Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models

By BRIAN J. REICH[a], JAMES S. HODGES[a], and VESNA ZADNIK[b] [1]

[a] *Division of Biostatistics, School of Public Health, University of Minnesota, 2221 University Ave SE, Suite 200, Minneapolis, Minnesota 55414, U.S.A.*

[b] *Epidemiology and Cancer Registries, Institute of Oncology Ljubljana, Zaloka 2, 1000 Ljubljana, Slovenia*

Correspondence Author: James S. Hodges
E-mail: hodges@ccbr.umn.edu
Telephone: (612) 626-9626
Fax: (612) 626-9054

May 12, 2005

[1] The authors thank Prof. Bradley Carlin of the University of Minnesota for his thoughtful discussion.

Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models

Abstract

Disease-mapping models for areal data often have fixed effects to measure the effect of spatially-varying covariates and random effects with a conditionally autoregressive (CAR) prior to account for spatial clustering. In such spatial regressions, the objective may be to estimate the fixed effects while accounting for the spatial correlation. But adding the CAR random effects can cause large changes in the posterior mean and variances of fixed effects compared to the non-spatial regression model. This paper explores the impact of adding spatial random effects on fixed-effect estimates and posterior variance. Diagnostics are proposed to measure posterior variance inflation from collinearity between the fixed effect covariates and the CAR random effects and to measure each region's influence on the change in the fixed effect's estimates from adding the CAR random effects. A new model that alleviates the collinearity between the fixed-effect covariates and the CAR random effects is developed and extensions of these methods to point-referenced data models are discussed.

Key Words: Conditional autoregressive prior; collinearity; diagnostics; disease mapping; spatial regression models.

1 Introduction

Spatially-referenced public health data sets have become more available in recent years. A common objective when analyzing these data sets is to estimate the effect of covariates on region-specific disease rates while accounting for spatial correlation. As a motivating example, consider analyzing of stomach cancer incidence rates in Slovenia's municipalities 1995-2001. The study's objective is to investigate the relationship between stomach cancer incidence and socioeconomic status.

Figure 1a plots the n=192 municipalities' observed standardized incidence ratio (SIR), i.e., the ratio of observed (O_i) to expected (E_i, computed using indirect standardization) number of cases. Each region's socioeconomic status has been placed into one of five ordered categories by the Institute of Macroeconomic Analysis and Development based on a composite score calculated from 1999 data. The centered version of this covariate (SEc) is plotted in Figure 1b. Both SIR and SEc exhibit strong spatial patterns. Western municipalities generally have low SIR and high SEc while Eastern municipalities generally have high SIR and low SEc. This suggests a negative relationship between stomach cancer rates and socioeconomic status. A simple Poisson regression ignoring the spatial structure, with log link and flat priors on the regression parameters (Model 1, Table 1), shows SEc is indeed related to SIR: the posterior median of SEc is -0.137 and its 95% posterior interval is (-0.17,-0.10).

To model these data correctly, the spatial structure must be incorporated into the analysis. A popular Bayesian disease-mapping model is the conditionally autoregressive (CAR) model of Besag et. al (1991). Under this model, the numbers of cases in each region follow conditionally independent Poisson distributions with $log(E(O_i)) = log(E_i) + \beta_{SEc} * SEc_i + S_i + H_i$. This model has one fixed effect, β_{SEc}, which is given a flat prior, and two types of random effects: S captures spatial clustering and H captures region-wide heterogeneity. The H_i are modelled as independent draws from a normal distribution with mean zero and precision τ_h. Spatial dependence is introduced through the prior (or model) on $S = (S_1, ..., S_n)'$. The CAR model with L_2 norm (also called a

Gaussian Markov random field) for S has improper density

$$p(S|\tau_s) \propto \tau_s^{(n-G)/2} \exp\left(-\frac{\tau_s}{2} S'QS\right), \tag{1}$$

where τ_s controls the smoothing induced by this prior, larger values smoothing more than smaller; G is the number of "islands" (disconnected groups of regions) in the spatial structure (Hodges et al., 2003); and Q is $n \times n$ with non-diagonal entries $q_{ij} = -1$ if regions i and j are neighbors and 0 otherwise, and diagonal entries q_{ii} equal to the number of region i's neighbors. This is a multivariate normal kernel, specified by its precision matrix $\tau_s Q$ instead of the usual covariance matrix. In this paper, we give the precision parameters τ_e and τ_h independent Gamma(0.01,0.01) priors (parameterized to have mean 1 and variance 100).

Insert Figure 1 about here

Insert Table 1 about here

Table 1 shows that despite the increase in model complexity (p_D increases from 2.0 to 62.3), the model including the heterogeneity and spatial random effects (Model 2) has smaller DIC (Spiegelhalter et. al, 2002) than the simple Poisson regression (1081.5 to 1153.0). Adding the random effects to the simple Poisson model has a dramatic effect on the posterior of β; its posterior mean changes from -0.137 to -0.022 and its posterior variance increases from 0.0004 to 0.0016.

This paper investigates the change in the posterior of the fixed effects from adding spatial random effects. Section 2 assumes the outcomes are normally distributed and characterizes the effect of adding CAR random effects on the posterior mean and variance of the fixed effects parameters. We propose diagnostics and identify combinations of covariates and spatial grids that are especially troublesome. Section 3 develops an approach to spatial regression that sidesteps collinearity problems by restricting the domain of the spatial random effects to the space orthogonal to the fixed-effect covariates. The methods developed in Sections 2 and 3 are extended to the generalized linear spatial regression model in Section 4. Section 5 illustrates these methods using the Slovenia stomach cancer data. Section 6 summarizes these results and briefly describes of how the methods could be extended to other spatial regression models, such as geostatistical models for point-referenced data.

2 Collinearity in the CAR model with normal observables

Let y be the n-vector of observed values and X be a known $n \times p$ matrix of covariates standardized so that each column sums to zero and has unit variance. The CAR spatial regression model is

$$\mathrm{y}|\boldsymbol{\beta}, S, \tau_e \sim N(X\boldsymbol{\beta} + S, \tau_e I_n) \tag{2}$$

$$S|\tau_s \sim N(0, \tau_s Q), \tag{3}$$

where $\boldsymbol{\beta}$ is a p-vector of fixed-effect regression parameters, S is an n-vector, $\tau_e I_n$ and $\tau_s Q$ are precision matrices, and Q is the known adjacency matrix described in Section 1. Since the rows and columns of Q sum to zero, the CAR model necessarily implies a flat prior on each island's average of S. A common solution to this impropriety is to add fixed effects for island's intercept and place a sum-to-zero constraint on each island's S. However, since collinearity between the intercept and the spatial random effects is not of interest, we let S remain unconstrained and assume X does not not have a column for the intercept, so that the intercept is implicitly present in S. To complete this Bayesian specification, $\boldsymbol{\beta}$ is given a flat prior and τ_e and τ_s are given independent Gamma(0.01,0.01) priors.

An equivalent representation of this model that highlights identification and collinearity concerns is

$$\mathbf{y}|\boldsymbol{\beta}, \mathrm{b}, \tau_e \sim N(X\boldsymbol{\beta} + Z\mathrm{b}, \tau_e I_n) \tag{4}$$

$$\mathrm{b}|\tau_s \sim N(0, \tau_s D) \tag{5}$$

where $\mathrm{b} = Z'S$ and $Q = ZDZ'$ is Q's spectral decomposition for $n \times n$ orthogonal matrix Z and $n \times n$ diagonal matrix D with positive diagonal elements $d_1 \geq ... \geq d_{n-G}$. The last G diagonal elements of D are zero. The corresponding elements of b represent the intercepts of the G islands and are implicit fixed effects. The mean of $\mathbf{y}$ (4) has $p+G$ fixed effects $(\boldsymbol{\beta}, b_{n-G+1}, \ldots, b_n)$ and $n-G$ random effects $(b_1, \ldots, b_{n-G})$ for a total of $n+p$ predictors; given that there are only n observations, this raises identifiability and collinearity concerns. Each column of X is a linear combination of the n orthogonal columns of Z. Therefore, ignoring b's prior, i.e., setting $\tau_s = 0$, the data cannot identify both $\boldsymbol{\beta}$ and b. Identification of $\boldsymbol{\beta}$ relies on smoothing of b, which is controlled by τ_s. As τ_s increases, b is smoothed closer to zero and the posterior of $\boldsymbol{\beta}$ becomes more similar to its posterior under the ordinary linear model (OLM).

To measure the influence on the posterior mean and variance of $\boldsymbol{\beta}$ from including and smoothing the spatial random effects, we investigate the posterior of $\boldsymbol{\beta}$ conditional on (τ_e, τ_s). After integrating out b, $\boldsymbol{\beta}$ is normal with

$$\begin{aligned} E\left(\boldsymbol{\beta}|\tau_e, \tau_s, \mathrm{y}\right) &= E\left(E\left(\boldsymbol{\beta}|\mathrm{b}, \tau_s, \tau_e, \mathrm{y}\right)\right) = E\left((X'X)^{-1}X'(\mathrm{y} - Z\mathrm{b})|\tau_s, \tau_e, \mathrm{y}\right) \\ &= (X'X)^{-1}X'(\mathrm{y} - Z\hat{\mathrm{b}}) = \hat{\boldsymbol{\beta}}_{OLM} - (X'X)^{-1}X'Z\hat{\mathrm{b}} \qquad (6) \\ \mathrm{Var}(\boldsymbol{\beta}|\tau_e, \tau_s, \mathrm{y}) &= \left(\tau_e X'Z\tilde{D}Z'X\right)^{-1} \qquad (7) \end{aligned}$$

where $\hat{\boldsymbol{\beta}}_{OLM} = (X'X)^{-1}X'\mathrm{y}$, $P^c = I - X(X'X)^{-1}X'$, $r = \tau_s/\tau_e$, $\tilde{D} = I - (I + rD)^{-1}$, and $\hat{\mathrm{b}} = E(\mathrm{b}|\tau_e, \tau_s, \mathrm{y}) = (Z'P^cZ + rD)^{-1} Z'P^c\mathrm{y}$ is b's posterior mean given (τ_s, τ_e) but *not* $\boldsymbol{\beta}$. The posterior mean and variance of $\boldsymbol{\beta}$ under the OLM with $\mathrm{b} \equiv 0$ are $\hat{\boldsymbol{\beta}}_{OLM}$ and $(\tau_e X'X)^{-1}$ respectively. By adding the random effects, $\boldsymbol{\beta}$'s posterior mean is shifted by $(X'X)^{-1}X'Z\hat{\mathrm{b}}$, (6), and $\boldsymbol{\beta}$'s posterior variance is increased, (7). These two effects are discussed separately in Sections 2.1 and 2.2.

2.1 Influence of spatial random effects on $\mathbf{E}(\boldsymbol{\beta}|\tau_e, \tau_s, \mathbf{y})$

This section provides diagnostics measuring each region's contribution to the difference between $E(\boldsymbol{\beta}|\tau_e, \tau_s, \mathbf{y})$ and $\hat{\boldsymbol{\beta}}_{OLM}$. Case influence on the estimates in the OLM has a long history, with two common diagnostics being leverage and the DFBETAS statistics (Hocking, 1996). The leverage of the j^{th} observation is the j^{th} diagonal element of $X(X'X)^{-1}X'$ and measures the potential for the j^{th} observation to be overly influential in fixed-effect estimation. The leverages are properties of the design and do not depend on the data; observations with outlying X_i typically have large leverage. By contrast, the DFBETAS statistics are functions of the observed data, y. They measure the influence of the j^{th} observation on the estimate of $\boldsymbol{\beta}_i$ by comparing the estimate of $\boldsymbol{\beta}_i$ with and without the j^{th} observation included in the analysis. This section develops measures of potential influence and observed influence analogous to leverage and DFBETA diagnostics.

The posterior mean of $\boldsymbol{\beta}$ given (τ_e, τ_s) under the spatial model is $(X'X)^{-1}X'(y - X\hat{\mathrm{b}})$ (Equation 6), the least squares estimate using the residuals $\mathrm{y} - Z\hat{\mathrm{b}}$ as the observables. Define the change in $\boldsymbol{\beta}$'s posterior mean given (τ_e, τ_s) due to adding the CAR random effects as $\Delta = \hat{\boldsymbol{\beta}}_{OLM} - E(\boldsymbol{\beta}|\tau_e, \tau_s, \mathbf{y}) = (X'X)^{-1}X'Z\hat{\mathrm{b}}$. Δ can be positive or negative, that is, adding the spatial random effects does not necessarily push the posterior mean of $\boldsymbol{\beta}$ towards zero as in the example in Section 1.

Δ can be written $C\mathrm{e}$ for the $p \times n$ matrix

$$C = \left[(X'X)^{-1}X'\right]\left[Z\left(Z'P^cZ + rD\right)^{-1}Z'\right]P^c = \left[(X'X)^{-1}X'\right]\left[(P^c + rQ)^{-1}\right]P^c, \tag{8}$$

where $\mathrm{e} = P^c\mathrm{y}$ is the n-vector of OLM residuals. Similar to leverage, the elements of C depend on the design (i.e., the spatial structure and the covariates) but do not depend on the data except through $r = \tau_s/\tau_e$, on which we have conditioned. Plotting the absolute values of the i^{th} row of C, which we call $|C_i|$, shows which regions have the potential to make a large contribution to the change in β_i when spatial effects are added to the OLM.

Figure 2 plots $|C|$ for the Slovenia spatial grid with $X = SEc$, with little spatial smoothing ($r = 0.1$, Figure 2a) and with substantial spatial smoothing ($r = 10$, Figure 2b). $|C|$ is generally smaller in Figure 2b than in Figure 2a because for large r the spatial random effects are smoothed to zero and $\boldsymbol{\beta}$'s posterior is similar to its posterior under the OLM. A careful comparison of Figure 1b and Figure 2 shows that regions with extreme SEc (i.e., high leverage in OLM) do not necessarily have large $|C_j|$; the regions with large $|C_j|$ are those with extreme SEc relative to the SEc of nearby regions. For example, most regions with large SEc are in Western Slovenia (Figure 1b). However, only western municipalities with moderate SEc have large $|C_j|$ because moderate SEc is unusual relative to other western regions.

Insert Figure 2 about here

Only regions with non-zero C_{ij} *and* non-zero e_j contribute to $\Delta_i = \sum_{j=1}^{n} \delta_{ij} = \sum_{j=1}^{n} C_{ij} e_j$. In the OLM, $\hat{\boldsymbol{\beta}}_i$ is computed assuming all associations between nearby regions can be explained by the fixed effects. C_{ij} measures the j^{th} region's potential to update this estimate when spatial clustering is introduced as a second explanation of association between nearby regions. Sites with covariates that are similar to neighboring sites (small $|C_{ij}|$) cannot distinguish between these competing explanations of association because both spatial clustering and the fixed effects predict the region will be similar to its neighbors. Regions with extreme covariates relative to nearby regions (large $|C_{ij}|$) can provide information to update $\hat{\boldsymbol{\beta}}$ because the two explanations of association predict different outcomes. The OLM predicts that y_j will be different than its neighbors. If y_j really is different from its neighbors (small $|e_j|$) this supports the assumption that all associations are caused by the fixed effects and an update to $\hat{\boldsymbol{\beta}}$ is not warranted. In contrast, if y_i is similar to the outcomes of nearby regions (large $|e_j|$), this is evidence that spatial clustering, not the fixed effects, is responsible for the similarity of nearby regions. This suggests a large change in $\boldsymbol{\beta}$, i.e., $|\Delta_{ij}|$ will be large. That is, if the simple linear regression fits poorly at regions where X_{ij} is different from its neighboring X_i, there will be a large adjustment to the posterior mean of $\boldsymbol{\beta}_i$ when the spatial structure is incorporated into the model.

The number of regions that qualify as "nearby" is determined by r. Consider the western municipality on the northern border that changes from turquoise in Figure 2a to dark blue in Figure 2b. For small r (Figure 2a), there is little spatial smoothing and this region's SEc is primarily compared to its neighboring regions. Its SEc is smaller than the SEc of the three adjacent municipalities (Figure 1b) so $|C_j|$ is moderately large. This municipality's $|C_j|$ is one of the largest in Figure 2b because for large r (strong spatial smoothing) the region's SEc is compared not only to adjacent regions, but all northwestern regions. Since this region has the smallest SEc in the entire Northwest and few other regions are extreme relative to such a large of a group of nearby regions, its $|C_j|$ is among the largest.

2.2 Effect of collinearity on Var($\boldsymbol{\beta}|\tau_e, \tau_s, \mathbf{y}$)

The variance inflation factor is commonly used in OLM theory to measure collinearity among covariates (Hocking, 1996). This diagnostic is defined as the actual variance of the i^{th} regression parameter divided by its variance assuming its covariate is uncorrelated with the other covariates. Although collinearity among fixed effects is an important issue, in this section we will inspect the increase in $\boldsymbol{\beta}_i$'s posterior variance from adding CAR random effects to the OLM.

Since the columns of Z are an orthogonal basis for $\Re^n$, we can write $X = \sqrt{N-1}Z\boldsymbol{\rho}$ for an $n \times p$ matrix $\boldsymbol{\rho}$ where ρ_{ij} is the correlation between the i^{th} column of Z and the j^{th} column of X and $\sum_{j=1}^{N} \rho_{ij}^2 = \text{Var}(X_i) = 1$. The posterior variance of $\boldsymbol{\beta}$ in (7) can be written $\text{Var}(\boldsymbol{\beta}|\tau_e, \tau_s, \mathbf{y}) = \left((N-1)\tau_e\boldsymbol{\rho}'\tilde{D}\boldsymbol{\rho}\right)^{-1}$ where $\tilde{D} = I - (I + rD)^{-1} = \text{diag}\left(rd_i/(1 + rd_i)\right)$. Under the OLM with $\tau_s = \infty$, $\text{Var}(\boldsymbol{\beta}|\tau_e, \mathbf{y}) = (\tau_e X'X)^{-1} = ((N-1)\tau_e\boldsymbol{\rho}'\boldsymbol{\rho})^{-1}$, so

$$VIF_i(r) = \frac{\left[\left(X'Z\tilde{D}Z'X\right)^{-1}\right]_{ii}}{\left[(X'X)^{-1}\right]_{ii}} = \frac{\left[\left(\boldsymbol{\rho}'\tilde{D}\boldsymbol{\rho}\right)^{-1}\right]_{ii}}{\left[(\boldsymbol{\rho}'\boldsymbol{\rho})^{-1}\right]_{ii}} \tag{9}$$

is the inflation in $\boldsymbol{\beta}_i$'s variance from adding the CAR random effects.

If X is a vector, (9) reduces to

$$\text{VIF}(r) = \left(1 - \sum_{j=1}^{n} \frac{\rho_j^2}{1 + rd_j}\right)^{-1} \geq 1 \text{ for all } r > 0. \tag{10}$$

For any spatial grid and for any covariate, $VIF(r) \to \infty$ as $r \to 0^+$, that is, $\boldsymbol{\beta}$ has infinite variance if the random effects b are not smoothed. As r increases, b is smoothed to zero and $VIF(r) \to 1$. The rate at which $VIF(r)$ descends to one depends on the ρ_j and the corresponding d_j. $VIF(r)$ approaches one less quickly if the directions with large ρ_j are associated with small d_j.

If $\rho_j \approx 1$, X is very similar to the j^{th} column of Z, so $\boldsymbol{\beta}$ and b_j must to compete to explain the same one-dimensional projection of y (4). Both parameters are identified if $r > 0$ because b_j's prior shrinks b_j towards its prior mean, zero (5). If b_j's prior is strong, $\boldsymbol{\beta}$ will be well-identified; if b_j's prior is weak, $\boldsymbol{\beta}$ will be weakly-identified. The strength of b_j's prior relative to the prior on the other elements of b is controlled by d_j. The inflation of $\boldsymbol{\beta}$'s variance due to collinearity with Z_j persists for larger r if X is highly correlated with Z_j corresponding to small d_j because small d_i means b_j's prior is less restrictive.

Reich and Hodges (2005) explored the properties of eigenvectors associated with large and small d_i. Eigenvectors associated with small d_i vary gradually in space ("low-frequency") while eigenvectors associated with large d_i vary abruptly ("high-frequency"). Figure 3 plots the eigenvectors associated with the smallest and largest eigenvalues of the adjacency matrix for Slovenia's municipalities. These eigenvectors measure the Southwest/Northeast gradient (Figure 3a) and the difference between the municipalities with the most neighbors and their respective neighbors (Figure 3b).

Insert Figure 3 about here

In summary, conditional on τ_e, $\boldsymbol{\beta}$'s posterior variance is always larger under the spatial model than the OLM for all $r > 0$. The variance inflation arising from collinearity between fixed effects and random effects is most troublesome when r is small and/or the covariates are highly correlated with low-frequency eigenvectors of the adjacency matrix. Of course, when τ_e is unknown its estimate will generally be different under spatial model than the OLM, which could also affect the marginal (over τ_e) variance inflation of $\boldsymbol{\beta}$.

3 Spatial smoothing orthogonal to the fixed effects

Several remedial measures have been proposed for collinearity in the OLM including variable deletion, principal component regression, and ridge regression (Hocking, 1996). If a pair of covariates are highly correlated, the natural reaction is variable deletion, i.e., to remove one of the correlated covariates from the model. The estimates of the remaining coefficients will be more precise but potentially biased. In situations where both of the correlated covariates are of scientific interest, it may be difficult to decide which covariate to remove and removing a scientifically relevant covariate may result in a model that is difficult to interpret.

Collinearity between the fixed effects and CAR random effects in Section 2's spatial regression model leads to many of the same problems as collinearity in the OLM, such as variance inflation and computational problems. However, the natural ordering of parameter importance may distinguish spatial regression from the OLM. The primary objective in a spatial regression may be to estimate the fixed effects; in such a situation, the CAR random effects are added to account for spatial correlation in the residuals when computing the posterior variance of the fixed effects and to improve predictions. With these objectives in mind, removing the combinations of CAR random effects that are collinear with the fixed effects may be a reasonable way to proceed.

The orthogonal projection matrices $P = X(X'X)^{-1}X'$ and $P^c = I - P$ have ranks p and $n-p$ respectively. Let K be the $p \times n$ matrix whose rows are the p eigenvectors of P that correspond to non-zero eigenvalues. Similarly, let L be the $(n-p) \times n$ matrix whose rows are the $n-p$ eigenvectors of P^c that correspond to non-zero eigenvalues. $\boldsymbol{\theta}_1 = KZb$ are the combinations of the CAR random effects in the span of X, while $\boldsymbol{\theta}_2 = LZb$ are the combinations of the CAR random effects that are orthogonal to X. Transforming from b to $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$ gives the model

$$\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \tau_e \sim N(X\boldsymbol{\beta} + K'\boldsymbol{\theta}_1 + L'\boldsymbol{\theta}_2, \tau_e I_n) \tag{11}$$

$$\boldsymbol{\theta}|\tau_s \sim N(0, \tau_s \tilde{Q}) \tag{12}$$

where $\tilde{Q} = (K' \; L')'Q(K' \; L')$ and $\tau_e I$ and $\tau_s \tilde{Q}$ are precision matrices.

Identifying both $\boldsymbol{\beta}$ and $\boldsymbol{\theta}_1$ will be entirely dependent on prior information because the spans of X and K' are the same. Since $\boldsymbol{\beta}$ is given a flat prior, it is free to explain all the variation in the data in the span of X or K', so it is not clear the data identify $\boldsymbol{\theta}_1$ at all. After integrating $\boldsymbol{\beta}$ out of ((11), (12)), the posterior of $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ given (τ_e, τ_s) can be written

$$p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\tau_e, \tau_s, \mathrm{y}) \propto \tau_e^{(n-p)/2} exp\left(-\frac{\tau_e}{2}(L\mathrm{y} - \boldsymbol{\theta}_2)'(L\mathrm{y} - \boldsymbol{\theta}_2)\right) \tag{13}$$

$$\times \; \tau_s^{(n-G)/2} exp\left(-\frac{\tau_s}{2}\begin{pmatrix}\boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2\end{pmatrix}'\begin{pmatrix}\tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}' & \tilde{Q}_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2\end{pmatrix}\right). \tag{14}$$

where $\tilde{Q}_{11} = KQK'$, $\tilde{Q}_{12} = KQL'$, and $\tilde{Q}_{22} = LQL'$. The likelihood (13) involves only Ly and $\boldsymbol{\theta}_2$; neither $\boldsymbol{\theta}_1$ nor the part of the data in the span of X (Ky) remain in the likelihood after integrating out the fixed effects. Assuming $\tilde{Q}_{12} \neq 0$, the data indirectly identify $\boldsymbol{\theta}_1$ through $\boldsymbol{\theta}_1$'s prior correlation with $\boldsymbol{\theta}_2$ (14), so the strength of $\boldsymbol{\theta}_1$'s identification depends on τ_s and $\tilde{Q}_{12}$. If τ_s is small or $\tilde{Q}_{12} \approx 0$, the identification of $\boldsymbol{\theta}_1$ is weak. In fact, if each column of X is an eigenvector of Q, $\tilde{Q}_{12} = 0$, and $\boldsymbol{\theta}_1$ is not identified.

As for variable deletion in the ordinary linear model, setting $\boldsymbol{\theta}_1$ to zero alleviates collinearity. With $\boldsymbol{\theta}_1 \equiv 0$, the model in ((11), (12)) becomes

$$\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}_2, \tau_e \sim N(X\boldsymbol{\beta} + L'\boldsymbol{\theta}_2, \tau_e I_n) \tag{15}$$

$$\boldsymbol{\theta}_2|\tau_s \sim N(0, \tau_s \tilde{Q}_{22}) \tag{16}$$

Conditional on the variances, $\boldsymbol{\beta}$'s posterior under this model is the same as its posterior under the OLM, normal with mean $\hat{\boldsymbol{\beta}}_{OLM}$ and variance $(\tau_e X'X)^{-1}$. However, compared to the OLM, in spatial regression, two factors can inflate β_{SEc}'s posterior variance: collinearity with the spatial random effects and reduction in the effective number of observations because of spatial correlation in the data. The reduced model in (15),(16) removes any increase due to collinearity, but the marginal posterior of $\boldsymbol{\beta}$ will have larger variance than the marginal posterior of $\boldsymbol{\beta}$ from OLM because τ_e is decreased by the spatial correlation of the residuals.

To see this, first transform from (τ_e, τ_s) to $(\tau_e, r = \tau_s/\tau_e)$. After integrating out $\boldsymbol{\beta}$ and b, the marginal posterior of τ_e given r is

$$\tau_e | r, \mathrm{y} \sim \mathrm{Gamma}(.5(n-p) + a_e,\ .5(L\mathrm{y})' \left(I - (I + r\tilde{Q}_{22})^{-1}\right)(L\mathrm{y}) + b_e), \tag{17}$$

where τ_e's prior is Gamma(a_e, b_e). The marginal posterior of τ_e from OLM is Gamma$(.5(n-p) + a_e,\ .5(L\mathrm{y})'(L\mathrm{y}) + b_e)$. Since $(L\mathrm{y})' \left(I - (I + r\tilde{Q}_{22})^{-1}\right)(L\mathrm{y}) < (L\mathrm{y})'(L\mathrm{y})$ for all r, the marginal posterior of the error precision in the reduced CAR model in (15), (16) is stochastically smaller than the marginal posterior of the error precision in the OLM.

4 Influence and collinearity in the CAR model with non-normal observables

As is often the case with disease-mapping models, the analysis of the Slovenia stomach cancer data in Section 1 assumed that outcomes followed Poisson distributions. Although the intuition gained from studying collinearity in the linear model with Gaussian errors is useful in studying collinearity in the generalized linear model, several authors have shown that the diagnostics used in linear models can be misleading when applied to generalized linear models (Mackinnon and Puterman 1998, Schaefer 1979).

Spatial models for non-Gaussian data were popularized by Clayton and Kaldor (1987). Assume the observations y_i are independent given η_i with conditional log-likelihood

$$log\left(p(y_i|\eta_i)\right) = y_i\eta_i - b(\eta_i) + c(y_i), \tag{18}$$

where $E(y_i|\eta_i) = \mu_i$, and $g(\mu_i) = \eta_i = X_i\boldsymbol{\beta} + S_i + H_i$ for some link function g. The linear predictor η is the sum of three vectors: $X\boldsymbol{\beta}$, where X is a known $n \times p$ matrix of covariates and $\boldsymbol{\beta}$ is a p-vector of fixed effects; the spatial random effects S, which follow a $CAR(\tau_s Q)$ distribution; and heterogeneity random effects H, which follow a Normal$(0, \tau_h I_n)$ distribution, where $\tau_h I_n$ is a precision matrix.

Following Section 2, the linear predictor can be rewritten to highlight collinearity issues. After reparameterizing from S to $\mathrm{b} = Z'S$ where $Q = ZDZ'$ for $n \times n$ orthogonal matrix Z and $n \times n$ diagonal matrix D with positive diagonal elements $d_1 \geq ... \geq d_{N-G}$, the linear predictor is $\eta = X\boldsymbol{\beta} + Z\mathrm{b} + H$ where b's prior is normal with mean zero and precision $\tau_s D$.

The diagnostics of Sections 2.1 and 2.2 required closed-form expressions of $\boldsymbol{\beta}$'s posterior mean and variance. Non-Gaussian likelihoods prohibit closed-form expressions but these diagnostics can be extended to the non-Gaussian case using approximate methods (Lee and Nelder, 1996). One method for computing the posterior mode of $p(\boldsymbol{\beta}, \mathrm{b}, H|\tau_h, \tau_s, \mathrm{y})$ is iteratively re-weighted least squares (IRLS). This method begins with an initial value, $(\boldsymbol{\beta}^{(1)}, \mathrm{b}^{(1)}, H^{(1)})$, and the new value at $(t+1)^{th}$ iteration is computed using the recurrence equation

$$\begin{pmatrix} \boldsymbol{\beta}^{(t+1)} \\ \mathrm{b}^{(t+1)} \\ H^{(t+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}^{(t)} \\ \mathrm{b}^{(t)} \\ H^{(t)} \end{pmatrix} - h(\boldsymbol{\beta}^{(t)}, \mathrm{b}^{(t)}, H^{(t)})^{-1} \begin{pmatrix} X' \\ Z' \\ I \end{pmatrix} \left(y - \mu(\boldsymbol{\beta}^{(t)}, \mathrm{b}^{(t)}, H^{(t)})\right), \tag{19}$$

where $\mu(\boldsymbol{\beta}, \mathrm{b}, H)_j = E(y_j|\boldsymbol{\beta}, \mathrm{b}, H)$,

$$h(\boldsymbol{\beta}, \mathrm{b}, H) = \begin{pmatrix} X'W(\boldsymbol{\beta}, \mathrm{b}, H)X & X'W(\boldsymbol{\beta}, \mathrm{b}, H)Z & X'W(\boldsymbol{\beta}, \mathrm{b}, H) \\ Z'W(\boldsymbol{\beta}, \mathrm{b}, H)X & Z'W(\boldsymbol{\beta}, \mathrm{b}, H)Z + \tau_s D & Z'W(\boldsymbol{\beta}, \mathrm{b}, H) \\ W(\boldsymbol{\beta}, \mathrm{b}, H)X & W(\boldsymbol{\beta}, \mathrm{b}, H)Z & W(\boldsymbol{\beta}, \mathrm{b}, H) + \tau_h I \end{pmatrix} \tag{20}$$

is the Hessian matrix, and $W(\boldsymbol{\beta}, \mathrm{b}, H)$ is diagonal with $W(\boldsymbol{\beta}, \mathrm{b}, H)_{jj} = Var(y_j|\boldsymbol{\beta}, \mathrm{b}, H)$.

To assess each region's contribution to the difference between the posterior mode of $p(\boldsymbol{\beta}|\tau_e, \tau_s, \mathrm{y})$ with ($\hat{\boldsymbol{\beta}}$) and without ($\boldsymbol{\beta}^*$) the random effects in the model, we approximate $\hat{\boldsymbol{\beta}}$ using one IRLS step with initial value $(\boldsymbol{\beta}, \mathrm{b}, H) = (\boldsymbol{\beta}^*, 0, 0)$, i.e.,

$$\begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathrm{b}} \\ \hat{H} \end{pmatrix} \approx \begin{pmatrix} \boldsymbol{\beta}^* \\ 0 \\ 0 \end{pmatrix} - h(\boldsymbol{\beta}^*, 0, 0)^{-1} \begin{pmatrix} X' \\ Z' \\ I \end{pmatrix} (y - \mu(\boldsymbol{\beta}^*, 0, 0)) . \tag{21}$$

Using this approximation, $\Delta = \boldsymbol{\beta}^* - \hat{\boldsymbol{\beta}} \approx C\,(\mathrm{y} - \mu(\boldsymbol{\beta}^*, 0, 0))$ where C is the $p \times n$ matrix whose rows are the first p rows of $h(\boldsymbol{\beta}^*, 0, 0)^{-1}(X\ Z\ I)'$. As with the diagnostic for normal data in Section 2.1, this approximation of Δ is the product of a $p \times n$ matrix and the n-vector of residuals from the fit without the spatial random effects. In Section 5 the $\delta_{ij} = C_{ij}(y_j - \mu(\boldsymbol{\beta}^*, 0, 0)_j)$ are plotted to search for municipalities of Slovenia that are highly influential in the change of β_{SEc} from the non-spatial to the spatial regression model.

Using Fisher's observed information, the posterior variance of $(\boldsymbol{\beta}, \mathrm{b}, H)$ given (τ_h, τ_s) can be approximated by $\mathrm{Var}(\boldsymbol{\beta}, \mathrm{b}, H|\tau_s, \tau_h, \mathrm{y}) \approx h(\hat{\boldsymbol{\beta}}, \hat{\mathrm{b}}, \hat{H})^{-1}$ where $(\hat{\boldsymbol{\beta}}, \hat{\mathrm{b}}, \hat{H})$ is the mode of $p(\boldsymbol{\beta}, \mathrm{b}, H|\tau_h, \tau_s, \mathrm{y})$. The normal-case variance inflation factor due to adding spatial effects (9) extends to the present non-Gaussian models as the ratio of $\left[h(\hat{\boldsymbol{\beta}}, \hat{\mathrm{b}}, \hat{H})^{-1}\right]_{ii}$ and the approximate posterior variance of $\boldsymbol{\beta}_i$ under the non-spatial model,

$$VIF_i(\tau_s, \tau_h) = \frac{\left[h(\hat{\boldsymbol{\beta}}, \hat{\mathrm{b}}, \hat{H})^{-1}\right]_{ii}}{\left[(X'W(\boldsymbol{\beta}^*, 0, 0)X)^{-1}\right]_{ii}}. \tag{22}$$

In the normal case, the variance inflation factor (9) was a function of only $r = \tau_s/\tau_e$ and did not depend on τ_e. In the non-normal case, the variance inflation factor (22) is a function of both τ_s and τ_h. The spatial model for non-normal outcomes in (18) adds both spatial and heterogeneity random effects to the simple generalized linear model. If either of these random effects is given a flat prior, i.e., if either τ_s or τ_h is zero, $\boldsymbol{\beta}_i$ will be perfectly correlated with some combination of the random effects and VIF_i will be infinite.

The reduced model in Section 3 diminished the variance inflation of the fixed effects caused by collinearity by deleting from the spatial smooth linear combinations of the spatial random effects that were correlated with the fixed effects. Extending this method to the non-normal case requires removing the combinations of both b and H that are correlated with $\boldsymbol{\beta}$. $h(\hat{\boldsymbol{\beta}}, \hat{\mathrm{b}}, \hat{H})$ resembles the Hessian of the linear mixed model with linear predictor $\hat{W}^{1/2}X\boldsymbol{\beta} + \hat{W}^{1/2}Z\mathrm{b} + \hat{W}^{1/2}H$ where $\mathrm{b} \sim N(0, \tau_s D)$, $H \sim N(0, \tau_h I)$, and $\hat{W} = W(\hat{\boldsymbol{\beta}}, \hat{\mathrm{b}}, \hat{H})$. To remove the collinearity between the fixed effects and the random effects in this model, we could delete the combinations of $\hat{W}^{1/2}Z\mathrm{b}$ and $\hat{W}^{1/2}H$ that are in the span of $\hat{W}^{1/2}X$. Following the steps of Section 3 to remove the collinear combinations of $\hat{W}^{1/2}Z\mathrm{b}$ and $\hat{W}^{1/2}H$ gives a reduced generalized linear mixed model analogous to the reduced Gaussian model in (15), (16), i.e.,

$$\eta = X\boldsymbol{\beta} + \hat{W}^{-1/2}L'\boldsymbol{\theta}_2 + \hat{W}^{-1/2}L'\boldsymbol{\gamma}_2 \tag{23}$$

$$\boldsymbol{\theta}_2 = L\hat{W}^{1/2}Z\mathrm{b} \sim N(0, \tau_s L\hat{W}^{1/2}Q\hat{W}^{1/2}L') \tag{24}$$

$$\boldsymbol{\gamma}_2 = L\hat{W}^{1/2}H \sim N(0, \tau_h L\hat{W}L') \tag{25}$$

where L is the $(n-p)\times p$ matrix whose rows are the eigenvectors of $I - \hat{W}^{1/2}X(X'\hat{W}X)^{-1}X'\hat{W}^{1/2}$ that correspond to non-zero eigenvalues.

Let $(\hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\gamma}}_2)$ be the mode of $p(\boldsymbol{\beta}, \boldsymbol{\theta}_2, \boldsymbol{\gamma}_2|\tau_h, \tau_s, \mathrm{y})$ and $\hat{W}_2 = W(\hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\gamma}}_2)$. Using Fisher's observed information approximation, under (23)–(25) $\mathrm{Var}(\boldsymbol{\beta}, \boldsymbol{\theta}_2, \boldsymbol{\gamma}_2|\tau_h, \tau_s, \mathrm{y})$ is approximately

$$\begin{pmatrix} X'\hat{W}_2X & X'\hat{W}_2\hat{W}^{-1/2}L' & X'\hat{W}_2\hat{W}^{-1/2}L' \\ L\hat{W}^{-1/2}\hat{W}_2X & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' + \tau_s L\hat{W}^{1/2}Q\hat{W}^{1/2}L' & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' \\ L\hat{W}^{-1/2}\hat{W}_2X & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' + \tau_h L\hat{W}L' \end{pmatrix}^{-1}. \tag{26}$$

The posterior of η should be similar under the full model (18) and the reduced model (23) – (25)

because the only difference in these models is that redundancies in η have been removed. Since W is a function of η, the estimate of W from the full model ($\hat{W}$) and reduced model ($\hat{W}_2$) should be similar. If $\hat{W} \approx \hat{W}_2$, $X'\hat{W}_2\hat{W}^{-1/2}L' \approx 0$ (since L is, by construction, orthogonal to $\hat{W}^{1/2}X$) and $\boldsymbol{\beta}$ is approximately uncorrelated with $\boldsymbol{\theta}_2$ and $\boldsymbol{\gamma}_2$.

5 Analysis of Slovenia data

In this section, the methods of Sections 2-4 are used to investigate influence and collinearity in the analysis of the Slovenia cancer data introduced in Section 1. As mentioned in Section 1, the posterior variance of β_{SEc} increases from 0.0004 under the simple Poisson regression model to 0.0016 under the spatial Poisson regression model. Equation 10 predicts that variance inflation due to the addition of CAR parameters will be large if there is little spatial smoothing of the CAR parameters and if the covariate is highly correlated with low-frequency eigenvectors of Q. Both of these criteria are met in the Slovenia analysis: the effective number of parameters in the model is $p_D = 62.3$ (Table 1) out of a possible n=192 and the correlation between SEc and the eigenvector of Q corresponding to the smallest eigenvalue — the Southwest/Northeast gradient plotted in Figure 3b — is 0.72.

Also, the posterior median of β_{SEc} shifts from -0.137 to -0.022 after adding the spatial terms. Figure 4a depicts the joint posterior of (β_{SEc}, τ_s) and shows how this shift in β_{SEc}'s median depends on τ_s, the parameter that controls smoothing of the CAR parameters. For small τ_s, the CAR parameters are unsmoothed and β_{SEc}'s median is close to zero. As τ_s increases, β_{SEc}'s median gradually tends towards β_{SEc}'s posterior median from the simple Poisson regression model (the middle horizontal dashed line). For the range of τ_s given posterior mass, β_{SEc}'s posterior median does not reach its posterior median from the simple Poisson regression model

Insert Figure 4 about here

The δ_{ij} statistics of Section 4 can be used to measure each region's contribution to the change in β's posterior mode from adding the spatial random effects. These statistics are plotted in Figure 5 for the Slovenia data with $(\tau_s, \tau_h) = (10.5, 125.9)$, their posterior medians under Model 2 (Table 1). $\delta_j < 0.010$ for each municipality except Murska Sobota and Ptuj, the two northeastern regions that are shaded blue in Figure 5. These regions have $\delta_j > 0.015$ and are largely responsible for the change in β_{SEc}. Murska Sobota and Ptuj are unusual because despite having higher SEc than their neighbors (Figure 1b), they do not have lower SIR than their neighbors (Figure 1a). Comparing these regions to their neighbors contradicts the general pattern that regions with high SEc have low SIR.

Insert Figure 5 about here

Another feature that distinguishes Murska Sobota and Ptuj from their neighbors is that they are two of the eleven municipalities the Slovenian government defines as urban. Since rural and urban neighbors are probably less similar than neighboring rural regions, smoothing these two types of neighbor pairs equally may not be appropriate. Smoothing all neighbors equally may be more sensible after including an urban indicator as a fixed effect to account for rural/urban differences. The model with a fixed effect for urban/rural is summarized in the third row of Table 1. Adding an urban fixed effect leads to only a small change in β_{SEc} and β_{urban}'s posterior 95% interval covers zero. Accounting for urban/rural differences does not mediate the effect of collinearity on β_{SEc}.

The model in (23)–(25) resolves the collinearity problem by removing the combinations of the CAR random effects that are collinear with the fixed effects. Under the usual model, β_{SEc}'s posterior depends on τ_s, the parameter controlling the smoothing of the CAR random effects (Figure 4a). Figure 4b shows that deleting the collinear spatial terms removes β_{SEc}'s dependence on τ_s; β_{SEc}'s median under the reduced model is similar for all τ_s. Removing the collinear terms also reduces β_{SEc}'s variance from 0.0016 to 0.0008, still greater than 0.0004 without the spatial or heterogeneity random effects.

Figure 6 plots e^{η_j} under the usual CAR model (Figure 6a) and the model without the collinear

CAR parameters (Figure 6b). Although removing the collinear terms from the spatial regression impacts the posterior of the fixed effects, the posterior mean relative risks (e^{η_j}) under the reduced model (Figure 6b) are similar to the full spatial model (Figure 6a). The remaining spatial random effects in Model 4 are smoothed more than the spatial random effects in Model 2 (the posterior median of τ_s increases from 10.5 under Model 2 to 15.0 under model 4). However, the remaining heterogeneity random effects in Model 4 are smoothed less than the heterogeneity random effects in Model 2, (the posterior median of τ_h decreases from 125.9 under Model 2 to 38.9 under model 4). As a result, the fit under the model without the collinear spatial terms is slightly less smooth ($p_D = 70.0$) than the usual spatial regression model ($p_D = 62.3$).

Insert Figure 6 about here

Table 1 shows that the posterior of β_{SEc} after removing the collinear spatial effects (Model 4) is more similar to its posterior under the simple Poisson model (Model 1) than its posterior under the usual spatial regression (Model 2), but β_{SEc}'s posterior is not identical under Models 1 and 4. Under Model 4, β_{SEc} posterior median is -0.120 compared to -0.137 under Model 1. In deriving Model 4 (Section 4), we assumed the posteriors modes of e^{η_j} were the same under Models 2 and 4. The slight difference in β_{SEc}'s posterior median under Models 1 and 4 may be due to slight differences in posteriors of e^{η_j} under Models 2 and 4 (Figure 6).

Also, β_{SEc}'s posterior variance is twice as large under Model 4 (0.0008) compared to Model 1 (0.0004). Two factors inflate β_{SEc}'s marginal posterior variance from Model 1 to Model 2: collinearity with the spatial random effects and reduction in the effective number of observations due to consideration of spatial correlation in the data. The variance of β_{SEc} under Model 4 is smaller than its variance under Model 2 because the collinearity effect is removed but larger than its variance under Model 1 because the spatial correlation in the data is taken into account.

6 Discussion

This paper investigated the effect of the spatial random effects on estimation of fixed effects in CAR models for disease mapping. Assuming normal outcomes, Section 2.1 proposed the C_{ij} and δ_{ij} statistics to investigate each region's contribution to the change in the fixed-effect estimate from including spatial random effects. The Slovenia example was used to illustrate that regions were highly influential if they have different covariates than nearby regions and also have large OLS residuals. Section 2.2 showed that the variance inflation of the fixed effects due to collinearity with the spatial random effects is most troublesome if there is little spatial smoothing and the covariates are highly correlated with low-frequency eigenvectors of the adjacency matrix Q. A model that removes the collinear spatial terms was developed in Sections 3 and 4 and applied to the Slovenia data in Section 5.

The results of this paper were developed for the CAR model of areal data but can be extended to other spatial models. A common geostatistical model for point-referenced data is

$$\begin{aligned} \mathrm{y}|\boldsymbol{\beta}, S, \tau_e &\sim N(X\boldsymbol{\beta} + S, \tau_e I_n) \\ S|\tau_s &\sim N(0, \tau_s Q(\phi)), \end{aligned} \tag{27}$$

where ϕ is a vector of unknown range parameters. Mechanically, the only difference between this model and the CAR model in (2),(3) is that S's prior precision depends on ϕ. This complicates matters because the geometry of the model, i.e., the eigenvectors and eigenvalues of S's prior precision, depend on ϕ. For example, the collinearity between the fixed effects and the random effects was illuminated in (4), (5) by a transformation that depended on Q's eigenvectors. This transformation would be more difficult to interpret if the eigenvectors themselves were functions of unknown parameters as in (27). However, since the methods developed in Sections 2-4 conditioned on the precisions, (τ_e, τ_s), these methods could be applied to the geostatistical model by simply conditioning on (τ_e, τ_s, ϕ).

References

Besag J, York JC, Mollié A (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, **43**, 1–59.

Best NG, Waller LA, Thomas A, Conlon EM, Arnold RA (1999). Bayesian models for disease and exposure data (with discussion). *Bayesian Statistics 6*: 131-156.

Clayton DG, Kaldor J (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrika*, **43**, 671–681.

Hocking RR (1996). Methods and Applications of Linear Models: Regression and the Analysis of Variance. New York, NY: Wiley Interscience.

Hodges JS, Carlin BP, Fan Q (2003). On the precision of the conditionally autoregressive prior in spatial models. *Biometrics*, **59**, 317-322.

Lee Y, Nelder JA (1996). Hierarchical Generalized Linear Models (with discussion and rejoinder). *J. Roy. Statist. Soc., Ser. B*, **58**, 619-678.

Mackinnon MJ, Puterman ML (1998). Collinearity in generalized linear models. *Commun. Statist. – Theory Meth*, **18(9)**, 3463-3472.

Reich BJ, Hodges JS (2005). Identification of the variance components in the general two-variance linear model. Research Report 2005-009, Division of Biostatistics, University of Minnesota, 2004. Submitted to *J. Roy. Statist. Soc., Ser. B.*

Schaefer RL (1979). Multicollinearity in logistic regression. PhD Thesis, University of Michigan.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002). Bayesian measures of model complexity and fit (with discussion and rejoinder) *J. Roy. Statist. Soc., Ser. B*, **64**, 583-639.

Table 1: DIC and p_D of various models along with 95% posterior confidence intervals for β (fixed effect parameters), τ_s (the prior precision of the CAR random effects), and τ_h (the prior precision of the heterogeneity random effects). The four models are (1) simple Poisson regression, (2) spatial regression with CAR and heterogeneity random effects and a fixed effect for SEc, (3) Model 2 with a fixed effect for Urban, and (4) Model 2 the CAR effects in the span of SEc removed.

Model	DIC	p_D	β_{SEc}	β_{Urban}	τ_s	τ_h
1	1153.0	2.0	(-0.175, -0.098)	—	—	—
2	1081.5	62.3	(-0.100, 0.057)	—	(4.7, 77.4)	(21.0, 1224.0)
3	1081.8	66.9	(-0.133 , 0.044)	(-0.06, 0.25)	(5.4, 89.2)	(18.9, 248.1)
4	1088.0	70.0	(-0.166, -0.069)	—	(4.9, 175.1)	(16.3, 236.8)

Figure 1: Plot of each municipality's observed standardized morality ratio ($SIR = O/E$) and centered socioeconomic status (SEc).

(a) SIR (b) SEc

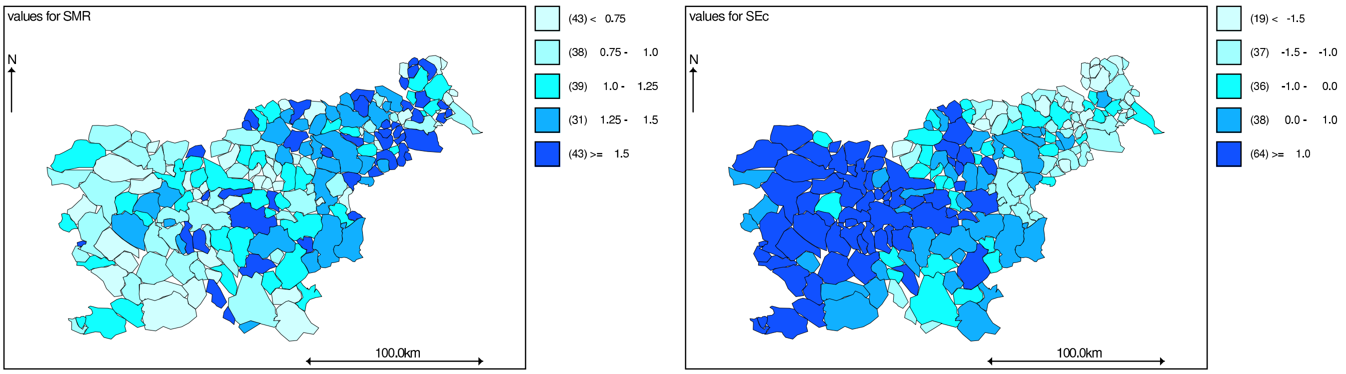

Figure 2: Plot of $|C|$ for the Slovenia grid with $X = SEc$ and (a) $r = 0.1$ and (b) $r = 10$.

(a) $r = 0.1$ (b) $r = 10$

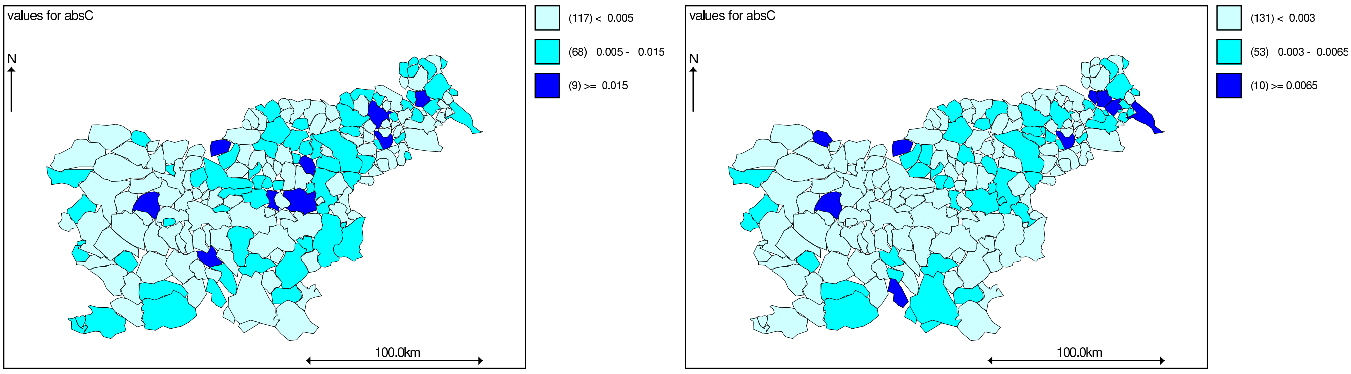

Figure 3: Eigenvectors associated with the smallest and largest eigenvalues of the municipalities of Slovenia adjacency matrix. Panel (a) shows the eigenvector associated with the smallest eigenvalue, $d_{193} = 0.03$, while panel (b) shows the eigenvector associated with the largest eigenvalue, $d_1 = 14.46$.

(a) (b)

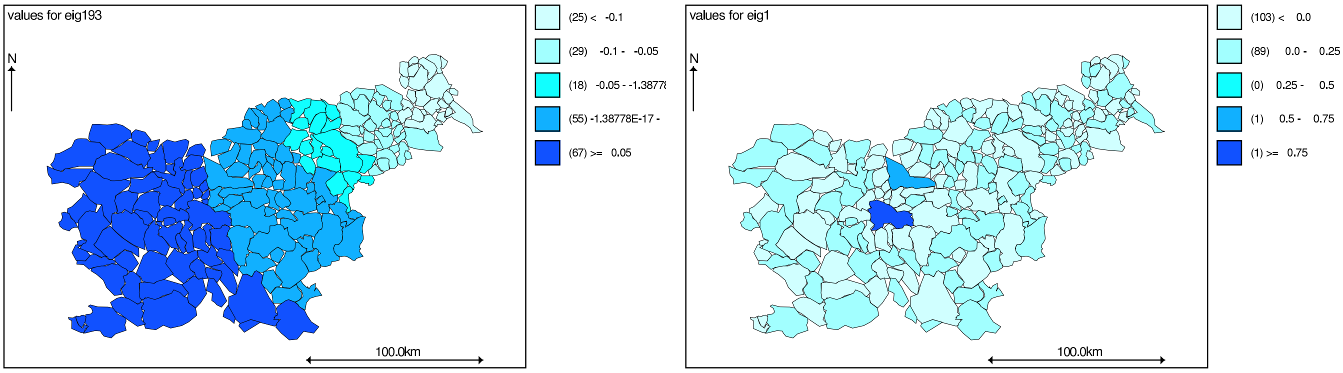

Figure 4: Plot of (β_{SEc}, τ_s)'s joint posterior with the combinations of the random effects that are collinear with SEc (a) included as in Section 1 and (b) excluded as in (23)–(25). The 30,000 samples of β_{SEc} are divided into bins according to the corresponding draws of τ_s. A box-plot is drawn for the samples of β_{SEc} in each bin with plot-width indicating the number of samples in the bin. The dashed lines represent β_{SEc}'s posterior median and 95% confidence interval from the Poisson regression without spatial terms.

(a) (b)

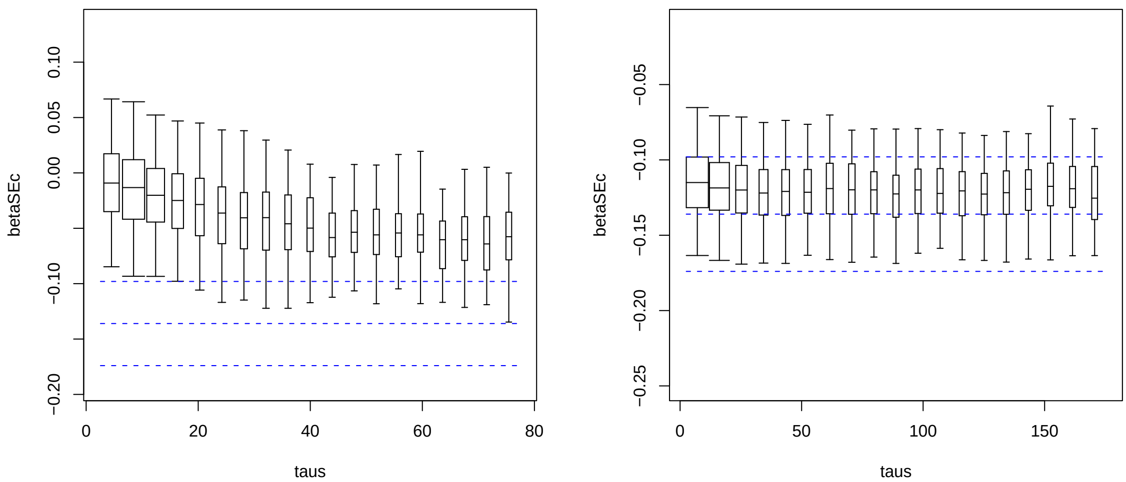

Figure 5: Plot of $|\delta|$, i.e., each region's contribution to the change in β's posterior mode due to the spatial terms, with (τ_s, τ_h) set to their posterior medians (10.5, 125.9).

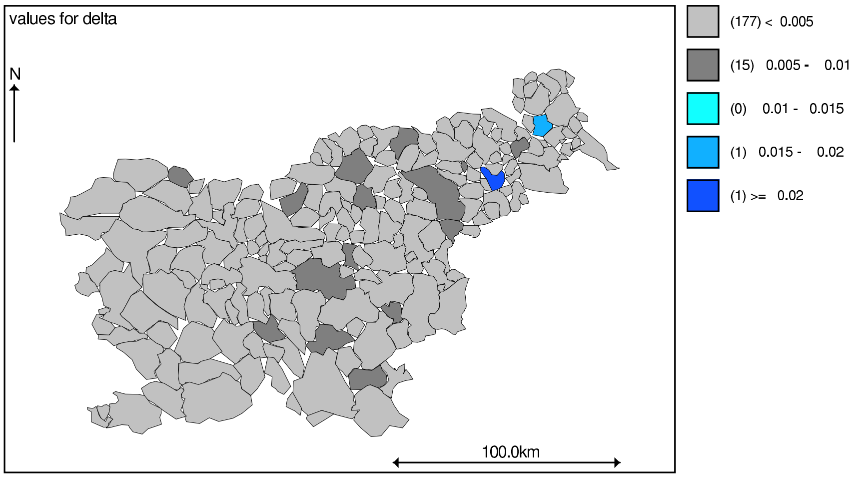

Figure 6: Plot of each municipality's posterior mean relative risk (e^{η_j}) from (a) the usual spatial regression model of Section 1 and (b) the spatial regression model of Section 4 that excludes the CAR parameters that are collinear with SEc.

(a) (b)