The Canadian Journal of Statistics
Vol. xx, No. yy, 2013, Pages 1–22
La revue canadienne de statistique

Objective Bayesian Analysis of Spatial Models with Separable Correlation Functions

Cuirong Ren[1]*, Dongchu Sun[2] and Sujit K Sahu[3]

[1]*Department of Plant Science, South Dakota State University, Brookings, SD 57007, USA*
[2]*School of Finance and Statistics, East China Normal University, 500 Dongchuan Road, Shanghai 200241, China and Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211, USA*
[3]*Southampton Statistical Sciences Research Institute, University of Southampton, SO17 1BJ, UK*

Key words and phrases: Geostatistics; correlation function; Jeffreys prior; reference prior; integrated likelihood; propriety of posterior.

MSC 2010: Primary 62C10; secondary 62F15

Abstract: This paper considers general linear models for Gaussian geostatistical data with multi-dimensional separable correlation functions involving multiple parameters. We derive various objective priors, such as the Jeffreys-rule, independence Jeffreys, and usual and exact reference priors for the model parameters. In addition, we relax and simplify the assumptions in Paulo [2005] for the propriety of the posteriors in the general setup. We show that the frequentist coverage of posterior credible intervals for a function of range parameters do not depend on the regression coefficient or error variance. These objective priors and a proper flat prior based on ML estimates are compared by examining the frequentist coverage of equal-tailed Bayesian credible intervals. An illustrative example is given from the field of complex computer model validations.
The Canadian Journal of Statistics xx: 1–22; 2013 © 2013 Statistical Society of Canada

Résumé: Insérer votre résumé ici. We will supply a French abstract for those authors who can't prepare it themselves. *La revue canadienne de statistique* xx: 1–22; 2013 © 2013 Société statistique du Canada

1. INTRODUCTION

The Bayesian approach for modeling spatial data and making inferences about spatially varying phenomena can be dated from Kitanidis [1986] and Le and Zidek [1992]. Despite its success, the Bayesian inference depends on the choice of prior distribution, which is supposed to summarize the information a researcher might have about the underlying distribution of the data. This is often a major task before any formal data analysis. However, in practice, it is difficult to accomplish a formal prior elicitation due to time constraints or a lack of prior information. Thus the prior elicitation was often a combination of intuition and ad-hoc methods. See De Oliveira [2010] for references. In addition, Diggle and Ribeiro [2007] pointed out that diffuse priors still have noticeable influence on the inferences even with several hundred observations in a data set, so "it remains a lingering concern because the prior does potentially influence the predictive distribution which we report for any target." Objective Bayesian analysis, however, avoids the prior elicitation and has received a lot of attention since 2001.

* *Author to whom correspondence may be addressed.*
E-mail: cuirong.ren@sdstate.edu

Berger et al. [2001] were perhaps the first who considered objective Bayesian analysis for geostatistical data with a single correlation parameter. Surprisingly, the authors found that the independence Jeffreys prior and the reference prior derived by using the asymptotic marginalization algorithm of Bernardo [1979] and Berger and Bernardo [1992] can fail to yield a proper posterior. Alternatively, they recommended the use of the exact reference prior developed by using exact marginalization in the reference prior algorithm. Since then, objective Bayesian analysis for geostatistical data has been extensively studied (e.g., Paulo 2005; De Oliveira 2007 and 2010).

Noting that there is a single (one-dimensional) range parameter assumed to be unknown in Berger et al. [2001], Paulo [2005] considered multi-dimensional separable correlation functions involving several parameters and found that these commonly used objective priors, which include independence Jeffreys and the usual reference priors, all yield proper posteriors under some assumptions. However, Paulo only considered a rather special linear model with either an intercept without any explanatory variables or one explanatory variable without intercept. This is quite restrictive and in practice, this model is rarely the case.

One of the motivations of this paper is to consider general linear models with spatially dependent errors. In addition, we relax and simplify the assumptions for the propriety of the posteriors in Paulo [2005]. The reference priors based on asymptotic marginalization in the reference prior algorithm of Bernardo [1979] and Berger and Bernardo [1992] are also considered and yield a proper posterior.

The paper is organized as follows. In Section 2, we set up the model and present the Jeffreys and reference priors. In Section 3, we introduce a class of priors and obtain the behavior of the integrated likelihood. We also develop the conditions ensuring the posterior propriety for these objective priors. Section 4 presents the frequentist coverage of the credible intervals for the function of parameter of interest, range parameters in our spatial model, is independent of regression coefficient and error variance. We also consider empirical Bayes method by placing proper flat priors for range parameters. The results of a simulation study are used to compare the frequentist coverage of equal-tailed Bayesian credible intervals for these priors. Finally, an example is used for illustration. The summary is given in Section 5.

2. MODELS AND OBJECTIVE PRIORS

2.1. The Model

The model considered here is of general interest for the field of geostatistics, especially for the analysis and validation of complex computer models. A prominent approach to the problem involves fitting a Gaussian process to the computer model output, and a separable correlation function involving several parameters. See, for example, Sacks et al. [1989] and Bayarri et al. [2007].

Let $\{y(\mathrm{s}), \mathrm{s} \in \mathcal{D}\}, \mathcal{D} \subset \mathbb{R}^l$, be the random field of interest, where $\mathbb{R}^l$ denotes l-dimensional Euclidean space and $l \geq 1$. In this study, we will assume the correlation matrix to be separable, and hence we only consider the case $l \geq 2$. Suppose that the data are observed at locations $\mathrm{s}_i \in$

$\mathcal{D}, i = 1, \cdots, n$ and that the joint distribution of $\mathbf{y} = (y(\mathbf{s}_1), y(\mathbf{s}_2), \ldots, y(\mathbf{s}_n))'$ follows

$$(\mathbf{y} \mid \theta, \delta_1, \beta) \sim N_n(\mathbf{X}\beta, \delta_1\mathbf{\Sigma}), \tag{1}$$

where $\mathbf{X}$ is the known $n \times p$ design matrix, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$ is the vector of unknown regression coefficients, $\delta_1 = \text{var}\{y(\mathbf{s})\}$ and $\mathbf{\Sigma}$ is an $n \times n$ matrix with elements $\mathbf{\Sigma}_{ij} = \text{corr}\{y(\mathbf{s}_i), y(\mathbf{s}_j)\}$. We assume that $\text{corr}\{y(\mathbf{s}_i), y(\mathbf{s}_j)\} = K_\theta(d(\mathbf{s}_i, \mathbf{s}_j))$, a known function of $d(\mathbf{s}_i, \mathbf{s}_j)$ and unknown parameter $\theta = (\theta_1, \ldots, \theta_r)'$. Here $d(\cdot, \cdot)$ is an Euclidean distance. The case where the correlation function depends only on the absolute distance between points is often termed isotropic. Thus, K_θ is an isotropic correlation function, and $\mathbf{\Sigma}$ is a function of unknown parameter vector θ.

Remark 1. *(a) When $r = 1$, objective Bayesian analysis for model (1) has been studied by Berger et al. [2001]. In this paper, we assume $r > 1$. (b) When $p = 1$, it has been studied by Paulo [2005].*

For the model (1), the likelihood function of $(\theta, \delta_1, \beta)$ is given by

$$L(\theta, \delta_1, \beta; \mathbf{y}) \propto \frac{1}{\delta_1^{\frac{n}{2}} |\mathbf{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\delta_1}(\mathbf{y} - \mathbf{X}\beta)'\mathbf{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\beta)\right\}. \tag{2}$$

A full Bayesian analysis would require a prior. As discussed in the introduction, we sought objective priors for the unknown parameter $(\theta, \delta_1, \beta)$. We first give the Jeffreys and usual reference priors, and then derive two exact reference priors. These objective priors are improper and the propriety of the posterior is studied under the separable conditions.

2.2. Jeffreys and Reference Priors

The following proposition gives the Fisher information matrix for $(\theta, \delta_1, \beta)$, denoted by $\mathbf{I}(\theta, \delta_1, \beta)$. This is crucial for finding the Jeffreys-rule and independence Jeffreys priors. The independence Jeffreys priors are obtained by treating the groups of parameters are independent, as it was studied by Berger et al. [2001] and Paulo [2005]. Parts (a)-(c) can also be found in Paulo [2005], whose proof is omitted. Some relationships among the independence Jeffreys priors and reference priors can be found in Remark 2.

Proposition 1. *(a) For $i, j = 1, \ldots, r$, write $\mathbf{U}_i = \frac{\partial \mathbf{\Sigma}}{\partial \theta_i}\mathbf{\Sigma}^{-1}$, $\psi_i = \text{tr}[\mathbf{U}_i]$, and $\mathbf{\Psi}_{ij} = \text{tr}[\mathbf{U}_i\mathbf{U}_j]$. Also, define the column vector $\psi = (\psi_\mathbf{1}, \ldots, \psi_\mathbf{r})'$ and the $r \times r$ matrix $\mathbf{\Psi} = (\mathbf{\Psi}_{ij})$. The Fisher information matrix of $(\theta, \delta_1, \beta)$ is given by*

$$\mathbf{I}(\theta, \delta_1, \beta) = \text{diag}\left(\frac{1}{2}\begin{pmatrix} \mathbf{\Psi} & \frac{\psi}{\delta_1} \\ \frac{\psi'}{\delta_1} & \frac{n}{\delta_1^2} \end{pmatrix}, \frac{1}{\delta_1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X}\right). \tag{3}$$

(b) The Jeffreys-rule prior of $(\theta, \delta_1, \beta)$ is then given by

$$\pi^J(\theta, \delta_1, \beta) \propto \frac{\sqrt{|\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X}|}\sqrt{|\mathbf{\Psi}|\,(n - \psi'\mathbf{\Psi}^{-1}\psi)}}{\delta_1^{1+\frac{p}{2}}}. \tag{4}$$

(c) The independence Jeffreys prior of $(\theta, \delta_1, \beta)$, obtained by assuming that (θ, δ_1) and β are independent a priori, is

$$\pi^{IJ1}(\theta, \delta_1, \beta) \propto \frac{\sqrt{|\boldsymbol{\Psi}| \left(n - \psi' \boldsymbol{\Psi}^{-1} \psi\right)}}{\delta_1}. \tag{5}$$

(d) The independence Jeffreys prior of $(\theta, \delta_1, \beta)$, obtained by assuming that θ, δ_1 and β are independent a priori, is

$$\pi^{IJ2}(\theta, \delta_1, \beta) \propto \frac{\sqrt{|\boldsymbol{\Psi}|}}{\delta_1}. \tag{6}$$

Remark 2. *One can show that the reference priors with the orderings $\{(\theta, \delta_{\mathbf{1}}), \beta\}$, $\{\beta, (\theta, \delta_{\mathbf{1}})\}$, $\{\theta, \delta_1, \beta\}$, $\{\beta, \theta, \delta_1\}$ and $\{\theta, \beta, \delta_1\}$ are the same as $\pi^{IJ1}(\theta, \delta_1, \beta)$. Here the ordering $\{\theta, \delta_1, \beta\}$ means that θ is the most important or the parameter of interest, δ_1 is less important, and β is least important. Also, the reference priors with the orderings $\{\theta, \delta_1, \beta\}$, $\{\delta_1, \beta, \theta\}$ and $\{\beta, \theta, \delta_1\}$ are the same as $\pi^{IJ2}(\theta, \delta_1, \beta)$.*

2.3. The Exact Reference Priors

Paulo [2005] also derived the Fisher information matrix based on the integrated likelihood of (θ, δ_1) by integrating out β under the constant prior and one "exact" reference prior for the parameters in model (1), and we briefly present the ideas and the results here.

In order to derive the "exact" reference priors for the parameters in model (1), we specify (θ, δ_1) as the parameter of interest and β as the nuisance parameter. The reference prior can be factored as $\pi^R(\theta, \delta_1, \beta) = \pi^R(\beta|\theta, \delta_1)\pi^R(\theta, \delta_1)$. Choose $\pi^R(\beta|\theta, \delta_1) = 1$ since the conditional Jeffreys-rule (or reference) prior for parameter vector β is constant when (θ, δ_1) is assumed to be known. The integrated likelihood of (θ, δ_1) with respect to this conditional prior is

$$L_*(\theta, \delta_1; \mathbf{y}) \propto \frac{1}{\delta_1^{\frac{n-p}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}} |\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}|^{\frac{1}{2}}} \exp\left(-\frac{S^2}{2\delta_1}\right), \tag{7}$$

where $S^2 = \mathbf{y}'\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})\mathbf{y}$. Here the matrix function $\boldsymbol{\Gamma}(\cdot, \cdot)$ is defined as

$$\boldsymbol{\Gamma}(\mathbf{A}, \mathbf{Z}) = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{A}^{-1}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{A}^{-1}, \tag{8}$$

where we assume $\mathbf{A}$ is an $n \times n$ and $\mathbf{Z}$ is $n \times p$, and both $\mathbf{A}$ and $\mathbf{Z}'\mathbf{A}^{-1}\mathbf{Z}$ are nonsingular so $\mathbf{Z}$ is a full column rank matrix.

Berger et al. [2001] pointed out that there is a particular transformation of the data that has sampling distribution proportional to (7) based on the result given by Harville [1974], so we can compute the associated Jeffreys-rule prior from the above integrated likelihood.

It follows from Paulo [2005] that the Fisher information matrix of (θ, δ_1) based on the integrated likelihood function L_* is given by

$$\mathbf{I}(\theta, \delta_1) = \frac{1}{2} \begin{pmatrix} \boldsymbol{\Xi} & \frac{\xi}{\delta_1} \\ \frac{\xi'}{\delta_1} & \frac{n-p}{\delta_1^2} \end{pmatrix}, \tag{9}$$

where ξ is the r-dimensional column vector defined by $\xi_i = \mathrm{tr}[\mathbf{V}_i]$ and Ξ is the $r \times r$ matrix defined by $\Xi_{ij} = \mathrm{tr}[\mathbf{V}_i \mathbf{V}_j]$; here $\mathbf{V}_i = \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X}), i = 1, \ldots, r$. Part (a) in the following proposition can be found in Proposition 2.1 of Paulo [2005]) and (b) is easy to derive.

Proposition 2. *Suppose that the sampling distribution is given by (1).*
(a) The exact reference prior with ordering $\{(\theta, \delta_1), \beta\}$ *is*

$$\pi^{R1}(\theta, \delta_1, \beta) \propto \frac{1}{\delta_1} \sqrt{|\Xi| \left(n - p - \xi' \Xi^{-1} \xi\right)}. \tag{10}$$

(b) The exact reference priors with orderings $\{\theta, \delta_1, \beta\}$ *and* $\{\delta_1, \theta, \beta\}$ *are given by*

$$\pi^{R2}(\theta, \delta_1, \beta) \propto \frac{1}{\delta_1} \sqrt{|\Xi|}. \tag{11}$$

2.4. Separability

We partition a general element of $\mathrm{s} \in \mathcal{D}$ into $\mathrm{s} = (\mathrm{s}^1, \ldots, \mathrm{s}^r)$, where $r \leq l$ and each subvector $\mathrm{s}^i, i = 1, \ldots r$ is of dimension l_i. Thus, $\sum_{i=1}^r l_i = l$. The correlation function K_θ is called *partially separable* if $K_\theta(d(\mathbf{s}, \mathbf{t})) = \prod_{i=1}^r K_i(d(\mathbf{s}^i, \mathbf{t}^i))$, where $\mathbf{s}, \mathbf{t} \in \mathcal{D}$ and $K_i(\cdot)$ are isotropic correlation functions in $\mathbb{R}^{\mathrm{l}_\mathrm{i}}$. If a separable correlation function is applied and if the set of locations at which the process is observed forms a Cartesian product, so that the parameter space is rectangle and the correlation matrix of the data is the Kronecker product of the individual correlation matrices associated with each dimension s^i. (Recall that the Kronecker product of two matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B}$ is defined by $\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B})$.) This is essential for us to study the analytical properties of the integrated likelihood and priors, along with establishing sufficient conditions for the propriety of the posterior. Thus, we will assume that the following conditions hold.

Assumption A. Separability of the correlation function:

$$K_\theta(d(\mathbf{s}, \mathbf{t})) = \prod_{i=1}^r K_{\theta_i}(d(\mathbf{s}^i, \mathbf{t}^i)),$$

where K_{θ_i} is a valid correlation function, $\mathrm{s} = (\mathrm{s}^1, \ldots, \mathrm{s}^r)$ and $l \geq r \geq 2$.
Assumption B. Cartesian product of the design set:

$$D = D_1 \times \ldots \times D_r,$$

where $D_i = \{\mathbf{x}_{1i}, \ldots, \mathbf{x}_{n_i,i}\} \subset \mathbb{R}^{l_i}$. By this, there are n_i elements in D_i. Thus, we have $n = \prod_{i=1}^r n_i$, where D has n elements.

Assumption C. Separable design matrix $\mathbf{X}$:

$$\mathbf{X} = \mathbf{X}_1 \otimes \mathbf{X}_2 \otimes \ldots \otimes \mathbf{X}_r, \tag{12}$$

where $\mathbf{X}_i$ are $n_i \times p_i, i = 1, 2, \ldots, r$. Thus, $p = \prod_{i=1}^r p_i$.

Remark 3. *Paulo [2005] considered $p_i = 1$ for $i = 1, \ldots r$.*

By Assumptions A and B, we have a very convenient Kronecker product expression for the correlation matrix of the data:

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_2 \otimes \ldots \otimes \boldsymbol{\Sigma}_r \equiv \otimes_{i=1}^r \boldsymbol{\Sigma}_i, \tag{13}$$

where $\boldsymbol{\Sigma}_i$ are the $n_i \times n_i$ correlation matrices associated with each of the separated dimensions and each $\boldsymbol{\Sigma}_i$ is assumed to depend only on θ_i for $i = 1, \ldots, r$.

Note that $\mathbf{U}_i = \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Sigma}^{-1}$. Under Assumptions A and B, we will have nice expressions for $\mathrm{tr}[\mathbf{U}_i]$ and $\mathrm{tr}[\mathbf{U}_i\mathbf{U}_j]$ by (13), accordingly the expressions of objective priors in Proposition 2.2, which are useful in deriving propriety of priors and the corresponding posteriors. In the following proposition, we only present the expression for $\pi^{IJ1}(\theta, \delta_1, \beta)$. The expressions for other priors such as $\pi^J(\theta, \delta_1, \beta)$, and $\pi^{IJ2}(\theta, \delta_1, \beta)$, can be derived similarly. The proof is in Appendix B.

Proposition 3. *Under Assumptions A and B, we have*

(a) The expressions of $\mathrm{tr}[\mathbf{U}_i]$ and $\mathrm{tr}[\mathbf{U}_i\mathbf{U}_j]$ are given by

$$\mathrm{tr}[\mathbf{U}_i] = \frac{n}{n_i} \mathrm{tr}\left[\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i} \boldsymbol{\Sigma}_i^{-1}\right], \tag{14}$$

$$\mathrm{tr}[\mathbf{U}_i\mathbf{U}_j] = \begin{cases} \frac{n}{n_i} \mathrm{tr}\left(\left[\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i} \boldsymbol{\Sigma}_i^{-1}\right]^2\right), & \text{if } i = j, \\ \frac{n}{n_i n_j} \mathrm{tr}\left[\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i} \boldsymbol{\Sigma}_i^{-1}\right] \mathrm{tr}\left[\frac{\partial \boldsymbol{\Sigma}_j}{\partial \theta_j} \boldsymbol{\Sigma}_j^{-1}\right]. & \text{if } i \neq j. \end{cases} \tag{15}$$

(b) The independence Jeffreys prior given in (5) has the form

$$\pi^{IJ1}(\theta, \delta_1, \beta) \propto \frac{1}{\delta_1} \prod_{i=1}^r \left\{ n_i \mathrm{tr}\left(\left[\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i} \boldsymbol{\Sigma}_i^{-1}\right]^2\right) - \left(\mathrm{tr}\left[\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i} \boldsymbol{\Sigma}_i^{-1}\right]\right)^2 \right\}^{1/2}. \tag{16}$$

(c) For any permutation $(i_1, \cdots, i_p)$ of $(1, \cdots, p)$ and any permutation $(j_1, \cdots, j_r)$ of $(1, \cdots, r)$, the one-at-a-time reference priors for $(\theta, \delta_1, \beta)$ with the ordering $\{\theta_{j_1} \cdots, \theta_{j_r}, \delta_1, \beta_{i_1}, \cdots, \beta_{i_p}\}$ or $\{\beta_{i_1}, \cdots, \beta_{i_p}, \theta_{j_1} \cdots, \theta_{j_r}, \delta_1\}$ are the same as π^{IJ1} in Part (b).

From equations (14) and (15), $\mathrm{tr}[\mathbf{U}_i]$ only depend on θ_i and $\mathrm{tr}[\mathbf{U}_i\mathbf{U}_j]$ only depend on both θ_i and θ_j. Part (b) shows that these priors can expressed as the product of r functions and each of them is a function of only $\theta_i, i = 1, \ldots, r$.

3. POSTERIOR PROPRIETY

The objective priors for $(\theta, \delta_1, \beta)$ in Propositions 1 and 2 are all improper because marginal priors for both δ_1 and β are improper. In order to use these priors, one has to verify the posterior propriety. The objective priors all belong to the following class of improper prior densities for $(\theta, \delta_1, \beta)$,

$$\pi(\theta, \delta_1, \beta) \propto \frac{\pi(\theta)}{\delta_1^a}, \tag{17}$$

where a is a fixed real value and $\pi(\theta)$ is the marginal prior for θ. Here the priors of θ, δ_1 and β are independent. We will see that the marginal prior for θ in Propositions 1 and 2 can be proper or improper.

Under the prior (17), we have

$$\int_0^\infty \int_{\mathbf{R}^p} L(\theta, \delta_1, \beta; \mathbf{y})\pi(\theta, \delta_1, \beta) d\delta_1 d\beta = L_{**a}(\theta; \mathbf{y})\pi(\theta),$$

where

$$L_{**a}(\theta; \mathbf{y}) \propto \frac{1}{|\mathbf{\Sigma}|^{\frac{1}{2}} |\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X}|^{\frac{1}{2}} (S^2)^{\frac{n-p}{2}+a-1}}. \tag{18}$$

Under Assumptions A, B and C, the integrated likelihood (18) can be written as

$$L_{**a}(\theta; \mathbf{y}) \propto \frac{1}{\prod_{i=1}^r \left\{ |\mathbf{\Sigma}_i|^{\frac{n}{2n_i}} |\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i|^{\frac{p}{2p_i}} \right\} (S^2)^{\frac{n-p}{2}+a-1}}. \tag{19}$$

Because the prior distributions are improper, we need to verify the posterior distribution propriety. This can be done under the following assumptions. These assumptions are not too restrictive and are satisfied by many commonly used isotropic models, especially four standard models: *Spherical, power exponential, rational quadratic* and *Matérn*. The details of these models can be found in Banerjee et al. [2004], for example.

Assumption D1. $K_{\theta_i}(d) = K_i(d/\theta_i)$, where K_i is a continuous correlation function satisfying $\lim_{u\to\infty} K_i(u) = 0$ for $i = 1, 2, \ldots, r$.

Assumption D2. As $\theta_i \to \infty$ for $i = 1, 2, \ldots, r$, correlation matrix $\mathbf{\Sigma}_i$ in (13) satisfies

$$\mathbf{\Sigma}_i = \mathbf{\Sigma}_i^* + o(\nu_i(\theta_i)), \tag{20}$$

$$\mathbf{\Sigma}_i^* = \mathbf{1}_{n_i}\mathbf{1}_{n_i}' + \nu_i(\theta_i)\mathbf{D}_i \tag{21}$$

where $\nu_i(\theta_i)(>0)$ is a continuous function of $\theta_i > 0$ satisfying $\lim_{\theta_i \to \infty} \nu_i(\theta_i) = 0$ and $\mathbf{D}_i$ is a fixed nonsingular matrix satisfying $\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i} \neq 0$.

Assumption D3. For each i, $\boldsymbol{\Sigma}_i$ is continuously differentiable and as $\theta_i \to \infty$,

$$\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i} = \nu'_i(\theta_i)\mathbf{D}_i(1 + o(1)). \tag{22}$$

Assumption D4. $n_i > p_i, i = 1, \ldots, r$.

Assumption D5. $(\mathrm{tr}[\mathbf{U}_i^2])^{1/2}$ is integrable at zero for $\theta_i, i = 1, \ldots, r$.

Remark 4. *The proof of the following (iii) and (iv) can be found in Ren [2008].*

(i) Although Paulo [2005] made the assumptions for $\boldsymbol{\Sigma}_i$ and $\frac{\partial \boldsymbol{\Sigma}_i}{\partial \theta_i}$ implying (20) and (22), he basically used our expressions of assumption in his verifications.

(ii) The following assumptions about $\mathbf{D}_i$ and $\mathbf{X}_i$ in Paulo [2005]:

(a) $\mathbf{X}'_i\mathbf{D}_i^{-1}\mathbf{X}_i$ is nonsingular;

(b) If $\mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i)$, then $\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i} \neq \mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{X}_i(\mathbf{X}'_i\mathbf{D}_i^{-1}\mathbf{X}_i)^{-1}\mathbf{X}'_i\mathbf{D}_i^{-1}\mathbf{1}_{n_i}$,

can be dropped where $\mathcal{C}(\mathbf{X}_i)$ denotes the column space of $\mathbf{X}_i$ consisting of all linear combinations of column vectors of $\mathbf{X}_i$, $i = 1, \ldots, r$. This generalizes Paulo's results.

(iii) If the correlation matrix $\boldsymbol{\Sigma}_i$ is positive definite, then the nonsingular $\mathbf{D}_i$ in Assumption $D2$ has exactly one negative eigenvalue and $n_i - 1$ positive eigenvalues.

(iv) Assume that the correlation matrix $\boldsymbol{\Sigma}_i$ is positive definite. An equivalent condition for which $\mathbf{D}_i$ is nonsingular and $\mathbf{1}_{n_i}{}'\mathbf{D}_i^{-1}\mathbf{1}_{n_i} \neq 0$ in Assumption $D2$ is that $\boldsymbol{\Sigma}_i^*$ is a positive definite correlation matrix for sufficiently large θ_i.

In the following, we first present the behavior of the integrated likelihood at zero and at infinity for each θ_i while the rest are fixed, which provides the key to determining whether the posterior is proper or not. The proof is in Appendix C. And then we give the results on the posterior propriety under the objective priors in Section 2, which is Theorem 3, and the proof is given in Appendix D. Finally, Theorem 3 gives the propriety of the posterior under the exact reference priors.

In order to simplify the expression of limiting matrices in the following proposition, we introduce some notations. For example, define $\boldsymbol{\Sigma}_{[-i]} = \otimes_{j \neq i}\boldsymbol{\Sigma}_j$, $\mathbf{X}_{[-i]} = \otimes_{j \neq i}\mathbf{X}_j$ and $\boldsymbol{\Phi}_i = \boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i(\mathbf{X}'_i\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i)^{-1}\mathbf{X}'_i\boldsymbol{\Sigma}_i^{-1}$. In addition, we define $\boldsymbol{\Sigma}_{0i}$ and $\boldsymbol{\Phi}_{0i}$ are the limit of $\boldsymbol{\Sigma}$ and $\otimes_{i=1}^r\boldsymbol{\Phi}_i$ when $\theta_i \to 0^+$, respectively. Under Assumption D1, we have

$$\boldsymbol{\Sigma}_{0i} = \boldsymbol{\Sigma}_1 \otimes \ldots \otimes \boldsymbol{\Sigma}_{i-1} \otimes \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_{i+1} \ldots \otimes \boldsymbol{\Sigma}_r,$$
$$\boldsymbol{\Phi}_{0i} = \boldsymbol{\Phi}_1 \otimes \ldots \otimes \boldsymbol{\Phi}_{i-1} \otimes \mathbf{X}_i(\mathbf{X}'_i\mathbf{X}_i)^{-1}\mathbf{X}'_i \otimes \boldsymbol{\Phi}_{i+1} \ldots \otimes \boldsymbol{\Phi}_r.$$

Proposition 4. *Under Assumptions $A, B, C, D1, D2$ and $D4$, we have $L_{**a}(\theta; \mathbf{y})$ is a continuous function and* *(a) For fixed* $\theta_{[-i]} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_r)'$*, when* $\theta_i \to 0^+$*,*

$$L_{**a}(\theta; \mathbf{y}) \propto \frac{1}{|\boldsymbol{\Sigma}_{[-i]}|^{\frac{n_i}{2}} |\mathbf{X}'_{[-i]} \boldsymbol{\Sigma}^{-1}_{[-i]} \mathbf{X}_{[-i]}|^{\frac{p_i}{2}} (S^2_{0,i})^{\frac{n-p}{2}+a-1}},$$

where $S^2_{0,i} = \mathbf{y}'(\boldsymbol{\Sigma}^{-1}_{0i} - \boldsymbol{\Phi}_{0i})\mathbf{y}$ *depends only on* $\theta_{[-i]}$*.*

(b) For fixed $\theta_{[-i]}$*, when* $\theta_i \to \infty$*,*

$$L_{**a}(\theta; \mathbf{y}) \propto \begin{cases} g(\theta_{[-i]})\nu_i(\theta_i)^{\frac{n}{2n_i}+a-1}, & \text{if } \mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i), \\ g(\theta_{[-i]})\nu_i(\theta_i)^{\frac{n}{2n_i}-\frac{p}{2p_i}+a-1}, & \text{if } \mathbf{1}_{n_i} \in \mathcal{C}(\mathbf{X}_i), \end{cases}$$

where

$$g(\theta_{[-i]}) = \frac{1}{|\boldsymbol{\Sigma}_{[-i]}|^{\frac{n_i}{2}} |\mathbf{X}'_{[-i]} \boldsymbol{\Sigma}_{[-i]} \mathbf{X}_{[-i]}|^{\frac{p_i}{2}} (S^2_{1,i})^{\frac{n-p}{2}+a-1}},$$

$$S^2_{1,i} = \mathbf{y}' \left\{ \otimes_{j=1}^{i-1} \boldsymbol{\Sigma}_j^{-1} \otimes \mathbf{E}_i \otimes_{j=i+1}^{r} \boldsymbol{\Sigma}_j^{-1} - \otimes_{j=1}^{i-1} \boldsymbol{\Phi}_j^{-1} \otimes \widetilde{\mathbf{E}}_i \otimes_{j=i+1}^{r} \boldsymbol{\Phi}_j^{-1} \right\} \mathbf{y},$$

$$\mathbf{E}_i = \mathbf{D}_i^{-1} - \frac{\mathbf{D}_i^{-1}\mathbf{1}_{n_i}\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}}{\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}}, \quad \widetilde{\mathbf{E}}_i = \begin{cases} \mathbf{E}_i\mathbf{X}_i(\mathbf{X}'_i\mathbf{E}_i\mathbf{X}_i)^{-1}\mathbf{X}'_i\mathbf{E}_i, & \text{if } \mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i), \\ \mathbf{E}_i - \boldsymbol{\Gamma}(\mathbf{D}_i, \mathbf{X}_i), & \text{if } \mathbf{1}_{n_i} \in \mathcal{C}(\mathbf{X}_i). \end{cases}$$

Theorem 1. *Consider the model with sampling distribution (2). Under Assumptions* $A, B, C, D1 - D5$*, the reference, Jeffreys-rule and independence Jeffreys priors yield proper posterior distributions.*

Since a Fisher information matrix is nonnegative definite, by Hadamard's inequality (e.g. Page 306 at Marshall et al. [2009]), $|\boldsymbol{\Xi}| \leq \prod_{i=1}^{r}(\mathrm{tr}[\mathbf{V}_i^2])^{1/2}$ and hence there exists a constant $c > 0$ such that

$$\pi^{Rj}(\theta) \leq c \prod_{i=1}^{r} (\mathrm{tr}[\mathbf{V}_i^2])^{1/2}, \quad j = 1, 2,$$

where $\pi^{Rj}(\theta)$ are the marginal reference priors for θ. By the definition of $\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})$, we have $\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X}) \leq \boldsymbol{\Sigma}^{-1}$, where two symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ satisfying $\mathbf{A} \leq \mathbf{B}$ mean that $\mathbf{B} - \mathbf{A}$ is nonnegative definite. Since $\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})$ is nonnegative definite, it can be written as $\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})^{1/2}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})^{1/2}$.

$$\begin{aligned} \mathrm{tr}[\mathbf{V}_i^2] &= \mathrm{tr}\left[\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})^{1/2} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X}) \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})^{1/2}\right] \\ &\leq \mathrm{tr}\left[\boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})^{1/2} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})^{1/2}\right] \\ &= \mathrm{tr}\left[\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Gamma}(\boldsymbol{\Sigma}, \mathbf{X})\right]. \end{aligned} \tag{23}$$

Since $\mathbf{\Sigma}^{-1}$ is a positive definite, it can be written as $\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{-1/2}$. Applying the above method to $\mathrm{tr}\left[\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\mathbf{\Gamma}(\mathbf{\Sigma},\mathbf{X})\right]$, we have

$$\mathrm{tr}[\mathbf{V}_i^2] \leq \mathrm{tr}[\mathbf{U}_i^2],$$

and hence we have the following results from the conclusion in Theorem 1.

Theorem 2. *Consider the model with sampling distribution (2). Under Assumptions $A, B, C, D1-D5$, the exact reference priors yield proper distributions.*

4. NUMERICAL COMPARISONS

In this section, we first introduce a simulation method that makes computation less expensive and more feasible. Second, we briefly describe the empirical Bayes method introduced by Paulo [2005], which is used to construct proper vague priors. Next, we will report the simulation studies of the frequentist properties of Bayesian inference for our parameter of interest θ based on the following priors: Empirical Bayes, the Jeffreys-rule prior π^J, two independence Jeffreys priors π^{IJ1} and π^{IJ2}, and two exact reference priors π^{R1} and π^{R2}. Finally, we present results for an example from Bayarri et al. [2007].

4.1. Frequentist Coverage of the Credible Intervals

Suppose $\eta = \eta(\xi)$, which is a function of the parameter vector $\xi = (\theta, \delta_1, \beta)$, is a parameter of interest. Note that η could be a function of θ if the components of θ are the only parameters of interest.

For the fixed $\xi = \xi^* \equiv (\theta^*, \delta_1^*, \beta^*)$, we will simulate the data based on $\mathbf{y} \mid \xi^*$. Denote the α-posterior quantile of η given $\mathbf{y}$ by $\eta_\alpha(\mathbf{y})$ where $0 < \alpha < 1$. That is

$$P(\eta^* < \eta_\alpha(\mathbf{y}) \mid \mathbf{y}) = \alpha, \ \forall\alpha \in (0,1), \tag{24}$$

where the probability is computed based on the marginal posterior distribution of η given $\mathbf{y}$.

Next, consider the frequentist coverage of the one-sided credible interval $(\eta_L, \eta_\alpha(\mathbf{y}))$, where

$$P_{\xi^*}(\eta^* < \eta_\alpha(\mathbf{y})), \tag{25}$$

and η_L is the low boundary of η and the probability is based on the distribution of $\mathbf{y}$ given ξ^*. It is possible that the coverage probability depends on the $\eta_\alpha(\mathbf{y})$, which is often hard to compute by itself. Alternatively, since

$$\eta^* < \eta_\alpha(\mathbf{y}) \ \text{ if and only if } \ F(\eta^*|\mathbf{y}) < \alpha,$$

where $F(\eta \mid \mathbf{y})$ is the marginal cumulative posterior distribution of η given $\mathbf{y}$,

$$P_{\xi^*}(\eta^* < \eta_\alpha(\mathbf{y})) = P_{\xi^*}(F(\eta^* \mid \mathbf{y}) < \alpha). \tag{26}$$

It shows that the frequentist coverage probabilities depend only on the posterior cumulative distribution function $F(\eta^* \mid \mathbf{y})$ at the true values. Finding $F(\eta^* \mid \mathbf{y})$ requires only integration, and there is no need to find posterior quantiles in simulations.

Theorem 3. *Assume that the prior (17) is used. If η is a function of θ, then the frequentist coverage probabilities in (26) depend only on θ^* and are independent of (δ_1, β).*

The proof is similar to the proof of Theorem 2 in Ren and Sun [2012], so it is omitted here.

Theorem 3 shows that the frequentist coverage probabilities of Bayesian credible intervals under a large class of priors will depend only on θ. Therefore, in the simulation study (δ_1, β) could be taken at any value. For simplicity, take (δ_1^*, β^*) to be $(1, \mathbf{0})$. Now, since one does not need to consider the choices of nuisance parameters, it can simplify and speed up computation tremendously.

Let's consider a special case when $r = 2$. That is, θ is a 2-dimensional vector, denoted by (θ_1, θ_2). It is easy to see that finding the marginal posterior cumulative distribution of θ_1 and θ_2 requires only 2-dimensional integration. In fact, define

$$g(\theta_1, \theta_2) = \frac{\pi(\theta_1, \theta_2)}{|\mathbf{G}|^{1/2}|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|^{1/2}(S^2)^{(n-p)/2+a-1}}.$$

If we use the transformations $s_1 = \theta_1/(\theta_1 + 1)$ and $s_2 = \theta_2/(\theta_2 + 1)$,

$$F_1(\theta_1^* \mid \mathbf{y}) \equiv P(\theta_1 < \theta_1^* \mid \mathbf{y}) = \frac{\displaystyle\int_0^{\frac{\theta_1^*}{\theta_1^*+1}}\int_0^1 g\left(\frac{s_1}{1-s_1}, \frac{s_2}{1-s_2}\right)\frac{1}{(1-s_1)^2(1-s_2)^2}ds_1 ds_2}{\displaystyle\int_0^1\int_0^1 g\left(\frac{s_1}{1-s_1}, \frac{s_2}{1-s_2}\right)\frac{1}{(1-s_1)^2(1-s_2)^2}ds_1 ds_2}, \tag{27}$$

$$F_2(\theta_2^* \mid \mathbf{y}) \equiv P(\theta_2 < \theta_2^* \mid \mathbf{y}) = \frac{\displaystyle\int_0^1\int_0^{\frac{\theta_2^*}{\theta_2^*+1}} g\left(\frac{s_1}{1-s_1}, \frac{s_2}{1-s_2}\right)\frac{1}{(1-s_1)^2(1-s_2)^2}ds_1 ds_2}{\displaystyle\int_0^1\int_0^1 g\left(\frac{s_1}{1-s_1}, \frac{s_2}{1-s_2}\right)\frac{1}{(1-s_1)^2(1-s_2)^2}ds_1 ds_2}. \tag{28}$$

Thus, if, for example, we take a random sample of size m, $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m)$, from the model (1) with the parameter $\xi^* = (\theta_1^*, \theta_2^*, 1, \mathbf{0})$, then the frequentist coverage probability $P_{\xi^*}(\theta_1^* < \theta_{1\alpha}(\mathbf{y}))$ can be estimated by

$$\frac{\#\{\mathbf{y}_i, i = 1, \cdots, m : F_1(\theta_1^* \mid \mathbf{y}_i) < \alpha\}}{m}.$$

4.2. Proper Vague Priors

In practice, some proper and vague priors or data-dependent priors have been used. An obvious advantage of using such proper priors is that the posterior is always proper. However, Berger et al. [2001] pointed out the problems by using vague proper priors. For example, it is extremely sensitive to the hyperparameters chosen. Paulo [2005] studied a data-dependent proper prior for range parameters. Here we describe his method, which we will use for comparisons in the next section.

Suppose $\theta_i, i = 1, \ldots, r$ have independent exponential priors with unknown mean parameters $\phi_i, i = 1, \ldots, r$. Let $\hat{\theta}_i$ be the MLE of θ_i. Then the priors of these parameters are exponential (denoted by π^{EB}) obtained by replacing ϕ_i by $c\hat{\theta}_i$, where c is a factor. Paulo [2005] chose $c = 10$. $\hat{\theta}_i$ are the solution of the following equations:

$$\frac{\partial l_*}{\partial \theta_i} = \frac{1}{2}\left\{\frac{\mathbf{y}'\mathbf{R}_{\boldsymbol{\Sigma}}\mathbf{V}_i\mathbf{y}}{\delta_1} - \mathrm{tr}[\mathbf{V}_i]\right\} = 0, i = 1, \ldots, r,$$
$$\frac{\partial l_*}{\partial \delta_1} = \frac{1}{2}\left(\frac{S^2}{\delta_1^2} - \frac{n-p}{\delta_1}\right) = 0,$$

where $l_* = \log L_*(\theta, \delta_1; \mathbf{y})$. Since there is no analytical expression of θ, Fisher's scoring method, a variant of the Newton-Raphson method that results from approximating the Hessian of the logarithm of integrated likelihood by its expected value is used. The $(s+1)$st iterate of the numerical method is given by

$$\theta^{(s+1)} = \theta^{(s)} + \lambda \Xi(\theta^{(s)})^{-1} \frac{\partial l_*}{\partial \theta}\bigg|_{\theta=\theta^{(s)}, \delta_1 = \hat{\delta}_1},$$

where $\Xi(\theta) \equiv \Xi$, $\hat{\delta}_1 = S^2/(n-p)$ and λ is the step size of the algorithm used to assure convergence.

4.3. Simulation Study

We perform a simulation experiment to investigate the frequentist coverage of equal-tailed Bayesian credible intervals for two parameters of interest: the range parameters θ_1 and θ_2, when one of the five objective priors or the empirical Bayes method is used. The closer to the nominal level this frequentist coverage is, the 'better' the prior.

The simulation study is conducted in the context of the power exponential correlation function with $\alpha_1 = \alpha_2 = 1$, that is

$$K((u_1, v_1), (u_2, v_2)) = \exp(-|u_1 - u_2|/\theta_1)\exp(-|v_1 - v_2|/\theta_2).$$

For a 5×5 equally spaced grid in $\mathcal{D} = [0, 1] \times [0, 1]$, we consider two different mean functions $\mathbb{E}\{\mathbf{y}(\mathbf{s})\}$, namely the constant $p_1 = p_2 = 1$ (dimension $p = 1$) or $p_1 = 2$ (linear) and $p_2 = 3$ (quadratic) (dimension $p = 6$). The three different values of each of θ_1 and θ_2: $0.2, 0.5$, and 1.0 are considered. $3,000$ replications are generated for each choice of (θ_1, θ_2) and compute the equal-tailed 95% credible intervals for both θ_1 and θ_2.

Table 1 gives frequentist coverage of Bayesian equal-tailed 95% credible intervals for both θ_1 and θ_2 corresponding to six priors. The results of these experiments based on the empirical Bayes method are comparatively poor performance for the general case compared with the model in Paulo [2005]. This further confirms that for those who tend to turn to "vague" or "diffuse" priors whenever a formal objective prior is not available, one should oppose this type of approach to objective Bayesian analysis for general case, while Paulo [2005] made it in special case. The coverage of the Bayesian credible intervals are similar to each other and reasonably close to the nominal 0.95 except the Jeffreys-rule prior when $p = 6$ for the rest priors. This was already reported by Berger et al. [2001]. As they pointed out, this is likely a consequence of the spurious degrees of freedom added by the Jeffreys-rule prior.

The present simulation study does not provide a way to compare the rest of the four priors. In addition, it is also too limited to ensure that these four priors generally yield inferences with satisfactory frequentist performance. However, this study does provide rather strong evidence that the Jeffreys-rule prior can be seriously inadequate in terms of frequentist coverage.

4.4. An Example

Consider a simplified version of a problem in Bayarri et al. [2007], which was also analyzed by Paulo [2005]. They considered the analysis of a computer model that simulates the crash of prototype vehicles against a barrier, recording the velocity curve from the point of impact until the vehicle stops. In this paper, we will envision that if we input an impact velocity and an instant in time, the computer model will return the velocity of the vehicle at that point in time after impact.

The data consist of the output of the computer model at 19 values of the instant in time, denoted by t, and 9 initial velocities, denoted by v, which corresponds to a 171-point design set that follows a Cartesian product. Paulo [2005] transformed the data by subtracting from each of the individual curves the initial velocity, so the curves would start at zero and decay at a rate roughly proportional to the initial velocity. Thus, he considered the mean function $\mathbb{E}y(v,t) = vt\beta$ for the transformed data. In this paper, we transform the data. When the original data are used, it seems reasonable to assume the mean function is separable and is given by

$$\begin{aligned}\mathbb{E}y(v,t) &= (\beta_{10} + \beta_{11}t)(\beta_{20} + \beta_{21}v) \\ &= \beta_{10}\beta_{20} + \beta_{11}\beta_{20}t + \beta_{10}\beta_{21}v + \beta_{11}\beta_{21}tv \\ &= \beta_0 + \beta_1 t + \beta_2 v + \beta_3 tv. \end{aligned} \tag{29}$$

For the correlation function, we will assume a two-dimensional separable power exponential function, which is the same as Paulo [2005], with the roughness parameters fixed at 2, and the parameterization

$$k((t_1, v_l), (t_2, v_2)) = \exp\{-\theta_1(t_1 - t_2)^2\}\exp\{-\theta_2(v_1 - v_2)^2\}.$$

Based on our methods, if the mean function is separable, we can test whether the model should include the higher order terms or not. For example, we can consider a separable mean function with a quadratic term. Since the estimates of the coefficients for the higher order terms based on the simulation results are almost equal to zero, we can ignore the quadratic term in the mean

TABLE 1: Frequentist Coverage of Bayesian Equal-Tailed 95% Credible Intervals for θ_1 and θ_2

			$p=1$			$p=6$		
			$\theta_1=.2$	$\theta_1=.5$	$\theta_1=1$	$\theta_1=.2$	$\theta_1=.5$	$\theta_1=1$
$\theta_2=.2$	θ_1	π^{R1}	.964	.947	.943	.951	.943	.939
		π^{R2}	.953	.946	.946	.936	.941	.944
		π^{IJ1}	.963	.948	.945	.955	.952	.948
		π^{IJ2}	.952	.945	.946	.940	.944	.945
		π^{J}	.979	.951	.941	.991	.713	.666
		π^{EB}	.911	.878	.869	.896	.853	.853
	θ_2	π^{R1}	.969	.970	.969	.956	.964	.968
		π^{R2}	.961	.965	.965	.944	.954	.961
		π^{IJ1}	.969	.969	.969	.958	.968	.972
		π^{IJ2}	.961	.964	.964	.943	.955	.964
		π^{J}	.983	.984	.984	.986	.991	.993
		π^{EB}	.920	.916	.917	.893	.893	.893
$\theta_2=.5$	θ_1	π^{R1}	.968	.946	.946	.960	.952	.950
		π^{R2}	.962	.945	.946	.953	.954	.956
		π^{IJ1}	.967	.947	.948	.962	.960	.961
		π^{IJ2}	.960	.945	.947	.957	.956	.958
		π^{J}	.982	.954	.946	.991	.655	.605
		π^{EB}	.904	.877	.866	.890	.842	.854
	θ_2	π^{R1}	.959	.958	.960	.954	.959	.960
		π^{R2}	.955	.957	.959	.948	.957	.960
		π^{IJ1}	.959	.959	.960	.958	.962	.962
		π^{IJ2}	.953	.955	.958	.948	.958	.960
		π^{J}	.956	.955	.957	.868	.825	.808
		π^{EB}	.880	.885	.881	.861	.859	.870
$\theta_2=1$	θ_1	π^{R1}	.967	.949	.945	.962	.955	.952
		π^{R2}	.964	.950	.946	.955	.955	.957
		π^{IJ1}	.966	.950	.947	.966	.964	.963
		π^{IJ2}	.961	.948	.947	.959	.959	.960
		π^{J}	.982	.953	.942	.991	.657	.597
		π^{EB}	.900	.874	.864	.879	.861	.858
	θ_2	π^{R1}	.948	.951	.951	.950	.954	.956
		π^{R2}	.953	.950	.951	.952	.955	.957
		π^{IJ1}	.949	.951	.952	.951	.955	.960
		π^{IJ2}	.954	.953	.953	.952	.956	.957
		π^{J}	.945	.942	.945	.858	.818	.791
		π^{EB}	.869	.869	.872	.864	.857	.862

function. Therefore, it is enough to consider the mean function given by (29).

In this example, four objective priors including two independence Jeffreys and two exact reference priors are used. We adopt a ratio-of-uniforms method by Wakefield et al. [1991] for sampling (θ_1, θ_2) from their marginal posterior distribution. Based on $3,000$ replications, Figure 1 shows the posterior densities based on these four priors and gives no noticeable difference among them.

5. SUMMARY AND COMMENTS

In this paper, we propose an objective Bayesian analysis method for spatial models with separable correlation functions and a separable design matrix. The reference priors and Jeffreys priors, including Jeffreys-rule and independence Jeffreys priors, are developed for these models. The properties of these priors and resulting posteriors are studied. Based on the simulation study, the exact reference priors and independence Jeffreys priors are recommended as default priors.

Berger et al. [2001] found both independence Jeffreys prior and reference prior failed to yield proper posteriors. Recently, Ren et al. [2012] studied the model for spatially correlated data with measurement errors, again for a single correlation parameter, and found these commonly used noninformative priors fail to yield proper posteriors. These results show the complications of using objective priors in spatially correlated data.

Appendix A: Lemmas

Lemma 1. *If $\mathbf{A}$ is a nonsingular matrix and $\mathbf{u}$ and $\mathbf{v}$ are vectors, then*

$$|\mathbf{A} + \mathbf{u}\mathbf{v}'| = |\mathbf{A}|(1 + \mathbf{v}'\mathbf{A}^{-1}\mathbf{u}).$$

If, furthermore, $\mathbf{v}'\mathbf{A}^{-1}\mathbf{u} \neq -1$, then $\mathbf{A} + \mathbf{u}\mathbf{v}'$ is nonsingular and

$$(\mathbf{A} + \mathbf{u}\mathbf{v}')^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}'\mathbf{A}^{-1}}{1 + \mathbf{v}'\mathbf{A}^{-1}\mathbf{u}}.$$

The following four lemmas can be found in Ren [2008].

Lemma 2. *Let $\mathbf{\Sigma}_n$ be an $n \times n$ positive definite correlation matrix. Then we have (i) $\mathbf{1}_n'\mathbf{\Sigma}_n^{-1}\mathbf{1}_n - 1 > 0$, and (ii) $\mathbf{A}$ is nonsingular and $1 + \mathbf{1}_n'\mathbf{A}^{-1}\mathbf{1}_n < 0$, where $\mathbf{A} = \mathbf{\Sigma}_n - \mathbf{1}_n\mathbf{1}_n'$.*

Lemma 3. *Suppose $\mathbf{\Sigma}_\theta = \mathbf{1}_n\mathbf{1}_n' + \nu(\theta)\mathbf{D}$ is an $n \times n$ positive definite correlation matrix, where $\nu(\theta) > 0$ is a continuous function of $\theta > 0$, and $\mathbf{D}$ is a fixed nonsingular matrix. Assume that $\mathbf{X}$ is an $n \times p$ full column rank matrix. Then, $\mathbf{X}'\mathbf{D}^{-1}\mathbf{X}$ is nonsingular if and only if $f(\mathbf{X}) \neq 0$, where*

$$f(\mathbf{X}) = 1 - \mathbf{1}_n'\mathbf{\Gamma}(\mathbf{\Sigma}_\theta, \mathbf{X})\mathbf{1}_n, \tag{30}$$

and $\mathbf{\Gamma}(\cdot, \cdot)$ is given in (8).

With the conclusion that $\mathbf{a}'(\mathbf{I} - \mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}')\mathbf{a} = 0$ if and only if $\mathbf{a} \in \mathcal{C}(\mathbf{T})$, where $\mathbf{T}$ is $n \times p$ and $\text{rank}(\mathbf{T}) = p$, one can easily verify the conclusion of the following lemma.

Lemma 4. *Suppose that* $\mathbf{A}$ *is an* $n \times n$ *positive definite matrix, and* $\mathbf{X}$ *is an* $n \times p$ *full column rank matrix. Then*

$$\mathbf{a}'\mathbf{\Gamma}(\mathbf{A}, \mathbf{X})\mathbf{a} = 0, \tag{31}$$

if and only if $\mathbf{a} \in \mathcal{C}(\mathbf{X})$*, where the function* $\mathbf{\Gamma}(\cdot, \cdot)$ *is given in (8).*

Lemma 5. *Define* $\mathbf{A} = \mathbf{B} - \mathbf{1}_n\mathbf{1}_n'$*. Assume that both* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $n \times n$ *symmetric and nonsingular, and* $\mathbf{X}$ *is an* $n \times p$ *full column rank matrix and* $\mathbf{1}_n'\mathbf{B}^{-1}\mathbf{1}_n \neq 1$*. Assume that both* $\mathbf{X}'\mathbf{A}^{-1}\mathbf{X}$ *and* $\mathbf{X}'\mathbf{B}^{-1}\mathbf{X}$ *are nonsingular. Then we have*

$$\mathbf{\Gamma}(\mathbf{A}, \mathbf{X}) = \mathbf{\Gamma}(\mathbf{B}, \mathbf{X}) + \frac{\mathbf{\Gamma}(\mathbf{B}, \mathbf{X})\mathbf{1}_n\mathbf{1}_n'\mathbf{\Gamma}(\mathbf{B}, \mathbf{X})}{1 - \mathbf{1}_n'\mathbf{\Gamma}(\mathbf{B}, \mathbf{X})\mathbf{1}_n}, \tag{32}$$

$$\mathbf{\Gamma}(\mathbf{B}, \mathbf{X}) = \mathbf{\Gamma}(\mathbf{A}, \mathbf{X}) - \frac{\mathbf{\Gamma}(\mathbf{A}, \mathbf{X})\mathbf{1}_n\mathbf{1}_n'\mathbf{\Gamma}(\mathbf{A}, \mathbf{X})}{1 + \mathbf{1}_n'\mathbf{\Gamma}(\mathbf{A}, \mathbf{X})\mathbf{1}_n}, \tag{33}$$

where the function $\mathbf{\Gamma}(\cdot, \cdot)$ *is given in (8).*

Lemma 6. *Suppose that* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $k \times l$ *matrices and* $\mathbf{C}$ *and* $\mathbf{D}$ *are* $m \times n$ *matrices. If* $\mathbf{A} \otimes \mathbf{C} = \mathbf{B} \otimes \mathbf{D}$*, then there exists a constant* c *such that* $\mathbf{A} = c\mathbf{B}$ *and* $\mathbf{D} = c\mathbf{C}$*.*

Appendix B: Proof of Proposition 3

(a) Because the correlation matrix $\mathbf{\Sigma}$ is separable and each $\mathbf{\Sigma}_i$ depends only on θ_i, we have

$$\mathbf{\Sigma}^{-1} = \otimes_{i=1}^{r}\mathbf{\Sigma}_i^{-1},$$
$$\frac{\partial\mathbf{\Sigma}}{\partial\theta_i} = \mathbf{\Sigma}_1 \otimes \ldots \otimes \mathbf{\Sigma}_{i-1} \otimes \frac{\partial\mathbf{\Sigma}_i}{\partial\theta_i} \otimes \mathbf{\Sigma}_{i+1} \otimes \ldots \otimes \mathbf{\Sigma}_r.$$

Thus, $\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\mathbf{\Sigma}^{-1} = \mathbf{I}_{n_1} \otimes \ldots \otimes \mathbf{I}_{n_{i-1}} \otimes \frac{\partial\mathbf{\Sigma}_i}{\partial\theta_i}\mathbf{\Sigma}_i \otimes \mathbf{I}_{n_{i+1}} \otimes \ldots \otimes \mathbf{I}_{n_r}$ and one can verify (14) and (15).
(b) Note that $\mathrm{tr}[\mathbf{U}_i\mathbf{U}_j] = \frac{1}{n}\mathrm{tr}[\mathbf{U}_i]\mathrm{tr}[\mathbf{U}_j]$ for $i \neq j$ from (a). Therefore, with some algebra, we can obtain the conclusion.
(c) We consider only the ordering $\{\theta_r \cdots, \theta_1, \delta_1, \beta_1, \cdots, \beta_p\}$ and the others are similar. Noting that the Fisher information of $(\beta_1, \cdots, \beta_p)$ given $(\theta_1 \cdots, \theta_r, \delta_1)$ does not depend on β, one can obtain that the conditional reference prior $(\beta_1, \cdots, \beta_p)$ given $(\theta_1 \cdots, \theta_r, \delta_1)$ is a constant. Then the conditional reference prior of δ_1 given $(\theta_1 \cdots, \theta_r)$ is $1/\delta_1$. For $i = 1, \cdots, r$, define

$$\psi_i = \begin{pmatrix} \mathrm{tr}[\mathbf{U}_1] \\ \mathrm{tr}[\mathbf{U}_2] \\ \vdots \\ \mathrm{tr}[\mathbf{U}_i] \end{pmatrix}, \quad \text{and} \quad \mathbf{\Psi}_i = \begin{pmatrix} \mathrm{tr}[\mathbf{U}_1^2] & \mathrm{tr}[\mathbf{U}_1\mathbf{U}_2] & \ldots & \mathrm{tr}[\mathbf{U}_1\mathbf{U}_i] \\ \mathrm{tr}[\mathbf{U}_1\mathbf{U}_2] & \mathrm{tr}[\mathbf{U}_2^2] & \ldots & \mathrm{tr}[\mathbf{U}_2\mathbf{U}_i] \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{tr}[\mathbf{U}_1\mathbf{U}_i] & \mathrm{tr}[\mathbf{U}_2\mathbf{U}_i] & \ldots & \mathrm{tr}[\mathbf{U}_i^2] \end{pmatrix}.$$

Note that the conditional reference prior of θ_i, given $\theta_{i-1} = (\theta_1, \cdots, \theta_{i-1})'$ is

$$\pi^r(\theta_i \mid \theta_{i-1}) \propto \begin{vmatrix} \mathbf{\Psi}_i & \frac{\psi_i}{\delta_1} \\ \frac{\psi_i'}{\delta_1} & \frac{n}{\delta_1^2} \end{vmatrix} \Bigg/ \begin{vmatrix} \mathbf{\Psi}_{i-1} & \frac{\psi_{i-1}}{\delta_1} \\ \frac{\psi_{i-1}'}{\delta_1} & \frac{n}{\delta_1^2} \end{vmatrix} \tag{34}$$

Using the results in (a), one can obtain

$$\pi^r(\theta_i \mid \theta_{i-1}) \propto n\mathrm{tr}[\mathbf{U}_i^2] - (\mathrm{tr}[\mathbf{U}_i])^2, \tag{35}$$

which does not depend on θ_{i-1}. The result then holds.

Appendix C: Proof of Proposition 4

It is easy that L_{**a} is a continuous function of $\theta > \mathbf{0}$. For (a), as $\theta_i \to 0^+$, $\mathbf{\Sigma}_i \to \mathbf{I}_{n_i}$. Thus, $\otimes_{j=1}^r \mathbf{\Sigma}_j \to \mathbf{\Sigma}_{0i}$, $\otimes_{j=1}^r \mathbf{\Phi}_j \to \mathbf{\Phi}_{0i}$, and

$$\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X} \to \otimes_{j=1}^{i-1} \mathbf{X}_j'\mathbf{\Sigma}_j^{-1}\mathbf{X}_j \otimes \mathbf{X}_i'\mathbf{X}_i \otimes_{j=i+1}^r \mathbf{X}_j'\mathbf{\Sigma}_j^{-1}\mathbf{X}_j.$$

Since $\mathbf{\Sigma}_i$ is $n_i \times n_i$, $|\mathbf{\Sigma}_{0i}|^{-1/2} = |\mathbf{\Sigma}_{[-i]}|^{-n_i/2}$ and

$$|\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X}|^{-1/2} \to |\mathbf{X}_i'\mathbf{X}_i|^{-p/(2p_i)} |\mathbf{X}_{[-i]}'\mathbf{\Sigma}_{[-i]}^{-1}\mathbf{X}_{[-i]}|^{-p_i/2}.$$

It is easy to see that $S^2 \to S_{0,i}^2$ as $\theta_i \to 0^+$. Part (a) then follows.

For (b), if $\theta_i \to \infty$, then by Assumption D2, we have

$$|\mathbf{\Sigma}_i| = \nu_i(\theta_i)^{n_i-1} |\mathbf{D}_i| \mathbf{1}_{n_i}' \mathbf{D}_i^{-1} \mathbf{1}_{n_i} (1 + o(1)). \tag{36}$$

For $|\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i|$, firstly, we assume that $\mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i)$. In the following discussion, from Remark 4(iv), we will ignore the $o(1)$ term, that is, we assume that $\mathbf{\Sigma}_i = \mathbf{1}_{n_i}\mathbf{1}_{n_i}' + \nu_i(\theta_i)\mathbf{D}_i$ is also a positive definite correlation matrix. By Lemma 1,

$$\mathbf{\Sigma}_i^{-1} = \frac{1}{\nu_i(\theta_i)} \left(\mathbf{D}_i^{-1} - \frac{\mathbf{D}_i^{-1}\mathbf{1}_{n_i}\mathbf{1}_{n_i}'\mathbf{D}_i^{-1}}{\nu_i(\theta_i) + \mathbf{1}_{n_i}'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} \right). \tag{37}$$

$\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}$ is a vector, so there exists an orthogonal matrix $\mathbf{P}$ such that

$$\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}\mathbf{1}_{n_i}'\mathbf{D}_i^{-1}\mathbf{X}_i = \mathbf{P}' \begin{pmatrix} a_i & \mathbf{0}' \\ \mathbf{0} & \mathbf{O} \end{pmatrix} \mathbf{P},$$

and $a_i = \mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{X}_i\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{1}_{n_i} \geq 0$. With (37), we have

$$\begin{aligned}\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i &= \frac{1}{\nu_i(\theta_i)}\left\{-\frac{\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{X}_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} + \mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i\right\} \qquad (38)\\ &= \frac{1}{\nu_i(\theta_i)}\mathbf{P}'\left\{\begin{pmatrix} -\frac{a_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} & \mathbf{0}' \\ \mathbf{0} & \mathbf{O}\end{pmatrix} + \mathbf{P}\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i\mathbf{P}'\right\}\mathbf{P}.\end{aligned}$$

Since $\boldsymbol{\Sigma}_i^{-1}$ is positive definite and $\mathbf{X}_i$ is full column rank matrix, so are $\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i$ and

$$\begin{pmatrix} -\frac{a_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} & \mathbf{0}' \\ \mathbf{0} & \mathbf{O}\end{pmatrix} + \mathbf{P}\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i\mathbf{P}' \hat{=} \begin{pmatrix} -\frac{a_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} + d_{11} & \mathbf{d}' \\ \mathbf{d} & \mathbf{D}_{22}\end{pmatrix}.$$

Thus, $\mathbf{D}_{22}$ is also positive definite. Since $\mathbf{P}$ is orthogonal, we have

$$|\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i| = \nu_i(\theta_i)^{-p_i}|\mathbf{D}_{22}|\left\{-\frac{a_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} + d_{11} - \mathbf{d}'\mathbf{D}_{22}^{-1}\mathbf{d}\right\}, \qquad (39)$$

and $-\frac{a_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} + d_{11} - \mathbf{d}'\mathbf{D}_{22}^{-1}\mathbf{d} > 0$. By applying Lemma 2 (ii) to the correlation matrix $\mathbf{1}_{n_i}\mathbf{1}'_{n_i} + \nu_i(\theta_i)\mathbf{D}_i$ one can obtain $\nu_i(\theta_i) + \mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i} < 0$, so if $d_{11} - \mathbf{d}'\mathbf{D}_{22}^{-1}\mathbf{d} = 0$, then a_i must be positive and by (39) we have

$$|\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i| = O(\nu_i(\theta_i)^{-p_i}). \qquad (40)$$

If $d_{11} - \mathbf{d}'\mathbf{D}_{22}^{-1}\mathbf{d} \neq 0$, that is, $\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i$ is nonsingular, then by (38) we have

$$|\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i| = \nu_i(\theta_i)^{-p_i}|\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i|\frac{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\mathbf{D}_i,\mathbf{X}_i)\mathbf{1}_i}{\nu_i(\theta_i)+\mathbf{1}'_{n_i}\mathbf{D}_i^{-1}\mathbf{1}_{n_i}}. \qquad (41)$$

Recall $\boldsymbol{\Sigma}_i = \mathbf{1}_{n_i}\mathbf{1}'_{n_i} + \nu_i(\theta_i)\mathbf{D}_i$. Since $\mathbf{X}_i'\{\nu_i(\theta_i)\mathbf{D}_i\}^{-1}\mathbf{X}_i = \nu_i(\theta_i)^{-1}\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i$, which is also nonsingular, by (32) in Lemma 5, we have the following

$$\nu_i(\theta_i)^{-1}\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\mathbf{D}_i,\mathbf{X}_i)\mathbf{1}_{n_i} = \mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}_i,\mathbf{X}_i)\mathbf{1}_{n_i} + \frac{(\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}_i,\mathbf{X}_i)\mathbf{1}_{n_i})^2}{1-\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}_i,\mathbf{X}_i)\mathbf{1}_{n_i}} = \frac{\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}_i,\mathbf{X}_i)\mathbf{1}_{n_i}}{1-\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}_i,\mathbf{X}_i)\mathbf{1}_{n_i}}.$$

By Lemma 4 and the assumption that $\mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i)$, we have $\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\boldsymbol{\Sigma}_i,\mathbf{X}_i)\mathbf{1}_{n_i} \neq 0$ and hence $\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\mathbf{D}_i,\mathbf{X}_i)\mathbf{1}_{n_i} \neq 0$ by above relationship between them. Noting that $\mathbf{1}'_{n_i}\boldsymbol{\Gamma}(\mathbf{D}_i,\mathbf{X}_i)\mathbf{1}_{n_i}$ is independent of θ_i, one can obtain (40).

If $\mathbf{1}_{n_i} \in \mathcal{C}(\mathbf{X}_i)$, then the matrix function in Lemma 3 $f(\mathbf{X}_i) = 1$ and hence $\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i$ is nonsingular and $\mathbf{1}_{n_i}'\mathbf{\Gamma}(\mathbf{D}_i, \mathbf{X}_i)\mathbf{1}_{n_i} = 0$. From (41), we have

$$|\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i| = \nu_i(\theta_i)^{-p_i+1}\frac{|\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i|}{\nu_i(\theta_i) + \mathbf{1}_{n_i}'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}} = O(\nu_i(\theta_i)^{-p_i+1}). \tag{42}$$

For S^2, first we assume that $\mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i)$. If $\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i$ is singular, then for sufficiently large θ_i, with the minor changes in the proof of finding the expression of $|\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i|$, one can show that $\mathbf{X}_i'\mathbf{E}_i\mathbf{X}_i$ is a positive definite matrix. If $\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i$ is nonsingular, then

$$|\mathbf{X}_i'\mathbf{E}_i\mathbf{X}_i| = \frac{|\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i|\mathbf{1}_{n_i}'\mathbf{\Gamma}(\mathbf{D}_i, \mathbf{X}_i)\mathbf{1}_{n_i}}{\mathbf{1}_{n_i}'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}}.$$

Similarly, one can verify that $|\mathbf{X}_i'\mathbf{E}_i\mathbf{X}_i| \neq 0$. That is, $\mathbf{X}_i'\mathbf{E}_i\mathbf{X}_i$ is nonsingular. Therefore, by ignoring the term $o(1)$, we have that S^2 is equal to

$$\nu_i(\theta_i)^{-1}\mathbf{y}'\big(\otimes_{j=1}^{i-1}\mathbf{\Sigma}_j^{-1}\otimes\mathbf{E}_i\otimes_{j=i+1}^{r}\mathbf{\Sigma}_j^{-1} - \otimes_{j=1}^{i-1}\mathbf{\Phi}_j^{-1}\otimes\mathbf{E}_i\mathbf{X}_i(\mathbf{X}_i'\mathbf{E}_i\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{E}_i\otimes_{j=i+1}^{r}\mathbf{\Phi}_j^{-1}\big)\mathbf{y}. \tag{43}$$

By Lemma 6, since $p_j < n_j$, rank$(\mathbf{\Sigma}_j^{-1}) >$ rank$(\mathbf{\Phi}_j)$ and $\mathbf{E}_i \neq \mathbf{O}$, we have

$$S^2 = O(\nu_i(\theta_i)^{-1}). \tag{44}$$

If $\mathbf{1}_{n_i} \in \mathcal{C}(\mathbf{X}_i)$, then $\mathbf{X}_i'\mathbf{D}_i^{-1}\mathbf{X}_i$ is nonsingular and S^2 is given in (43) by replacing $\mathbf{E}_i\mathbf{X}_i(\mathbf{X}_i'\mathbf{E}_i\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{E}_i$ with $\mathbf{E}_i - \mathbf{\Gamma}(\mathbf{D}_i, \mathbf{X}_i)$. The equation (44) holds by the same argument.

Finally, when $\mathbf{1}_{n_i} \notin \mathcal{C}(\mathbf{X}_i)$, using (36), (40) and (44) in (19) results in the conclusion, whereas when $\mathbf{1}_{n_i} \in \mathcal{C}(\mathbf{X}_i)$, using (36), (42) and (44) in (19) results in the conclusion.

Appendix D: Proof of Theorem 1

We prove the posterior propriety under π^{IJ1} only. The others are similar. In this case, $a = 1$. The same notation is used to denote the marginal prior of θ. From Proposition 3, we have

$$\pi^{IJ1}(\theta) \propto \prod_{i=1}^{r}\left\{n_i\mathrm{tr}[\mathbf{U}_i^2] - (\mathrm{tr}[\mathbf{U}_i])^2\right\}^{1/2} \leq n^{1/2}\prod_{i=1}^{r}(\mathrm{tr}[\mathbf{U}_i^2])^{1/2}.$$

Thus, we obtain

$$\int_{(0,\infty)^r} L_{**a}(\theta;\mathbf{y})\pi^{IJ1}(\theta)d\,\theta \leq n^{1/2}\int_{(0,\infty)^r} L_{**a}(\theta;\mathbf{y})\prod_{i=1}^{r}(\mathrm{tr}[\mathbf{U}_i^2])^{1/2}d\,\theta. \tag{45}$$

By Assumption $D5$ and Proposition 4, the integration for the right-hand side of (45) is finite at the interval $(\mathbf{0}, \xi]$, where $\xi = (\xi_1, \ldots, \xi_r)$ and $\xi_i > 0$ and ξ_i can take any positive value.

As $\theta_i \to \infty$, by Assumptions $D2$ and $D3$, we obtain

$$\mathrm{tr}[\mathbf{U}_i^2] = \frac{n\nu_i'(\theta_i)^2}{n_i}\mathrm{tr}([\mathbf{D}_i\mathbf{\Sigma}_i^{-1}]^2)(1+o(1)).$$

Note that

$$\mathbf{\Sigma}_i^{-1} = \frac{1}{\nu_i(\theta_i)}\left(\mathbf{D}_i^{-1} - \frac{\mathbf{D}_i^{-1}\mathbf{1}_{n_i}\mathbf{1}_{n_i}'\mathbf{D}_i^{-1}}{\mathbf{1}_{n_i}'\mathbf{D}_i^{-1}\mathbf{1}_{n_i}}\right)(1+o(1)),$$

so we obtain

$$\mathrm{tr}[\mathbf{U}_i^2] = \left\{\frac{\nu_i'(\theta_i)}{\nu_i(\theta_i)}\right\}^2\left(\frac{n(n_i-1)}{n_i}\right)(1+o(1)) \propto \left\{\frac{\nu_i'(\theta_i)}{\nu_i(\theta_i)}\right\}^2.$$

Let both A and B be subsets of $\{1,\ldots,r\}$ and defined as follows. For A, if $\mathbf{1}_{n_k} \notin \mathcal{C}(\mathbf{X}_k)$, then $k \in A$, and otherwise, $k \notin A$. For B, if $\theta_k > \xi_k$, then $k \in B$, where ξ_k is chosen to be large; otherwise, $k \notin B$. In the last case, we have $0 < \theta_k \leq \xi_k$. Then we have

$$L_{**a}(\theta;\mathbf{y})\prod_{i=1}^{r}(\mathrm{tr}[\mathbf{U}_i^2])^{1/2} \propto \prod_{i\in A\cap B}|\nu_i'(\theta_i)|\nu_i(\theta_i)^{\frac{n}{2n_i}-1}\prod_{i\in\bar{A}\cap B}|\nu_i'(\theta_i)|\nu_i(\theta_i)^{\frac{n}{2n_i}-\frac{p}{2p_i}-1}\prod_{i\notin B}(\mathrm{tr}[\mathbf{U}_i^2])^{1/2}.$$

It is easy to see that the integration in right-hand side of above equation is finite. Therefore, the prior $\pi^{IJ1}(\theta,\delta_1,\beta)$ yields proper posterior distribution.

Acknowledgements

D. Sun's research was supported by the National Science Foundation grants SES-1024080, SES-1260806, and DMS-1007874. The authors would like to thank Professor Paulo for his generosity in providing us with the data for the computer model. The authors also thank the editor, the associate editor, and two anonymous referees for constructive comments on earlier versions of the paper.

BIBLIOGRAPHY

Banerjee, S., Carlin, B. and Gelfand, A. [2004], *Hierarchical modeling and analysis for spatial data*, Boca Raton: Chapman & Hall.

Bayarri, M., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C. H. and Tu, J. [2007], 'A framework for validation of computer models', *Technometrics* **49**, 138–154.

Berger, J. O. and Bernardo, J. M. [1992], On the development of reference priors (Disc: p49-60), *in* 'Bayesian statistics 4. Proceedings of the Fourth Valencia International Meeting', pp. 35–49.

Berger, J. O., De Oliveira, V. and Sansó, B. [2001], 'Objective Bayesian analysis of spatially correlated data', *Journal of the American Statistical Association* **96**, 1361–1374.

Bernardo, J. M. [1979], 'Reference posterior distributions for Bayesian inference (C/R p128-147)', *Journal of the Royal Statistical Society, Series B, Methodological* **41**, 113–128.

De Oliveira, V. [2007], 'Objective Bayesian analysis of spatial data with measurement error', *The Canadian Journal of Statistics* **35**, 283–301.

De Oliveira, V. [2010], Objective Bayesian analysis for Gaussian random fields. in frontiers of statistical decision making and Bayesian analysis, *in* 'In Frontiers of Statistical Decision Making and Bayesian Analysis—In Honor of James O. Berger', pp. 497–511.

Diggle, P. and Ribeiro, P. [2007], *Model-Based Geostatistics*, New York: Springer.

Harville, D. A. [1974], 'Bayeisan inference for variance components using only error contrasts', *Biometrika* **61**, 383–385.

Kitanidis, P. K. [1986], 'Parameter uncertainty in estimation of spatial functions: Bayesian analysis', *Water Resources Research* **22**, 449–507.

Le, N. D. and Zidek, J. V. [1992], 'Interpolation with uncertain spatial covariances: A Bayesian alternative to kriging', *Journal of Multivariate Analysis* **43**, 351–374.

Marshall, A., Olkin, I. and Arnold, B. [2009], *Inequalities: Theory of Majorization and Its Applications (Springer Series in Statistics)*, New York: Springer.

Paulo, R. [2005], 'Default priors for Gaussian processes', *The Annals of Statistics* **33**, 556–582.

Ren, C. [2008], 'A note of the approximation of likelihood function for a spatial model', *Journal of Interdisciplinary Mathematics* **11**, 395–403.

Ren, C. and Sun, D. [2012], 'Objective Bayesian analysis for CAR models', *Annals of the Institute of Statistical Mathematics* **In press**.

Ren, C., Sun, D. and He, Z. [2012], 'Objective Bayesian analysis for a spatial model with nugget effects', *Journal of Statistical Planning and Inference* **142**, 1933–1946.

Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P. [1989], 'Design and analysis of computer experiments (with discussion)', *Statistics Science* **4**, 409–435.

Wakefield, J. C., Gelfand, A. E. and Smith, A. F. M. [1991], 'Efficient generation of random variates via the ratio-of-uniforms method', *Statistics and Computing* **1**, 129–133.

Received 5 October 2011
Accepted 2 April 2013

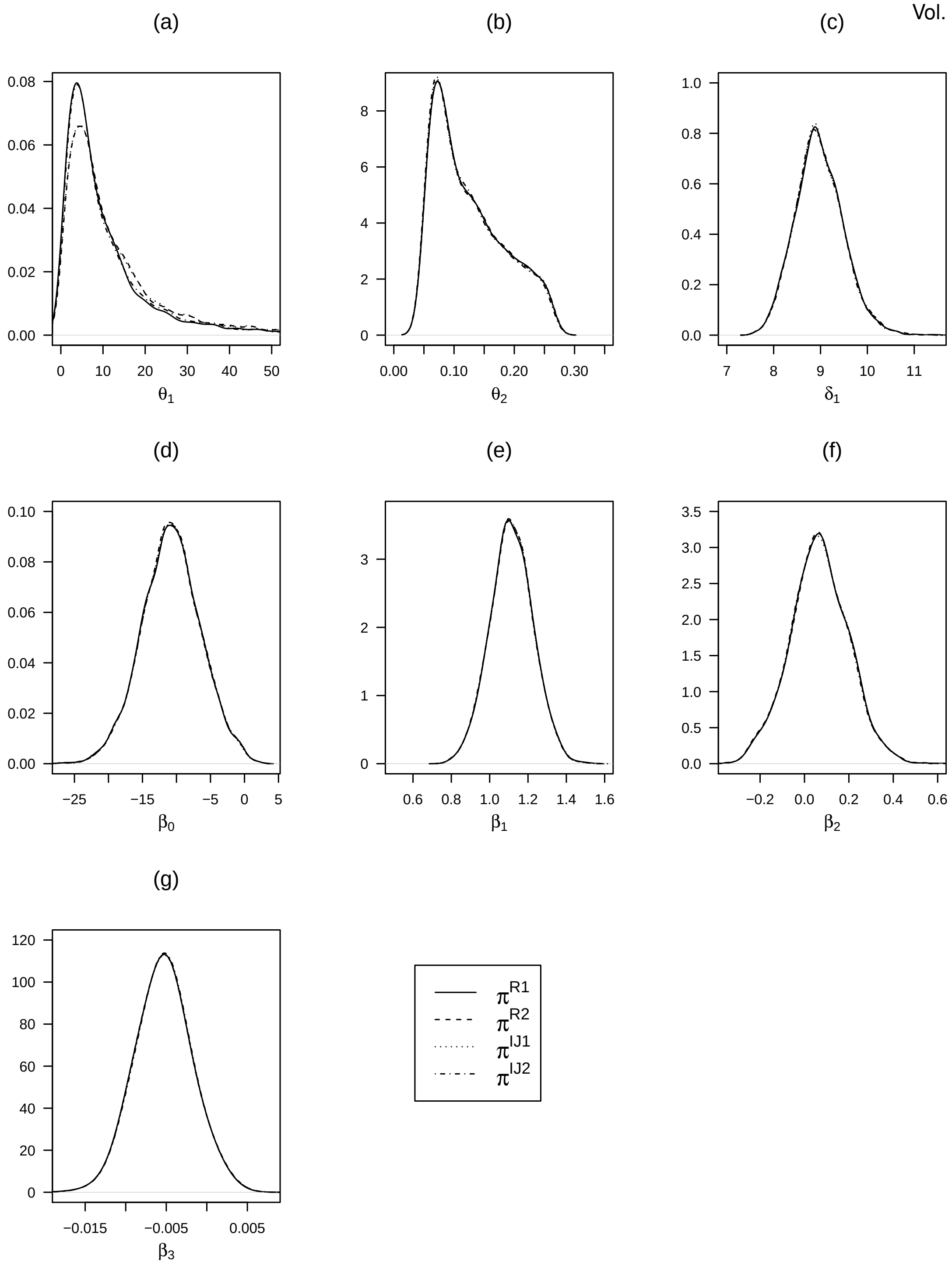

FIGURE 1: The posterior densities of $(\theta_1, \theta_2, \delta_1, \beta_0, \beta_1, \beta_2, \beta_3)$ based on the four priors π^{IJ1} (dotted line), π^{IJ2} (broken line and dots), π^{R1} (solid line), and π^{R2} (dashed line) for the simulated data in Bayarri et al. [2007].