Journal of Sciences, Islamic Republic of Iran 33(2): 139 - 148 (2022) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

An Approximate Bayesian Inference for Beta Regression Models

K. Gholizadeh, M. Mohammadzadeh*

Department of Statistics, Tarbiat Modares University, Tehran, Islamic Republic of Iran

Received: 21 August 2021 / Revised: 23 January 2022 / Accepted: 19 February 2022

Abstract

In modeling the variables related to each other, regression models are usually used assuming that the response variable is Normal. But in problems dealing with data such as the rate or ratio of an event distributed in the (0,1) interval, these models may provide out-of-range predictions for the response variable. In addition rate or ratio data are oftem asymmetrically distributed, and the use of symmetric distributions leads to invalid results. In such cases, the Beta regression model is used, in which the distribution of the response variable is in the Beta family. Bayesian analysis of these models generally requires the calculation of multiple integrals. The use of MCMC algorithms sometimes encounters long computation times and divergence. This work presents approximate methods for obtaining posterior distributions for Bayesian analysis of Beta regression models. Then the Integrated Nested Laplace Approximation will be offered for getting the posterior distributions in the Bayesian analysis of these models. Moreover, these models' application is illustrated on a real data set.

Keywords: Beta regression models; Integrated nested Laplace; Approximation Bayesian inference.

Introduction

A simple standard method to analyze data related to some other variables can be a regression model. In practice, there are many situations in which the response is restricted to the interval (0,1), such as percentages, proportions, rates, and fractions. Since linear regression models may predict the response variable out-of-range, they are not appropriate for such situations.

Although transforming the dependent variable on the real line is possible, the model parameters cannot be easily interpreted. Moreover, proportions typically display asymmetry, and hence inference based on the normality assumption can be misleading [1]. The Beta distribution provides a useful tool for modeling data restricted to the interval (0,1). The density function of a Beta distribution, Beta(p,q), is given by

$$f(y|p,q)=\frac{1}{B(p,q)}y^{(p-1)}(1-y)^{(q-1)}, 0<y<1,$$
$$p,q>0,$$

where $B(p,q)=\int_0^1 y^{p-1}(1-y)^{q-1}dy$. The mean and variance of this distribution are, respectively given by

$$\mathrm{E}(Y)=\mu=\frac{p}{p+q},\ \ \mathrm{Var}(Y)=\frac{pq}{(p+q)^2(p+q+1)}.$$

Since the probability density function of the Beta distribution can have quite different shapes depending on the values of p and q, it is very flexible for modeling proportions (Figure 1). This distribution family includes left or right-skewed, and alsosymmetric distributions.

* Correspomding author: Tel: +989122066712; Email: mohsen_m@modares.ac.ir

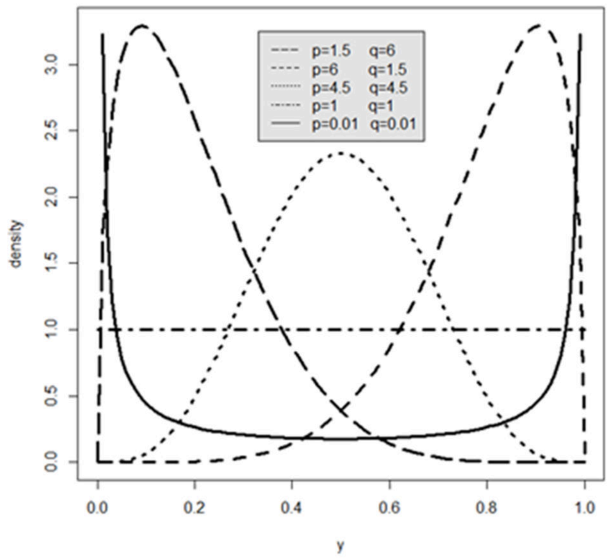


Figure 1. Beta distribution form for different parameters.

[1] proposed a Beta regression model assumimg that the response is a Beta distributed random variable. In such a model, the mean of the response variable relates to a linear predictor through a known link function. The linear predictor contains some covariates and unknown regression parameters, assuming that the precision parameter φ is constant. In many applications, the precision parameter φ may not be constant over all observations, so it is proposed modeling of φ. Simultaneous modelling of two parameters of μ and φ complicates the model and therefore, more time-consuming calculations, so we used the INLA method to run the model faster.

Bayesian Beta Regression Models

The density function of a Beta distribution with parameters p and q can be re-parameterized in terms of its mean, $\mu = E(Y)$, and precision parameter, $\varphi = p + q$, as

$f(y|p,q) = \frac{\Gamma(\varphi)}{\Gamma(\varphi\mu)\Gamma((1-\mu)\varphi)} y^{(\mu\varphi-1)}(1-y)^{(1-\mu)\varphi-1}, 0 < y < 1.$

In this reparameterization $p = \mu\varphi$, $q = \varphi(1-\mu)$ and $\mathrm{Var}(Y) = \frac{\mu(1-\mu)}{1+\varphi}$. Let $Y_1, \cdots, Y_n$ be independent random variables such that $Y_i \sim \mathrm{Beta}(\mu_i, \varphi)$ and the mean of Y_i, μ_i, is related to a linear predictor, η_i, through a twice differentiable strictly monotonic link function $g(0,1) \to \mathcal{R}$. A general model for μ_i, is given by $g(\mu_i) = \boldsymbol{z_i\beta}$, where $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_m)^T$ is the regression parameters and $\boldsymbol{z_i} = (z_{i1}, \cdots, z_{im})$ are observations on m covariates. The most common link functions for μ_i are the logit, probit, and complementary log–log links [2, 3].

In many applications, the precision parameter φ may not be constant over the all observations, so [4] proposed modeling of φ, thereby setting up a second regression model. To guarantee the positivity of the variance, the log-link function for the precision parameter is considered as $h(\varphi_i) = \ln(\varphi_i) = \boldsymbol{w_i\gamma}$, where $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_k)^T$ is the regression parameters and $\boldsymbol{w_i} = (w_{i1}, \cdots, w_{ik})$ are observations on k covariates.

[5] proposed a model that involved random effects for Beta-distributed dependent variables. Generalized linear mixed models (GLMMs) were developed to handle dependencies in spatial, clustered, longitudinal data, etc. They considered the mean and precision model

$$\mathrm{logit}(\mu_i) = \boldsymbol{z_i\beta}, + \boldsymbol{\nu_i b}, \quad \ln(\varphi_i) = \boldsymbol{w_i\gamma} + \boldsymbol{\upsilon_i d},$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\upsilon}$ are the regressors for random effects $\boldsymbol{b}$ and $\boldsymbol{d}$, respectively.

Bayesian Estimation of Beta Regression Models

The likelihood function, with independent data, is given by

$$\ell(\boldsymbol{\beta},\boldsymbol{\gamma}|\boldsymbol{y}) = \prod_{i=1}^{n} \frac{\Gamma(h^{-1}(\varphi_i))}{\Gamma(g^{-1}(\mu_i)h^{-1}(\varphi_i))\Gamma((1-g^{-1}(\mu_i))h^{-1}(\varphi_i))} \times y_i^{g^{-1}(\mu_i)h^{-1}(\varphi_i)-1}(1-y_i)^{\left(1-g^{-1}(\mu_i)\right)h^{-1}(\varphi_i)-1}$$

and the joint posterior density is given by

$$\pi(\boldsymbol{\beta},\boldsymbol{\gamma}|\boldsymbol{y}) \propto \boldsymbol{\ell}(\boldsymbol{\beta}|\boldsymbol{\gamma};\boldsymbol{y}) \times \pi(\boldsymbol{\beta},\boldsymbol{\gamma}).$$

It is analytically intractable to find and to generate samples from posterior distribution $\pi(\boldsymbol{\beta},\boldsymbol{\gamma}|\boldsymbol{y})$. The

MCMC technique is an excellent solution to solve this problem [6, 7] presented the MCMC algorithm to estimate the parameters of a Beta regression model. They built two Normal transition kernels, q1 and q2, to apply a Metropolis-Hastings algorithm for generating posterior samples of β and γ. This algorithm can be implemented using the Bayesianbetareg package in R developed by [8]. Furthermore, there is some software for performing MCMC algorithms such as OpenBUGS and BayesX, etc. Although BayesX is faster than other software, it uses a repetitive method. Selecting the number of repetitions affects computing time.

The Integrated Nested Laplace Approximation (INLA)

This approximation is a novel numerical inference approach to perform fast Bayesian inference accurate approximation of the marginal posterior densities of hyperparameters and latent variables in the latent Gaussian models [9, 10]. This statistical model classifies several well-known regression models such as generalized linear models, spatial and spatio-temporal models, geostatistical and geoadditive models. These models have been successfully implemented using the R package INLA.

The INLA method has two advantages over the MCMC techniques. The most important one is computational. We may analyze models that included high dimensional latent random field by INLA method in a few minutes, while this task may take several hours by the MCMC algorithms. Besides, the parallel execution of INLA package in R has been prepared, making it possible to use the modern multi-core processors. The second advantage, since the INLA method treats the latent Gaussian models in a unified manner, makes it possible to automate the inference process of the model [11].

Let $Y_i \sim \text{Beta}\big(\mu_i\varphi_i, \varphi_i(1-\mu_i)\big),\ i = 1, \cdots, n$. Suppose the mean and precision parameters are modeled as

$$\text{logit}(\mu_i) = \boldsymbol{z_i\beta} + \boldsymbol{\nu}_i, \quad \ln(\varphi_i) = \boldsymbol{w_i\gamma},$$

where $\boldsymbol{\beta} = (\beta_0, \cdots, \beta_m)^T$ and $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_k)^T$ are the mean and precision regression coefficients, $\boldsymbol{z_i}$s are the mean explanatory variables, and $\boldsymbol{\nu_i}$ is structured random effects in modeling mean. The linear effect of covariates $\boldsymbol{z}$ in modeling precision parameter, $\boldsymbol{\gamma}$, is considered as a hyper parameter. A Gaussian prior is assigned to $\boldsymbol{\beta}$ and $\boldsymbol{\nu}$, and let $\mathbf{x}$ denote the vector of all the latent Gaussian variables, $\mathbf{x} = \{\boldsymbol{\nu}, \boldsymbol{\beta}\}$, with the density that is Gaussian with zero mean and precision matrix $\boldsymbol{Q(\theta_1)}$ with hyperparameters $\boldsymbol{\theta_1}$ with $\dim(\boldsymbol{\theta_1}) = l$. It is assumed that the latent field $\mathbf{x}$ admits conditional independence properties. Hence, the latent field is a Gaussian Markov Random Field (GMRF) with a sparse precision matrix $\boldsymbol{Q(\theta_1)}$.

The distribution for $\boldsymbol{y} = (y_1, \dots, y_n)$ is denoted by $\pi(\boldsymbol{y}|\mathbf{x}, \boldsymbol{\gamma})$ and we assume that $\{y_i : i = 1, \dots, n\}$ are conditionally independent given $\mathbf{x}$ and $\boldsymbol{\gamma}$. For simplicity, denote by $\boldsymbol{\theta} = (\boldsymbol{\gamma^T}, \boldsymbol{\theta_1^T})$, with $\dim(\boldsymbol{\theta}) = m + l$. Then the joint posterior of $\mathbf{x}$ and $\boldsymbol{\theta}$ is given by

$$\pi(\mathbf{x}, \boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta}_1)\prod_{i=1}^{n}\pi(y_i|\mathrm{x}_i, \boldsymbol{\gamma}),$$

$$\propto \pi\,(\boldsymbol{\theta})|\,\boldsymbol{Q}(\boldsymbol{\theta}_1)|^{\frac{1}{2}}\exp\{\frac{-1}{2}\mathbf{x}^T\boldsymbol{Q}(\boldsymbol{\theta}_1)\mathbf{x} + \sum_{i=1}^{n}\log\{\,\pi(y_i|\mathrm{x}_i, \boldsymbol{\gamma})\}\}.$$

In the INLA computation, the main interest are the marginal posterior distributions $\pi(x_i|\boldsymbol{y})$ and $\pi(\theta_j|\boldsymbol{y})$, that can be written as

$$\pi(x_i|\boldsymbol{y}) = \int \pi(x_i|\boldsymbol{y}, \boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta},$$

$$\pi(\theta_j|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}_{-j}. \qquad (1)$$

The INLA method exploits the assumptions of the model to produce a numerical approximation for the posteriors based on the Laplace approximation in two steps [12]. The first step is the computation of an approximate posterior of the hyperparameters as

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\mathbf{x}, \boldsymbol{\theta}|\boldsymbol{y})}{\pi(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y})},$$

$$\propto \frac{\pi(\boldsymbol{y}|\mathbf{x}, \boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y})},$$

$$= \frac{\pi(\boldsymbol{y}|\mathbf{x}, \boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\tilde{\pi}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y})}\Big|_{\mathrm{x}=\mathrm{x}^*(\boldsymbol{\theta})},$$

where $\tilde{\pi}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y})$ is the Gaussian approximation to the full conditional of $\mathbf{x}$, and $\mathrm{x}^*(\boldsymbol{\theta})$ is the mode of the full conditional of $\mathbf{x}$, for a given $\boldsymbol{\theta}$. In order to the Gaussian approximate, it can be written

$$\pi(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y}) \propto \ \pi(\mathbf{x}|\boldsymbol{\theta})\prod_{i=}^{n}\pi(y_i|x_i, \boldsymbol{\gamma}),$$

$$\propto \exp\{\frac{-1}{2}\mathbf{x}^T\boldsymbol{Q}(\boldsymbol{\theta}_1)\mathbf{x} + \sum_{i=1}^{n}\log\pi(y_i|x_i, \boldsymbol{\gamma})\}.$$

Now, a second-order Taylor expansion of $\sum_{i=1}^{n}\log\pi(y_i|x_i, \boldsymbol{\gamma})$ around, $\mathbf{x}^*(\boldsymbol{\theta})$ is used. To be specific,

$$\pi(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y}) \propto \exp(\frac{-1}{2}\mathbf{x}^T\boldsymbol{Q}(\boldsymbol{\theta}_1)\mathbf{x} + \sum_{i=1}^{n}a_i + b_i x_i - \frac{1}{2}c_i x_i^2),$$

$$\propto \exp(\frac{-1}{2}\mathbf{x}^T\boldsymbol{Q}(\boldsymbol{\theta}_1)\mathbf{x} + \mathrm{diag}(\boldsymbol{c}))\mathbf{x} + \boldsymbol{b}^T\mathbf{x}),$$

Table 1. The accuracy of the parameter estimates using MCMC and INLA methods (Simulation A)

	Parameter	True value	Estimate	SD	Quantile 0.025	Quantile 0.975
INLA	β_x	1	0.996	0.232	0.541	1.451
	β_1	0	-0.002	0.040	-0.081	0.077
	β_2	1	0.995	0.050	0.896	1.093
	β_3	1.5	1.495	0.047	1.401	1.588
	β_4	2	1.992	0.052	1.888	2.095
	γ_1	25	24.76	3.039	18.969	31.961
	γ_2	20	19.95	2.615	15.235	25.263
	γ_3	30	29.473	3.842	22.891	37.714
	γ_4	35	34.102	4.748	25.923	44.828
MCMC	β_x	1	0.999	0.229	0.551	1.448
	β_1	0	-0.002	0.040	-0.080	0.076
	β_2	1	0.999	0.064	0.873	1.123
	β_3	1.5	1.500	0.062	1.379	1.621
	β_4	2	2.00	0.065	1.872	2.128
	γ_1	25	25.22	3.553	19.100	33.185
	γ_2	20	20.13	2.747	15.226	25.737
	γ_3	30	30.36	4.219	23.222	39.559
	γ_4	35	35.528	5.331	26.552	47.741

Table 2. DIC, WAIC, LS and computing time

Model	DIC	WAIC	Time(second)
INLA	-927.203	-927.415	2200
MCMC	-927.282	-927.728	3400

which is in the form of canonical parametrization $N(\boldsymbol{b}, \boldsymbol{Q}(\boldsymbol{\theta}) + \text{diag}(\boldsymbol{c}))$ [13].

For the next step, the posterior conditional distributions $\pi(x_i|\boldsymbol{y}, \boldsymbol{\theta})$ can be approximated directly as the marginals from $\tilde{\pi}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{y})$. The Gaussian approximation often gives suitable results, but the skewness of the distribution may not be captured [9]. A more accurate approach would be used the Laplace approximation, that is given by

$$\tilde{\pi}_{LA}(x_i|\boldsymbol{\theta}, \boldsymbol{y}) \propto \frac{\pi(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{y})\pi(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\tilde{\pi}_{GG}(\mathbf{x}_{-\mathbf{i}}|x_i, \boldsymbol{\theta}, \boldsymbol{y})}\Big|_{\mathrm{x}_{-\mathrm{i}}=\mathrm{x}^*_{-i}(x_i,\boldsymbol{\theta})},$$

where $\tilde{\pi}_{GG}$ is the Gaussian approximation to $\tilde{\pi}_G(\mathbf{x}_{-\mathbf{i}}|x_i, \boldsymbol{\theta}, \boldsymbol{y})$ and $\mathrm{x}^*_{-i}(x_i, \boldsymbol{\theta})$ is its mode.

Once two steps are computed, the marginal posterior distributions $\pi(x_i|\boldsymbol{y})$, the equation (1), would be calculated numerically as

$$\tilde{\pi}(x_i|\boldsymbol{y}) = \sum_k \tilde{\pi}(x_i|\boldsymbol{\theta}_k, \boldsymbol{y})\tilde{\pi}(\boldsymbol{\theta}_k|\boldsymbol{y})\Delta_k,$$

where Δ_k is the corresponding weight of the integration point $\boldsymbol{\theta}_k$.

Simulation Study

This section performs two simulation studies to evaluate the INLA method's performances for joint modeling of the mean and precision parametersin Beta regression model.

Simulation A: The purpose is to evaluate the adequacy of using MCMC and INLA to fit the Beta regression models and compare the performance of these two methods and their computational time. The data sets were simulated from the Beta distributions, $Y_{st} \sim \text{Beta}(\mu_{st}, \varphi_t)$, at the grid points $s \in \{1, \cdots, 100\}$, and the time points $t \in \{1, \cdots, 4\}$, where

$$\text{logit}(\mu_{st}) = \boldsymbol{i'}_t\boldsymbol{\beta} + \beta_x\, x_{st}, \quad \log(\varphi_t) = \boldsymbol{i'}_t\boldsymbol{\gamma}, \quad (2)$$

and $\boldsymbol{i'}_t$ is a vector of indicator variables corresponding to the sampling time. The covariates values were generated from a normal distribution and the true values of the parameters were set as $\boldsymbol{\beta} = (\beta_1, \dots, \beta_4) = (0, 1, 1.5, 2)$, $\beta_x = 1$ and $\gamma = (\gamma_1, \dots, \gamma_4) = (25,20,30,35)$. We fitted the model in (2) using both MCMC and INLA in 1000 replications. The results in Table 1 show the estimates' closeness from these two approaches. To assess the performance of INLA, also the output will be compared with results MCMC from BayesX based on DIC, WAIC, and computing time (Table 2).

Table 2 suggests that these two approaches produced similar results based on DIC and WAIC, but in the computation times in seconds, MCMC is very time-consuming. Thus, in applications where processing time is critical, INLA can be an excellent alternative to MCMC.

Simulation B: The data sets were generated as simulation A to evaluate the effect of modeling the precision parameter in the Beta regression models' result when there are four-time points with different variances. We simulated 1000 data sets and fitted two models for each data set. Both models have the same mean model as $\text{logit}(\mu_{st}) = \boldsymbol{i}_t'\boldsymbol{\beta} + \beta_x x_1$. In the Model 1, the precision parameter is considered to be constant, φ, but in the Model 2, it is modeled as $\log(\varphi_t) = \boldsymbol{i}_t'\boldsymbol{\gamma}$ (i.e., variance of the response variable is different at four time points). Table 3 shows the averages, standard deviations, 2.5% and 97.5% quantiles of the estimates across 100 replications. Although both models' estimates are close to the true value, the credibility interval of β_4 in Model 1 does not include the true value. The preference of Model 2 is also accorded by DIC's, WAIC's and LS's, as shown in Figure 2. Therefore, assuming the precision parameter to be constant in cases where it is not constant causes the other parameters not to be estimated well.

Application

The data in this section is about soil-crop sensor technologies to analyze the spatial and temporal variations of soil-plant relationships because of management practices. The goal is to model the relationship between plant growth and soil characteristics as measured by two sensors in southern Italy. The response and explanatory variables are the Normalized Density Vegetation Index (NDVI) and the Electrical soil Resistivity (ER).

The spatial resolution contains a 2574 cells square

Table 3. The accuracy of the parameter estimates using MCMC and INLA methods (Simulation B)

					Quantile	
Model	Parameter	True value	Estimate	SD	0.025	0.975
1	β_x	1	1.00	0.04	1.00	1.09
	β_1	0	0.00	0.01	-0.01	0.02
	β_2	1	1.01	0.01	1.00	1.03
	β_3	1.5	1.49	0.01	1.47	1.51
	β_4	2	1.97	0.01	1.94	1.99
	φ	27.5	26.16	0.36	25.47	26.7
2	β_x	1	1.00	0.05	0.91	1.09
	β_1	0	0.00	0.01	-0.01	0.02
	β_2	1	1.00	0.01	0.98	1.02
	β_3	1.5	1.50	0.01	1.48	1.52
	β_4	2	2.00	0.01	1.98	2.02
	γ_1	25	25.01	0.68	23.69	26.38
	γ_2	20	20.03	0.55	18.97	21.13
	γ_3	30	30.00	0.83	28.39	31.66
	γ_4	35	35.01	0.98	33.12	36.96

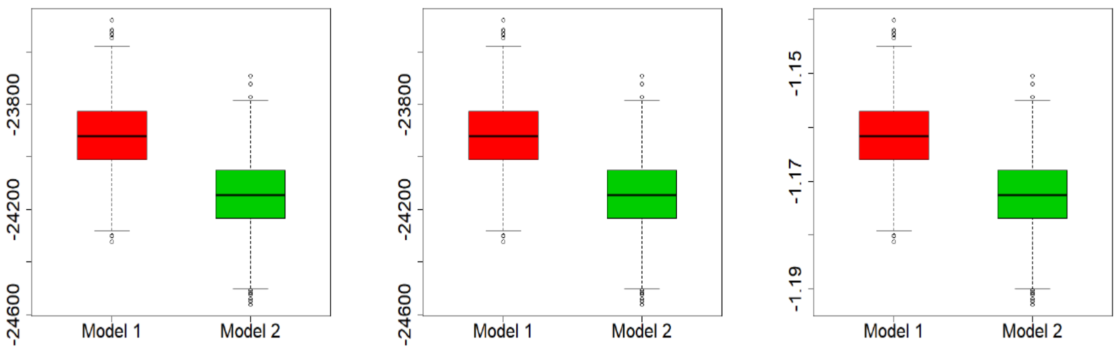
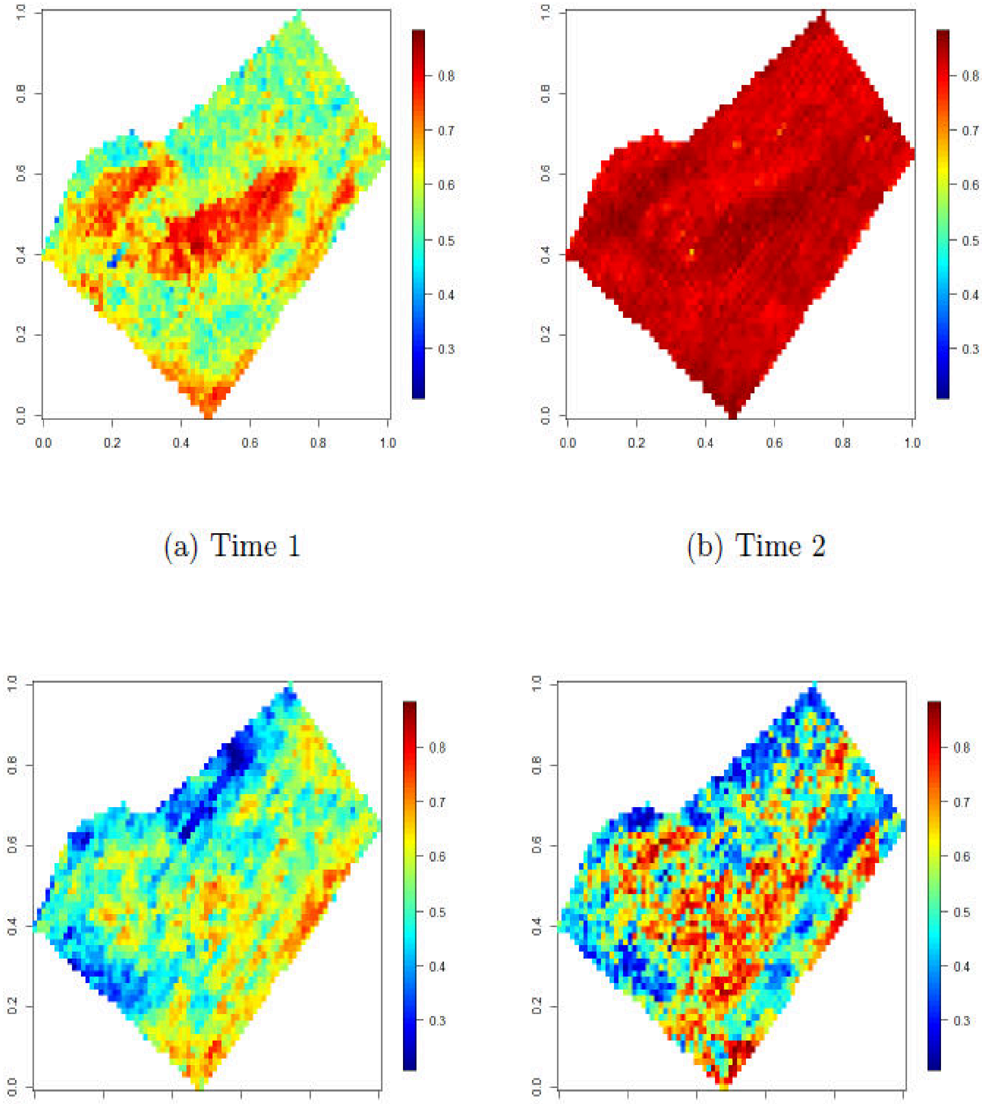


Figure 2. Boxplots of DIC (left), WAIC (middle) and LS (right) for 1000 simulations of Models 1 and 2.



(a) Time 1 (b) Time 2

(c) Time 3 (d) Time 4

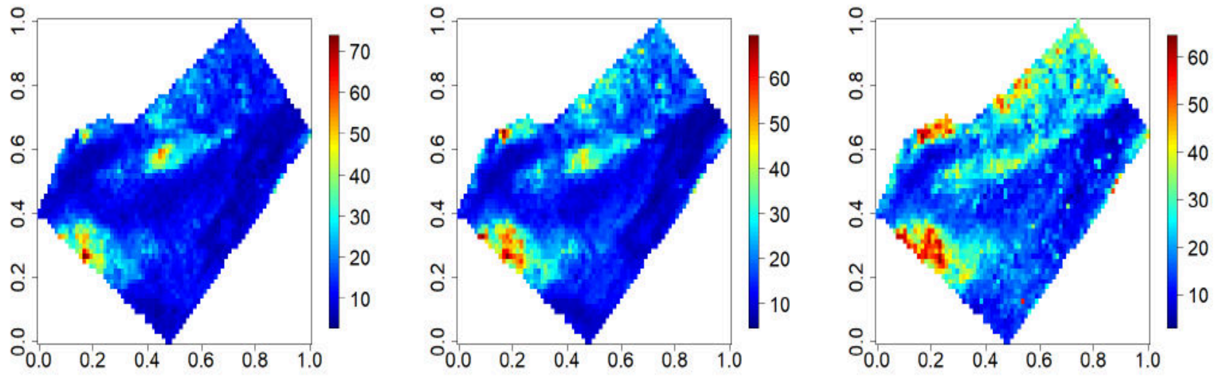
Figure 3. NDVI at 4 time points.

lattice. The NDVIs were measured at four-time points in different seasons, and ERs were taken once at three depth layers: 0.5m, 1m, and 2m. Figure 3 suggests that NDVI measurements have a higher position and minor variability at the second temporal point. Figure 4 shows that ER has not robust systematic variation along with depth; for this reason, the mean of ER along three depth layers are used in modeling. The NDVI data introduced are analyzed using a Beta regression model with INLA. For the grid points $s \in \{1, \cdots, 2574\}$ and the time points: $t \in \{1, \cdots, 4\}$, we considered two differen Beta regression models: (a) a model with the same variance for 4 time points, (b) a model with different variance for 4 time points.

a. Same Variance: The model is given by

$$\text{logit}(\mu_{st}) = \beta_0^{\mu} + \boldsymbol{i}_t' \boldsymbol{\beta}_1^{\mu} + f_1^{\mu}(\text{ER}_s) + f_2^{\mu}(s),$$

where $\boldsymbol{i}_t$ is a seasonal indicator variable corresponding to the sampling times, $f_1(\cdot)$ is a nonlinear smooth function of latent replicate-free (mean of three replicate) ER and $f_2(\cdot)$ is a nonlinear smooth



(a) depth=0.5 (b) depth=1 (c) depth=1

Figure 4. ER at 3 depth layers.

Table 4. DIC, WAIC and LS

Model	DIC	WAIC	LS
1	-33063.44	-32806.71	-1.5591
2	-37834.22	-37950.26	-1.8186

Table 5. The parameter estimates of the Models 2

			Quantile	
Parameter	Estimate	SD	0.025	0.975
$\boldsymbol{\beta_1}$	0.247	0.191	-0.127	0.621
$\boldsymbol{\beta_2}$	0.124	0.264	0.605	1.642
$\boldsymbol{\beta_3}$	-0.405	0.270	-0.935	0.124
$\boldsymbol{\beta_4}$	-0.080	0.279	-0.628	0.469
$\boldsymbol{\gamma_1}$	161.858	6.177	150.044	174.340
$\boldsymbol{\gamma_2}$	1350.986	57.49	1241.367	1467.499
$\boldsymbol{\gamma_3}$	142.313	5.888	131.061	154.227
$\boldsymbol{\gamma_4}$	26.773	0.861	25.156	28.532

function of situation s for estimating spatial effect. Here $f_1(\cdot)$ and $f_2(\cdot)$ follow an second-order random walk model and an 2-dimensional random walk model, Respectively. In this model, for every 4 time points, the same variance is assumed, so the precision parameter φ is constant.

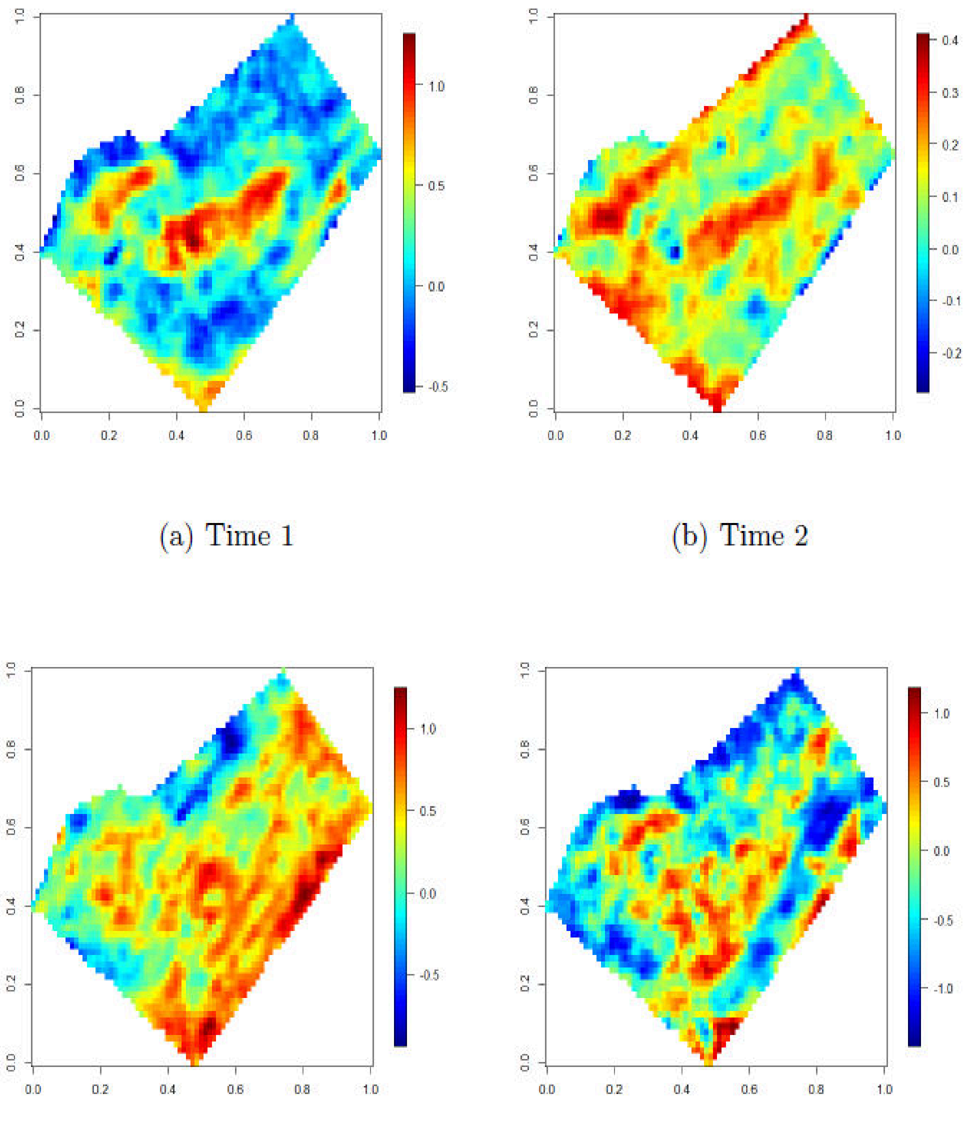
b. Different Variance: Concerning Figure 3 (right), the variance of time points is different. In the second model, we consider the effects of heteroscedasticity of NDVI seasonal recordings in the precision model. The mean model is the same (3), and the precision model is defined as $\ln(\varphi_t) = \boldsymbol{i}_t'\boldsymbol{\gamma}$.

We represent the same variance with Model 1 and a different variance with Model 2. The INLA package performed the computations of both models in R. Note that Models 1 and 2 are similar to BM1 and BM2 mentioned by [14], where the difference is that they implemented the models using MCMC, while we have used INLA. Table 4 shows three model choice criteria for the two models using INLA methods: the Deviance Information Criterion (DIC), the Watanabe-Akaike Information Criterion (WAIC), and the Logarithmic Score (LS). According to the results, the three criteria values have decreased for a model with different variances. Therefore, Model 2 is the best suitable one. Figures 5 and 6 display the posterior mean of the spatial effects of Models 1 and 2, respectively. The spatial results are smoother in Model 2. Table 5 shows the estimates of the Model 2 parameters. As can be seen, the precision parameter in time point 2, γ_2, is higher than the others, which is entirely consistent with the data.

The INLA computational time of this application is about 2 hours. In contrast, MCMC computational time is multiplied due to the model complexity and the size of the data.

Results and Discussion

The Beta regression models, with joint modeling of the mean and precision parameters, are described and illustrated with two simulation examples and real data using INLA. In the first simulation, we compared the performance of MCMC and INLA approach. The INLA approach's accuracy is as good as the MCMC approach, but it is much faster than the MCMC. Thus, in applications where processing time is crucial, INLA can be an excellent alternative to MCMC. As shown in the second simulation, the modeling of dispersion is



(a) Time 1 (b) Time 2

(c) Time 3 (d) Time 4

Figure 5. The spatial effects of Model 1.

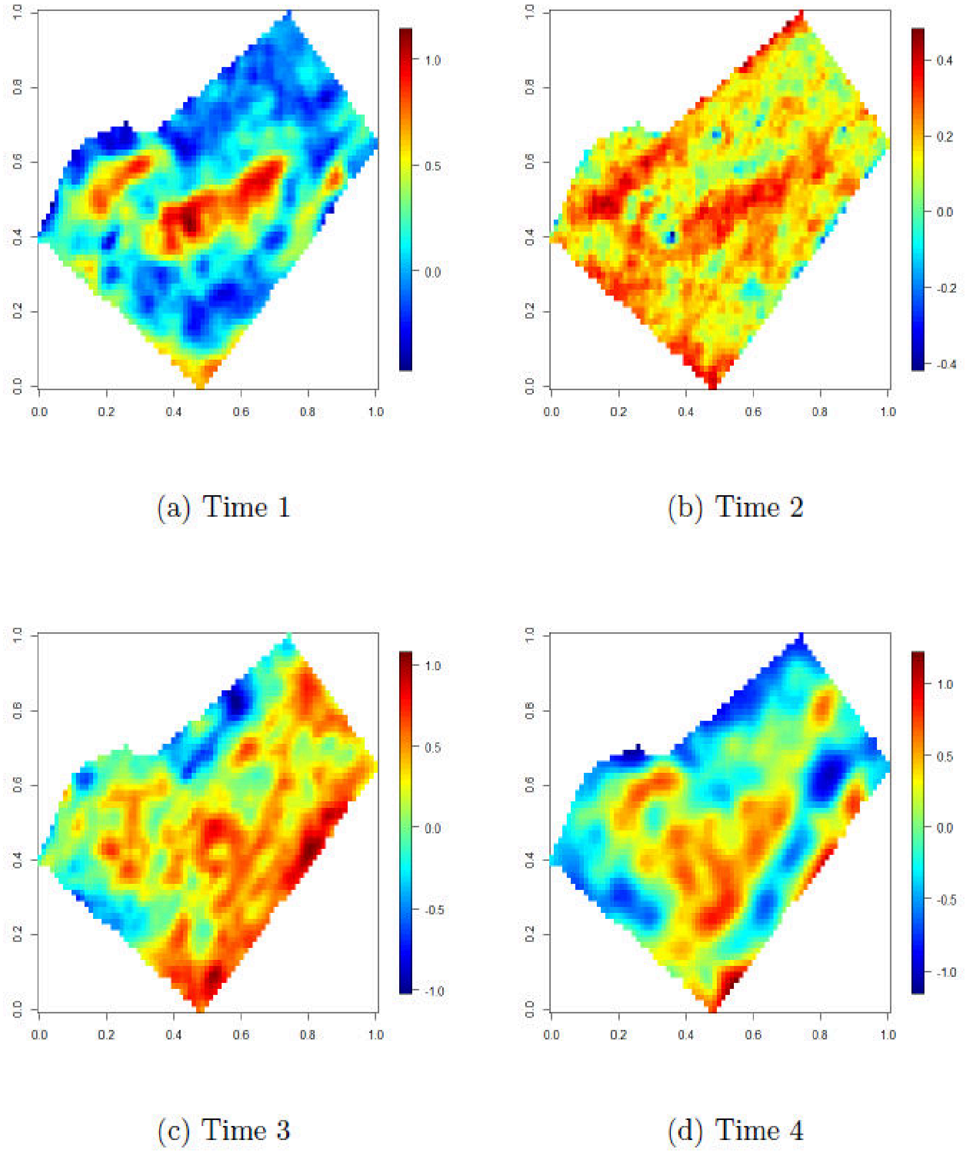


Figure 6. The spatial effects of Model 2.

essential because, in addition to the model accuracy, it also affects the estimation of the mean model. As one can see, a more complex model including spatial effects was used for real data modeling. Implementation of this model using the INLA method took about 2 hours. According to the results, using the MCMC method will take more time.

Acknowledgements

We thank Giovanna Jona Lasinio and Sara Martino for their guidance and help in this paper. We are also grateful to Alessio Pollice, Roberta Rossi, and Mariana Amato for permission to use the sensor data on soil-plant variability. Receiving support from the Center of Excellence in Analysis of Spatio-Temporal Correlated Data at Tarbiat Modares University is acknowledged.

References

1. Bonat WH and Ribeiro Jr PJ. Bayesian analysis for a class of beta mixed models. Chilean J Stat. 2014; 6(1): 3–13.
2. Bonat WH, Ribeiro Jr PJ and Zeviani WM Likelihood analysis for a class of beta mixed models. J Appl Stat. 2015; 42(2): 252-256.
3. Ferrari SLP, and Cribari-Neto F. Beta Regression for Modelling Rates and Proportions. J Appl Stat. 2004; 31: 799-815.

4. Smithson M and Verkuilen J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychol Methods. 2006 Mar; 11: 54-71.
5. Verkuilen Jand Smithson M. Mixed and mixture regression models for continuous bounded responses using the Beta distribution. J Educ Behav Stat. 2012; 37: 82-113.
6. Cepeda E and Gammerman D. Bayesian modeling of variance heterogeneity in normal regression models. Brazil J Probab Stat. 2001; 14: 207-221.
7. Cepeda-Cuervo E. Modeling Variability in Generalized Linear Models, Unpublished Ph.D. Thesis. Mathematics Institute, Universidade Federal do Rio de Janeiro. 2001; https://repositorio.unal.edu.co/handle/unal/11839.
8. Cepeda-Cuervo E, Jaimes D, Marin Mand Rojas J. Bayesian beta regression with Bayesianbetareg R-package. Computational Statistics. 2016; 31: 165-187.
9. Rue H and. Martino S. Approximate bayesian inference for hierarchical gaussian markov random fields models. J Stat Plan Infer. 2007; 137: 3177-3192.
10. Rue H and Martino S, Chopin N. Approximate bayesian inference for latent gaussian models using integrated nested laplace approximations (with discussion). J Royal Stat Soc. 2009; 71: 319-392.
11. Martino Sand Rue H. Case studies in bayesian computation using INLA, In: Mantovan P., Secchi P. (eds) complex data modeling and computationally intensive statistical methods. Contributions to Statistics. 2010; Milano: Springer. p. 99-114.
12. Tierney L and Kadane J. Accurate approximations for posterior moments and marginal densities. J Am Stat Assoc. 1986; 81: 82-86.
13. Rue H and Held I. Gaussian Markov Random Field: Theory and Application, vol 104 of Monographs on Statistics and Applied Probability, Chapman and Hall, London; 2005.
14. Pollice A, Jona Lasinio G, Rossi R, Amato M, Kneib T and Lang S. Bayesian measurement error correction in structured additive distributional regression with an application to the analysis of sensor data on soil–plant variability. Stochast Environ Res Risk Assess. 2019; 33: 747-763.