Journal of Sciences, Islamic Republic of Iran 29(2): 173 - 185 (2018) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Bayesian Inference for Spatial Beta Generalized Linear Mixed Models

L. Kalhori Nadrabadi[1, 2], and M. Mohhamadzadeh[*1]

[1] *Department of Statistics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Islamic Republic of Iran*
[2] *Statistical Research and Training Center, Tehran, Islamic Republic of Iran*

Received: 18 October 2017 / Revised: 15 November 2017 / Accepted: 3 January 2018

Abstract

In some applications, the response variable assumes values in the unit interval. The standard linear regression model is not appropriate for modelling this type of data because the normality assumption is not met. Alternatively, the beta regression model has been introduced to analyze such observations. A beta distribution represents a flexible density family on (0, 1) interval that covers symmetric and skewed families. In this paper, a beta generalized linear mixed model with spatial random effect is proposed emphasizing on small values of the spatial range parameter and small sample sizes. Then some models with both fixed and varying precision parameter and different combinations of priors and sample sizes are discussed. Next, the Bayesian estimation of the model parameters is evaluated in an intensive simulation study. Selected priors improved the Bayesian estimation of the parameters, especially for small sample sizes and small values of range parameter. Finally, an application of the proposed model on data provided by Household Income and Expenditure Survey (HIES) of Tehran city is presented.

Keywords: Bayesian estimation; Beta regression model; Household income and expenditure data; Spatial random effect.

Introduction

Regression models have been widely used in statistical analysis when the basic assumptions are satisfied, and normality of response variable is one of the main assumptions. However, in many practical studies, we have encountered data with the realization of response variables lying in (0, 1) interval. There are common examples of rates and proportions, such as unemployment rate, illiteracy rate, fertility rate, the fraction of income spent on food, the proportion of time devoted to an activity, the percentage of a land covered by special vegetation, the proportion of people suffering from cancer, and so forth. So, in these situations, the standard regression models are rather restrictive and inaccurate for modelling large bodies of authentic data with limited range.

A possible solution is to transform the dependent variable in a way that the transformed response follows a normal distribution, and then model the mean of the transformed response. This approach can solve the normality problem, but some new drawbacks may emerge, for instance, the model parameters cannot be

* Corresponding author: Tel: +982182882008; Fax: +982182883017; Email: mohsen_m@modares.ac.ir

easily interpreted in terms of the original response. On the other hand, measures of proportions may be asymmetric hence any inference based on the assumption of normality results in the departure from reality. In order to overcome these drawbacks, [1] introduced the beta regression model which is suitable for modelling response variables restricted to the (0, 1) interval. In this frequentist approach, the beta distribution is reparametrized in terms of its mean and a positive parameter that can be regarded as a precision parameter. They have linked the mean parameter to a regression structure while assuming that the precision parameter is fixed. Likelihood-based inference in beta regression may be misleading for small sample sizes. A well-adjusted likelihood ratio statistics for small sample sizes was introduced by [2].

A Bayesian approach for modelling both the mean and the precision parameter, which has been linked to a linear regression structure through logit and logarithm link functions was proposed by [3,4] respectively. Under the Bayesian paradigm, [5] implemented a semiparametric beta regression model using penalized splines to study the proportion of nucleotides that differ by a given sequence or gene. Incorporation of a nonlinear regression structure to the mean model is developed by [6], as well as a regression structure for the precision parameter which may also be nonlinear. [7] added a random effect in the mean model of beta regression and applied it to study the reaction time of old people in a longitudinal study. Mixed beta regression models for both the mean and precision parameters were proposed by [8]. Both maximum likelihood and Bayesian MCMC mixed beta regression models were elaborated by [9]. Recently [10] proposed a partially linear model with correlated disturbances from a Bayesian perspective for modelling Brazilian and Chilean monthly unemployment rate.

A new class of spatial models based on the biparametric exponential family of distributions proposed by [11], in which the spatial effect was included in the model through the distance of points as an explanatory variable. This model was applied to study the quality of education in Columbia by [12]. Gholizadeh et al., ([13]) have developed a spatial analysis of structured additive regression model using Integrated Nested Laplace Approximation (INLA) and modelled crime rate data in Tehran as an application of their model. On the other hand, [14] proposed a Bayesian approach on beta regression with spatial dependence structure given by exponential covariance function and has suggested a square beta as prior distribution both for spatial range and spatial variance, i.e.$(aB)^2$, where B ~Beta (1 + ε, 1 + ε), for given positive values of a and ε. However, the Fustos approach overestimates the small values of the spatial range. In addition, [15] worked on spherical covariance function assumed that the spatial range parameter gets large values. Recently, [16] introduced a spatial beta regression model in which the correlation existing in the data was considered through an explanatory random variable in the mean model.

This paper proposes a Bayesian analysis for the spatial beta generalized linear mixed model with a new prior elicitation for the spatial dependence structure, emphasizing on small values of the spatial range parameter and small sample sizes. We consider two cases in turn. First, we assume that the precision parameter is fixed; next, we suppose that it is varying over observations and for both cases, we consider different scenarios for parameter estimations. This proposal is evaluated through Markov Chain Monte Carlo (MCMC) experiments and implemented via Gibbs sampling. The Multivariate Proportional scale Reduction Factor (MPRF) by [17] and [18] tests were used to check the convergence of the Gibbs samplers. Sensitivity analysis implementing more non-informative priors also declares our proposed priors are reliable. The Household Income and Expenditure Survey (HIES) is one of the main surveys which is conducted annually by the Statistical Center of Iran (SCI). The information provided by this survey is used for calculation of poverty line and national accounts. We calculate the proportion of monthly expenses spent on food to the entire expenditure of a household. This proportion can be used for understanding the welfare situation of a household. We aim to study the factors that affect this proportion as our response variable and Deviance Information Criterion (DIC) is used for models evaluations. This paper is arranged in the following order, the material and method part includes five sections. In the first section, the beta regression model is reviewed. Then, the beta generalized linear mixed model with spatial random effects is introduced in section two. In the third section, the motivation of selecting the priors along with model fitting by using Gibbs sampling method are presented. A simulation study is also performed for model evaluations in the fourth section. In the fifth section, we illustrated how to apply the proposed model to a real data set. Finally, the paper is closed with discussion and results.

Materials and Methods

Beta Regression Model

The Beta distribution is very flexible and adapts different shapes in regard to values of the parameters.

For instance, Beta (1, 1) is equivalent to U(0, 1), other shapes involve the J shape, U shape, symmetric and skewed shapes. If Y~Beta(a, b) then its probability density function is given by

$$f(y, a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}, 0 < y < 1,$$

where a>0, b>0 and $\Gamma(\cdot)$ is the gamma function. The mean and variance of Y are given by $E(Y) = \frac{a}{a+b}$ and $Var(Y) = \frac{ab}{(a+b)^2(1+a+b)}$. In the sake of modeling the mean parameter, [8] introduced a reparametrized beta density in a way that E(Y) $=\mu$ and $Var(Y) = \frac{\mu(1-\mu)}{\phi+1}$. In this case a $=\mu\phi$ and b$=(1-\mu)\phi$. Since ϕ is inversely related to the variance of Y, it can be interpreted as a precision parameter. Obviously for a fixed value of μ, larger values of ϕ result in smaller values of variance. The density function of the reparametrized Beta distribution is given by

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1} \quad 0 < y < 1$$

where 0 <μ < 1 and ϕ > 0. Situations where the response is limited to a known interval (c, d) are also accommodated through the transformation $y^* = \frac{(y-c)}{(d-c)}$ where c, d > 0.

The Beta density function (1) is due to Ferrari and Cribari-Neto's (2004) aim for modelling the mean parameter of the Beta distribution. If $Y_1, \ldots, Y_n$ are independent random variables from Beta$(\mu_i\phi, (1-\mu_i)\phi)$, i=1,…, n, then the beta regression model can be defined as

$$g(\mu_i) = \sum_{j=0}^{p-1} x_{ij}\beta_j,$$

where $\beta = (\beta_0, \ldots, \beta_{p-1})^T$ is a vector of regression coefficients, x_{ij} is the known value of covariate j for sample unit i, and $g(\cdot)$ is a continuous twice differentiable link function. In this model the parameter ϕ is assumed to be fixed over all observations.

The reparametrized Beta distribution has two parameters, the mean and the precision parameter. Modelling both parameters of the Beta distribution have been proposed by [3], [4] and [6]. If $Y_1, \ldots, Y_n$ are independent random variables from Beta$(\mu_i\phi_i, (1-\mu_i)\phi_i)$, i=1,…,n, then

$$h(\phi_i) = \sum_{\ell=0}^{m-1} z_{i\ell}\delta_\ell \quad (2)$$

as before $\delta = (\delta_0, \ldots, \delta_{m-1})^T$ is the vector of regression parameters, $z_{i\ell}$ known as a value of covariate ℓ for the sample unit i, and $h(\cdot)$ is a continuous twice differentiable link function. In the existing literature, the common link function for ϕ is the logarithm function. Note that covariates for modelling the mean and precision parameters could be exactly the same, completely different, or a combination of similar and different features.

Spatial Beta Generalized Linear Mixed Model

When the location of sampling units affects the response variable, there would be a spatial correlation. Initially, [11] introduced spatial beta regression model in which the spatial dependency was captured through an explanatory variable, which was a multiple of the response variable and corresponding spatial weights. Their model is given by

$$\text{logit}(\mu_i) = x_i\beta + \rho WY$$

$$\log(\phi_i) = z_i\delta + \lambda WY,$$

where $x_i=(x_{i0}, \ldots, x_{i(p-1)})$ and $z_i=(z_{i0}, \ldots, z_{i(m-1)})$ are vectors of non-stochastic regressors. In addition, ρ and λ are regression parameters, W is the spatial weight matrix and y = $(y_1, \ldots, y_n)$ represents the vector of response variables. The structure of the spatial correlation is not considered in the [11] model.

In this paper, we aim to incorporate the spatial correlation structure of the response variable into the model. This can be achieved by adding a random component to the mean model. Models of this kind belong to the class of Spatial Generalized Linear Mixed Models (SGLMM) introduced by [19]. Suppose y(s) = (y(s_1),…, y(s_n)) are realizations of random variables Y(s)= (Y(s_1),…,Y(s_n)) at n distinct locations s_1,…, s_n. For simplicity assume y_i and Y_i denote $y(s_i)$ and $Y(s_i)$ respectively, in the following parts. We are interested in modelling beta distributed spatially dependent random variables, Y_1,…, Y_n, where their geographical associations are referenced by a Gaussian Random Field (GRF). Without loss of generality, the term can be extended for modelling spatially correlated response variables, restricted to the known positive interval (c, d). Let the random vector $\tau(s) = (\tau(s_1), \ldots, \tau(s_n))$ denote a GRF. Conditionally on $\tau(s)$, the spatial random fields Y(s) are independent and beta distributed, i.e., $Y(s)|\tau(s) \sim \text{Beta}(\mu(s)\phi(s), (1-\mu(s))\phi(s))$, which is obtained by replacing μ and ϕ with μ(s) and ϕ(s) in (1). Then, using the idea of SGLMM ([19]), we define Spatial Beta Generalized Linear Mixed Model (SBGLMM) as follows; which is a linear function both in the fixed effects and random effect

$$g(E(y_i|\tau_i)) = g(\mu_i) = x_i\beta + \tau_i = \eta_i \quad i = 1, \dots, n \ (3)$$

where $x_i = (x_{i0}, \dots, x_{i(p-1)})$ and $\beta = (\beta_0, \dots, \beta_{p-1})^T$ are vectors of covariates and related regression coefficients respectively, and τ_i denote $\tau(s_i)$ is the random effect capturing the spatial correlation. The random process $\tau = (\tau_1, \dots, \tau_n)$ follows a multivariate normal distribution, i.e., $\tau(s) \sim N_n(0, \Sigma_\tau)$, where

$$(\Sigma_\tau)_{ij} = \text{Cov}\left(\tau(s_i), \tau(s_j)\right) = \sigma^2 \text{Corr}\left(\tau(s_i), \tau(s_j)\right),$$

and describe the spatial correlation structure. Following the previous studies we choose the logit link function, so (3) can be rewritten as below

$$\text{logit}(\mu_i) = x_i\beta + \tau_i \quad i=1,\dots, n. \qquad (4)$$

The precision parameter ϕ can also be modelled using a suitable link function and a linear predictor, e.g. $h(\phi_i) = z_i\delta = \omega_i$. According to existing studies on beta regression models, a suitable link function would be the logarithm link function. So, we assume

$$\log(\phi_i) = z_i\delta = \omega_i \qquad i = 1, \dots, n;$$

where $z_i = (z_{i0}, \dots, z_{i(m-1)})$ is the vector of covariates and $\delta = (\delta_0, \dots, \delta_{m-1})^T$ is the vector of regression parameters. In cases where ϕ is fixed, the corresponding model is readily obtained by assuming $\phi_i = \phi$, $i = 1, \dots, n$, $\delta = \delta_0$, and $z = 1$. In the following section, we will consider SBGLMM in two cases and intend to estimate the parameters involved in the proposed model using the Bayesian approach. First, we assume that the precision parameter is fixed and then assume a linear structure for the varying precision parameter as well.

Bayesian Estimation of the Model Parameters

Consider the spatial beta regression model given by

$$y_i|\beta, \phi_i, \tau_i \sim Beta(\mu_i\phi_i, (1-\mu_i)\phi_i), i = 1, \dots, n;$$

$$\tau|\upsilon \sim N_n(0, \Sigma_\tau) \qquad (5$$

where υ is the vector of structural parameters related to the spatial covariance function and $\text{logit}(\mu_i) = x_i\beta + \tau_i = \eta_i$ as in (4). The precision parameter ϕ_i could be fixed or varying. We will discuss both cases in turn.

In order to complete the Bayesian specification of the spatial beta regression model, elicitation of prior distributions for all unknown parameters is required. Multivariate normal prior distributions are typically considered for the regression coefficients involved in the mean model, i.e., we propose $\beta \sim N_p(0, \Sigma_\beta)$, where Σ_β is a diagonal matrix with large values of variances, which provides a vague prior. In the Bayesian context, a popular choice for the prior distribution of the variance would be inverse gamma distribution. Since ϕ is the precision parameter and inversely related to the variance of the Beta distribution, it can be assumed that ϕ is gamma distributed with small positive values of parameters to avoid using an informative prior. In the case of a varying dispersion parameter ϕ_i as in (2), we have specified a convenient prior for fixed effects given by $\delta \sim N_m(0, \Sigma_\delta)$, where Σ_δ is a diagonal matrix with large values of variance components.

We assume that the spatial dependence between two geographical points is given by the exponential correlation function i.e.,

$$\rho(d, \psi) = \exp\left(-\frac{d}{\psi}\right) \quad (6)$$

where ψ^{-1} is the spatial range and $d = |s_i - s_j|$ denotes the distance between sample units i and j which are sited at locations s_i and s_j respectively, for $i, j=1, \dots, n$. Therefore $(\Sigma_\tau)_{ij} = \sigma^2 \exp(-\psi^{-1} d)$ and σ^2 describe the spatial variance. Thus, through this study we assume that $\upsilon = (\sigma^2, \psi^{-1})$.

In this paper, we focus our attention on the exponential covariance function, which has been used in various applications [20]. We have not yet investigated the estimation and properties of Matern covariance models [21], which include the exponential model as a special case, and this issue needs further research. Note that the variance component of the exponential covariogram is called spatial variance and its inverse will be called spatial precision in the following parts.

An inverse gamma distribution is set as prior for spatial variance σ^2 which is a typical prior for variance in the Bayesian context. The range parameter is commonly given an inverse gamma or a bounded uniform prior distribution. The bounded interval is essential to avoid improper priors. Because using improper prior distributions for the parameters of a GRF without attention may result in improper posteriors, [22].

The joint posterior distribution is obtained by combining the likelihood function of beta distribution (1) with the prior information. We now present the joint posterior distribution for the varying precision parameter model and omit the observable vectors x and z in the notation since these are non-random and already known. Let $y = (y_1, \dots, y_n)^T$ be the observed spatial variable, under the assumption that the parameters

Box-1

$$f(\beta,\delta,\alpha,\sigma^2,\lambda,\tau\,|\,\mathbf{y}) \propto \prod_{i=1}^{n} f(y_i\,|\,\tau_i,\beta,\delta)\prod_{i=1}^{n} f(\tau_i\,|\,\beta,\alpha,\sigma^2,\lambda)\times f(\beta)f(\alpha)f(\sigma^2\,|\,\lambda)f(\lambda)f(\delta).$$

β, δ, σ^2 and ψ^{-1} are prior independent, the joint posterior density is given by

$$f(\beta,\delta,\sigma^2,\psi^{-1},\tau\,|\,\mathbf{y}) \propto \prod_{i=1}^{n} f(y_i\,|\,\tau_i,\beta,\delta)\prod_{i=1}^{n} f(\tau_i\,|\,\beta,\sigma^2,\psi^{-1}) \times f(\beta)f(\sigma^2)f(\psi^{-1})f(\delta).$$

Note that, in cases where the precision parameter ϕ is constant, the posterior distribution is obtained by replacing prior distributions of δ with ϕ. The joint posterior distribution is complicated, so the Gibbs sampler [23] can be utilized to generate samples from the joint posterior density. The Gibbs sampler in this context involves iteratively sampling from the full conditional distributions, and this procedure is implemented by means of a MCMC scheme.

Next, in order to evaluate the performances of the specified priors, a simulation study for the fixed precision parameter case was conducted. Outcomes of the simulation study reveal that the above mentioned inverse gamma and uniform priors for the spatial range provide overestimation. Regarding this issue, we tried to choose a prior distribution for the transformed parameter. Suppose that the spatial range is the growing amount of a quantity, this can be reduced by logarithm transformation. Therefore a uniform distribution is utilized as prior distribution for logarithm of the spatial range according to [22]. Assuming that $\psi^{-1} \in (0.01, 10)$, then a bounded uniform prior on the logarithm of the range is given by $\alpha = -\mathrm{Ln}(\psi) \sim U(-4.5, 2.3)$.

Implementing an inverse gamma prior for the spatial variance led to a slight underestimation. Therefore we tried to find another prior to achieve the desired improvement in our estimation. Based on the idea of using a prior distribution that has the property of rapid growing values, we assume that σ^2 is following an exponential distribution, i.e., $\sigma^2 \sim \exp(\lambda)$. But there is no clue about the value of λ, so we set a hierarchical prior and assume that $\lambda \sim N(0, \sigma_\lambda^2) I(0, \infty)$, where σ_λ^2 takes large values to have a non-informative prior. Note that the truncated normal distribution was chosen due to the support of the exponential distribution parameter. Results of the simulation over these recent priors and typical priors for spatial parameters will be presented simultaneously in the next section.

Implementing the new proposed priors under the assumption that the parameters $\beta, \delta, \alpha, \sigma^2$ and λ are priori independent, the joint posterior density is given by (Box-1).

To deal with this complicated posterior distribution, again the Gibbs sampler has been applied to generate samples from the full conditional distributions which are presented in Appendix I. Posterior inferences on $\beta, \delta, \alpha, \sigma^2$ and λ are readily obtained using WinBUGS through the R2WinBUGS package [24] in R [25]. When conditional distributions are nonstandard, it can't sample directly from them using Gibbs sampling. Therefore, Metropolis-within-Gibbs algorithm is implemented by WinBUGS to sample from difficult full conditional distributions.

Hypothesis testing regarding the regression coefficients and mean responses are also straightforward. The program codes are available from the authors upon request. When the MCMC implementation is applied to the simulated data (see the Simulation Study section), the convergence of the MCMC samples is assessed using standard tools within WinBUGS such as trace plots and autocorrelation function (ACF) plots, as well as the Gelman-Rubin [17] convergence diagnostic.

Simulation Study

In this section, we study through some intensive simulation experiments, the behavior of the Bayesian estimators based on the square root of MSE and relative bias. We performed the simulation of the spatial beta regression model by assuming a GRF whose covariance structure is given by (6) while incorporating different scenarios for the spatial parameter. Consider the model

$$y_i|\beta,\phi_i,\tau_i \sim Beta(\mu_i\phi_i,(1-\mu_i)\phi_i), \qquad i=1,\dots,n,$$
$$\tau|\sigma^2,\psi^{-1} \sim N_n(0,\Sigma_\tau),$$

where $\beta = (\beta_0, \beta_1, \beta_2)^T$ and $\log = \left(\frac{\mu_i}{1-\mu_i}\right) = x_i\beta + \tau_i = \eta_i$. The values of the covariates x_{1i} and x_{2i} were generated from a uniform distribution on the unit interval referring to [8], and we set, $\beta = (-1, 2, -1.5)^T$ and ϕ=50. On the other hand, we consider the spatial variance $\sigma^2 = 0.5$ and different spatial range settings; ψ^{-1} =0.1, 0.45 and 0.9. These values are considered after an initial study of the spatial parameters and we find out that the natural prior for the range parameter overestimates the aimed parameter in the case of spatial beta regression, so we attempt to estimate small values of the range.

Additionally, we consider different settings for the precision parameter ϕ_i, generating two possible models.

Model 1: ϕ_i=ϕ,

Model 2: $\log(\phi_i) = \delta_0 + \delta_1 z_{1i}$

In order to evaluate our proposal priors, a simulation study was conducted for two sets of priors, and parameters of prior distributions were specified in a way that would provide noninformative priors. Priors for each scenario are defined as follow.

Set 1. $\beta_j \sim N(0,100)$ for *j*=0,1,2, $\phi \sim Gamma(0.01,0.01)$ whereas for spatial parameter we have $\sigma^2|\lambda \sim exp(\lambda), \lambda \sim N(0, \sigma_\lambda^2)I(0, \infty)$. We attempt to find a prior for small values of the range and have our aimed value for ψ^{-1} lying in the (0.01, 10) interval, therefore $\alpha = -Ln(\psi) \sim U(-4.5, 2.3)$.

Set 2. $\beta_j \sim N(0,100)$ for j=0,1,2, $\phi \sim Gamma(0.01,0.01)$ with spatial parameter given by $\sigma^{-2} \sim Gamma(0,0.01)$ and $\psi^{-1} \sim U(0.01, 10)$.

To illustrate the effect of sample size on Bayesian estimators, we consider different locations defined on a regularly spaced grid, which consist of [0,5]× [0,5], [0, 10]×[0, 10] and [0, 15] × [0, 15].

First, we study the model with a fixed precision parameter (Model 1). For the generated data set of size 225, we simulate one chain of size 100,000 for each parameter, disregarding the first 50,000 iterations to eliminate the effect of the initial values. Further, to avoid correlation problems we considered spacing of size 100. Thus we obtained an effective sample of size 500, and the posterior inference is based on this. The chains are mixed slowly especially for the intercept of the model,β_0, so we need to consider more iterations for smaller sample sizes to reach convergence.

Diagnostic tests of convergence were implemented and there was no evidence of divergence of the chains. To validate this assertion, we have used the multivariate proportional reduction factor. Two chains with different random initial values were generated simultaneously. The resulted MPRF is equal to 1.02 which is lower than 1.2, indicating that the chains are convergent. Geweke's statistics [26] also show the convergence of the results of Gibbs samplers for each parameter. The reference diagnostics were done using the Coda package in R. Plots of autocorrelation in Figure 1 show independence between distinct replication for estimating parameters.

Figure 2 displays another tool employed to assess the Markov chain convergence on all parameters. Trace

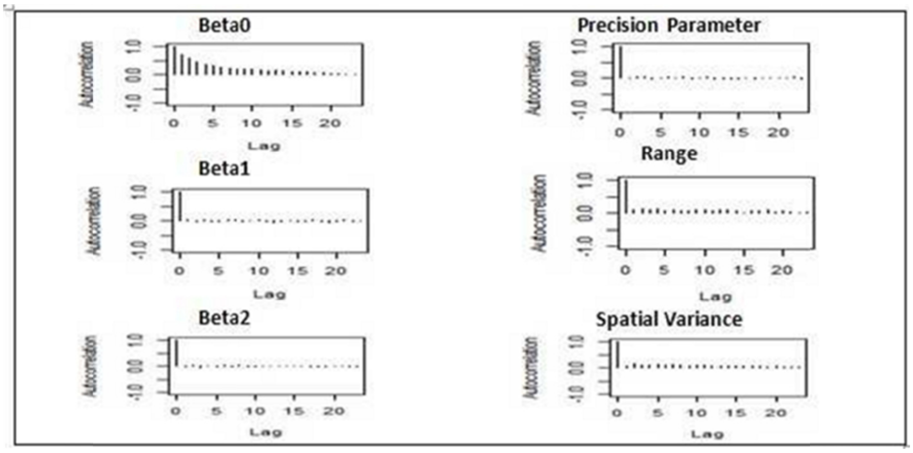

Figure 1. ACF plots of parameters

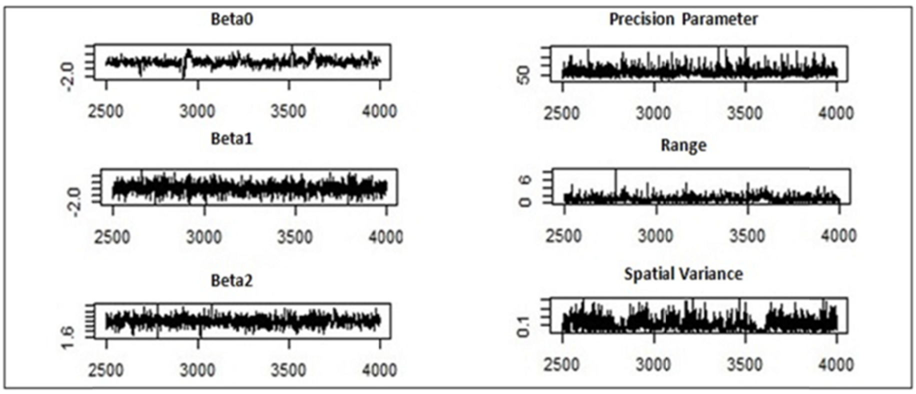

Figure 2. Trace plots of the estimated parameters

plots showed that the chains have a stable performance around the true parameter value.

We computed the relative bias (RelBias) and the square root of MSE (RMSE) for each parameter over the 50 simulated samples. They are defined as

$\mathrm{RelBias}(\theta)=\frac{1}{50}\sum_{i=1}^{50}\left(\frac{\hat{\theta}_i}{\theta_i}-1\right)$ and $\mathrm{RMSE}(\theta)=\left\{\frac{1}{50}\sum_{i=1}^{50}\left(\hat{\theta}_i-\theta_i\right)^2\right\}^{1/2}$

Where $\theta=(\beta_0,\beta_1,\beta_2,\phi,\sigma^2,\psi^{-1})$ and $\hat{\theta}_i$ is the posterior estimate of θ_i for the ith sample. Table [1] presents the summary results for the estimation of all the parameters. It seems that the spatial variance is overestimated or underestimated according to the first and second sets of priors respectively. Nevertheless, it is worth mentioning that using a hierarchical prior for spatial variance, resulted in a reduction in the RelBias and RMSE values when the sample size is increased. Moreover, bounded uniform distribution on the logarithm of the range resulted in proper estimates. This prior has solved the problem of over estimation and has also reduced the RelBias and RMSE in comparison with uniform prior density, especially for small sample sizes.

Another important aspect that should be assessed is the performance of the Bayesian estimators for other values of the spatial range when the sample size is small. The data generation scheme is similar to the simulation study described above, whereas other values for the range parameter are considered. Table 2 summarizes the numerical results of our simulation study over a 5×5 lattice, where the RelBias and the RMSE of the parameter estimators of Set 1 are smaller than Set 2. Hence we can conclude that our method

Table 1. Estimates of the parameters of Model 1 for different priors and sample sizes

			Prior Set 1				**Prior Set 2**			
n	Par.	Value	Mean	Bias	Sd.	RMSE	Mean	Bias	Sd.	RMSE
	β_0	-1	-1.043	0.043	0.191	0.442	-1.156	0.156	0.212	0.263
	β_1	2	2.093	0.046	0.215	0.483	2.225	0.112	0.219	0.313
25	β_2	-1.5	-1.542	0.028	0.219	0.472	-1.383	0.078	0.242	0.269
	ϕ	50	60.527	0.211	17.472	4.516	75.281	0.506	22.782	34.032
	σ^2	0.5	0.724	0.449	0.323	0.627	0.503	0.006	0.542	0.541
	ψ^{-1}	0.1	0.218	1.183	0.072	0.372	9.353	92.532	5.873	10.959
	β_0	-1	-0.912	0.088	0.167	0.189	-1.173	0.173	0.540	0.567
	β_1	2	2.075	0.038	0.165	0.181	2.022	0.011	0.102	0.104
100	β_2	-1.5	-1.528	0.019	0.121	0.124	-1.513	0.009	0.191	0.192
	ϕ	50	64.779	0.295	23.665	27.90	67.319	0.346	20.107	26.538
	σ^2	0.5	0.685	0.370	0.235	0.298	0.311	0.376	0.242	0.306
	ψ^{-1}	0.1	0.161	0.616	0.052	0.080	1.047	9.475	2.065	2.272
	β_0	-1	-0.930	0.069	0.338	0.345	-0.922	0.077	0.382	0.389
	β_1	2	1.972	0.013	0.067	0.072	1.99	0.005	0.106	0.106
225	β_2	-1.5	-1.486	0.01	0.111	0.112	-1.518	0.012	0.098	0.100
	ϕ	50	59.612	0.192	12.018	15.389	65.347	0.307	16.728	22.702
	σ^2	0.5	0.579	0.159	0.130	0.152	0.376	0.246	0.205	0.239
	ψ^{-1}	0.1	0.158	0.586	0.039	0.070	0.252	1.525	0.133	0.142

Table 2. Estimates of the parameters of Model 1 for different priors and different values of spatial range

		Prior Set 1				**Prior Set 2**			
Par.	Value	Mean	Bias	Sd.	RMSE	Mean	Bias	Sd.	RMSE
$\boldsymbol{\beta_0}$	-1	-0.903	0.097	0.223	0.243	-1.696	0.131	0.166	0.257
$\boldsymbol{\beta_1}$	2	1.939	0.030	0.266	0.273	2.792	0.396	0.262	0.835
$\boldsymbol{\beta_2}$	-1.5	-1.568	0.045	0.218	0.228	-0.901	0.399	0.215	0.636
$\boldsymbol{\phi}$	50	33.233	0.335	9.899	19.470	44.288	0.114	11.558	12.893
$\boldsymbol{\sigma^2}$	0.5	0.637	0.273	0.263	0.296	0.123	0.754	0.059	0.382
$\boldsymbol{\psi^{-1}}$	0.45	0.361	0.199	0.166	0.188	14.046	30.213	3.914	14.148
$\boldsymbol{\beta_0}$	-1	-0.814	0.186	0.282	0.338	-0.798	0.202	0.191	0.279
$\boldsymbol{\beta_1}$	2	1.532	0.234	0.359	0.590	1.932	0.034	0.177	0.190
$\boldsymbol{\beta_2}$	-1.5	-1.354	0.097	0.362	0.390	-1.713	0.142	0.247	0.326
$\boldsymbol{\phi}$	50	37.112	0.258	13.577	18.720	31.347	0.373	8.282	20.409
$\boldsymbol{\sigma^2}$	0.5	0.949	0.898	0.606	0.754	0.095	0.810	0.002	0.405
$\boldsymbol{\psi^{-1}}$	0.9	0.819	0.091	0.330	0.340	22.303	22.477	3.804	21.690

exhibits good performances in the estimation of the range parameter.

To explore how Bayesian estimates are affected by less informative priors, a sensitivity analysis was conducted. So applying vague priors, we use the following elicitation of prior distributions $\beta_j \sim N(0,1000)$ for j=0,1,2, $\phi \sim Gamma(0.001,0.001)$ and $\lambda \sim N(0,1000)I(0,\infty)$.

The results of this assessment are given in Table [3]. It can be observed that estimation of the parameters is not significantly affected by the use of less informative priors. In conclusion, the sensitivity analysis, convergence tests, plots and results over the 50 simulated data sets indicate that our results are reliable and can be applied to analyze real data sets when the basic circumstances of the model are met.

After investigating Model 1, we expended considerable effort to develop an SBGLMM model, assuming that the precision parameter is not fixed and then going on to examine different scenarios for parameter estimation. Consider the spatial beta regression model presented in (5) and suppose that $\log(\phi_i) = z_i\delta$. As mentioned before, in Model 2 we have $\log(\phi_i) = \delta_0 + \delta_1 z_{1i}$. Table 4 shows the outcomes of our simulation study for Model 2 assuming $\delta_j \sim N(0,100)$ for j = 0, 1 and a prior distribution for other parameters are the same as when assuming the precision parameter is fixed. In this state, we consider two samples in the regular grid, namely, 10 ×10 and 15× 15, then RelBias and RMSE of each parameter are computed over the 50 simulated samples.

From Table 4, we find out that δ_0 and δ_1 are not estimated properly for the sample of size 100. In order to obtain better estimates for these parameters, we examine different priors such as non-informative uniform prior distribution for each δ_j, and a zero mean t-student distribution with large variance and 3 degrees of freedom [27], i.e., we assume that $\delta_j \sim t(0,100,3)$. Our motivation to use these priors was implementing a flat prior and a prior with heavier tails in comparison with normal distribution in order to treat extreme values if they exist. These suggested priors were not achieving proper estimates, so for the sake of brevity, outcomes are not presented here.

[8] utilized an exponential prior for degrees of freedom in t-student distribution, while setting a hierarchical prior for the random effect in the mixed beta regression model. Borrowing this idea for fixed effects, we assume a hierarchical prior for regression coefficients, that is $\delta_j \sim t(0,100,\varsigma_j)$ where $\varsigma_j \sim \exp(0.1)$. Considering these priors the joint posterior distribution is given by (Box-2).

Table 3. Summary results of sensitive analysis for estimation of the parameters of Model 1

		***n* = 100**				***n* = 225**			
Par.	Value	Mean	Bias	Sd.	RMSE	Mean	Bias	Sd.	RMSE
β_0	-1	-0.830	0.170	0.337	0.377	-1.042	0.042	0.152	0.158
β_1	2	2.067	0.033	0.161	0.175	2.040	0.020	0.125	0.132
β_2	-1.5	-1.551	0.034	0.130	0.140	-1.509	0.006	0.123	0.123
ϕ	50	90.146	0.803	39.974	56.654	80.184	0.604	39.184	49.462
σ^2	0.5	0.679	0.357	0.218	0.282	0.411	0.178	0.187	0.207
ψ^{-1}	0.1	0.364	2.643	0.213	0.340	0.192	0.923	0.087	0.127

Table 4. Estimates of the parameters of Model 2 for different priors and sample sizes

			Prior Set 1				**Prior Set 2**			
n	Par.	Value	Mean	Bias	Sd.	RMSE	Mean	Bias	Sd.	RMSE
	β_0	-1	-0.999	0.0004	0.241	0.240	-1.047	0.047	0.168	0.174
	β_1	2	2.016	0.008	0.162	0.163	2.069	0.034	0.169	0.182
100	β_2	-1.5	-1.463	0.025	0.187	0.191	-1.468	0.021	0.169	0.171
	δ_0	4	5.819	0.454	3.080	3.577	6.154	0.538	2.278	3.135
	δ_1	-0.4	1.780	5.451	3.116	3.803	0.053	0.053	3.180	3.212
	σ^2	0.5	0.463	0.072	0.235	0.237	0.382	0.234	0.325	0.345
	ψ^{-1}	0.1	0.663	5.629	1.058	1.198	1.501	14.016	1.422	2.538
	β_0	-1	-0.701	0.299	0.245	0.387	-0.994	0.006	0.322	0.322
	β_1	2	2.013	0.007	0.083	0.084	2.066	0.033	0.112	0.130
225	β_2	-1.5	-1.525	0.016	0.080	0.084	-1.491	0.006	0.068	0.069
	δ_0	4	4.013	0.003	0.282	0.283	4.385	0.096	0.493	0.626
	δ_1	-0.4	-0.390	0.026	0.406	0.406	-0.159	0.602	1.222	1.246
	σ^2	0.5	0.510	0.020	0.224	0.224	0.402	0.196	0.078	0.125
	ψ^{-1}	0.1	0.188	0.880	0.083	0.121	0.234	1.340	0.092	0.163

Box-2

$$f(\beta,\delta,\alpha,\sigma^2,\lambda,\zeta,\tau\,|\,\mathbf{y})\propto$$
$$\prod_{i=1}^{n} f(y_i\,|\,\tau_i,\beta,\delta)\prod_{i=1}^{n} f(\tau_i\,|\,\beta,\alpha,\sigma^2,\lambda)\times f(\beta)f(\alpha)f(\sigma^2\,|\,\lambda)f(\lambda)f(\delta\,|\,\zeta)f(\zeta).$$

Posterior inference is made up by applying the Gibbs sampling method on related full conditional distributions. Since these latest prior distributions do not work properly for the sample of sizes 100, only the simulation results for n = 225 are presented in Table [5]. Values of posterior mean, RelBias and RMSE indicate that the suggested priors are able to estimate δ appropriately.

From Table 4 and 5, it can be seen that by applying Normal, t-student and hierarchical t-student prior distributions for δ, the obtained estimates are close to the corresponding true values of the parameter. In addition, RelBias and RMSE are relatively similar. Therefore, it can be concluded that these prior distributions do not make significant differences in estimation of δ. Accordingly, it is suggested to apply the Normal distribution for the sake of simplicity. Finding a suitable prior for regression coefficients of the precision parameter model requires further research. The results of the sensitivity analysis indicate that the proposed priors can be used for estimation of the parameters involved in an SBGLMM. Hence, these priors are implied to analyze a real data set provided by HIES as an application of our proposed model.

Application on HIES data

As an application, we consider the information available from the Household Income and Expenditure Survey (HIES) of Iran, which is an important survey to provide information about national accounts. The HIES aims to provide estimates of the average income and expenditure for urban and rural households at provincial and national levels. Making it possible to learn about the households income, expenditure composition, and distributional patterns, the information provided by HIES has been applied to calculate poverty line and study the impurity in household income and facilities. The HIES has a three-stage cluster sampling method for its survey which is conducted annually with a 0-5 rotating panel design and its target population includes all private and collective settled households in urban and rural areas. In a 0-5 rotating panel design, each sampling unit remains in the sample for five sequential surveys and then goes out of the survey forever. In order to obtain more representative estimates of the whole year, the samples are distributed across the months of the year, so each month some samples are considered. The working data set is a selected subset of the data for Tehran provided by HIES. The dataset is not publicly available due to the privacy policy of SCI but it is in access from the corresponding author on a reasonable request. The response variable is the proportion of expenses spent on food to the total expenditure of a house-hold during the reference month of sampling. This proportion can be used for exploring the welfare situation of households. We expect that our desired response variable gets lower values for households with a better economic status. The potentially useful covariates provided by HIES questionnaire are the Decile of Income (DI) , the Household Size (HS), Area of Housing Unit (AHU), Household Income (HI) and Number of Employed Members of the household (NEM). The Moran-I test showed significant spatial dependency, so we adopted SBGLMM for modelling the data. Figure [3] shows the spatial map of the response variable in Tehran.

In order to extract the best subset of effective covariates on the response variable, we proceed using the backward elimination procedure starting by the full model and at each step remove a non-significant covariate until all remaining covariates are significant. We initially suggest using the following full spatial

Table 5. Summary results of modelling both parameters of Beta density

		t- student				**Hierarchical t-student**			
Par.	Value	Mean	Bias	Sd.	RMSE	Mean	Bias	Sd.	RMSE
$\boldsymbol{\beta_0}$	-1	-0.921	-0.079	0.089	0.119	-0.846	-0.154	0.394	0.423
$\boldsymbol{\beta_1}$	2	2.105	0.052	0.083	0.133	1.996	-0.002	0.100	0.100
$\boldsymbol{\beta_2}$	-1.5	-1.540	0.027	0.079	0.089	-1.501	0.001	0.132	0.132
$\boldsymbol{\delta_0}$	4	4.148	0.037	0.441	0.465	4.209	0.052	1.103	1.123
$\boldsymbol{\delta_1}$	-0.4	-0.369	-0.076	0.730	0.731	-0.235	-0.413	1.325	1.335
$\boldsymbol{\sigma^2}$	0.5	0.204	-0.592	0.101	0.313	0.664	0.328	0.271	0.316
$\boldsymbol{\psi^{-1}}$	0.1	0.434	3.336	0.219	0.399	0.111	0.115	0.065	0.066

Box-3

$$\text{Full Model}: \quad \log(\frac{\mu_i}{1-\mu_i}) = \beta_0 + \beta_1 DI_i + \beta_2 HS_i + \beta_3 AHU_i + \beta_4 HI_i + \beta_5 NEM_i + \tau_i, \text{i} = 1,\ldots, 57$$

model (Box-3).

The Variance Inflation Factor (VIF) is calculated to check the multicollinearity problem between the regressors. The largest VIF value is 3.33(<5) when all the assumed covariates are included in the model, so there is no significant evidence of multicollinearity. As mentioned in results of the simulation study, sensitive analysis and convergence tests indicate that we can use the proposed priors to analyze SBLGMM in small sample sizes. So we use the same priors in the application, i.e., $\beta \sim N_5(0, 10^2 I_5)$ and assume that the precision parameter is fixed with $\phi \sim \text{Gamma}(0.01, 0.01)$. In addition, for the spatial parameter, we have $\sigma^2 | \lambda \sim \exp(\lambda)$, $\lambda \sim N(0, \sigma_\lambda^2) I(0, \infty)$, and $\alpha = -\text{Ln}(\psi) \sim U(-4.5, 2.3)$.

To estimate the parameters, we burned-in 500,000 of the 800,000 values of the chain considering spacing of size 100. Thus, we have a total of 3,000 samples upon which the posterior inference is based on. The following model is the outcome of the backward elimination procedure.

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \beta_0 + \beta_1 DI_i + \beta_2 HS_i + \tau_i$$

To evaluate the effectiveness of including spatial correlation structure in the study, we omit the spatial random effect from the model in (7) and fit it again.

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \beta_0 + \beta_1 DI_i + \beta_2 HS_i$$

The deviance information criterion (DIC) is used for evaluating the performance of the fitted models. The DIC criteria for models (7) and (8) are -160.52 and -138.007, respectively. So according to the DIC criteria, the model containing the spatial random effect better fits the data.

Table [6] shows the posterior mean and the 95% credible interval (CI) for the β, ϕ, σ^2, and ψ^{-1} parameters. It can be seen that the two covariates household size and decile of income are significant. Increasing household size results in increasing the proportion of expenses spent on food to the total expenditure of a household. By increasing the decile of income the proportion will be reduced. It means that households with more income allocate a smaller proportion of their expenses to food in comparison with households with a weaker economic

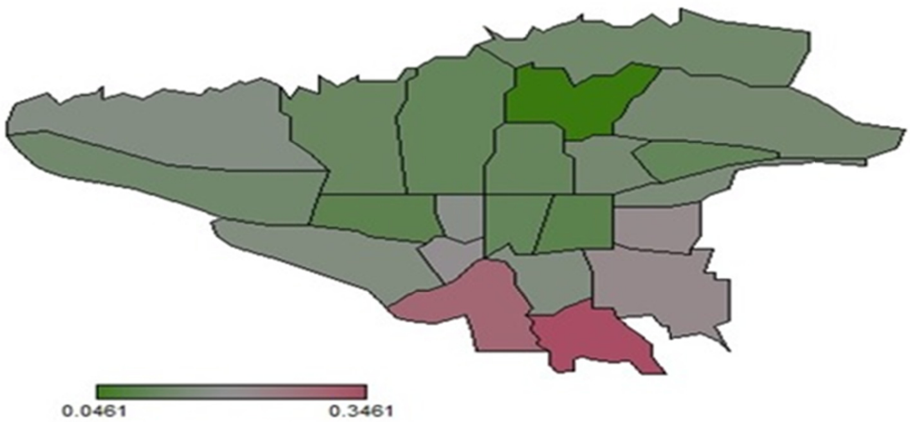

Figure 3. Spatial plot of response variable in Tehran city

Table 6. Parameter estimates, Sd. and 95% CI for Model Step 4

Par.	Estimates	Sd.	95%CI
β_0	-1.574	0.627	(-3.255, -0.556)
β_1	-0.165	0.075	(-0.317, -0.019)
β_2	0.172	0.061	(0.058, 0.293)
ϕ	69.94	26.87	(33.48, 139.2)
σ^2	1.496	4.808	(0.079, 9.695)
ψ^{-1}	5.282	6.475	(0.109, 24.68)

status.

It is important to note that adding spatial effect is necessary to study the response variable. Because ignoring the spatial correlation results in increasing the DIC value, it means that if the spatial dependency of the response variable is neglected, the model is not well fitted.

Results and Discussion

In cases of working with spatially correlated data, this property should be considered in the data analysis to prevent arriving at misleading results. Our proposed SBGLMM is applicable when the spatial response variable is beta distributed. The reparametrized beta distribution has two parameters namely, the mean and the precision parameter. We consider SBGLMM in two situations. First, we assumed that the precision parameter is fixed and then the model was extended for a varying precision parameter status. The spatial correlation structure was included in the model using a random effect in the mean model. After doing an initial study and realizing that typical priors in the Bayesian context are unable to estimate small values of the spatial range parameter, we made an effort to find a suitable prior for this parameter. To analyze the sensitivity of priors, an intensive simulation study carried out upon which the proper values of hyper parameters were assigned. The Bayesian approach was applied to estimate the parameters involved in the proposed model. As the posterior distribution is complicated, the Gibbs sampler was run to fit the model using MCMC scheme, while facing slow mixing chains. The further inference was performed considering convergence tests and graphical diagnostic tools. Outcomes of intensive simulation studies over different sample sizes and results of sensitivity analysis, demonstrated that parameter estimations are reliable based on the proposed priors which are working properly, especially for small values of spatial range even for small sample sizes.

Additionally, we targeted to fit a model while a varying precision parameter is supposed. The linear pattern including fixed effects is assumed for the precision parameter. Finding proper priors for these regression coefficients was challenging and time consuming. Although Normal prior distribution does not provide appropriate estimates for the regression coefficients when the sample size is small, results of simulations confirm that Normal prior distribution is able to estimate regression parameters of the precision model when the sample size is increased.

SBGLMM provides us with a useful tool for modelling spatially correlated rates and proportions which are common in many areas such as official statistics. As an application, we used our model to study HIES data in Tehran, the capital of Iran. The results show that the proportion of expenses spent on food in a household is affected by decile of income and household size. It is worth to mention that due to the spatial dependency of the response variable, adding the spatial effect to the model yielded to get better results.

There is a wide vicinity for future work on this issue. For instance, this can be carried out by extending the models including some other random effects, working on other spatial correlation structures, or studying spatiotemporal issues.

Acknowledgement

The authors are thankful to the referees for their many helpful comments that greatly improved this paper. We also wish to acknowledge for the support from Center of Excellence of Spatio-Temporal Data Analysis in Tarbiat Modares University.

References

1. Ferrari S., Cribari-Neto F. Beta Regression for Modelling Rates and Proportions. *J. Appl. Stat*. **31**: 799-815 (2004).
2. Ferrari S.L., Pinheiro E.C. Improved Likelihood Inference in Beta Regression. *J. Stat. Comput. Simul.* **81:** 431-443 (2011).
3. Cepeda-Cuervo E., Gamerman D. Bayesian Methodology for Modelling Parameters in the two Parameter Exponential Family. *Rev. Estad.* **57**: 93-105 (2005).
4. Smithson M., Verkuilen J. A Better Lemon Squeezer? Maximum-Likelihood Regression with Beta Distributed Dependent Variables. *Psychol. Methods*. **11**: 54-71 (2006).
5. Branscum A.J., Johnson W.O., Thurmond M.C. Bayesian Beta Regression: Applications to Household Expenditure Data and Genetic Distance Between Foot-and-Mouth Disease Viruses. *Aust. N. Z. J. Stat.* **49**: 287-301 (2007).
6. Simas A.B., Barreto-Souza W., Rocha A.V. Improved Estimators for a General Class of Beta Regression Models. *Comput. Stat. Data Anal.* **54**: 348-366 (2010).
7. Zimprich D. Modelling Change in Skewed Variables using Mixed Beta Regression Models*. Res Hum Dev.* **7**: 9-26 (2010).
8. Figueroa-Zúñiga J.I., Arellano-Valle R.B., Ferrari S.L. Mixed Beta Regression: A Bayesian Perspective. *Comput. Stat. Data Anal***. 61**: 137-147 (2013).
9. Verkuilen J., Smithson M. Mixed and Mixture Regression Models for Continuous Bounded Responses using the Beta Distribution. *J. Educ. Behav. Stat.* **37**: 82-113 (2012).
10. Ferreira G., Figueroa-Zúñiga J.I., de Castro M. Partially Linear Beta regression Model with Autoregressive Errors. *TEST*. **24**: 752-775 (2015).
11. Cepeda-Cuervo E., Urdinola B.P., Rodriguez D. Double

Generalized Spatial Econometric Models. *Commun. Stat. Simul. Comput.* **41**: 671-685 (2012).
12. Cepeda-Cuervo, E., Nunez-Anton V. Spatial Double Generalized Beta Regression Models Extensions and Application to Study Quality of Education in Colombia. *J. Educ. Behav. Stat.* **38**: 604-628 (2013).
13. Gholizadeh K., Mohammadzadeh M., Ghayyomi Z. Spatial Analysis of Structured Additive Regression and Modelling of Crime Data in Tehran City Using Integrated Nested Laplace Approximation. *Journal of Statistical Society*. **7**: 103-124 (2013).
14. Fustos R. *Modelo Lineal Generalizado Espacial Con Variable Respuesta Beta*. Engineer's Degree Dissertation. Department of Statistics. University of Concepcion, Chile. (2013).
15. Lagos-Alvarez B.M., Fustos-Toribio R., Figueroa-Zúñiga J., and Mateu, J. Geostatistical Mixed Beta Regression: A Bayesian Approach. *SERRA*. **31**: 571-584 (2016).
16. Kalhori L., Mohammadzadeh M. Spatial Beta Regression Model with Random Effect. *J. SRI*. **13**: 214-230 (2016).
17. Brooks S.P., Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. *J. Comput. Graph. Stat.* **7**: 434-455 (1998).
18. Heidelberger P., Welch P.D. A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations. *Commun. ACM*. **24**: 233-245 (1981).
19. Diggle P.J., Tawn J., Moyeed R. Model-Based Geostatistics. *J. Royal. Stat. Soc: Ser. C Appl. Stat.* **47**: 299-350 (1998).
20. Mark S. Handcock M.L.S. A Bayesian Analysis of Kriging. *Technometrics*. **35**: 403-410 (1993).
21. Stein M.L. Interpolation of Spatial Data: Some Theory for Kriging, *Springer Science & Business Media, New York.* (2012).
22. Stein M.L. Interpolation of Spatial Data: Some Theory for Kriging, *Springer Science & Business Media, New York.* (2012).
23. Chen M.H., Shao Q.M., Ibrahim J. Monte Carlo Methods in Bayesian Computation, *Springer*, *New York*. (2000).
24. Sturtz S., Ligges U., Gelman A. R2winbugs: A Package for Running Winbugs from R. *J. Stat. Softw*. **12**: 1-16 (2005).
25. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, *Austria, Vienna.* (2013).
26. Geweke J. *Evaluating the Accuracy of Sampling Based Approaches to the Calculation of Posterior Moments*. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA. **196**: (1991).
27. Huang X., Li G., Elashoff R.M. A Joint Model of Longitudinal and Competing Risks Survival Data with Heterogeneous Random Effects and Outlying Longitudinal Measurements. *Stat. Its Interface*. **3**: 185 (2010).

Appendix I:
Full conditional distributions are as follow

$$\pi(\beta \mid \alpha,\sigma_0^2,\lambda,\tau,\phi,\mathrm{y}) \propto \pi(\beta \mid \tau,\phi,\mathrm{y}) \propto \prod_{i=1}^{n} f(y_i \mid \tau_i,\beta,\phi)\pi(\beta)$$
$$= \exp(\sum_{i=1}^{n} \log(f(y_i \mid \tau_i,\beta,\phi)) - \frac{(\beta-\beta_0)'\Sigma_0^{-1}(\beta-\beta_0)}{2})$$

$$\pi(\sigma_0^2 \mid \beta,\alpha,\lambda,\tau,\phi,\mathrm{y}) \propto \pi(\sigma_0^2 \mid \alpha,\lambda,\tau)$$
$$\propto \pi(\tau \mid \alpha,\lambda,\sigma_0^2)\pi(\sigma_0^2 \mid \lambda)\pi(\lambda)$$
$$= \exp\frac{-1}{2}(\tau'\Sigma_\tau^{-1}\tau)\lambda \exp(-\lambda\sigma_0^2)\exp\frac{-1}{2}(\frac{\lambda-\mu_\lambda}{\sigma_\lambda})^2$$

$$\pi(\alpha \mid \lambda,\sigma_0^2,\tau) \propto \pi(\tau \mid \alpha,\lambda,\sigma_0^2)\pi(\sigma_0^2 \mid \lambda)\pi(\alpha)$$
$$= \exp\frac{-1}{2}(\tau'\Sigma_\tau^{-1}\tau)\ell_2 - \ell_1$$
$$= \exp\frac{-1}{2}(\tau'\Sigma_\tau^{-1}\tau + 2\ln(\ell_2 - \ell_1))$$

$$\pi(\lambda \mid \beta,\sigma_0^2,\alpha,\tau,\phi,\mathrm{y}) \propto \pi(\lambda \mid \sigma_0^2,\alpha,\tau)$$
$$\propto \pi(\tau \mid \alpha,\lambda,\sigma_0^2)\pi(\lambda) \quad \}$$
$$= \exp\frac{-1}{2}\{(\tau'\Sigma_\tau^{-1}\tau) + (\frac{\lambda-\mu_\lambda}{\sigma_\lambda})^2$$

$$
\begin{aligned}
\pi(\phi \mid \beta, \sigma_0^2, \alpha, \lambda, \tau, \mathrm{y}) &\propto \pi(\phi \mid \beta, \tau, y) \\
&\propto \prod_{i=1}^{n} f(y_i \mid \tau_i, \beta, \phi)\pi(\phi) \\
&= \exp\Big(\sum_{i=1}^{n} \log(f(y_i \mid \tau_i, \beta, \phi) - \phi \grave{o}_2)\frac{1}{\grave{o}_1} \grave{o}_2^{\grave{o}_1} \phi^{\grave{o}_1 - 1}.
\end{aligned}
$$

$$
\begin{aligned}
\pi(\tau_k \mid \tau_{-k}, \beta, \alpha, \sigma_0^2, \lambda, \phi, y) &\propto \pi(\tau_k \mid \tau_{-k}, \alpha, \sigma_0^2, \lambda, y) \\
&\propto \prod_{i=1}^{n} f(y_i \mid \tau_i, \beta, \phi)\pi(\tau_k \mid \tau_{-k}, \alpha, \sigma_0^2, \lambda) \quad 1 \le k \le n
\end{aligned}
$$

where $\tau_{-k} = (\tau_1, \ldots, \tau_{k-1}, \tau_{k+1}, \ldots, \tau_n)$.