




Article

New Flexible Asymmetric Log-Birnbaum–Saunders Nonlinear Regression Model with Diagnostic Analysis

Guillermo Martínez-Flórez ^{1,*}, Inmaculada Barranco-Chamorro ^{2,*} and Héctor W. Gómez ^{3,†}

¹ Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Córdoba 230027, Colombia

² Departamento de Estadística e I.O., Facultad de Matemáticas, Universidad de Sevilla, 41012 Sevilla, Spain

³ Departamento de Estadística y Ciencias de Datos, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile; hector.gomez@uantof.cl

* Correspondence: guillermomartinez@correo.unicordoba.edu.co (G.M.-F.); chamorro@us.es (I.B.-C.)

† These authors contributed equally to this work.

Abstract: A nonlinear log-Birnbaum–Saunders regression model with additive errors is introduced. It is assumed that the error term follows a flexible sinh-normal distribution, and therefore it can be used to describe a variety of asymmetric, unimodal, and bimodal situations. This is a novelty since there are few papers dealing with nonlinear models with asymmetric errors and, even more, there are few able to fit a bimodal behavior. Influence diagnostics and martingale-type residuals are proposed to assess the effect of minor perturbations on the parameter estimates, check the fitted model, and detect possible outliers. A simulation study for the Michaelis–Menten model is carried out, covering a wide range of situations for the parameters. Two real applications are included, where the use of influence diagnostics and residual analysis is illustrated.

Keywords: flexible log-Birnbaum–Saunders; flexible sinh-normal; influence diagnostics; Michaelis–Menten model; nonlinear regression

MSC: 62E15; 62J02; 62J20



Citation: Martínez-Flórez, G.; Barranco-Chamorro, I.; Gómez, H.W. New Flexible Asymmetric Log-Birnbaum–Saunders Nonlinear Regression Model with Diagnostic Analysis. *Axioms* **2024**, *13*, 576. <https://doi.org/10.3390/axioms13090576>

Academic Editor: Hans J. Haubold

Received: 20 July 2024

Revised: 17 August 2024

Accepted: 21 August 2024

Published: 23 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Regression models are one of the most common statistical techniques used to explain the relationship between a response or output continuous variable Y and a given set of explanatory covariates or predictors X_1, \dots, X_p , where p denotes the number of predictors. In the case where Y is a continuous random variable, linear models stand out as the most frequently used in practice. However, in certain situations, the relationship between the response variable and the set of predictors is nonlinear. In these cases, it is quite difficult to determine the nonlinear behavior of the response variable and the set of covariates. Nonlinear models usually arise from an underlying theory about the relationships between the variables under study. By definition, given a set of covariates X_1, \dots, X_p and a function $f(x_1, x_2, \dots, x_p; \beta_1, \beta_2, \dots, \beta_p)$, where $\beta_1, \beta_2, \dots, \beta_p$ are unknown parameters, f is linear in β_j , $j = 1, 2, \dots, p$, if and only if the first derivative of f with respect to β_j does not depend on β_j for $j = 1, 2, \dots, p$; otherwise, f is nonlinear. That is, in the nonlinear case, some of β_j 's parameters appear in f a nonlinear way.

One very general form of nonlinear regression model considers additive errors. Generally speaking, a nonlinear regression model with additive errors is defined as

$$Y_i = \psi(\beta, x_i) + \epsilon_i, \quad i = 1, 2, \dots, n, \quad (1)$$

where β is a vector of unknown parameters, x is a vector of known covariates, and $\psi(\cdot; \cdot)$, is a nonlinear, injective, and continuous function, twice differentiable with respect to the elements of β . The error terms ϵ_i are independent and identically distributed random

variables. Usually it is assumed that $\epsilon_i \sim N(0, \sigma^2)$ with unknown $\sigma^2 > 0$. However, quite often, the assumption of normality of errors does not hold. Other alternative models should be considered. Among the most relevant ones, we can cite Cancho et al. [1], who proposed a nonlinear model where the error term follows Azzalini's skew normal distribution introduced in [2], $\epsilon \sim SN(0, \sigma, \lambda)$ with skewness parameter $\lambda \in \mathbb{R}$, and scale parameter $\sigma > 0$. Another nonlinear regression model with alpha power distributed errors was proposed in Martínez-Flórez et al. [3–5]. Details about the alpha power model can be seen in Durrans [6].

We highlight that all these precedents are unimodal. However, in practice, other situations are possible. As an introductory example of interest, consider, for instance, the distribution of RNA in patients with human immunodeficiency virus (HIV) undergoing highly active antiretroviral therapy (HAART). It can be seen in Li et al. [7] that the logarithm of RNA is an asymmetric bimodal distribution, which can be adjusted by a function of a given set of covariates, x_1, x_2, \dots, x_p (for instance: level of certain biomarkers, sex, age, kind of diet, employment, etc.), and by using a nonlinear regression model with asymmetric and bimodal errors. In this context, we focus on nonlinear regression models based on the extension of the Birnbaum–Saunders (BS) distribution introduced in Martínez-Flórez et al. [8]. Our proposal may be superior for considering a mixture of distributions, since it involves fewer parameters. Influence diagnostics tools are also given. Following a traditional approach, different kinds of perturbations for the parameters in the model are introduced, which can be used to detect outliers and to assess the sensitivity of parameter estimates. Martingale-type residuals are proposed to check the fit provided by the model. All of these aim to illustrate the use of asymmetric nonlinear models to describe lifetime data.

The outline of this paper is as follows. In Section 2, the background of our proposal is established. In Section 3, the nonlinear flexible log-Birnbaum–Saunders regression model is introduced. Results from inferences based on the maximum likelihood method are given. The observed information Fisher matrix is obtained, along with its applications to this regression model. Due to the possible complexity of nonlinear models, a key point is to study the possible deficiencies in the fitted model. Therefore, Section 4 is devoted to influence diagnostics. First, the Cook generalized distance is considered. Later, local influence criteria are given to assess the effect of minor perturbations in the data and/or the proposed model on the statistical summaries. Different perturbation schemes are considered: case-weight as well as perturbation for the response variable, for explanatory variables, and for the scale parameter. Finally, martingale-type residuals are presented to detect deficiencies in the structural part of the model and detect possible outliers. A complete simulation study is given in Section 6. The two-parameter Michaelis–Menten model, widely used in kinetic chemistry, is considered. A variety of situations and error terms are covered. Two applications to real datasets can be seen in Section 7. There, a thoughtful discussion of different regression models is carried out. Moreover, the use of diagnostic criteria and residuals to improve the fitted nonlinear model is illustrated. Final conclusions are presented in Section 8. Technical details are provided in Appendices A.1–A.3.

2. Materials and Methods

Recall that the BS distribution was introduced by Birnbaum and Saunders [9] in order to model the lifetime of certain structures under dynamic load and that a random variable (RV) T follows a BS distribution, $T \sim BS(\alpha, \tau)$, if its probability density function (PDF) is

$$f_T(t) = \phi(a_t)A_t, \quad t > 0, \quad (2)$$

with $a_t = \frac{1}{\alpha} \left(\sqrt{\frac{t}{\tau}} - \sqrt{\frac{\tau}{t}} \right)$, $A_t = \frac{t^{-3/2}(t+\tau)}{2\alpha\sqrt{\tau}}$, and $\phi(\cdot)$ denoting the PDF of the $N(0, 1)$ distribution. In Equation (2), $\alpha > 0$ is a shape parameter, and $\tau > 0$ is a scale parameter and the median of the distribution.

A number of regression models related to the BS distribution can be found in the literature. We first refer to the pioneering work carried out by Rieck and Nedelman [10],

where the log-linear-Birnbaum–Saunders regression model was introduced. There, it was assumed that the output variable Y follows a sinh-normal (SHN) distribution, $SHN(\alpha, \mu, 2)$, whose PDF is

$$\varphi(y) = \frac{2}{\alpha} \frac{\cosh\left(\frac{y-\mu}{2}\right)}{2} \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\mu}{2}\right)\right), \quad y \in \mathbb{R}, \tag{3}$$

where $\alpha > 0$ is a shape parameter and $\mu \in \mathbb{R}$ is a location parameter. Given a random sample $T_i \sim BS(\alpha, \tau)$ for $i = 1, 2, \dots, n$, let $Y_i = \log(T_i)$. Rieck and Nedelman [10] proposed the linear regression model

$$Y_i = \beta' x_i + \epsilon_i, \tag{4}$$

where x_i is a covariate vector, β is a vector of unknown parameters, and the errors, ϵ_i , follow an SHN distribution, $\epsilon_i \sim SHN(\alpha, 0, 2)$, $i = 1, 2, \dots, n$. In this case, $E(Y_i) = \mu_i = x_i' \beta$, and the errors ϵ_i are symmetric with respect to zero. Asymmetric extensions were proposed by Leiva et al. [11], who considered a skewed sinh-normal model and provided applications to pollution data in Santiago, Chile. Later, Lemonte et al. [12] proposed another asymmetric extension based on the skew-normal, and Martínez-Flórez et al. [3] introduced the asymmetric extension based on the alpha power model of Durrans [6].

Another kind of asymmetric generalization of the sinh-normal model was proposed in Martínez-Flórez et al. [8,13]. This is based on the flexible skew-normal distribution introduced in Gómez et al. [14], and it is known as the flexible sinh-normal (FSHN) distribution, $Y \sim FSHN(\alpha, \xi, \sigma, \delta, \lambda)$, whose PDF is given by

$$\varphi(y) = \frac{c_\delta}{\sigma} \frac{2}{\alpha} \cosh\left(\frac{y-\xi}{\sigma}\right) \phi\left(\left|\frac{2}{\alpha} \sinh\left(\frac{y-\xi}{\sigma}\right)\right| + \delta\right) \Phi\left(\lambda \frac{2}{\alpha} \sinh\left(\frac{y-\xi}{\sigma}\right)\right), \quad y \in \mathbb{R}, \tag{5}$$

with $c_\delta = (1 - \Phi(\delta))^{-1}$ and $\Phi(\cdot)$ denoting the cumulative distribution function (CDF) of the $N(0, 1)$ distribution. In Equation (5), $\alpha \in \mathbb{R}^+$ is a shape parameter, $\xi \in \mathbb{R}$ is location, $\sigma \in \mathbb{R}^+$ is scale, $\delta \in \mathbb{R}$ is related to the bimodality, and $\lambda \in \mathbb{R}$ is skewness.

It can be seen in [8] that particular cases of interest in the FSHN model are:

- If $\delta = 0$, then the FSHN model reduces to the skewed sinh-normal distribution introduced by Leiva et al. [11].
- If $\lambda = 0$, then a symmetric model denoted as $FSHN_{\lambda=0}(\alpha, \xi, \sigma, \delta)$ is obtained, which allows us to model symmetric bimodal data; see [8].
- If $\lambda = \delta = 0$, then the FSHN model reduces to the sinh-normal distribution introduced by Rieck and Nedelman [10].

Furthermore, the following properties are proven in [8], which will be used subsequently.

Lemma 1. Let $Y \sim FSHN(\alpha, \xi, \sigma, \delta, \lambda)$. Then

1. $Y = \xi + \sigma \operatorname{arcsinh}\left(\frac{\alpha}{2} Z\right)$ with $Z \sim \text{flexible skew-normal}(\delta, \lambda)$.
2. $aY \sim FSHN(\alpha, a\xi, a\sigma, \delta, \lambda)$, $a > 0$.
3. $Y + b \sim FSHN(\alpha, \xi + b, \sigma, \delta, \lambda)$, $\forall b \in \mathbb{R}$.
4. $-Y \sim FSHN(\alpha, -\xi, \sigma, \delta, -\lambda)$.

Remark 1. Properties given in Lemma 1 allow us to obtain features of $Y \sim FSHN(\alpha, \xi, \sigma, \delta, \lambda)$ in terms of $Z \sim \text{flexible skew-normal}(\delta, \lambda)$. For instance,

1. The p -th quantile of Y , y_p , can be obtained from the p -th quantile, z_p , of Z :

$$y_p = \xi + \sigma \operatorname{arcsinh}\left(\frac{\alpha}{2} z_p\right), \quad 0 < p < 1. \tag{6}$$

2. The moments of Y can be expressed in terms of the moments of the RV $\operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right)$ with $Z \sim \text{flexible skew-normal}(\delta, \lambda)$. This fact will be explicitly introduced in the notation. So, we will write

$$\mathbb{E}_{FSN}\left(\operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right)\right) = c_\delta \int_{-\infty}^{\infty} \operatorname{arcsinh}\left(\frac{\alpha z}{2}\right) \phi(|z| + \delta) \Phi(\lambda z) dz.$$

Additional details can be seen in Martínez-Flórez et al. [8] and Gómez et al. [14].

The FSHN model was used in Martínez-Flórez et al. [8] to introduce the flexible Birnbaum–Saunders (FBS) linear regression model, which is based on the fact that, if $T \sim \text{FBS}(\alpha, \tau, \delta, \lambda)$, then

$$Y = \log(T) \sim \text{FSHN}(\alpha, \xi, 2, \delta, \lambda), \quad \text{with } \xi = \log(\tau). \tag{7}$$

So, given a random sample from Equation (7), $Y_i = \log(T_i) \sim \text{FSHN}(\alpha, \xi_i, 2, \delta, \lambda)$, covariates can be considered to explain the response variable in a natural way, that is,

$$\xi_i = \boldsymbol{\beta}^T \mathbf{x}_i, \quad \text{for } i = 1, 2, \dots, n.$$

This is called the flexible log-linear Birnbaum–Saunders regression model, whose details can be seen in [8].

In this paper, we consider a nonlinear extension of the model introduced in [8]. The interest of our proposal is based on the fact that there exist few papers dealing with nonlinear extensions of the log-BS regression model. We can cite the BS nonlinear regression model proposed by Lemonte and Cordeiro [15], the study on diagnostics and influence analysis techniques in nonlinear log-BS models with asymmetric models carried out by Lemonte [12], and the paper by Martínez-Flórez et al. [3] on the nonlinear log-BS exponentiated model.

3. Nonlinear Flexible Log-Birnbaum–Saunders

3.1. Regression Model

In this subsection, the flexible log-Birnbaum–Saunders nonlinear regression model is introduced. So, let us consider T_1, T_2, \dots, T_n independent RVs with $T_i \sim \text{FBS}(\alpha_i, \tau_i, \delta_i, \lambda_i)$. Suppose now that the distribution of T_i depends on a set of p explanatory variables, denoted by $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$, in the following way:

1. $\tau_i = \exp(f(\boldsymbol{\beta}, \mathbf{x}_i))$, $i = 1, 2, \dots, n$, where $\boldsymbol{\beta}^T = (\beta_1, \beta_2, \dots, \beta_p)$ is a p -dimensional vector of unknown parameters, and f is an injective and continuous nonlinear function, twice differentiable with respect to the elements of $\boldsymbol{\beta}$.
2. The shape parameters do not involve \mathbf{x}_i ; that is, $\alpha_i = \alpha$, $\delta_i = \delta$, $\lambda_i = \lambda$, and $i = 1, 2, \dots, n$.

Let $Y_i = \log(T_i)$. Then, the nonlinear flexible log-Birnbaum–Saunders model is defined by

$$Y_i = f(\boldsymbol{\beta}, \mathbf{x}_i) + \epsilon_i, \tag{8}$$

where it is assumed that the error terms, ϵ_i , are independent and identically distributed, $\epsilon_i \sim \text{FSHN}(\alpha, 0, 2, \delta, \lambda)$. Since $\epsilon_i \sim \text{FSHN}(\alpha, 0, 2, \delta, \lambda)$, by applying results given in [8]

$$\mathbb{E}(\epsilon_i) = 2c_1(\alpha, \delta, \lambda), \tag{9}$$

$$\text{Var}(\epsilon_i) = 4V(\alpha, \delta, \lambda), \tag{10}$$

where

$$c_1(\alpha, \delta, \lambda) = \mathbb{E}_{FSN}\left(\operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right)\right) = c_\delta \int_{-\infty}^{\infty} \operatorname{arcsinh}\left(\frac{\alpha z}{2}\right) \phi(|z| + \delta) \Phi(\lambda z) dz,$$

and $V(\alpha, \delta, \lambda)$ is the variance of the RV $\operatorname{arcsinh}(\alpha Z/2)$, with $Z \sim \text{Flexible Skew-Normal}(\delta, \lambda)$. The existence of $c_1(\alpha, \delta, \lambda)$ and $V(\alpha, \delta, \lambda)$ follow from $\operatorname{arcsinh}(x) = \log(x + \sqrt{1 + x^2})$, $\forall x \in \mathbb{R}$, where $\log(\cdot)$ denotes the Napierian logarithm and the existence of moments of the skew-normal distribution [16]. Additional details can be seen in Appendix A.1.

By applying Lemma 1 to Equation (8), we have that $Y_i \sim \text{FSHN}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), 2, \delta, \lambda)$. Moreover, $\mathbb{E}(Y_i) = f(\boldsymbol{\beta}, \mathbf{x}_i) + \mathbb{E}(\epsilon_i)$, where $\mathbb{E}(\epsilon_i)$ was given in Equation (9).

Corollary 1. *Relevant nonlinear regression models that can be obtained as particular cases of the nonlinear flexible log-Birnbaum–Saunders are the following ones:*

1. *If $\delta = 0$, then the log-BS nonlinear regression model with asymmetric errors proposed by Lemonte and Cordeiro [15] is obtained.*
2. *If $\lambda = 0$, then a nonlinear log-BS regression submodel with flexible sinh-normal errors is obtained. We highlight that this submodel may fit bimodal data for $\delta < 0$.*
3. *If $\delta = \lambda = 0$, then the nonlinear extension of the Rieck and Nedelman regression model is obtained [10].*

Proof. Note that the nonlinear Flexible log-Birnbaum–Saunders model is built on

$$Y_i \sim \text{FSHN}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), 2, \delta, \lambda).$$

From Equation (5), its PDF reduces to

$$f_Y(y_i) = \frac{c_\delta}{\alpha} \cosh(w_i) \phi\left(\left|\frac{2}{\alpha} \sinh(w_i)\right| + \delta\right) \Phi\left(\lambda \frac{2}{\alpha} \sinh(w_i)\right), \tag{11}$$

with $w_i = \frac{y_i - f(\boldsymbol{\beta}, \mathbf{x}_i)}{2}$ and $c_\delta = (1 - \Phi(\delta))^{-1}$. By considering the different values for the parameters in Equation (11), the proposed results follow. \square

Remark 2. *From now on, the following considerations must be taken into account:*

- *The general model with a scale parameter $\sigma > 0$ is considered, $Y_i \sim \text{FSHN}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), \sigma, \delta, \lambda)$, whose PDF was given in Equation (5). Also note that $\mathbb{E}(\epsilon_i) = \sigma c_1(\alpha, \delta, \lambda)$ and $\text{Var}(\epsilon_i) = \sigma^2 V(\alpha, \delta, \lambda)$.*
- *To emphasize the origin of our proposal, that is, the use of logarithm of a flexible Birnbaum–Saunders distribution as a regression model, the notation flexible log-Birnbaum–Saunders $Y_i \sim \text{FLBS}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), \sigma, \delta, \lambda)$ will be used instead of $Y_i \sim \text{FSHN}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), \sigma, \delta, \lambda)$. Both notations are equivalent.*

3.2. Inference

In this subsection, the maximum likelihood method is applied to estimate the parameters in the $Y_i \sim \text{FLBS}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), \sigma, \delta, \lambda)$ model. To simplify the exposition of results, let us denote by $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}, \sigma, \delta, \lambda)^T$

$$\zeta_{i1} = \zeta_{i1}(\boldsymbol{\theta}) = \frac{2}{\alpha} \cosh\left(\frac{y_i - f(\boldsymbol{\beta}, \mathbf{x}_i)}{\sigma}\right), \quad \zeta_{i2} = \zeta_{i2}(\boldsymbol{\theta}) = \frac{2}{\alpha} \sinh\left(\frac{y_i - f(\boldsymbol{\beta}, \mathbf{x}_i)}{\sigma}\right). \tag{12}$$

Proposition 1. *Let Y_1, \dots, Y_n be a random sample of $Y \sim \text{FLBS}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}), \sigma, \delta, \lambda)$. Then, the log-likelihood function is*

$$\ell(\boldsymbol{\theta}) = n \log(c_\delta) - n \log(\sigma) + \sum_{i=1}^n \log(\zeta_{i1}) - \frac{1}{2} \sum_{i=1}^n (\zeta_{i2}^2 + 2\delta \operatorname{sgn}(\zeta_{i2}) \zeta_{i2} + \delta^2) + \sum_{i=1}^n \log(\Phi(\lambda \zeta_{i2})), \tag{13}$$

where $c_\delta = (1 - \Phi(\delta))^{-1}$, and $\operatorname{sgn}(\cdot)$ denotes the sign function.

Proof. By using the notation introduced in Equation (12), note that

$$f_Y(y_i) = \frac{c_\delta}{\sigma} \zeta_{i1} \phi(|\zeta_{i2}| + \delta) \Phi(\lambda \zeta_{i2}). \tag{14}$$

Since we have a random sample, the log-likelihood function for $\theta = (\alpha, \beta, \sigma, \delta, \lambda)^T$ is

$$\ell(\theta) = \sum_{i=1}^n \log f_Y(y_i; \theta). \tag{15}$$

Equation (13) follows from Equations (14) and (15) and from the fact that $\phi(\cdot)$ is the PDF of an $N(0, 1)$ distribution. \square

Corollary 2. Let Y_1, \dots, Y_n be a random sample of $Y \sim FLBS(\alpha, f(\beta, x_i), \sigma, \delta, \lambda)$. Then:

1. The score functions are

$$\begin{aligned} U(\alpha) &= -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \zeta_{i2}^2 - \frac{\lambda}{\alpha} \sum_{i=1}^n \zeta_{i2} W_i + \frac{\delta}{\alpha} \sum_{i=1}^n \text{sgn}(\zeta_{i2}) \zeta_{i2}, \\ U(\beta_j) &= \frac{1}{\sigma} \sum_{i=1}^n d_{i,j} \left(\zeta_{i1} \zeta_{i2} - \frac{\zeta_{i2}}{\zeta_{i1}} \right) - \frac{\lambda}{\sigma} \sum_{i=1}^n d_{i,j} \zeta_{i1} W_i + \frac{\delta}{\sigma} \sum_{i=1}^n d_{i,j} \text{sgn}(\zeta_{i2}) \zeta_{i1}, \quad j = 1, 2, \dots, p, \\ U(\sigma) &= -\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^n z_i \left(\zeta_{i1} \zeta_{i2} - \frac{\zeta_{i2}}{\zeta_{i1}} \right) - \frac{\lambda}{\sigma} \sum_{i=1}^n z_i \zeta_{i1} W_i + \frac{\delta}{\sigma} \sum_{i=1}^n \text{sgn}(\zeta_{i2}) z_i \zeta_{i1}, \\ U(\delta) &= \frac{n\phi(\delta)}{1 - \Phi(\delta)} - \sum_{i=1}^n \text{sgn}(\zeta_{i2}) \zeta_{i2} - n\delta, \\ U(\lambda) &= \sum_{i=1}^n \zeta_{i2} W_i, \end{aligned}$$

where $\text{sgn}(\cdot)$ denotes the sign function, and

$$d_{i,j} = \frac{\partial f(\beta, x_i)}{\partial \beta_j}, \quad z_i = \frac{y_i - f(\beta, x_i)}{\sigma}, \quad W_i = \frac{\phi(\lambda \zeta_{i2})}{\Phi(\lambda \zeta_{i2})}, \quad j = 1, 2, \dots, p; \quad i = 1, \dots, n.$$

2. Maximum likelihood estimators for the regression parameter β and parameters $\alpha, \sigma, \delta,$ and λ are obtained as solutions for $U(\theta) = \mathbf{0}$, which require numerical procedures.

Proof. It is straightforward by applying standard calculus techniques. \square

Observed Information Matrix

Let us consider the matrix

$$J(\theta) = -\mathbf{H}(\theta), \tag{16}$$

where $\mathbf{H}(\theta)$ is the Hessian matrix of the log-likelihood function $l(\theta)$. The elements of $J(\theta)$ are denoted by $j_{\alpha\alpha}, j_{\beta_j\beta_j}, j_{\alpha\beta_j}, \dots, j_{\delta\delta}, j_{\delta\lambda}, j_{\lambda\lambda}$. Their expressions are given in Appendix A.2.

Recall that the Fisher (or expected) information matrix, $I_F(\theta)$, is given by the expected values of the elements in $J(\theta)$. For large samples and under regularity conditions, the MLE of $\theta, \hat{\theta}$, is asymptotically normal, and its asymptotic covariance matrix is the inverse of the Fisher information matrix, $I_F^{-1}(\theta)$. Explicitly,

$$(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} N_{p+4}(\mathbf{0}, I_F^{-1}(\theta)), \tag{17}$$

where \mathcal{L} denotes convergence in law or in distribution; see [17].

Let $J(\hat{\theta})$ be the observed information matrix, which is obtained by replacing the unknown parameters in Equation (16) by their MLEs. Since, for large n , $J(\hat{\theta})$ converges

in probability to $I_F(\theta)$, $J(\hat{\theta}) \xrightarrow{P} I_F(\theta)$, in practice, Equation (17) is applied, taking $J^{-1}(\hat{\theta})$ instead of $I_F^{-1}(\theta)$.

As for the existence of previous matrices, recall that the flexible skew-normal (FSN) and its properties are well-established. Our model is obtained by applying the $\text{arcsinh}(\cdot)$ transformation to an FSN variable. Therefore, taking into account that $\text{sinh}(\cdot)$ and $\text{cosh}(\cdot)$ functions are continuous and derivable, and given the good analytical properties of the FSN model, the existence of derivatives of log-likelihood in the nonlinear FLBS model follows. Moreover, for $\lambda = \delta = 0$, the rows (or columns) of the information matrix are linearly independent. All these facts support that the MLEs' regularity conditions are met in practice.

Moreover, for large n , and due to the convergence in probability of $J(\hat{\theta})$ to $I_F(\theta)$, the inverse of submatrix in Equation (16), which corresponds to β , $J^{-1}(\beta)$, can be used to obtain the asymptotic variance of $\hat{\mu}_i = f(\hat{\beta}, x_i)$. Specifically,

$$\text{Var}(\hat{\mu}_i) = \text{trace}(d_i d_i' J^{-1}(\hat{\beta})),$$

where $D = \frac{\partial \mu}{\partial \beta}$ is a $p \times n$ matrix with $\mu = (\mu_1, \dots, \mu_n)$, and d_i denotes the i -th column of the matrix D evaluated at $\hat{\beta}$.

4. Influence Diagnostics

In this section, influence diagnostic tools are proposed. Specifically, Cook's generalized distance is considered in Section 4.1. This is a case-deletion kind of influence diagnostic that can be used to detect influential observations on parameter estimates. Later, local influence measures are introduced in Section 4.2. Four perturbation schemes are proposed, which may be used to carry out a sensitivity study and detect influential cases affecting the obtained inferential results.

First, the notation is introduced. Following Cook [18], let $\ell(\theta)$ be the log-likelihood corresponding to the model proposed in Equation (8), with θ being the vector of unknown parameters. Perturbations into the model can be introduced through a $q \times 1$ vector ω , which is restricted to some open subset $\omega \in \Omega \subseteq \mathbb{R}^q$. In practice, ω is a given perturbation scheme of the initial model.

Let us now consider $\ell(\theta|\omega)$ to be the log-likelihood associated with the perturbed model for certain $\omega \in \Omega$. It is assumed that there exists an $\omega_0 \in \Omega$ such that $\ell(\theta|\omega_0) = \ell(\theta)$ for all θ ; that is, ω_0 represents no perturbation into the model. It is also assumed that $\ell(\theta|\omega)$ is twice continuously differentiable at (θ', ω') , where θ' and ω' denote the transpose of θ and ω . Let us denote by $\hat{\theta}$ and $\hat{\theta}_\omega$ the MLEs of the unknown parameters under $\ell(\theta)$ and $\ell(\theta|\omega)$, respectively.

The diagnostic curvature was introduced by Cook [18] as

$$C(u) = 2|u' H u|, \tag{18}$$

where u is an eigenvector of H with $\|u\| = 1$, and H is the $q \times q$ matrix with elements $h_{ii'} = \frac{\partial^2 \ell(\theta|\omega)}{\partial \omega_i \partial \omega_{i'}} |_{\hat{\theta}_\omega}$. In practice, the elements of H can be obtained from the relationship

$$H = \Delta' J^{-1}(\hat{\theta}) \Delta \tag{19}$$

evaluated at $\omega = \omega_0$, with $J(\theta)$ being the observed information matrix evaluated at $\theta = \hat{\theta}$ and $\Delta = \frac{\partial^2 \ell(\theta|\omega)}{\partial \theta \partial \omega}$ evaluated at $\theta = \hat{\theta}$ and $\omega = \omega_0$.

Based on Equation (18), the influence diagnostic analysis of maximum curvature, C_{max} , can be carried out. So, the eigenvector u_{max} associated with the largest eigenvalue of the H matrix can be used to assess the local influence on the estimates of parameters in the log-BS nonlinear model. The effect of locally influential observations is determined by the perturbation of the data in the u_{max} direction.

The previous tools can also be used to assess the local change in $\hat{\theta}$, due to the influence of ω , by using the likelihood displacement defined as $LD(\omega) = 2\{\ell(\hat{\theta}) - \ell(\hat{\theta}_\omega)\}$, which compares $\hat{\theta}_\omega$ and $\hat{\theta}$ with respect to the non-perturbed log-likelihood.

To conclude, it is worth mentioning that, in order to study influential observations, Poon and Poon [19] proposed the conformal normal curvature, B_l , defined by

$$B_l = - \frac{u' \Delta' (\ddot{L}^{-1}) \Delta u}{\sqrt{\text{tr}[u' \Delta' (\ddot{L}^{-1}) \Delta u]^2}} \Big|_{\theta = \hat{\theta}, \omega = \omega_0},$$

where $\text{tr}(A)$ is the trace of matrix A . Moreover, B_l and LD are computationally equivalent. The conformal normal curvature in the direction ω_0 is invariant under reparameterization and for any direction l , $0 \leq |B_l| \leq 1$, implying that B_l is a normalized measure, which allows us to compare two curvatures.

4.1. Cook Generalized Distance

For the vector of parameters in the log-BS nonlinear model, $\theta = (\alpha, \beta, \sigma, \delta, \lambda)'$, Cook's generalized distance (GDC) measures the global effect on MLEs of θ when an observation is removed; details can be seen in [18]. If the number of regression coefficients in the fitted model is p , then the global influential statistic for the log-FBS nonlinear regression model is given by

$$GDC_i(\theta) = \frac{1}{p + 4} \left[(\hat{\theta} - \hat{\theta}_{(i)})' \hat{\Sigma}_{\hat{\theta}}^{-1} (\hat{\theta} - \hat{\theta}_{(i)}) \right], \quad i = 1, \dots, n, \tag{20}$$

where $\hat{\Sigma}_{\hat{\theta}}$ is the estimated covariance matrix of $\hat{\theta}$, and $\hat{\theta}_{(i)}$ is the MLE of θ when the i -th observation is removed.

If we focus on the vector of regression coefficients β , then we only need the subvector that corresponds to these coefficients, and Cook's generalized distance reduces to

$$GDC_i(\beta) = \frac{1}{p} \left[(\hat{\beta} - \hat{\beta}_{(i)})' \hat{\Sigma}_{\hat{\beta}}^{-1} (\hat{\beta} - \hat{\beta}_{(i)}) \right], \quad i = 1, \dots, n,$$

where $\hat{\Sigma}_{\hat{\beta}}$ is the submatrix of $\hat{\Sigma}_{\hat{\theta}}$ associated with vector $\hat{\beta}$.

4.2. Local Influence Measurements

In this subsection, we deal with the effect of minor perturbations on the data, since they may cause a considerable effect on the estimates of parameters in the fitted model [20]. The local influence diagnostics are useful to check the model's assumptions and assess the effects of minor perturbations in the dataset or in the proposed model on estimates of regression parameters, scale parameters, and other parameters. In this way, problems with the error distribution assumptions or the fitted regression model can be detected. The following perturbation schemes are addressed:

- Case-weight;
- Perturbation of the response variable;
- Perturbation of an explanatory variable;
- Perturbation of the scale parameter.

4.2.1. Case-Weight Perturbation

Let ω be the $n \times 1$ vector of case-weights for the log-FBS nonlinear regression model introduced in Equation (8). The relevant part of the log-likelihood for the perturbed model is

$$\ell(\theta|\omega) = -\sum_{i=1}^n \omega_i (\log(1 - \Phi(\delta)) + \log(\sigma)) + \sum_{i=1}^n \omega_i \log(\xi_{i1}) \tag{21}$$

$$-\frac{1}{2} \sum_{i=1}^n \omega_i (\xi_{i2}^2 + 2\delta \operatorname{sgn}(\xi_{i2}) \xi_{i2} + \delta^2) + \sum_{i=1}^n \omega_i \log(\Phi(\lambda \xi_{i2})). \tag{22}$$

Then, for the i -th observation and the j -th coefficient β_j , it follows that

$$\Delta_{ij} = \frac{1}{\sigma} d_{i,j} \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) - \frac{\lambda}{\sigma} d_{i,j} \xi_{i1} W_i + \frac{\delta}{\sigma} d_{i,j} \operatorname{sgn}(\xi_{i2}) \xi_{i1}, \quad j = 1, 2, \dots, p.$$

Therefore, the Δ_β matrix is given by:

$$\Delta_\beta = \mathbf{D} \operatorname{diag}\{\kappa_1, \kappa_2, \dots, \kappa_n\}_{i=1,2,\dots,n} \tag{23}$$

where $\mathbf{D} = \{d_{ij}\}$ is a $p \times n$ matrix and

$$\kappa_i = \frac{1}{\sigma} \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) - \frac{\lambda}{\sigma} \xi_{i1} W_i + \frac{\delta}{\sigma} \operatorname{sgn}(\xi_{i2}) \xi_{i1}.$$

Moreover, $\Delta_{\alpha,\sigma,\delta,\lambda}$ is a $4 \times n$ matrix given by

$$\Delta_{\alpha,\sigma,\delta,\lambda} = (v_1, v_2, \dots, v_n),$$

where

$$v_i = \begin{pmatrix} -\frac{1}{\alpha} + \frac{1}{\alpha} \xi_{i2}^2 - \frac{\lambda}{\alpha} W_i \xi_{i2} + \frac{\delta}{\alpha} \operatorname{sgn}(\xi_{i2}) \xi_{i2} \\ \frac{1}{\sigma} z_i \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) - \frac{\lambda}{\sigma} z_i \xi_{i1} W_i + \frac{\delta}{\sigma} \operatorname{sgn}(\xi_{i2}) z_i \xi_{i1} \\ \frac{\Phi(\delta)}{1-\Phi(\delta)} - \operatorname{sgn}(\xi_{i2}) \xi_{i2} - \delta \\ \xi_{i2} W_i \end{pmatrix}, \quad i = 1, \dots, n.$$

4.2.2. Perturbation of the Response Variable

An additive perturbation scheme on the response variable is proposed as follows:

$$Y_{i\omega} = Y_i + \omega_i s_y, \quad i = 1, \dots, n, \tag{24}$$

where s_y is the standard deviation of the response variable, and $\omega_i \in \mathbb{R}$. So, the relevant part of the perturbed log-likelihood is given by

$$\ell(\theta|\omega) = -n \log(1 - \Phi(\delta)) - n \log(\sigma) + \sum_{i=1}^n \log(\xi_{i1\omega}) \tag{25}$$

$$-\frac{1}{2} \sum_{i=1}^n (\xi_{i2\omega}^2 + 2\delta \operatorname{sgn}(\xi_{i2\omega}) \xi_{i2\omega} + \delta^2) + \sum_{i=1}^n \log(\Phi(\lambda \xi_{i2\omega})), \tag{26}$$

where

$$\xi_{i1\omega} = \frac{2}{\alpha} \cosh\left(\frac{y_{i\omega} - f(\boldsymbol{\beta}, \mathbf{x}_i)}{\sigma}\right), \quad \xi_{i2\omega} = \frac{2}{\alpha} \sinh\left(\frac{y_{i\omega} - f(\boldsymbol{\beta}, \mathbf{x}_i)}{\sigma}\right).$$

In this case, the elements of the Δ matrix are given by

$$\begin{aligned} \Delta_{ij}(\boldsymbol{\beta}) &= \frac{s_y}{\sigma^2} d_{i,j} \left[2\bar{\zeta}_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\bar{\zeta}_{i2}^2}{\bar{\zeta}_{i2}^2 + 4/\alpha^2} \right] + \frac{\lambda s_y}{\sigma^2} d_{i,j} W_i \left[-\bar{\zeta}_{i2} + \lambda \bar{\zeta}_{i1}^2 (\lambda \bar{\zeta}_{i2} + W_i) \right] \\ &\quad + \frac{\delta s_y}{\sigma^2} d_{i,j} \text{sgn}(\bar{\zeta}_{i2}) \bar{\zeta}_{i2}, \quad i = 1, \dots, n, j = 1, \dots, p. \\ \Delta_{ij}(\alpha, \sigma, \delta, \lambda) &= (v_{1i}, v_{2i}, v_{3i}, v_{4i}), \quad i = 1, \dots, n, j = 1, \dots, 4, \end{aligned}$$

where

$$\begin{aligned} v_{1i} &= \frac{2s_y}{\alpha\sigma} \bar{\zeta}_{i1} \bar{\zeta}_{i2} + \frac{\lambda s_y}{\alpha\sigma} \bar{\zeta}_{i1} W_i [-1 + \lambda \bar{\zeta}_{i2} (\lambda \bar{\zeta}_{i2} + W_i)] + \frac{\delta s_y}{\alpha\sigma} \text{sgn}(\bar{\zeta}_{i2}) \bar{\zeta}_{i1}, \\ v_{2i} &= \frac{s_y}{\sigma^2} \left[\bar{\zeta}_{i1} \bar{\zeta}_{i2} - \frac{\bar{\zeta}_{i2}}{\bar{\zeta}_{i1}} \right] + \frac{s_y}{\sigma^2} z_i \left[2\bar{\zeta}_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\bar{\zeta}_{i2}^2}{\bar{\zeta}_{i2}^2 + 4/\alpha^2} \right] \\ &\quad - \frac{\lambda s_y}{\sigma^2} W_i \left[\bar{\zeta}_{i1} - z_i (-\bar{\zeta}_{i2} + \lambda \bar{\zeta}_{i1}^2 (\lambda \bar{\zeta}_{i2} + W_i)) \right] + \frac{\delta s_y}{\sigma^2} (\bar{\zeta}_{i1} + z_i \bar{\zeta}_{i2}), \\ v_{3i} &= -\frac{s_y}{\sigma} \text{sgn}(\bar{\zeta}_{i2}) \bar{\zeta}_{i1}, \\ v_{4i} &= -\frac{s_y}{\sigma} \bar{\zeta}_{i1} W_i (-1 + \lambda \bar{\zeta}_{i2} (\lambda \bar{\zeta}_{i2} + W_i)). \end{aligned}$$

4.2.3. Perturbation of an Explanatory Variable

Let us now consider the following additive perturbation for the explanatory variable x_q :

$$x_{iq\omega} = x_{iq} + \omega_i s_{x_q}, \quad i = 1, \dots, n, \tag{27}$$

where s_{x_q} is the standard deviation of the x_q variable, and $\omega_i \in \mathbb{R}$. Then, the relevant part of the perturbed log-likelihood is

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -n \log(1 - \Phi(\delta)) - n \log(\sigma) + \sum_{i=1}^n \log(\bar{\zeta}_{i1q\omega}) \tag{28}$$

$$- \frac{1}{2} \sum_{i=1}^n (\bar{\zeta}_{i2q\omega}^2 + 2\delta \text{sgn}(\bar{\zeta}_{i2q\omega}) \bar{\zeta}_{i2q\omega} + \delta^2) + \sum_{i=1}^n \log(\Phi(\lambda \bar{\zeta}_{i2q\omega})), \tag{29}$$

where

$$\bar{\zeta}_{i1q\omega} = \frac{2}{\alpha} \cosh\left(\frac{y_i - f(\boldsymbol{\beta}_w, \mathbf{x}_i)}{\sigma}\right), \quad \bar{\zeta}_{i2q\omega} = \frac{2}{\alpha} \sinh\left(\frac{y_i - f(\boldsymbol{\beta}_w, \mathbf{x}_i)}{\sigma}\right),$$

and $\boldsymbol{\beta}_w$ is $(x_1, x_2, \dots, x_{i(q-1)}, x_{iq} + \omega_i s_{x_q}, x_{i(q+1)}, \dots, x_p)$.

The elements of $\Delta(\boldsymbol{\beta})$ are

$$\begin{aligned} \Delta_{ij}(\boldsymbol{\beta}) &= \frac{s_{x_q}}{\sigma} r_{ijq} \left[\left(\bar{\zeta}_{i1} \bar{\zeta}_{i2} - \frac{\bar{\zeta}_{i2}}{\bar{\zeta}_{i1}} \right) - \lambda \bar{\zeta}_{i1} W_i + \delta \text{sgn}(\bar{\zeta}_{i2}) \bar{\zeta}_{i1} \right] - \frac{\delta s_{x_q}}{\sigma^2} d_{i,j} r_{iq} \text{sgn}(\bar{\zeta}_{i2}) \bar{\zeta}_{i2} \\ &\quad - \frac{s_{x_q}}{\sigma^2} d_{i,j} r_{iq} \left[\left(2\bar{\zeta}_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\bar{\zeta}_{i2}^2}{\bar{\zeta}_{i2}^2 + 4/\alpha^2} \right) + \lambda W_i (-\bar{\zeta}_{i2} + \lambda \bar{\zeta}_{i1}^2 (\lambda \bar{\zeta}_{i2} + W_i)) \right], \end{aligned}$$

with $r_{ijq} = \frac{\partial^2 f(\boldsymbol{\beta}_w, \mathbf{x})}{\partial \beta_q \partial x_{iq\omega}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \omega=0}$ and $r_{iq} = \frac{\partial f(\boldsymbol{\beta}_w, \mathbf{x})}{\partial x_{iq\omega}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \omega=0}$.

Analogously, for parameters α, σ, δ , and λ , we have that

$$\Delta(\alpha, \sigma, \delta, \lambda) = (\hat{c}_1, \hat{c}_2, \dots, \hat{c}_n)$$

where

$$\begin{aligned}
 c_{i\alpha} &= -\frac{2s_{x_q}}{\alpha\sigma}r_{iq}\xi_{i1}\xi_{i2} - \frac{\lambda s_{x_q}}{\alpha\sigma}r_{iq}\xi_{i1}W_i[-1 + \lambda\xi_{i2}(\lambda\xi_{i2} + W_i)] - \frac{\delta s_{x_q}}{\alpha\sigma}r_{iq}\text{sgn}(\xi_{i2})\xi_{i1}, \\
 c_{i\sigma} &= -\frac{s_{x_q}}{\sigma^2}r_{iq}\left[\left(\xi_{i1}\xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}}\right) - \lambda\xi_{i1}W_i + \delta\text{sgn}(\xi_{i2})\xi_{i1}\right] \\
 &\quad - \frac{s_{x_q}}{\sigma^2}r_{iq}z_i\left[\left(2\xi_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\xi_{i2}^2}{\xi_{i2}^2 + 4/\alpha^2}\right) + \lambda W_i(-\xi_{i2} + \lambda\xi_{i1}^2(\lambda\xi_{i2} + W_i)) + \delta\text{sgn}(\xi_{i2})\xi_{i2}\right], \\
 c_{i\delta} &= \frac{s_{x_q}}{\sigma}\text{sgn}(\xi_{i2})r_{iq}\xi_{i1}, \\
 c_{i\lambda} &= \frac{s_{x_q}}{\sigma}r_{iq}\xi_{i1}W_i[-1 + \lambda\xi_{i2}(\lambda\xi_{i2} + W_i)].
 \end{aligned}$$

4.2.4. Perturbation of a Scale Parameter

Now, we focus on the effect of a minor perturbation on the scale parameter σ , which may cause heteroscedasticity. In this subsection, the effect of this perturbation on the MLEs is studied. Let us assume that the error term in Equation (8) is distributed as $\epsilon_i \sim FSHN(\alpha, 0, \sigma/w_i, \delta, \lambda)$, $i = 1, 2, \dots, n$. Then, the perturbed log-likelihood is

$$\begin{aligned}
 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= -n\log(1 - \Phi(\delta)) - \sum_{i=1}^n \log(\sigma/w_i) + \sum_{i=1}^n \log(\xi_{i1\omega_1}) \\
 &\quad - \frac{1}{2} \sum_{i=1}^n (\xi_{i2\omega_1}^2 + 2\delta\text{sgn}(\xi_{i2\omega_1})\xi_{i2\omega_1} + \delta^2) + \sum_{i=1}^n \log(\Phi(\lambda\xi_{i2\omega_1})),
 \end{aligned}$$

where

$$\xi_{i1\omega_1} = \frac{2}{\alpha} \cosh(z_{i\omega_1}), \quad \xi_{i2\omega_1} = \frac{2}{\alpha} \sinh(z_{i\omega_1}), \quad z_{i\omega_1} = \frac{y_i - f(\boldsymbol{\beta}, \mathbf{x}_i)}{\sigma/w_i}.$$

In this case, we have

$$\begin{aligned}
 \Delta_{ij}(\boldsymbol{\beta}) &= \frac{1}{\sigma}d_{i,j}z_i\left[\left(2\xi_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\xi_{i2}^2}{\xi_{i2}^2 + 4/\alpha^2}\right) + \lambda W_i(-\xi_{i2} + \lambda\xi_{i1}^2(\lambda\xi_{i2} + W_i))\right] \\
 &\quad + \frac{\delta}{\sigma}\text{sgn}(\xi_{i2})d_{i,j}z_i\xi_{i2},
 \end{aligned}$$

and $\Delta_i(\alpha, \sigma, \delta, \lambda) = (\hat{r}_1, \hat{r}_2, \dots, \hat{r}_n)$ with

$$\begin{aligned}
 r_{i\alpha} &= \frac{2}{\alpha}z_i\xi_{i1}\xi_{i2} + \frac{\lambda}{\alpha}z_i\xi_{i1}W_i[-1 + \lambda\xi_{i2}(\lambda\xi_{i2} + W_i)] + \frac{\delta}{\alpha}z_i\text{sgn}(\xi_{i2})\xi_{i1}, \\
 r_{i\sigma} &= \frac{1}{\sigma}z_i\left[\left(\xi_{i1}\xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}}\right) - \lambda\xi_{i1}W_i + \delta\text{sgn}(\xi_{i2})\xi_{i1}\right] \\
 &\quad + \frac{1}{\sigma}z_i^2\left[\left(2\xi_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\xi_{i2}^2}{\xi_{i2}^2 + 4/\alpha^2}\right) + \lambda W_i(-\xi_{i2} + \lambda\xi_{i1}^2(\lambda\xi_{i2} + W_i)) + \delta\text{sgn}(\xi_{i2})\xi_{i2}\right], \\
 r_{i\delta} &= -\text{sgn}(\xi_{i2})z_i\xi_{i1}, \\
 r_{i\lambda} &= -z_i\xi_{i1}W_i[-1 + \lambda\xi_{i2}(\lambda\xi_{i2} + W_i)].
 \end{aligned}$$

5. Residual Analysis

In this section, martingale-type residuals are considered to detect deficiencies in the fitted FBS nonlinear regression model with respect to the error distributional assumptions and to detect possible outliers. The study is based on the deviance component residual built on the martingale-type residuals proposed in Therneau et al. [21]. For the nonlinear log-FBS model, the martingale residuals can be obtained as

$$r_{M_i} = 1 + \ln(1 - \{F_{FSHN}(\hat{z}_i)\}), \quad i = 1, 2, \dots, n,$$

where $F_{FSHN}(\cdot)$ is the CDF of the FSHN distribution, and $\hat{z}_i = \zeta_{i2}(\hat{\theta})$.

Therneau et al. [21] proposed the deviance component residual as a transformation of the martingale-type residual. For non-censored data, they can be taken as

$$r_{MT_i} = \text{sgn}(r_{M_i}) \{-2[r_{M_i} + \ln(1 - r_{M_i})]\}^{1/2}, \quad i = 1, 2, \dots, n.$$

Here, r_{MT_i} can be used as martingale-type residuals since they are symmetrically distributed around zero.

On the other hand, Ortega et al. [22] proposed the consideration of the standardized residuals, $r_{MT_i}^*$. For the log-BS nonlinear model, they are given by

$$r_{MT_i}^* = \frac{r_{MT_i}}{\sqrt{1 - GL_{ii}(\hat{\theta})}}, \quad i = 1, \dots, n,$$

where $GL_{ii}(\hat{\theta})$ is the i -th principal component for the generalized leverage matrix evaluated at $\hat{\theta}$; details can be seen in Wei et al. [23].

The generalized leverage matrix is defined by

$$GL(\theta) = D_{\theta}(-\ddot{L})^{-1}\ddot{L}_{\theta y},$$

where $D_{\theta} = \frac{\partial f(\theta)}{\partial \theta'} = (D, \mathbf{0})$, \ddot{L} is a Hessian matrix, and $\ddot{L}_{\theta y} = (\ddot{L}_{\alpha y}, \ddot{L}_{\beta y}, \ddot{L}_{\sigma y}, \ddot{L}_{\delta y}, \ddot{L}_{\lambda y})'$, with

$$\ddot{L}_{\alpha y} = \frac{2}{\alpha\sigma} \zeta_{i1}\zeta_{i2} + \frac{\lambda}{\alpha\sigma} \zeta_{i1} W_i [-1 + \lambda\zeta_{i2}(\lambda\zeta_{i2} + W_i)] + \frac{\delta}{\alpha\sigma} \text{sgn}(\zeta_{i2})\zeta_{i1},$$

$$\begin{aligned} \ddot{L}_{\beta y} &= \frac{1}{\sigma^2} d_{i,j} \left[\left(2\zeta_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\zeta_{i2}^2}{\zeta_{i2}^2 + 4/\alpha^2} \right) + \lambda W_i (-\zeta_{i2} + \lambda\zeta_{i1}^2(\lambda\zeta_{i2} + W_i)) \right] \\ &\quad + \frac{\delta}{\sigma^2} \text{sgn}(\zeta_{i2}) d_{i,j} \zeta_{i2}, \end{aligned}$$

$$\begin{aligned} \ddot{L}_{\sigma y} &= \frac{1}{\sigma^2} \left[\left(\zeta_{i1}\zeta_{i2} - \frac{\zeta_{i2}}{\zeta_{i1}} \right) - \lambda\zeta_{i1} W_i + \delta \text{sgn}(\zeta_{i2})\zeta_{i1} \right] \\ &\quad + \frac{1}{\sigma^2} z_i \left[\left(2\zeta_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\zeta_{i2}^2}{\zeta_{i2}^2 + 4/\alpha^2} \right) + \lambda W_i (-\zeta_{i2} + \lambda\zeta_{i1}^2(\lambda\zeta_{i2} + W_i)) + \delta \text{sgn}(\zeta_{i2})\zeta_{i2} \right], \end{aligned}$$

$$\ddot{L}_{\delta y} = -\frac{1}{\sigma} \text{sgn}(\zeta_{i2})\zeta_{i1},$$

and

$$\ddot{L}_{\lambda y} = -\frac{1}{\sigma} \zeta_{i1} W_i [-1 + \lambda\zeta_{i2}(\lambda\zeta_{i2} + W_i)].$$

In general, the distributions of martingale residuals and deviance component residuals are unknown. Based on the properties of deviance component residuals and the suggestions proposed in Atkinson [24], the residual analysis must be based on envelopes with normality plots. We will follow this recommendation in the practical applications carried out in Section 7.

6. Simulation

In order to study the performance of MLEs in the log-FBS nonlinear regression model, two simulation studies have been carried out. Both suggest the good performance of our proposal.

6.1. Simulation for the Two-Parameter Michaelis–Menten Model

Next, the two-parameter Michaelis–Menten model is introduced. This model is widely used in kinetic chemistry to describe the concentration of a substrate (Y) in terms of the reaction rate (X). In practice, it is applied to enzyme-catalyzed reactions of one substance and is used in a variety of situations, such as antigen–antibody binding, DNA–DNA hybridization, and protein–protein interactions.

The Michaelis–Menten model with log-FBS errors is given by

$$y_i = \frac{\beta_2 x_i}{\beta_1 + x_i} + \epsilon_i, \quad i = 1, \dots, n, \quad (30)$$

where the reaction rate $x > 0$, $\beta_2 > 0$ represents the maximum rate of reaction, $\beta_1 > 0$ is a kinetic constant, and $\epsilon \sim FSHN(\alpha, 0, \sigma, \delta, \lambda)$.

Properties of Equation (30) were studied by Cysneiros and Vanegas [25] by assuming that $\beta_1 = 0.0645$, $\beta_2 = 212$, and the error term ϵ follows a Student's t distribution. In our simulation, we consider Equation (30) with these values of β_1 and β_2 , but the error term is FSHN distributed as $\epsilon \sim FSHN(\alpha, 0, \sigma, \delta, \lambda)$.

Without loss of generality, $\sigma = 2$ is fixed. For the shape parameters in the FSHN model, we take $\alpha = 0.75, 2.75$, $\lambda = 1, 2.5$, and $\delta = -1.5, -0.75, 0.75, 1.5$. We highlight that these scenarios cover a variety of situations for the shapes of the FSHN distribution (unimodal and bimodal), as can be seen in [8].

As for the sample size, for every scenario, we consider $n = 30, 50, 100, 200$. For the explanatory variable X , a random sample of a $U(0, 1)$ distribution was generated. In every setting, 2000 simulations were carried out by using the `maxLik` function of the R software and by applying the BFGS method. As statistical summaries, the standard deviation (sd) of MLEs, the absolute value of bias ($|bias|$), and the root of the mean squared error (\sqrt{MSE}) are given in Tables 1–4. In these tables, we can observe that, when the sample size increases, the standard deviations, biases, and \sqrt{MSE} decrease in all scenarios. It can also be determined that the bias and \sqrt{MSE} of $\hat{\beta}_1$ and $\hat{\beta}_2$ are negligible, which confirms the asymptotic unbiasedness of these estimators in the Michaelis–Menten model.

Also note that we obtained satisfactory results in our estimates for $\delta < 0$, suggesting the good performance of our model in bimodal settings.

For $(\alpha, \sigma, \delta, \lambda)$ parameters, we highlight that the greater bias was obtained for $\delta > 0$ and $n = 50$. However, if n increases, then $|bias|$ decreases. In all settings, the best results are obtained for negative values of δ , even for $n = 30$.

Since these simulations cover a variety of situations and shapes of the FSHN distribution, they suggest that the estimators of the parameters are consistent when the sample size increases.

As for the interest of our results in practice, we highlight the importance of the Michaelis–Menten model. Recall that, in kinetic chemistry, this model is applied in the classical case of an enzyme substrate mechanism in which the reaction timescale of the enzyme must be faster than the reaction timescale of the substrate.

Table 1. $|Bias|$, sd , and \sqrt{MSE} for the MLEs in the nonlinear $FSHN(\alpha, \beta_1 = 0.0645, \beta_2 = 212, \sigma = 2, \delta = -1.5, \lambda)$ regression model.

α	λ	θ_j	$n = 30$			$n = 50$			$n = 100$			$n = 200$		
			$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}
0.75	1	α	1.0008	2.2462	2.4581	0.5139	1.4779	1.5642	0.3016	1.06	1.1018	0.154	0.6548	0.6725
		β_1	0.0000	0.0005	0.0005	0.0000	0.0004	0.0004	0	0.0003	0.0003	0	0.0002	0.0002
		β_2	0.0820	0.3662	0.3751	0.0301	0.2839	0.2854	0.0076	0.1980	0.1980	0.009	0.1447	0.1450
		σ	0.1063	1.5869	1.5898	0.2153	1.4548	1.4701	0.3131	1.3863	1.4208	0.2194	1.0837	1.1054
		δ	0.0437	0.7784	0.7793	0.0239	0.5766	0.5769	0.0132	0.4391	0.4392	0.0076	0.3088	0.3088
		λ	0.1901	0.9566	0.9749	0.1634	0.7098	0.7281	0.1330	0.5332	0.5494	0.0749	0.3549	0.3626
	2.5	α	0.8865	1.8557	2.0552	0.4839	1.3727	1.4549	0.2467	0.8404	0.8756	0.1551	0.7013	0.7180
		β_1	0.0001	0.0005	0.0005	0.0000	0.0004	0.0004	0.0000	0.0003	0.0003	0.0000	0.0002	0.0002
		β_2	0.1426	0.3531	0.3805	0.0594	0.2775	0.2837	0.0137	0.2094	0.2098	0.0036	0.1827	0.1827
		σ	0.1177	1.6055	1.6085	0.2733	1.5332	1.5567	0.2657	1.3060	1.3324	0.2249	1.2441	1.2643
		δ	0.3248	0.8475	0.9069	0.1141	0.6391	0.6489	0.0497	0.4878	0.4902	0.0207	0.4241	0.4245
		λ	0.5084	2.1027	2.1632	0.4425	1.5452	1.6066	0.3441	1.1168	1.1682	0.2063	0.8370	0.8618
2.75	1	α	0.3120	2.4598	2.4784	0.1955	2.2903	2.2978	0.2532	2.1052	2.1203	0.0390	1.8850	1.8854
		β_1	0.0001	0.0009	0.0009	0.0001	0.0006	0.0006	0	0.0004	0.0004	0.0000	0.0003	0.0003
		β_2	0.2337	0.5774	0.6227	0.1574	0.4361	0.4635	0.0641	0.3065	0.3130	0.0296	0.2195	0.2214
		σ	0.7659	1.4792	1.6652	0.5457	1.1195	1.245	0.2845	0.7086	0.7634	0.1109	0.4229	0.4371
		δ	0.4424	0.7825	0.8986	0.3415	0.6329	0.7189	0.1809	0.4946	0.5265	0.0636	0.4025	0.4074
		λ	0.1496	0.5012	0.5229	0.1098	0.4537	0.4667	0.0327	0.4008	0.4020	0.0201	0.3471	0.3476
	2.5	α	0.6689	2.5377	2.6244	0.5444	2.0995	2.1679	0.3647	1.7431	1.7802	0.0185	1.5518	1.5514
		β_1	0.0004	0.0009	0.0010	0.0003	0.0007	0.0007	0.0001	0.0005	0.0005	0	0.0004	0.0004
		β_2	0.5897	0.7349	0.9419	0.4189	0.5920	0.725	0.2282	0.5054	0.5544	0.0845	0.3681	0.3776
		σ	0.8094	1.6277	1.8169	0.6858	1.3942	1.5531	0.3438	0.7213	0.7988	0.1377	0.4321	0.4534
		δ	0.2255	0.8603	0.8893	0.1995	0.7052	0.7325	0.1525	0.5421	0.5629	0.0648	0.4146	0.4195
		λ	0.6643	1.5532	1.6884	0.5497	1.3291	1.4377	0.2237	1.2552	1.2745	0.0556	1.1871	1.1881

Table 2. $|Bias|$, sd , and \sqrt{MSE} for the MLEs in the nonlinear $FSHN(\alpha, \beta_1 = 0.0645, \beta_2 = 212, \sigma = 2, \delta = -0.75, \lambda)$ regression model.

α	λ	θ_j	$n = 30$			$n = 50$			$n = 100$			$n = 200$		
			$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}
0.75	1	α	1.2619	2.5259	2.8222	0.6735	1.7690	1.8921	0.2822	1.1092	1.1441	0.1241	0.6143	0.6265
		β_1	0.0000	0.0006	0.0006	0.0000	0.0004	0.0004	0.0000	0.0003	0.0003	0.0000	0.0002	0.0002
		β_2	0.0609	0.3885	0.3930	0.0282	0.3460	0.3470	0.0338	0.2469	0.2491	0.0268	0.1948	0.1966
		σ	0.7200	1.7957	1.9347	0.4471	1.6701	1.7289	0.3592	1.4841	1.5264	0.3280	1.2659	1.3073
		δ	0.1714	0.8152	0.8325	0.1928	0.6807	0.7071	0.1197	0.4928	0.5070	0.0694	0.3554	0.3620
	λ	0.3075	0.9914	1.0379	0.1689	0.9837	0.9976	0.1980	0.9147	0.9356	0.1046	0.5340	0.5440	
	2.5	α	0.8614	1.8335	2.0242	0.5484	1.4235	1.5246	0.3236	1.0391	1.0879	0.1646	0.7811	0.7980
		β_1	0.0000	0.0005	0.0005	0.0000	0.0004	0.0004	0.0000	0.0002	0.0002	0.0000	0.0002	0.0002
		β_2	0.0394	0.2951	0.2975	0.0249	0.2172	0.2186	0.0185	0.1739	0.1748	0.0035	0.1399	0.1399
		σ	0.5831	1.7424	1.8374	0.3804	1.5203	1.5672	0.2747	1.3811	1.4076	0.1787	1.3684	1.3800
δ		0.0222	0.7953	0.7949	0.0316	0.7477	0.7483	0.0135	0.6622	0.6621	0.0090	0.5516	0.5516	
2.75	1	λ	0.5665	2.1875	2.2577	0.7411	2.0010	2.1326	0.6756	1.6576	1.7893	0.4951	1.2955	1.3865
		α	0.5663	3.1198	3.1694	0.4971	2.8185	2.8620	0.5086	2.5052	2.5556	0.3869	2.2428	2.2759
		β_1	0.0000	0.0010	0.0010	0.0000	0.0008	0.0008	0.0000	0.0005	0.0005	0.0000	0.0004	0.0004
		β_2	0.2157	0.6739	0.7073	0.1263	0.5217	0.5366	0.0645	0.3705	0.3759	0.0317	0.2634	0.2653
		σ	0.6392	1.6206	1.7414	0.5644	1.3796	1.4901	0.2792	0.9033	0.9452	0.1480	0.5875	0.6057
	2.5	δ	0.3471	0.8917	0.9565	0.2720	0.7695	0.8159	0.1118	0.6381	0.6477	0.0359	0.5442	0.5452
		λ	0.1501	0.6607	0.6773	0.1247	0.6139	0.6264	0.0853	0.5434	0.5498	0.0576	0.4816	0.4849
		α	0.8108	2.7504	2.8674	0.5540	2.1766	2.2450	0.4585	1.9212	1.9745	0.3650	1.5705	1.6119
		β_1	0.0002	0.0010	0.0010	0.0001	0.0008	0.0008	0.0000	0.0006	0.0006	0.0000	0.0004	0.0004
		β_2	0.3392	0.7320	0.8064	0.2300	0.6258	0.6665	0.1061	0.5086	0.5193	0.0267	0.4458	0.4465
2.5	σ	0.8629	1.7817	1.9787	0.8173	1.3572	1.5837	0.6250	1.1474	1.3062	0.3488	0.6221	0.7130	
	δ	0.3800	0.8719	0.9506	0.3974	0.7350	0.8352	0.3090	0.6264	0.6982	0.2297	0.5059	0.5554	
	λ	0.6113	1.7079	1.8130	0.5432	1.5239	1.6171	0.2708	1.4400	1.4647	0.0918	1.2792	1.2821	

Table 3. $|Bias|$, sd , and \sqrt{MSE} for the MLEs of parameters in the nonlinear $FSHN(\alpha, \beta_1 = 0.0645, \beta_2 = 212, \sigma = 2, \delta = 0.75, \lambda)$ regression model.

α	λ	θ_j	$n = 30$			$n = 50$			$n = 100$			$n = 200$		
			$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}
0.75	1	α	2.1108	2.9608	3.6309	1.4057	2.3245	2.7123	0.9875	1.9073	2.1446	0.8043	1.7183	1.8944
		β_1	0.0001	0.0005	0.0005	0.0001	0.0004	0.0004	0.0001	0.0003	0.0003	0.0001	0.0002	0.0002
		β_2	0.2369	0.3454	0.4188	0.1023	0.3046	0.3207	0.0836	0.2342	0.2487	0.0744	0.1890	0.2028
		σ	0.4017	1.7876	1.8284	0.1440	1.7734	1.7755	0.1072	1.7118	1.7120	0.1068	1.5550	1.5587
		δ	0.6162	1.7212	1.8246	0.1996	2.0235	2.0291	0.1920	1.8514	1.8402	0.1982	1.5857	1.5951
	λ	1.8053	2.8485	3.3724	1.4874	2.4124	2.8341	1.3146	1.7855	2.2146	0.6068	1.0620	1.2214	
	2.5	α	1.2944	2.2954	2.6307	0.7644	1.6713	1.8349	0.5569	1.2543	1.3723	0.4382	1.3193	1.3885
		β_1	0.0000	0.0004	0.0004	0.0001	0.0003	0.0003	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001
		β_2	0.0501	0.2343	0.2391	0.0761	0.1870	0.2016	0.0773	0.1352	0.1557	0.0763	0.1013	0.1267
		σ	0.3180	1.7369	1.7619	0.3031	1.9523	1.9756	0.2879	1.6463	1.6688	0.2814	1.6088	1.6310
δ		0.6808	1.4775	1.6238	0.5812	1.4641	1.5726	0.5628	1.2540	1.3728	0.2694	1.2843	1.3105	
2.75	1	λ	1.9729	2.6751	3.3239	1.5213	2.5374	2.9543	1.3813	1.9897	2.4221	1.2252	1.6178	2.0276
		α	2.3567	2.8848	3.7181	0.8979	2.6963	2.8419	0.6777	2.0076	2.1169	0.2569	1.5997	1.6201
		β_1	0.0000	0.0004	0.0004	0.0001	0.0010	0.0010	0.0001	0.0007	0.0007	0.0000	0.0006	0.0006
		β_2	0.4284	0.9595	1.0508	0.2046	0.7624	0.7893	0.1121	0.6519	0.6608	0.0254	0.5195	0.5196
		σ	1.6940	1.6090	2.3363	1.5132	1.6300	2.2241	1.0485	1.5881	1.9017	0.9707	1.3685	1.6767
	2.5	δ	1.5242	2.2937	2.7539	1.0093	1.0623	1.4646	0.9449	1.0496	1.4114	0.8077	0.8722	1.1882
		λ	0.2632	1.6498	1.6656	0.1691	1.0744	1.0866	0.1070	1.0530	1.0584	0.0810	0.9354	0.9380
		α	1.6058	2.8996	3.3145	1.4957	2.3819	2.8126	1.0571	1.7376	2.0322	0.9787	1.6344	1.9038
		β_1	0.0001	0.0010	0.0010	0.0001	0.0008	0.0008	0.0002	0.0005	0.0006	0.0002	0.0004	0.0005
		β_2	0.7109	0.9153	1.1589	0.3315	0.6603	0.7388	0.3033	0.5026	0.5865	0.2466	0.4437	0.5073
2.5	σ	1.5889	1.6265	2.2738	1.4450	1.7168	2.2440	1.4456	1.6668	2.2050	1.2058	1.4130	1.8566	
	δ	1.3270	1.8647	2.2887	1.0963	1.3638	1.7482	1.1484	0.9965	1.5198	1.0000	0.9666	1.3903	
	λ	0.6418	2.8511	2.9224	0.4011	2.0858	2.1218	0.4122	2.0135	2.0552	0.3777	1.8670	1.9033	

Table 4. $|Bias|$, sd , and \sqrt{MSE} for the MLEs of parameters in the nonlinear $FSHN(\alpha, \beta_1 = 0.0645, \beta_2 = 212, \sigma = 2, \delta = 1.5, \lambda)$ regression model.

α	λ	θ_j	$n = 30$			$n = 50$			$n = 100$			$n = 200$		
			$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}
0.75	1	α	1.3321	3.1923	3.4564	1.2461	2.0285	2.3745	1.0868	1.7757	2.0769	1.0086	1.5979	1.8858
		β_1	0.0001	0.0013	0.0013	0.0000	0.0003	0.0003	0.0001	0.0002	0.0002	0.0000	0.0001	0.0001
		β_2	0.2190	0.8386	0.8667	0.1585	0.2494	0.2947	0.0721	0.1668	0.1813	0.0231	0.1041	0.1064
		σ	0.1885	1.6809	1.6914	0.0990	2.2292	2.2314	0.0229	1.8572	1.8512	0.0144	1.6128	1.6084
		δ	0.9510	2.2002	2.3969	0.4849	2.1557	2.2020	0.2764	2.0100	2.0223	0.1950	1.6797	1.6909
	λ	1.2860	2.6139	2.9131	1.2695	2.3624	2.6744	0.9981	1.5862	1.8697	0.5423	0.9422	1.0849	
	2.5	α	0.8852	2.9406	3.0709	0.7970	1.6712	1.8464	0.6649	1.3687	1.5184	0.5923	1.2810	1.4088
		β_1	0.0001	0.0012	0.0012	0.0001	0.0002	0.0002	0.0000	0.0002	0.0002	0.0000	0.0001	0.0001
		β_2	0.0180	0.7550	0.7541	0.0798	0.1517	0.1709	0.0809	0.1235	0.1474	0.0513	0.0936	0.1065
		σ	0.6500	1.9598	2.0620	0.1369	2.0962	2.1006	0.1155	1.4954	1.4960	0.0726	1.6704	1.6719
δ		1.1244	1.9715	2.2696	0.8033	1.8345	1.9970	0.4552	1.7969	1.8492	0.0588	1.7381	1.7353	
2.75	1	λ	1.6103	2.8710	3.2917	1.5760	2.5588	2.9978	1.7609	2.1633	2.7850	1.1074	1.6314	1.9688
		α	1.6704	3.0504	3.4729	0.2643	2.7861	2.7921	0.2461	2.4365	2.4427	0.0770	2.5228	2.5187
		β_1	0.0001	0.0013	0.0013	0.0002	0.0010	0.0010	0.0003	0.0007	0.0007	0.0001	0.0005	0.0005
		β_2	0.6137	0.8675	1.0626	0.2631	0.8215	0.8608	0.2425	0.6921	0.7317	0.1712	0.4836	0.5121
		σ	1.8340	1.9599	2.6842	1.6571	1.7980	2.4452	1.0247	1.6313	1.9229	0.9110	1.6398	1.8729
	2.5	δ	1.7730	1.8202	2.5410	1.3264	1.7092	2.1603	0.9127	1.8085	2.0217	0.3544	1.7783	1.8096
		λ	0.8348	1.9410	2.1129	0.5945	1.6926	1.7940	0.4609	1.6122	1.6729	0.3565	1.1140	1.1675
		α	1.1420	3.0652	3.2640	0.7401	2.5240	2.6303	0.6398	2.2713	2.3559	0.6194	2.1003	2.1866
		β_1	0.0001	0.0011	0.0011	0.0001	0.0007	0.0007	0.0003	0.0005	0.0006	0.0002	0.0004	0.0004
		β_2	0.5882	0.6913	0.9077	0.3236	0.5841	0.6677	0.2770	0.4772	0.5511	0.2510	0.3822	0.4568
2.5	σ	0.2023	1.8242	1.8309	1.9826	1.9431	2.7760	1.3813	1.8272	2.2880	1.2014	1.5971	1.9965	
	δ	1.2485	1.8292	2.2146	1.0625	1.8562	2.1350	1.1734	1.7224	2.0817	0.9119	1.6783	1.9077	
	λ	0.9592	2.9735	3.1244	0.8472	2.7257	2.8483	0.8469	2.3657	2.5089	0.6073	1.8613	1.9551	

6.2. Simulation for a Nonlinear Regression Model Proposed in [15]

A second simulation study is carried out for the nonlinear regression model with $p = 3$ covariates, 4 regression coefficients, and sinh-normal errors discussed in Lemonte and Cordeiro [15]. The model is

$$\mu_i = \beta_1 x_{i,1} + \beta_2 x_{i,2} + \beta_3 \exp(\beta_4 x_{i,3}), \quad i = 1, \dots, n.$$

Next, the MLE properties of parameters in this model are studied under the assumption of FSHN errors.

As statistical summaries, again we give: the standard error (sd) of estimates, $|bias|$, and \sqrt{MSE} .

The scenarios under consideration are: $\beta_1 = 4$, $\beta_2 = 5$, $\beta_3 = 3$, $\beta_4 = 1.5$, and $\sigma = 3$. As for the parameters for the error term, we considered $\alpha = 0.75, 1.5$, $\lambda = 1, 2.5$, and $\delta = -2.5, -0.75$. The sample sizes in our simulations were $n = 50, 100, 150$. The covariates X_1, X_2, X_3 were random variables generated as uniform $(0, 1)$. In every setting, the simulations were repeated 2000 times. The maxLik function of the R software and the BFGS method were applied.

Results are listed in Tables 5 and 6. In general, we can see that the standard errors, the absolute bias, and the square root of the mean squared error decrease for all the parameters if the sample size increases.

We highlight that the estimates of parameters $\beta_1, \beta_2, \beta_3$, and β_4 , that is, for those involved in the regression model, behave well, which suggests their asymptotic unbiasedness and consistency.

Table 5. $|Bias|$, empirical sd, and \sqrt{MSE} for the MLEs of the parameters in the nonlinear $FSHN(\alpha, 4, 5, 3, 1.5, 3, -0.75, \lambda)$ regression model.

α	λ	θ_j	$n = 50$			$n = 100$			$n = 150$		
			$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}
0.75	1	α	1.3278	2.6594	2.9712	0.4640	1.4611	1.5324	0.2054	0.9908	1.0115
		β_1	0.0127	0.5236	0.5235	0.0061	0.3558	0.3558	0.0041	0.2850	0.2850
		β_2	0.0087	0.5465	0.5463	0.0049	0.3624	0.3624	0.0038	0.2789	0.2789
		β_3	0.0692	0.4888	0.4935	0.0136	0.3781	0.3783	0.0008	0.3075	0.3077
		β_4	0.0058	0.1499	0.1499	0.0069	0.1145	0.1146	0.0083	0.0909	0.0913
		δ	0.1084	1.0004	1.0062	0.0967	0.6975	0.7042	0.0882	0.4929	0.5005
		λ	0.2565	0.9038	0.9395	0.1744	0.9060	0.9223	0.1045	0.6923	0.6999
	2.5	α	0.6796	1.8382	1.9585	0.3686	1.2026	1.2573	0.2519	1.0130	1.0435
		β_1	0.0120	0.4337	0.4336	0.0129	0.2889	0.2891	0.0039	0.2368	0.2367
		β_2	0.0065	0.4461	0.4458	0.0050	0.2952	0.2951	0.0010	0.2213	0.2213
		β_3	0.0439	0.3829	0.3851	0.0037	0.2651	0.2650	0.0011	0.2315	0.2317
		β_4	0.0050	0.1142	0.1142	0.0044	0.0800	0.0801	0.0030	0.0681	0.0681
		δ	0.0846	0.8939	0.8978	0.0409	0.7071	0.7080	0.0294	0.6409	0.6413
		λ	0.4954	2.0678	2.1248	0.6857	1.7562	1.8845	0.5697	1.5043	1.6080
1.5	1	α	1.1758	2.9068	3.1345	0.4951	2.0550	2.1132	0.3055	1.6689	1.6961
		β_1	0.0069	0.8074	0.8071	0.0060	0.5288	0.5288	0.0038	0.4244	0.4244
		β_2	0.0337	0.7960	0.7967	0.0264	0.5343	0.5350	0.0087	0.4157	0.4156
		β_3	0.1620	0.6968	0.7151	0.0589	0.5360	0.5390	0.0381	0.4312	0.4328
		β_4	0.0181	0.2146	0.2153	0.0012	0.1636	0.1636	0.0000	0.1337	0.1337
		δ	0.1501	0.9077	0.9197	0.1218	0.6588	0.6698	0.0895	0.5521	0.5591
		λ	0.0617	0.6945	0.6970	0.0372	0.6878	0.6886	0.0232	0.6023	0.6025
	2.5	α	0.6311	2.4365	2.5155	0.0921	1.5692	1.5713	0.0219	1.2922	1.2920
		β_1	0.0192	0.6908	0.6911	0.0088	0.4680	0.4681	0.0033	0.3818	0.3816
		β_2	0.0130	0.7076	0.7077	0.0049	0.4630	0.4630	0.0024	0.3851	0.3849
		β_3	0.1391	0.6510	0.6653	0.0349	0.4880	0.4890	0.0208	0.4139	0.4143
		β_4	0.0107	0.1873	0.1875	0.0040	0.1445	0.1445	0.0029	0.1201	0.1201
		δ	0.1416	0.9404	0.9510	0.1367	0.6979	0.7109	0.1350	0.6132	0.6279
		λ	0.2127	2.1304	2.1409	0.1602	1.6150	1.6222	0.1497	1.5055	1.5124

Table 6. $|Bias|$, empirical sd, and \sqrt{MSE} for the MLEs of the parameters in the nonlinear $FSHN(\alpha, 4, 5, 3, 1.5, 3, -2.5, \lambda)$ regression model.

α	λ	θ_j	$n = 50$			$n = 100$			$n = 150$		
			$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}	$ bias $	sd	\sqrt{MSE}
0.75	1	α	0.3506	1.2033	1.2529	0.1560	0.7021	0.7190	0.1113	0.5457	0.5568
		β_1	0.0059	0.4137	0.4136	0.0033	0.2683	0.2684	0.0011	0.2221	0.2220
		β_2	0.0085	0.4150	0.4150	0.0062	0.2770	0.2770	0.0036	0.2216	0.2215
		β_3	0.0896	0.4543	0.4629	0.0356	0.2860	0.2881	0.0088	0.2373	0.2374
		β_4	0.0113	0.1354	0.1358	0.0050	0.0824	0.0825	0.0004	0.0695	0.0695
		δ	0.0307	0.6869	0.6874	0.0260	0.4574	0.4580	0.0204	0.3664	0.3670
		λ	0.1476	0.7338	0.7482	0.1447	0.6229	0.6394	0.0954	0.3625	0.3748
	2.5	α	0.9213	1.8964	2.1068	0.4460	1.0851	1.1727	0.2751	0.8086	0.8539
		β_1	0.0146	0.4238	0.4237	0.0336	0.2802	0.2820	0.0007	0.2264	0.2263
		β_2	0.0225	0.4202	0.4204	0.0165	0.2822	0.2825	0.0045	0.2319	0.2319
		β_3	0.3120	0.6527	0.7229	0.2020	0.4219	0.4676	0.1725	0.3376	0.3790
		β_4	0.0538	0.1903	0.1976	0.0413	0.1136	0.1209	0.0381	0.0894	0.0972
		δ	0.5141	1.0195	1.1409	0.3482	0.7174	0.7971	0.2598	0.5429	0.6016
		λ	0.7986	1.8937	2.0552	0.3364	1.5904	1.6248	0.1472	1.4668	1.4742
1.5	1	α	0.4957	1.5216	1.6003	0.2721	1.3524	1.3791	0.2153	1.1273	1.1474
		β_1	0.0243	0.5179	0.5183	0.0003	0.3448	0.3447	0.0001	0.2774	0.2774
		β_2	0.0085	0.4935	0.4934	0.0064	0.3465	0.3465	0.0031	0.2753	0.2752
		β_3	0.2468	0.5959	0.6447	0.1023	0.4279	0.4398	0.0441	0.3325	0.3354
		β_4	0.0450	0.1725	0.1782	0.0178	0.1209	0.1221	0.0057	0.0943	0.0945
		δ	0.2537	0.7339	0.7762	0.0787	0.5017	0.5077	0.0582	0.4081	0.4122
		λ	0.2088	0.7408	0.7697	0.1201	0.6318	0.6430	0.0887	0.4707	0.4789
	2.5	α	0.5712	2.0248	2.1017	0.4065	1.4279	1.4839	0.3128	1.1008	1.1439
		β_1	0.0544	0.4883	0.4908	0.0241	0.3454	0.3460	0.0069	0.2796	0.2796
		β_2	0.0073	0.5443	0.5438	0.0070	0.3423	0.3422	0.0019	0.2705	0.2705
		β_3	0.5942	0.9954	1.1583	0.3861	0.6357	0.7435	0.2816	0.5181	0.5895
		β_4	0.0964	0.2568	0.2741	0.0748	0.1695	0.1852	0.0559	0.1393	0.1500
		δ	0.2298	0.9835	1.0099	0.1968	0.6376	0.6670	0.1721	0.4951	0.5239
		λ	0.2078	1.8451	1.8547	0.1331	1.6260	1.6305	0.0266	1.5968	1.5970

7. Real Applications

In this section, two illustrations of the log-FBS nonlinear regression model are given. In Section 7.1, a discussion of linear and nonlinear regression models is carried out. The residual analysis techniques proposed in Section 5 are applied to check the model’s adequacy. On the other hand, in Section 7.2, the emphasis is put on the use of diagnostic influence techniques proposed in Section 4.

7.1. Illustration I

7.1.1. Classical Discussion of Models

Let us consider the Australian Institute of Sport (AIS) dataset available from the *sn* package of R [26]. It consists of $n = 202$ measurements from high-performance athletes on various characteristics of the blood. We aim to explain the Hematocrit (Hc) variable in terms of Hemoglobin (Hg). A linear model with log-FBS errors and two nonlinear models with log-BS and log-FBS errors are considered for $Y = \log(Hc)$. First, it is established that a nonlinear model is better. Later, by using a likelihood ratio test (LRT), it is seen that a nonlinear model with log-FBS errors must be preferred. Its adequacy is checked by envelope plots for the martingale-type residuals.

The regression models under consideration are:

1. Linear model for $Y = \log(Hc)$:

$$Y_i = \log(Hc_i) = \beta_1 + \beta_2 x_i + \epsilon_i, \quad i = 1, 2, \dots, n$$

where

$$\epsilon_i \sim \text{Log-FBS}(\alpha, 0, 2, \delta, \lambda).$$

2. Nonlinear model for $Y = \log(Hc)$ with log-BS errors:

$$Y_i = \log(Hc_i) = \beta_1 x_i^{\beta_2} + \epsilon_i, \quad i = 1, 2, \dots, n$$

where

$$\epsilon_i \sim \text{Log-BS}(\alpha, 0, 2).$$

3. Nonlinear model for $Y = \log(Hc)$ with log-FBS errors:

$$Y_i = \log(Hc_i) = \beta_1 x_i^{\beta_2} + \epsilon_i, \quad i = 1, 2, \dots, n$$

where

$$\epsilon_i \sim \text{Log-FBS}(\alpha, 0, 2, \delta, \lambda).$$

Table 7 provides the parameter estimates by using maximum likelihood, along with the estimated standard errors (in parentheses) and the p -values of tests for the coefficients. The proposed models are compared by using the Akaike information criterion (AIC) and corrected AIC (AICc) defined as

$$AIC = -2\ell(\hat{\theta}) + 2k \quad \text{and} \quad AICc = AIC + \frac{2k(k+1)}{n-k-1},$$

where k is the number of parameters in the model. The better model is the one with the lowest AIC or AICc.

From AIC and AICc in Table 7, the nonlinear log-BS and nonlinear log-FBS models provide the better fit.

Since nonlinear log-BS and nonlinear log-FBS are nested models, they can be compared with a likelihood ratio test (LRT), where

$$H_0 : (\delta, \lambda) = (0, 0) \quad \text{versus} \quad H_1 : (\delta, \lambda) \neq (0, 0),$$

and the test statistic is

$$\Delta = \frac{L_{LBS}(\tilde{\alpha}, \tilde{\beta}_1, \tilde{\beta}_2)}{L_{LFBS}(\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2, \hat{\delta}, \hat{\lambda})}$$

By taking $\alpha = 0.05$ and using the estimates in Table 7, we obtain

$$-2 \log(\Delta) = -2(451.8546 - 457.0297) = 10.3502,$$

which is greater than the critical point $\chi^2_{2,0.95} = 5.9914$. Therefore, the null hypothesis is rejected, and it can be stated that the nonlinear log-FBS model provides a better fit than the nonlinear log-BS one.

Moreover, Figure 1a,b give the envelopes for the martingale-type residuals for both nonlinear models. Again, these plots suggest that the nonlinear log-FBS model provides a better fit than the nonlinear log-BS one.

Table 7. Parameter estimates (standard error) and p -values for nonlinear log-BS, linear log-FBS, and nonlinear log-FBS models, along with AIC and AICc values.

Parameter Estimates	Nonlinear Log-BS	Linear Log-FBS	Nonlinear Log-FBS
α	0.0258 (0.0012)	0.0260 (0.0016)	0.0244 (0.0014)
p -value	(0.0000)	(0.0000)	(0.0000)
β_1	2.0343(0.0281)	2.9460 (0.0000)	2.0822 (0.0001)
p -value	(0.0000)	(0.0000)	(0.0000)
β_2	0.2295 (0.0051)	0.0593 (0.0001)	0.2259 (0.0002)
p -value	(0.0000)	(0.0000)	(0.0000)
δ		-1.9303 (0.0070)	-2.0949 (0.0062)
p -value		(0.0000)	(0.0000)
λ		-1.5232 (0.0071)	-1.3293 (0.0053)
p -value		(0.0000)	(0.0000)
AIC	-897.70	-893.27	-904.05
AICc	-897.58	-892.96	-903.75

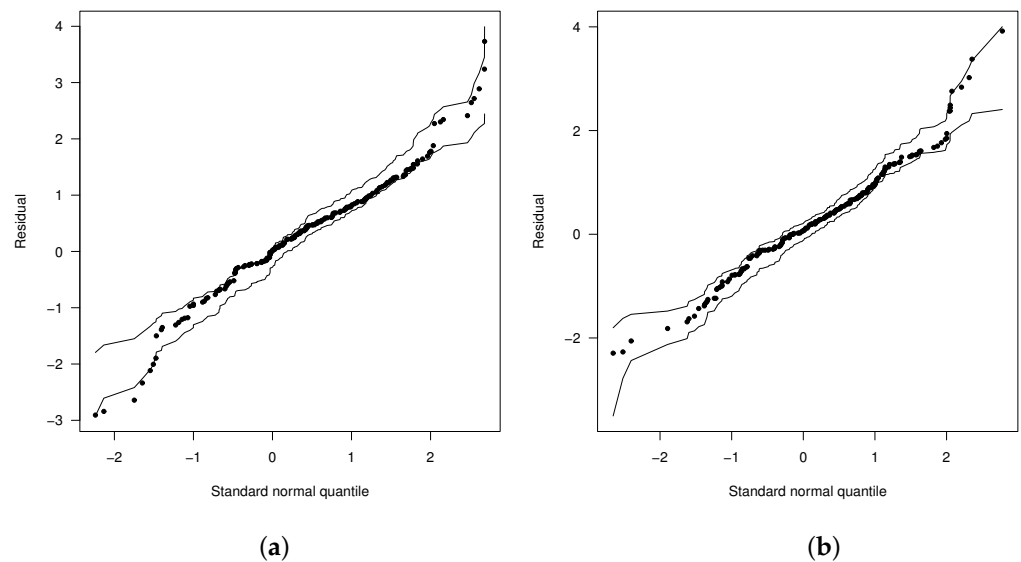


Figure 1. Normal probability plots for r_{MT_i} with envelopes of Q-plots for the scaled residuals for the fitted models: (a) nonlinear log-BS model and (b) nonlinear log-FBS model.

7.1.2. Discussion of Models Based on Validation Techniques

An evaluation of regression models previously proposed is next carried out by using validation techniques. Uses of cross-validation (CV) and generalized CV criteria to select models or optimal ridge parameters in partial linear regression models can be seen in [27,28], among others. In our case, we follow a validation approach.

We randomly split the 202 observations into two sets: a training set containing 101 of the data points, and a validation set with the remaining 101 observations. The regression model fitted to the training sample is later evaluated on the validation sample by using the mean squared error (MSE) as a measure of error. This process is repeated 100 times, that is, using 100 different random splits of the original sample into training and validation sets. In summary, the mean of the $MSE_i, i = 1, \dots, 100$, is given in Table 8.

Table 8. MSE obtained for linear log-FBS, nonlinear log-BS, and nonlinear log-FBS models by using validation approach (50%, 50%), 100 times.

	Linear Log-FBS	Nonlinear Log-BS	Nonlinear Log-FBS
MSE	1.107259	0.8222637	0.004790429

Since the nonlinear log-FBS model exhibits the smallest MSE in Table 8, this model would be preferred to describe our dataset.

As for goodness-of-fit tests to decide if the sample comes from a specific model, the Anderson–Darling test, implemented in the goftest R package, Version 1.2-3 is used [29]. The p -values of these tests are listed in Table 9. These summaries suggest that the nonlinear log-FBS model provides a good fit for this dataset.

Table 9. Results for the Anderson–Darling goodness-of-fit test.

	Linear Log-FBS	Nonlinear Log-BS	Nonlinear Log-FBS
MSE	0.0052	0.0000	0.6750

7.2. Illustration II: Model Discussion with Emphasis on the Use of Diagnostic Influence Techniques

The real dataset under consideration consists of 46 independent observations of a metal specimen subject to cyclic stress. This dataset has been previously studied by Rieck and Nedelman [10], Galea et al. [30] and Xie and Wei [31]. The variables involved in this study are: the response variable (T), number of cycles to failure, and the explanatory covariable (X), which corresponds to the work by cycle (mJ/m^3). The aim is to check the number of cycles before a failure (crack) happens. Rieck and Nedelman [10] proposed to fit the linear regression model

$$Y_i = \log(T_i) = \beta_1 + \beta_2 \log(X_i) + \epsilon_i, \tag{31}$$

where $\epsilon_i \sim SHN(\alpha, 0, 2)$.

Note that Equation (31) is a linear model in the logarithm of the explanatory variable X . This fact may render cumbersome the interpretation of parameters in the fitted model. An alternative is to keep the original explanatory variable X and try to fit a nonlinear model. As for what nonlinear model to propose, the scatter plot of $\{(x_i, y_i)\}$ may be useful. In this case, an exponential model is suggested. So, we propose to fit the nonlinear regression model

$$Y_i = \log(T_i) = \beta_1 + \beta_2 \exp(\beta_3 / X_i) + \epsilon_i,$$

where $\epsilon_i \sim FSHN(\alpha, 0, \sigma, \delta, \lambda)$.

Note that

$$\mathbb{E}(Y_i) = \mu_i = \beta_1 + \beta_2 \exp(\beta_3 / X_i) + \mathbb{E}(\epsilon_i) = \beta_1^* + \beta_2 \exp(\beta_3 / X_i), \tag{32}$$

with $\beta_1^* = \beta_1 + \mathbb{E}(\epsilon)$ and $\epsilon \sim FSHN(\alpha, 0, \sigma, \delta, \lambda)$.

Remark 3. A brief discussion to illustrate that, in this case, the nonlinear proposal is superior to the linear one is given in Appendix A.3.

The maximum likelihood estimates for the parameters in Equation (32) along with their standard errors (in parentheses) are given in Table 10.

Table 10. Estimates of parameters in Equation (32).

$\hat{\alpha} = 2.3061$ (1.3846)	$\hat{\beta}_1 = 10.1725$ (1.4170)	$\hat{\beta}_2 = -5.8413$ (1.0844)	$\hat{\beta}_3 = -17.9475$ (8.5355)
$\hat{\sigma} = 0.9499$ (0.3100)	$\hat{\delta} = 1.2984$ (0.6121)	$\hat{\lambda} = -5.6195$ (2.6171)	

In Figure 2a,b, martingale-based residuals and the diagonal elements in matrix GL_{ii} are plotted. These plots suggest that certain observations may be potentially influential. These are cases #1, #2, #3, #4, #5, #12, #32, and #46, among others.

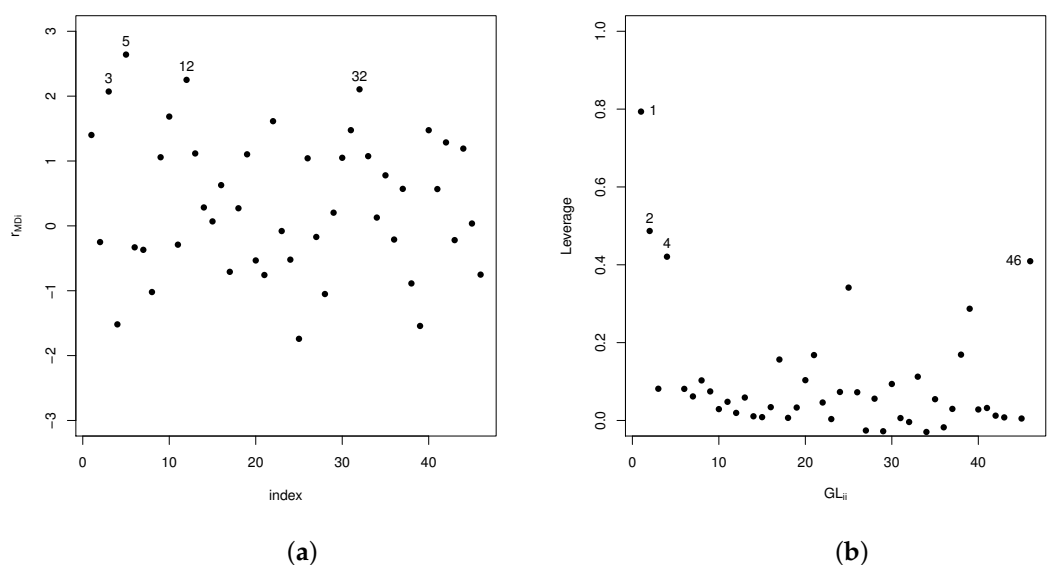


Figure 2. (a) Plots of deviance component residuals r_{MT_i} . (b) Elements in the diagonal of the generalized Leverage matrix GL_{ii} .

Analogously, diagnostic analyses have been carried out for the vector of parameter β and the remaining parameters in the model $\alpha, \sigma, \delta, \lambda$. The statistic $GDC_i(\theta)$ introduced in Section 4.1 is used in the following settings: case-weights, perturbation of the response variable, perturbed covariables, and perturbed scale parameters. Our results show that, in the case-weights setting, cases #1, #3, and #4 are detected as potentially influential, whereas for the perturbed response variable and the perturbed scale parameter setting, only case #44 was detected as potentially influential. As for the perturbed covariable setting, cases #1, #2, #3, #8, #13, and #19 are detected. The plots given in Figures 3a,b, 4a,b, 5a,b and 6a,b show the results of these analyses.

Next, the influence of previously mentioned cases is studied through the relative change statistic, defined as

$$RC_{\theta_j} = \left| \frac{\hat{\theta}_j - \hat{\theta}_{j(I)}}{\hat{\theta}_j} \right| \times 100,$$

where $\hat{\theta}_j$ denotes the maximum likelihood estimate for the parameter θ_j , including all observations, and $\hat{\theta}_{j(I)}$ denotes the estimate of the same parameter, deleting the influential observations.

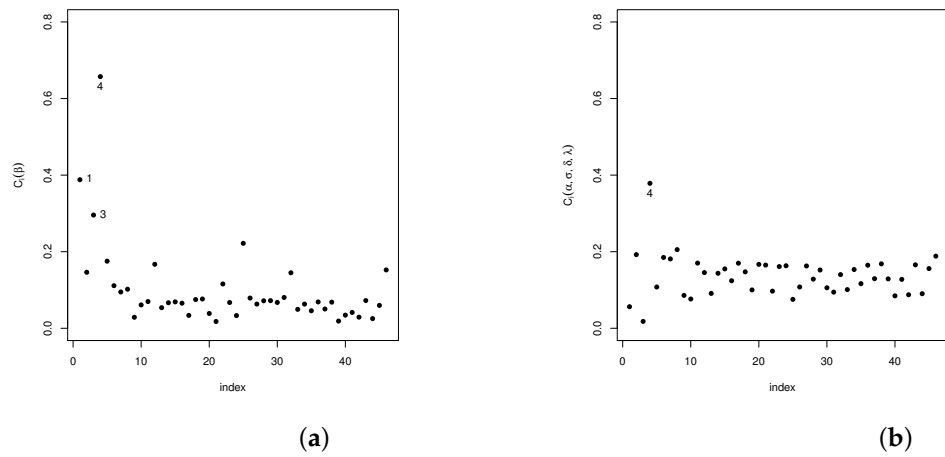


Figure 3. Plots for the $C_i(\cdot)$ index in the case-weights scheme: (a) for β and (b) for $\theta_1 = (\alpha, \sigma, \delta, \lambda)$.

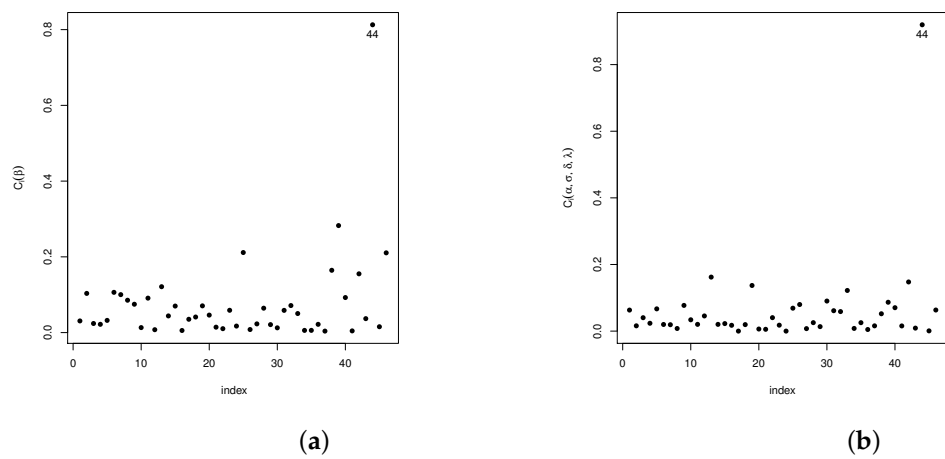


Figure 4. Plots for the $C_i(\cdot)$ index in the perturbation of the response variable scheme: (a) for β and (b) for $\theta_1 = (\alpha, \sigma, \delta, \lambda)$.

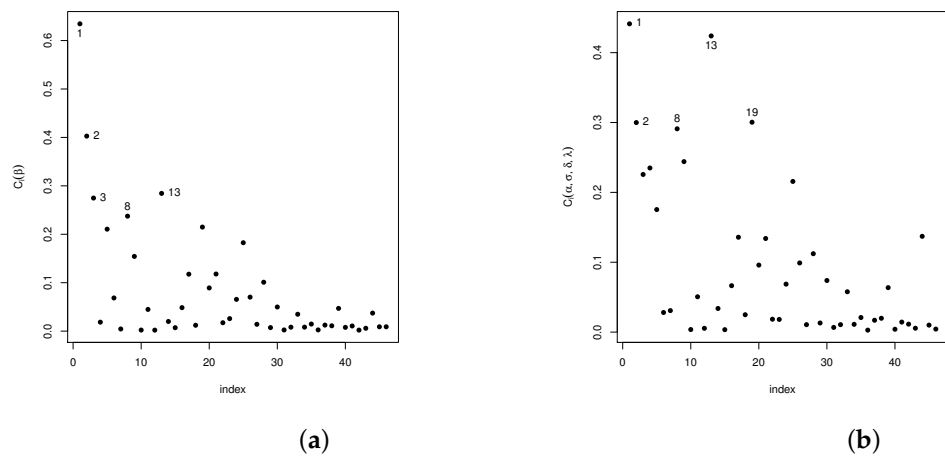


Figure 5. Plots for the $C_i(\cdot)$ index in the perturbed covariate scheme: (a) for β and (b) for $\theta_1 = (\alpha, \sigma, \delta, \lambda)$.

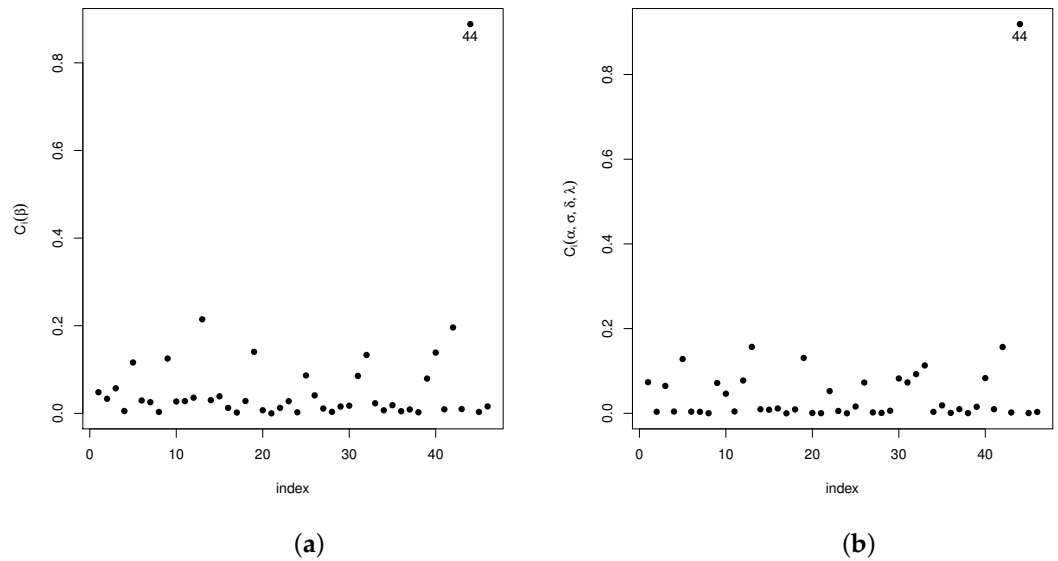


Figure 6. Plots for the $C_i(\cdot)$ index for the perturbation of the scale parameter scheme: (a) for β and (b) for $\theta_1 = (\alpha, \sigma, \delta, \lambda)$.

The approximation for the standard error (SE) of RC proposed in Santana et al. [32] will be used:

$$RC(\widehat{SE}(\theta_j)) = \left| \frac{\widehat{SE}(\hat{\theta}_j) - \widehat{SE}(\hat{\theta}_{j(I)})}{\widehat{SE}(\hat{\theta}_j)} \right| \times 100,$$

where $\widehat{SE}(\hat{\theta}_j)$ denotes the SE for the maximum likelihood estimate for the parameter θ_j , including all observations, and $\widehat{SE}(\hat{\theta}_{j(I)})$ denotes the SE for the estimate of the same parameter, deleting the influential observations.

Table 11 provides the RC(%), their $RC(\widehat{SE}(\theta_j))$ in parentheses, and the p -values associated with test $H_0 : \beta_j = 0$ versus $H_1 : \beta_j \neq 0$ for $j = 1, 2, 3$.

From the results in Table 11, we may conclude that, individually, cases #2, #44, and #46 affect the estimates of the parameters $\beta_1, \beta_2,$ and β_3 in the nonlinear regression model, as well as the scale (σ), shape (δ), and skewness (λ) parameters. It can also be seen that cases #2 and #46 have influence on the ML-estimator of α .

Moreover, the joint influence of cases $\{2, 44\}, \{2, 46\}, \{44, 46\},$ and $\{2, 44, 46\}$ is analyzed. It can be seen that, jointly, these observations have influence on the estimates of the model parameters. As for the hypothesis tests for the parameters in the nonlinear regression model, none of the observations (#2, #44, and #46) affect their significance.

Therefore, cases #2, #44, and #46 were eliminated. The new fitted model is

$$\hat{\mu}_i = \hat{y}_i | x_i = 9.1327 - 5.3787e^{-21.3189/x_i}, \tag{33}$$

whose error terms are distributed as $\epsilon_i \sim FSHN(2.2841, 0, 0.9028, 1.0741, -4.6699)$.

The plots in Figure 7a,b show the envelope plot for the martingale-type residuals and the dispersion plot for the final fitted model given in Equation (33). Both plots suggest that this is a good choice for describing this dataset.

Table 11. Relative change, $RC\%(RC(\hat{SE}(\theta_j)))$, and p -values for model parameters.

Dropped Case		$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\lambda}$
{1}	RC	65.06 (373.35)	3.15 (95.07)	1.58 (92.37)	8.77 (98.23)	18.43 (27.66)	1.14 (392.14)	188.39 (726.20)
	p -value		<0.0001	<0.0001	<0.0001			
{2}	RC	52.71 (1.38)	6.78 (0.46)	10.31 (0.47)	20.22 (0.18)	20.22 (0.12)	42.08 (2.27)	6.76 (0.40)
	p -value		<0.0001	<0.0001	0.0019			
{3}	RC	32.49 (400.94)	0.56 (31.20)	1.80 (32.00)	0.64 (31.48)	15.48 (100.93)	28.76 (447.30)	123.62 (623.66)
	p -value		<0.0001	<0.0001	0.0022			
{4}	RC	21.11 (195.41)	7.75 (77.04)	15.98 (81.52)	28.77 (41.05)	1.92 (223.74)	101.39 (230.88)	665.89 (2330.91)
	p -value		<0.0001	<0.0001	<0.0001			
{5}	RC	57.45 (209.88)	0.82 (31.88)	3.41 (35.85)	5.17 (23.69)	27.15 (14.58)	34.87 (209.86)	133.96 (398.36)
	p -value		<0.0001	<0.0001	0.0037			
{8}	RC	59.70 (420.19)	0.17 (94.88)	3.32 (91.86)	2.06 (97.94)	23.12 (42.53)	25.68 (379.91)	120.30 (575.81)
	p -value		<0.0001	<0.0001	<0.0001			
{12}	RC	61.67 (565.47)	3.47 (51.15)	6.86 (55.68)	14.91 (32.77)	24.02 (72.43)	15.74 (551.91)	130.18 (751.91)
	p -value		<0.0001	<0.0001	0.0003			
{13}	RC	77.45 (519.92)	1.38 (37.44)	4.55 (41.19)	5.92 (27.83)	21.63 (43.23)	3.74 (510.58)	175.01 (841.64)
	p -value		<0.0001	<0.0001	0.0020			
{19}	RC	50.87 (375.77)	1.77 (37.60)	4.93 (42.11)	7.76 (25.55)	17.93 (52.82)	16.93 (391.25)	146.89 (643.19)
	p -value		<0.0001	<0.0001	0.0023			
{32}	RC	11.88 (109.88)	2.75 (94.60)	6.83 (91.06)	10.26 (97.26)	13.98 (16.50)	54.41 (149.02)	77.57 (205.20)
	p -value		<0.0001	<0.0001	<0.0001			
{44}	RC	3.87 (2.66)	5.88 (0.96)	8.57 (0.94)	17.81 (0.98)	7.59 (2.53)	49.73 (4.33)	17.77 (0.85)
	p -value	0.6362	0.0000	0.0000	0.0000	0.3500	0.5512	0.3396
{46}	RC	7.67 (1.42)	5.43 (0.95)	7.63 (0.93)	18.59 (0.98)	8.77 (0.67)	10.22 (2.17)	18.89 (0.28)
	p -value		<0.0001	<0.0001	<0.0001			
{2, 44}	RC	71.66 (2.59)	6.96 (0.98)	11.05 (0.97)	19.89 (0.99)	21.76 (0.02)	67.37 (3.55)	3.05 (0.94)
	p -value		<0.0001	<0.0001	<0.0001			
{2, 46}	RC	134.88 (4.40)	6.81 (0.96)	10.86 (0.90)	19.11 (0.96)	32.51 (0.25)	85.00 (4.63)	12.13 (1.63)
	p -value		<0.0001	<0.0001	<0.0001			
{44, 46}	RC	3.27 (2.16)	5.76 (0.98)	4.09 (0.94)	18.50 (0.99)	21.85 (3.44)	115.02 (5.22)	48.58 (0.09)
	p -value		<0.0001	<0.0001	<0.0001			
{2, 44, 46}	RC	0.95 (0.12)	5.47 (0.22)	7.91 (0.29)	18.78 (0.12)	4.95 (0.02)	17.26 (0.53)	16.89 (0.13)
	p -value		<0.0001	<0.0001	0.0262			

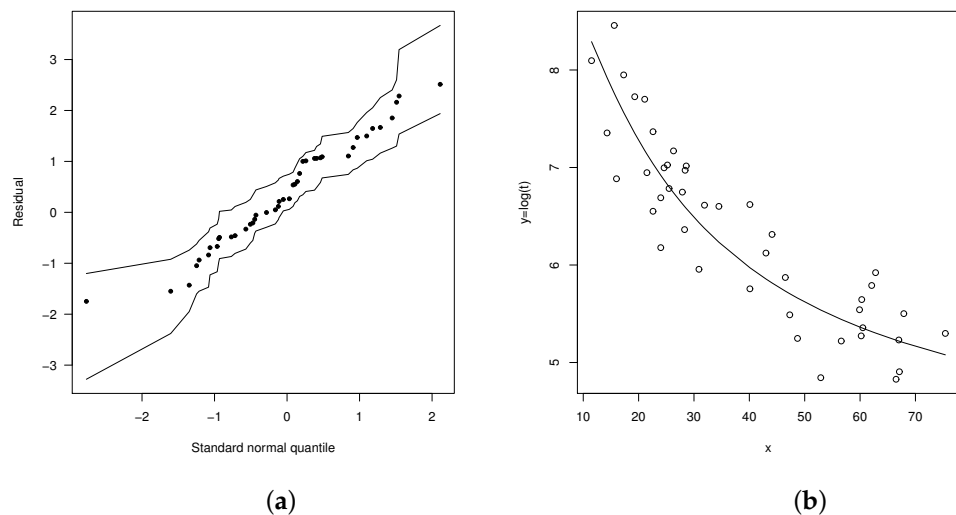


Figure 7. (a) Normal probability plots for rMT_i with envelopes of Q-qplots for the scaled residuals from the fitted model and (b) dispersion diagram for nonlinear model FSHN with errors.

8. Conclusions

In this paper, a nonlinear log-Birnbaum–Saunders regression model with additive errors is introduced. It is assumed that the errors follow the flexible sinh-normal distribution introduced in [8]. As an advantage, we highlight that this model can be used to describe a continuous response variable with asymmetric, unimodal, or bimodal behavior by using a

nonlinear function of covariates. Its theoretical properties are studied, along with results from the maximum likelihood estimation of the parameters. Influence diagnostics are introduced in order to provide tools to detect influential observations and deficiencies in the fitted model. Statistics based on Cook's generalized distance and local influence diagnostics are considered. Several perturbation schemes are proposed. Specifically, these are: case-weight, perturbations of the response variable, perturbations of an explanatory variable, and perturbations of the scale parameter. Martingale-type residual analysis is presented to detect deficiencies in the fitted model. Simulation studies are included, where the Michaelis–Menten model of interest in kinetic chemistry is considered. The simulations cover a variety of situations as for the range of parameters under consideration. In particular, we highlight that negative and positive values for the parameter that controls the unimodality or bimodality in our model are considered. Our results suggest that the estimators of the parameters are consistent and show the good performance in both unimodal and bimodal settings. To conclude, applications to real datasets are given in order to illustrate how to use our methodology.

We highlight that, in the first application, a discussion of linear and nonlinear regression models is carried out, first from a classical point of view. Several information criteria and the likelihood ratio test are applied. Martingale-type residual analysis techniques are applied to check the model's adequacy. Second, this discussion is also carried out by using cross-validation techniques. In both cases, our proposal is superior to the other regression models under consideration. On the other hand, in the second real application, the emphasis is on the use of diagnostic influence techniques to detect influential observations.

As important novelties and advantages of our model, we highlight the nonlinearity of our proposal and the kind of distribution used for the errors, since it is assumed that they follow a flexible sinh-normal distribution. Therefore, it is useful for modeling unimodal and bimodal settings. Also, recall that it can be applied to nonnegative data and that it is of interest in reliability and medicine. Moreover, our approach is more general and reduces to other nonlinear regression models previously introduced in the literature, such as the log-BS nonlinear regression model with asymmetric errors introduced in [15].

As future research, we will consider the extension of these models to a censored continuous response variable Y . This kind of study is relevant, for instance, in medicine, where, quite often, we have data that can only be recorded above or below a certain threshold and that depend on the sensitivity of a laboratory test.

Author Contributions: Conceptualization, G.M.-F., I.B.-C. and H.W.G.; formal analysis, G.M.-F., I.B.-C. and H.W.G.; investigation, G.M.-F., I.B.-C. and H.W.G.; methodology, G.M.-F., I.B.-C. and H.W.G.; software, G.M.-F.; supervision, H.W.G.; validation, G.M.-F. and I.B.-C.; visualization, H.W.G. All authors have read and agreed to the published version of the manuscript.

Funding: The research of H.W. Gómez was supported by Grant SEMILLERO UA-2024 (Chile). The research of I. Barranco-Chamorro was supported by IOAP of the University of Seville, Spain. The research of G. Martínez-Flórez was supported by the Vice-rectorate for Research of the Universidad de Córdoba, Colombia, project grant FCB-06-22.

Data Availability Statement: Details about data availability have been given in Section 7. The numerical code is available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations for distributions are used in this manuscript:

BS	Birnbaum–Saunders
SHN	sinh-normal
FSHN	flexible sinh-normal
FBS	flexible Birnbaum–Saunders
FLBS	flexible log-Birnbaum–Saunders

Appendix A

Appendix A.1

The following lemma was used to establish the existence of Equations (9) and (10).

Lemma A1. Let $\epsilon \sim \text{FSHN}(\alpha, 0, 2, \delta, \lambda)$. Then,

$$\begin{aligned} \mathbb{E}(\epsilon) &= 2c_1(\alpha, \delta, \lambda), \\ \text{Var}(\epsilon) &= 4V(\alpha, \delta, \lambda), \end{aligned}$$

where

$$c_1(\alpha, \delta, \lambda) = \mathbb{E}_{\text{FSN}}\left(\text{arcsinh}\left(\frac{\alpha Z}{2}\right)\right) = c_\delta \int_{-\infty}^{\infty} \text{arcsinh}\left(\frac{\alpha z}{2}\right) \phi(|z| + \delta) \Phi(\lambda z) dz,$$

and $V(\alpha, \delta, \lambda)$ denotes the variance of the RV $\text{arcsinh}(\alpha Z/2)$, with $Z \sim \text{Flexible Skew - Normal}(\delta, \lambda)$.

Proof. Since $\epsilon \sim \text{FSHN}(\alpha, 0, 2, \delta, \lambda)$, it can be seen in [8] that $\epsilon = 2\text{arcsinh}(\alpha Z/2)$, with $Z \sim \text{Flexible Skew - Normal}(\delta, \lambda)$. Recall that the PDF of Z is

$$f_Z(z) = c_\delta \phi(|z| + \delta) \Phi(\lambda z), \quad z \in \mathbb{R}.$$

Therefore,

$$\mathbb{E}(\epsilon) = 2 \mathbb{E}_{\text{FSN}}\left(\text{arcsinh}\left(\frac{\alpha Z}{2}\right)\right) = 2 c_\delta \int_{\mathbb{R}} \text{arcsinh}\left(\frac{\alpha z}{2}\right) \phi(|z| + \delta) \Phi(\lambda z) dz .$$

Taking into account the relationship $\text{arcsinh}(x) = \log(x + \sqrt{x^2 + 1})$, $\forall x \in \mathbb{R}$, and the fact that, if $w = x + \sqrt{x^2 + 1} > 0$, then $\log(w) \leq w - 1$ (see Love [33]), then we have that

$$\int_{\mathbb{R}} \text{arcsinh}\left(\frac{\alpha z}{2}\right) \phi(|z| + \delta) \Phi(\lambda z) dz$$

can be expressed in terms of incomplete moments of the skew-normal distribution. The existence of moments of the skew-normal distribution can be seen in Haas [16]; therefore, $c_1(\alpha, \delta, \lambda)$ also exists.

Note that we can proceed similarly for $\text{Var}(\epsilon)$ or $\mathbb{E}_{\text{FSN}}\left(\text{arcsinh}\left(\frac{\alpha Z}{2}\right)^j\right)$ with $j \in \mathbb{Z}^+$, since the moments of the skew-normal distribution exist; see Haas [16]. Therefore, the moments of $\mathbb{E}_{\text{FSN}}\left(\text{arcsinh}\left(\frac{\alpha Z}{2}\right)^j\right)$ with $j \in \mathbb{Z}^+$ also exist, in particular, $V(\alpha, \delta, \lambda)$. \square

Appendix A.2

Next, the elements of the observed information matrix are given.

Corollary A1. Let Y_1, \dots, Y_n be a random sample of $Y \sim \text{FLBS}(\alpha, f(\boldsymbol{\beta}, \mathbf{x}_i), \sigma, \delta, \lambda)$. Then, the elements of the observed information matrix are:

$$\begin{aligned}
 j_{\alpha\alpha} &= -\frac{n}{\alpha^2} + \frac{3}{\alpha^2} \sum_{i=1}^n \bar{\xi}_{i2}^2 + \frac{\lambda}{\alpha^2} \sum_{i=1}^n W_i \bar{\xi}_{i2} [-2 + \lambda \bar{\xi}_{i2} (\lambda \bar{\xi}_{i2} + W_i)] + \frac{2\delta}{\alpha^2} \sum_{i=1}^n \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i2}, \\
 j_{\alpha\beta_j} &= \frac{2}{\alpha\sigma} \sum_{i=1}^n d_{i,j} \bar{\xi}_{i1} \bar{\xi}_{i2} + \frac{\lambda}{\alpha\sigma} \sum_{i=1}^n d_{i,j} W_i \bar{\xi}_{i1} [-1 + \lambda \bar{\xi}_{i2} (\lambda \bar{\xi}_{i2} + W_i)] + \frac{\delta}{\alpha\sigma} \sum_{i=1}^n d_{i,j} \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i1}, \\
 j_{\alpha\delta} &= -\frac{1}{\alpha} \sum_{i=1}^n \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i2}, \\
 j_{\alpha\lambda} &= -\frac{1}{\alpha} \sum_{i=1}^n \bar{\xi}_{i2} W_i [-1 + \lambda \bar{\xi}_{i2} (\lambda \bar{\xi}_{i2} + W_i)], \\
 j_{\alpha\sigma} &= \frac{2}{\alpha\sigma} \sum_{i=1}^n z_i \bar{\xi}_{i1} \bar{\xi}_{i2} - \frac{\lambda}{\alpha\sigma} \sum_{i=1}^n z_i \bar{\xi}_{i1} W_i [-1 + \lambda \bar{\xi}_{i2} (\lambda \bar{\xi}_{i2} + W_i)] + \frac{\delta}{\alpha\sigma} \sum_{i=1}^n \text{sgn}(\bar{\xi}_{i2}) z_i \bar{\xi}_{i1}, \\
 j_{\beta_i\beta_j} &= -\frac{1}{\sigma} \sum_{i=1}^n g_{i,jj'} \left[\bar{\xi}_{i1} \bar{\xi}_{i2} - \lambda \bar{\xi}_{i1} W_i + \delta \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i1} - \frac{\bar{\xi}_{i2}}{\bar{\xi}_{i1}} \right] + \frac{\delta}{\sigma^2} \sum_{i=1}^n d_{i,j} d_{i,j'} \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i2} \\
 &+ \frac{1}{\sigma^2} \sum_{i=1}^n d_{i,j} d_{i,j'} \left[2\bar{\xi}_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\bar{\xi}_{i2}^2}{\bar{\xi}_{i2}^2 + 4/\alpha^2} \right] + \frac{\lambda}{\sigma^2} \sum_{i=1}^n d_{i,j} d_{i,j'} W_i \left[-\bar{\xi}_{i2} + \lambda \bar{\xi}_{i1}^2 (\lambda \bar{\xi}_{i2} + W_i) \right], \\
 j_{\beta_j\sigma} &= \frac{1}{\sigma^2} \sum_{i=1}^n d_{i,j} \left[\bar{\xi}_{i1} \bar{\xi}_{i2} - \frac{\bar{\xi}_{i2}}{\bar{\xi}_{i1}} \right] + \frac{\delta}{\sigma^2} \sum_{i=1}^n d_{i,j} \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i1} + \frac{\delta}{\sigma^2} \sum_{i=1}^n d_{i,j} \text{sgn}(\bar{\xi}_{i2}) z_i \bar{\xi}_{i2} - \frac{\lambda}{\sigma^2} \sum_{i=1}^n d_{i,j} \bar{\xi}_{i1} W_i \\
 &+ \frac{1}{\sigma^2} \sum_{i=1}^n d_{i,j} z_i \left[2\bar{\xi}_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\bar{\xi}_{i2}^2}{\bar{\xi}_{i2}^2 + 4/\alpha^2} \right] + \frac{\lambda}{\sigma^2} \sum_{i=1}^n d_{i,j} z_i W_i \left[-\bar{\xi}_{i2} + \lambda \bar{\xi}_{i1}^2 (\lambda \bar{\xi}_{i2} + W_i) \right], \\
 j_{\sigma\sigma} &= \frac{2}{\sigma^2} \sum_{i=1}^n z_i \left[\bar{\xi}_{i1} \bar{\xi}_{i2} - \frac{\bar{\xi}_{i2}}{\bar{\xi}_{i1}} \right] + \frac{\delta}{\sigma^2} \sum_{i=1}^n \text{sgn}(\bar{\xi}_{i2}) z_i [2\bar{\xi}_{i1} + z_i \bar{\xi}_{i2}] - \frac{2\lambda}{\sigma^2} \sum_{i=1}^n z_i \bar{\xi}_{i1} W_i - \frac{n}{\sigma^2} \\
 &+ \frac{1}{\sigma^2} \sum_{i=1}^n z_i^2 \left[2\bar{\xi}_{i2}^2 + \frac{4}{\alpha^2} - 1 + \frac{\bar{\xi}_{i2}^2}{\bar{\xi}_{i2}^2 + 4/\alpha^2} \right] + \frac{\lambda}{\sigma^2} \sum_{i=1}^n z_i^2 W_i \left[-\bar{\xi}_{i2} + \lambda \bar{\xi}_{i1}^2 (\lambda \bar{\xi}_{i2} + W_i) \right], \\
 j_{\beta_i\lambda} &= -\frac{1}{\sigma} \sum_{i=1}^n d_{ij} W_i \bar{\xi}_{i1} [-1 + \lambda \bar{\xi}_{i2} (\lambda \bar{\xi}_{i2} + W_i)], & j_{\lambda\lambda} &= \sum_{i=1}^n \bar{\xi}_{i2}^2 W_i (\lambda \bar{\xi}_{i2} + W_i), \\
 j_{\sigma\lambda} &= \frac{1}{\sigma} \sum_{i=1}^n z_i W_i \bar{\xi}_{i1} - \frac{\lambda}{\sigma} \sum_{i=1}^n z_i \bar{\xi}_{i1} \bar{\xi}_{i2} W_i (\lambda \bar{\xi}_{i2} + W_i), & j_{\sigma\delta} &= -\frac{1}{\sigma} \sum_{i=1}^n \text{sgn}(\bar{\xi}_{i2}) z_i \bar{\xi}_{i1}, \\
 j_{\beta_j\delta} &= -\frac{1}{\sigma} \sum_{i=1}^n d_{i,j} \text{sgn}(\bar{\xi}_{i2}) \bar{\xi}_{i1}, & j_{\lambda\delta} &= 0, & j_{\delta\delta} &= n(W_\delta(\delta - W_\delta) + 1),
 \end{aligned}$$

where $W_\delta = \phi(\delta)/(1 - \Phi(\delta))$.

Proof. It is straightforward by applying standard calculus techniques. \square

Appendix A.3

In this appendix, it is shown that the nonlinear model proposed in Section 7.2 is superior to a linear model. So, for the dataset studied in Section 7.2, let us consider the linear model with sinh-normal errors or log-BS

$$Y_i = \log(T_i) = \beta_1 + \beta_2 x_i + \epsilon_i,$$

with $\epsilon_i \sim SHN(\alpha, 0, 2)$. Summaries for this model and the nonlinear one proposed in Equation (32) are given in Table A1. From AIC and AICc in Table A1, it can be concluded that the nonlinear flexible log-BS provides a better fit to this dataset than the linear log-BS model.

Table A1. Parameter estimates (standard errors).

Estimates	Log-BS	Flexible Log-BS
$\hat{\alpha}$	0.5202 (0.0543)	2.3061 (1.3846)
p -value	(0.0000)	(0.0479)
$\hat{\beta}_1$	7.9849 (0.1546)	10.1725 (1.4170)
p -value	(0.0000)	(0.0000)
$\hat{\beta}_2$	−0.0406 (0.0033)	−5.8413 (1.0844)
p -value	(0.0000)	(0.0000)
$\hat{\beta}_3$		−17.9475 (8.5355)
p -value		(0.0354)
$\hat{\sigma}$		0.9499 (0.3100)
p -value		(0.0021)
$\hat{\delta}$		1.2984 (0.6121)
p -value		(0.0339)
$\hat{\lambda}$		−5.6195 (2.6171)
p -value		(0.0317)
AIC	75.5638	53.6077
AICc	76.5394	56.5550

References

- Cancho, V.G.; Lachos, V.H.; Ortega, E.M.M. A nonlinear regression model with skew-normal errors. *Stat. Papers* **2010**, *51*, 547–558. [\[CrossRef\]](#)
- Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
- Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. The Log-Linear Birnbaum–Saunders Power Model. *Methodol. Comput. Appl. Probab.* **2017**, *19*, 913–933. [\[CrossRef\]](#)
- Bolfarine, H.; Martínez-Flórez, G.; Salinas, H.S. Bimodal symmetric-asymmetric families. *Commun. Stat. Theory Methods* **2018**, *47*, 259–276. [\[CrossRef\]](#)
- Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. An alpha-power extension for the Birnbaum–Saunders distribution. *Statistics* **2014**, *48*, 896–912. [\[CrossRef\]](#)
- Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [\[CrossRef\]](#)
- Li, X.; Chu, H.; Gallant, J.E.; Hoover, D.R.; Mack, W.J.; Chmiel, J.S.; Muñoz, A. Bimodal virologic response to antiretroviral therapy for HIV infection: An application using a mixture model with left censoring. *J. Epidemiol. Commun. Health* **2006**, *60*, 811–818. [\[CrossRef\]](#)
- Martínez-Flórez, G.; Barranco-Chamorro, I.; Gómez, H.W. Flexible Log-Linear Birnbaum–Saunders Model. *Mathematics* **2021**, *9*, 1188. [\[CrossRef\]](#)
- Birnbaum, Z.W.; Saunders, S.C. A New Family of Life Distributions. *J. Appl. Probab.* **1969**, *6*, 319–327. [\[CrossRef\]](#)
- Rieck, J.R.; Nedelman, J.R. A log-linear model for the Birnbaum–Saunders distribution. *Technometrics* **1991**, *33*, 51–60.
- Leiva, V.; Vilca-Labra, F.; Balakrishnan, N.; Sanhueza, A. A skewed sinh-normal distribution and its properties and application to air pollution. *Commun. Stat. Theory Methods* **2010**, *39*, 426–443. [\[CrossRef\]](#)
- Lemonte, A.J. A log-Birnbaum–Saunders regression model with asymmetric errors. *J. Stat. Comput. Simul.* **2012**, *82*, 1775–1787. [\[CrossRef\]](#)
- Martínez-Flórez, G.; Barranco-Chamorro, I.; Bolfarine, H.; Gómez, H.W. Flexible Birnbaum–Saunders Distribution. *Symmetry* **2019**, *11*, 1305. [\[CrossRef\]](#)
- Gómez, H.W.; Elal-Olivero, D.; Salinas, H.S.; Bolfarine, H. Bimodal extension based on the skew-normal distribution with application to pollen data. *Environmetrics* **2011**, *22*, 50–62. [\[CrossRef\]](#)
- Lemonte, A.J.; Cordeiro, G. Birnbaum–Saunders nonlinear regression models. *Comput. Stat. Data Anal.* **2009**, *53*, 4441–4452. [\[CrossRef\]](#)
- Haas, M. A Note on the Moments of the Skew-Normal Distribution. *Econ. Bull.* **2012**, *32*, 3306–3312.
- Lehman, L.E. *Elements of Large-Sample Theory*; Springer: New York, NY, USA, 1999.
- Cook, R.D. Assessment of Local Influence. *J. R. Stat. Society Ser. B (Methodol.)* **1986**, *48*, 133–169. [\[CrossRef\]](#)
- Poon, W.Y.; Poon, Y.S. Conformal normal curvature and assessment of local influence. *J. R. Statist. Soc. B* **1999**, *61*, 51–61. [\[CrossRef\]](#)
- Barros, M.; Galea, M.; Gonzalez, M.; Leiva, V. Influence diagnostics in the tobit censored response' model. *Stat. Methods Appl.* **2010**, *19*, 379–397. [\[CrossRef\]](#)
- Therneau, T.; Grambsch, P.; Fleming, T. Martingale based residuals for survival models. *Biometrika* **1990**, *77*, 147–160. [\[CrossRef\]](#)
- Ortega, E.M.; Bolfarine, H.; Paula, G.A. Influence diagnostics in generalized log-gamma regression models. *Comput. Stat. Data Anal.* **2003**, *42*, 165–186. [\[CrossRef\]](#)
- Wei, B.C.; Hu, Y.Q.; Fung, W.K. Generalized leverage and its applications. *Scand. J. Stat.* **1998**, *25*, 25–37. [\[CrossRef\]](#)
- Atkinson, A.C. Two graphical displays for outlying and influential observations in regression. *Biometrika* **1981**, *68*, 13–20. [\[CrossRef\]](#)

25. Cysneiros, F.J.A.; Vanegas, L.H. Residuals and their statistical properties in symmetrical nonlinear models. *Stat. Probab. Lett.* **2008**, *78*, 3269–3273. [[CrossRef](#)]
26. Azzalini, A. The R Package 'sn': The Skew-Normal and Related Distributions such as the Skew-t and the SUN (Version 2.1.1). Available online: <https://cran.r-project.org/package=sn> (accessed on 15 August 2024).
27. Wasserman, L. *All of Nonparametric Statistics*; Springer Science+Business Media Inc.: New York, NY, USA, 2006.
28. Amini, M.; Roozbeh, M. Optimal partial ridge estimation in restricted semiparametric regression models. *J. Multivar. Anal.* **2015**, *136*, 26–40. [[CrossRef](#)]
29. Baddeley, B. Gofstest: Classical Goodness-of-Fit Tests for Univariate Distributions. R Package Version 1.2-3. Available online: <https://cran.r-project.org/web/packages/gofstest> (accessed on 5 June 2024).
30. Galea, M.; Leiva-Sánchez, V.; Paula, G.A. Influence diagnostics in log-Birnbaum–Saunders regression models. *J. Appl. Stat.* **2004**, *31*, 1049–1064. [[CrossRef](#)]
31. Xie, F.C.; Wei, B.C. Diagnostics analysis for log-Birnbaum–Saunders regression models. *Comput. Statist. Data Anal.* **2007**, *51*, 4692–4706. [[CrossRef](#)]
32. Santana, L.; Vilca, F.; Leiva, V. Influence analysis in skew-Birnbaum–Saunders regression models and applications. *J. Appl. Stat.* **2011**, *38*, 1633–1649. [[CrossRef](#)]
33. Love, E.R. Some Logarithm Inequalities. *Math. Gaz.* **1980**, *64*, 55–57. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.