

A Bayesian approach to errors-in-variables beta regression

Jorge Figueroa-Zúñiga^a, Jalmar M. F. Carrasco^b,
Reinaldo Arellano-Valle^c and Silvia L. P. Ferrari^d

^aUniversity of Concepción

^bFederal University of Bahia

^cPontifical Catholic University of Chile

^dUniversity of São Paulo

Abstract. Beta regression models have been widely used for the analysis of limited-range continuous variables. Here, we consider an extension of the beta regression models that allows for explanatory variables to be measured with error. Then we propose a Bayesian treatment for errors-in-variables beta regression models. The specification of prior distributions is discussed, computational implementation via Gibbs sampling is provided, and two real data applications are presented. Additionally, Monte Carlo simulations are used to evaluate the performance of the proposed approach.

1 Introduction

Beta regression models have been widely employed for modeling limited-range continuous outcomes. A remarkable feature of these models is that the beta distribution is very flexible. The beta family includes left or right skewed, symmetric, J-shaped, and inverted J-shaped distributions. In addition, these models are parameterized in such a way as to allow for modeling of the mean response and a precision parameter in terms of covariates and unknown parameters; see Ferrari and Cribari-Neto (2004) and Smithson and Verkuilen (2006).

An alternative strategy to model limited-range continuous variables, such as rates and proportions, is to apply a homoscedastic normal model after transforming the dependent variable to the real line. However, this strategy is not without shortcomings. First, the model parameters cannot be easily interpreted as characteristics of the original response. Second, measures of proportions are naturally heteroscedastic and often considerably skewed, and hence it may be difficult or even impossible to find a transformation to correct for both skewness and heteroscedasticity. Beta regression models naturally accommodate asymmetry and heteroscedasticity, and allow easy parameter interpretation.

Beta regression models assume that the covariates are measured exactly. In practice, however, this assumption is not always met because some covariates may be

Key words and phrases. Bayesian analysis, beta distribution, beta regression, continuous proportions, errors-in-variables models.

Received May 2016; accepted January 2017.

not directly observable or are subject to measurement error. If errors in covariates are not properly taken into account in the inferential process, biased and inconsistent estimates may be obtained and erroneous conclusions may be reached. There is a vast literature on measurement error models. Some book-length relevant references include Buonaccorsi (2010), Cheng and JW (1999), Fuller (1987) and Carroll et al. (2006).

Recently, Carrasco, Ferrari and Arellano-Valle (2014) proposed a classical extension of the beta regression models to allow for explanatory variables to be measured with error, where identifiability conditions are necessary to obtain consistent estimates. Errors-in-variables beta regression models are specified in such a way that the distribution of the response variable, y , is assumed to depend on covariates, x , which are imprecisely measured (Carroll et al., 2006, Chapter 9), and observable variables, w , are seen as surrogates for the unobservable true covariates.

In this paper, we address errors-in-variables beta regression models under a Bayesian approach, where the classical identifiability conditions on the model parameters are replaced by the specification of prior distributions. Following Ferrari and Cribari-Neto (2004), our proposed model uses a parameterization of the beta law in terms of its mean and an additional positive parameter that can be regarded as a precision parameter. The mean of the response variable, and possibly the precision parameter, are conveniently linked with regression structures by link functions. In both cases, one or more covariates are allowed to be measured with error.

We address the issues of model fitting via Gibbs sampling, the choice of prior distributions, and model selection based on different information criteria. Simulated and real data analyses are presented for illustration. In addition, a Monte Carlo simulation study is performed to investigate the performance of the proposed approach. Various pieces of BUGS code used for fitting errors-in-variables beta regression models are presented in an Appendix.

2 Bayesian errors-in-variables beta regression

Beta regression models assume that the response variable, y , has a beta distribution with probability density function

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma[(1-\mu)\phi]} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, \quad (2.1)$$

where $\Gamma(\cdot)$ is the gamma function, $0 < \mu < 1$ and $\phi > 0$, and we write $y \sim \text{beta}(\mu\phi, (1-\mu)\phi)$. Here, $\mu = E(y)$ and ϕ is interpreted as a precision parameter because $\text{Var}(y) = \mu(1-\mu)/(1+\phi)$. The beta regression model introduced by Ferrari and Cribari-Neto (2004) assumes that y_1, y_2, \dots, y_n are n independent random variables such that each y_i follows a beta density (2.1) with mean μ_i and unknown fixed precision parameter ϕ , with

$$g(\mu_i) = \mathbf{z}_i^\top \boldsymbol{\alpha}, \quad (2.2)$$

where $\alpha \in \mathbb{R}^{p_\alpha}$ is a column vector of unknown parameters, and $\mathbf{z}_i = (z_{i1}, \dots, z_{ip_\alpha})^\top$ is a vector of p_α fixed covariates ($p_\alpha < n$). The link function $g(\cdot) : (0, 1) \rightarrow \mathbb{R}$ is assumed to be a continuous, strictly monotone and twice differentiable function. Possible choices for $g(\cdot)$ include the logit link, $g(\mu_i) = \log[\mu_i/(1 - \mu_i)]$, the probit link, $g(\mu_i) = \Phi^{-1}(\mu_i)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, and the complementary log-log link, $g(\mu_i) = \log[-\log(1 - \mu_i)]$.

The beta regression model (2.1)–(2.2) has been extended in various directions. [Smithson and Verkuilen \(2006\)](#) present a beta regression model for which the precision parameter varies across observations. [Simas, Barreto-Souza and Rocha \(2010\)](#) include non-linear structures for the regression specification of the mean and the precision parameter. Furthermore, [Zimprich \(2010\)](#) and [Figueroa-Zúniga, Arellano-Valle and Ferrari \(2013\)](#) address inference in beta regression models with mixed effects under the frequentist and Bayesian approaches, respectively.

Let $\gamma \in \mathbb{R}^{p_\gamma}$ ($p_\alpha + p_\gamma < n$) be a column vector of unknown parameters, and $\mathbf{v}_i = (v_{i1}, \dots, v_{ip_\gamma})^\top$ be a vector of fixed covariates. The beta regression model with linear specification for the transformed mean and nonconstant precision parameter assumes that each y_i follows a beta density (2.1) with mean μ_i satisfying (2.2), and precision parameter ϕ_i such that

$$h(\phi_i) = \mathbf{v}_i^\top \boldsymbol{\gamma}. \tag{2.3}$$

A natural choice for $h(\cdot)$ is $h(\phi_i) = \log(\phi_i)$.

The beta regression models described above do not involve covariates measured with error. Recently, [Carrasco, Ferrari and Arellano-Valle \(2014\)](#) proposed an errors-in-variables beta regression model to accommodate the situation in which some covariates are imprecisely measured. They considered a linear structure for the measurement error mechanism and developed a classical likelihood-based inference for the proposed model under some standard identifiability conditions on the nuisance model parameters. Here we introduce a Bayesian errors-in-variables beta regression model, in which the standard identifiability conditions are replaced by the elicitation of prior distributions, and the maximization of complex likelihood functions is unnecessary.

Following [Carrasco, Ferrari and Arellano-Valle \(2014\)](#), we propose to replace the mean submodel (2.2) and the precision submodel (2.3) by

$$g(\mu_i) = \mathbf{z}_i^\top \boldsymbol{\alpha} + \mathbf{x}_i^\top \boldsymbol{\beta}, \tag{2.4}$$

$$h(\phi_i) = \mathbf{v}_i^\top \boldsymbol{\gamma} + \mathbf{m}_i^\top \boldsymbol{\lambda}, \tag{2.5}$$

respectively, where $\boldsymbol{\beta} \in \mathbb{R}^{p_\beta}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{p_\lambda}$ are column vectors of unknown parameters, $\mathbf{x}_i = (x_{i1}, \dots, x_{ip_\beta})^\top$ and $\mathbf{m}_i = (m_{i1}, \dots, m_{ip_\lambda})^\top$ ($p_\alpha + p_\beta + p_\gamma + p_\lambda < n$) are unobservable (latent) covariates, in the sense that they are observed with error. The vectors of covariates measured without error, \mathbf{z}_i and \mathbf{v}_i , may contain variables in common, and likewise, \mathbf{x}_i and \mathbf{m}_i . Typically, $h(\phi_i)$ is specified as a submodel

of $g(\mu_i)$, so that \mathbf{v}_i and \mathbf{m}_i are subvectors of \mathbf{z}_i and \mathbf{x}_i , respectively. Let \mathbf{s}_i be the p -vector containing all the unobservable covariates \mathbf{x}_i and \mathbf{m}_i , $i = 1, \dots, n$. For $i = 1, \dots, n$, a random vector \mathbf{w}_i is observed in place of \mathbf{s}_i , and we consider the classical additive model for the measurement error

$$\mathbf{w}_i = \mathbf{s}_i + \mathbf{e}_i, \quad (2.6)$$

where \mathbf{e}_i is a vector of random errors which is assumed to be independent of \mathbf{s}_i . More specifically, in this work we assume for the measurement error model in (2.6) that

$$\mathbf{s}_i \stackrel{\text{i.i.d.}}{\sim} N_p(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \quad \mathbf{e}_i \stackrel{\text{i.i.d.}}{\sim} N_p(\mathbf{0}, \boldsymbol{\Sigma}_e), \quad (2.7)$$

where $\boldsymbol{\mu}_s \in \mathbb{R}^p$ is unknown, and $\boldsymbol{\Sigma}_s$ and $\boldsymbol{\Sigma}_e$ are unknown $p \times p$ covariance matrices. Thus, assuming (2.6) and (2.7), a convenient hierarchical representation for the proposed model is

$$\begin{aligned} y_i | \mathbf{s}_i, \mu_i, \phi_i &\stackrel{\text{ind}}{\sim} \text{beta}(\mu_i \phi_i, (1 - \mu_i) \phi_i), \\ \mathbf{w}_i | \mathbf{s}_i, \boldsymbol{\Sigma}_e &\stackrel{\text{ind}}{\sim} N_p(\mathbf{s}_i, \boldsymbol{\Sigma}_e), \\ \mathbf{s}_i | \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s &\stackrel{\text{i.i.d.}}{\sim} N_p(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \end{aligned}$$

where μ_i and ϕ_i are given in (2.4) and (2.5). Note from (2.4) and (2.5) that, for each $i = 1, \dots, n$, the response y_i is conditionally independent of the observed covariate \mathbf{w}_i given the true covariate \mathbf{s}_i .

At this point, it should be noted that the classical distinction between structural and functional models, namely whether the unobserved true covariates are considered random vectors or fixed (incidental) parameters, is blurred in the Bayesian paradigm, in which all unknown quantities are treated as random variables. In addition, for the measurement error model a distinction is usually made between the classical model defined by (2.6) and the Berkson-type model. Unlike (2.6), the Berkson model assumes that $\mathbf{s}_i = \mathbf{w}_i + \mathbf{e}_i$, and considers the observed covariates \mathbf{w}_i as known constants. [Dellaportas and Stephens \(1995\)](#) studied linear and non-linear regression models considering classical and Berkson measurement error models under a Bayesian paradigm. These authors emphasize that choice between these two models can have different consequences under a frequentist approach. Under a Bayesian formulation, however, the distinction between classical and Berkson models is unnecessary (except purely for reasons of model specification), since the unobserved covariates are treated as a set of unknown parameters. As indicated above, in this work the classical model given in (2.6) is considered.

To complete the Bayesian specification of the errors-in-variables beta regression model described above, 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 submodel, i.e., $\boldsymbol{\alpha} \sim N_{p_\alpha}(\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha)$,

$\beta \sim N_{p_\beta}(\mu_\beta, \Sigma_\beta)$, and for the mean vector of the true covariate s_i , i.e., $\mu_s \sim N_p(\xi_s, \Omega_s)$. Vague priors are usually specified by taking large values for the prior variances. However, the impact of the scale choice under the normal model cannot be neglected. We propose to replace the multivariate normal priors by multivariate t distributions, that is,

$$\alpha \sim t_{p_\alpha}(v_\alpha, \mu_\alpha, \Sigma_\alpha), \quad \beta \sim t_{p_\beta}(v_\beta, \mu_\beta, \Sigma_\beta), \quad \mu_s \sim t_p(v_s, \xi_s, \Omega_s),$$

and to specify appropriate values for the degrees of freedom parameters. Prior distributions for the covariance matrices Σ_s and Σ_e are chosen, mainly for computational simplicity, to be inverted Wishart distributions as in Fong, Rue and Wakefield (2010), that is,

$$\Sigma_s \sim IW_p(\Psi_s, a_s), \quad \Sigma_e \sim IW_p(\Psi_e, a_e).$$

If $\Sigma_s = \text{diag}\{\sigma_{s1}^2, \dots, \sigma_{sp}^2\}$ and $\Sigma_e = \text{diag}\{\sigma_{e1}^2, \dots, \sigma_{ep}^2\}$, then we consider a product of inverted gamma prior distributions for σ_{si}^2 and σ_{ei}^2 , that is, we assume that $\sigma_{s1}^2, \dots, \sigma_{sn}^2$ (similarly, $\sigma_{e1}^2, \dots, \sigma_{en}^2$) are independent and have an inverted gamma prior distribution, $IG(\varepsilon_1, \varepsilon_2)$. Typically, $\varepsilon_1 = \varepsilon_2 = \varepsilon$, a small positive value, to obtain a slightly informative prior. Following Figueroa-Zúñiga, Arellano-Valle and Ferrari (2013), we propose an alternative, less informative prior distribution for σ_{si}^2 (similarly for σ_{ei}^2) given by

$$\sigma_{si}^2 \stackrel{d}{=} (aB)^2, \quad B \sim \text{beta}(1 + \varepsilon, 1 + \varepsilon),$$

where $\stackrel{d}{=}$ represents equality in distribution, with a positive value for a and a small positive value for ε (see Figure 1). Note that the smaller the value of ε , the flatter the prior density will be.

We now turn to the specification of prior distributions for the precision parameter. We consider the following errors-in-variables beta regression models.

Model 1—Constant precision case. This model considers the errors-in-variables regression model given by (2.1) and (2.4) with a common precision parameter ϕ for all the observations. If a slightly informative prior is required, it can be assumed that $\phi \sim IG(\varepsilon, \varepsilon)$, with a small fixed positive value for ε . A less informative prior distribution for ϕ is

$$\phi \stackrel{d}{=} (aB)^2, \quad B \sim \text{beta}(1 + \varepsilon, 1 + \varepsilon),$$

for a positive value for a and a small positive value for ε . As suggested by Gelman (2006) the prior distribution $\phi = U^2$ with $U \sim U(0, a)$ with large a ($a = 50$ for example) is less informative than an inverse gamma prior; the prior distribution for ϕ we suggest is then a more flexible version of Gelman’s prior as shown in Figure 1. Our simulations indicate, however, that a more convenient prior for ϕ is a log-normal distribution, say $\phi \sim \text{LN}(\mu_\phi, \sigma_\phi^2)$.

Model 2—Varying precision case. This model considers the errors-in-variables regression model given by (2.1), (2.4) and (2.5). Here, the specification of prior distributions for γ and λ is similar to that used for α and β .

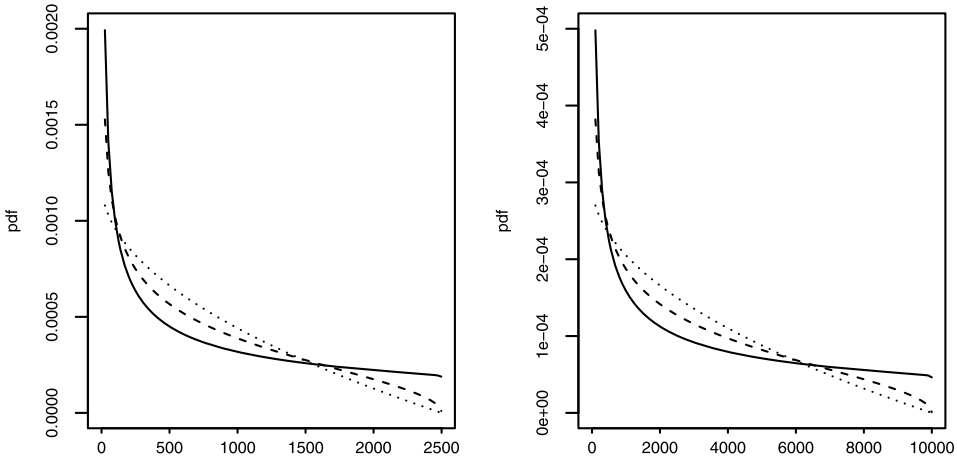


Figure 1 Probability density function of $(aB)^2$, where $B \sim \text{beta}(1 + \varepsilon, 1 + \varepsilon)$, with $a = 50$ (left hand) and $a = 100$ (right hand), and different values for ε : $\varepsilon = 0.01$ (solid line), $\varepsilon = 0.5$ (dashed line) and $\varepsilon = 1$ (dotted line).

3 Model fitting using Markov chain Monte Carlo sampling

Let $\mathbf{y} = (y_1, \dots, y_n)^\top$ and $\mathbf{w} = (\mathbf{w}_1^\top, \dots, \mathbf{w}_n^\top)^\top$ be the observable variables, and consider the vector $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)^\top$, where $\eta_i = g(\mu_i)$ is given in (2.4). We omit the observable vectors \mathbf{z}_i and \mathbf{v}_i in the notation as they are non-random and known. From (2.4) and (2.7) it can be noticed that, conditionally on $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\mu}_s$ and $\boldsymbol{\Sigma}_s$, the η_i s are independent and normally distributed.

We now present the joint posterior distribution under Models 1 and 2 described in the previous section. Under Model 1, for which $\mathbf{s}_i = \mathbf{x}_i$, and the assumption that the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\mu}_s$, $\boldsymbol{\Sigma}_s$, $\boldsymbol{\Sigma}_e$ and ϕ are independent, the joint posterior density is

$$\begin{aligned}
 f(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s, \boldsymbol{\Sigma}_e, \boldsymbol{\eta}, \phi \mid \mathbf{y}, \mathbf{w}) &\propto \left[\prod_{i=1}^n f(y_i, \mathbf{w}_i \mid \eta_i, \phi, \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s, \boldsymbol{\Sigma}_e) \right] \\
 &\times f(\boldsymbol{\eta} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s) f(\boldsymbol{\alpha}) f(\boldsymbol{\beta}) \\
 &\times f(\boldsymbol{\mu}_s) f(\boldsymbol{\Sigma}_s) f(\boldsymbol{\Sigma}_e) f(\phi).
 \end{aligned}$$

Gibbs sampling can be used to generate a Monte Carlo sample from the joint posterior density, $f(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s, \boldsymbol{\Sigma}_e, \boldsymbol{\eta}, \phi \mid \mathbf{y}, \mathbf{w})$. The Gibbs sampler in this context involves iteratively sampling from the full conditional distributions which can be implemented in the WinBUGS software. Posterior inferences on $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\mu}_s$, $\boldsymbol{\Sigma}_s$, $\boldsymbol{\Sigma}_e$, and ϕ and on the mean responses μ_i , for $i = 1, \dots, n$, are readily obtained in WinBUGS.

We now turn to Model 2. Let $\boldsymbol{\rho} = (\rho_1, \dots, \rho_n)^\top$, where $\rho_i = h(\phi_i)$ is given in (2.5). From (2.5) and (2.7), we have that, conditionally on $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}_s$ and $\boldsymbol{\Sigma}_s$, the

(η_i, ρ_i) s are independent and normally distributed. Assuming prior independence of $\alpha, \beta, \mu_s, \Sigma_s, \gamma, \lambda$ and Σ_e , we obtain the posterior density given by

$$\begin{aligned}
 & f(\alpha, \beta, \mu_s, \Sigma_s, \gamma, \lambda, \Sigma_e, \eta, \rho \mid \mathbf{y}, \mathbf{w}) \\
 & \propto \left[\prod_{i=1}^n f(y_i, \mathbf{w}_i \mid \eta_i, \rho_i, \mu_s, \Sigma_s, \Sigma_e) \right] \\
 & \quad \times f(\eta \mid \alpha, \beta, \mu_s, \Sigma_s) f(\rho \mid \eta, \gamma, \lambda, \mu_s, \Sigma_s) \\
 & \quad \times f(\alpha) f(\beta) f(\mu_s) f(\Sigma_s) f(\gamma) f(\lambda) f(\Sigma_e).
 \end{aligned}$$

Similarly to Model 1, the Gibbs sampling can be used to generate a Monte Carlo sample from $f(\alpha, \beta, \mu_s, \Sigma_s, \gamma, \lambda, \Sigma_e, \eta, \rho \mid \mathbf{y}, \mathbf{w})$. In this case, again the Gibbs sampler involves iteratively sampling from the full conditional distributions which can be implemented in the WinBUGS software. Thus, posterior inferences on $\alpha, \beta, \mu_s, \Sigma_s, \gamma, \lambda, \Sigma_e$, and ϕ_i for $i = 1, \dots, n$, and on the mean responses μ_i , for $i = 1, \dots, n$, are again easily obtained in WinBUGS.

4 Illustration via simulations

To illustrate the proposed methodology, we consider the following errors-in-variables beta regression model with simulated data (Model 1):

$$\begin{aligned}
 & y_i \mid x_i, w_i, \alpha_1, \alpha_2, \beta, \phi \sim \text{beta}(\mu_i \phi, (1 - \mu_i) \phi), \\
 & \log\left(\frac{\mu_i}{1 - \mu_i}\right) = \alpha_1 + \alpha_2 z_i + \beta x_i, \\
 & w_i = x_i + e_i, \quad x_i \sim N(\mu_x, \sigma_x^2), e_i \sim N(0, \sigma_e^2), i = 1, \dots, 100
 \end{aligned}$$

with x_i and $e_{i'}$, for $i, i' = 1, \dots, 100$, being independent.

For our simulation study, the values of the covariates z_i were generated from a uniform distribution in the unit interval, and we set $\alpha_1 = 2, \alpha_2 = 1, \beta = -0.4, \mu_x = 5, \sigma_x^2 = 2, \phi = 50$ and 300 . Typically, the measurement error variance, σ_e^2 , or the reliability ratio, $k = \sigma_x^2 / (\sigma_x^2 + \sigma_e^2)$, is considered known. At the outset, we consider σ_e^2 known and equal to 0.1, 0.66 and 2. The corresponding values for k are approximately 0.95 (low measurement error), 0.75 (moderate measurement error) and 0.5 (high measurement error), respectively. As proposed, we adopt the following prior specifications for $\alpha_1, \alpha_2, \beta$ and μ_x :

$$\begin{aligned}
 & \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \sim t_2\left(5, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 20 & 0 \\ 0 & 20 \end{pmatrix}\right), \\
 & \beta \sim t(5, 0, 20), \quad \mu_x \sim t(5, 0, 20).
 \end{aligned}$$

We first analyze 100 simulated datasets under different prior specifications for the precision parameter ϕ and the variance parameter σ_x^2 . The following prior distributions for ϕ and σ_x^2 are considered:

- (i) Model 1a: ϕ and $\sigma_x^2 \sim \text{IG}(\varepsilon, \varepsilon)$, with $\varepsilon = 0.01$;
- (ii) Model 1b: $\phi \stackrel{d}{=} (50B)^2$ and $\sigma_x^2 \stackrel{d}{=} (50B)^2$, where $B \sim \text{beta}(1 + \varepsilon, 1 + \varepsilon)$, with $\varepsilon = 0.1$;
- (iii) Model 1c: $\phi \sim \text{LN}(\mu_\phi, \sigma_\phi^2)$ with $\mu_\phi = 4$, $\sigma_\phi^2 = 1$ and $\sigma_x^2 \stackrel{d}{=} (50B)^2$, where $B \sim \text{beta}(1 + \varepsilon, 1 + \varepsilon)$, with $\varepsilon = 0.1$. Note that this prior specification for ϕ implies that $E(\phi) \cong 90.02$ and $\text{Var}(\phi) \cong 13,923.38$ and that the 1st and 99th percentiles of ϕ are approximately 5.33 and 559.11. These hyperparameter values were carefully chosen in order to produce a disperse prior distribution and to avoid very small estimates for ϕ .

All the numerical results presented in this paper were obtained using WinBUGS by considering 140,000 Monte Carlo iterations and discarding the first 50,000 as burn-in. For each of the different prior distributions considered, we have not observed problems of autocorrelation, convergence and stationarity in the chain.¹

Tables 1 and 2 report the average of the deviance information criterion (DIC) proposed by Spiegelhalter et al. (2002). This criterion is based on the posterior mean of the deviance and it can be approximated by $\overline{D} = \sum_{q=1}^Q D(\theta_q)/Q$, where $\theta_1, \dots, \theta_Q$ is a sample of size Q drawn from $\pi(\theta | y, w)$ after the burn-in period and $D(\theta) = -2 \sum_{i=1}^n \log[f(y_i | w_i, \theta)]$. The DIC criterion can be estimated using the MCMC output by $\widehat{\text{DIC}} = \overline{D} + p_D = 2\overline{D} - \widehat{D}$, where p_D is the effective number of parameters, defined by $p_D = E\{D(\theta)\} - D\{E(\theta)\}$, with $D\{E(\theta)\}$ being the deviance evaluated at the posterior mean, and $\widehat{D} = \sum_{q=1}^Q \theta_q/Q$. For comparison of two alternative models, the model that better fits a data set is the model with the smallest value of the DIC.

In addition, Tables 1 and 2 report the average of the expected Akaike information criterion (EAIC) introduced by Brooks (2002) and the expected Bayesian information criterion (EBIC) given in Carlin and Louis (2001), which can be estimated by means of $\widehat{\text{EAIC}} = \overline{D} + 2p_D$ and $\widehat{\text{EBIC}} = \overline{D} + p_D \log(n)$, respectively. The averages are based on the 100 simulated datasets for the fitted models with different prior distributions for ϕ and for σ_x^2 . The goal is to minimize the information criteria to identify the best model. We observe that the different proposed priors lead to similar DICs, EAICs and EBICs when the measurement error is low ($k = 0.95$). However, the differences among the criteria increase as k decreases, and for moderate and high measurement error ($k = 0.75$ and $k = 0.5$, respectively) the three criteria indicate that Model 1c provides a better fit than the other proposals.

Tables 3 and 4 present the relative bias (RelBias) and the root-mean-squared error ($\sqrt{\text{MSE}}$) for each parameter estimator over the 100 simulated samples under

¹Pieces of BUGS code used for fitting errors-in-variables beta regression models in the simulated data example are presented in the Appendix.

Table 1 Mean DIC, EAIC and EBIC based on 100 simulated datasets with $\phi = 50$, for the fitted models with different prior specifications for the precision parameter; Model 1

k	Model	Prior for ϕ and σ_x^2	DIC	EAIC	EBIC
0.95	Model 1a	$\phi, \sigma_x^2 \sim \text{IG}(0.01, 0.01)$	-118.41	-205.94	-190.31
	Model 1b	$\phi, \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	-118.34	-206.32	-190.69
	Model 1c	$\phi \sim \text{LN}(4, 1), \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	-118.29	-205.92	-190.29
0.75	Model 1a	$\phi, \sigma_x^2 \sim \text{IG}(0.01, 0.01)$	39.86	-38.86	-23.23
	Model 1b	$\phi, \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	33.86	-38.99	-23.36
	Model 1c	$\phi \sim \text{LN}(4, 1), \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	24.49	-58.36	-42.73
0.50	Model 1a	$\phi, \sigma_x^2 \sim \text{IG}(0.01, 0.01)$	153.31	79.59	95.22
	Model 1b	$\phi, \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	140.42	73.12	88.75
	Model 1c	$\phi \sim \text{LN}(4, 1), \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	128.04	54.31	69.94

Table 2 Mean DIC, EAIC and EBIC based on 100 simulated datasets with $\phi = 300$, for the fitted models with different prior specifications for the precision parameter; Model 1

k	Model	Prior for ϕ and σ_x^2	DIC	EAIC	EBIC
0.95	Model 1a	$\phi, \sigma_x^2 \sim \text{IG}(0.01, 0.01)$	-127.32	-210.98	-195.33
	Model 1b	$\phi, \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	-127.33	-211.10	-195.47
	Model 1c	$\phi \sim \text{LN}(4, 1), \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	-127.42	-211.43	-195.80
0.75	Model 1a	$\phi, \sigma_x^2 \sim \text{IG}(0.01, 0.01)$	38.71	-37.52	-21.89
	Model 1b	$\phi, \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	24.61	-30.42	-14.79
	Model 1c	$\phi \sim \text{LN}(4, 1)$ and $\sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	17.57	-50.02	-34.39
0.50	Model 1a	$\phi, \sigma_x^2 \sim \text{IG}(0.01, 0.01)$	177.24	97.21	112.84
	Model 1b	$\phi, \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	168.41	93.85	109.48
	Model 1c	$\phi \sim \text{LN}(4, 1), \sigma_x^2 \stackrel{d}{=} (50B)^2, B \sim \text{beta}(1.1, 1.1)$	145.10	76.61	92.24

the different settings. They are defined as

$$\text{RelBias}(\theta) = \frac{1}{100} \sum_{i=1}^{100} \left(\frac{\hat{\theta}^{(i)} - \theta}{\theta} \right) \quad \text{and} \quad \text{MSE}(\theta) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}^{(i)} - \theta)^2,$$

where θ represents any particular parameter, and $\hat{\theta}^{(i)}$ is the posterior estimate of θ for the i th sample. The figures in Tables 3 and 4 show that the different proposed priors achieve reasonable estimates of the model parameters. However, the prior distributions considered in Model 1c, namely $\phi \sim \text{LN}(4, 1)$ and $\sigma_x^2 \stackrel{d}{=} (50B)^2$, where $B \sim \text{beta}(1.1, 1.1)$, produces the smallest $\sqrt{\text{MSE}}$.

Table 3 Relative bias and root-mean-square error based on 100 simulated datasets with $\phi = 50$ and different prior specifications for the precision parameter; Model 1

k	Model		Posterior inference					
			β	α_1	α_2	$\log(\phi)$	μ_x	σ_x^2
0.95	Model 1a	RelBias	0.000	-0.001	0.009	0.021	-0.001	0.010
		$\sqrt{\text{MSE}}$	0.022	0.129	0.114	0.204	0.146	0.286
	Model 1b	RelBias	0.001	0.000	0.008	0.025	-0.001	0.022
		$\sqrt{\text{MSE}}$	0.022	0.129	0.114	0.212	0.146	0.292
	Model 1c	RelBias	0.002	0.004	-0.007	0.021	-0.004	0.029
		$\sqrt{\text{MSE}}$	0.022	0.130	0.117	0.200	0.138	0.284
0.75	Model 1a	RelBias	0.061	0.056	0.013	0.274	-0.004	-0.021
		$\sqrt{\text{MSE}}$	0.051	0.265	0.156	2.273	0.154	0.378
	Model 1b	RelBias	0.062	0.059	0.010	0.143	-0.004	-0.006
		$\sqrt{\text{MSE}}$	0.049	0.258	0.154	0.824	0.153	0.378
	Model 1c	RelBias	0.038	0.038	-0.003	0.078	-0.006	0.010
		$\sqrt{\text{MSE}}$	0.041	0.219	0.153	0.520	0.150	0.351
0.5	Model 1a	RelBias	0.210	0.204	0.016	1.138	-0.006	-0.089
		$\sqrt{\text{MSE}}$	0.121	0.603	0.211	5.802	0.175	0.623
	Model 1b	RelBias	0.224	0.218	0.015	0.439	-0.007	-0.066
		$\sqrt{\text{MSE}}$	0.150	0.764	0.192	1.843	0.175	0.627
	Model 1c	RelBias	0.067	0.065	-0.004	0.114	-0.008	0.017
		$\sqrt{\text{MSE}}$	0.065	0.328	0.186	0.625	0.177	0.504

In order to investigate the prior’s impact on the model selection for smaller samples, we repeated the above analysis for $n = 30$ and $n = 50$. The results are shown in a supplement available from the authors upon request. Overall, the results are qualitatively the same as those obtained for $n = 100$.

In Tables 5 and 6, we report the parameter estimates for one dataset generated from Model 1c ($n = 100$). The estimated parameters obtained from the Bayesian methodology proposed here are similar to the true values of the model parameters. In addition, the necessary diagnostic tests (such as convergence, autocorrelation, history) were performed, from which desirable behaviors were observed in the chains (for brevity detailed numerical results are not shown but are commented on below). We also conducted a sensitivity analysis with respect to the specifications of the parameter k . In each case, the posterior inferences were not appreciably altered, but the standard deviation (s.d.) for each parameter increases when k decreases.

The multivariate version of Gelman and Rubin’s convergence diagnostic proposed by Brooks and Gelman (1998) indicates that the chain is convergent because the multivariate proportional scale reduction factor (m.p.r.f.) equals 1.01 (< 1.2).

Table 4 Relative bias and root-mean-square error based on 100 simulated datasets with $\phi = 300$ and different prior specifications for the precision parameter; Model 1

k	Model		Posterior inference					
			β	α_1	α_2	$\log(\phi)$	μ_x	σ_x^2
0.95	Model 1a	RelBias	0.011	0.010	-0.001	0.141	-0.001	0.007
		$\sqrt{\text{MSE}}$	0.016	0.084	0.062	2.271	0.147	0.286
	Model 1b	RelBias	0.008	0.005	0.000	0.041	-0.004	0.034
		$\sqrt{\text{MSE}}$	0.031	0.155	0.116	0.373	0.153	0.376
	Model 1c	RelBias	0.003	0.003	-0.003	0.015	-0.004	0.030
		$\sqrt{\text{MSE}}$	0.015	0.082	0.063	0.349	0.139	0.285
0.75	Model 1a	RelBias	0.037	0.030	0.009	0.684	-0.005	-0.001
		$\sqrt{\text{MSE}}$	0.037	0.183	0.131	4.862	0.158	3.897
	Model 1b	RelBias	0.037	0.031	0.001	0.080	-0.007	0.038
		$\sqrt{\text{MSE}}$	0.064	0.323	0.156	0.578	0.177	0.575
	Model 1c	RelBias	-0.017	-0.018	-0.004	-0.057	-0.006	0.052
		$\sqrt{\text{MSE}}$	0.029	0.155	0.112	0.468	0.150	0.369
0.5	Model 1a	RelBias	0.032	0.025	0.007	0.860	-0.006	0.037
		$\sqrt{\text{MSE}}$	0.053	0.278	0.171	6.139	0.186	0.569
	Model 1b	RelBias	0.189	0.173	0.007	0.337	-0.013	-0.114
		$\sqrt{\text{MSE}}$	0.118	0.581	0.159	1.026	0.176	0.601
	Model 1c	RelBias	-0.053	-0.056	-0.009	-0.141	-0.008	0.108
		$\sqrt{\text{MSE}}$	0.050	0.259	0.154	0.862	0.177	0.555

Also, for each parameter, we checked that the convergence is achieved for each chain.

We now use the same simulated datasets from the beginning of this section to fit Model 2 with three different regression structures for the precision parameter. Note that the true (unknown) model is a beta regression model with constant precision, and hence only Model 2a (see Tables 7 and 8 below) corresponds to the true model. Note also that the ϕ_i s are not assumed to be constant over the observations under Models 2b and 2c (see Tables 9 and 10 below), and we then report $\overline{\log(\phi)} = 1/100 \sum_{i=1}^{100} \log(\phi_i)$ to allow comparisons with the real value of $\log(\phi)$.

Prior distributions for $\beta, \alpha_1, \alpha_2, \mu_x$ are the same as those proposed for Model 1 and the prior distribution for σ_x^2 is the same as the one proposed for Models 1b and 1c. For model specifications that include measurement error for the precision parameter, we assume that the parameter λ has the same distribution as the parameter β , namely $t(\nu_\beta, 0, \sigma_\beta^2)$ and the parameters γ_1 and γ_2 have the same distribution as the parameters α_1 and α_2 , namely $t(\nu_{\alpha_1}, 0, \sigma_{\alpha_1}^2)$, and $t(\nu_{\alpha_2}, 0, \sigma_{\alpha_2}^2)$ respectively.

Tables 7 and 8 report the mean DIC, EAIC and EBIC based on the 100 simulated datasets for the three fitted models using simulated data for $\phi = 50$ and $\phi = 300$,

Table 5 True mean and estimated posterior means and standard deviation, with $\phi = 50$ for Model 1c; simulated dataset

k	Parameter	Posterior inference		
		True	Mean	s.d.
0.95	β	-0.4	-0.397	0.027
	α_1	2.0	2.002	0.143
	α_2	1.0	1.011	0.101
	$\log(\phi)$	3.912	4.132	0.174
	μ_x	5.0	4.954	0.131
	σ_x^2	2.0	1.610	0.251
0.75	β	-0.4	-0.400	0.039
	α_1	2.0	2.087	0.212
	α_2	1.0	0.896	0.141
	$\log(\phi)$	3.912	4.138	0.348
	μ_x	5.0	5.057	0.171
	σ_x^2	2.0	2.224	0.426
0.5	β	-0.4	-0.378	0.066
	α_1	2.0	1.978	0.363
	α_2	1.0	0.830	0.170
	$\log(\phi)$	3.912	3.969	0.500
	μ_x	5.0	5.147	0.201
	σ_x^2	2.0	2.003	0.557

respectively. Model 2a, the model under which the data were simulated and which is equivalent but less informative than Model 1c, provides a better fit than the other proposals when $\phi = 50$, and when $\phi = 300$ and the measurement error is low ($k = 0.95$). When $\phi = 300$ and the measurement error is moderate or high ($k = 0.75$ and $k = 0.5$), DIC and EAIC are slightly smaller for Model 2b than for Model 2a, but the smallest EBIC corresponds to Model 2a.

Tables 9 and 10 present the relative bias and the root-mean-squared error for each parameter estimator over the simulated samples under the different settings. We observe that the different proposed priors achieve reasonable estimates of the model parameters, but when the value of k is not very large, Model 2a exhibits more stability in the estimation of ϕ .

Tables 11 and 12 report the parameter estimates for one dataset under Model 2a for each value of k . The estimates obtained through the Bayesian methodology proposed here are similar to the corresponding true values of the parameters. Diagnostic tools suggest that the chain for each parameter is convergent, not correlated and stationary. Hence, our estimates can be regarded as reliable.

In the simulation studies reported above, the measurement error variance, σ_e^2 , is assumed to be known. Next, we study the estimates of the parameters of Model 1c,

Table 6 True mean and estimated posterior means and standard deviation, with $\phi = 300$ for Model 1c; simulated dataset

k	Parameter	Posterior inference		
		True	Mean	s.d.
0.95	β	-0.4	-0.413	0.017
	α_1	2.0	1.984	0.073
	α_2	1.0	0.960	0.055
	$\log(\phi)$	5.704	5.609	0.303
	μ_x	5.0	5.000	0.125
	σ_x^2	2.0	2.181	0.336
0.75	β	-0.4	-0.374	0.027
	α_1	2.0	1.992	0.140
	α_2	1.0	0.873	0.107
	$\log(\phi)$	5.704	5.719	0.526
	μ_x	5.0	4.967	0.170
	σ_x^2	2.0	2.286	0.415
0.5	β	-0.4	-0.411	0.084
	α_1	2.0	2.200	0.434
	α_2	1.0	0.924	0.145
	$\log(\phi)$	5.704	5.291	0.675
	μ_x	5.0	5.080	0.205
	σ_x^2	2.0	2.409	0.581

assuming that σ_e^2 is unknown. The prior distribution for σ_e^2 is taken as the same prior distribution used for σ_x^2 in Models 1b and 1c.

In Tables 13 and 14, we report the mean, the relative bias and the root-mean-squared error for each parameter estimator over 100 simulated samples under the different settings. We observe that for different values for σ_e^2 and k , good estimates of all the parameters are achieved both for $\phi = 50$ and $\phi = 300$.

5 Real data applications

To illustrate our Bayesian approach to errors-in-variables beta regression in practice, we apply the proposed methods to two real datasets. First, we consider the data studied by Coakley and Rust (1968), who investigated 39 sediment samples taken at different depths from an arctic lake and then classified according to their relative amounts of sand, silt, and clay. We will focus on the analysis of the relative amount of clay or, in other words, the proportion of clay (y), which can be seen as a continuous variable with limited range (0, 1). According to the authors, “water depths were obtained from a winch cable meter, because a hammer seismograph designed for this purpose was inefficient due to ice reverberation.” It is reasonable

Table 7 Mean DIC, EAIC and EBIC based on 100 simulated datasets for the fitted models with $\phi = 50$ and different prior specifications for the precision parameter under Model 2; simulated dataset

k	Model	Precision submodel	DIC	EAIC	EBIC
0.95	Model 2a	$\log(\phi_i) = \gamma_1$	-118.44	-206.01	-190.37
	Model 2b	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i$	-118.42	-205.77	-187.53
	Model 2c	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i + \lambda x_i$	-115.40	-198.64	-177.80
0.75	Model 2a	$\log(\phi_i) = \gamma_1$	29.28	-44.54	-23.70
	Model 2b	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i$	33.00	-41.72	-23.49
	Model 2c	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i + \lambda x_i$	71.74	-2.90	17.94
0.5	Model 2a	$\log(\phi_i) = \gamma_1$	115.75	43.84	59.47
	Model 2b	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i$	118.85	45.92	64.16
	Model 2c	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i + \lambda x_i$	239.14	155.83	176.68

Table 8 Mean DIC, EAIC and EBIC based on 100 simulated datasets for the fitted models with $\phi = 300$ and different prior specifications for the precision parameter; Model 2

k	Model	Precision submodel	DIC	EAIC	EBIC
0.95	Model 2a	$\log(\phi_i) = \gamma_1$	-127.49	-211.85	-196.22
	Model 2b	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i$	-127.33	-211.67	-193.43
	Model 2c	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i + \lambda x_i$	-122.85	-203.00	-182.16
0.75	Model 2a	$\log(\phi_i) = \gamma_1$	16.43	-51.11	-35.48
	Model 2b	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i$	16.01	-51.79	-33.56
	Model 2c	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i + \lambda x_i$	52.69	-12.71	8.13
0.5	Model 2a	$\log(\phi_i) = \gamma_1$	142.10	73.94	89.57
	Model 2b	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i$	141.85	72.92	91.16
	Model 2c	$\log(\phi_i) = \gamma_1 + \gamma_2 z_i + \lambda x_i$	214.31	142.24	163.08

to admit that the water depth is measured with error. Our goal is to model the proportion of clay (y) using water depth (x , in meters) as a (latent) covariate measured with error. The water depth measured with the winch cable is denoted by depth.

The model (hereafter Model 3) under consideration is defined as follows:

$$\begin{aligned}
 y_i | x_i, \alpha, \beta, \gamma, \lambda &\overset{\text{ind}}{\sim} \text{beta}(\mu_i \phi_i, (1 - \mu_i) \phi_i), \\
 \log\left(\frac{\mu_i}{1 - \mu_i}\right) &= \alpha + \beta x_i, \\
 \log(\phi_i) &= \gamma + \lambda x_i, \\
 \text{depth}_i | x_i, \sigma_e^2 &\overset{\text{ind}}{\sim} N(x_i, \sigma_e^2), \\
 x_i | \mu_x, \sigma_x^2 &\overset{\text{ind}}{\sim} N(\mu_x, \sigma_x^2), \quad \text{for } i = 1, \dots, 39.
 \end{aligned}$$

Table 9 Relative bias and root-mean-square error based on 100 simulated datasets with $\phi = 50$ and different structures for the precision parameter; Model 2

k	Model		Posterior inference					
			β	α_1	α_2	$\overline{\log(\phi)}$	μ_x	σ_x^2
0.95	Model 2a	RelBias	0.001	0.000	0.008	0.021	-0.001	0.022
		$\sqrt{\text{MSE}}$	0.022	0.130	0.114	0.198	0.146	0.293
	Model 2b	RelBias	0.000	-0.002	0.012	0.014	-0.004	0.017
		$\sqrt{\text{MSE}}$	0.022	0.130	0.120	0.178	0.139	0.301
	Model 2c	RelBias	0.002	0.000	0.009	0.020	-0.004	0.015
		$\sqrt{\text{MSE}}$	0.023	0.132	0.124	0.192	0.139	0.301
0.75	Model 2a	RelBias	0.034	0.030	0.011	0.080	-0.004	0.012
		$\sqrt{\text{MSE}}$	0.041	0.212	0.150	0.532	0.153	0.377
	Model 2b	RelBias	0.045	0.041	0.021	0.102	-0.005	0.005
		$\sqrt{\text{MSE}}$	0.046	0.243	0.159	0.638	0.146	0.396
	Model 2c	RelBias	0.116	0.111	0.024	0.471	-0.006	-0.045
		$\sqrt{\text{MSE}}$	0.073	0.368	0.167	2.752	0.150	0.421
0.5	Model 2a	RelBias	0.067	0.060	0.011	0.142	-0.007	0.024
		$\sqrt{\text{MSE}}$	0.067	0.344	0.185	0.788	0.174	0.541
	Model 2b	RelBias	0.060	0.055	0.026	0.168	-0.006	0.037
		$\sqrt{\text{MSE}}$	0.066	0.343	0.187	0.911	0.168	0.584
	Model 2c	RelBias	0.202	0.192	0.030	1.200	-0.009	-0.051
		$\sqrt{\text{MSE}}$	0.112	0.538	0.205	5.332	0.183	0.618

As before, we assume that, a priori, all the parameters are independent and that α , β , γ , λ and μ_x follow a $t(5, 0, 20)$ distribution, $\sigma_x^2 \stackrel{d}{=} (50B)^2$ and $\sigma_e^2 \stackrel{d}{=} (50B)^2$, where $B \sim \text{beta}(1.1, 1.1)$; for the constant precision case, $\phi \sim \text{LN}(4, 1)$.

Table 15 gives the DIC, EAIC and EBIC for the model fitting with different specifications for the precision parameter. Results reported under the label “Model 1 (naïve)” refer to the situation in which one ignores the measurement error. The smallest DIC, EAIC and EBIC are achieved for Model 2, that is, the model that assumes that the precision parameter is constant and takes the measurement error into account, that is, according to these criterion Model 2 is the “best” among its competitors.

Table 16 gives the posterior estimates of the parameters associated with Model 2, which provides the best fit for the data. The estimated reliability ratio, $k = 0.836$, indicates the presence of moderate measurement error in the covariate. Because $\hat{\beta}$ is positive, there is an indication that the mean proportion of clay increases with depth. Furthermore, as the logit link is employed, $\exp(c\beta)$ has an interpretation as odds ratios when the regressor is increased in c units. Hence, be-

Table 10 Relative bias and root-mean-square error based on 100 simulated datasets with $\phi = 300$ and different structures for the precision parameter; Model 2

k	Model		Posterior inference					
			β	α_1	α_2	$\overline{\log(\phi)}$	μ_x	σ_x^2
0.95	Model 2a	RelBias	0.004	0.001	0.006	0.025	-0.002	0.018
		$\sqrt{\text{MSE}}$	0.015	0.079	0.062	0.425	0.150	0.314
	Model 2b	RelBias	0.008	0.008	-0.001	0.064	-0.004	0.016
		$\sqrt{\text{MSE}}$	0.016	0.087	0.064	0.866	0.138	0.301
	Model 2c	RelBias	0.020	0.020	-0.003	0.274	-0.004	0.005
		$\sqrt{\text{MSE}}$	0.019	0.098	0.064	2.903	0.139	0.299
0.75	Model 2a	RelBias	-0.012	-0.014	-0.007	-0.002	-0.003	0.042
		$\sqrt{\text{MSE}}$	0.031	0.158	0.115	0.577	0.152	0.387
	Model 2b	RelBias	0.003	0.001	0.005	0.072	-0.005	0.030
		$\sqrt{\text{MSE}}$	0.032	0.169	0.111	0.852	0.145	0.399
	Model 2c	RelBias	0.041	0.035	0.021	0.891	-0.007	0.023
		$\sqrt{\text{MSE}}$	0.041	0.205	0.146	6.548	0.155	0.493
0.5	Model 2a	RelBias	-0.052	-0.058	-0.007	-0.110	-0.007	0.111
		$\sqrt{\text{MSE}}$	0.052	0.262	0.157	0.757	0.172	0.590
	Model 2b	RelBias	-0.035	-0.041	0.010	-0.059	-0.007	0.097
		$\sqrt{\text{MSE}}$	0.054	0.283	0.144	0.602	0.167	0.651
	Model 2c	RelBias	0.031	0.021	0.033	0.893	-0.009	0.072
		$\sqrt{\text{MSE}}$	0.066	0.378	0.168	5.728	0.209	0.703

cause $\exp(10\hat{\beta}) \cong 1.42$, $\hat{\beta}$ being the posterior mean of β , the estimated increase in the average proportion of clay relative to the average proportion of other components (i.e., sand and silt) when the depth is increased in 10 meters is 42%; (25%, 70%) is the corresponding 95% highest posterior density (HPD) credible interval.

Some technical details on the model fit are now in order. We ran 140,000 Monte Carlo iterations and discarded the first 50,000 iterations. The multivariate version of Gelman and Rubin’s convergence diagnostic (Brooks and Gelman, 1998) indicates that the chain is convergent (m.p.r.f. = 1.02 < 1.2). Additionally, diagnostic plots (not shown) suggest that the chain for each parameter is not correlated and stationary. We can then assume that the estimates reported in Table 16 are reliable.

We now present an application of our results to data taken from a sample of 408 young males between 8 and 18 years (Machado, Oikawa and Barbanti, 2013). The observed variables under consideration are: fat mass percentage (y), height (cm), weight (kg), horizontal abdominal skinfold (SkHab, mm) and year for peak height velocity (phv, years). Usually, measures of skinfolds are subject to measurement

Table 11 True mean and estimated posterior means and standard deviation, with $\phi = 50$; Model 2a

k	Parameter	Posterior inference		
		True	Mean	s.d.
0.95	β	-0.4	-0.399	0.026
	α_1	2.0	2.014	0.135
	α_2	1.0	1.011	0.101
	$\overline{\log(\phi)}$	3.912	4.134	0.174
	μ_x	5.0	4.953	0.130
	σ_x^2	2.0	2.177	0.335
0.75	β	-0.4	-0.385	0.033
	α_1	2.0	1.787	0.176
	α_2	1.0	1.220	0.126
	$\overline{\log(\phi)}$	3.912	4.285	0.328
	μ_x	5.0	4.970	0.168
	σ_x^2	2.0	2.184	0.412
0.5	β	-0.4	-0.383	0.047
	α_1	2.0	2.059	0.281
	α_2	1.0	0.897	0.148
	$\overline{\log(\phi)}$	3.912	3.975	0.573
	μ_x	5.0	4.864	0.202
	σ_x^2	2.0	2.137	0.600

error. We therefore assume that the observed value of SkHab is a surrogate for the true (latent) horizontal abdominal skinfold (x). We assume that y_1, y_2, \dots, y_n are independent observations of the percentage fat mass of the $n = 408$ individuals in the sample, and such that y_i follows a beta distribution with mean μ_i and precision parameter ϕ_i , with

$$\log\left(\frac{\mu_i}{1 - \mu_i}\right) = \alpha_1 + \alpha_2\text{phv}_i + \alpha_3\text{height}_i + \alpha_4\text{weight}_i + \beta x_i,$$

$$\log(\phi_i) = \gamma_1 + \gamma_2\text{phv}_i + \gamma_3\text{height}_i + \gamma_4\text{weight}_i + \lambda x_i,$$

$$\text{SkHab}_i = x_i + e_i,$$

$i = 1, \dots, n$, $x_i \sim N(\mu_x, \sigma_x^2)$ and $e_i \sim N(0, \sigma_e^2)$. The specifications for the prior distributions of the parameters are as follows: $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \lambda$ and μ_x have a $t(5, 0, 20)$ distribution; σ_x^2 , and σ_e^2 follow the same distribution as $(50B)^2$, where $B \sim \text{beta}(1 + \varepsilon, 1 + \varepsilon)$, with $\varepsilon = 0.1$; finally, for the constant precision case, $\phi \sim \text{LN}(4, 1)$.

Table 17 gives the DIC, EAIC and EBIC for the model fitting with different prior specifications for the precision parameter. Model 1 ignores the measurement

Table 12 True mean and estimated posterior means and standard deviation, with $\phi = 300$; Model 2a

k	Parameter	Posterior inference		
		True	Mean	s.d.
0.95	β	-0.4	-0.404	0.013
	α_1	2.0	1.981	0.074
	α_2	1.0	1.025	0.060
	$\overline{\log(\phi)}$	5.7	5.585	0.286
	μ_x	5.0	4.932	0.151
	σ_x^2	2.0	2.059	0.315
0.75	β	-0.4	-0.382	0.024
	α_1	2.0	1.937	0.142
	α_2	1.0	0.989	0.113
	$\overline{\log(\phi)}$	5.7	5.367	0.673
	μ_x	5.0	5.052	0.174
	σ_x^2	2.0	2.357	0.421
0.5	β	-0.4	-0.373	0.046
	α_1	2.0	1.879	0.254
	α_2	1.0	1.016	0.145
	$\overline{\log(\phi)}$	5.7	5.164	1.030
	μ_x	5.0	5.060	0.206
	σ_x^2	2.0	2.434	0.603

errors, and Models 2–5 take the measurement errors into account. It can be noticed that the smallest DIC, EAIC and EBIC are achieved for Model 2, the constant precision model with measurement error.

Table 18 gives the posterior estimates of the parameters associated with Model 2, which provides the best fit for the data. It can be noticed that the 95% HPD credible interval for α_3 suggests that this parameter could be removed from the model and the estimated reliability ratio $k_1 = 0.822$ indicates that the measurement error in the horizontal abdominal skinfold is non-negligible. All the convergence diagnostic tools (results not shown) suggest that the estimates reported in Table 18 are reliable.

We now evaluate the estimated impact of the horizontal abdominal skinfold in the mean fat mass percentage. Here, $\exp(c\beta)$ has an interpretation as the odds ratio when SkHab is increased in c units. Therefore, because $\exp(10(\hat{\beta})) \cong 1.70$, $\hat{\beta}$ being the posterior mean of β , the estimated increase in the mean fat mass percentage relative to the mean non-fat mass percentage when the horizontal abdominal skinfold is increased in 10 mm is 70%; (63%, 79%) is the corresponding 95% HPD credible interval.

Table 13 Summary results based on 100 simulated datasets for the fitted models with $\phi = 50$ and prior specifications given in Model 1c; σ_e^2 unknown

k		Posterior inference						
		β	α_1	α_2	$\log(\phi)$	μ_x	σ_x^2	σ_e^2
0.95	True	-0.4	2.0	1.0	3.912	5.0	2.0	0.1
	mean	-0.430	2.147	1.018	4.397	4.996	1.917	0.223
	RelBias	0.076	0.073	0.018	0.124	-0.001	-0.042	-0.888
	$\sqrt{\text{MSE}}$	0.040	0.218	0.121	0.528	0.147	0.322	1.778
0.75	True	-0.4	2.0	1.0	3.912	5.0	2.0	0.66
	mean	-0.405	2.026	0.995	4.184	4.972	2.106	0.610
	RelBias	0.013	0.013	-0.005	0.070	-0.006	0.053	-0.076
	$\sqrt{\text{MSE}}$	0.041	0.223	0.153	0.374	0.149	0.380	0.192
0.5	True	-0.4	2.0	1.0	3.912	5.0	2.0	2.0
	mean	-0.381	1.903	0.986	4.024	4.957	2.406	1.648
	RelBias	-0.047	-0.049	-0.014	0.029	-0.009	0.203	1.498
	$\sqrt{\text{MSE}}$	0.066	0.333	0.189	0.339	0.179	0.737	1.091

Table 14 Summary results based on 100 simulated datasets for the fitted models with $\phi = 300$ and prior specifications given in Model 1c; σ_e^2 unknown

k		Posterior inference						
		β	α_1	α_2	$\log(\phi)$	μ_x	σ_x^2	σ_e^2
0.95	True	-0.4	2.0	1.0	5.7	5.0	2.0	0.1
	mean	-0.391	1.957	0.996	5.421	4.979	2.108	0.053
	RelBias	-0.022	-0.022	-0.004	-0.050	-0.004	0.054	-0.468
	$\sqrt{\text{MSE}}$	0.017	0.089	0.063	0.339	0.139	0.297	0.05
0.75	True	-0.4	2.0	1.0	5.7	5.0	2.0	0.66
	mean	-0.351	1.753	0.988	4.601	4.974	2.362	0.350
	RelBias	-0.123	-0.124	-0.012	-0.193	-0.005	0.181	-0.470
	$\sqrt{\text{MSE}}$	0.056	0.285	0.114	1.122	0.149	0.509	0.327
0.5	True	-0.4	2.0	1.0	5.7	5.0	2.0	2.0
	mean	-0.316	1.575	0.982	4.231	4.956	2.758	1.295
	RelBias	-0.210	-0.213	-0.018	-0.258	-0.009	0.379	0.962
	$\sqrt{\text{MSE}}$	0.096	0.489	0.154	1.494	0.178	0.975	0.721

6 Concluding remarks

Beta regression models have become a popular tool for modeling limited range continuous data. The constant precision beta regression model proposed by Ferrari and Cribari-Neto (2004) and its non-constant precision version studied by Smithson and Verkuilen (2006) have been extended in various directions. Recently,

Table 15 *DIC, EAIC, EBIC for the fitted models with different specifications of the precision parameter; Arctic lake data*

Model	Precision submodel	DIC	EAIC	EBIC
Model 1 (naïve)	$\log(\phi) = \gamma$	105.7	116.5	126.5
Model 2	$\log(\phi) = \gamma$	97.5	104.3	114.3
Model 3	$\log(\phi_i) = \gamma + \lambda x_i$	148.8	144.1	155.7

Table 16 *Estimated posterior medians and means for Model 2; Arctic lake data*

Parameter	Posterior inference			
	Mean	s.d.	Median	95% HPD credible interval
α	-2.659	0.404	-2.600	(-3.568, -1.955)
β	0.035	0.008	0.033	(0.022, 0.053)
$\log(\phi)$	2.776	0.435	2.705	(2.048, 3.624)
μ_x	45.135	4.934	45.290	(35.050, 54.490)
σ_x^2	723.467	224.489	703.600	(309.600, 1174.000)
σ_e^2	142.280	115.197	124.950	(0.113, 355.600)

Table 17 *DIC, EAIC, EBIC for the fitted models with different specifications of the precision parameter; fat mass data*

Model	Precision submodel	DIC	EAIC	EBIC
Model 1 (naïve)	$\log(\phi) = \gamma_1$	191.7	230.2	282.35
Model 2	$\log(\phi) = \gamma_1$	106.7	148.1	200.25
Model 3	$\log(\phi_i) = \gamma_1 + \gamma_2 \text{phv}_i + \gamma_3 \text{height}_i + \gamma_4 \text{weight}_i + \lambda x_i$	301.8	350.2	418.39
Model 4	$\log(\phi_i) = \gamma_1 + \lambda x_i$	299.6	348.7	404.86
Model 5	$\log(\phi_i) = \gamma_1 + \gamma_2 \text{phv}_i + \gamma_3 \text{height}_i + \gamma_4 \text{weight}_i$	334.1	380.5	444.68

Table 18 *Estimated posterior medians and means for Model 2; fat mass data*

Parameter	Posterior inference			
	Mean	s.d.	Median	95% HPD credible interval
α_1	-4.7393	0.208	-4.770	(-5.083, -4.304)
α_2	-0.060	0.013	-0.061	(-0.085, -0.036)
α_3	0.001	0.001	-0.001	(-0.002, 0.003)
α_4	0.022	0.002	0.022	(0.019, 0.026)
β	0.053	0.002	0.053	(0.049, 0.058)
$\log(\phi)$	6.702	0.384	6.690	(5.974, 7.436)
μ_x	17.513	0.566	17.520	(16.420, 18.630)
σ_x^2	104.807	9.059	104.400	(88.160, 122.900)
σ_e^2	22.630	3.089	22.710	(16.670, 28.840)

Carrasco, Ferrari and Arellano-Valle (2014) extended both models for the case in which some of the covariates are recorded with measurement error, and proposed different likelihood-based methods to estimate the model parameters. The proposed methods require the optimization of the likelihood or pseudo-likelihood functions that involve analytically intractable integrals. In the present paper, we present a Bayesian approach for errors-in-variables beta regression models that does not require an intensive computational evaluation of these integrals. A careful study was conducted to find prior distributions that have little influence in the posterior inferences. Our approach can easily be implemented when the measurement error variance is either known or unknown by using simple and accessible software, such as WinBUGS.

Appendix: BUGS codes for the errors-in-variables beta regression

This appendix presents various pieces of BUGS code used for fitting the errors-in-variables beta regression in the simulated data example.

Log-normal prior for ϕ , known measurement error variance

```
model
{
for( i in 1 : n ) {
Y[i] ~ dbeta(a1[i] , a2[i])
a1[i] <- mu[i]*phi
a2[i] <- (1-mu[i])*phi
logit(mu[i]) <- z[i,1]*alpha1+z[i,2]*alpha2+x[i]*beta
w[i] ~ dnorm(x[i] , sigmaeinv0)
x[i] ~ dnorm(mux , sigmaxinv)
}
mux ~ dt(mu0,sigma0inv,nu0)
sigmaxx ~ dbeta(a0,a0)
sigma2e0<-1/sigmaeinv0
sigma2x <- (sigmaxx*b0)*(sigmaxx*b0)
sigmaxinv<-1/sigma2x
sigma0inv<-1/sigma20
alpha1 ~ dt(mualpha1,sigmaalpha1inv,nualpha1)
alpha2 ~ dt(mualpha2,sigmaalpha2inv,nualpha2)
sigmaalpha1inv<-1/sigmaalpha1
sigmaalpha2inv<-1/sigmaalpha2
beta ~ dt(mubeta,sigmabetainv,nubeta)
sigmabetainv<-1/sigma2beta
logphi ~ dnorm(muphi,sigmaphiinv)
sigmaphiinv<-1/sigmaphi
phi<-exp(logphi)
}
```

Submodel for ϕ , known measurement error variance

```

model
{
for( i in 1 : n ) {
Y[i] ~ dbeta(a1[i] ,a2[i])
a1[i] <- mu[i]*phi[i]
a2[i] <- (1-mu[i])*phi[i]
logit(mu[i]) <- z[i,1]*alpha1+z[i,2]*alpha2+x[i]*beta
log(phi[i])<- z[i,1]*gamma1+z[i,2]*gamma2+x[i]*lambda
w[i] ~ dnorm(x[i] , sigmaeinv0)
x[i] ~ dnorm(mux , sigmaxinv)
}
mux ~ dt(mu0,sigma0inv,nu0)
sigmaxx ~ dbeta(a0,a0)
sigma2e0<-1/sigmaeinv0
sigma2x <- (sigmaxx*b0)*(sigmaxx*b0)
sigmaxinv<-1/sigma2x
sigma0inv<-1/sigma20
alpha1 ~ dt(mualpha1,sigmaalpha1inv,nualpha1)
alpha2 ~ dt(mualpha2,sigmaalpha2inv,nualpha2)
gamma1 ~ dt(mugamma1,simagamma1inv,nugamma1)
gamma2 ~ dt(mugamma2,simagamma2inv,nugamma2)
sigmaalpha1inv<-1/sigmaalpha1
sigmaalpha2inv<-1/sigmaalpha2
sigmagamma1inv<-1/simagamma1
sigmagamma2inv<-1/simagamma2
beta ~ dt(mubeta,sigmabetainv,nubeta)
lambda ~ dt(mulambda,sigmalambdainv,nulambda)
sigmabetainv<-1/sigma2beta
sigmalambdainv<-1/sigma2lambda
meanphi<-mean(phi[])
}

```

Log-normal prior for ϕ , unknown measurement error variance

```

model
{
for( i in 1 : n ) {
Y[i] ~ dbeta(a1[i] ,a2[i])
a1[i] <- mu[i]*phi
a2[i] <- (1-mu[i])*phi
logit(mu[i]) <- z[i,1]*alpha1+z[i,2]*alpha2+x[i]*beta
w[i] ~ dnorm(x[i] , sigmaeinv)
x[i] ~ dnorm(mux , sigmaxinv)
}

```

```

mux ~ dt(mu0, sigma0inv, nu0)
sigmaee ~ dbeta(a0, a0)
sigma2e <- (sigmaee*b0)*(sigmaee*b0)
sigmaeinv<-1/sigma2e
sigmaxx ~ dbeta(a0, a0)
sigma2x <- (sigmaxx*b0)*(sigmaxx*b0)
sigmaxinv<-1/sigma2x
sigma0inv<-1/sigma20
alpha1 ~ dt(mualpha1, sigmaalpha1inv, nualpha1)
alpha2 ~ dt(mualpha2, sigmaalpha2inv, nualpha2)
sigmaalpha1inv<-1/sigmaalpha1
sigmaalpha2inv<-1/sigmaalpha2
beta ~ dt(mubeta, sigmabetainv, nubeta)
sigmabetainv<-1/sigma2beta
logphi ~ dnorm(muphi, sigmaphiinv)
sigmaphiinv<-1/sigmaphi
phi<-exp(logphi)
}

```

Acknowledgments

This research was partially supported by grant FONDECYT 1120121/1150325-Chile, by Facultad de Matemáticas y Vicerrectoría de Investigación (VRI) of the Pontificia Universidad Católica de Chile, CNPq-Brazil and FAPESP-Brazil. J. I. Figueroa-Zúñiga acknowledges funding support by grant FONDECYT 11130483, from the Chilean government.

References

- Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde. *Journal of the Royal Statistical Society, Series B* **64**, 616–618.
- Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**, 434–455. [MR1665662](#)
- Buonaccorsi, J. P. (2010). *Measurement Error: Models, Methods and Applications*. London: Chapman & Hall. [MR2682774](#)
- Carlin, B. P. and Louis, T. A. (2001). *Bayes and Empirical Bayes Methods for Data Analysis*. Boca Raton: Chapman & Hall. [MR1427749](#)
- Carrasco, J. M. F., Ferrari, S. L. P. and Arellano-Valle, R. B. (2014). Errors-in-variables beta regression models. *Journal of Applied Statistics* **41**, 1530–1547. [MR3292031](#)
- Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. New York: Chapman & Hall. [MR2243417](#)
- Cheng, C. and JW, V. (1999). *Statistical Regression with Measurement Error*. London: Oxford University Press. [MR1719513](#)

- Coakley, J. P. and Rust, B. R. (1968). Sedimentation in an Arctic lake. *Journal of Sedimentary Petrology* **4**, 1290–1300.
- Dellaportas, P. and Stephens, D. A. (1995). Bayesian analysis of errors-in-variables regression models. *Biometrics* **51**, 1085–1095.
- Ferrari, S. L. P. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics* **31**, 799–815. [MR2095753](#)
- Figueroa-Zúñiga, J. I., Arellano-Valle, R. B. and Ferrari, S. L. (2013). Mixed beta regression: A Bayesian perspective. *Computational Statistics & Data Analysis* **61**, 137–147. [MR3063006](#)
- Fong, Y., Rue, H. and Wakefield, J. (2010). Bayesian inference for generalized linear mixed models. *Biostatistics* **11**, 397–412.
- Fuller, W. A. (1987). *Measurement Error Models*. New York: Wiley. [MR0898653](#)
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* **1**, 515–533. [MR2221284](#)
- Machado, D., Oikawa, S. and Barbanti, V. (2013). The multicomponent anthropometric model for assessing body composition in a male pediatric population: A simultaneous prediction of fat mass, bone mineral content, and lean soft tissue. *Journal of Obesity*. **2013**, Article ID 428135.
- Simas, A. B., Barreto-Souza, W. and Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis* **54**, 348–366. [MR2756431](#)
- Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods* **11**, 54–71.
- Spiegelhalter, D., Best, N., Carlin, B. and Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* **64**, 583–639. [MR1979380](#)
- Zimprich, D. (2010). Modeling change in skewed variables using mixed beta regression models. *Research in Human Development* **7**, 9–26.

J. Figueroa-Zúñiga
 Department of Statistics
 University of Concepción
 Concepción
 Chile
 E-mail: jifiguer@gmail.com

R. Arellano-Valle
 Department of Statistics
 Pontifical Catholic University of Chile
 Santiago
 Chile
 E-mail: reivalle@mat.puc.cl

J. M. F. Carrasco
 Department of Statistics
 Federal University of Bahia
 Salvador
 Brazil
 E-mail: carrascojalmar@gmail.com

S. L. P. Ferrari
 Department of Statistics
 University of São Paulo
 São Paulo
 Brazil
 E-mail: silviaferrari.usp@gmail.com