local dependencies. Our approach is Bayesian and has the ability to incorporate prior information on model parameters and derive inference using MCMC techniques.

The paper proceeds as follows. In Section 2, we introduce the spatial gamma process for positive random variables and incorporate that into the ZI formulation of a GLMM. Section 3 details the Bayesian inference, related prior and posterior distributions. A simulation study is conducted in Section 4 to compare the performance of our SGP model with CAR alternatives. Section 5 applies our methodology to analyze the motivating data on maternity mortality. Finally, some discussions and concluding comments appear in Section 6.

2. MODEL

2.1. SPATIAL GAMMA PROCESS

Let η_i be a nonnegative random variable associated with area i. If i and j are neighboring areas, we represent it as $i \sim j$. Let ∂_i denote the set of all indices corresponding to the neighbors of area i. We introduce dependence between two neighboring areas, say η_i and η_j, via a latent edge υ_{ij}. The latent edges $\{\upsilon_{ij}\}$ form a set of exchangeable random variables, that is, conditionally independent given a common variable ϕ. This is formally defined in a three-level hierarchical model for the components $\phi \to \{\upsilon_{ij}\} \to \{\eta_i\}$, through three nested gamma distributions. The first level is given by

$$\eta_i \mid (\upsilon_{ij}, j \in \partial_i) \overset{\text{ind}}{\sim} \text{Ga}\left(\alpha + \sum_{j\in\partial_i} \kappa_{ij}, \beta + \sum_{j\in\partial_i} \upsilon_{ij}\right) \tag{2.1}$$

for $i = 1, \ldots, n$. The second level is, $\upsilon_{ij} \mid \phi \overset{\text{ind}}{\sim} \text{Ga}(\kappa_{ij}, \phi)$ where we define $\upsilon_{ij} \equiv \upsilon_{ji}$ and take $\kappa_{ij} = \kappa_{ji}$, for $i \sim j \in \{1, \ldots, n\}$. Finally, the third and last level is $\phi \sim \text{Ga}(\alpha, \beta)$. Using notations, we have $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n) \sim \text{SGP}(\alpha, \beta, \boldsymbol{\kappa})$, where $\boldsymbol{\kappa} = \{\kappa_{ij}\}$ with $\alpha, \beta, \kappa_{ij} > 0$. The first level (2.1) defines the risks η_i's to be conditionally independent given the set of latent edges υ_{ij}'s. The latents $\{\upsilon_{ij}\}$ are a collection of exchangeable random variables, whose joint distribution is described by the second and third levels. An important difference with Nieto-Barajas (2008) is that they consider independent distributions for the υ_{ij}, whereas in this new proposal, any pair $(\upsilon_{ij}, \upsilon_{i'j'})$ are positively correlated with the correlation given by $\sqrt{\kappa_{ij}\kappa_{i'j'}/\{(\alpha+\kappa_{ij}-1)(\alpha+\kappa_{i'j'}-1)\}}$ and the induced marginal distribution follows gamma-gamma (Gga) distribution. In notation, if $X|Y \sim \text{Ga}(c, Y)$ and $Y \sim \text{Ga}(a, b)$, then $X \sim \text{Gga}(a, b, c)$. Thus, $\upsilon_{ij} \sim \text{Gga}(\alpha, \beta, \kappa_{ij})$. These latent edges have an appealing interpretation, they determine the strength of the dependence among the risks $\{\eta_i\}$ in all regions. In fact, the larger the parameter κ_{ij} (and thus the latent υ_{ij}), the stronger the dependence between η_i and η_j. In particular, if $\kappa_{ij} \to 0$ for all i and j then the distribution of the latents υ_{ij} degenerate to a point mass at zero with probability one, which renders the η_i's be independent.

A nice property of the spatial gamma process defined this way is that $\{\eta_i\}$ follows a strictly gamma process with marginals $\text{Ga}(\alpha, \beta)$. This follows from the fact that the latent υ_{ij}'s are conditionally independent (given ϕ) with gamma distribution with common scale ϕ, and hence $\sum_{j\in\partial_i} \upsilon_{ij} \mid \phi$ is also gamma with shape $\sum_{j\in\partial_i} \kappa_{ij}$ and scale ϕ. Now, integrating out ϕ we have (marginally) $\sum_{j\in\partial_i} \upsilon_{ij} \sim \text{Gga}(\alpha, \beta, \sum_{j\in\partial_i} \kappa_{ij})$. From (2.1), we can integrate out $\sum_{j\in\partial_i} \upsilon_{ij}$ using this marginal distribution and obtain $\eta_i \sim \text{Ga}(\alpha, \beta)$ marginally. Unfortunately, the correlation between the risk of two neighboring regions is not available in closed form, except for trivial cases. In the hypothetical setting that a territory consists of only two regions, the correlation between η_1 and η_2 is given by $\text{Corr}(\eta_1, \eta_2) = \kappa_{12}/(\alpha + \kappa_{12} + 1)$. So the dependence is not a function of the scale parameter β, it only depends on the shape parameter α and the association parameter κ_{12}. This confirms the fact that a stronger dependence is achieved when κ_{12} is large relative to α. In general, the dependence between any two adjacent neighbors, say λ_i and λ_j, is controlled

by the parameter κ_{ij}, thus allowing for different degrees of dependence for different neighbors. This distinctive feature of our model is not present in, for example, the popular CAR model where the dependence is controlled by a single parameter ρ.

2.2. ZI-SGP MODELS

We assume that the maternity deaths Y_i are generated independently from a ZI setting where the zeroes correspond to two distinct (latent) health-states; the first state (low-risk group) occurs with probability θ_i and produces structural zeroes, while the other state (high-risk group) occurs with probability $(1-\theta_i)$ and produces random nonnegative counts having intensity λ_i and expected rate E_i. The zeroes from the second state are the sampling zeroes and contributes to the total haul of observed zero counts. This produces a two-component mixture formulation. For convenience, let us consider the Poisson case first. The probability mass function of a ZI Poisson model is given by

$$P(Y_i = y_i) = \begin{cases} \theta_i + (1-\theta_i)e^{-\lambda_i E_i}, & \text{if } y_i = 0, \\ (1-\theta_i)\frac{e^{-\lambda_i E_i}(\lambda_i E_i)^{y_i}}{y_i!}, & \text{if } y_i > 0 \end{cases} \tag{2.2}$$

with $0 \le \theta_i \le 1$ and $\lambda_i, E_i > 0$, for $i = 1, \ldots, n$. It is straightforward to see that $E(Y_i) = (1-\theta_i)\lambda_i E_i$.

For defining the Negative Binomial case, let us recall a connection between Poisson and a Negative Binomial distributions. It is well known (Bernardo and Smith 2000) that if $Y|T \sim \text{Po}(cT)$ and $T \sim \text{Ga}(a,b)$ then marginally $T \sim \text{Pga}(a,b,c)$, that is, Y has a Poisson-gamma distribution. Moreover, if a is an integer then $\text{Pga}(a,b,c) \equiv \text{NB}(a, b/(b+c))$. We note that the Negative Binomial has only two parameters whereas the Poisson-gamma has three parameters, so the Negative Binomial in terms of a Poisson-gamma is overparameterized and therefore we can set one of the parameters, b or c, to a fixed value. Let us take $b = 1$. Now, using the previous ideas, we can construct the Negative Binomial model in terms of a Poisson model as follows: introduce a latent variable T_i such that $Y_i|T_i \sim \text{Po}(\lambda_i T_i)$ with $T_i \sim \text{Ga}(E_i, 1)$ then marginally $Y_i \sim \text{NB}(E_i, 1/(1+\lambda_i))$ as desired. Furthermore, the ZI Negative Binomial model can be reconstructed from the ZIP model (2.2) by replacing E_i by T_i and again taking $T_i \sim \text{Ga}(E_i, 1)$.

In either model, ZIP or ZINB, we connect the relative risk λ_i and the inflation probability θ_i to the set of covariates $\mathbf{x}_i$ and $\mathbf{z}_i$ (not necessarily the same) of dimensions p and q, respectively, through a logarithmic link and a logistic link (McCullagh and Nelder 1989). The total residual heterogeneity (not explained by covariates/confounders) is decomposed into 'spatially' correlated, $\{\eta_i\}$, and uncorrelated, $\{\xi_i\}$ and $\{\zeta_i\}$, random effects, often defined as a 'convolution prior' (Lawson 2009) and log-transformed version of both of them enters additively in two regressions. In general, these two random effects are not identifiable (Eberley and Carlin 2000), however, this parametrization explains the total effect of unobserved confounding. We thus have

$$\begin{aligned} \log(\lambda_i) &= \boldsymbol{\gamma}'\mathbf{x}_i + \log(\xi_i) + \log(\eta_i), \\ \text{logit}(\theta_i) &= \boldsymbol{\delta}'\mathbf{z}_i + \log(\zeta_i). \end{aligned} \tag{2.3}$$

Here, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are coefficient vectors for the fixed-effects; ξ_i and ζ_i capture area-wise heterogeneity, such that $\xi_i \overset{\text{iid}}{\sim} \text{Ga}(\beta_\xi, \beta_\xi)$ and $\zeta_i \overset{\text{iid}}{\sim} \text{Ga}(\beta_\zeta, \beta_\zeta)$, independently and η_i captures regional spatial clustering, such that $\{\eta_i\} \sim \text{SGP}(\alpha, \alpha, \boldsymbol{\kappa})$. Note that for identifiability issues, we constrain all random effects such that $\text{E}(\xi_i) = \text{E}(\zeta_i) = \text{E}(\eta_i) = 1$ a priori. Re-writing the expressions (2.3) for λ_i and θ_i in terms of its original scale, we obtain the following multiplicative model:

$$\lambda_i = \eta_i \xi_i e^{\boldsymbol{\gamma}' \mathbf{x}_i} \quad \text{and} \quad \theta_i = \zeta_i e^{\boldsymbol{\delta}' \mathbf{z}_i} / \left(1 + \zeta_i e^{\boldsymbol{\delta}' \mathbf{z}_i}\right). \tag{2.4}$$

3. BAYESIAN INFERENCE

3.1. HYPER-PRIOR DISTRIBUTIONS

Our model is fully specified for Bayesian inference once we determine the values for the hyper-parameters $\boldsymbol{\vartheta} = (\beta_\xi, \beta_\zeta, \alpha, \boldsymbol{\kappa}, \boldsymbol{\gamma}, \boldsymbol{\delta})$, with $\boldsymbol{\kappa} = \{\kappa_{ij} : i < j = 1, \ldots, n\}$. Instead of fixing them, we assign hyper-priors to all of them to pull strength across all regions. In particular we consider that all parameters $\boldsymbol{\vartheta}$ are independent a priori, with prior distributions given as $\beta_\xi \sim \text{Ga}(a_\beta, b_\beta)$, $\beta_\zeta \sim \text{Ga}(a_\beta, b_\beta)$, $\alpha \sim \text{Ga}(a_\alpha, b_\alpha)$, $\kappa_{ij}|\omega \sim \text{Ga}(1, \omega)$ and $\omega \sim \text{Ga}(a_\omega, b_\omega)$ for $i \sim j \in \{1, \ldots, n\}$. Moreover, $\gamma_s \sim \text{N}(0, \sigma_\gamma^2)$ and $\delta_r \sim \text{N}(0, \sigma_\delta^2)$ independently for $s = 1, \ldots, p$ and $r = 1, \ldots, q$.

3.2. POSTERIOR

Since the ZINB model can be written as a conditionally ZIP model, we present posterior inference for the ZIP model first. The sampling distribution of the ZIP model in (2.2) has a mixture representation and posterior inference of mixture models is greatly simplified by introducing an auxiliary variable that indicates from what mixture component the data are coming from. If we let $W_i \sim \text{Ber}(\theta_i)$, for $i = 1, \ldots, n$ independently, then the joint distribution of $\mathbf{Y}$ and $\mathbf{W}$ becomes

$$f(\mathbf{y}, \mathbf{w}|\boldsymbol{\lambda}, \boldsymbol{\theta}) = \prod_{i=1}^{n} \left\{\theta_i I(y_i = 0)\right\}^{w_i} \left\{(1 - \theta_i)(1/y_i!) e^{-\lambda_i E_i} (\lambda_i E_i)^{y_i}\right\}^{1 - w_i}, \tag{3.1}$$

where λ_i and θ_i are given in (2.4). This joint distribution will be used as extended likelihood for obtaining the posterior distribution of the components of λ_i's and the θ_i's. Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)$ and $\boldsymbol{\upsilon} = \{\upsilon_{ij}, i < j = 1, \ldots, n\}$, then the joint prior distribution for the whole parameter vector is

$$f(\boldsymbol{\eta}, \boldsymbol{\upsilon}, \phi, \boldsymbol{\xi}, \boldsymbol{\zeta}, \boldsymbol{\vartheta}) = f(\boldsymbol{\eta}, \boldsymbol{\upsilon}, \phi|\boldsymbol{\vartheta}) f(\boldsymbol{\zeta}_\lambda) f(\boldsymbol{\xi}|\boldsymbol{\vartheta}) f(\boldsymbol{\zeta}|\boldsymbol{\vartheta}) f(\boldsymbol{\vartheta})$$

with

$$f(\boldsymbol{\eta}, \boldsymbol{\upsilon}, \phi \mid \boldsymbol{\vartheta}) = \left\{\prod_{i=1}^{n} \text{Ga}\left(\eta_i \,\middle|\, \alpha + \sum_{j \in \partial_i} \kappa_{ij}, \alpha + \sum_{j \in \partial_i} \upsilon_{ij}\right)\right\} \times \left\{\prod_{i<j} \text{Ga}(\upsilon_{ij} \mid \kappa_{ij}, \phi)\right\} \text{Ga}(\phi \mid \alpha, \alpha),$$

$f(\boldsymbol{\xi}) = \prod_{i=1}^{n} \text{Ga}(\xi_i|\beta_\xi, \beta_\xi)$, $f(\boldsymbol{\zeta}) = \prod_{i=1}^{n} \text{Ga}(\zeta_i|\beta_\zeta, \beta_\zeta)$ and $f(\boldsymbol{\vartheta})$ is given by the product of the hyper prior distributions presented in Section 3.1.

Posterior characterization of the model will be based on the posterior conditional distributions for which we will require a Gibbs sampling (Smith and Roberts 1993) strategy. Outline of the conditional posteriors appear in the Appendix. For derivation of the conditionals for parameters related to θ_i, we introduced an extra latent variable V_i. We recall that for fitting a ZI Negative Binomial we replace the number of exposed E_i by a latent variable T_i. The conditional distributions of V_i and T_i are also included in the appendix. The full conditional distribution of W_i, V_i, T_i, η_i, ξ_i, ζ_i and ϕ are in standard (closed) form so they can be simulated straightforwardly using Gibbs sampling. However, for simulating from the others, we will require Metropolis–Hastings steps (Tierney 1994).

4. SIMULATION STUDY

In this section, we first assess the performance of our model and compare it with a CAR model alternative via a simulation study. We consider non-ZIP and ZIP scenarios equipped with SGP and CAR spatial random effects as ground truth. Thus, we have a total of four scenarios: P + CAR, ZIP + CAR, P + SGP and ZIP + SGP. The ZIP model together with SGP as spatial random effects is described in Equations (2.2) and (2.3). For the Poisson model, we simply set $\theta_i = 0$ in (2.2). For the (convoluted) CAR specifications, we replace Equations (2.3) by the following setting:

$$\begin{aligned} \log(\lambda_i) &= \boldsymbol{\gamma}'\mathbf{x}_i + \xi_i^* + \eta_i^*, \\ \text{logit}(\theta_i) &= \boldsymbol{\delta}'\mathbf{z}_i + \zeta_i^*, \end{aligned} \tag{4.1}$$

where $\xi_i^* \overset{\text{iid}}{\sim} \text{N}(0, \sigma_\xi^2)$ and $\zeta_i^* \overset{\text{iid}}{\sim} \text{N}(0, \sigma_\zeta^2)$, independently for $i = 1, \ldots, n$, and $\{\eta_i^*\} \sim \text{CAR}(\tau^2, \rho)$. The CAR model is defined as a multivariate normal distribution with mean vector zero and variance-covariance matrix $\Sigma = \tau^2(D_W - \rho W)^{-1}$, where $W = (w_{ij})$ is the proximity matrix such that $w_{ij} = I(i \sim j)$, and $D_W = \text{diag}(m_1, \ldots, m_n)$ with $m_i = \sum_{j=1}^{n} w_{ij}$.

We consider two explanatory variables, one in each of the regressions given by $\boldsymbol{\gamma}'\mathbf{x}_i = \gamma_0 + \gamma_1 x_i$ and $\boldsymbol{\delta}'\mathbf{z}_i = \delta_0 + \delta_1 z_i$, respectively. Model specifications for these fixed effects are those proposed in Jung, Jhun, and Lee (2005), i.e., $x_i \sim \text{Unif}(0, 1)$ and $z_i \sim \text{Unif}(0, 0.5)$ independently, and the parameter values are set as $\gamma_0 = 1.2$, $\gamma_1 = 0.5$, $\delta_0 = -1.5$, $\delta_1 = 1$. These specifications imply that the intensities and the zero inflation probability will be around 4.26 and 0.22, respectively, for all regions.

The variance of the area specific random effects was fixed to σ^2 in both the convoluted CAR and SGP, that is, we take $\sigma_\xi^2 = \sigma_\zeta^2 = \sigma^2$ for the CAR and $\beta_\xi = \beta_\zeta = 1/\sigma^2$ for the SGP. Different values of σ^2, say 0.1 and 1, were considered for comparison. The variance of the spatial effects, $\{\eta_i^*\}$ for the CAR and $\{\eta_i\}$ for the SGP, was fixed to one. This is achieved by taking $\tau^2 = 1$ and $\alpha = 1$, in the corresponding model. For the association

Table 1. Maternity mortality in Mexico. The columns are State name (State), number of deaths due to hypertensive disorder (Y), number of births in 100 thousands (E), number of medical units (X_1), proportion of pregnant women with social security (X_2), proportion of pregnant women who were seen by a physician in the first trimester of pregnancy (X_3), public expenditure in health per capita in thousands of Mexican pesos (X_4), poverty index (Z_1), proportion of births in clinics or hospitals (Z_2), sector indicator (S), edge indicator (ED), posterior estimates of λ ($\hat{\lambda}$) and θ ($\hat{\theta}$) from the ZIP+SGP model, and adjacency relationships.

State	Y	E	X_1	X_2	X_3	X_4	Z_1	Z_2	S	ED	$\hat{\lambda}$	$\hat{\theta}$	Adjacent states
1. Aguascalientes	2	0.14	0.13	0.60	0.42	3.59	−0.91	0.98	1	0	8.81	0.006	14, 32
2. Baja California	0	0.23	0.29	0.63	0.42	3.21	−1.14	0.59	1	1	5.18	0.542	3, 26
3. Baja California Sur	0	0.07	0.15	0.66	0.41	5.51	−0.68	0.95	1	1	7.01	0.007	2
4. Campeche	1	0.09	0.25	0.46	0.47	4.99	0.43	0.86	3	1	9.75	0.000	23, 27, 31
5. Coahuila	0	0.18	0.41	0.72	0.37	3.23	−1.14	0.92	1	1	4.45	0.051	8, 10, 19, 32
6. Colima	0	0.08	0.19	0.57	0.47	4.27	−0.78	0.98	2	1	5.77	0.011	14, 16
7. Chiapas	4	0.39	1.65	0.19	0.37	2.11	2.32	0.26	3	1	12.18	0.028	20, 27, 30
8. Chihuahua	2	0.22	0.59	0.65	0.35	3.37	−0.52	0.75	1	1	7.37	0.034	5, 10, 25, 26
9. Distrito Federal	8	0.50	0.56	0.54	0.32	8.06	−1.48	0.96	2	0	16.59	0.031	15, 17
10. Durango	1	0.16	0.51	0.53	0.36	3.25	0.05	0.81	1	0	8.13	0.001	5, 8, 18, 25, 32
11. Guanajuato	2	0.62	0.69	0.40	0.45	2.59	0.06	0.83	2	0	5.05	0.001	14, 16, 22, 24, 32
12. Guerrero	8	0.34	1.13	0.26	0.32	2.61	2.53	0.58	2	1	21.06	0.003	15, 16, 17, 20, 21
13. Hidalgo	6	0.25	0.96	0.33	0.33	2.56	0.66	0.75	2	0	19.02	0.001	15, 21, 22, 24, 29, 30
14. Jalisco	3	0.65	1.05	0.51	0.43	2.98	−0.82	0.87	2	1	5.20	0.016	1, 6, 11, 16, 18, 24, 32
15. Mexico	18	1.14	1.44	0.45	0.33	2.35	−0.55	0.94	2	0	14.10	0.002	9, 12, 13, 16, 17, 21, 22, 29
16. Michoacan	1	0.40	1.05	0.29	0.41	2.47	0.53	0.84	2	0	6.60	0.000	6, 11, 12, 14, 15, 22
17. Morelos	5	0.17	0.29	0.38	0.40	3.05	−0.27	0.84	2	0	18.15	0.003	9, 12, 15, 21
18. Nayarit	0	0.12	0.41	0.48	0.44	3.47	0.12	0.84	1	1	5.94	0.001	10, 14, 25, 32
19. Nuevo Leon	0	0.29	0.60	0.68	0.45	3.19	−1.38	0.96	1	1	3.04	0.051	5, 24, 28, 32
20. Oaxaca	4	0.29	1.47	0.24	0.35	2.43	2.15	0.57	2	1	14.63	0.001	7, 12, 21, 30
21. Puebla	1	0.48	1.17	0.29	0.35	2.02	0.71	0.80	2	0	6.53	0.001	12, 13, 15, 17, 20, 29, 30
22. Queretaro	3	0.19	0.30	0.52	0.45	2.68	−0.26	0.95	2	0	8.58	0.001	11, 13, 15, 16, 24
23. Quintana Roo	0	0.13	0.23	0.57	0.36	3.04	−0.42	0.90	3	1	6.66	0.005	4, 31
24. San Luis Potosi	3	0.24	0.62	0.45	0.40	2.47	0.56	0.86	1	0	9.80	0.000	11, 13, 14, 19, 22, 28, 30, 32
25. Sinaloa	1	0.27	0.51	0.57	0.37	3.26	−0.26	0.95	1	1	6.32	0.001	8, 10, 18, 26
26. Sonora	3	0.25	0.43	0.64	0.44	3.37	−0.70	0.95	1	1	6.65	0.004	2, 8, 25
27. Tabasco	7	0.34	0.69	0.33	0.43	4.08	0.47	0.75	3	1	16.00	0.001	4, 7, 30
28. Tamaulipas	3	0.33	0.57	0.59	0.37	3.28	−0.72	0.94	1	1	7.93	0.004	19, 24, 30
29. Tlaxcala	2	0.15	0.22	0.31	0.38	2.69	−0.15	0.92	2	0	14.94	0.001	13, 15, 21
30. Veracruz	10	0.68	1.70	0.29	0.42	2.89	1.08	0.72	2	1	12.99	0.001	7, 13, 20, 21, 24, 27, 28
31. Yucatan	0	0.15	0.33	0.50	0.42	3.51	0.42	0.92	3	1	6.05	0.001	4, 23
32. Zacatecas	1	0.18	0.47	0.38	0.48	2.86	0.10	0.92	1	0	6.97	0.001	1, 5, 10, 11, 14, 18, 19, 24

parameters we took $\rho = 0.5$ for the CAR and vector $\boldsymbol{\kappa}$ defined as

$$\kappa_{ij} = \begin{cases} 4 & \text{if } s_i = s_j = 1, \\ 2 & \text{if } s_i = s_j = 2, \\ 1 & \text{if } s_i = s_j = 3, \\ 0.25 & \text{if } s_i \neq s_j \end{cases}$$

Table 2. Mean square error of model parameters averaged over 100 repetitions of the experiment. For the parameters, the superscripts (1) and (0) denote boundary and non-boundary states, while free of superscripts indicate estimates for the full dataset. The first column denotes the experiment simulated (E), and the first row denotes the model fitted (F).

E/F	(1) P+CAR		(2) ZIP+CAR				(3) P+SGP		(4) ZIP+SGP			
	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	δ_0 $\theta^{(0)}$	δ_1 $\theta^{(1)}$	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	δ_0 $\theta^{(0)}$	δ_1 $\theta^{(1)}$
						$\sigma = 0.1$						
(1)	0.20	0.42	0.15	0.37	–	–	0.80	0.42	0.72	0.37	–	–
	9.04	17.59	8.20	15.28	0.20	0.20	8.85	18.05	10.57	17.28	0.05	0.05
(2)	0.65	0.93	0.35	0.77	3.01	9.28	1.19	0.88	1.10	0.63	3.00	9.13
	10.53	15.02	19.21	44.23	0.05	0.05	11.65	15.51	14.04	20.82	0.05	0.05
(3)	1.44	0.63	0.81	0.57	–	–	1.82	0.57	2.10	0.65	–	–
	13.11	14.51	15.15	19.57	0.22	0.22	13.26	13.66	11.56	11.89	0.08	0.08
(4)	2.32	1.46	1.90	1.17	3.56	9.65	2.91	1.15	2.31	0.95	3.82	9.57
	17.58	20.12	42.00	74.10	0.05	0.05	16.12	16.77	28.25	29.18	0.05	0.05
						$\sigma = 1$						
(1)	0.41	1.19	0.39	0.93	–	–	1.07	1.17	1.08	0.99	–	–
	108.66	174.80	157.06	581.42	0.21	0.21	157.54	137.43	138.03	166.10	0.07	0.08
(2)	1.76	2.30	0.79	1.88	3.38	10.03	1.36	1.76	1.31	1.64	4.61	10.08
	98.22	100.77	1053.66	8733.43	0.05	0.05	72.45	97.74	108.18	174.48	0.05	0.05
(3)	2.73	1.49	0.96	1.21	–	–	2.54	1.74	2.30	1.46	–	–
	38.50	44.56	132.30	303.19	0.25	0.25	28.87	36.14	26.27	33.86	0.14	0.14
(4)	3.72	2.32	2.28	2.06	4.06	10.77	3.17	2.09	2.64	1.81	4.66	10.07
	22.63	31.02	1197.00	5529.26	0.06	0.07	31.93	38.78	69.83	69.10	0.06	0.06

for the SGP, where the entire spatial study region (described later) is divided into three sectors (1 = north, 2 = center and 3 = south) representing varying degrees (high, medium and low) of spatial association, respectively, and the s_i's are indicators determining the sector membership for region i. Neighboring regions belonging to different sectors will have minimum interaction. The study region for our spatial test-bed is chosen according to the adjacency relationships generated by the 32 states of the country of Mexico. For the SGP, the spatial domain was divided into three big sectors say, north, center and south-east. The adjacency relationships as well as the s_i indicators are available in Table 1.

Under this set-up, we generated ZIP data in the following way: if a generated Unif(0, 1) random variable is less than θ_i then Y_i is set as 0, else Y_i is generated from $\text{Po}(\lambda_i E_i)$ distribution with $E_i = 1$, for $i = 1, \ldots, 32$. For the non-ZIP scenario, we directly sample from the Poisson distribution. This set-up was replicated 100 times. For each replication, we fitted the four models we are considering, that is, (1) P + CAR, (2) ZIP + CAR, (3) P + SGP and (4) ZIP + SGP.

Prior specifications of our model based on SGP and for the model based on CAR rely on non-informative (vague) priors for all the parameters. In particular, we took $a_\alpha = b_\alpha = a_\beta = b_\beta = a_\omega = b_\omega = 0.1$ for the SGP model, and $a_\xi = b_\xi = a_\zeta = b_\zeta = a_\omega = b_\omega = 0.1$ for the CAR model. In all cases, normal priors were taken for the coefficients γ_r and δ_r for

Table 3. Coverage of 95 % posterior CI of model parameters computed over 100 repetitions of the experiment. For the parameters, the superscripts (1) and (0) denote boundary and non-boundary states, while free of superscripts indicate estimates for the full dataset. The first column denotes the experiment simulated (E), and the first row denotes the model fitted (F).

E/F	(1) P+CAR		(2) ZIP+CAR				(3) P+SGP		(4) ZIP+SGP			
	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	δ_0 $\theta^{(0)}$	δ_1 $\theta^{(1)}$	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	γ_0 $\lambda^{(0)}$	γ_1 $\lambda^{(1)}$	δ_0 $\theta^{(0)}$	δ_1 $\theta^{(1)}$
					$\sigma = 0.1$							
(1)	96	97	96	96	62	–	97	94	99	95	90	–
	88	83	91	88	62	62	89	83	90	88	90	90
(2)	88	100	95	96	98	100	97	98	98	100	99	100
	67	63	91	88	98	98	65	62	92	90	99	99
(3)	46	97	57	97	59	–	74	97	80	98	85	–
	51	52	60	66	59	59	52	52	67	67	85	85
(4)	55	97	61	98	100	100	66	97	78	98	98	100
	39	40	60	61	100	100	39	42	60	60	99	99
					$\sigma = 1$							
(1)	96	97	91	95	60	–	95	92	97	95	86	–
	65	61	76	74	60	60	61	61	75	73	86	86
(2)	71	94	95	94	98	100	95	94	98	93	99	100
	37	40	76	78	99	99	42	43	80	77	99	99
(3)	48	95	69	97	56	–	78	93	79	98	75	–
	36	35	70	71	56	56	42	42	69	70	75	75
(4)	43	96	68	97	99	100	73	97	80	94	100	100
	32	31	63	67	100	100	33	32	65	67	100	100

$r = 0, 1$. These priors were centered at zero with prior variances $\sigma_\gamma^2 = \sigma_\delta^2 = 100$. For fitting the (convoluted) CAR model, we put hyper-priors on σ_ξ^2, σ_ζ^2 and τ^2 as $1/\sigma_\xi^2 \sim \text{Ga}(a_\xi, b_\xi)$, $1/\sigma_\zeta^2 \sim \text{Ga}(a_\zeta, b_\zeta)$ and $1/\tau^2 \sim \text{Ga}(a_\tau, b_\tau)$ independently. Additionally, the prior for the association parameter was $\rho \sim \text{Unif}(0, 1)$. We ran a MCMC chain of 50,000 iterations with a burn-in size of 5000. To reduce auto-correlation, we considered a spacing of size 5. For each experiment, we computed the posterior mean square error (MSE) and 95 % posterior credible intervals (CI) for each of the parameters $\gamma_0, \gamma_1, \delta_0, \delta_1$. The average MSE across the 100 replicates of the experiment as well as the empirical coverage of the 95 % posterior CI are reported in Tables 2 and 3.

Interestingly, 19 of the 32 states of Mexico (corresponding to 59 %) lie on the sea boundary. As pointed out by a referee, it is worth studying possible edge effects in the estimation process (see for example, Griffith 1983). To assess this, we study the coverage of the 95 % CI and the MSE when estimating λ_i and θ_i for $i = 1, \ldots, 32$. The average MSE over the 100 replicates as well as the empirical coverage was summarized for those states that lie on the boundary, and for those that do not, by taking the average. These numbers are also reported as the second row for each scenario in Tables 2 and 3. We note that when the ground truth is a non ZIP scenario (cases 1 and 3) and the fitted model is a ZIP model (cases 2 and 4), the mean square error and the empirical coverage with respect to γ_0 and γ_1

are not given. However, when the simulated scenario is a non ZIP case, the parameter θ is set to zero, and when we fit a ZIP model we can actually provide an estimate of its value. Average over posterior estimates of $1-\theta$ is reported instead of the coverage of γ_0 and γ_1, and instead of the coverage of $\theta^{(0)}$ and $\theta^{(1)}$ in Table 3.

From Table 2, we observe that for $\sigma = 0.1$ (which corresponds to a small noise in the individual random effects), the mean square errors are comparable between CAR and SGP cases when sampling from and fitting a non ZIP model (scenarios 1 and 3). However, when sampling from and fitting a ZIP model (scenarios 2 and 4), the mean square errors are smaller for the SGP, regardless whether the ground truth was SGP or CAR. The effect is more clear specially for $\lambda^{(0)}$ and $\lambda^{(1)}$. This discrepancies are emphasized when setting a moderate noise in the individual random effects ($\sigma = 1$). This implies that when the uncertainty in the specific random effects becomes larger, the posterior distributions of the parameters induced by the CAR model are less precise, producing larger MSE's. However, for the SGP, the posterior distribution of the parameters is less sensitive to the amount of uncertainty in the specific random effects.

In terms of the coverage (see Table 3), the values are comparable between CAR and SGP fittings when sampling from the two CAR scenarios. However, when sampling from the SGP scenarios (3 and 4), the parameter γ_0 (first value of first row in each block) shows a consistently smaller coverage when fitting a CAR model compared with that obtained when fitting a SGP model. This behavior is consistent regardless of σ being small or moderate. Another important aspect of this table is that we can compare how good a ZI model is to estimate the probability of non ZI when the data were actually generated from a non ZI scenario. This case corresponds to sampling scenarios 1 and 3 and fitting models 2 and 4. The third number in first row corresponds to the average of posterior estimates for the non ZI probability $1-\theta$. Likewise, the third and fourth numbers in second row correspond to estimates of the same probability but for non-boundary and boundary states, respectively. We see that the estimates obtained with the SGP models are a lot higher (as it should be) than those obtained with the CAR models, with the differences ranging from 19 to 26 percentage points according to the sampling scenario. Therefore, the SGP models show a higher power to distinguish between ZI and non ZI datasets.

In some cases the coverage values reported in Table 3 are smaller than the nominal value of 95 %, the problem intensifies for the SGP sampling scenarios and for σ larger. We believe this is due to the small size (32 states) of the test bed chosen.

Regarding the non boundary and boundary states, for all scenarios considered in Table 2, the mean square errors when estimating λ is smaller for those states in-land ($\lambda^{(0)}$) than for those on the sea boundary ($\lambda^{(1)}$). On the other hand, no edge effect is appreciated when estimating θ where the MSE's are the same. In terms of the coverage (Table 3), no edge effect is appreciated since the coverage shows practically no difference for non boundary and boundary states when estimating λ and θ.

5. EXAMPLE: MEXICAN MATERNITY MORTALITY DATA

We now apply our proposed methodology to the motivating and publicly available data on maternity mortality in Mexico due to hypertensive disorder mentioned in Sec-

tion 1 and presented in Table 1. Reported are the observed death counts of mothers (Y) and the number of births (E) in 100 thousands. By choosing this as the expected count, it turns out that λ becomes the maternity mortality rate (number of maternity deaths for every 100 thousand births). Additionally, for each of the 32 states of Mexico, information on state-level covariates are available, which includes the number of medical units (X_1) (including hospitals and clinics), proportion of pregnant women with social security (X_2), proportion of pregnant women who were seen by a physician in the first trimester of pregnancy (X_3), public expenditure in health per capita in thousands of Mexican pesos (X_4), poverty index (Z_1), and proportion of births in clinics and hospitals (Z_2). The dataset is available from the National System on Health Information of Mexico (*http://www.sinais.salud.gob.mx/publicaciones/index.html*). With 25 % zero counts, a naive Poisson fitting to this dataset (assuming a homogeneous sample) produces a maximum likelihood (ML) estimate of zero counts be 4.53 %. Considering this underestimation, a ZI model might be appropriate here, which would explain the two latent regimes of the origin of zeroes.

We implemented our SGP model using specifications in both the regressions for λ_i and θ_i as described in Section 2. Note that in the descriptions of the covariates, in the previous paragraph, there was implicitly a division of them into groups $\mathbf{x}$ and $\mathbf{z}$. We first took the same group of covariates in both regressions but realized that choosing separate groups achieves a better fit. The best selection is the one taken here. We considered both the P and NB choices. Additionally, we also compare our models with (convoluted) CAR random effects (individual + spatial) as described by (4.1). To test for zero inflation, we also fitted non ZI and ZI P and NB models.

To avoid numerical problems, the number of medical units (X_1) was divided by 1000 and log transformed. We considered non-informative (vague) priors for all the parameters in the models. In particular, we choose $a_\alpha = 3$, $b_\alpha = a_\beta = b_\beta = a_\omega = b_\omega = 0.1$ for our SGP model, and $a_\xi = b_\xi = a_\zeta = b_\zeta = a_\tau = b_\tau = 0.1$, together with $\rho \sim \text{Unif}(0, 1)$, for the CAR model. Additionally, $\sigma_\gamma^2 = \sigma_\delta^2 = 100$ were the prior variances of the fixed effects. We ran MCMC chains for 100,000 iterations with a burn-in of 10,000. The posterior estimates were computed with a spacing of 10 to reduce auto-correlations.

To compare among the different competing models, we computed two sets of goodness of fit (g.o.f) statistics, named conditional predictive ordinates (CPO) (Gelfand, Dey, and Chang 1992) and the L-measure (Ibrahim and Laud 1994) based on the posterior predictive density. Because the ZI models are mixture propositions, we avoided using the deviance information criterion (DIC) (Spiegelhalter et al. 2002) for model comparisons. The CPO's are defined as follows: if we denote by $D^{(-i)}$ the data with the ith observation removed, then for each i, $\text{CPO}_i = f(y_i|D^{(-i)})$. Since the CPO is a g.o.f measure for each observation, we can summarize it for all of the data via a single value given by $\text{ALPML} = (1/n)\sum_i \log(\text{CPO}_i)$, where ALPML is called average log-pseudo marginal likelihood. The ALPML is computationally stable and well-defined even under improper priors, and is a very popular tool for Bayesian model comparison and assessments. Larger values of CPO's and ALPML indicates better fit. On the other hand, the L-measure is based on a function of the variance and bias of the predictive distribution of each y_i, defined as $\text{L}(\nu) = (1/n)\sum_{i=1}^n \text{Var}(y_i^F|\mathbf{y}) + (\nu/n)\sum_{i=1}^n \{\text{E}(y_i^F|\mathbf{y}) - y_i\}^2$ where y_i^F is the

Table 4. Goodness of fit statistics for different competing models for the maternity mortality dataset. Reported are ALPML, L-measure for $\nu = 0, 1/2, 1$, and Bias $= \mathrm{L}(1) - \mathrm{L}(0)$ for the full data (rows 1–5), non-boundary states$^{(0)}$ (rows 6–10) and boundary states$^{(1)}$ (rows 11–15).

g.o.f.	P	ZIP	NB	ZINB	P	ZIP	NB	ZINB
	SGP				CAR			
AMPML	−1.98	−1.97	−2.12	−2.11	−2.03	−2.02	−2.09	−2.08
L(0)	5.19	5.27	5.98	6.06	5.12	5.35	5.98	6.13
L(1/2)	5.67	5.77	6.01	6.08	5.74	5.94	6.00	6.15
L(1)	6.16	6.27	6.03	6.10	6.35	6.53	6.02	6.16
Bias	0.97	0.99	0.04	0.04	1.22	1.18	0.04	0.03
ALPML$^{(0)}$	−2.47	−2.45	−2.64	−2.67	−2.52	−2.55	−2.63	−2.62
L(0)$^{(0)}$	6.72	6.99	7.86	7.98	6.72	7.08	7.80	8.12
L(1/2)$^{(0)}$	7.49	7.76	7.90	8.02	7.75	8.09	7.84	8.14
L(1)$^{(0)}$	8.27	8.53	7.94	8.06	8.78	9.10	7.87	8.16
Bias$^{(0)}$	1.55	1.54	0.08	0.08	2.07	2.01	0.07	0.04
ALPML$^{(1)}$	−1.65	−1.64	−1.76	−1.73	−1.69	−1.65	−1.69	−1.73
L(0)$^{(1)}$	4.14	4.10	4.70	4.74	4.04	4.17	4.73	4.77
L(1/2)$^{(1)}$	4.43	4.41	4.71	4.75	4.36	4.48	4.74	4.78
L(1)$^{(1)}$	4.72	4.71	4.72	4.76	4.68	4.78	4.75	4.79
Bias$^{(1)}$	0.58	0.62	0.02	0.02	0.65	0.61	0.02	0.02

predicted value for y_i and $\nu \in (0, 1)$ is a weighting term which determines a trade-off between variance and bias. The bias itself can be obtained from the L-measure through $\mathrm{Bias} = \mathrm{L}(1) - \mathrm{L}(0)$. Smaller values of the L-measure and the Bias indicate better fit.

Note that the L-measure defined here is an average, which results of dividing by n the original measure presented in Ibrahim and Laud (1994). The reason for doing this is to assess edge effects and compare L-measures considering only those states that are on the boundary (L-measure$^{(1)}$), and those that are not on the boundary (L-measure$^{(0)}$). With the same spirit we compute ALPML$^{(0)}$ and ALPML$^{(1)}$ as an average of log(CPO)'s for those states that are not on the boundary and for those that are on the boundary, respectively. Table 4 summarizes the values of these statistics for the competing models.

Several conclusions can be drawn from Table 4. According to the ALPML, denoted by A in the table, the two Poisson scenarios (non ZI and ZI) together with SGP show a better fit to the data than the corresponding Poisson fits with CAR alternatives. On the other hand, for the two Negative Binomial cases (non ZI and ZI), the fitting is better when using CAR compared with SGP. Overall, the best fit is achieved by the Poisson models when using SGP, with an slight advantage for the ZI case over the non ZI. This advantage is slightly larger in the non boundary states.

In terms of the L-measures, the smallest and the largest predictive variance ($\nu = 0$) are obtained by the P and ZINB models, respectively, both with CAR random effects. In all cases, the variances produced by the NB models are larger than that of the P models. However, as suggested by Ibrahim and Laud (1994), the best way of comparing models is to use $\nu = 1/2$, which gives the best trade-off between variance and bias. The models that produce the smallest L-measure with $\nu = 1/2$ are the two Poisson models, non ZI and ZI,

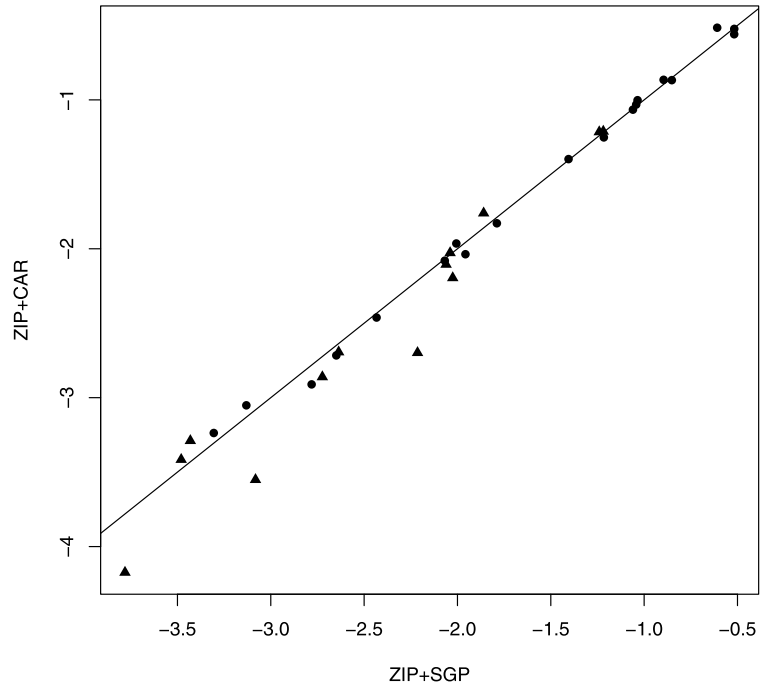

Figure 1. Scatter plot of log(CPO)'s when fitting ZIP+SGP vs. ZIP+CAR to the maternity mortality data. The boundary and non-boundary states are represented by (●) and (△) respectively.

with SGP as random effects. In terms of the bias, the best models are the NB regardless of it being non ZI or ZI and regardless of SGP or CAR random effects. In general for all cases and considering all g.o.f measures, the fitting is better for those states that are on the boundary than for those that are not on the boundary.

The P model produces smaller ALPML values and smaller variances than the NB model. However, the bias obtained with the NB is remarkably smaller. The gain in bias by the NB model is counterbalanced with producing a huge predictive variance, we therefore prefer the more conservative P model which in the end produces a better overall fitting given by the ALPML statistic. Finally, the need for a ZI proposition is consistently supported by the ALPML values. In summary, the best fitting is attained by the ZIP + SGP model.

Before we proceed to the interpretation of the analysis of the Mexican mortality dataset, we graphically compare the fitting of the SGP vs. the CAR cases in the ZIP model. Figure 1 presents a scatter plot of the log(CPO)'s of these two competing models. Additionally, a different symbol is given to the points to assess edge differences. Dots denote boundary states whereas non boundary states are denoted by a triangle. As we can see, most of the points lie on the 45 degrees line with a larger discrepancy for the non boundary states. Specially there are three triangles that lie clearly below the line showing that the fitting is better (larger log(CPO)) with the SGP than with the CAR alternative. In fact 18 out of 32 points (approximately 56 %) lie below the line, confirming the superiority of the SGP model for this dataset.

Table 5. Posterior summaries of regression parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ under ZIP models with SGP and CAR convoluted random effects. Reported are mean, 95 % credible interval and P($\cdot < 0$). Bold numbers denote significant effect at 10 % (i.e. if the parameter has at least 0.90 probability of being positive or negative).

		ZIP+SGP			ZIP+CAR		
Coef.	Covariate	Mean	95 % CI	P($\cdot < 0$)	Mean	95 % CI	P($\cdot < 0$)
γ_0	intercept	4.70	(1.68, 7.17)	**0.00**	4.27	(0.84, 7.58)	**0.01**
γ_1	num.med.u.	−0.14	(−0.73, 0.47)	0.67	−0.11	(−0.74, 0.50)	0.64
γ_2	soc.sec.	−2.70	(−5.44, 0.04)	**0.97**	−2.65	(−5.75, 0.64)	**0.96**
γ_3	1st.trim.	−4.56	(−10.42, 2.11)	**0.93**	−3.80	(−11.36, 2.70)	0.79
γ_4	expend.pc	0.13	(−0.11, 0.46)	0.16	0.13	(−0.19, 0.44)	0.19
δ_0	intercept	−5.58	(−22.88, 7.34)	0.77	−6.68	(−22.96, 5.15)	0.83
δ_1	pov.index	−8.12	(−21.19, 3.54)	**0.91**	−8.74	(−21.62, 2.64)	**0.93**
δ_2	births.hosp.	−7.97	(−21.34, 2.19)	**0.93**	−8.38	(−21.15, 1.43)	**0.95**

Posterior summaries of covariate effects for ZIP models with the two competing spatial propositions (SGP and CAR) are presented in Table 5. Reported are the posterior means as point estimates, 95 % posterior credible intervals (CIs), as well as posterior probabilities of below zero to better assess significant coefficients. From the table, we observe that the two spatial models detect four common significant covariates (at 10 % level). These are the intercept and X_2, which explain the mortality ratio λ, and Z_1 and Z_2 which explain the zero inflation probability θ. Although the numerical estimates for the parameters are not exactly the same, the conclusions arisen lead to the same direction. One important difference is that the SGP model detects significant a covariate not detected by the CAR model, this is X_3, the proportion of pregnant women who were seen by a physician in the first trimester of pregnancy. Since the best fitting model is the SGP, we consider this later covariate as significant and proceed to the interpretation.

We conclude that apart from the intercept, the percentage of pregnant women with access to social security and the percentage of pregnant women seen by a physician in the first trimester have both a negative significant effect in the mortality ratio. An increment of 10 percentage points in the number of women with access to social security would reduce the risk in almost 25 %. Moreover, if 10 percentage more of pregnant women were seen by a physician, the relative risk of dying would reduce in 37 %. Something interesting is that the number of medical units in the state and the government expenditure on health showed no significant effect on the risk. Thus, more investment in health is not the solution to the problem, but the penetration of the health services to more people and the actual use of them by pregnant women, specially in the first trimester.

Something that adds up to validate the need for a ZI model is the detection of significant covariates explaining the structural zeroes regime. The poverty index is one of this significant covariates. This variable is a standardized index that ranges (roughly) from −1.5 to 2.5. An increment of 0.1 units in this poverty scale would lead to a 55 % reduction in the odds of observing a structural zero. Similarly, increasing 10 percent the number of births in clinics or hospitals would reduce 55 % the odds of a structural zero. That is, less poor

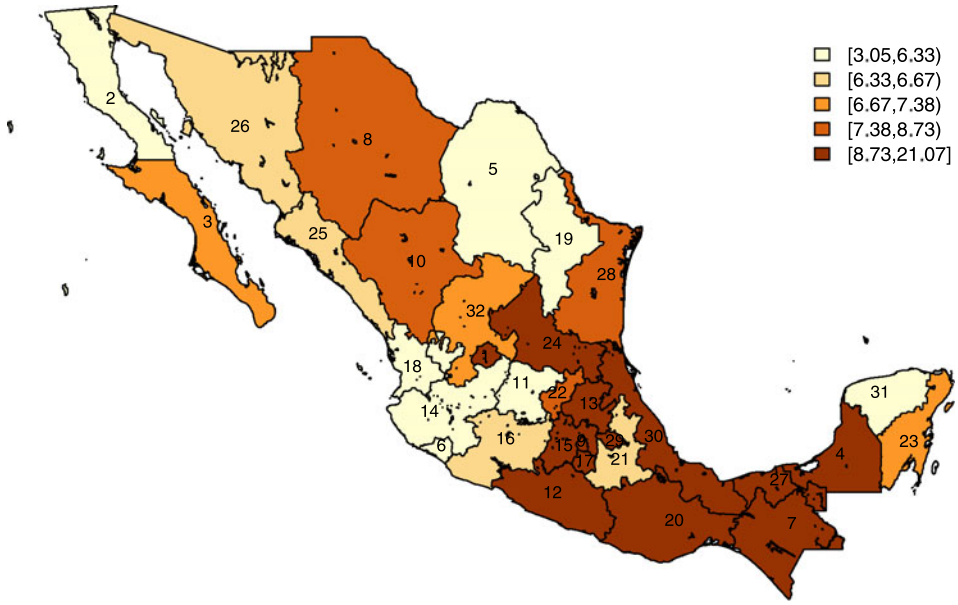

Figure 2. Posterior estimates of maternity mortality rate λ_i, $i = 1, \ldots, 32$, obtained after fitting the ZIP+SGP model. The actual numbers are given in Table 1.

Figure 3. Posterior estimates of zero inflation probability θ_i, $i = 1, \ldots, 32$, obtained after fitting the ZIP+SGP model. The actual numbers are given in Table 1.

states or a lower percentage of births in clinics or hospitals induce a higher probability of observing a structural zero in the number of maternity deaths.

Finally, we present a map of posterior estimates of relative risk or maternity mortality ratio (due to hypertensive disorder) λ_i and probability of experiencing structural zero deaths θ_i for state i, respectively in Figures 2 and 3. The actual estimates pertaining to every Mexican state are listed in Table 1. From Figure 2, we observe that high mortality ratios (>8.7) predominate clustered within the central-southern states, excluding two states from the Yucatan Peninsula (States 23 and 31) and the state of Puebla (21). This high risk region includes the poorest states of Mexico like Chiapas (7), Guerrero (12) and Oaxaca (20), moderately poor states like Aguascalientes (1), San Luis Potosi (24), Hidalgo (13), Tlaxcala (29), Veracruz (30), Tabasco (27) and Campeche (4), as well as some of the most developed and affluent states like Distrito Federal (9), Mexico (15) and Morelos (17). It is not surprising that eight out of these 13 states that belong to this high risk group are in the 33th percentile of the states with the lowest percentage of women with social security. On the other hand, states with the lowest mortality ratio are not all clustered geographically as the high risk group. The lowest risk group is divided into two well defined clusters and two scattered states, one in each of the peninsulas, Baja California (2) and Yucatan (31). One of the well defined clusters of lowest risk lies in the center of the country and includes Colima (6), Jalisco (14), Nayarit (18) and Guanajuato (11), and the second cluster is in the north and includes Coahuila (5) and Nuevo Leon (19). Among these eight states that form the lowest risk group, seven lie above the median of the percentage of women who were seen by a physician in the first trimester of pregnancy. An important conclusion that follows from these risk estimates is that economic development do not trigger high (or low) maternity mortality.

In Figure 3, we observe the geographical distribution of structural zero probabilities. We first note that most of the states show a very low probability (<0.01) of observing a structural zero, that is, for these states the number of deaths could be well explained by a simple P model without ZI. On the other hand, the state of Baja California (2) presents a distinctive high probability (>0.5) of presenting a structural zero in the number of deaths. Additionally, not as high as Baja California, two more states have a probability between 0.04 and 0.06 of a structural zero. They are Coahuila (5) and Nuevo Leon (19), which together with Baja California, have zero deaths observed in the dataset. Four other states show a probability of structural zeroes in $(0.01, 0.04)$. It is not surprising that, apart from Chiapas (7), the rest of the states that belong to the highest risk group (in Figure 2) belong to the group with the lowest probability of structural zeroes (in Figure 3). Conversely, the state with the highest probability of a structural zero (Baja California (2)) belongs to the group of the lowest risk of maternity mortality.

6. CONCLUSIONS

In this paper, we introduced a new Markov gamma random field model for analyzing ZI geo-referenced health count data via a three-level hierarchical structure. The dependence structure is modeled through latent edges and a set of association parameters $\kappa = \{\kappa_{ij}\}$. Our approach is Bayesian, and provides easy implementation of the underlying MCMC sampling schemes.

Our proposal is different from that in Nieto-Barajas (2008) in several aspects. If we denote by SGP_1 the 2008's process and by SGP_2 the current process, the differences are: (1) SGP_1 incorporates the latent variables $\{\nu_i\}$ via the shape parameter of the gamma density, whereas SGP_2 does it through the scale parameter; (2) for SGP_1, $\{\nu_i\}$ is a discrete process (Poisson-gamma), whereas for SGP_2, $\{\nu_i\}$ is a continuous process (gamma); (3) for SGP_1 the ν_i's are independent, whereas for SGP_2 the ν_i's form a dependent (exchangeable) sequence; and 4) SGP_1 is not strictly stationary, it only satisfies second order stationarity properties, and marginally is not gamma distributed, on the contrary, SGP_2 is a strictly stationary process with gamma marginals.

The model for this paper was implemented using Fortran and is available upon request from the first author, however, our model can also be easily implemented in `OpenBugs` (*http://www.openbugs.info/*). Although CAR models dominate the field of disease mapping, simulation studies and application to a real dataset on maternity mortality reveal that our proposition outperforms the CAR models to a certain degree, and is justifiably a close competitor. Our model showed to be less sensitive to high uncertainty in the individual random effects and offered a higher ability to detect ZI regimes when they are (or not) present in the data.

Additional advantages of our approach are: First, we work in the same space as the relative risk (λ) and do not need any transformation (like log) as in CAR models. Second, our spatial model is always proper and stationary with gamma marginals. In contrast, the CAR model is improper for $\rho = 1$ and requires further constraints (such as $\sum_{i=1}^{n} \eta_i^* = 0$) to ensure identifiability of the intercept term in a mixed effects model. Finally, our model is more flexible in capturing different degrees of dependence by considering pair-wise association parameters κ_{ij}, whereas the dependence in the CAR model relies on a single association parameter ρ.

Our current spatial model can be extended in several ways. The most immediate one, for instance, is to a spatio-temporal setting where spatial health information is collected longitudinally. Other extensions are in the context of spatial survival models, or cure rate models. These are currently under investigation.

APPENDIX. OUTLINE OF CONDITIONAL POSTERIORS

The conditional posterior distributions for all parameters in the model are described as follows.

(i) The conditional distribution of the auxiliary variable w_i, for $i = 1, \ldots, n$, has the form

$$f(w_i \mid \mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\theta}) = \text{Ber}\big(w_i \mid \theta_i^*\big), \quad \text{where}$$
$$\theta_i^* = \frac{\theta_i}{\theta_i + (1-\theta_i)\text{Po}(y_i \mid \lambda_i E_i)/I(y_i = 0)}.$$

For updating the components of λ_i we obtained the following full conditionals:

(ii) The conditional posterior distribution for η_i, $i = 1, \ldots, n$, is

$$f(\eta_i \,|\, \text{rest}) = \text{Ga}\left(\eta_i \;\middle|\; \alpha + \sum_{j \in \partial_i} \kappa_{ij} + y_i(1-w_i), \alpha + \sum_{j \in \partial_i} \upsilon_{ij} + \xi_i E_i (1-w_i) e^{\boldsymbol{\gamma}' \mathbf{x}_i}\right).$$

(iii) The conditional posterior distribution for υ_{ij}, for $\upsilon_{ij} > 0$ and $i \sim j \in \{1, \ldots, n\}$, is

$$f(\upsilon_{ij} \,|\, \text{rest}) \propto \left(\alpha + \sum_{k \in \partial_i} \upsilon_{ik}\right)^{\alpha + \sum_{k \in \partial_i} c_{ik}} \left(\alpha + \sum_{k \in \partial_j} \upsilon_{jk}\right)^{\alpha + \sum_{k \in \partial_j} c_{jk}} \times \upsilon_{ij}^{\kappa_{ij} - 1} e^{-\upsilon_{ij}(\eta_i + \eta_j + \phi)}.$$

(iv) The conditional posterior distribution for ϕ is

$$f(\phi \,|\, \text{rest}) = \text{Ga}\left(\phi \;\middle|\; \alpha + \sum_{i<j}^{n} \kappa_{ij}, \alpha + \sum_{i<j}^{n} \upsilon_{ij}\right).$$

(v) The conditional posterior distribution for α is

$$f(\alpha \,|\, \text{rest}) \propto \frac{[\{\prod_{i=1}^{n} (\alpha + \sum_{j \in \partial_i} \upsilon_{ij})\}]^{\alpha + \sum_{j \in \partial_i} \kappa_{ij}} \{(\prod_{i=1}^{n} \eta_i e^{-\eta_i}) \phi e^{-\phi}\}^{\alpha}}{\Gamma(\alpha) \prod_{i=1}^{n} \Gamma(\alpha + \sum_{j \in \partial_i} \kappa_{ij})} \times \text{Ga}(\alpha \mid a_\alpha, b_\alpha).$$

(vi) The conditional posterior distribution for κ_{ij} is

$$f(\kappa_{ij} \,|\, \text{rest}) \propto \frac{\{(\alpha + \sum_{k \in \partial_i} \upsilon_{ik})(\alpha + \sum_{k \in \partial_j} \upsilon_{jk}) \upsilon_{ij} \eta_i \eta_j \phi\}^{\kappa_{ij}}}{\Gamma(\alpha + \sum_{k \in \partial_i} \kappa_{ik}) \Gamma(\alpha + \sum_{k \in \partial_j} \kappa_{jk}) \Gamma(\kappa_{ij})} \text{Ga}(\kappa_{ij} \mid 1, \omega).$$

(vii) The conditional posterior distribution for ω is

$$f(\omega \mid \text{rest}) = \text{Ga}\left(\omega \;\middle|\; 1 + \frac{a_\omega}{2} \sum_{i=1}^{n} m_i, b_\omega + \sum_{i<j} \kappa_{ij}\right),$$

with $m_i = \sum_{j=1}^{n} w_{ij}$.

(viii) The conditional posterior distribution for ξ_i, $i = 1, \ldots, n$, is

$$f(\xi_i \,|\, \text{rest}) = \text{Ga}\big(\xi_i \;\big|\; \beta_\xi + y_i(1-w_i), \beta_\xi + \eta_i E_i (1-w_i) e^{\boldsymbol{\gamma}' \mathbf{x}_i}\big).$$

(ix) The conditional posterior distribution for β_ξ is

$$f(\beta_\xi \,|\, \text{rest}) \propto \frac{(\beta_\xi)^{n\beta_\xi}}{\Gamma^n(\beta_\xi)} \left(\prod_{i=1}^{n} \xi_i\right)^{\beta_\xi} e^{-\beta_\xi \sum_{i=1}^{n} \xi_i} \text{Ga}(\beta_\xi \mid a_\beta, b_\beta).$$

(x) The conditional posterior distribution for γ_s, $s = 1, \ldots, p$, is

$$f(\gamma_s \,|\, \text{rest}) \propto \exp\left\{\gamma_s \sum_{i=1}^{n} y_i (1-w_i) x_{is} - \sum_{i=1}^{n} \eta_i \xi_i E_i (1-w_i) e^{\boldsymbol{\gamma}' \mathbf{x}_i}\right\} \text{N}\big(\gamma_s \mid 0, \sigma_\gamma^2\big).$$

For the updating of the components of θ_i we require a further arrangement of the likelihood. We note that the joint distribution (5) can be written as $f(\mathbf{y}, \mathbf{w}|\boldsymbol{\lambda}, \boldsymbol{\theta}) = f(\mathbf{y}|\boldsymbol{\lambda}, \mathbf{w}) f(\mathbf{w}|\boldsymbol{\theta})$. So inference on $\boldsymbol{\theta}$ relies entirely on the joint distribution of the auxiliary variables $\mathbf{w}$. This latter distribution, using the definition for θ_i given in (4), has the form

$$f(\mathbf{w}|\boldsymbol{\theta}) = \prod_{i=1}^{n} \theta_i^{w_i} (1-\theta_i)^{1-w_i} = \prod_{i=1}^{n} \frac{(\zeta_i)^{w_i} e^{(\boldsymbol{\delta}' \mathbf{z}_i) w_i}}{1 + \zeta_i e^{\boldsymbol{\delta}' \mathbf{z}_i}}.$$

Again, to simplify posterior updating, we introduce another set of independent auxiliary variables, $v_i > 0$, such that the new extended likelihood becomes

$$f(\mathbf{w}, \mathbf{v}|\boldsymbol{\delta}, \boldsymbol{\zeta}) = \prod_{i=1}^{n} (\zeta_i)^{w_i} \exp\{(\boldsymbol{\delta}' z_i) w_i - v_i (1 + \zeta_i e^{\boldsymbol{\delta}' z_i})\}.$$

By integrating out v_i we get back $f(\mathbf{w}|\boldsymbol{\theta})$ as desired. Therefore,

(x) The conditional distribution of the auxiliary variable v_i, $i = 1, \ldots, n$, is

$$f(v_i \,|\, \text{rest}) = \text{Ga}(v_i \mid 1, 1 + \zeta_i e^{\boldsymbol{\delta}' \mathbf{z}_i}).$$

(xi) The conditional posterior distribution for ζ_i, $i = 1, \ldots, n$, is

$$f(\zeta_i \,|\, \text{rest}) = \text{Ga}(\zeta_i \mid \beta_\zeta + w_i, \beta_\zeta + v_i e^{\boldsymbol{\delta}' \mathbf{z}_i}).$$

(xii) The conditional posterior distribution for δ_r, $r = 1, \ldots, q$, is

$$f(\delta_r \,|\, \text{rest}) \propto \exp\left\{ \delta_r \sum_{i=1}^{n} w_i z_{ir} - \sum_{i=1}^{n} v_i \zeta_i e^{\boldsymbol{\delta}' \mathbf{z}_i} \right\} \text{N}(\delta_r \mid 0, \sigma_\delta^2).$$

The posterior conditional distributions for β_θ is given in (viii) with the appropriate sub-index ζ placed instead of ξ. If a ZI Negative Binomial model is desired, we replace E_i by T_i and sample this later one from its full conditional distribution.

(xiii) The conditional posterior distribution for T_i is

$$f(t_i \,|\, \text{rest}) = \text{Ga}(w_i \mid E_i + y_i(1 - w_i), 1 + \lambda_i).$$

ACKNOWLEDGEMENTS

We thank the associate editor and two anonymous referees whose constructive comments led to a much improved presentation of the manuscript. Nieto-Barajas acknowledges support from Grant I130991 from the National Council for Science and Technology of Mexico (CONACYT). Bandyopadhyay acknowledges funding support from Grants P20RR017696-06, R03-DE020114 and R03-DE021762 from the US National Institutes of Health.

[Received February 2012. Accepted January 2013.]

REFERENCES

Annual Statistics Book. (2009), *National System on Health Information*. Mexico. Available at *http://www.sinais.salud.gob.mx/publicaciones/index.html*.

Banerjee, S., Wall, M. M., and Carlin, B. P. (2003), "Frailty Modeling for Spatially Correlated Survival Data, With Application to Infant Mortality in Minnesota," *Biostatistics*, 4, 123–142.

Bernardo, J. M., and Smith, A. F. M. (2000), *Bayesian Theory*, New York: Wiley.

Best, N., Richardson, S., and Thomson, A. (2005), "A Comparison of Bayesian Spatial Models for Disease Mapping," *Statistical Methods in Medical Research*, 14, 35–59.

Böhning, D., Dietz, E., Schlattmann, P., Mendonca, L., and Kirchner, U. (1999), "The Zero-Inflated Poisson Model and the Decayed, Missing and Filled Teeth in Dental Epidemiology," *Journal of the Royal Statistical Society, Series A*, 162 (2), 195–209.

Cheung, Y. B. (2002), "Zero-Inflated Models for Regression Analysis of Count Data: A Study of Growth and Development," *Statistics in Medicine*, 21, 1461–1469.

Denison, D. G. T., and Holmes, C. C. (2001), "Bayesian Partitioning for Estimating Disease Risk," *Biometrics*, 57, 143–149.

Eberley, L., and Carlin, B. P. (2000), "Identifiability and Convergence Issues for Markov Chain Monte Carlo Fitting of Spatial Models," *Statistics in Medicine*, 19, 2279–2294.

Gelfand, A. E., Dey, D. K., and Chang, H. (1992), "Model Determinating Using Predictive Distributions With Implementation via Sampling-Based Measures" (with discussion), in *Bayesian Statistics 4*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, pp. 147–167.

Green, P., and Richardson, S. (2002), "Hidden Markov Models and Disease Mapping," *Journal of the American Statistical Association*, 97, 1055–1070.

Griffith, D. A. (1983), "The Boundary Value Problem in Spatial Statistical Analysis," *Journal of Regional Science*, 23, 377–387.

Ibrahim, J. G., and Laud, P. W. (1994), "A Predictive Approach to the Analysis of Designed Experiments," *Journal of the American Statistical Association*, 89, 309–319.

Ickstadt, K., and Wolpert, R. L. (1999), "Spatial Regression for Marked Point Processes," in *Bayesian Statistics 6*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, pp. 323–341.

Jung, B. C., Jhun, M., and Lee, J. W. (2005), "Bootstrap Tests for Overdispersion in a Zero-Inflated Poisson Regression Model," *Biometrics*, 61, 626–629.

Knorr-Held, L., and Rasser, G. (2000), "Bayesian Detection of Clusters and Discontinuities in Disease Maps," *Biometrics*, 56 (1), 13–21.

Lam, K. F., Xue, H., and Cheung, Y. B. (2006), "Semiparametric Analysis of Zero-Inflated Count Data," *Biometrics*, 62, 996–1003.

Lambert, D. (1992), "Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing," *Technometrics*, 34, 1–14.

Lawson, A. B. (2009), *Bayesian Disease Mapping: Hierarchical Models in Spatial Epidemiology*, Boca Raton, FL: Chapman and Hall/CRC.

Marshall, R. J. (1991), "A Review of Methods for the Statistical Analysis of Spatial Patterns of Disease," *Journal of the Royal Statistical Society, Series A*, 154, 421–441.

McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models*, Boca Raton: Chapman and Hall.

Nieto-Barajas, L. E. (2008), "A Markov Gamma Random Field for Modeling Disease Mapping Data," *Statistical Modelling*, 8, 97–114.

Nieto-Barajas, L. E., and Walker, S. G. (2002), "Markov Beta and Gamma Processes for Modelling Hazard Rates," *Scandinavian Journal of Statistics*, 29, 413–424.

Smith, A. F. M., and Roberts, G. O. (1993), "Bayesian Computations via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society, Series B*, 55, 3–23.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. (2002), "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society, Series B*, 64, 583–639.

Street, A., Jones, A., and Furuta, A. (1999), "Cost Sharing and Pharmaceutical Utilisation and Expenditure in Russia," *Journal of Health Economics*, 18, 459–472.

Tierney, L. (1994), "Markov Chains for Exploring Posterior Distributions," *Annals of Statistics*, 22, 1701–1722.

Wakefield, J. (2007), "Disease Mapping and Spatial Regression With Count Data," *Biostatistics*, 8 (2), 158–183.

Wolpert, R. L., and Ickstadt, K. (1998), "Poisson/Gamma Random Field Models for Spatial Statistics," *Biometrika*, 85, 251–267.

Women and Men in Mexico. (2009), National Institute of Statistics and Geography, Mexico. Available at *http://www.inegi.org.mx/*.

Yau, K. K. W., Wang, K., and Lee, A. H. (2003), "Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data With Excess Zeros," *Biometrical Journal*, 45, 437–452.