Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Multivariate Analysis 97 (2006) 246–273

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Bayesian multivariate spatial models for roadway traffic crash mapping

J.J. Song[a], M. Ghosh[b], S. Miaou[c], B. Mallick[a,*]

[a]*Department of Statistics, TAMU 3143, Texas A&M University, College Station, TX 77843-3143, USA*
[b]*Department of Statistics, University of Florida, Gainesville, USA*
[c]*Texas Transportation Institute, Texas A&M University, College Station, USA*

Received 16 March 2004
Available online 24 May 2005

Abstract

We consider several Bayesian multivariate spatial models for estimating the crash rates from different kinds of crashes. Multivariate conditional autoregressive (CAR) models are considered to account for the spatial effect. The models considered are fully Bayesian. A general theorem for each case is proved to ensure posterior propriety under noninformative priors. The different models are compared according to some Bayesian criterion. Markov chain Monte Carlo (MCMC) is used for computation. We illustrate these methods with Texas Crash Data.

AMS 1991 subject classification: 91B72

Keywords: Hierarchical models; Markov chain Monte Carlo; Multivariate CAR; Noninformative priors; Posterior propriety

1. Introduction

The analysis of spatially correlated data has been an increasingly active area of both methodological and applied statistical research. Sophisticated computer software known as Geographical Information system (GIS) has allowed the scientists to incorporate geographical information about the phenomenon being studied. This motivated the statisticians

* Corresponding author. Fax: +979 845 3144. Bani K. Mallick, Department of Statistics, TAMU.
E-mail address: bmallick@stat.tamu.edu (B.K. Mallick).

0047-259X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2005.03.007

to develop and analyze models that can account for spatial clustering and variation. The conditional autoregressive (CAR) models [2] have found wide application for the analysis of spatial data in a univariate setup.

In this paper, we will explore the extension of spatial models to a multivariate setup. Such models are necessary to analyze more than one type of events simultaneously, since a number of different events may share the same set of risk factors. We will propose four multivariate models to improve risk estimates. In the first two models, the correlation among the regions is induced by a random error term and this is a spatial analog of "shared component" models proposed by [18]. The third model is a multivariate CAR model following a suggestion of [20]. The fourth model will be based on the correlated CAR structure where the correlation is induced through the scale parameters of the CAR model.

Since the CAR prior is usually improper, it is imperative to check the propriety of the joint posterior. Usually a remedy to this problem is to introduce a suitably constrained autoregression parameter which ensures a proper joint distribution for the resulting CAR model (see Cressie [34], Sun et al [35] for univariate models; Gelfand and Vounatsu [36] for multivariate CAR models). Instead we have considered here improper priors. One of the main contributions of this paper is to provide sufficient conditions to obtain a proper posterior distribution for each of these four popular multivariate models. [14] provided sufficient conditions to obtain a proper posterior for the univariate CAR prior. We generalize their result in several multivariate setups.

The motivating example of our paper is from traffic crash mapping. Transportation-related injuries and deaths cause major health problems in the United States. Injuries and fatalities occur in all transportation modes. In particular crashes involving motor vehicles account for almost 95 percent of all transportation fatalities and most injuries. Despite the progress made in highway safety in the past three decades, tens of thousands of people are still killed and millions of people are injured in motor vehicle crashes each year in this country. For example, 42,000 people were killed in traffic crashes, and almost 3.2 million more were injured in 1998. This makes motor vehicle fatalities the leading cause of injury deaths in the United States, accounting for about 29 percent of all injury deaths in recent years. They are also responsible for as many preretirement years of life lost as cancer and heart disease, about 1.2 million years annually. Societal economic losses from these crashes are huge, estimated by the National Highway Traffic Society (NHTSA) to exceed \$150 billion annually. Thus, much work remains to be done to develop a better understanding of the cause of vehicle crashes and to prevent them [5].

Motor vehicle crashes are complex events involving the interactions of five major factors: the driver, the traffic, the road, the vehicle and the environment (weather and lighting condition) [23]. Among these factors, "driver error" has been identified as the main contributing factor to a great percentage of vehicle crashes, and many research efforts are being undertaken to better understand the human factors that cause or facilitate crashes. Recognizing that "to err is human" and that driver behavior is affected by virtually all elements of the roadway environment, highway engineers are constantly designing and building (or redesigning and rebuilding) roadways to meet higher standards that are safer for the traveler. This includes building roadways that are more "forgiving" when an error is made, more conforming to the physical and operational demands of the vehicle, and better meeting the riding expectancies of the driver [32]. The remarkable low fatality rate on the

Interstate Highway System is an evidence of the impact of good roadway design on highway safety [10].

However, many impediments keep highway engineers from achieving the desired design and operational goals. These include lack of resources and a vast highway system that needs to be built, maintained, operated and improved. Highway engineers continue to juggle available resources to make incremental safety improvements, which often require them to make difficult decisions on the trade-off between cost and safety. Statistical models that describe the relationship between vehicle crash frequency and different factors are one of the major tools used by highway engineers to make their cost-safety trade-off decisions. These models are typically named vehicle crash (accident) prediction models in the highway safety literature.

Though highway safety community is a latecomer in the application of generalized linear models (GLM) in data analysis, recently there a surge of applications of GLM in highway safety research. The use of overdispersed Poisson, including the negative-binomial regression models and their variations has become very popular. Examples include Morris et al. [28], Hauer [16], Miaou et al. [24], Miaou and Lum [26], Miaou [22], Miaou [23], Bonneson and McCoy [4], Maher and Summersgill [19], Shankar et al. [30], and Vogt and Bared [33]. Adjusting for the regression-to-the-mean and local effect has been an important problem surrounding many "before-after" safety evaluation and problem site identification studies using empirical-Bayes estimators [16,7,12]. Also, the use of logistic and ordered probit regression models has now become fairly common in studying the factors that affect the crash severity [9,21].

Most of the above-mentioned papers ignore the spatial dependence among the crash data. A very recent exception is Miaou et al. [27] who studied the geographical pattern of accidents in the state of Texas. The analysis of spatially referenced data has been an increasingly active area of both methodological and applied statistical research. Sophisticated computer programs known as GIS have revolutionized the analysis and display of such data sets, through their ability to "layer" multiple data sources over a common study area. Finally, Markov chain Monte Carlo (MCMC) algorithms enable the fitting of complex hierarchical models in a full Bayesian framework, permitting full posterior inference for underlying parameters in complex model settings. That way one can avoid the naive empirical Bayes analysis which usually underestimates uncertainties related to the model.

Roeleven et al. [29] considered a modeling situation with multiple type of accidents and exploited GLM (Binomial regression) to obtain the risk estimates. As the risk set is large and the chance of accident is small, usually the binomial regression is well-approximated by the Poisson regression in most situations. However, in the present situation, this approximation has several limitations. First, for these type of data, variation in the observed number of accidents exceeds that expected from regular Poisson inference. In a given area, variation in the observed number of accidents is partly due to Poisson sampling, but it is also subject to extra-Poisson variation arising from variability in the accident rate within the area, which results from heterogeneity in individual risk levels within the area. Bayesian inference is more suitable for these problems as in addition to the observed events in each area, prior information is used on the variability of accident rates in the overall map. Indeed, Bayesian estimates of area-specific accident rates integrate these two types of information. Bayesian estimates are close to the area-specific standardized accident rates, when based on a large number of accidents within the area, but with a few accidents, prior information over all the areas will dominate, thereby shrinking standardized rates towards the overall mean accident rate. Also, Roeleven et al. ignored the correlation among the types of accidents and fitted independent GLM models for each type of accident. Thus, they failed to borrow strength or to share information from similar sources, as well as from the most directly available sources to improve the risk estimates and the variance of their risk estimates are inflated. They also did not consider the spatial pattern in accidents, i.e. the tendency for geographically close areas to have similar accident rates. Expressing the geographical dependence through prior information, a Bayesian approach shrinks the usual estimates of the risk and other model parameters in an area towards a local mean, thereby producing more stable estimates. Ignoring this spatial effect fails to provide a risk map, and also the risk and other parameter estimates can have much larger variances due to this lack of shrinkage.

We analyze traffic crash data involving multivariate measurements which are the different types of crashes. Multivariate spatial models are used. These models are necessary to analyze more than one type of crash simultaneously, since a number of different crashes may share the same set of risk factors. As an example, for different types of crashes, the risk factor could be the excessive curvature or the bad condition of the road. The main purpose of this work is to borrow strength or share information from similar sources, as well as use the most directly available sources, to improve crash risk estimates. Estimation of crash risk for a particular crash type may be improved by using information from other types of crash. A Bayesian criterion is used to choose the best fitted model for our data.

The outline of the remaining sections is as follows. In Section 2 of this paper, we review briefly the univariate hierarchical Bayesian model. Several versions of multivariate hierarchical Bayesian models are introduced in Section 3. Data analysis based on both the univariate and multivariate models is carried out in Section 4. Some concluding remarks are made in Section 5. The proofs of some of the technical results are deferred to the Appendix.

2. Univariate hierarchical model

Let $y_1, y_2, \ldots, y_n$ denote the number of crashes in a given period of time for the n regions. Conditional on $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^T$, $y_1, \ldots, y_n$ are assumed to be independent with pdfs

$$p(y_i|\theta_i) = \exp(y_i\theta_i - \Psi(\theta_i))h(y_i).$$

This is the one-parameter exponential family model.

Ghosh et al. [14] developed a hierarchical model as $\theta_i = q_i + x_i^T\boldsymbol{\beta} + \eta_i + e_i$ for $i = 1, \ldots, n$, where q_i is a known parameter. The x_i are region-level covariates , having parameter coefficient $\boldsymbol{\beta}$. The e_i capture region-wide heterogeneity via an exchangeable normal prior. Finally, the η_i are the parameters that make this a truly spatial model by capturing regional clustering. The spatial random effects η_i and the random errors e_i are assumed to be mutually independent, Also the η_i have a *pairwise difference prior* with joint pdf

$$p(\boldsymbol{\eta}) \propto (\sigma_\eta^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{i\neq j} w_{ij}(\eta_i-\eta_j)^2\right\}, \tag{2.1}$$

where $w_{ij} = w_{ji}$ are the known weights. Such priors were introduced, and used quite extensively in Besag et al. [3]. The errors e_i were assumed to be iid with 0 mean and variance σ_e^2. Finally, β, σ_e^2, and σ_η^2 were mutually independent and $\beta \sim$ Uniform(R^p), $(\sigma_e^2)^{-1} \sim G(a/2, b/2)$, and $(\sigma_\eta^2)^{-1} \sim G(c/2, d/2)$. Throughout this paper, a random variable Z is said to have a $G(\alpha, p)$ distribution if it has a pdf of the form $f(z) \propto \exp(-\alpha z)z^{p-1}$. The joint posterior under the given prior is

$$\begin{aligned}\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{e}, r_\eta, r_e|\boldsymbol{y}) \;\propto\; & \prod_i p(y_i|\theta_i) \\ & \times\; r_e^{n/2} \exp\left\{-\frac{r_e}{2}\sum_{i=1}^n(\theta_i - q_i - x_i^T\boldsymbol{\beta} - \eta_i)^2\right\} \\ & \times\; r_\eta^{n/2} \exp\left\{-\frac{r_\eta}{2}\sum_{1\leqslant i<l\leqslant n} w_{il}(\eta_i-\eta_l)^2\right\} \\ & \times\; r_\eta^{(d/2)-1} \exp\left(-\frac{cr_\eta}{2}\right) r_e^{(b/2)-1} \exp\left(-\frac{ar_e}{2}\right),\end{aligned} \tag{2.2}$$

where $r_\eta = \sigma_\eta^{-2}$ and $r_e = \sigma_e^{-2}$.

Ghosh et al. [14] provided sufficient conditions to ensure propriety of the posterior. The Bayesian analysis to obtain samples from the posterior distributions of the unknown parameters was implemented by the MCMC technique. The full conditionals needed for such implementation are available in Ghosh et al. [14].

3. Multivariate hierarchical models

3.1. Introduction

In this section, we propose four multivariate hierarchical Bayesian spatial models. Let $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iq})^T, i = 1, \ldots, n$ denote the n response vectors. For our specific example, the responses are the numbers of crashes at n regions due to q different causes. Analogous to the previous section, we begin with the one-parameter exponential family model

$$p(y_{ij}|\theta_{ij}) = \exp[\theta_{ij}y_{ij} - \psi(\theta_{ij})]h(y_{ij}), \tag{3.1}$$

$j = 1, \ldots, q; i = 1, \ldots, n$. In the next stage, we model the θ_{ij} as

$$\theta_{ij} = \boldsymbol{x}_{ij}^T\boldsymbol{\beta}_j + \eta_{ij} + e_{ij} \;\; (j = 1, \ldots, q; i = 1, \ldots, n), \tag{3.2}$$

where the $\boldsymbol{x}_{ij}$ are p-component column vectors. We may note that the regression coefficients are assumed to be cause-specific rather than common across all causes of crash. Writing

$\boldsymbol{\theta}_i = (\theta_{i1}, \ldots, \theta_{iq})^T$, $\boldsymbol{\eta}_i = (\eta_{i1}, \ldots, \eta_{iq})^T$, $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{iq})^T$, $\boldsymbol{\beta}^T = (\boldsymbol{\beta}_1^T, \ldots, \boldsymbol{\beta}_q^T)$ and $\boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}_{i1}^T \cdots \boldsymbol{0}^T \\ \cdots\cdots\cdots \\ \boldsymbol{0}^T \cdots \boldsymbol{x}_{iq}^T \end{pmatrix}$, we can rewrite (3.2) as

$$\boldsymbol{\theta}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\eta}_i + \boldsymbol{e}_i, \; i = 1, \ldots, n. \tag{3.3}$$

Note that the $\boldsymbol{X}_i$ are matrices of order $q \times pq$. In the above, the errors $\boldsymbol{e}_i$ and the spatial effects $\boldsymbol{\eta}_i$ are assumed to be mutually independent. Throughout this paper, we assume that $\boldsymbol{e}_i \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_e)$ and $\text{rank}(\sum_{i=1}^n (\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T(\boldsymbol{X}_i - \bar{\boldsymbol{X}})) = pq$. We will introduce various spatial priors for the $\boldsymbol{\eta}_i$ in the next four subsections. In particular, we will consider various CAR priors for the $\boldsymbol{\eta}_i$. We will label these priors as CAR priors I–IV.

3.2. *CAR Prior I*

We first consider the case when $\boldsymbol{\eta}_i = \eta_i \boldsymbol{1}_q, i = 1, \ldots, n$. This amounts to the assumption that all the components of the spatial vector $\boldsymbol{\eta}_i$ in a given region are equal, i.e. the spatial influence is not cause-specific. For $\eta_1, \ldots, \eta_n$, we consider the pairwise difference prior as given in (2.1). At the final stage of the hierarchical model, it is assumed that $\boldsymbol{\beta}$, r_η and $\boldsymbol{\Sigma}_e$ are mutually independent with $\boldsymbol{\beta} \sim \text{uniform}(R^p)$, $r_\eta \sim G(a/2, b/2)$, and $\boldsymbol{\Sigma}_e$ has an inverse Wishart distribution with pdf

$$\pi(\boldsymbol{\Sigma}_e) \propto |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp[-(1/2)\text{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})].$$

This distribution will be written symbolically as IW($\boldsymbol{A}$, γ). Now writing $\boldsymbol{y} = (y_{11}, \ldots, y_{1q}, \ldots, y_{n1}, \ldots, y_{nq})^T$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)^T$ and $\boldsymbol{\theta}^T = (\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_n^T)$, the joint posterior is given by

$$\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, r_\eta | \boldsymbol{y}) \\
&\quad \propto \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
&\quad \times |\boldsymbol{\Sigma}_e|^{-n/2} \exp\left\{ -\frac{1}{2} \sum_{i=1}^n (\boldsymbol{\theta}_i - \eta_i \boldsymbol{1}_q - \boldsymbol{X}_i\boldsymbol{\beta})^T \boldsymbol{\Sigma}_e^{-1} (\boldsymbol{\theta}_i - \eta_i \boldsymbol{1}_q - \boldsymbol{X}_i\boldsymbol{\beta}) \right\} \\
&\quad \times r_\eta^{n/2} \exp\left\{ -\frac{r_\eta}{2} \sum_{1 \leqslant i < l \leqslant n} w_{il}(\eta_i - \eta_l)^2 \right\} \\
&\quad \times |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left\{ -\frac{1}{2} \text{tr} \boldsymbol{\Sigma}_e^{-1} \boldsymbol{A} \right\} \\
&\quad \times r_\eta^{(b/2)-1} \exp\left(-\frac{a r_\eta}{2} \right).
\end{aligned} \tag{3.4}$$

The prior for β is improper. We present a general theorem ensuring propriety of the posterior.

Theorem 3.1. *Assume $a > 0$, $n + b > 0$, and $n > pq + q - \gamma + 1$. Then, if*

$$\int_{-\infty}^{\infty} \exp\{y_{ij}\theta - \psi(\theta)\}d\theta < \infty \tag{3.5}$$

for all y_{ij}, the joint posterior pdf of the θ_{ij} given $\boldsymbol{y}$ is proper.

The proof of the theorem is deferred to Appendix A.

Remark 1. Assumption (3.5) holds for the Poisson model if and only if $y_{ij} > 0$ for all *i* and *j*. This is specifically suited for the present application related to highway traffic crashes. In the case when y_{ij} is Binomial(n_{ij}, p_{ij}), (3.5) holds when $1 \leqslant y_{ij} \leqslant n_{ij} - 1$ for all *i* and *j*. We just point out that (3.5) may not hold in all small area estimation problems. For example, if y_{ij} denotes the count of deaths in a small geographical area due to a rare disease (such as lung cancer), it is possible to have $y_{ij} = 0$ for certain areas in a small time interval.

Direct evaluation of the posterior of the θ_{ij} given $\boldsymbol{y}$ involves high-dimensional numerical integration and is not computationally feasible. Instead the Gibbs sampler is used requiring generation of samples from the full conditional distributions of the parameters. These conditionals are given by

$$r_\eta|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{y} \sim Gamma\left(\frac{1}{2}\left(\sum_{1\leqslant i<l\leqslant n} w_{il}(\eta_i - \eta_l)^2 + a\right), \frac{n+b}{2}\right),$$

$$\boldsymbol{\Sigma}_e|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, r_\eta, \boldsymbol{y} \sim \mathrm{IW}\left(\boldsymbol{A} + \sum_{i=1}^{n}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T, n + \gamma\right),$$

$$\boldsymbol{\beta}|\boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, r_\eta, \boldsymbol{y} \sim N_p(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta),$$

$$\boldsymbol{\mu}_\beta = \left(\sum_{i=1}^n \boldsymbol{X}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{X}_i\right)^{-1}\left(\sum_{i=1}^n \boldsymbol{X}_i^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \eta_i\mathbf{1}_q)\right), \boldsymbol{\Sigma}_\beta = \left(\sum_{i=1}^n \boldsymbol{X}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{X}_i\right)^{-1},$$

$$\begin{aligned}&\eta_i|\boldsymbol{\theta}, \boldsymbol{\beta}, \eta_l(l \neq i), \boldsymbol{\Sigma}_e, r_\eta, \boldsymbol{y}\\ &\quad\sim N\left(\frac{(\boldsymbol{\theta}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q + r_\eta w_{i+}\bar{\eta}_i}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q + r_\eta w_{i+}}, \frac{1}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q + r_\eta w_{i+}}\right),\end{aligned}$$

where $w_{i+} = \sum_{l\neq i} w_{li}$ and $\bar{\eta}_i = \sum_{l\neq i} w_{li}\eta_l/w_{i+}$,

$$\begin{aligned}\pi(\boldsymbol{\theta}_i|\boldsymbol{\theta}_l(l \neq i), \boldsymbol{\beta}, \eta, \boldsymbol{\Sigma}_e, r_\eta, \boldsymbol{y}) \;\propto\; &\prod_j p(y_{ij}|\theta_{ij})\\ &\times\ \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}_i - \eta_i\mathbf{1}_q - \boldsymbol{X}_i\boldsymbol{\beta})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \eta_i\mathbf{1}_q - \boldsymbol{X}_i\boldsymbol{\beta})\right\}.\end{aligned}$$

The full conditionals for r_η, $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\beta}$ are standard, and it is easy to generate samples from them. Also, the conditionals of the $\boldsymbol{\theta}_i$ are log-concave , so that one can use the adaptive rejection sampling [15] to generate samples from these conditionals.

3.3. CAR Prior II

The model considered in the previous subsection is based on the assumption that all the components of $\boldsymbol{\eta}_i$, the ith the spatial effect vector are the same ($i = 1, \ldots, n$). In this subsection, we consider the situation when the vectors $(\eta_{1j}, \ldots, \eta_{nj})$ $(j = 1, \ldots, q)$ are mutually independent, and $\eta_{1j}, \ldots, \eta_{nj}$ have the joint prior

$$\pi(\eta_{1j}, \ldots, \eta_{nj}|r_{\eta_j}) \propto r_{\eta_j}^{n/2} \exp\left\{-\frac{r_{\eta_j}}{2} \sum_{1 \leqslant i < l \leqslant n} w_{il}(\eta_{ij} - \eta_{lj})^2\right\}.$$

Also, we assign the same prior distributions for all the other parameters as in the previous section. Then the joint posterior is given by

$$\begin{aligned}
\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{r}_\eta|\boldsymbol{y}) \;\propto\; & \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
\times\; & |\boldsymbol{\Sigma}_e|^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T \boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})\right\} \\
\times\; & \prod_{j=1}^{q}\left[r_{\eta_j}^{n/2} \exp\left\{-\frac{r_{\eta_j}}{2} \sum_{1 \leqslant i < l \leqslant n} w_{il}(\eta_{ij} - \eta_{lj})^2\right\}\right] \\
\times\; & |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\} \\
\times\; & \prod_{j=1}^{q} r_{\eta_j}^{b_j/2-1} \exp\left(-\frac{a_j r_{\eta_j}}{2}\right),
\end{aligned}$$

where $\boldsymbol{r}_\eta = (r_{\eta_1}, \ldots, r_{\eta_q})^T$. The following theorem is provided to ensure posterior propriety under diffuse prior for $\boldsymbol{\beta}$.

Theorem 3.2. *Assume $a_j > 0, n + b_j > 0, j = 1, \ldots, q,$ and $n > pq + q - \gamma + 1$. Then, if* (3.5) *holds, the joint posterior pdf of the θ_{ij} given $\boldsymbol{y}$ is proper.*

The proof of the theorem is deferred to Appendix B.

The full conditionals required for Gibbs sampling are given by

$$r_{\eta_j}|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{y} \sim \text{Gamma}\left(\frac{1}{2}\left(\sum_{1 \leqslant i < l \leqslant n} w_{il}(\eta_{ij} - \eta_{lj})^2 + a_j\right), \frac{n + b_j}{2}\right),$$

$$\boldsymbol{\Sigma}_e|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, r_\eta, \boldsymbol{y} \sim \text{IW}\left(\boldsymbol{A} + \sum_{i=1}^{n}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T, n + \gamma\right),$$

$$\boldsymbol{\beta}|\boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, r_\eta, \boldsymbol{y} \sim N_p(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta),$$

where

$$\mu_\beta = \left(\sum_{i=1}^n X_i^T \Sigma_e^{-1} X_i\right)^{-1} \sum_{i=1}^n X_i^T \Sigma_e^{-1} (\theta_i - \eta_i)$$
$$\text{and } \Sigma_\beta = \left(\sum_{i=1}^n X_i^T \Sigma_e^{-1} X_i\right)^{-1},$$

$$\eta_i | \theta, \beta, \eta_l(\neq i), \Sigma_e, r_\eta, y \sim N(\mu_\eta, \Sigma_\eta),$$

where

$$\mu_\eta = \left((\Sigma_e^{-1} + w_{i+} R)^{-1} \left(\Sigma_e^{-1}(\theta_i - X_i\beta) + \frac{R w_{i+} \eta_{i+}}{2}\right)\right),$$

$$\Sigma_\eta = \left(\Sigma_e^{-1} + R w_{i+}\right)^{-1},$$

and $R = Diag(r_{\eta_1}, \ldots, r_{\eta_q})$.

$$\begin{aligned}\pi(\theta_i | \theta_j (j \neq i)), \beta, \eta, \Sigma_e, r_\eta, y) \;\propto\; & \prod_j p(y_{ij} | \theta_{ij}) \\ & \times \exp\left\{-\frac{1}{2}(\theta_i - \eta_i - X_i\beta)^T \Sigma_e^{-1} (\theta_i - \eta_i - X_i\beta)\right\}.\end{aligned}$$

3.4. CAR Prior III

The first two spatial models do not induce correlation among the types of crashes directly. In this section, we consider a different Bayesian version of a multivariate CAR model first introduced by Mardia [20]. Carlin and Banerjee [6] considered a special case which is what we consider. Under this framework, conditional on V, the spatial effect is given by $V^{-1} = (D - \alpha W) \otimes \Lambda$. Here $\otimes$ is the Kronecker product, $D = \text{Diag}(m_1, \ldots, m_n)$, m_i being the number of neighbors for the ith region; W is the adjacency matrix; Λ^{-1} describe the relative variability and covariance relationships between the different crashes given the neighboring sites; $\alpha \in (0, 1)$ is the propriety parameter for V to prevent any possible singularities in it. Thus, V^{-1} may be looked upon as the Kronecker product of two partial precision matrices: $D - \alpha W$ for spatial components, and Λ for variation across crashes.

We assume a beta (c, d) prior for α and a Wishart (s, B) prior for Λ. Other prior specifications remain the same as in the previous section. Then the joint posterior is given by

$$\begin{aligned}\pi(\theta, \beta, \eta, \Sigma_e, \alpha, \Lambda | y) \;\propto\; & \prod_{i,j} p(y_{ij} | \theta_{ij}) \\ & \times |\Sigma_e|^{-n/2} \exp\left[-\frac{1}{2}\sum_{i=1}^n (\theta_i - \eta_i - X_i\beta)^T \Sigma_e^{-1} (\theta_i - \eta_i - X_i\beta)\right] \\ & \times |D - \alpha W|^{q/2} |\Lambda|^{n/2} \exp\left(-\frac{1}{2}\eta^T V^{-1} \eta\right) \alpha^{c-1} (1-\alpha)^{d-1}\end{aligned}$$

$$\times \ |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right]$$
$$\times \ |\boldsymbol{\Lambda}|^{(s-q-1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Lambda}\boldsymbol{B})\right].$$

The following theorem is proved to ensure the propriety of the posterior.

Theorem 3.3. *Suppose* $n + s > q, n + \gamma > q$, *and* (3.5) *holds. Then the joint posterior pdf of the* θ_{ij} *given* $\boldsymbol{y}$ *is proper.*

The proof of the theorem is deferred to Appendix C.

In this case, the full conditionals are given by

$$\boldsymbol{\Sigma}_e|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \alpha, \boldsymbol{\Lambda}, \boldsymbol{y} \sim \mathrm{IW}\left(\boldsymbol{A} + \sum_{i=1}^{n}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T, n+\gamma\right),$$

$$\boldsymbol{\beta}|\boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \alpha, \boldsymbol{\Lambda}, \boldsymbol{y} \sim MN(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta),$$

where

$$\boldsymbol{\mu}_\beta = \left(\sum\nolimits_{i=1}^{n} \boldsymbol{X}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{X}_i\right)^{-1}\left(\sum\nolimits_{i=1}^{n} \boldsymbol{X}_i^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i)\right)$$
$$\text{and } \boldsymbol{\Sigma}_\beta = \left(\sum\nolimits_{i=1}^{n} \boldsymbol{X}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{X}_i\right)^{-1},$$

$$\boldsymbol{\eta}_i|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}_{-i}, \boldsymbol{\Sigma}_e, \alpha, \boldsymbol{\Lambda}, \boldsymbol{y} \sim N(\boldsymbol{\mu}_\eta^*, \boldsymbol{\Sigma}_\eta^*),$$

where

$$\boldsymbol{\Sigma}_\eta^* = [\boldsymbol{\Sigma}_e^{-1} + (m_i - \alpha w_{ii})\boldsymbol{\Lambda}]^{-1}$$
$$\text{and } \boldsymbol{\mu}_\eta^* = \boldsymbol{\Sigma}_\eta^*\left[\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \boldsymbol{X}_i\boldsymbol{\beta} + \frac{1}{2}\sum_{j(\neq i)}(\alpha w_{ij})\boldsymbol{\Lambda}\eta_{ij}\right],$$

$$\pi(\alpha|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{\Lambda}, \boldsymbol{y}) \propto |\boldsymbol{D} - \alpha\boldsymbol{W}|^{q/2}\alpha^{c-1}(1-\alpha)^{d-1},$$

$$\boldsymbol{\Lambda}|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \alpha, \boldsymbol{y} \sim \mathrm{Wishart}(\boldsymbol{B}, n+s),$$

$$\pi(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{j(j\neq i)}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{V}, \boldsymbol{y}) \propto \prod_j p(y_{ij}|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{V})$$
$$\times \ \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i\boldsymbol{\beta})\right\}.$$

3.5. CAR prior IV

In this section, we develop a novel correlated CAR (CCAR) priors for spatial random effects where the scale parameters, say, r_{η_j} vary across the different components $j = 1, \ldots, q$. Also, we assume that the logarithms of the scale parameters have a joint multivariate normal distribution. Writing $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_q)^T = (\log r_{\eta_1}, \ldots, \log r_{\eta_q})^T$, we assume that $\boldsymbol{\rho} \sim N_q(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$. Now the spatial models for different crash types are correlated through the scale parameters and we can measure the strength of the correlation as well. The other components of the model remain the same as in the previous section. We first prove the following theorem which provides sufficient conditions for the propriety of the posterior.

Theorem 3.4. *Assume $n + \gamma > 0$ and (3.5) holds. Then the joint posterior pdf of the θ_{ij} given $\mathbf{y}$ is proper.*

The proof of the theorem is deferred to Appendix D.

The full conditionals are given by

$$\begin{aligned}
&\rho_j | \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \rho_l (l \neq j), \mathbf{y} \\
&\quad \propto \exp\left\{ \frac{\rho_j n - \exp(\rho_j) \sum_{1 \leqslant i < l \leqslant n} w_{il} (\eta_{ij} - \eta_{lj})^2 - \boldsymbol{\rho}^T \boldsymbol{\Sigma}_\eta^{-1} \boldsymbol{\rho}}{2} \right\},
\end{aligned}$$

$$\boldsymbol{\Sigma}_e | \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\rho}, \mathbf{y} \sim \mathrm{IW}\left(\boldsymbol{A} + \sum_{i=1}^{n} (\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i \boldsymbol{\beta})(\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i \boldsymbol{\beta})^T, n + \gamma \right),$$

$$\boldsymbol{\beta} | \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{\rho}, \mathbf{y} \sim MN(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta),$$

where $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\Sigma}_\beta$ are the same as in the previous section.

$$\boldsymbol{\eta}_i | \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\eta}_{-i}, \boldsymbol{\Sigma}_e, \boldsymbol{\rho}, \mathbf{y} \sim N(\boldsymbol{\mu}_\eta, \boldsymbol{\Sigma}_\eta),$$

where $\boldsymbol{\mu}_\eta$ and $\boldsymbol{\Sigma}_\eta$ are the same as in the previous section, and $\boldsymbol{R} = \mathrm{Diag}(\exp(\rho_1), \ldots, \exp(\rho_q))$.

$$\begin{aligned}
\pi(\boldsymbol{\theta}_i | \boldsymbol{\theta}_{j(j \neq i)}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Sigma}_e, \boldsymbol{\rho}, \mathbf{y}) \;\propto\; & \prod_j p(y_{ij} | \theta_{ij}) \\
\times\; & \exp\left\{ -\frac{1}{2} (\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i \boldsymbol{\beta})^T \boldsymbol{\Sigma}_e^{-1} (\boldsymbol{\theta}_i - \boldsymbol{\eta}_i - \boldsymbol{X}_i \boldsymbol{\beta}) \right\}.
\end{aligned}$$

3.6. Model choice

Spiegelhalter et al. [31] proposed the deviance information criterion (DIC) to compare complex hierarchical models and DIC reflects the fit and complexity of the models in which the number of parameters are not obviously specified. DIC is a generalization of the well-known Akaike Information Criterion (AIC) and is based on the posterior distribution of the deviance statistic

$$D(\boldsymbol{\theta}) = -2\log(p(\boldsymbol{y}|\boldsymbol{\theta})) + 2\log(f(\boldsymbol{y})),$$

where $p(\boldsymbol{y}|\boldsymbol{\theta})$ is the likelihood function for the observed data vector $\boldsymbol{y}$ given the parameter vector $\boldsymbol{\theta}$, and $f(\boldsymbol{y})$ is some standardizing function of the data alone. For the Poisson model, $f(\boldsymbol{y})$ is usually set as the saturated likelihood, i.e., $f(\boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{\mu} = \boldsymbol{y})$, where $\boldsymbol{\mu}$ is a vector of the statistical means of vector $\boldsymbol{y}$. Note that $\boldsymbol{\theta}$ is the lowest-level parameter vector in a hierarchical model.

The fit and complexity of hierarchical models are shown by $\bar{D}$ the posterior mean of the deviance D and the effective number of parameters p_D respectively. The posterior expectation of the deviance for the fit of a model is defined as

$$\bar{D} = \mathrm{E}_{\theta|y}(D).$$

The number of effective parameters p_D for the complexity of a model is defined as

$$p_D = \mathrm{E}_{\theta|y}(D) - D(\mathrm{E}_{\theta|y}(\theta)) = \bar{D} - D(\bar{\theta}).$$

Finally, a deviance information criterion (DIC) is defined as

$$\mathrm{DIC} = \bar{D} + p_D.$$

DIC can be easily obtained at end of MCMC analysis by monitoring θ and $D(\theta)$ during the simulation and smaller value of DIC represents better fitting of model.

4. Data analysis

Texas Department of Transportation is structured into 25 districts and each contains 6 to 17 counties. The 254 counties are divided among the districts. Since climates and soil conditions in Texas vary considerably, design, maintenance, planning of roads are primarily accomplished locally. The department has maintained the crash data by separating four types of crash based on a location in which a traffic crash occurs:

- *Intersection crash*: A traffic crash which occurs within the limits of an intersection.
- *Intersection-related crash*: A traffic crash which (1) occurs on an approach to or exit from an intersection and (2) result from an activity, behavior or control related to the movement of traffic units through the intersection.
- *Driveway access crash*: A traffic crash (1) occurs a driveway access or (2) involves a road vehicle entering or leaving another roadway by way of on a driveway access.
- *Non-intersection crash*: A traffic crash that is not intersection crash, intersection-related crash, and driveway access crash.

We also consider three covariates in this paper:

- *Wet*: A surrogate variable to represent weather variations.
- *Curve*: A surrogate variable to capture spatial variations in percent of sharp horizontal curves.

- *Obj*: A surrogate variable indicating spatial variations in roadside conditions.

The interaction terms between covariates are involved in the model.

Let Y_{ij} be the number of jth type of reported KAB crashes in county i, $i = 1, \ldots, n(= 251)$, $j = 1, \ldots, q(= 4)$. At the first level of hierarchy, conditional on mean μ_{ij}, Y_{ij} are assumed to be mutually independent and

$$Y_{ij} \sim \text{Poisson}(\mu_{ij}).$$

The mean of the Poisson is modeled

$$\mu_{ij} = v_{ij}\lambda_{ij},$$

where v_{ij} is an offset (in million of vehicle-miles traveled, or MVMT) and λ_{ij} is the KAB crash rate. Since the rate has to be nonnegative, it is structured as

$$\theta_{ij} = \log(\mu_{ij}) = \log(\lambda_{ij}) + \log(v_{ij}) = \log(v_{ij}) + \boldsymbol{x}_i^T \boldsymbol{\beta}_j + \eta_{ij} + e_{ij},$$

where $\boldsymbol{x}_i$ is the ith covariates, $\boldsymbol{\beta}_j$ is the jth regression coefficient, η_{ij} is the spatial random effect for the ith county and jth type of crash, and e_{ij}'s are exchangeable random effects. For simplicity of notation, we can rewrite the expression as

$$\boldsymbol{\theta}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\eta}_i + \boldsymbol{e}_i,$$

where $\boldsymbol{X}_i = \boldsymbol{I}_q \otimes \boldsymbol{x}_i^T$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \ldots, \boldsymbol{\beta}_q^T)^T$, and $\boldsymbol{x}_i^T$ is a $p \times 1$ row vector for county i. $\boldsymbol{\theta}_i$'s can be expressed as a $N(= n \times q) \times 1$ column matrix $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_n^T)^T$ and the model is given by

$$\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{e},$$

where $\boldsymbol{X}^T = (\boldsymbol{X}_1^T, \ldots, \boldsymbol{X}_n^T)^T$, $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^T, \ldots, \boldsymbol{\eta}_n^T)^T$, and $\boldsymbol{e} = (\boldsymbol{e}_1^T, \ldots, \boldsymbol{e}_n^T)^T$. In this study, known weights in the CAR models are given by $w_{ij} = 1/d_{ij}$, where d_{ij} is the Great Circle distance between the centroid of county i and j.

Prior distributions of all parameters in the model are specified as those in previous section and four types of spatial priors for multivariate models are considered in this analysis. Posterior propriety for each proposed spatial prior is ensured through the theorems with the integrability of the likelihood. Let $\theta_{ij} = \log(\mu_{ij})$ and $\psi(\theta_{ij}) = \exp(\theta_{ij})$. Then, the integral in the theorems is replaced by

$$\int_0^{\infty} \xi_{ij}^{y_{ij}-1} \exp(-\xi_{ij}) d\xi_{ij} < \infty,$$

which hold when $y_{ij} = 1, 2, \ldots$. Therefore, all proposed theorems hold for Poisson models with additional requirement $y_{ij} = 1, 2, \ldots$.

As mentioned earlier, posterior inference is carried out by MCMC and Gibbs sampler is implemented for most of the parameters whose full conditionals are available in closed form. The rest of them are sampled using Metropolis-Hastings algorithm. It is only necessary to replace the one-parameter exponential family density by the Poisson density in the full conditionals and note that sampling step for $\boldsymbol{\theta}$ depends on the likelihood function. For specification of the hyperparameters, we use gamma priors with $a_j = 0.2$ and $b_j = 0.2$ (mean=1, variance=10) are placed on r_{η_j}, inverse Wishart prior $\gamma = 4 = q$ and $\boldsymbol{A}$ with

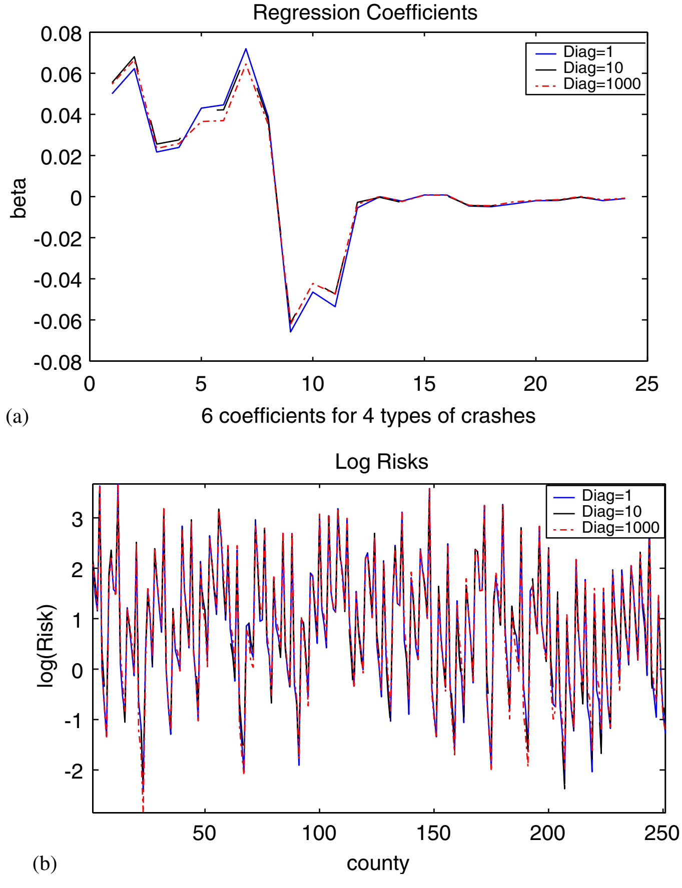

Fig. 1. Sensitivity analysis for the choice of hyper-parameters of the inverse Wishart distribution using the estimates of the regression and risk parameters.

diagonals 1000 and off-diagonals 0.005 is specified for the prior of $\boldsymbol{\Sigma}_e$, and Beta(18, 2) prior distribution is assigned to α.

We have used all of our models to fit the data and performed a model comparison to select the best one using the criterion DIC values which are presented in Table 1. In the table, for model 3, "single" indicates that α_j's are common over different types of crashes and "multiple" means that α_j's vary over different types of crashes. It is clear that correlated CAR and multivariate CAR with unknown α is performing well. We will present other results based on the correlated CAR model, Model 4.

We also compared our model with the independent Poisson regression models (with the spirit of Roeleven et al. [29]) without spatial and within response correlations. As expected, the standard deviations of all the estimated parameters are much larger and number of significant variables are less than the results obtained from any of the Bayesian models. It identifies some of the important main effects like Wet(Int), curve (Dri), obj(int) as

Table 1
DIC and effective number of parameters p_D for competing models

	p_D	DIC
Model 1	516.8	1480.3
Model 2	481.1	1399.6
Model 3 (with fixed $\alpha = 1$)	462.8	1391.0
Model 3 (with single, unknown α)	459.6755	1391.2
Model 3 (with multiple, unknown α)	457.5247	1385.9
Model 4	455.0	1385.2

Model 1 = Model with same Spatial effect, Model 2 = Model with independent CAR, Model 3 = Model with multivariate CAR and different choices of $\boldsymbol{\alpha}$, and Model 4 = Model with correlated CAR.

Table 2
Posterior summaries for regression coefficients for Model 4

Regression coefficient	2.5%	50%	97.5%
Wet(Int)*	0.002368	0.030970	0.059490
Wet(Int-Re)*	0.004350	0.040030	0.074600
Wet(Dri)	−0.049490	−0.001993	0.04473
Wet(Non)	−0.001146	0.013580	0.027570
Curve (Int)*	0.011050	0.022120	0.035830
Curve (Int-Re)*	0.004510	0.018090	0.035070
Curve (Dri)*	0.023950	0.03993	0.058360
Curve (Non)*	0.018280	0.026390	0.034630
Obj (Int)*	−0.094790	−0.080800	−0.067920
Obj (Int-Re)*	−0.080970	−0.063400	−0.046900
Obj (Dri)*	−0.095990	−0.074590	−0.053610
Obj (Non)*	−0.019450	−0.011260	−0.003554
Wet*Curve (Int)*	−0.014910	−0.011450	−0.007936
Wet*Curve (Int-Re)*	−0.020610	−0.016320	−0.011930
Wet*Curve (Dri)*	−0.022040	−0.016220	−0.01019
Wet*Curve (Non)*	−0.006721	−0.005007	−0.003294
Wet*Obj (Int)*	−0.023780	−0.019280	−0.014720
Wet*Obj (Int-Re)*	−0.028210	−0.022510	−0.016770
Wet*Obj (Dri)*	−0.031700	−0.024480	−0.017030
Wet*Obj (Non)*	−0.010770	−0.008458	−0.006040
Curve*Obj (Int)	−0.002536	−0.000774	0.000840
Curve*Obj (Int-Re)	−0.001353	0.000743	0.002670
Curve*Obj (Dri)	−0.003448	−0.000680	0.001893
Curve*Obj (Non)	−0.001201	−0.000343	0.000535

* Indicates that the HPD set does not cover 0 and it is significant one.

nonsignificant. Also the DIC corresponding to this model is 3069.2, much higher than any of the proposed Bayesian models.

Sensitivity analyses with respect to several prior specifications have been conducted for model 4 and the results appeared to be robust. We illustrate one of these sensitivity analyses for the prior of $\boldsymbol{\Sigma}_e$. For the degree of freedom (DF) for inverse Wishart hyperprior, we

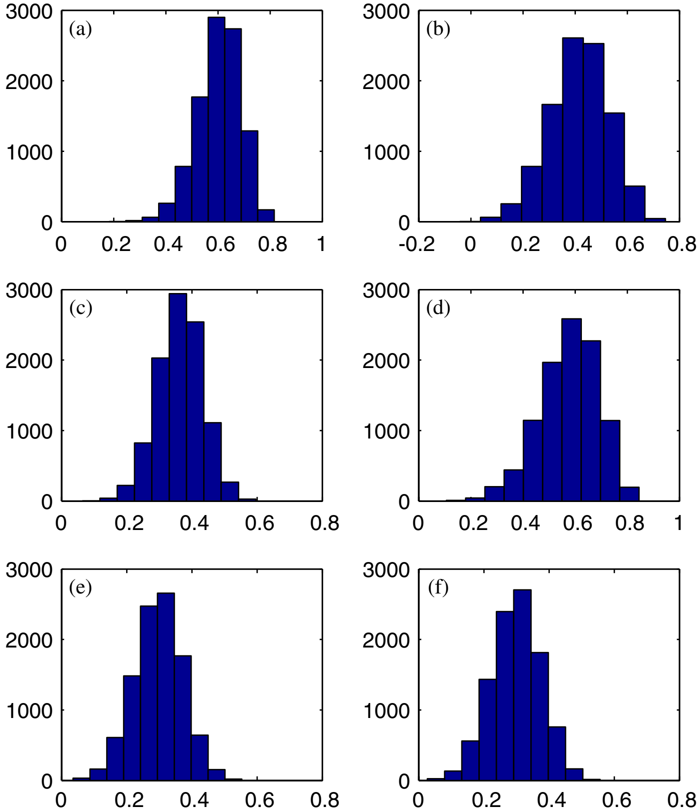

Fig. 2. Plot of the posterior distributions of the correlation coefficients between the responses. Correlation coefficients corresponding to (a) intersection and intersection-related, (b) intersection vs. driveway access, (c) intersection vs. nonintersection, (d) intersection-related vs. driveway access, (e) intersection-related vs. nonintersection, (f) driveway access vs. nonintersection.

adapted the least informative choice corresponding to $\gamma = 4$, the number of variables. The scale matrix $\boldsymbol{A}$ is specified to vary the diagonals. The off-diagonals of the scale matrix are given by 0.005. Three values of diagonals are considered, 1, 10, and 1000. The results are shown in Fig. 1. The first plot is for estimated regression coefficients and the second plot is for estimated log crash risk rate of a type of crash, Intersection. It is clear from the plots that the estimates are not that sensitive to the change of prior specifications.

We present the posterior summaries of the regression parameters corresponding to the covariates and their interactions for each of the responses in Table 2.

From Table 2 it is clear that the covariates wet, curve, and obj has significant effect on intersection and intersection-related crashes. Also the interaction between wet–curve and wet–obj is significant. For driveway and nonintersection crash, two main covariates, curve and obj, are only significant and similarly two significant interactions, wet–curve and wet–obj, are given. An interaction, curve–obj, is not significant over all types of crashes.

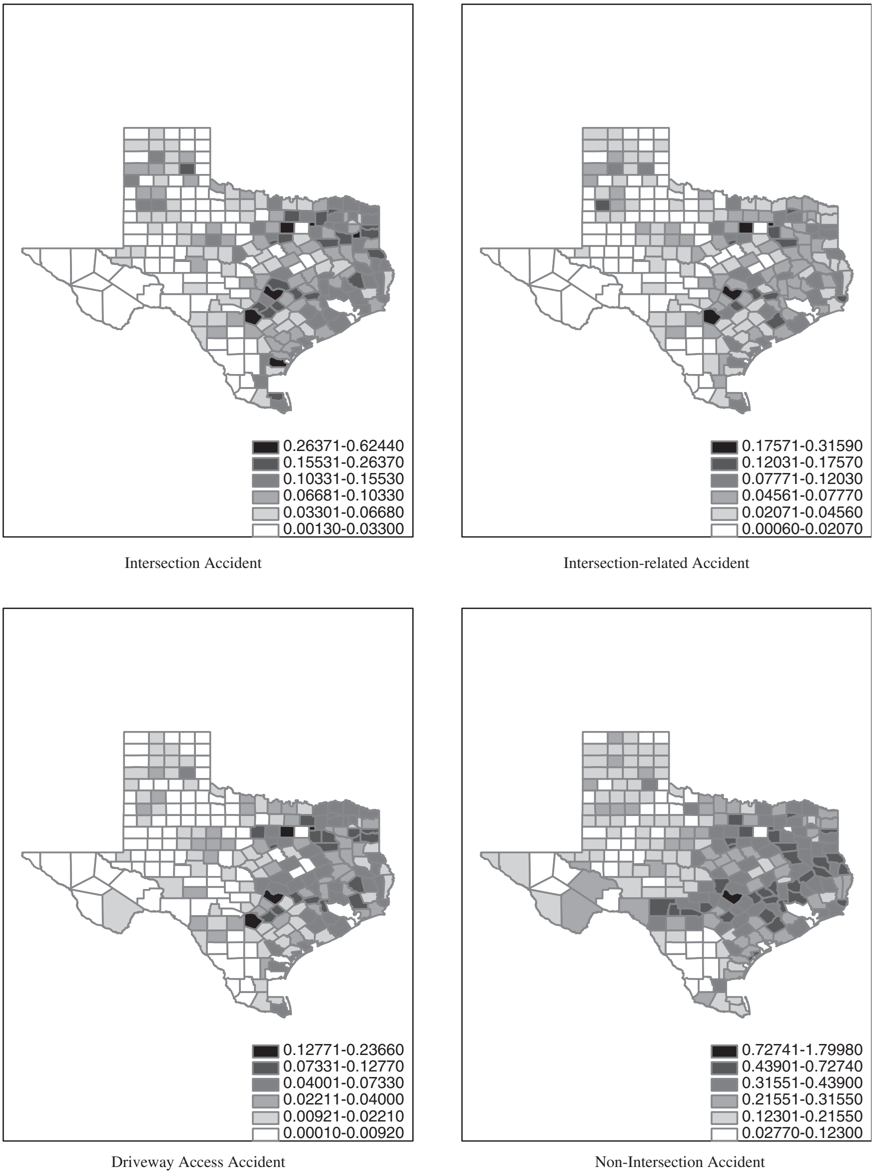

Fig. 3. Predicted map for different types of crash.

Altogether the covariate wet is not always significant over different crashes and curve–obj is not significant over all crashes, the other two covariates curve and obj and two interactions, wet–curve and wet–obj, have significant effect.

We have also plotted the posterior distributions of the correlations of the spatial scale parameters $\boldsymbol{\rho}$ in Fig. 2. All of the parameters have significant positive correlations which is expected. Higher correlation has been seen among intersection, intersection-related and driveway crashes. All these responses have lower but significant positive correlation with nonintersection crash.

Form the MCMC simulation results we can obtain the posterior distribution of the risk parameters λ. The predicted risk maps based on the posterior mean estimates of λ at different regions has been presented in Fig. 3. It is clear that east Texas has higher crash risk than the west. By further investigation of the maps, we found the high-risk sites for each type of crashes are rural areas near to the big cities like Dallas, Austin, San Antonio, and Fort Worth. Limited by the rolling terrain in the eastern counties, roadways in rural area tend to have less driver-friendly characteristics with, e.g., more horizontal and vertical curves, restricted sight-distance, and less forgiving roadside development (e.g. tree closer to the travelway and steeper side-slopes). In addition, with more and larger urbanized areas in the ease, rural roads tend to have higher roadside development scores, higher access density, and narrow lanes and/or shoulder [11] which may be the reason of high-risk accidents at those areas.

5. Summary and conclusion

In this paper, we have provided several spatial priors for Bayesian multivariate hierarchical models and sufficient conditions to ensure posterior propriety under noninformative prior for all the proposed models. These models are useful tools for spatially correlated multivariate data in order to analyze all response variables simultaneously. As future work, it is an interesting topic to apply statistical ranking criteria to identify sites on a road network for further engineering inspection and safety improvement. A future study in transportation application is to explore some of the issues raised regarding ranking methodology in the light of recent statistical developments in spatio-temporal generalized linear mixed models.

Acknowledgments

This research is partially supported by NSF DMS 0327713 and Bureau of Transportation Statistics, US Department of Transportation, via transportation statistics research grant number DTTS-00-G-B005-TX and by Southwest Regional University Transportation Center FY 2002 Research Program (Project Number 167424). The authors are, however, solely responsible for the contents and views expressed in this paper.

Appendix A

Proof of Theorem 1. Let $z_i = \eta_i - \eta_n$ $(i = 1, \ldots, n-1)$ and $z_n = 0$. Then, the transformed posterior is

$$\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{z}, \eta_n, \boldsymbol{\Sigma}_e, r_\eta | \boldsymbol{y}) \propto \prod_{i,j} p(y_{ij}|\theta_{ij})$$

$$
\begin{aligned}
&\times \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\theta}_i - \boldsymbol{X}_i\boldsymbol{\beta} - (z_i+\eta_n)\mathbf{1}_q)^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i - \boldsymbol{X}_i\boldsymbol{\beta} - (z_i+\eta_n)\mathbf{1}_q)\right\}\\
&\times \exp\left\{-\frac{r_\eta}{2}\sum_{1\leqslant i<l\leqslant n} w_{il}(z_i-z_l)^2\right\} r_\eta^{n/2}\\
&\times|\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}\\
&\times r_\eta^{(b/2)-1}\exp\left(-\frac{ar_\eta}{2}\right),
\end{aligned}
$$

where $\boldsymbol{z} = (z_1,\dots,z_{n-1})^T$. Writing $\boldsymbol{c}_i = \boldsymbol{\theta}_i - \boldsymbol{X}_i\boldsymbol{\beta} - z_i\mathbf{1}_q$, $\bar{\boldsymbol{c}} = n^{-1}\sum_{i=1}^n \boldsymbol{c}_i$, one has

$$
\begin{aligned}
\sum_{i=1}^{n}(\eta_n\mathbf{1}_q-\boldsymbol{c}_i)^T\boldsymbol{\Sigma}_e^{-1}(\eta_n\mathbf{1}_q-\boldsymbol{c}_i) &= n\eta_n^2(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q) - 2n\eta_n(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\bar{\mathbf{c}}) + \sum_{i=1}^n \boldsymbol{c}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{c}_i\\
&= n(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q)\left(\eta_n - \frac{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\bar{\boldsymbol{c}}}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q}\right)^2 + \sum_{i=1}^n \boldsymbol{c}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{c}_i\\
&\quad - \frac{n(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\bar{\boldsymbol{c}})^2}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q}.
\end{aligned}
$$

Now integrating with respect to η_n,

$$
\begin{aligned}
\pi(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{z},\boldsymbol{\Sigma}_e,r_\eta|\boldsymbol{y}) \;\propto\; &\prod_{i,j} p(y_{ij}|\theta_{ij})\\
\times\;& (\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q)^{-1/2}\exp\left\{-\frac{1}{2}\left(\sum_{i=1}^n \boldsymbol{c}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{c}_i - \frac{n(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\bar{\boldsymbol{c}})^2}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q}\right)\right\}\\
\times\;& \exp\left[-\frac{r_\eta}{2}\left\{a + \sum_{1\leqslant i<l\leqslant n} w_{il}(z_i-z_l)^2\right\}\right] r_\eta^{((n+b)/2)-1}\\
\times\;& |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}.
\end{aligned}
$$

Next, by the inequality

$$
\begin{aligned}
&\sum_{i=1}^n \boldsymbol{c}_i^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{c}_i - \frac{n(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\bar{\boldsymbol{c}})^2}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q}\\
&\quad = \sum_{i=1}^n(\boldsymbol{c}_i-\bar{\boldsymbol{c}})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{c}_i-\bar{\boldsymbol{c}}) + n\left[\bar{\boldsymbol{c}}^T\boldsymbol{\Sigma}_e^{-1}\bar{\boldsymbol{c}} - \frac{(\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\bar{\boldsymbol{c}})^2}{\mathbf{1}_q^T\boldsymbol{\Sigma}_e^{-1}\mathbf{1}_q}\right]\\
&\quad \geqslant \sum_{i=1}^n(\boldsymbol{c}_i-\bar{\boldsymbol{c}})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{c}_i-\bar{\boldsymbol{c}})\\
&\quad = \sum_{i=1}^n[\boldsymbol{g}_i - (\boldsymbol{X}_i-\bar{\boldsymbol{X}})\boldsymbol{\beta}]^T\boldsymbol{\Sigma}_e^{-1}[\boldsymbol{g}_i - (\boldsymbol{X}_i-\bar{\boldsymbol{X}})\boldsymbol{\beta}],
\end{aligned}
$$

where $\boldsymbol{g}_i = \boldsymbol{\theta}_i - \bar{\boldsymbol{\theta}} - (z_i - \bar{z})\mathbf{1}_q \ (i = 1, \ldots, n)$, one gets

$$\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{z}, \boldsymbol{\Sigma}_e, r_\eta|\boldsymbol{y}) \\
&\quad \leqslant K \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
&\quad \times (\mathbf{1}_q^T \boldsymbol{\Sigma}_e^{-1} \mathbf{1}_q)^{-1/2} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\{\boldsymbol{g}_i - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}\}^T \boldsymbol{\Sigma}_e^{-1}\{\boldsymbol{g}_i - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}\}\right] \\
&\quad \times \exp\left[-\frac{r_\eta}{2}\left\{a + \sum_{1\leqslant i<l\leqslant n} w_{il}(z_i - z_l)^2\right\}\right] r_\eta^{((n+b)/2)-1} \\
&\quad \times |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\},
\end{aligned}$$

where in the above and in what follows, $K\,(>0)$ is a generic constant. Next integrating with respect to $\boldsymbol{\beta}$,

$$\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\Sigma}_e, r_\eta|\boldsymbol{y}) \\
&\quad \leqslant K \prod_{i,j} p(y_{ij}|\theta_{ij})(\mathbf{1}_q^T \boldsymbol{\Sigma}_e^{-1} \mathbf{1}_q)^{-1/2}|\boldsymbol{\Sigma}_*|^{1/2} \\
&\quad \times \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\{\boldsymbol{g}_i - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}_*\}^T \boldsymbol{\Sigma}_e^{-1}\{\boldsymbol{g}_i - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}_*\}\right] \\
&\quad \times \exp\left[-\frac{r_\eta}{2}\left\{a + \sum_{1\leqslant i<l\leqslant n} w_{il}(z_i - z_l)^2\right\}\right] r_\eta^{((n+b)/2)-1} \\
&\quad \times |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\},
\end{aligned}$$

where $\boldsymbol{\Sigma}_*^{-1} = \sum_{i=1}^{n}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})$ and $\boldsymbol{\beta}_* = \boldsymbol{\Sigma}_*^{-1}[\sum_{i=1}^{n}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{g}_i]$. The above is bounded above by

$$\begin{aligned}
\pi(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\Sigma}_e, r_\eta|\boldsymbol{y}) \ \leqslant\ & K \prod_{i,j} p(y_{ij}|\theta_{ij})(\mathbf{1}_q^T \boldsymbol{\Sigma}_e^{-1} \mathbf{1}_q)^{-1/2}|\boldsymbol{\Sigma}_*|^{1/2} \\
& \times\ \exp\left[-\frac{r_\eta}{2}\left\{a + \sum_{1\leqslant i<l\leqslant n} w_{il}(z_i - z_l)^2\right\}\right] r_\eta^{((n+b)/2)-1} \\
& \times\ |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}.
\end{aligned}$$

Next observe that $\sum\sum_{1\leqslant i<l\leqslant n} w_{il}(z_i - z_l)^2 = \boldsymbol{z}^T\boldsymbol{W}^*\boldsymbol{z}$, where

$$\boldsymbol{W}^* = \begin{bmatrix} \sum_{l=1}^{n} w_{1,l} & -w_{1,2} & \cdots & -w_{1,n-1} \\ -w_{2,1} & \sum_{l=1}^{n} w_{2,l} & \cdots & -w_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ -w_{n-1,1} & w_{n-1,2} & \cdots & \sum_{l=1}^{n} w_{n-1,l} \end{bmatrix}.$$

Hence, integrating first with respect to z and then with respect to r_η, one gets

$$\pi(\boldsymbol{\theta}, \boldsymbol{\Sigma}_e|\boldsymbol{y}) \leqslant K \prod_{i,j} p(y_{ij}|\theta_{ij})(\mathbf{1}_q^T \boldsymbol{\Sigma}_e^{-1} \mathbf{1}_q)^{-1/2} |\boldsymbol{\Sigma}_*|^{1/2} \times |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}.$$

Let $\zeta_1 \leqslant \zeta_2 \leqslant \cdots \zeta_q$ denote the eigenvalues of $\boldsymbol{\Sigma}_e$. Then, ζ_q^{-1} is the smallest eigenvalue of $\boldsymbol{\Sigma}_e^{-1}$. Now by the inequalities $(\mathbf{1}_q^T \boldsymbol{\Sigma}_e^{-1} \mathbf{1}_q)^{-1/2} \leqslant \zeta_q^{1/2} q^{-1/2}$ and $|\boldsymbol{\Sigma}_*|^{1/2} = |\sum_{i=1}^n (\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T \boldsymbol{\Sigma}_e^{-1} (\boldsymbol{X}_i - \bar{\boldsymbol{X}})|^{-1/2} \leqslant \zeta_q^{pq/2} |\sum_{i=1}^n (\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T (\boldsymbol{X}_i - \bar{\boldsymbol{X}})|^{-1/2}$,

$$\begin{aligned}
&\mathbf{1}_q^T \boldsymbol{\Sigma}_e^{-1} \mathbf{1}_q)^{-1/2} |\boldsymbol{\Sigma}_\beta|^{1/2} |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q+1)/2} \\
&\leqslant K \zeta_q^{(pq+1)/2} \left(\prod_{j=1}^q \zeta_j\right)^{-(n+q+\gamma+1)/2} \\
&= \left(\prod_{j=1}^{q-1} \zeta_j^{-1}\right)^{(n+q+\gamma+1)/2} \qquad (\zeta_q^{-1})^{(n+q+\gamma-pq)/2} \\
&= K \left(\prod_{j=1}^{q} \zeta_j^{-1}\right)^{(n+q+\gamma-pq)/2} \quad \left(\prod_{j=1}^{q-1} \zeta_j^{-1}\right)^{(pq+1)/2} \\
&\leqslant K |\boldsymbol{\Sigma}_e|^{-(n+q+\gamma-pq)/2} \left(\frac{1}{q-1} \sum_{j=1}^{q-1} \zeta_j^{-1}\right)^{(q-1)(pq+1)/2} \\
&\leqslant K |\boldsymbol{\Sigma}_e|^{-(n+q+\gamma-pq)/2} \left(\sum_{j=1}^{q} \zeta_j^{-1}\right)^{(q-1)(pq+1)/2} \\
&= K |\boldsymbol{\Sigma}_e|^{-(n+q+\gamma-pq)/2} [\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1})]^{(q-1)(pq+1)/2}.
\end{aligned}$$

Since $n > pq + q - \gamma + 1$, but for a constant multiple, $|\boldsymbol{\Sigma}_e|^{-(n+q+\gamma-pq)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}$ represents the pdf of an inverse Wishart matrix ($\boldsymbol{\Sigma}_e$ being inverse Wishart) with $n+\gamma-pq-1$ degrees of freedom (see [1, p. 268]). Hence,

$$\int [\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1})]^{(q-1)(pq+1)/2} |\boldsymbol{\Sigma}_e|^{-(n+q+\gamma-pq)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right] d\boldsymbol{\Sigma}_e < \infty$$

since the integral is the positive moment of the trace of a Wishart matrix ($\boldsymbol{\Sigma}_e^{-1}$ being Wishart) times a constant. Thus,

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \leqslant K \prod_{i,j} p(y_{ij}|\theta_{ij}).$$

The propriety of the posterior now follow from the assumption that $\int p(y_{ij}|\theta) d\theta < \infty$ for all i and j. $\square$

Appendix B

Proof of Theorem 2. By the transformation $z_i = \boldsymbol{\eta}_i - \boldsymbol{\eta}_n$ $(i = 1, \ldots, n-1)$ and $z_n = \mathbf{0}$, writing $z^T = (z_1^T, \ldots, z_{n-1}^T)$, the joint posterior is given by

$$
\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, z, \boldsymbol{\eta}_n, \boldsymbol{\Sigma}_e, \boldsymbol{r}_\eta | \boldsymbol{y}) \\
&\quad \propto \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
&\quad \times |\boldsymbol{\Sigma}_e|^{-n/2} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{\theta}_i - z_i - \boldsymbol{\eta}_n - X_i \boldsymbol{\beta})^T \boldsymbol{\Sigma}_e^{-1} (\boldsymbol{\theta}_i - z_i - \boldsymbol{\eta}_n - X_i \boldsymbol{\beta}) \right\} \\
&\quad \times |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left\{ -\frac{1}{2} \mathrm{tr} \boldsymbol{\Sigma}_e^{-1} A \right\} \\
&\quad \times \prod_{j=1}^{q} r_{\eta_j}^{n/2} \exp\left\{ -\frac{r_{\eta_j}}{2} \sum_{1 \leqslant i < l \leqslant n} w_{il} (z_{ij} - z_{lj})^2 \right\} \\
&\quad \times \prod_{j=1}^{q} r_{\eta_j}^{(b_j/2)-1} \exp\left(-\frac{a_j r_{\eta_j}}{2} \right).
\end{aligned}
$$

Let $\boldsymbol{c}_i = \boldsymbol{\theta}_i - z_i - X_i \boldsymbol{\beta}$, $\bar{\boldsymbol{c}} = n^{-1} \sum_{i=1}^{n} \boldsymbol{c}_i$. Then integrating with respect to $\boldsymbol{\eta}_n$,

$$
\begin{aligned}
\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, z, \boldsymbol{\Sigma}_e, \boldsymbol{r}_\eta | \boldsymbol{y}) \propto\ & \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
\times\ & |\boldsymbol{\Sigma}_e|^{-(n-1)/2} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{c}_i - \bar{\boldsymbol{c}})^T \boldsymbol{\Sigma}_e^{-1} (\boldsymbol{c}_i - \bar{\boldsymbol{c}}) \right] \\
\times\ & \prod_{j=1}^{q} r_{\eta_j}^{n/2} \exp\left\{ -\frac{r_{\eta_j}}{2} \sum_{1 \leqslant i < l \leqslant n} w_{il} (z_{ij} - z_{lj})^2 \right\} \\
\times\ & |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left\{ -\frac{1}{2} \mathrm{tr} \boldsymbol{\Sigma}_e^{-1} A \right\} \\
\times\ & \prod_{j=1}^{q} r_{\eta_j}^{(b_j/2)-1} \exp\left(-\frac{a_j r_{\eta_j}}{2} \right).
\end{aligned}
$$

Now writing $\boldsymbol{c}_i - \bar{\boldsymbol{c}} = \boldsymbol{g}_i - \bar{\boldsymbol{g}} - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}$, where $\boldsymbol{g}_i = \boldsymbol{\theta}_i - z_i$ and $\bar{\boldsymbol{X}} = n^{-1} \sum_{i=1}^{n} \boldsymbol{X}_i$, integration with respect to $\boldsymbol{\beta}$ yields

$$
\begin{aligned}
&\pi(\boldsymbol{\theta}, z, \boldsymbol{\Sigma}_e, \boldsymbol{r}_\eta | \boldsymbol{y}) \\
&\quad \propto \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
&\quad \times \exp\left[-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{g}_i - \bar{\boldsymbol{g}} - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}_*)^T \boldsymbol{\Sigma}_*^{-1} (\boldsymbol{g}_i - \bar{\boldsymbol{g}} - (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{\beta}_*) \right] |\boldsymbol{\Sigma}_*|^{1/2}
\end{aligned}
$$

$$\times|\boldsymbol{\Sigma}_e|^{-(\gamma+n+q-p)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}A\right\}$$
$$\times\prod_{j=1}^{q}\left[r_{\eta_j}^{((n+b_j)/2)-1}\exp\left\{-\frac{r_{\eta_j}}{2}\left(a_j+\sum_{1\leqslant i<l\leqslant n}w_{il}(z_{ij}-z_{lj})^2\right)\right\}\right],$$

where as before $\boldsymbol{\Sigma}_*^{-1}=\sum_{i=1}^n(\boldsymbol{X}_i-\bar{\boldsymbol{X}})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{X}_i-\bar{\boldsymbol{X}})$ and $\boldsymbol{\beta}_*=\boldsymbol{\Sigma}_*^{-1}[\sum_{i=1}^n(\boldsymbol{X}_i-\bar{\boldsymbol{X}})^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{g}_i]$. Now, writing $K(>0)$ once again for a generic constant,

$$\pi(\boldsymbol{\theta},\boldsymbol{z},\boldsymbol{\Sigma}_e,\boldsymbol{r}_\eta|\boldsymbol{y})\leqslant K\prod_{i,j}p(y_{ij}|\theta_{ij})|\boldsymbol{\Sigma}_*|^{1/2}|\boldsymbol{\Sigma}_e|^{-(\gamma+n+q)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}$$
$$\times\prod_{j=1}^{q}\left[r_{\eta_j}^{((n+b_j)/2)-1}\exp\left\{-\frac{r_{\eta_j}}{2}\left(a_j+\sum_{1\leqslant i<l\leqslant n}w_{il}(z_{ij}-z_{lj})^2\right)\right\}\right].$$

Next, integrating first with respect to $\boldsymbol{z}$ and then with respect to $\boldsymbol{r}_\eta$, one gets

$$\pi(\boldsymbol{\theta},\boldsymbol{\Sigma}_e|\boldsymbol{y})\leqslant K\prod_{i,j}p(y_{ij}|\theta_{ij})|\boldsymbol{\Sigma}_*|^{1/2}|\boldsymbol{\Sigma}_e|^{-(\gamma+n+q)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}A\right\}.$$

Arguing as in the previous section, we get

$$\pi(\boldsymbol{\theta},\boldsymbol{\Sigma}_e|\boldsymbol{y})\leqslant K\prod_{i,j}p(y_{ij}|\theta_{ij})[\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1})]^{(q-1)pq/2}|\boldsymbol{\Sigma}_e|^{-(n+q+\gamma-pq)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}.$$

This leads to

$$\pi(\boldsymbol{\theta}|\boldsymbol{y})\leqslant K\prod_{i,j}p(y_{ij}|\theta_{ij}),$$

after integration with respect to $\boldsymbol{\Sigma}_e$. The result follows now from the condition of the theorem. □

Appendix C

Proof of Theorem 3. Let $\boldsymbol{g}_i=\boldsymbol{\theta}_i-\boldsymbol{X}_i\boldsymbol{\beta}$ $(i=1,\ldots,n)$, and $\boldsymbol{g}^T=(\boldsymbol{g}_1^T,\ldots,\boldsymbol{g}_n^T)$. Now we write

$$\begin{aligned}&\sum_{i=1}^n(\boldsymbol{\theta}_i-\boldsymbol{X}_i\boldsymbol{\beta}-\boldsymbol{\eta}_i)^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i-\boldsymbol{X}_i\boldsymbol{\beta}-\boldsymbol{\eta}_i)+\boldsymbol{\eta}^T\boldsymbol{V}^{-1}\boldsymbol{\eta}\\
&\quad=\boldsymbol{\eta}^T(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1}+\boldsymbol{V}^{-1})\boldsymbol{\eta}-2\boldsymbol{g}^T(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})\boldsymbol{\eta}+\boldsymbol{g}^T(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})\boldsymbol{g}\\
&\quad=[\boldsymbol{\eta}-(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1}+\boldsymbol{V}^{-1})^{-1}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})\boldsymbol{g}]^T(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1}+\boldsymbol{V}^{-1})\\
&\quad\times[\boldsymbol{\eta}-(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1}+\boldsymbol{V}^{-1})^{-1}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})\boldsymbol{g}]\\
&\qquad+\boldsymbol{g}^T[(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})-(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1}+\boldsymbol{V}^{-1})^{-1}\\
&\quad\times(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1})]\boldsymbol{g}.\end{aligned}$$

Noting that $(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1}) - (\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1})(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1} + \boldsymbol{V}^{-1})^{-1}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1}) = [(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1})^{-1} + \boldsymbol{V}]^{-1} = \boldsymbol{C}$, say, integration with respect to $\boldsymbol{\eta}$ yields

$$
\begin{aligned}
\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e, \alpha, \boldsymbol{\Lambda}|\boldsymbol{y}) \;\propto\; & \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
& \times\; |\boldsymbol{\Sigma}_e|^{-n/2} |\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1} + \boldsymbol{V}^{-1}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{g}^T \boldsymbol{C}\boldsymbol{g}\right) \\
& \times\; |\boldsymbol{D} - \alpha \boldsymbol{W}|^{q/2} |\boldsymbol{\Lambda}|^{n/2} \alpha^{c-1} (1-\alpha)^{d-1} \\
& \times\; |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right] \\
& \times\; |\boldsymbol{\Lambda}|^{(s-q-1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Lambda}\boldsymbol{B})\right].
\end{aligned}
$$

Next writing $\boldsymbol{X}^T = (\boldsymbol{X}_1^T, \ldots, \boldsymbol{X}_n^T)$, and noting that rank $(\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X}) = \mathrm{rank}(\boldsymbol{X}) = pq$, $\boldsymbol{\theta}^T = (\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_n^T)$,

$$
\begin{aligned}
\boldsymbol{g}^T\boldsymbol{C}\boldsymbol{g} &= \boldsymbol{\beta}^T(\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X})\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{\theta}^T\boldsymbol{C}\boldsymbol{\theta} \\
&= [\boldsymbol{\beta} - (\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{\theta}]^T(\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X})[\boldsymbol{\beta} - (\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{\theta}] \\
&\quad + \boldsymbol{\theta}^T[\boldsymbol{C} - \boldsymbol{C}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{C}]\boldsymbol{\theta}.
\end{aligned}
$$

Hence, integrating with respect to $\boldsymbol{\beta}$, one gets

$$
\begin{aligned}
\pi(\boldsymbol{\theta}, \boldsymbol{\Sigma}_e, \alpha, \boldsymbol{\Lambda}|\boldsymbol{y}) \;\propto\; & \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
& \times\; |\boldsymbol{\Sigma}_e|^{-n/2} |\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1} + \boldsymbol{V}^{-1}|^{-1/2} |\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X}|^{-1/2} \\
& \times\; \exp\left[-\frac{1}{2}\boldsymbol{\theta}^T\{\boldsymbol{C} - \boldsymbol{C}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{C}\}\boldsymbol{\theta}\right] \\
& \times\; |\boldsymbol{D} - \alpha \boldsymbol{W}|^{q/2} |\boldsymbol{\Lambda}|^{n/2} \alpha^{c-1} (1-\alpha)^{d-1} \\
& \times\; |\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right] \\
& \times\; |\boldsymbol{\Lambda}|^{(s-q-1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Lambda}\boldsymbol{B})\right].
\end{aligned}
$$

Hence, writing $K(>0)$ for a generic constant which does not depend on any unknown parameters,

$$
\begin{aligned}
\pi(\boldsymbol{\theta}, \boldsymbol{\Sigma}_e, \alpha, \boldsymbol{\Lambda}|\boldsymbol{y}) \;\leqslant\; & K \prod_{i,j} p(y_{ij}|\theta_{ij}) \\
& \times\; |\boldsymbol{\Sigma}_e|^{-(n+\gamma+q+1)/2} |\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_e^{-1} + \boldsymbol{V}^{-1}|^{-1/2} |\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X}|^{-1/2}
\end{aligned}
$$

$$\times\ |\boldsymbol{D}-\alpha \boldsymbol{W}|^{q/2}\alpha^{c-1}(1-\alpha)^{d-1}\ \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right]$$
$$\times\ |\boldsymbol{\Lambda}|^{(n+s-q-1)/2}\ \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Lambda}\boldsymbol{B})\right].$$

But, $|\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}_e^{-1}+\boldsymbol{V}^{-1}|^{-1/2}\leqslant|\boldsymbol{V}^{-1}|^{-1/2}=|\boldsymbol{V}|^{1/2}=|\boldsymbol{D}-\alpha\boldsymbol{W}|^{-q/2}|\boldsymbol{\Lambda}|^{-n/2}$. Also, $|\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X}|^{-1/2}\leqslant|\boldsymbol{X}^T\boldsymbol{V}\boldsymbol{X}|^{-1/2}$. Thus,

$$\begin{aligned}\pi(\boldsymbol{\theta},\boldsymbol{\Sigma}_e,\alpha,\boldsymbol{\Lambda}|\boldsymbol{y}) \ \leqslant\ & K\prod_{i,j}p(y_{ij}|\theta_{ij})\\ &\times\ |\boldsymbol{\Sigma}_e|^{-(n+\gamma+q+1)/2}|\boldsymbol{X}^T\boldsymbol{C}\boldsymbol{X}|^{-1/2}\\ &\times\ \alpha^{c-1}(1-\alpha)^{d-1}\ \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right]\\ &\times\ |\boldsymbol{\Lambda}|^{(s-q-1)/2}\ \exp\left[-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Lambda}\boldsymbol{B})\right].\end{aligned}$$

One now integrate out both α and Λ. The rest of the proof is the same as in previous sections. □

Appendix D

Proof of Theorem 4. The joint posterior is given by

$$\begin{aligned}&\pi(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\eta},\boldsymbol{\Sigma}_e,\boldsymbol{\rho},|\boldsymbol{y})\\ &\quad\propto\prod_{i,j}p(y_{ij}|\theta_{ij})\\ &\quad\times|\boldsymbol{\Sigma}_e|^{-n/2}\ \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\theta}_i-\boldsymbol{\eta}_i-\boldsymbol{X}_i\boldsymbol{\beta})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{\theta}_i-\boldsymbol{\eta}_i-\boldsymbol{X}_i\boldsymbol{\beta})\right\}\\ &\quad\times\prod_{j=1}^{q}\left[\exp(\rho_j n/2)\ \exp\left\{-\frac{\exp(\rho_j)}{2}\sum_{1\leqslant i<l\leqslant n}w_{il}(\eta_{ij}-\eta_{lj})^2\right\}\right]\\ &\quad\times|\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2}\ \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right\}\\ &\quad\times\exp\left\{-\frac{1}{2}\boldsymbol{\rho}^T\boldsymbol{\Sigma}_\eta^{-1}\boldsymbol{\rho}\right\}.\end{aligned}$$

As in the previous section, writing $\boldsymbol{z}_i=(z_{i1},\ldots,z_{iq})^T=\boldsymbol{\eta}_i-\boldsymbol{\eta}_n\ (i=1,\ldots,n-1)$, $\boldsymbol{z}^T=(\boldsymbol{z}_1^T,\ldots,\boldsymbol{z}_{n-1}^T)$, $\boldsymbol{g}_i=(\boldsymbol{\theta}_i-\bar{\boldsymbol{\theta}})-(\boldsymbol{z}_i-\bar{\boldsymbol{z}})$, integration with respect to $\boldsymbol{\eta}_n$ yields

$$\begin{aligned}&\pi(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{z},\boldsymbol{\Sigma}_e,\boldsymbol{\rho},|\boldsymbol{y})\\ &\quad\propto\prod_{i,j}p(y_{ij}|\theta_{ij})\\ &\quad\times|\boldsymbol{\Sigma}_e|^{-(n-1)/2}\ \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{g}_i-(\boldsymbol{X}_i-\bar{\boldsymbol{X}})\boldsymbol{\beta})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{g}_i-(\boldsymbol{X}_i-\bar{\boldsymbol{X}})\boldsymbol{\beta})\right\}\end{aligned}$$

$$
\begin{aligned}
&\times \prod_{j=1}^{q}\left[\exp(\rho_j n/2)\ \exp\left\{-\frac{\exp(\rho_j)}{2}\sum_{1\leqslant i<l\leqslant n} w_{il}(z_{ij}-z_{lj})^2\right\}\right]\\
&\times|\boldsymbol{\Sigma}_e|^{-(\gamma+q+1)/2}\ \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right\}\\
&\times\exp\left\{-\frac{1}{2}\boldsymbol{\rho}^T\boldsymbol{\Sigma}_\eta^{-1}\boldsymbol{\rho}\right\}.
\end{aligned}
$$

Next integrating with respect to $\boldsymbol{\beta}$ and writing $K(>0)$ for a generic constant, one gets

$$
\begin{aligned}
\pi(\boldsymbol{\theta},\boldsymbol{z},\boldsymbol{\Sigma}_e,\boldsymbol{\rho},|\boldsymbol{y}) \ \leqslant\ & K\prod_{i,j} p(y_{ij}|\theta_{ij})|\boldsymbol{\Sigma}_*|^{1/2}\\
&\times \prod_{j=1}^{q}\left[\exp(\rho_j n/2)\ \exp\left\{-\frac{\exp(\rho_j)}{2}\sum_{1\leqslant i<l\leqslant n} w_{il}(z_{ij}-z_{lj})^2\right\}\right]\\
&\times\ |\boldsymbol{\Sigma}_e|^{-(\gamma+n+q)/2}\ \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right\}\\
&\times\ \exp\left\{-\frac{1}{2}\boldsymbol{\rho}^T\boldsymbol{\Sigma}_\eta^{-1}\boldsymbol{\rho}\right\},
\end{aligned}
$$

where as before $\boldsymbol{\Sigma}_*^{-1}=\sum_{i=1}^{n}(\boldsymbol{X}_i-\bar{\boldsymbol{X}})^T\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{X}_i-\bar{\boldsymbol{X}})$.

Now integrating with respect to $\boldsymbol{z}$, one gets

$$
\begin{aligned}
\pi(\boldsymbol{\theta},\boldsymbol{\Sigma}_e,\boldsymbol{\rho}|\boldsymbol{y}) \ \leqslant\ & K\prod_{i,j} p(y_{ij}|\theta_{ij})|\boldsymbol{\Sigma}_*|^{1/2}\prod_{j=1}^{q}\exp(\rho_j/2)\\
&\times\ \exp(-\boldsymbol{\rho}^T\boldsymbol{\Sigma}_\eta^{-1}\boldsymbol{\rho}/2)|\boldsymbol{\Sigma}_e|^{-(\gamma+n+q)/2}\ \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A})\right\}.
\end{aligned}
$$

Next integrating with respect to ρ, and using the finiteness of the mgf of a multivariate normal distribution, one gets

$$
\pi(\boldsymbol{\theta},\boldsymbol{\Sigma}_e|\boldsymbol{y}) \ \leqslant\ K\prod_{i,j} p(y_{ij}|\theta_{ij})|\boldsymbol{\Sigma}_*|^{1/2}|\boldsymbol{\Sigma}_e|^{-(\gamma+n+q)/2}\ \exp\left\{-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{A}\right\}
$$

The rest of the proof is the same as in the previous sections. □

References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, second ed., Wiley, New York, 1984.
[2] J. Besag, Spatial interaction and the statistical analysis of lattice systems (with discussion), J. Roy. Statistical Society B 36 (1974) 192–236.
[3] J. Besag, P. Green, D. Higdon, K. Mengersen, Bayesian computation and stochastic systems (with discussion), Statistical Sci. 10 (1995) 3–66.
[4] J.A. Bonneson, P.T. McCoy, Effect of median treatment on urban arterial safety: an accident prediction model, Paper presented at the 72nd Annual Meeting of the Transportation Research Board, National Research Council, Washington, DC, 1996.

[5] Bureau of Transportation Statistics (BTS), Tranaportation statistics annual report. United States Department of Transportation, 1996–1999.
[6] B.P. Carlin, S. Banerjee, Hierarchical multivariate CAR models for spatio-temporally correlated survival data, in: J.M. Bernardo, et al., (Eds.), Bayesian Statistics, vol. 7, Oxford University Press, Oxford, 2002, to appear.
[9] PC. Duncan, A. Khattak, F. Council, Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions, Transportation Research Record, National Research Council, 1998.
[10] L. Evans, Traffic Safety and the Drivers, Van Nostrand Reinhold, New York, 1991.
[11] K. Fitzpatrick, A.H. Parham, M.A. Brewer, S-P. Miaou, Characteristics of and potential treatments for crashes on low-volume, rural two-lane highways in Texas, Report Number TX01/4048-1, Texas Department of Transportation, Austin, TX, 2002.
[12] R.J. Flowers, L.I. Griffin, Development of a plan for identifying highway locations that may be over represented in accident frequency and/or severity, Technical Report, Texas Transportation Institute, 1992.
[14] M. Ghosh, K. Natarajan, L.A. Waller, D. Kim, Hierarchical Bayes GLMs for analysis of spatial data: an application to disease mapping, J. Statistical Planning Inference 75 (1999) 305–318.
[15] W.R. Gilks, P. Wild, Adaptive rejection sampling for Gibbs sampling, J. Roy. Statistical Soc. C 41 (1992) 337–348.
[16] E. Hauer, Empirical Bayes approach to the estimation of unsafety: the multivariate regression approach, Accident Anal. Prevention 24 (1992) 456–478.
[18] L. Knorr-Held, N.G. Best, A shared component model for detecting joint and selective clustering of two diseases, J. Roy. Statistical Soc. Ser. A 164 (2000) 73–85.
[19] M.J. Maher, L. Summersgill, A comprehensive methodology for the fitting of predictive accident models, Accident Anal. Prevention 28 (1996) 281–296.
[20] K.V. Mardia, Multi-dimensional multivariate Gaussian Markov random fields with application to image processing, J. Multivariate Anal. 24 (1988) 265–284.
[21] R.G. McGinnis, L. Wissinger, R. Kelly, C. Acuna, Estimating the influence of driver, highway, and environmental factors on run-off-road crashes using logistic regression, Transportation Research Board 78th Annual Meeting, National Research Council, Washington, DC, 1999.
[22] S-P. Miaou, The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions, Accident Anal. Prevention 26 (1994) 471–482.
[23] S-P. Miaou, Measuring the goodness-of-fit of accident prediction models, FHWA-RD-96-040, Federal Highway Administration, US Department of Transportation, 1996, p. 121.
[24] S.-P. Miaou, P.S. Hu, T. Wright, A.K. Rathi, S.C. Davis, Relationship between truck accidents and geometric design: a Poisson regression approach, Transportation Research Record, Transportation Research Board, National Research Council, vol. 1376, 1992, pp. 10–18.
[26] S-P. Miaou, H. Lum, Modeling vehicle accidents and highway geometric design relationships, Accident Anal. Prevention 28 (1993) 689–709.
[27] S.-P. Miaou, J.J. Song, B. Mallick, Roadway traffic-crash mapping: a space-time modeling approach, J. Transportation Statistics 6 (2003) 33–57.
[28] C.N. Morris, C.L. Christiansen, O.J. Pendleton, Application of New Accident Analysis Methodologies, vol. III: Theoretical Development, Publication No. FHWA-RD-91-015, Federal Highway Administration, 1991.
[29] D. Roeleven, M. Kok, H.L. Stipdonk, W.A. de Vries, Inland waterway transport: modelling the probability of accidents, Safety Sci. 19 (1995) 191–202.
[30] V.N. Shankar, J.C. Milton, F.L. Mannering, Modeling statewide accident frequencies as zero-inflated probability processes: an empirical inquiry, Accident Anal. Prevention 29 (1997) 829–837.
[31] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, A. van der Linde, Bayesian measures of model complexity and fit (with discussion), J. Roy. Statistical Soc. B 64 (2002) 583–639.
[32] Transportation Research Board (TRB), Designing Safer Roads: Practices for Resurfacing, Restoration, and Rehabilitation, Special Report 214, National Research Council, Washington, DC, 1987.
[33] A. Vogt, J.G. Bared, Accident models for two-lane rural roads: segments and intersections, Report no. FHWA-RD-98-133, Office of Safety and Traffic Operations R&D, Federal Highway Administration, 1998.
[34] N.A.C. Cressie, Statistics for Spatial Data, Wiley, New York, 1993.
[35] D. Sun, R. Tsutakawa, H. Kim, Z. He, Spatio-temporal interaction with disease mapping, Statistics in Medicine 19 (2000) 2015–2035.

[36] A.E. Gelfand, P. Vounatsou, Proper multivariate conditional autoregressive models for spatial data analysis, Biostatistics 4 (2003) 11–25.

Further reading

[25] S-P. Miaou, D. Lord, Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical bayes, Transportation Research Record 1840 (2003) 31–40