RESEARCH ARTICLE

Beta regression model nonlinear in the parameters with additive measurement errors in variables

Daniele de Brito Trindade¹, Patrícia Leone Espinheira^{2*}, Klaus Leite Pinto Vasconcellos², Jalmir Manuel Farfán Carrasco³, Maria do Carmo Soares de Lima²¹ Instituto Federal de Educação Ciência e Tecnologia Baiano Guanambi, Guanambi, BA, Brazil,² Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil,³ Departamento de Estatística, Universidade Federal da Bahia, Salvador, Bahia, Brazil* patespipa@de.ufpe.br

Abstract

We propose in this paper a general class of nonlinear beta regression models with measurement errors. The motivation for proposing this model arose from a real problem we shall discuss here. The application concerns a usual oil refinery process where the main covariate is the concentration of a typically measured in error reagent and the response is a catalyst's percentage of crystallinity involved in the process. Such data have been modeled by nonlinear beta and simplex regression models. Here we propose a nonlinear beta model with the possibility of the chemical reagent concentration being measured with error. The model parameters are estimated by different methods. We perform Monte Carlo simulations aiming to evaluate the performance of point and interval estimators of the model parameters. Both results of simulations and the application favors the method of estimation by maximum pseudo-likelihood approximation.

OPEN ACCESS

Citation: de Brito Trindade D, Espinheira PL, Pinto Vasconcellos KL, Farfán Carrasco JM, do Carmo Soares de Lima M (2021) Beta regression model nonlinear in the parameters with additive measurement errors in variables. PLoS ONE 16(7): e0254103. <https://doi.org/10.1371/journal.pone.0254103>

Editor: Fabio Rapallo, Università degli Studi di Genova, ITALY

Received: January 12, 2021

Accepted: June 18, 2021

Published: July 29, 2021

Copyright: © 2021 de Brito Trindade et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Funding: This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (<https://www.capes.gov.br/>), received by DBT through a PhD Scholarship. No additional external funding was received for this study.

1 Introduction

Regression models for dependent variables that assume values in the unit interval have been proving quite important in the literature, with a special highlight to the beta regression model proposed by [1] and generalized by [2] whom proposed the nonlinear beta regression model. In the above mentioned models traditionally all covariates are considered fixed, not random, without measurement error. In practice, covariates may not be observed directly or may be subject to measurement errors. It is important to emphasize that if this assumption is not respected, unreliable inferential results shall be obtained [3]. In these circumstances, regression models with measurement errors are usually defined and structured so that the mean response is explained by covariates x_i whose measurements are inaccurate. Thus, instead of the true value of x_i , the value of another predictor variable, w_i , which is associated with a measurement error, is considered.

It is reasonable that in regression models for dependent variables assuming values in the unit interval, some of the covariates are not observed directly, but acquired with possible

Competing interests: The authors have declared that no competing interests exist.

measurement errors. Thus, [4] proposed the beta regression model with measurement errors, in which covariates measured with errors can enter in the mean and precision submodels described by linear predictors. A possible extension to the model proposed by [4] is to consider a nonlinear modeling for the parameters. In fact, a real problem has shown the necessity of this modeling proposal. The interest of the problem lies in modeling the percentage of crystallinity of a chemical reagent based on different concentrations of vanadium and steam, considering two values for temperature. It is expected that the higher the concentrations of vanadium and steam, the lower the percentage of crystallinity. The loss of crystallinity is a negative effect of the process and precise knowledge of how vanadium concentration destroys this crystallinity is one of the most chief goals of this regression modeling. As we shall describe in the application, the measurement of vanadium concentration may be inaccurate, which characterizes the possibility of this covariate to present measurement error. These data were modeled by [5] based on beta and simplex nonlinear regressions in which the vanadium was taken as a fixed covariate. In our application the nonlinearity is related with steam, that is here treated as a fixed covariate.

Nonlinear models with measurement errors have been being recently studied, as shown in the literature; see [6–9]. Our aim here is propose a particular class of nonlinear beta regression with measured errors, in which the nonlinearity is on a parameter related to a fixed covariate. We consider three methods for the estimation of the parameters, namely: approximate maximum likelihood, approximate pseudo maximum likelihood and regression calibration. We compare these three methods with the estimation of the naive model, which considers that the regression does not have covariates measured in error, that is, it considers the classical nonlinear beta regression model. We evaluate the properties of point and interval estimators for the four estimation methods considered, based on Monte Carlo simulations and different scenarios. Both the simulations and our application pointed out that the estimation method by approximate pseudo maximum likelihood is the one that presents the best performance. Furthermore, recently [10] proposed the simplex regression models with measurement error, and as a future research we shall extend this proposal to nonlinear case.

2 Model and estimation methods

Consider n independent random variables $y_1, \dots, y_n$ such that each y_t , $t = 1, \dots, n$ is beta distributed with density

$$f(y_t; \mu_t, \phi_t) = \Gamma(\phi_t) \Gamma(\mu_t \phi_t) \Gamma(1 - \mu_t \phi_t)^{-1} y_t^{\mu_t \phi_t - 1} (1 - y_t)^{(1 - \mu_t) \phi_t - 1}, \quad 0 < y_t < 1, \quad (1)$$

$0 < \mu_t < 1$, $\phi_t > 0$, $E(y_t) = \mu_t$ and $\text{Var}(y_t) = \mu_t(1 - \mu_t)/1 + \phi_t$. The class of nonlinear beta regression models here with measurement errors here proposed considers

$$g_1(\mu_t) = \eta_{1t} = f_1(\mathbf{z}_t^\top, \mathbf{x}_t^\top; \alpha, \beta), \quad g_2(\phi_t) = \eta_{2t} = f_2(\mathbf{v}_t^\top, \mathbf{m}_t^\top; \gamma, \lambda), \quad (2)$$

where $\alpha = (\alpha_1, \dots, \alpha_p)^\top \in \mathbb{R}^p$, $\beta = (\beta_1, \dots, \beta_r)^\top \in \mathbb{R}^r$, $\gamma = (\gamma_1, \dots, \gamma_{\tilde{q}}) \in \mathbb{R}^{\tilde{q}}$, $\lambda = (\lambda_1, \dots, \lambda_s)^\top \in \mathbb{R}^s$ are unknown parameters, $(p + r) + (\tilde{q} + s) < n$, $\mathbf{z}_t^\top = (z_{t1}, \dots, z_{tp})$ and $\mathbf{v}_t^\top = (v_{t1}, \dots, v_{t\tilde{q}})$ are fixed and unknown vectors of observations and $\mathbf{x}_t^\top = (x_{t1}, \dots, x_{tr})$, $\mathbf{m}_t^\top = (m_{t1}, \dots, m_{ts})$ are vectors of covariates not directly observed or associated to measurement errors. The link functions $g_1(\cdot): (0, 1) \rightarrow \mathbb{R}$ and $g_2(\cdot): (0, \infty) \rightarrow \mathbb{R}$ are strictly monotonic, continuous and twice differentiable. Further, $f_1(\cdot)$ and $f_2(\cdot)$ are differentiable functions with Jacobian matrices $\mathbf{F}_1 = \partial \eta_1 / \partial \alpha$, $\mathbf{F}_2 = \partial \eta_1 / \partial \beta$, $\mathbf{F}_3 = \partial \eta_2 / \partial \gamma$ and $\mathbf{F}_4 = \partial \eta_2 / \partial \lambda$ having ranks p , r , $\tilde{q}$ and s , respectively. We should point out that the method proposed here is not applicable to models with random effects. Finally, the parameter defining the degree of variability of precision is

here defined as $\delta = \max_{\{t=1,\dots,n\}} \phi_t / \min_{\{t=1,\dots,n\}} \phi_t$. For the model with constant precision, $\phi_1 = \dots = \phi_n = \phi$, and, therefore, $\delta = 1$.

Following [4] we consider that the variables with measurement errors in the submodels for mean and precision are the same. Also, the vector of random measurements $\mathbf{x}_t$ is not directly observed. The observed vector is $\mathbf{w}_t = (w_{t1}, \dots, w_{tr})^\top$, which we here suppose related to $\mathbf{x}_t$ in the form $\mathbf{w}_t = \tau_1 + \tau_2 \circ \mathbf{x}_t + \mathbf{e}_t$, where $\mathbf{e}_t$ is a vector of random errors, $\tau_1 = (\tau_{11}, \dots, \tau_{1r})^\top \in \mathbb{R}^r$ and $\tau_2 = (\tau_{21}, \dots, \tau_{2r})^\top \in \mathbb{R}^r$ are vectors of unknown parameters and “ $\circ$ ” represents the Hadamard product; for two vectors $\mathbf{u}_1 = (u_{11}, \dots, u_{1r})^\top$ and $\mathbf{u}_2 = (u_{21}, \dots, u_{2r})^\top$ of same dimension, $\mathbf{u}_1 \circ \mathbf{u}_2 = (u_{11} u_{21}, \dots, u_{1r} u_{2r})^\top$. We may assume τ_1 to be a vector of zeros, as we may assume τ_2 to be a vector of ones. However, we can also assume them to be unknown parameters to be estimated, as we shall see in the application.

Hence, our vector of unknown parameters is $\Psi = (\theta^\top, \xi^\top, \sigma_e^{2\top})^\top$, where $\theta = (\alpha^\top, \beta^\top, \gamma^\top, \lambda^\top)^\top$ and $(\xi^\top, \sigma_e^{2\top})^\top = (\mu^\top, \sigma_x^{2\top}, \sigma_e^{2\top})^\top$ are, respectively, the vectors of interest and nuisance parameters. The joint density function of $(y_t, \mathbf{w}_t)$ is

$$f(y_t, \mathbf{w}_t; \Psi) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_t | \mathbf{x}_t; \theta) f(\mathbf{w}_t | \mathbf{x}_t; \sigma_e^2) f(\mathbf{x}_t; \xi) d\mathbf{x}_t. \quad (3)$$

In (3), $f(y_t | \mathbf{x}_t; \theta)$ represents a density for a beta distribution, $f(\mathbf{w}_t | \mathbf{x}_t; \sigma_e^2)$ is a conditional density of $\mathbf{w}_t$ given $\mathbf{x}_t$ and $f(\mathbf{x}_t; \xi)$ is a density for the explanatory variable $\mathbf{x}_t$. Observe that, to obtain (3), we must assume that y_t is independent of $\mathbf{w}_t$ given $\mathbf{x}_t$, that is, the distribution of y_t given $(\mathbf{w}_t, \mathbf{x}_t)$ involves only $\mathbf{x}_t$ [8]. Therefore, the log-likelihood function for the n independent observations is

$$\begin{aligned} \ell(\Psi) &= \sum_{t=1}^n \log \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_t | \mathbf{x}_t; \theta) f(\mathbf{x}_t | \mathbf{w}_t; \xi, \sigma_e^2) f(\mathbf{w}_t; \xi, \sigma_e^2) d\mathbf{x}_t \\ &= \sum_{t=1}^n \log f(\mathbf{w}_t; \xi, \sigma_e^2) + \sum_{t=1}^n \log \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_t | \mathbf{x}_t; \theta) f(\mathbf{x}_t | \mathbf{w}_t; \xi, \sigma_e^2) d\mathbf{x}_t, \end{aligned} \quad (4)$$

The log-likelihood above involves integrals that are analytically unsolvable, which means that approximation methods need to be used. We will consider here three different approaches to estimate the parameters and, aiming to describe each one of the three methods, we will suppose, for illustration purposes, that only x_t is measured with error. However, it is important to stress that the methods here used can be easily generalized to r covariates with measurement errors.

From now on, we make the following assumptions:

$$y_t | x_t \sim \text{Beta}(\mu_t, \phi_t), \quad w_t = x_t + e_t, \quad x_t \stackrel{\text{ind}}{\sim} \mathcal{N}(\mu_x, \sigma_x^2), \quad e_t \stackrel{\text{ind}}{\sim} \mathcal{N}(0, \sigma_e^2),$$

and, finally, x_t and $e_{t'}$, with $t, t' = 1, \dots, n$, are independent. This is a *structural model* for which the vector of non-observed variables follows a normal distribution. It is important here to stress that the models for μ_t and ϕ_t , $t = 1, \dots, n$ are based in (2). From the above assumptions, we have

$$w_t \stackrel{\text{ind}}{\sim} \mathcal{N}(\mu_x, \sigma_x^2 + \sigma_e^2), \quad x_t | w_t \stackrel{\text{ind}}{\sim} \mathcal{N}(\mu_{x_t|w_t}, \sigma_{x_t|w_t}^2), \quad (5)$$

$$\mu_{x_t|w_t} = \mu_x + k_x(w_t - \mu_x), \quad \sigma_{x_t|w_t}^2 = \sigma_e^2 k_x, \quad k_x = \sigma_x^2 / (\sigma_x^2 + \sigma_e^2), \quad (6)$$

where k_x is known as the **reliability coefficient**. An additional assumption is that the variance σ_e^2 of the measurement error is known. This assumption is necessary to avoid identifiability

problems. When we have replicas for w_t , it is possible to use the sample variance of w_t to estimate the variance of the measurement error.

In order to establish the log-likelihood for the model to be studied, we use the results in (5) and also that $\sigma_e^2 = \hat{\sigma}_e^2$, from which we obtain

$$\begin{aligned}\ell(\Psi) &= \sum_{t=1}^n \ell_{1t}(\xi, \hat{\sigma}_e^2) + \sum_{t=1}^n \ell_{2t}(\theta, \xi, \hat{\sigma}_e^2), \\ \ell_{1t}(\xi, \hat{\sigma}_e^2) &= \log f(w_t; \xi, \hat{\sigma}_e^2) = -\frac{1}{2} \log [2\pi(\sigma_x^2 + \hat{\sigma}_e^2)] - \frac{1}{2(\sigma_x^2 + \hat{\sigma}_e^2)} (w_t - \mu_x)^2,\end{aligned}\quad (7)$$

$$\ell_{2t}(\theta, \xi, \hat{\sigma}_e^2) = \log \int_{-\infty}^{\infty} f(y_t | x_t; \theta, \xi, \hat{\sigma}_e^2) \frac{1}{\sqrt{2\pi\sigma_{x_t|w_t}^2}} \exp \left[-\frac{(x_t - \mu_{x_t|w_t})^2}{2\sigma_{x_t|w_t}^2} \right] dx_t. \quad (8)$$

We also suppose that the log-likelihood is a concave function of the parameters to be estimated. The estimators we shall propose have no closed form. Therefore, we shall use the nonlinear optimization method Fisher's scoring with initial values proposed by [11] to the nonlinear beta regression models.

2.1 Estimation by approximate maximum likelihood

The integral in (8) can be approximated using the Gauss-Hermite quadrature, given by

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{q=1}^Q v_q f(s_q), \quad (9)$$

where s_q and v_q are, respectively, the q -th zero and weight of the Q -th order orthogonal Hermite polynomial, Q being the number of points for quadrature [12]. Using the change of variable $u_t = (x_t - \mu_{x_t|w_t}) / \sqrt{2\sigma_{x_t|w_t}^2}$ in (8), that is, $x_t = \mu_{x_t|w_t} + \sqrt{2\sigma_{x_t|w_t}^2} u_t$, we obtain an approximation for the integral in (8) based in (9), which yields the approximation for the log-likelihood function given by

$$\ell_a(\Psi) = \sum_{t=1}^n \ell_{1t}(\xi, \sigma_e^2) + \sum_{t=1}^n \log \left(\sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp [\ell_{iq}(\mu_{iq}, \phi_{iq})] \right), \quad (10)$$

where

$$\begin{aligned}\ell_{iq}(\mu_{iq}, \phi_{iq}) &= \log \Gamma(\phi_{iq}) - \log \Gamma(\mu_{iq} \phi_{iq}) - \log \Gamma[(1 - \mu_{iq}) \phi_{iq}] \\ &+ (\mu_{iq} \phi_{iq} - 1) \log y_t + [(1 - \mu_{iq}) \phi_{iq} - 1] \log (1 - y_t),\end{aligned}\quad (11)$$

here, μ_{iq} and ϕ_{iq} are the parameters of the beta distribution of y_t when the explanatory value x_t takes the value $x_{iq}^* = \mu_{x_t|w_t} + \sqrt{2\sigma_{x_t|w_t}^2} s_q$, $\mu_{x_t|w_t}$ and $\sigma_{x_t|w_t}^2$ being defined in (6) and s_q being the q -th zero of the Q -th order orthogonal Hermite polynomial. Consequently, from (2), we have

$$g_1(\mu_{iq}) = f_1(\mathbf{z}_t^\top, x_{iq}^*; \alpha, \beta), \quad g_2(\phi_{iq}) = f_2(\mathbf{v}_t^\top, x_{iq}^*; \gamma, \lambda).$$

Typically in the beta linear models with errors in covariates the MLE's asymptotic distribution is normal with mean Ψ and covariance matrix $J_a^{-1}(\Psi)$ for n and Q large enough [4]. Here, we also shall consider this result to build asymptotic confidence intervals for the interest parameters of the model defined in (2) with a level α . We define $J_a(\theta)$ as the partition of $J_a(\Psi)$

related with the interest parameters in the Appendix. Previous to $J_a(\theta)$ matrix, we also present in the Appendix the score functions for the parameters of the model. Here, MLEa is the approximate maximum likelihood estimator.

2.2 Estimation by approximate pseudo maximum likelihood (pseudo-likelihood estimation)

This estimation method considers maximizing a function depending only of the parameters of interest, with, following [13], the nuisance parameters being replaced by consistent estimators in the original likelihood function defined in (3). Pseudo-likelihood estimation consists of first obtaining the optimal point $\hat{\xi}$ of the log-likelihood corresponding only to the nuisance parameters, which is given by $\ell_r(\xi) = \sum_{t=1}^n \ell_{1t}(\xi)$, with $\ell_{1t}(\xi)$ as in (7). Once $\hat{\xi}$ is obtained, the pseudo log-likelihood is defined as $\ell_\theta(\theta, \hat{\xi}) = \sum_{t=1}^n \ell_{1t}(\hat{\xi}) + \sum_{t=1}^n \ell_{2t}(\theta, \hat{\xi})$. It is important to emphasize that the ℓ_{2t} integral defined in (8) will also be approximated by Gauss-Hermite quadrature. However, since $\ell_{2t}(\theta, \hat{\xi})$ involves only θ , its approximation will be a function only of the interest parameters. Let $\ell_{2t}^\dagger(\theta, \hat{\xi})$ be the approximation for $\ell_{2t}(\theta, \hat{\xi})$ obtained with (9). From this, the approximate pseudo log-likelihood function for the nonlinear beta regression model with unidimensional error measures is defined as $\ell_p(\theta, \hat{\xi}) = \sum_{t=1}^n \ell_{1t}(\hat{\xi}) + \sum_{t=1}^n \ell_{2t}^\dagger(\theta, \hat{\xi})$ in the form

$$\begin{aligned} \ell_p(\theta, \hat{\xi}) = & -\frac{1}{2} \log [2\pi(\hat{\sigma}_x^2 + \hat{\sigma}_e^2)] - \frac{1}{2(\hat{\sigma}_x^2 + \hat{\sigma}_e^2)} (w_t - \mu_x)^2 \\ & + \sum_{t=1}^n \log \left(\sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp [\ell_{iq}(\mu_{iq}^r; \hat{\phi}_{iq}^r)] \right), \end{aligned} \quad (12)$$

where $\ell_{iq}(\mu_{iq}; \phi_{iq})$, defined in (11), is evaluated at $\hat{\xi}$, that is $\mu_{iq}^r = g_1^{-1}(f_1(\mathbf{z}_t^\top \hat{\mathbf{x}}_t^*; \alpha, \beta))$, $\hat{\phi}_{iq}^r = g_2^{-1}(f_2(\mathbf{v}_t^\top, \hat{\mathbf{x}}_t^*; \gamma, \lambda))$ and $\hat{\mathbf{x}}_t^* = \hat{\mu}_{x_t|w_t} + \sqrt{2\hat{\sigma}_{x_t|w_t}^2} s_q$.

Let $\tilde{\theta}$ be the approximate pseudo maximum likelihood estimator of θ , obtained by maximizing (12). Then, under assumptions that are usually satisfied in practice, it can be shown that the asymptotic distribution of $\sqrt{n}(\tilde{\theta} - \theta)$ is normal with zero mean and covariance matrix

$$\Sigma = I_{\theta\theta}^{-1} + I_{\theta\theta}^{-1} I_{\theta\xi} \Sigma_{\xi\xi}^{-1} I_{\theta\xi}^\top I_{\theta\theta}^{-1},$$

with $I_{\theta\theta} = -\sum_{t=1}^n E\left\{\frac{\partial^2 \ell_{pt}(\theta, \xi)}{\partial \theta \partial \theta^\top}\right\}$, $I_{\theta\xi} = -\sum_{t=1}^n E\left\{\frac{\partial^2 \ell_{pt}(\theta, \xi)}{\partial^2 \theta \partial \xi^\top}\right\}$ and $\Sigma_{\xi\xi} = \sum_{t=1}^n E\left\{\frac{\partial^2 \ell_{qt}(\xi)}{\partial \xi \partial \xi^\top}\right\}$, where $\ell_{rt}(\xi)$ and $\ell_{pt}(\theta, \xi)$ are, respectively, the t -th element of the log-likelihood restricted to perturbation parameters $\ell_r(\xi) = \sum_{t=1}^n \ell_{1t}(\xi)$, and of $\ell_p(\theta, \xi)$ given in (12). For the nonlinear beta regression model with measurement errors, the matrices I_θ , I_ξ and I_θ have no explicit forms. Since the integral is approximated numerically, simply the second derivatives of $\ell_{rt}(\xi)$ and $\ell_{pt}(\theta, \xi)$ are used, thus the expected information matrix is replaced by the observed information matrix [14].

2.3 Estimation by regression calibration

When we estimate by regression calibration, the central idea is to replace the non-observed variable x_t by an estimate of the expected value of x_t given w_t , $E(x_t|w_t)$ in the likelihood function. If only one variable is measured with error, the calibration function is $r(w_t; \xi, \sigma_e^2) = \mu_x + k_x(w_t - \mu_x)$. Since $w_t \sim \mathcal{N}(\mu_x, \sigma_x^2 + \sigma_e^2)$, then, $\bar{w} = \sum_{t=1}^n w_t/n$ and

$s_w^2 = \sum_{t=1}^n (w_t - \bar{w})^2 / (n - 1)$ are optimal estimators of μ_x and $\sigma_x^2 + \sigma_e^2$, respectively. However, it is not possible to estimate σ_x^2 from the observed data $w_1, \dots, w_n$. We assume that k_x or σ_e^2 to be known, or, alternatively, that σ_e^2 can be estimated since it is possible to observe replicas of w_t .

Replacing the estimated calibration function in the probability density function of y_t given x_t , that is, in $f(y_t|x_t; \alpha, \beta, \gamma, \lambda)$ given in (1), we obtain the modified log-likelihood as

$$\ell_{rc}(\theta) = \sum_{t=1}^n \ell(\check{\mu}_t, \check{\phi}_t),$$

where $\ell(\mu_t; \phi_t) = \log \Gamma(\phi_t) - \log \Gamma(\mu_t \phi_t) - \log \Gamma[(1 - \mu_t)\phi_t] + (\mu_t \phi_t - 1) \log y_t + [(1 - \mu_t)\phi_t - 1] \log(1 - y_t)$, $g(\check{\mu}_{tq}) = f_1(\mathbf{z}_t^\top, \check{x}_t^*; \alpha, \beta)$, $h(\check{\phi}_{tq}) = f_2(\mathbf{v}_t^\top, \check{x}_t^*; \gamma, \lambda)$ and $\check{x}_t^* = \bar{w} + \hat{k}_x(w_t - \bar{w})$, $\hat{k}_x$ being the known or estimated value of k_x . Observe that the log of the modified likelihood involves only the interest parameter θ and is the same as that of the usual beta regression model, since x_t^* plays the role of an explanatory variable that is measured with no error. Standard errors of these estimators can be estimated via nonparametric bootstrap.

3 Simulations

In order to check the performances of the estimation methods that were described in Section (2), we have carried out simulation studies with 10,000 Monte Carlo replications each. We consider beta regression models where some of the covariates are measured with error and, at the same time, is nonlinear with respect to the parameters associated to the covariates that are measured with no error. This was the structure we have used in our application. The model for which we perform our simulations is a particular case in the class of models we propose in Section (2).

Model parameters are estimated with the naive model, (ι_{naive}), which is the traditional beta likelihood function, approximated maximum likelihood (ι_a), approximated pseudo maximum likelihood (ι_p) and regression calibration (ι_{rc}). In all optimization processes, we have used the nonlinear BFGS quasi-Newton algorithm with numerical derivatives that is implemented in the programming language Ox [15]. Integral approximations were tried by using $Q = 12, 30, 50$ and 80 points of quadrature, these four different numbers of points yielding very similar results. We present here results for $Q = 50$.

We have made simulations for $n = 40, 80$ and 160. At each Monte Carlo replication, the n fixed observations of the covariate with no measurement errors, denoted by z_{t1} , were obtained as independent draws from a uniform distribution on $(0.2, 1.2)$, $t = 1, \dots, n$. Furthermore,

$$y_t|x_t \sim \text{Beta}(\mu_t, \phi_t), \quad w_t = x_t + e_t, \quad x_t \stackrel{\text{ind}}{\sim} \mathcal{N}(\mu_x, \sigma_x^2), \quad e_t \stackrel{\text{ind}}{\sim} \mathcal{N}(0, \sigma_e^2),$$

x_t and e_t being independent, $t, t' = 1, \dots, n$. The simulations were performed for three different values of k_x , namely: $k_x = 0.50$ ($\sigma_e^2 = \sigma_x^2$, high measurement error), $k_x = 0.75$ ($\sigma_e^2 = \sigma_x^2/3$, moderate measurement error) and $k_x = 0.95$ ($\sigma_e^2 = \sigma_x^2/19$, low measurement error). Finally, we consider $g_1(\mu_t) = \log \frac{\mu_t}{1-\mu_t}$, that is the logit link function and $g_2(\phi_t) = \log(\phi_t)$, $t = 1, \dots, n$.

3.1 Scenario 1—Constant precision

The model is defined as

$$g_1(\mu_t) = \alpha_1 + z_{t1}^{\alpha_2} + \beta_1 x_{t1}, \quad g_2(\phi) = \gamma_1, \quad t = 1, \dots, n.$$

The true values of the parameters submodels are $\alpha_1 = -0.6$, $\alpha_2 = 2.4$ and $\beta_1 = 0.8$. Aiming to compare the performances of the competing estimation schemes presented in Section (2), for different precision magnitudes, we have chosen three values for γ_1 , namely: 2.8, 4.0 and 5.7 leading to $\phi = 16.44$, $\phi = 54.60$ and $\phi = 298.87$, typically these ϕ 's values in beta regression represent three degree of precision model, namely: substantially low, low to medium and medium to high, respectively. Concerns to the nuisance parameters we chose $\mu_x = 0$ and $\sigma_x^2 = 1$. Thus, for $k_x = 0.50$ ($\sigma_e^2 = \sigma_x^2 = 1.0$, high measurement error), $k_x = 0.75$ ($\sigma_e^2 = 0.333$, moderate measurement error) and $k_x = 0.95$ ($\sigma_e^2 = 0.053$, low measurement error).

3.2 Scenario 2—Varying precision I

In this scenario there is a covariate measured with error in both submodels. However, the non-linearity is only in the mean submodel, such that

$$g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}, \quad g_2(\phi_t) = \gamma_1 + \lambda_1 x_{t1}, \quad t = 1, \dots, n.$$

The true values of the parameters are $\alpha_1 = -0.6$, $\alpha_2 = 2.4$, $\beta_1 = 0.8$, $\gamma_1 = 2.5$, $\lambda_1 = 0.9$. In this scenario we have $\mu_x = 0$ and $\sigma_x^2 = 1$ yielding the same values to σ_e^2 of the Scenario 1, resulting in $k_x = 0.50$ ($\sigma_e^2 = 1.0$), $k_x = 0.75$ ($\sigma_e^2 = 0.333$) and $k_x = 0.95$ ($\sigma_e^2 = 0.053$).

3.3 Scenario 3—Varying precision II

This is the most complex scenario. The beta regression model with varying precision, and the both submodels present measurement error and nonlinearity, such that the complete model is defined as

$$g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}, \quad g_2(\phi_t) = \gamma_1 + v_{t1}^2 + \lambda_1 x_{t1}, \quad t = 1, \dots, n.$$

We admit that $v_{t1} = z_{t1}$, for $t = 1, \dots, n$. The true values of the parameters are $\alpha_1 = 0.7$, $\alpha_2 = 2.0$, $\beta_1 = -1.5$, $\gamma_1 = 1.5$, $\gamma_2 = 2.0$ and $\lambda_1 = 1.3$. Here, $\mu_x = 1.5$ and $\sigma_x^2 = 0.5$. Thus we have used $k_x = 0.50$, with $\sigma_e^2 = 0.5$, $k_x = 0.75$ with $\sigma_e^2 = 0.167$ and $k_x = 0.95$ with $\sigma_e^2 = 0.026$.

3.4 Simulation results

Table 1 displays the RMSE (root mean square error) of the Scenario 1's estimators. From this table, overall, we note that the RMSE decreases considerably when the precision model, the k_x value and the sample size increase, which is the expected behavior.

However, this behavior does not hold for the γ_1 's estimators, in special when we use the naive and regression calibration methods. It is noteworthy how the $\hat{\gamma}_1$'s RMSE for these two methods become larger when ϕ increases. We shall focus on $k_x = 0.50$ and $n = 160$. Be $\phi = 16.4$ and $\phi = 298.9$, the $\hat{\gamma}_1$'s RMSE for the naive (t_{naive}), regression calibration (t_{rc}), approximate likelihood (t_a) and pseudo-likelihood (t_p) estimation methods are (0.7796, 0.7796, 0.7784, 0.7149) and (3.0352, 3.0352, 0.9895, 0.9565), respectively. Furthermore, the RMSE of $\hat{\gamma}_1$ does not decrease for the naive and regression calibration methods when the sample size increases, notably for $k_x = 0.50$ and $k_x = 0.75$.

The best RMSE-related performances of the estimators are due the approximated likelihood schemes, overall. The largest RMSEs are of the $\hat{\gamma}_1$. When $\phi = 298.9$ these RMSEs are larger than when $\phi = 16.4$. However, these RMSEs decrease substantially when the sample size and reliability coefficient increase. Furthermore, the $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\beta}_1$ do not display large RMSEs when the precision is so low. Lets focus in the t_p scheme, pseudo-likelihood and in the $\hat{\beta}_1$, the estimator related to measurement error. Be $k_x = 0.75$ and $n = 40$, when $\phi = 16.4$ and $\phi = 298.9$ their RMSEs are 0.1825 and 0.1059, respectively. Finally, it is important to note that only when

Table 1. Root mean square error (RMSE) for the estimators of α_1 , α_2 , β_1 and γ_1 for the nonlinear beta regression model with constant precision. Scenario 1.

		$g_1(\mu_t) = \alpha_1 + z_{t1}^{x_2} + \beta_1 x_{t1}$ and $g_2(\phi) = \gamma_1, t = 1, \dots, n$							
ϕ		16.4				298.9			
		$n = 40$							
k_x		α_1	α_2	β_1	γ_1	α_1	α_2	β_1	γ_1
0.95	t_{naive}	0.0945	0.7119	0.1051	0.2319	0.0370	0.2872	0.0613	1.0742
	t_{rc}	0.0951	0.7119	0.1027	0.2319	0.0377	0.2872	0.0411	1.0742
	t_a	0.0955	0.7199	0.1060	0.2982	0.0377	0.2872	0.0410	0.8217
	t_p	0.0955	0.7199	0.1060	0.2982	0.0377	0.2872	0.0410	0.8213
0.75	t_{naive}	0.1128	0.8215	0.2403	0.4402	0.0709	0.5379	0.2329	2.3364
	t_{rc}	0.1234	0.8215	0.1607	0.4402	0.0836	0.5379	0.1201	2.3364
	t_a	0.1266	0.8605	0.1862	0.8649	0.0860	0.5683	0.1112	1.5487
	t_p	0.1263	0.8603	0.1825	0.8310	0.0856	0.5667	0.1059	1.4792
0.50	t_{naive}	0.1312	0.9275	0.4357	0.7332	0.0970	0.7186	0.4340	2.9636
	t_{rc}	0.2887	0.9275	2.9256	0.7332	0.4810	0.7186	2.1507	2.9636
	t_a	0.2027	1.0201	0.4436	1.8390	0.1579	0.7832	0.3009	1.4812
	t_p	0.2079	1.0164	0.5249	1.6914	0.1699	0.7773	0.3853	1.3890
		$n = 160$							
k_x		α_1	α_2	β_1	γ_1	α_1	α_2	β_1	γ_1
0.95	t_{naive}	0.0485	0.3805	0.0651	0.1411	0.0194	0.1561	0.0502	1.0949
	t_{rc}	0.0486	0.3805	0.0496	0.1411	0.0196	0.1561	0.0208	1.0949
	t_a	0.0487	0.3839	0.0507	0.1295	0.0195	0.1537	0.0200	0.3876
	t_p	0.0487	0.3839	0.0507	0.1295	0.0195	0.1537	0.0200	0.3876
0.75	t_{naive}	0.0585	0.4568	0.2275	0.4634	0.0376	0.3148	0.2286	2.3981
	t_{rc}	0.0613	0.4568	0.0736	0.4634	0.0416	0.3148	0.0582	2.3981
	t_a	0.0615	0.4618	0.0786	0.2301	0.0427	0.3119	0.0479	0.9346
	t_p	0.0615	0.4618	0.0785	0.2297	0.0426	0.3113	0.0481	0.9203
0.50	t_{naive}	0.0692	0.5462	0.4306	0.7796	0.0526	0.4500	0.4321	3.0352
	t_{rc}	0.0832	0.5462	0.1379	0.7796	0.0696	0.4500	0.1260	3.0352
	t_a	0.0859	0.5504	0.1676	0.7784	0.0718	0.4470	0.0904	0.9828
	t_p	0.0854	0.5484	0.1606	0.7149	0.0718	0.4411	0.0937	0.9565

<https://doi.org/10.1371/journal.pone.0254103.t001>

$n = 40$, $k_x = 0.50$ and for the parameter β_1 , RMSE-related, the estimation by approximate likelihood outperforms the pseudo-likelihood method.

Fig 1 presents plots for the biases of the parameter estimators of the Scenario 1, when $n = 40, 80, 160$, $\phi \approx 17, 55, 300$ and $k_x = 0.50$. It is noteworthy as the naive structure and regression calibration perform poorly, in especial for the precision estimation. As the ϕ value increases, the biases of the γ_1 estimators moving considerably away from zero showing absolute values close to three. Both estimation schemes also show the worst performances for estimating the nonlinearity parameter α_2 , since their biases moving further away from zero as the sample size increases. However, it is important to note that the absolute biases of $\hat{\alpha}_2$, equally close to 0.2, for both methods are considerably small compared to the actual value of the parameter, $\alpha_2 = 2.4$.

In fact, it is remarkable how badly the naive β_1 estimator behaves. Its bias is constant for any value of n and ϕ and considerably high, equal to 0.4 while $\beta_1 = 0.8$. Since β_1 is the coefficient of the covariate measured with error, this behavior considerably disfavors the naive structure.

Regarding methods based on approximate likelihood (t_a and t_p), the Fig 1 leads to important conclusions. In general, their biases are not markedly high, especially when the sample

Fig 1. Biases of estimators of α_1 , α_2 , β_1 and γ_1 , for $\phi = 16.4, 54.6, 298.9$. l_a (square), l_p (circle), l_{rc} (triangle) and l_{naive} (star). Scenario 1: $g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}$, $g_2(\phi) = \gamma_1$, $t = 1, \dots, n$, $n = 40, 80, 160$. $k_x = 0.50$.

<https://doi.org/10.1371/journal.pone.0254103.g001>

size increases, even when ϕ is quite low. As an example, for β_1 , the coefficient of the covariate measured with error, the largest biases of these two estimators occur when $n = 40$ and $\phi \approx 17$, and are close to -0.1 , getting closer and closer to zero for $\phi \approx 55$ and $\phi \approx 300$ (Fig 1(g)–1(i)). Nevertheless, be α_2 , the nonlinearity parameter. We note that the biases of the approximate likelihood and the pseudo-likelihood estimators for this parameter are close to zero when $n = 160$ (Fig 1(d)–1(f)).

For significant value of the reliability coefficient, i.e., $k_x = 0.50$ only when the ι_a and ι_p methods are used, the biases of the estimators for all model parameters decrease as the sample size increases. and their performances are quite similar. However, for $\hat{\gamma}$, the ι_p method exhibits slightly lower bias than the ι_p method, particularly for $n = 40$.

The fact described above is the first evidence of a better performance of the ι_p method, typically. The smaller the $\hat{\gamma}$'s bias, the smaller the bias of the variances of the model estimators and the better the performance of the confidence intervals and hypothesis tests related to the model parameters.

In what follows we shall discuss estimation by confidence interval with respect to Scenario 1. The Table 2 displays the coverage rates of the estimated asymptotically 95% confidence intervals for $k_x = 0.50, 0.75, 0.95$, $n = 80$ and $\phi = 16.4, 298.9$.

At this point, it is interesting to note how the high $\hat{\gamma}_1$ biases negatively affected the performance of the interval estimation. In particular, the considerable overestimation of γ_1 when we use the naive and regression calibration methods leads to variances of the estimators highly underestimated. Especially the larger the value of ϕ . Therefore, we have interval lengths that are too small, to the point of not coverage the true values of the parameters. For example, when $\phi = 298.8$ the coverage rates for the parameter γ_1 based on the two methods are equal to zero, and 100% of the estimated values are greater than the true value of the parameter (Table 2).

We turn now to the coefficient for the covariate with measurement error, i.e., the parameter β_1 . Since $\hat{\beta}_1$ relative to the naive method considerably overestimates that parameter, regardless of the value of ϕ (Fig 1(j)–1(l)), in association with the underestimation of variances already commented above, its β_1 interval estimator produces values that are 100% higher than the true value of the parameter. The performance of the interval estimator based on the regression calibration method is considerably better than the naive related to β_1 . However, it is still considerably poorer than that of the ι_a and ι_b methods. For $k_x = 0.50$ and $\phi = 298.9$, the coverage rates and average lengths for ι_{rc} , ι_a and ι_p are (66.59%, 73.27%, 93.38%) and (0.40, 0.75, 2.02), respectively. Here we note the outperformance of the ι_p scheme. Its coverage rate is the closest the nominal level 95% and the average length equal to 2.02 decreases to 0.31, for $k_x = 0.75$, ensuring a coverage rate closer to the nominal level (Table 2).

We shall now evaluate the interval estimators of α_2 . Based on the Table 2 we note that for $\phi = 16.4$, the schemes: ι_{naive} , ι_{rc} , ι_a and ι_p show coverage rates near each other and equal to (90.12%, 88.94%, 91.65%, 89.45%), respectively, for $k_x = 0.75$. These coverage rates are still closer when $k_x = 0.50$ and $k_x = 0.95$. However, when $\phi = 298.9$ the behavior of the four methods differs considerably. When $k_x = 0.75$ those coverage rates above become equal to (90.93%, 89.82%, 72.78%, 75.20%). When $k_x = 0.50$, these coverage rates differ further, in especial related to ι_a estimator, being equal to (87.91%, 87.26%, 61.40%, 75.29%). Here it is necessary to increase the sample size to better understand the behavior of the estimators.

Thus, Fig 2 displays the coverage rates of the interval estimators of the beta regression model with measurement errors and nonlinearity with constant precision equal to 16.4. We consider $n = 40, 80, 160$ and $k_x = 0.95, 0.75, 0.50$.

Table 2. Coverage rates and average lengths of the nominal 95% confidence interval estimators. $k_x = 0.50, 0.75, 0.95$. **Scenario 1:** $g_1(\mu_t) = \alpha_1 + z_{t1}^{x_2} + \beta_1 x_{t1}$, $g_2(\phi) = \gamma_1$, $t = 1, \dots, n$, $n = 80$.

Parameters	Interval	$k_x = 0.95$				$k_x = 0.75$				$k_x = 0.50$			
		inf	right	sup	length	inf	right	sup	length	inf	right	sup	length
		$\phi = 16.4$											
α_1	l_{naive}	2.59	94.18	3.23	0.28	2.13	93.92	3.95	0.33	1.73	93.67	4.60	0.38
	l_{rc}	2.65	93.43	3.92	0.28	2.78	91.88	5.34	0.34	4.47	86.79	8.74	0.39
	l_a	2.61	93.43	2.99	0.28	2.45	94.49	3.06	0.35	2.55	94.06	3.39	0.49
	l_p	2.94	93.43	3.59	0.27	2.79	93.81	3.40	0.35	2.16	94.71	3.13	0.87
α_2	l_{naive}	2.14	92.02	5.84	2.23	1.61	90.12	8.27	2.55	1.29	87.30	11.41	2.83
	l_{rc}	3.24	90.46	6.30	2.25	2.34	88.94	8.72	2.59	1.70	87.50	10.80	2.89
	l_a	2.38	92.38	5.24	2.25	2.68	91.65	5.67	2.67	6.97	84.68	8.35	2.90
	l_p	3.55	90.41	6.04	2.19	3.92	89.45	6.63	2.62	5.92	86.39	7.69	4.36
β_1	l_{naive}	0.43	88.45	11.12	0.26	0.00	9.50	90.50	0.26	0.00	0.00	100.00	0.23
	l_{rc}	3.50	91.87	4.63	0.26	4.49	83.77	11.74	0.30	11.12	66.59	22.29	0.40
	l_a	3.07	94.05	2.88	0.28	2.18	95.00	2.82	0.43	10.63	85.23	4.14	0.75
	l_p	3.40	93.46	3.14	0.27	2.54	94.25	3.21	0.43	4.36	91.10	4.54	2.02
γ_1	l_{naive}	1.28	92.08	6.64	0.60	0.00	20.94	79.06	0.60	0.00	0.28	99.72	0.59
	l_{rc}	1.84	90.61	7.55	0.61	0.00	22.78	77.22	0.61	0.00	0.56	99.44	0.61
	l_a	5.24	93.60	1.16	0.70	0.82	97.39	1.79	1.28	10.63	85.69	3.68	2.49
	l_p	6.27	92.16	1.57	0.68	1.91	95.93	2.16	1.30	8.86	86.96	4.18	10.61
		$\phi = 298.9$											
α_1	l_{naive}	2.34	94.42	3.24	0.11	2.01	93.95	4.04	0.21	1.71	93.83	4.46	0.29
	l_{rc}	2.79	93.22	3.99	0.11	3.55	90.38	6.07	0.22	6.02	83.24	10.74	0.30
	l_a	2.80	94.39	2.81	0.11	7.79	84.62	7.59	0.20	5.76	87.34	6.90	0.34
	l_p	2.87	93.87	3.26	0.11	6.11	87.72	6.17	0.26	3.03	92.76	4.21	0.93
α_2	l_{naive}	2.08	93.37	4.55	0.92	1.28	90.93	7.79	1.69	0.97	87.91	11.12	2.19
	l_{rc}	2.60	91.96	5.44	0.93	1.98	89.82	8.20	1.71	1.48	87.26	11.26	2.24
	l_a	2.93	93.32	3.75	0.92	13.91	72.78	13.31	1.33	19.73	61.40	18.87	1.47
	l_p	3.81	91.81	4.38	0.89	12.72	75.20	12.08	1.57	12.08	75.29	12.63	4.07
β_1	l_{naive}	0.00	56.37	43.63	0.10	0.00	0.14	99.86	0.17	0.00	0.00	100.00	0.18
	l_{rc}	1.95	92.34	5.71	0.10	2.99	81.55	15.46	0.22	10.09	64.87	25.04	0.37
	l_a	2.39	95.27	2.34	0.11	8.28	85.77	5.95	0.23	15.00	73.27	11.73	0.36
	l_p	2.79	94.52	2.69	0.11	4.12	90.92	4.96	0.31	0.60	93.38	6.02	2.12
γ_1	l_{naive}	0.00	0.00	100.00	0.62	0.00	0.00	100.00	0.61	0.00	0.00	100.00	0.60
	l_{rc}	0.00	0.00	100.00	0.62	0.00	0.00	100.00	0.62	0.00	0.00	100.00	0.60
	l_a	0.60	97.20	2.20	2.38	36.37	56.13	7.50	3.14	48.92	39.38	11.70	2.80
	l_p	1.26	95.96	2.78	2.31	33.86	58.07	8.07	4.14	29.86	57.34	12.80	12.64

<https://doi.org/10.1371/journal.pone.0254103.t002>

Those results of coverage rates are related to the biases of the point estimators. In this case, the l_{naive} and l_{rc} methods have been incorrectly favored for being more biased. Their respective point estimators slightly overestimate α_2 , this fact associate with the considerably underestimate the variance, in particular when $\phi = 298.9$ (Fig 1(l)) leads to interval estimators that cover the true value of the parameters without loss of accuracy. However, this fact changes as the sample size increases. We can see in Fig 2(f) that when $n = 160$, the coverage rates of the interval estimators for α_2 , related to l_a and l_p methods, are close to the nominal level of 95%, while those of the l_{naive} and l_{rc} methods remain constant for all sample sizes and close to 90%.

The l_a method is the most illustrative and true in its performance, even for $n = 80$ and $k_x = 0.50$ ($\phi = 298.9$). The average bias of its respective $\hat{\gamma}_1$ can be considered equal to zero. Besides,

Fig 2. Coverage rates of the nominal 95% confidence intervals for the parameters α_1 , α_2 , β_1 and γ_1 . $k_x = 0.95, 0.75, 0.50$. Scenario 1: $g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}$, $g_2(\phi) = \gamma_1$, $t = 1, \dots, n$. $n = 40, 80, 160$ $\phi = 16.4$.

<https://doi.org/10.1371/journal.pone.0254103.g002>

its largest $\hat{\alpha}_2$ average bias is less than -0.1 while $\alpha_2 = 2.4$. Thus, the chance that their interval estimates cover the true value of the parameter only depend on the goodness of approximation of the point estimator's distribution to the normal distribution. For the t_a estimators this approximation seems to improve as the sample size increases, once that coverage rate equal to 84.68% cited above, for $n = 80$ and $\phi = 16.4$, increases and remains close to the nominal level, in this case 95%. This behavior can be seen in the Fig 2(f).

From this figure, we confirm what we had discussed in the last paragraph. The interval estimators of the t_a and t_p methods are the ones that present the closest coverage rates to the nominal level of 95%, taking into account all parameters of the model. Most importantly, these rates get closer to the nominal level when the sample size increases. For the nonlinear beta regression with measurement error and constant precision, we should conclude that the best performance regarding interval estimation, concerns the approximate pseudo maximum likelihood method.

Table 3 presents the RMSE (root mean square error) of estimators for second and third scenarios. We begin by considering the second scenario. We observe that, in practically all situations, this RMSE decreases when the sample size increases. The exceptions lie in the estimation of γ_1 of the precision model. When we use the naive model and regression calibration with significant values of the reliability coefficient k_x , i.e., $k_x = 0.75$ and $k_x = 0.50$. For example, when using the naive model for $k_x = 0.50$, we obtained $\text{RMSE}(\hat{\gamma}_1) = (0.6956, 0.7475, 0.7660)$, respectively, for $n = (40, 80, 160)$.

In fact, for smaller values of k_x it is reasonable that for at least one estimator (in our case it was $\hat{\gamma}_1$) we should obtain a larger RMSE when n increases when running the naive model. This is because, for small values of k_x , the naive model represents a poor approximation for the true model and larger values of n represent more information about data. Therefore, larger values of n will detect the approximation is not good. This is translated in a larger RMSE for at least one estimator. Also, since regression calibration does not work with the full information about data distribution, but only with moments, it is expected that this loss of information that is intrinsic to regression calibration will be felt in the estimation process if more data are available. Anyway, the poor performance of the naive model is a clear indication of the necessity, in practice, of the model here proposed.

We turn now to the parameter β_1 , which is the coefficient for the covariate that has error measurement. Estimation by approximate maximum likelihood (t_a) and approximate pseudo maximum likelihood (t_p) present better performances. This is expected, since they use the full information about data distribution. We then investigate if calibration regression has comparable performance, and we can see it does not. For example, in the first scenario, for $k_x = 0.50$ and $n = 40$, we have for the $\hat{\lambda}_1$ estimator that $\text{RMSE}(\hat{\lambda}_1) = 7.549$, while the values of the RMSE are typically smaller than 1 for the other estimators.

In general, with respect to the complete estimation of the model, when we have large measurement errors, $k_x = 0.50$, and the sample size increases, the performances of the approximate pseudo maximum likelihood and the approximate maximum likelihood are comparable. Therefore, the pseudo maximum likelihood method is here recommended.

From the simulation results in Table 3, we can compare, for different reliability coefficients, the performances of the different estimators in terms of RMSE for the beta regression model with measurement errors in one covariate and nonlinearity in one parameter both for mean and precision submodel, i.e., **Scenario 3**. For the coefficient β_1 of the covariate with measurement error, the naive model is the one presenting the worst performance, in particular, when the reliability coefficient is $k = 0.50$. For example, for $n = 80$ and $k = 0.50$, we have the RMSEs of the estimators t_{naive} , t_{rc} , t_a and t_p for β_1 given, respectively, by 0.8355, 0.4412, 0.3088 and

Table 3. Root mean square error (RMSE) for the estimators of α_1 , α_2 , β_1 , γ_1 , γ_2 and λ_1 for the nonlinear beta regression model with nonconstant precision. Scenario 2 and Scenario 3.

			$\mathbf{g_1(\mu_i) = \alpha_1 + z_{i1}^{z_2} + \beta_1 x_{i1} \text{ and } g_2(\phi_i) = \gamma_1 + \lambda_1 x_{i1}}$					$\mathbf{g_1(\mu_i) = \alpha_1 + z_{i1}^{z_2} + \beta_1 x_{i1} \text{ and } g_2(\phi_i) = \gamma_1 + v_{i1}^{z_2} + \lambda_1 x_{i1}}$					
k_x	n		α_1	α_2	β_1	γ_1	λ_1	α_1	α_2	β_1	γ_1	γ_2	λ_1
0.95	40	t_{naive}	0.1132	0.7682	0.1163	0.2426	0.2903	0.2042	0.4090	0.1247	0.4926	1.2504	0.4405
		t_{rc}	0.1137	0.7682	0.1123	0.2433	0.2955	0.1483	0.4090	0.0825	0.5342	1.2504	0.4149
		t_a	0.1138	0.7717	0.1135	0.3348	0.3141	0.1569	0.4231	0.0878	0.8864	1.3151	0.5074
		t_p	0.1138	0.7717	0.1135	0.3348	0.3141	0.1568	0.4231	0.0877	0.8856	1.3140	0.5069
	80	t_{naive}	0.0858	0.6358	0.0867	0.1687	0.2162	0.1847	0.3510	0.1160	0.4179	1.2012	0.4039
		t_{rc}	0.0861	0.6358	0.0742	0.1691	0.2015	0.1250	0.3510	0.0705	0.4620	1.2012	0.3730
		t_a	0.0858	0.6377	0.0746	0.2042	0.1979	0.1261	0.3646	0.0701	0.6522	1.1742	0.4113
		t_p	0.0858	0.6377	0.0746	0.2042	0.1979	0.1260	0.3646	0.0700	0.6518	1.1736	0.4111
	160	t_{naive}	0.0581	0.4101	0.0723	0.1368	0.1857	0.1567	0.2457	0.1030	0.2996	1.0786	0.3632
		t_{rc}	0.0583	0.4101	0.0528	0.1374	0.1592	0.0879	0.2457	0.0505	0.3489	1.0786	0.3225
		t_a	0.0585	0.4122	0.0530	0.1343	0.1381	0.0877	0.2611	0.0488	0.3965	0.9533	0.2649
		t_p	0.0585	0.4122	0.0530	0.1343	0.1381	0.0877	0.2611	0.0488	0.3965	0.9529	0.2649
0.75	40	t_{naive}	0.1340	0.8891	0.2553	0.4274	0.4763	0.6845	0.6917	0.4640	0.6493	1.6562	0.7947
		t_{rc}	0.1430	0.8893	0.1803	0.4352	0.4232	0.3143	0.6917	0.1964	0.9770	1.6562	0.6683
		t_a	0.1409	0.9013	0.1836	0.7116	0.5336	0.3219	0.8580	0.1938	2.1235	2.2013	1.0610
		t_p	0.1404	0.9006	0.1799	0.6811	0.5071	0.3113	0.8562	0.1855	2.0538	2.1559	1.0282
	80	t_{naive}	0.1038	0.7448	0.2397	0.4547	0.4547	0.6668	0.6070	0.4563	0.6135	1.6753	0.7780
		t_{rc}	0.1094	0.7448	0.1154	0.4594	0.3353	0.2680	0.6070	0.1699	0.9286	1.6753	0.6339
		t_a	0.1035	0.7381	0.1173	0.3801	0.3202	0.2757	0.8135	0.1654	1.7196	2.1238	0.8920
		t_p	0.1035	0.7380	0.1166	0.3745	0.3130	0.2673	0.8126	0.1593	1.6758	2.0838	0.8670
	160	t_{naive}	0.0705	0.5000	0.2347	0.4661	0.4480	0.6412	0.4455	0.4473	0.5971	1.6298	0.7493
		t_{rc}	0.0749	0.5000	0.0802	0.4703	0.3052	0.1878	0.4455	0.1283	0.8971	1.6298	0.5794
		t_a	0.0702	0.4840	0.0797	0.2170	0.2022	0.1844	0.6788	0.1107	0.9473	1.8018	0.5794
		t_p	0.0703	0.4840	0.0798	0.2174	0.2015	0.1813	0.6784	0.1085	0.9265	1.7809	0.5679
0.50	40	t_{naive}	0.1594	1.0040	0.4475	0.6956	0.6656	1.2176	0.8865	0.8412	0.8002	1.7773	1.0005
		t_{rc}	1.1022	1.0040	4.0169	2.2007	7.5492	0.8632	0.8865	0.5608	1.6009	1.7772	0.8471
		t_a	0.2124	1.0420	0.4163	1.3432	0.9974	0.5995	1.0594	0.3769	2.0546	2.3897	1.1270
		t_p	0.2259	1.0317	0.5150	1.2253	0.9892	0.5769	1.0520	0.3627	1.9312	2.3062	1.0705
	80	t_{naive}	0.1276	0.8540	0.4410	0.7475	0.6615	1.2049	0.7907	0.8355	0.7923	1.8151	0.9887
		t_{rc}	0.1557	0.8540	0.2777	0.7582	0.4910	0.6787	0.7907	0.4412	1.4915	1.8151	0.7611
		t_a	0.1389	0.8471	0.2259	0.7885	0.5571	0.4929	0.9886	0.3088	1.7557	2.3761	0.9814
		t_p	0.1391	0.8443	0.2337	0.7464	0.5162	0.4861	0.9837	0.3032	1.6658	2.2847	0.9384
	160	t_{naive}	0.0891	0.6079	0.4388	0.7660	0.6598	1.1822	0.6115	0.8312	0.8025	1.7750	0.9793
		t_{rc}	0.1065	0.6079	0.1521	0.7737	0.4313	0.3725	0.6115	0.5561	1.3753	1.7750	0.6858
		t_a	0.0913	0.5566	0.1456	0.4936	0.3439	0.3167	0.8012	0.2008	1.1880	2.0265	0.6806
		t_p	0.0915	0.5557	0.1470	0.4785	0.3287	0.3085	0.8004	0.1953	1.1401	1.9649	0.6557

<https://doi.org/10.1371/journal.pone.0254103.t003>

0.3032. Thus, the approximate maximum likelihood and approximate pseudo maximum likelihood estimators present much better performances. The much worst performance of the naive estimator indicates again the usefulness of our proposed model. Also, we observe that the approximate maximum likelihood and approximate pseudo maximum likelihood are comparable, which means that pseudo-likelihood can be a useful tool in our specific proposal.

Figs 3 and 4 present plots for the biases of the estimators of the parameters in Scenario 2, for $n = 40, 80, 160$. It is possible to see that the naive structure and regression calibration

Fig 3. Biases of estimators of α_1 , α_2 and β_1 for $k_x = 0.95, 0.75, 0.50$. t_a (square), t_p (circle), t_{rc} (triangle) and t_{naive} (star). **Scenario 2:** $g_1(\mu_t) = \alpha_1 + z_{t1}^{z_2} + \beta_1 x_{t1}$, $g_2(\phi_t) = \gamma_1 + \lambda_1 x_{t1}$, $t = 1, \dots, n$, $n = 40, 80, 160$.

<https://doi.org/10.1371/journal.pone.0254103.g003>

present much worst biases. Not only that, but it was observed for these two last estimators, that absolute values of biases increase with measurement error. In particular, consider $\hat{\alpha}_2$ and $\hat{\beta}_1$, estimators we are very interested in, since they are related with the nonlinearity and measurement error of the mean submodel. Fig 3(d)–3(i) show how biases of both estimators are near

Fig 4. Biases of estimators of γ_1 and λ_1 for $k_x = 0.95, 0.75, 0.50$. ι_a (square), ι_p (circle), ι_{rc} (triangle) and ι_{naive} (star). **Scenario 2:** $g_1(\mu_t) = \alpha_1 + z_{t1}^{x_2} + \beta_1 x_{t1}$, $g_2(\phi_t) = \gamma_1 + \lambda_1 x_{t1}$, $t = 1, \dots, n$, $n = 40, 80, 160$.

<https://doi.org/10.1371/journal.pone.0254103.g004>

zero when we use approximate likelihood (ι_a) and approximate pseudo likelihood (ι_p) for all sample sizes.

It is remarkable how bias of $\hat{\alpha}_2$, (related to nonlinearity) increases with measurement error for the regression calibration method. (Fig 3(d)–3(f)). This same behavior can be observed for the precision submodel parameters, in particular for $\hat{\lambda}_1$, which estimates the coefficient of the covariate with measurement error. In this case, the biases of the estimators using regression calibration increase considerably for large sample sizes and measurement errors. (Fig 4(d)–4(f)). In fact only approximate maximum likelihood and approximate pseudo maximum likelihood produce estimators with biases tending to zero as the sample size increases.

Figs 5 and 6 present the plots of biases of the estimators of the parameters in **Scenario 3**, namely: nonlinear predictors with measurement error on both submodels. Approximate maximum likelihood (ι_a) and approximate pseudo maximum likelihood (ι_p) present much better performances than those of naive and regression calibration methods (ι_{rc}) for $\hat{\beta}_1$, the estimator of the coefficient of the covariate with measurement error (mean submodel). (Fig 5(g)–5(i)). Estimators of the parameters related to nonlinearity are more biased, but, even so, the performances of ι_a and ι_p methods are slightly superior to those of the ι_{Naive} and ι_{rc} , in particular for

Fig 5. Biases of estimators of α_1 , α_2 and β_1 for $k_x = 0.95, 0.75, 0.50$. l_a (square), l_p (circle), l_{rc} (triangle) and l_{naive} (star). **Scenario 3:** $g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}$, $g_2(\phi_t) = \gamma_1 + \nu_t^2 + \lambda_1 x_{t1}$, $t = 1, \dots, n$, $n = 40, 80, 160$.

<https://doi.org/10.1371/journal.pone.0254103.g005>

larger values of the measurement error (Fig 5(f)). We can also check that, for the parameters of the precision submodel, the maximum likelihood based methods are clearly superior (Fig 6). In fact, once more, only the l_a and l_p methods yield estimators with biases decreasing when sample size increases.

Fig 6. Biases of estimators of γ_1 , γ_2 and λ_1 for $k_x = 0.95, 0.75, 0.50$. l_a (square), l_p (circle), l_{rc} (triangle) and l_{naive} (star). **Scenario 3:** $g_1(\mu_t) = \alpha_1 + z_{t1}^{x_2} + \beta_1 x_{t1}$, $g_2(\phi_t) = \gamma_1 + \gamma_2^2 + \lambda_1 x_{t1}$, $t = 1, \dots, n$, $n = 40, 80, 160$.

<https://doi.org/10.1371/journal.pone.0254103.g006>

Figs 7 and 8 present the coverage rates for the different sample sizes and reliability coefficients of the interval estimators of the beta regression model with measurement errors and nonlinearity in both submodels (Scenario 3). From those figures, it becomes evident how the approximate pseudo maximum likelihood method is a good alternative to estimate the parameters of the models.

Fig 7. Coverage rates of the nominal 95% confidence intervals for the parameters α_1 , α_2 and β_1 . l_a (square), l_p (circle), l_{rc} (triangle) and l_{naive} (star). Scenario 3: $g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}$, $g_2(\phi_t) = \gamma_1 + v_t^2 + \lambda_1 x_{t1}$, $t = 1, \dots, n$, $n = 40, 80, 160$.

<https://doi.org/10.1371/journal.pone.0254103.g007>

It is still important to stress how inference based on interval estimation for β_1 can be affected when we do not consider that the covariate is measured with error, that is, when we use the naive method. That becomes more remarkable if we think this is perhaps the parameter we are more interested in, the coefficient of the covariate in the mean submodel. Fig 7(g)–7(i)

Fig 8. Coverage rates of the nominal 95% confidence intervals for the parameters γ_1 , γ_2 and λ_1 . t_a (square), t_p (circle), t_{rc} (triangle) and t_{naive} (star). Scenario 3: $g_1(\mu_t) = \alpha_1 + z_{t1}^2 + \beta_1 x_{t1}$, $g_2(\phi_t) = \gamma_1 + v_{t2}^2 + \lambda_1 x_{t1}$, $t = 1, \dots, n$, $n = 40, 80, 160$.

<https://doi.org/10.1371/journal.pone.0254103.g008>

clearly shows how we lose with the naive estimation, even when the measurement error is small, $k_x = 0.95$, since coverage rates for the naive estimation become considerably distant from the nominal 95% level. This is particular true for large sample sizes (Fig 7(g)). The sample size $n = 160$ is information enough to detect that we may be dealing with a wrong asymptotic

normal distribution for the parameters, or, perhaps, even with nonconsistent parameters, due to the wrong assumption that the covariate has no measurement error. This behavior tends to become more seriously large as the measurement error increases, as expected (Fig 7(h) and 7(i)).

Fig 8 presents interval estimation of the parameters of the precision submodel. Here we emphasize, in particular, the very good performance of approximate pseudo maximum likelihood estimation. Considering all parameters, measurement errors and sample sizes, this seems to be the recommended method. This is an important aspect, since good estimation of the precision parameters of the observations is directly linked with the efficiency of the estimators and robustness of hypotheses tests about the model parameters.

4 An application: Fluid Catalytic Cracking (FCC)

Fluid Catalytic Cracking (FCC) is an important chemical process used in petroleum refineries to convert heavy hydrocarbons in small molecules with high commercial value. This process is accomplished by contact of those hydrocarbons with a catalyst having as main component a mineral that is known as zeolite Y [16]. A chemical element that can be found in FCC is vanadium. This chemical element takes part in the process by reducing, among other characteristics, the crystallinity of zeolite Y, in particular when water steam is present. The chemical reaction also depends on temperature, which must be near 720 °C (Salazar, 2005). Hence, at the end of the process, it is important to study how crystallinity fraction of zeolite Y is influenced by different concentrations of vanadium, by water steam and by temperature. The data set is constituted of $n = 28$ observations. The authors report that the response is moderately concentrated in the upper part of the unit interval. According to them, 75% of the observations are not smaller than 0.77. This characteristic will be considered as an important factor in our modeling.

[5] have modeled the response, considering the beta and simplex distributions, by using a nonlinear predictor for the mean but without measurement errors in the covariates. However, following [16], the concentration of vanadium for each observation was determined by a spectrophotometric method, with the aid of a chemical complex that includes the element and using a calibration curve. This calibration curve was defined with the same spectrophotometric method, by using another complex, that includes vanadium and a different reagent. Therefore, the concentration of vanadium here can be considered a covariate that is obtained for each case with measurement error. Extending the nonlinear model in [5], we propose the nonlinear beta regression model with measurement errors given by

$$g_1(\mu_t) = \alpha_1 + \alpha_2 \frac{z_{1t}}{z_{1t} + \alpha_3} + \alpha_4 z_{2t} + \beta_1 x_{1t}, \quad \log(\phi) = \gamma_1, \quad t = 1, \dots, 28, \quad (13)$$

where x_{1t} represents vanadium concentration, z_{1t} denotes water steam and z_{2t} is a categorical variable indicating in which of the two temperatures the experiment was done ($0 = 700$ °C and $1 = 760$ °C). Since x_{1t} is a covariate with measurement error, we need to previously estimate the nuisance parameters μ_x , σ_x^2 and σ_e^2 . Thus, we perform a simple linear regression for the relation between the original x and observed w chemical complexes, as $w_t = \tau_1 + \tau_2 x_t + e_t$, with $t = 1, \dots, 15$, $e_t \sim N(0, \sigma_e^2)$. The obtained estimates were $\hat{\tau}_1 = 0.181$ (0.123), $\hat{\tau}_2 = 0.713$ (0.089) and $\hat{\sigma}_e^2 = 0.03810$, with $R^2 = 0.98$. Also, we have $\hat{\sigma}_x^2 = 0.68$, supposing $x_t \sim N(\mu_x, \sigma_x^2)$. This yields an estimate of the reliability coefficient $k_x = \sigma_x^2 / (\sigma_x^2 + \sigma_e^2)$ of $\hat{k}_x = 0.9436$, a low measurement error.

We have tried two different link functions: the logit and the complementary log-log. According to [17], when the mean of the response variable is near 1, the complementary

Table 4. Lower Limits (LL), upper limits (UL) and length (length) of interval estimations obtained considering 90%, 95% and 99% for confidence level, using as estimation methods t_{Naive} , regression calibration (t_{rc}), approximate maximum likelihood (t_a) and approximate pseudo maximum likelihood (t_p). Logit link function. Fluid Catalytic Cracking (FCC) application.

Methods	Parameters	CI 90%			CI 95%			CI 99%		
		LL	UL	length	LL	UL	length	LL	UL	length
t_{Naive}	α_1	2.05259	2.47829	0.42570	2.01182	2.51907	0.50725	1.93212	2.59876	0.66664
	α_2	-0.15852	-0.04847	0.11005	-0.16906	-0.03793	0.13113	-0.18966	-0.01733	0.17234
	α_3	-30.85159	-21.42127	9.43032	-31.75489	-20.51797	11.23692	-33.52034	-18.75252	14.76782
	α_4	-0.53894	-0.16029	0.37865	-0.57521	-0.12402	0.45119	-0.64610	-0.05314	0.59296
	β_1	-0.72235	-0.39372	0.32863	-0.75383	-0.36225	0.39159	-0.81536	-0.30072	0.51463
	γ_1	3.82316	4.69967	0.87650	3.73921	4.78362	1.04442	3.57512	4.94771	1.37260
t_{rc}	α_1	2.07344	2.50673	0.43329	2.03194	2.54823	0.51630	1.95082	2.62935	0.67853
	α_2	-0.18173	-0.02527	0.15646	-0.19671	-0.01028	0.18643	-0.22600	0.01901	0.24502
	α_3	-30.34481	-21.92805	8.41676	-31.15103	-21.12183	10.02919	-32.72673	-19.54613	13.18059
	α_4	-0.61799	-0.08124	0.53675	-0.66940	-0.02983	0.63957	-0.76989	0.07065	0.84054
	β_1	-0.74831	-0.44067	0.30764	-0.77778	-0.41120	0.36658	-0.83537	-0.35361	0.48176
	γ_1	3.69926	4.82357	1.12431	3.59157	4.93126	1.33969	3.38109	5.14174	1.76066
t_a	α_1	2.07119	2.51090	0.43971	2.02907	2.55302	0.52395	1.94675	2.63534	0.68859
	α_2	-0.15887	-0.04902	0.10985	-0.16939	-0.03850	0.13089	-0.18996	-0.01793	0.17202
	α_3	-30.83006	-21.37182	9.45824	-31.73603	-20.46584	11.27019	-33.50671	-18.69517	14.81154
	α_4	-0.54651	-0.16333	0.38319	-0.58322	-0.12662	0.45660	-0.65496	-0.05489	0.60007
	β_1	-0.76658	-0.41297	0.35361	-0.80045	-0.37910	0.42135	-0.86665	-0.31290	0.55375
	γ_1	3.85072	4.79551	0.94480	3.76022	4.88601	1.12579	3.58334	5.06288	1.47954
t_p	α_1	2.04896	2.53316	0.48420	2.00258	2.57954	0.57696	1.91193	2.67018	0.75825
	α_2	-0.14984	-0.05805	0.09179	-0.15863	-0.04926	0.10938	-0.17582	-0.03207	0.14375
	α_3	-29.30830	-22.89357	6.41473	-29.92275	-22.27912	7.64362	-31.12365	-21.07822	10.04542
	α_4	-0.53200	-0.17784	0.35415	-0.56592	-0.14392	0.42200	-0.63222	-0.07762	0.55460
	β_1	-0.76217	-0.41743	0.34474	-0.79519	-0.38441	0.41078	-0.85973	-0.31987	0.53986
	γ_1	3.86546	4.78077	0.91531	3.77779	4.86844	1.09066	3.60643	5.03980	1.43337

<https://doi.org/10.1371/journal.pone.0254103.t004>

log-log link function for the mean softens the impact of influent points in maximum likelihood estimation. In this sense, we consider here the logit and the complementary log-log link functions and check the performances of both modelings.

Tables (4) and (5) present the lower and upper limits of the estimated confidence intervals for the parameters of Model (13) obtained for the confidence levels of 90%, 95% and 99%, by considering the logit and complementary log-log link functions, respectively. We also present the lengths of the interval estimates. From both tables, it can be seen that the only interval estimate having zero is that for α_2 at the nominal level of 99% using the method of regression calibration and with the logit model for the mean. Note that by accepting the hypothesis that $\alpha_2 = 0$, we eliminate the term $\alpha_2 \frac{z_{1t}}{z_{1t} + \alpha_3}$ from the model and, consequently, not only the covariate z_{1t} , but also nonlinearity. Also, observe that the length 0.245 of the confidence interval is quite large, more than twice the length of 0.0936 for the confidence interval of this same parameter and method of estimation, but using the complementary log-log as the link function.

In fact, it is remarkable how the complementary log-log model presents better interval estimation with confidence intervals having lengths much smaller than the corresponding ones for the logit model, in all cases. For example, for α_3 , the coefficient related to temperature, when the nominal confidence level is 95% and we use approximate pseudo maximum likelihood, the lengths of the interval estimates are 0.4220 and 0.2099 for the logit and complementary log-log models, respectively. Comparing the estimation methods, it is also remarkable that

Table 5. Lower Limits (LL), upper limits (UL) and length (length) of interval estimations obtained considering 90%, 95% and 99% for confidence level, using as estimation methods t_{Naive} , regression calibration (t_{rc}), approximate maximum likelihood (t_a) and approximate pseudo maximum likelihood (t_p). Complementary log-log link function. Fluid Catalytic Cracking (FCC) application.

Methods	Parameters	CI 90%			CI 95%			CI 99%		
		LL	UL	length	LL	UL	length	LL	UL	length
t_{Naive}	α_1	0.80588	0.98914	0.18326	0.78833	1.00670	0.21837	0.75402	1.04101	0.28698
	α_2	-0.07247	-0.02467	0.04781	-0.07705	-0.02009	0.05696	-0.08600	-0.01114	0.07486
	α_3	-30.96676	-22.02752	8.93923	-31.82302	-21.17126	10.65176	-33.49653	-19.49775	13.99878
	α_3	-0.27015	-0.09198	0.17818	-0.28722	-0.07491	0.21231	-0.32058	-0.04155	0.27902
	β_1	-0.35046	-0.19772	0.15274	-0.36509	-0.18309	0.18200	-0.39368	-0.15450	0.23919
	γ_1	3.88034	4.75715	0.87681	3.79636	4.84114	1.04478	3.63221	5.00528	1.37307
t_{rc}	α_1	0.81184	1.00739	0.19555	0.79311	1.02612	0.23301	0.75651	1.06273	0.30623
	α_2	-0.07846	-0.01868	0.05979	-0.08419	-0.01295	0.07124	-0.09539	-0.00176	0.09363
	α_3	-30.67078	-22.32350	8.34728	-31.47034	-21.52395	9.94640	-33.03303	-19.96125	13.07178
	α_4	-0.31937	-0.04276	0.27661	-0.34586	-0.01627	0.32960	-0.39765	0.03552	0.43316
	β_1	-0.36819	-0.21580	0.15240	-0.38279	-0.20120	0.18159	-0.41132	-0.17267	0.23865
	γ_1	3.78844	4.84905	1.06060	3.68685	4.95064	1.26379	3.48830	5.14919	1.66090
t_a	α_1	0.81567	1.00411	0.18844	0.79762	1.02216	0.22454	0.76235	1.05744	0.29509
	α_2	-0.07258	-0.02485	0.04773	-0.07715	-0.02028	0.05687	-0.08609	-0.01135	0.07474
	α_3	-30.93200	-21.97158	8.96042	-31.79029	-21.11329	10.67700	-33.46777	-19.43581	14.03196
	α_4	-0.27103	-0.09222	0.17881	-0.28815	-0.07509	0.21306	-0.32163	-0.04162	0.28001
	β_1	-0.37361	-0.20958	0.16403	-0.38932	-0.19387	0.19546	-0.42003	-0.16316	0.25687
	γ_1	3.92316	4.88551	0.96235	3.83098	4.97769	1.14671	3.65082	5.15786	1.50703
t_p	α_1	0.80695	1.01283	0.20588	0.78723	1.03256	0.24533	0.74869	1.07110	0.32241
	α_2	-0.06756	-0.02987	0.03769	-0.07117	-0.02626	0.04492	-0.07823	-0.01920	0.05903
	α_3	-29.56210	-23.34146	6.22064	-30.15796	-22.74561	7.41235	-31.32252	-21.58104	9.74148
	α_4	-0.26970	-0.09354	0.17616	-0.28658	-0.07667	0.20991	-0.31956	-0.04369	0.27587
	β_1	-0.37307	-0.21012	0.16296	-0.38868	-0.19451	0.19417	-0.41919	-0.16400	0.25519
	γ_1	3.93373	4.87495	0.94122	3.84357	4.96511	1.12153	3.66737	5.14131	1.47395

<https://doi.org/10.1371/journal.pone.0254103.t005>

approximate pseudo maximum likelihood yields interval estimates with lengths, that are, in general, smaller than those of the interval estimates obtained with the other estimation methods. For example, using the complementary log-log link function, the confidence intervals for α_3 , the parameter that makes the model nonlinear, the associated lengths for t_p are (6.22, 7.41, 9.74), while those associated to t_a , t_{rc} and t_{naive} are, respectively, (8.96, 10.68, 14.03); (8.35, 9.95, 13.07); (8.94, 10.65, 13.99).

It is important here to observe that approximate pseudo maximum likelihood presents much better performance than the naive model, that which does not take measurement error into account, even for a very low measurement error, that is $\hat{k}_x \approx 0.95$. This shows how important to consider measurement errors can be when estimating the beta nonlinear regression models.

From all the results here presented, we select for this particular data set the nonlinear beta regression model for which the covariate x_{1t} , the vanadium concentration, has measurement error, and the regression structure defined by

$$\log[-\log(1 - \mu_t)] = \alpha_1 + \alpha_2 \frac{z_{1t}}{z_{1t} + \alpha_3} + \alpha_4 z_{2t} + \beta_1 x_{1t} \quad \log(\phi) = \gamma_1, \quad t = 1, \dots, 28.$$

Inference for measurement error in x_{1t} , produces $\hat{\sigma}_e^2 = 0.03810$, $\hat{\sigma}_x^2 = 0.68$, $\hat{k}_x \approx 0.95$ and estimation based on approximate pseudo maximum likelihood yields $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\beta}_1, \hat{\gamma}_1$ equal to 0.910, -0.049, -26.451, -0.182, -0.292 and 4.405, respectively, with standard errors given by 0.063, 0.011, 1.891, 0.054, 0.050 and 0.286, all parameters being significant at the 1% level. Finally, note that $\hat{\phi} = \exp(\hat{\gamma}_1) \approx 82$ which for the beta regression can be considered a moderate precision, neither too low, but far from being considered high. That is, the application shows similarities with Scenario 1 of the simulations, for $\phi = 55$ and $K_x = 0.95$. The results of the application only confirm those of the simulations in which, among the methods evaluated, the approximate pseudo maximum likelihood method presents the best performance regarding interval estimation.

5 Conclusions

We have proposed in this work a beta regression model with parameter nonlinearity and covariates measurement with errors for mean and precision submodels. Log-likelihood for this kind of model are written in terms of integrals with difficult solutions. For this reason, we have proposed to approximate those integrals using a Gaussian quadrature. This yields the estimators which we have referred to in this paper as the approximate maximum likelihood estimators. Because the approximate likelihood can become a very difficult function to maximize, we have also considered an approximate pseudo maximum likelihood, where we first estimate the nuisance parameters. The advantage of this two-phase method is to reduce the dimension of the problem, which makes maximization problem easier. We have also tried a regression calibration method, where the nonobserved variable is replaced by its conditional expectation in the likelihood function.

Numerical simulations have shown that the approximate pseudo maximum likelihood proposal is a very good alternative to estimate the model parameters, competing with the original maximum likelihood estimators. On the other hand, regression calibration is not a good proposal, which means that too much information about the observations can be truncated when we use this method. Therefore, regression calibration is not recommended here, although we strongly recommend approximate pseudo maximum likelihood.

An application has also shown that approximate pseudo maximum likelihood presents much better performance than the naive model, which does not take measurement error into account, even for very low measurement error. This shows how important to consider measurement errors can be when estimating beta nonlinear regression models. Furthermore, noteworthy are the superior performances related to the use of complementary log-log link function, in a sample whose responses are close to the upper limit of the unit interval.

6 Appendix: Approximate maximum likelihood

Differentiating the approximate log-likelihood $\ell_a(\Psi)$ defined in (10) with respect to each one of the interest parameters $\theta = (\alpha^\top, \beta, \gamma^\top, \lambda)^\top$, we obtain the approximate score function, given by $U_a(\theta^\top) = (U_\alpha^\top, U_\beta, U_\gamma^\top, U_\lambda)^\top$ with $U_\alpha(\theta) = \mathbf{F}_1^\top \mathbf{B} \mathbf{D}_1$, $U_\beta(\theta) = \mathbf{D}_2^\top b$, $U_\gamma(\theta) = \mathbf{F}_3^\top \mathbf{B} \mathbf{D}_3$ and $U_\lambda(\theta) = \mathbf{D}_4^\top b$. Also, B is a $n \times n$ diagonal matrix in which the t -th diagonal element is given by

$$b_t = \left(\sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{tq}(\mu_{tq}, \phi_{tq}) \} \right)^{-1}, \quad \mathbf{D}_1 = [(E \circ P) \circ H_1] V_q, \quad \mathbf{D}_3 = [(E \circ A) \circ H_2] V_q, \quad (14)$$

with $\mathbf{E}$, $\mathbf{A}$, $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{P}$ being $n \times Q$ matrices with entries defined as $e_{tq} = \exp\{\ell_{tq}(\mu_{tq}, \phi_{tq})\}$,

$$a_{tq} = \mu_{tq}(y_t^* - \mu_{tq}^*) + \log(1 - y_t) - \psi((1 - \mu_{tq})\phi_{tq}) + \psi(\phi_{tq}) \quad \text{with} \quad (15)$$

$$y_t^* = \log\left(\frac{y_t}{1 - y_t}\right), \quad \mu_{tq}^* = \psi(\mu_{tq}\phi_{tq}) - \psi[(1 - \mu_{tq})\phi_{tq}], \quad (16)$$

$h_{1tq} = 1/g'_1(\mu_{tq})$, $h_{2tq} = 1/g'_2(\phi_{tq})$ and $p_{tq} = \phi_{tq}(y_t^* - \mu_{tq}^*)$, respectively, where $\psi(\cdot)$ represents the digamma function (first derivative of log of gamma function). Additionally, $\mathbf{V}_q$ is a Q -dimensional vector with q -th element given by $v_q/\sqrt{\pi}$ where v_t is the weight for the orthogonal Hermite polynomial of order Q in (9). Finally, we have $D_2 = [(E \circ P) \circ (H_1 \circ F_2)]V_q$ and $D_4 = [(E \circ A) \circ (H_2 \circ F_4)]V_q$.

The approximate score vector $U_a(\xi^\top) = (U_{\mu_x}, U_{\sigma_x^2})^\top$ for the nuisance parameters is given by

$$\begin{aligned} U_{\mu_x} &= \frac{1}{(\sigma_x^2 + \sigma_e^2)} \sum_{t=1}^n (w_t - \mu_x) + \sum_{t=1}^n b_t a_{t\mu_x}, \quad U_{\sigma_x^2} = -\frac{1}{2(\sigma_x^2 + \sigma_e^2)} \sum_{t=1}^n \left[1 - \frac{(w_t - \mu_x)^2}{(\sigma_x^2 + \sigma_e^2)} \right] + \sum_{t=1}^n b_t a_{t\sigma_x^2}, \\ a_{t\mu_x} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp\{\ell_{tq}(\mu_{tq}, \phi_{tq})\} \left[\frac{\phi_{tq}(y_t^* - \mu_{tq}^*)}{g'_1(\mu_{tq})} \frac{\partial \eta_{1t}}{\partial \mu_x} + \frac{a_{tq}}{g'_2(\phi_{tq})} \frac{\partial \eta_{2t}}{\partial \mu_x} \right], \\ a_{t\sigma_x^2} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp\{\ell_{tq}(\mu_{tq}, \phi_{tq})\} \left[\frac{\phi_{tq}(y_t^* - \mu_{tq}^*)}{g'_1(\mu_{tq})} \frac{\partial \eta_{1t}}{\partial \sigma_x^2} + \frac{a_{tq}}{g'_2(\phi_{tq})} \frac{\partial \eta_{2t}}{\partial \sigma_x^2} \right]. \end{aligned}$$

We have that the approximate information matrix for the interest parameters is given by $-J_a(\theta)$, with

$$J_a(\theta) = \begin{bmatrix} \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha \partial \alpha^\top} & \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha \partial \gamma^\top} & \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha \partial \lambda} \\ & \frac{\partial^2 \ell_a(\Psi)}{d\beta d\beta} & \frac{\partial^2 \ell_a(\Psi)}{d\beta d\gamma^\top} & \frac{\partial^2 \ell_a(\Psi)}{d\beta d\lambda} \\ & & \frac{\partial^2 \ell_a(\Psi)}{\partial \gamma \partial \gamma^\top} & \frac{\partial^2 \ell_a(\Psi)}{\partial \gamma d\lambda} \\ & & & \frac{\partial^2 \ell_a(\Psi)}{d\lambda d\lambda} \end{bmatrix}_{n(\theta) \times n(\theta)},$$

in which $n(\theta) = (p + 1) + (\check{q} + 1) < n$ and for $i, j = 1, 2, \dots, p$ (first row), $i = 1, 2, \dots, p; j = 1, 2, \dots, \check{q}$ (second row), $j = 1, 2, \dots, \check{q}$ (third row), $k = 1, 2, \dots, \check{q}$ (the last rows) we have

that

$$\begin{aligned}
 \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha_i \partial \alpha_j} &= J_{a(\alpha_i \alpha_j)} = \sum_{t=1}^n \left[\dot{b}_{t(\alpha_j)} a_{t\alpha_i} + b_t \dot{a}_{t\alpha_i(\alpha_j)} \right], & \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha_i \partial \beta} &= J_{a(\alpha_i \beta)} = \sum_{t=1}^n \left[\dot{b}_{t(\beta)} a_{t\alpha_i} + b_t \dot{a}_{t\alpha_i(\beta)} \right], \\
 \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha_i \partial \lambda} &= J_{a(\alpha_i \lambda)} = \sum_{t=1}^n \left[\dot{b}_{t(\lambda)} a_{t\alpha_i} + b_t \dot{a}_{t\alpha_i(\lambda)} \right], & \frac{\partial^2 \ell_a(\Psi)}{\partial \alpha_i \partial \gamma_j} &= J_{a(\alpha_i \gamma_j)} = \sum_{t=1}^n \left[\dot{b}_{t(\gamma_j)} a_{t\alpha_i} + b_t \dot{a}_{t\alpha_i(\gamma_j)} \right] \\
 \frac{\partial^2 \ell_a(\Psi)}{d\beta d\beta} &= J_{a(\beta\beta)} = \sum_{t=1}^n \left[\dot{b}_{t(\beta)} a_{t\beta} + b_t \dot{a}_{t\beta(\beta)} \right], & \frac{\partial^2 \ell_a(\Psi)}{d\beta d\gamma_j} &= J_{a(\beta\gamma_j)} = \sum_{t=1}^n \left[\dot{b}_{t(\gamma_j)} a_{t\beta} + b_t \dot{a}_{t\beta(\gamma_j)} \right], \\
 \frac{\partial^2 \ell_a(\Psi)}{d\beta d\lambda} &= J_{a(\beta\lambda)} = \sum_{t=1}^n \left[\dot{b}_{t(\lambda)} a_{t\beta} + b_t \dot{a}_{t\beta(\lambda)} \right], & \frac{\partial^2 \ell_a(\Psi)}{\partial \gamma_k \partial \gamma_j} &= J_{a(\gamma_k \gamma_j)} = \sum_{t=1}^n \left[\dot{b}_{t(\gamma_j)} a_{t\gamma_k} + b_t \dot{a}_{t\gamma_k(\gamma_j)} \right] \\
 \frac{\partial^2 \ell_a(\Psi)}{\partial \gamma_k d\lambda} &= J_{a(\gamma_k \lambda)} = \sum_{t=1}^n \left[\dot{b}_{t(\lambda)} a_{t\gamma_k} + b_t \dot{a}_{t\gamma_k(\lambda)} \right], & \frac{\partial^2 \ell_a(\Psi)}{d\lambda d\lambda} &= J_{a(\lambda\lambda)} = \sum_{t=1}^n \left[\dot{b}_{t(\lambda)} a_{t\lambda} + b_t \dot{a}_{t\lambda(\lambda)} \right].
 \end{aligned} \quad (17)$$

Many expressions in (17) differ only in the parameter of interest. More specifically, they differ as to the derivatives of η_1 and η_2 with respect to these parameters. Thus, we defined a general argument ϑ varying among the interest parameters, namely: $\alpha_i, \gamma_j, \beta, \lambda$. Furthermore, we also defined the following functions of ϑ :

$$\begin{aligned}
 \dot{\ell}_{iq(\vartheta)}(\mu_{iq}, \phi_{iq}) &= \phi_{iq} [y_t^* - \mu_{iq}^*] \frac{1}{g_1'(\mu_{iq})} \frac{\partial \eta_{1t}}{\partial \vartheta}, \quad \text{for } \vartheta = \alpha_i \text{ and } \beta; \\
 &= a_{iq} \frac{1}{g_2'(\phi_{iq})} \frac{\partial \eta_{2t}}{\partial \vartheta}, \quad \text{for } \vartheta = \gamma_j \text{ and } \lambda. \\
 a_{t\vartheta} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{iq}(\mu_{iq}, \phi_{iq}) \} \dot{\ell}_{iq(\vartheta)}(\mu_{iq}, \phi_{iq}), \quad \text{for } \vartheta = \alpha_i, \beta, \gamma_j, \lambda. \\
 \dot{b}_{t(\vartheta)} &= -b_t^2 \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{iq}(\mu_{iq}, \phi_{iq}) \} \dot{\ell}_{iq(\vartheta)}(\mu_{iq}, \phi_{iq}), \quad \text{for } \vartheta = \alpha_i, \beta, \gamma_j, \lambda. \\
 \dot{a}_{t\alpha_i(\vartheta)} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{iq}(\mu_{iq}, \phi_{iq}) \} \left[\dot{\ell}_{iq(\alpha_i)}(\mu_{iq}, \phi_{iq}) \dot{\ell}_{iq(\vartheta)}(\mu_{iq}, \phi_{iq}) + \ddot{\ell}_{iq(\alpha_i)(\vartheta)}(\mu_{iq}, \phi_{iq}) \right], \text{ with} \\
 \ddot{\ell}_{iq(\alpha_i)(\vartheta)}(\mu_{iq}, \phi_{iq}) &= -\phi_{iq} \left[u_{iq} + (y_t^* - \mu_{iq}^*) \frac{g_1''(\mu_{iq})}{[g_1'(\mu_{iq})]^3} \right] \frac{\partial \eta_{1t}}{\partial \alpha_i} \frac{\partial \eta_{1t}}{\partial \vartheta}; \quad \text{for } \vartheta = \alpha_j \text{ and } \beta, \quad i, j = 1, \dots, p. \\
 &= \left[(y_t^* - \mu_{iq}^*) - c_{iq} \right] \frac{1}{g_1'(\mu_{iq}) g_2'(\phi_{iq})} \frac{\partial \eta_{1t}}{\partial \alpha_i} \frac{\partial \eta_{2t}}{\partial \vartheta}; \quad \text{for } \vartheta = \gamma_j \text{ and } \lambda, \quad j = 1, \dots, \tilde{q}. \\
 \dot{a}_{t\beta(\vartheta)} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{iq}(\mu_{iq}, \phi_{iq}) \} \left[\dot{\ell}_{iq(\beta)}(\mu_{iq}, \phi_{iq}) \dot{\ell}_{iq(\vartheta)}(\mu_{iq}, \phi_{iq}) + \ddot{\ell}_{iq(\beta)(\vartheta)}(\mu_{iq}, \phi_{iq}) \right], \text{ with} \\
 \ddot{\ell}_{iq(\beta)(\vartheta)}(\mu_{iq}, \phi_{iq}) &= -\phi_{iq} \left[u_{iq} + (y_t^* - \mu_{iq}^*) \frac{g_1''(\mu_{iq})}{[g_1'(\mu_{iq})]^3} \right] \frac{\partial \eta_{1t}}{\partial \beta} \frac{\partial \eta_{1t}}{\partial \vartheta}; \quad \text{for } \vartheta = \beta. \\
 &= \left[(y_t^* - \mu_{iq}^*) - c_{iq} \right] \frac{1}{g_1'(\mu_{iq}) g_2'(\phi_{iq})} \frac{\partial \eta_{1t}}{\partial \beta} \frac{\partial \eta_{2t}}{\partial \vartheta}; \quad \text{for } \vartheta = \gamma_j \text{ and } \lambda, \quad j = 1, \dots, \tilde{q}. \\
 \dot{a}_{t\gamma_k(\vartheta)} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{iq}(\mu_{iq}, \phi_{iq}) \} \left[\dot{\ell}_{iq(\gamma_k)}(\mu_{iq}, \phi_{iq}) \dot{\ell}_{iq(\vartheta)}(\mu_{iq}, \phi_{iq}) + \ddot{\ell}_{iq(\gamma_k)(\vartheta)}(\mu_{iq}, \phi_{iq}) \right], \text{ with} \\
 \ddot{\ell}_{iq(\gamma_k)(\vartheta)}(\mu_{iq}, \phi_{iq}) &= \left\{ d_{iq} - a_{iq} \frac{g_2''(\phi_{iq})}{[g_2'(\phi_{iq})]^3} \right\} \frac{\partial \eta_{2t}}{\partial \gamma_k} \frac{\partial \eta_{2t}}{\partial \vartheta}; \quad \text{for } \vartheta = \lambda \text{ and } \gamma_j, \quad k, j = 1, \dots, \tilde{q}. \\
 \dot{a}_{t\lambda(\lambda)} &= \sum_{q=1}^Q \frac{v_q}{\sqrt{\pi}} \exp \{ \ell_{iq}(\mu_{iq}, \phi_{iq}) \} \left[\dot{\ell}_{iq(\lambda)}(\mu_{iq}, \phi_{iq}) \dot{\ell}_{iq(\lambda)}(\mu_{iq}, \phi_{iq}) + \ddot{\ell}_{iq(\lambda)(\lambda)}(\mu_{iq}, \phi_{iq}) \right], \text{ with} \\
 \ddot{\ell}_{iq(\lambda)(\lambda)}(\mu_{iq}, \phi_{iq}) &= \left\{ d_{iq} - a_{iq} \frac{g_2''(\phi_{iq})}{[g_2'(\phi_{iq})]^3} \right\} \frac{\partial \eta_{2t}}{\partial \lambda} \frac{\partial \eta_{2t}}{\partial \lambda}.
 \end{aligned} \quad (18)$$

In (18) we have that a_t and b_t are defined in (14) and (15), respectively and

$$u_{tq} = \phi_{tq} \{ \psi'(\mu_{tq} \phi_{tq}) + \psi'[(1 - \mu_{tq}) \phi_{tq}] \} [g_1'(\mu_{tq})]^{-2}, c_{tq} = \phi_{tq} [\psi'(\mu_{tq} \phi_{tq}) \mu_{tq} + \psi'[(1 - \mu_{tq}) \phi_{tq}] (1 - \mu_{tq})]$$

and $d_{tq} = \mu_{tq} \{ -\mu_{tq} \psi'(\mu_{tq} \phi_{tq}) - (1 - \mu_{tq}) \psi'[(1 - \mu_{tq}) \phi_{tq}] \} + \psi'(\phi_{tq}) - (1 - \mu_{tq}) \psi'[(1 - \mu_{tq}) \phi_{tq}]$.

Supporting information

S1 File. We create a directory in to allocate the supplementary information, namely: “S1 File”. As supplementary material we placed two simulation programs, the application program, the data set, the file of quadrature points, and the description of the data set. (ZIP)

Acknowledgments

We thank two anonymous referees for comments and suggestions that led to a much improved manuscript.

Author Contributions

Conceptualization: Daniele de Brito Trindade, Patrícia Leone Espinheira, Klaus Leite Pinto Vasconcellos, Jalmar Manuel Farfán Carrasco.

Data curation: Patrícia Leone Espinheira.

Formal analysis: Daniele de Brito Trindade, Patrícia Leone Espinheira, Klaus Leite Pinto Vasconcellos.

Investigation: Daniele de Brito Trindade, Patrícia Leone Espinheira, Klaus Leite Pinto Vasconcellos.

Methodology: Daniele de Brito Trindade, Patrícia Leone Espinheira, Jalmar Manuel Farfán Carrasco.

Software: Daniele de Brito Trindade, Patrícia Leone Espinheira, Jalmar Manuel Farfán Carrasco.

Supervision: Patrícia Leone Espinheira, Klaus Leite Pinto Vasconcellos.

Validation: Daniele de Brito Trindade, Patrícia Leone Espinheira.

Visualization: Maria do Carmo Soares de Lima.

Writing – review & editing: Maria do Carmo Soares de Lima.

References

1. Ferrari SLP, Cribari-Neto F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*. 2004; 31(7):799–815. <https://doi.org/10.1080/0266476042000214501>
2. Simas AB, Barreto-Souza W, Rocha AV. Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis*. 2010; 54(2):348–366. <https://doi.org/10.1016/j.csda.2009.08.017>
3. Stefanski LA. The effects of measurement error on parameter estimation. *Biometrika*. 1985; 72(3):583–592. <https://doi.org/10.1093/biomet/72.3.583>
4. Carrasco JMF, Ferrari SLP, Arellano VRB. Errors-in-variables beta regression models. *Journal of Applied Statistics*. 2014; 41(7):1530–1547. <https://doi.org/10.1080/02664763.2014.881784>
5. Espinheira PL, Silva AO. Residual and influence analysis to a general class of simplex regression. *Test*. 2020; 29(2):523–552. <https://doi.org/10.1007/s11749-019-00665-3>

6. Huwang L, Huang YHS. On errors-in-variables in polynomial regression—Berkson case. *Statistics Sinica*. 2000; 10:923–936.
7. Wang L. Estimation of nonlinear models with Berkson measurement errors. *The Annals of Statistics*. 2004; 32(6):2559–2579. <https://doi.org/10.1214/009053604000000670>
8. Carroll RJ, Ruppert D, Stefanski L, Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd ed. New York: Chapman & Hall/CRC; 2006.
9. Schennach SM. Measurement error in nonlinear models—a review. *Advances in Economics and Econometrics, Theory and Applications: Tenth World Congress of the Econometric Society*. 2013; 3:296–337. <https://doi.org/10.1017/CBO9781139060035.009>
10. Carrasco JMF, Reid N. Simplex regression models with measurement error. *Communications in Statistics—Simulation and Computation*. 2019; p. 1–16. <https://doi.org/10.1080/03610918.2019.1626881>
11. Espinheira PL, Santos EG, Cribari-Neto F. On nonlinear beta regression residuals. *Biometrical Journal*. 2017; 59(3):445–461. <https://doi.org/10.1002/bimj.201600136> PMID: 28128858
12. Abramowitz M, Stegun I. *Handbook of Mathematical Functions*. New York: Dover Publications; 1972.
13. Gong G, Samaniego FJ. Pseudo maximum likelihood estimation: theory and applications. *The Annals of Statistics*. 1981; 9(4):861–869. <https://doi.org/10.1214/aos/1176345526>
14. Buonaccorsi JP, Tosteson TD. Correcting for nonlinear measurement errors in the dependent variable in the general linear model. *Communications in Statistics—Theory and Methods*. 1993; 22:2687–2702. <https://doi.org/10.1080/03610929308831179>
15. Doornik JA. *An Object-Oriented Matrix Language Ox 6*. London: Timberlake Consultants Press; 2009.
16. Salazar SMG. *Contribución al estudio de la reacción de decomposición de la Zeolita Y em presencia de vapor de agua y vanadio*. Universidad Nacional de Colombia; 2005.
17. Espinheira PL, Silva LCM, Silva AO, Ospina R. Model selection criteria on beta regression for machine learning. *Machine Learning and Model Extraction*. 2019; 1(1):427–449. <https://doi.org/10.3390/make1010026>