Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/140509

A New Bimodal Maxwell Regression Model With Engineering Applications

Fábio Prataviera[1,*], *Edwin M. M. Ortega*[1] *and Gauss M. Cordeiro*[2]

[1]Departamento de Ciências Exatas, ESALQ, Universidade de São Paulo, Piracicaba, SP, Brazil
[2]Departamento de Estatística, Universidade Federal de Pernambuco, Recife, PE, Brazil

Received: 1 Apr. 2020, Revised: 29 Jul. 2020, Accepted: 7 Aug. 2020
Published online: 1 Sep. 2020

Abstract: We define the generalized odd log-logistic Maxwell distribution and propose a parametric regression model based on the new distribution with three systematic components for its parameters. Some properties and maximum likelihood estimation are addressed and various simulations for different parameter settings, systematic components and sample sizes are performed. Three applications to real data sets in engineering (strength and mechanics of materials) empirically prove the usefulness of the proposed models.

Keywords: Maximum likelihood, Maxwell distribution, Odd log-logistic distribution, Regression model, Simulation

The Maxwell (or Maxwell-Boltzmann) distribution is an important model in physics, chemistry and statistical mechanics. It forms the basis of the kinetic energy of gases and explains several fundamental properties of gases including pressure and diffusion. In statistical mechanics, it is related to properties of molecules in thermal equilibrium from the microscopic perspective. The Maxwell distribution is also important in kinetic translational energies for molecules. For example, [1] discussed this distribution for chemical reactions in gases and [2] addressed its deviations in granular gases with constant coefficient of restitution.

Recently, the Maxwell distribution has been used to model failure times in survival and reliability analysis and some of its extended forms have been investigated. Krishna and Malik [3] addressed reliability estimation in the Maxwell distribution with progressively type-II censored data, [4] explored a heterogeneous population by means of two mixture components of Maxwell distributions, [5] considered point and interval estimation procedures for the Maxwell distribution in the presence of type-I progressively hybrid censored data, [6] presented its structural properties and different methods of estimation, [7] defined the gamma-Maxwell distribution, and more recently [8] proposed the transmuted exponentiated Maxwell distribution. However, none of these papers deal with bimodality as real data and do not even present regression models for the extensions of the Maxwell distribution. We aim to fill up this gap.

Many scientific studies involve data with bimodal characteristics of continuous random variables, in which the usual choice is to employ a mixture of distributions. However, the most common mixtures of distributions have a large number of parameters, so their estimation is complicated. For example, engineers study materials to learn their properties and the problems they cause. In this respect, ceramic materials are generally composed of a combination of metallic and non-metallic elements (forming oxides, nitrides and carbides), and are more resistant to high temperatures and severe environments than metals and polymers. In the applications, we study chemical compounds with four levels (ZrO2, ZrO2-TiB2, Si3N4 and glass). The behavior of each level is shown in Figure 1, where the asymmetry and bimodality of the data can be noted.

Cordeiro et al. [9] proposed a family of distributions called *generalized odd log-logistic* family which is very flexible and has been used to analyze and model different sets of data. For example, [10] introduced a new three-parameter exponential distribution called the modified generalized odd log-logistic exponential distribution, [11] proposed the generalized odd log-logistic flexible Weibull regression with applications in repairable systems and [12] defined the generalized

* Corresponding author e-mail: fabio_prataviera@usp.br

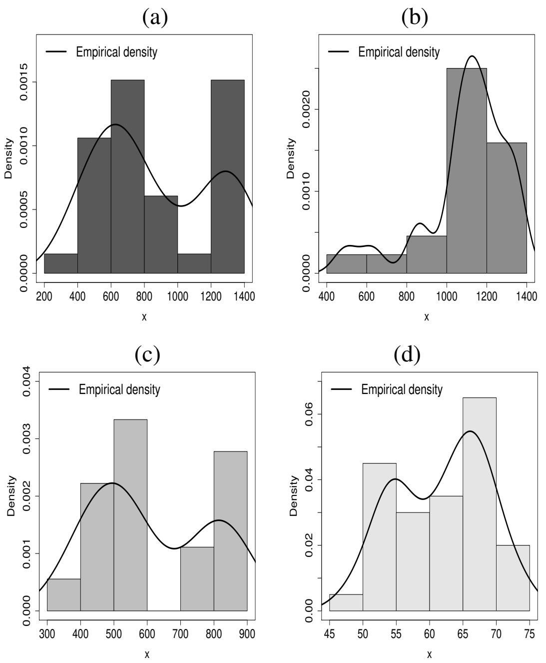

Fig. 1: Histograms and empirical densities. (a) ZrO2. (b) ZrO2-TiB2. (c) Si3N4. (d) Glass.

odd log-logistic exponential distribution based on complete and censored samples. Recently, Prataviera [13] proposed the generalized odd log-logistic log-Weibull regression to modelling non-proportional hazards for survival data and [14] proposed the additive partial linear regressions based on the generalized odd log-logistic log-normal distribution.

In this context, we define a new distribution called the *generalized odd log-logistic Maxwell* ("GOLLMax" for short), whose main advantage related to other competitive distributions as modeling bimodal, asymmetric and heavy tails data. It includes as special cases the Maxwell, exponentiated Maxwell (EMax) and odd log-logistic Maxwell (OLLMax) distributions. Note that the proposed GOLLMax distribution with three parameters has the advantage to model different forms of failure rate, such as increasing, decreasing, unimodal, U-shape and bimodal.We derive some mathematical properties of the proposed distribution. In practice, it is quite common situations where there are some explanatory variables associated with the response random variable. For example, in industry, the failure time of an equipment can be influenced by the voltage level to which the equipment is subjected. In the medical field, a patient's survival time can be related to the type of tumor and the amount of hemoglobin in the blood. In general, we study the effects of these explanatory variables on the response variable by means of a regression model.

In the second part of this paper, we propose a regression model based on the new distribution and present some global influence measures. In addition, we develop residual analysis based on quantile residuals. For different parameter settings and sample sizes, various simulation studies are performed and the empirical distribution of these residuals is compared with the standard normal distribution. The simulation results indicate that the empirical distribution of the quantile residuals is consistent with the standard normal distribution.

The rest of the paper is organized as follows: In Section 1, we introduce the new distribution. We propose the GOLLMax regression model with three systematic structures in Section 2. In Section 3, we evaluate the performance of the maximum likelihood estimators (MLEs) by means of a simulation study. In Section 4, we investigate the case-deletion diagnostic measure and define quantile residuals for the fitted model. In Section 5, we provide three applications to real data to illustrate the flexibility of the new models. Some concluding remarks are offered in Section 6. Finally, we obtain some mathematical properties of the GOLLMax distribution in the Appendix 1.

1 The model definition

The cumulative distribution function (cdf) of the Maxwell random variable W is given by

$$G(w;\mu) = \gamma_1\left(\frac{3}{2},\frac{w^2}{\mu^2}\right),\ x>0, \tag{1}$$

where w denotes the molecule speed, $\mu > 0$ is a scale parameter depending on three quantities (Boltzmann constant, temperature and mass of a molecule), $\gamma_1(p,y) = \gamma(p,y)/\Gamma(p)$ is the incomplete gamma function ratio, $\gamma(p,y) = \int_0^u v^{p-1}\mathrm{e}^{-v}\mathrm{d}v$ is the incomplete gamma function and $\Gamma(\cdot)$ is the gamma function.

The probability density function (pdf) of W is given by

$$g(w;\mu) = \frac{4}{\sqrt{\pi}}\,\frac{w^2}{\mu^3}\exp\left(-\frac{w^2}{\mu^2}\right). \tag{2}$$

The expectation and the variance of W are $E(W) = 2\mu\sqrt{\pi}$ and $Var(W) = (3\pi-8)\mu^2/(2\pi)$.

The idea of the GOLLMax distribution follows the generator. Let $G(x;\gamma)$ be a baseline cdf having a $p\times 1$ vector γ of unknown parameters. [9] defined the cdf of a wider generator called the *generalized odd log-logistic-G* ("GOLL-G") family by integrating the log-logistic density function, namely

$$F(x;\sigma,\nu,\gamma)=\int_0^{\frac{G(x;\gamma)^\sigma}{1-G(x;\gamma)^\sigma}}\frac{\nu w^{\nu-1}}{(1+w^\nu)^2}dw=\frac{G(x;\gamma)^{\sigma\nu}}{G(x;\gamma)^{\sigma\nu}+[1-G(x;\gamma)^\sigma]^\nu}, \tag{3}$$

where $\sigma > 0$ and $\nu > 0$ are two extra shape parameters. Equation (3) includes as special cases the *odd* *log-logistic-G* (OLL-G) family introduced by [15] and the exponentiated-G (exp-G) class when $\sigma = 1$ and $\nu = 1$, respectively. Furthermore, the G distribution is the basic exemplar when $\sigma = \nu = 1$.

We define the cdf of the three-parameter GOLLMax distribution by inserting (1) in equation (3)

$$F(x;\mu,\sigma,\nu)=\frac{\gamma_1^{\sigma\nu}(3/2,x^2/\mu^2)}{\gamma_1^{\sigma\nu}(3/2,x^2/\mu^2)+\left[1-\gamma_1^{\sigma}(3/2,x^2/\mu^2)\right]^{\nu}}. \qquad (4)$$

Henceforth, if X is a random variable with cdf (4), we write $X \sim$GOLLMax(μ,σ,ν).

By differentiating (3) and inserting (1) and (2), we obtain the density function of X (for $x > 0$) as

$$f(x;\mu,\sigma,\nu) = \frac{4\sigma\nu}{\sqrt{\pi}\mu^3} x^2 \exp\left(-\frac{x^2}{\mu^2}\right)\times \frac{\gamma_1^{\sigma\nu-1}(3/2,x^2/\mu^2)\left[1-\gamma_1^{\sigma}(3/2,x^2/\mu^2)\right]^{\nu-1}}{\left\{\gamma_1^{\sigma\nu}(3/2,x^2/\mu^2)+\left[1-\gamma_1^{\sigma}(3/2,x^2/\mu^2)\right]^{\nu}\right\}^2}. \qquad (5)$$

The hazard rate function (hrf) of X is $h(x) = f(x)/[1-F(x)]$. The GOLLMax model contains as special cases the following distributions:

–For $\sigma = 1$, it gives the (new) OLLMax distribution.
–For $\nu = 1$, it yields the (new) EMax distribution.
–The Maxwell distribution is as a basic exemplar when $\sigma = \nu = 1$.

Some plots of the density and hrf of X for selected parameter values, including well-known distributions, are displayed in Figures 2 and 3, respectively. A characteristic of the proposed distribution is that its hrf can be bathtub shaped, monotonically (increasing or decreasing), unimodal, increasing-decreasing-increasing shaped depending basically on the parameter values.

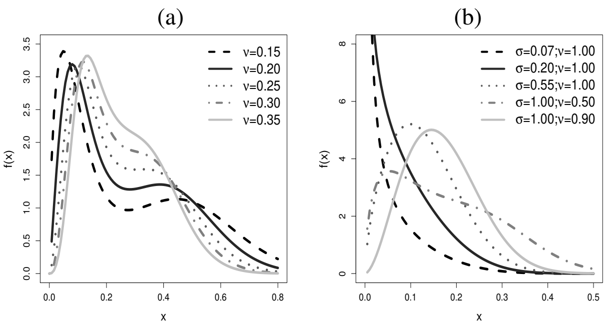

Fig. 2: Plots of the GOLLMax density for some parameter values. (a) For different values of ν with $\sigma = 3.45$ and $\mu = 0.15$. (b) For different values of ν and σ with $\mu = 0.15$.

Equation (4) has tractable properties especially for simulations, since its quantile function (qf) takes the simple form

$$x = Q_{\text{Max}}\left(\varepsilon_{\sigma;\nu}(u);\mu\right), \qquad (6)$$

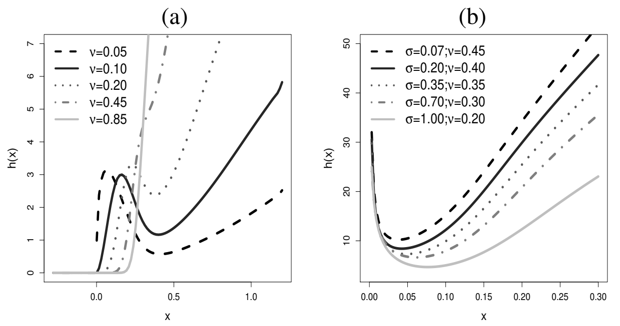

Fig. 3: Plots of the GOLLMax hrf for some parameter values. (a) For different values of ν with $\sigma = 9.00$ and $\mu = 0.20$. (b) For different values of ν and σ with $\mu = 0.07$.

where $Q_{\text{Max}}(\varepsilon_{\sigma;\nu}(u);\mu) = G^{-1}(\varepsilon_{\sigma;\nu}(u);\mu)$ is the qf of the Maxwell distribution and $\varepsilon_{\sigma,\nu}(u) = \left[\frac{u^{1/\nu}}{(1-u)^{1/\nu}+u^{1/\nu}}\right]^{1/\sigma}$.

Further, we can write

$$x = Q_{\text{Max}}(\varepsilon_{\sigma,\nu}(u);\mu) = \mu\sqrt{\gamma_1^{-1}(3/2,\varepsilon_{\sigma,\nu}(u))}, \qquad (7)$$

where $\gamma_1^{-1}(3/2,\varepsilon_{\sigma,\nu}(u))$ is the inverse of the upper gamma regularized function. For more details, see http://functions.wolfram.com/GammaBetaErf/InverseGammaRegularized/.

In Appendix 1, we derive some mathematical properties of the GOLLMax distribution including a linear representation for its density function and explicit expressions for the ordinary and incomplete moments, mean deviations and generating function.

2 The GOLLMax regression model

Regression analysis involves specifications of the distribution of X given a vector $\mathbf{v} = (v_1,\cdots,v_p)^T$ of explanatory variables. In this section, we adopt systematic components for the three parameters in density (5) to allow them to vary across the observations (for $i = 1,\ldots,n$) as

$$g_1(\mu_i) = \mathbf{v}_i^T\beta_1, \quad g_2(\sigma_i) = \mathbf{v}_i^T\beta_2, \quad g_3(\nu_i) = \mathbf{v}_i^T\beta_3, \qquad (8)$$

where $g_k : [0,\infty) \to \mathbb{R}$ for $k = 1,2,3$ are known one-to-one link functions continuously twice differentiables, $\mathbf{v}_i^T = (v_{i1},\cdots,v_{ip})$ is a vector of known explanatory variables for the ith observation, and $\beta_1 = (\beta_{11},\cdots,\beta_{1p})^T$, $\beta_2 = (\beta_{21},\cdots,\beta_{2p})^T$ and $\beta_3 = (\beta_{31},\cdots,\beta_{3p})^T$ are parameter vectors of dimension p. Then, $g_1(\mu) = \mathbf{V}\beta_1$, $g_2(\sigma) = \mathbf{V}\beta_2$, $g_3(\nu) = \mathbf{V}\beta_3$, where $\mu = (\mu_1,\cdots,\mu_n)^T$, $\sigma = (\sigma_1,\cdots,\sigma_n)^T$, $\nu = (\nu_1,\cdots,\nu_n)^T$, and $\mathbf{V} = (\mathbf{v}_1,\cdots,\mathbf{v}_n)^T$ is a specified $n\times p$ matrix of full column rank $p < n$. It is assumed that β_1, β_2 and β_3 are functionally independent. The

GOLLMax regression aims to select the explanatory variables in $\mathbf{V}$ which model μ, σ and ν.

Consider a sample $(x_1, \mathbf{v}_1), \cdots, (x_n, \mathbf{v}_n)$ of n independent observations. Conventional likelihood estimation techniques can be applied here. The total log-likelihood function for the vector of parameters $\psi = (\beta_1^T, \beta_2^T, \beta_3^T)^T$ from model (8) takes the form

$$l(\psi) = n\log\left(\frac{4}{\sqrt{\pi}}\right) + \sum_{i=1}^{n}\log(\nu_i) + \sum_{i=1}^{n}\log(\sigma_i) + \sum_{i=1}^{n}\log\left(\frac{x_i^2}{\mu_i^3}\right) - \sum_{i=1}^{n}\left(\frac{x_i^2}{\mu_i^2}\right) + (\sigma_i\,\nu_i - 1)\sum_{i=1}^{n}\log\left[\gamma_1(3/2, x_i^2/\mu_i^2)\right] + (\nu_i - 1)\sum_{i=1}^{n}\log\left[1 - \gamma_1(3/2, x_i^2/\mu_i^2)\right] - 2\sum_{i=1}^{n}\log\left\{a_i + [1 - b_i]^{\nu_i}\right\}, \quad (9)$$

where $a_i = \gamma_1^{\sigma_i\,\nu_i}(3/2, x_i^2/\mu_i^2)$ and $b_i = \gamma_1^{\sigma_i}(3/2, x_i^2/\mu_i^2)$.

The MLE $\widehat{\psi}$ of ψ can be calculated by maximizing the log-likelihood (9). The numerical maximization of (9) can be done in `R` software (`gamlss` package). Moreover, we can construct likelihood ratio (LR) statistics for comparing some sub-models with the GOLLMax regression model in the classical way.

3 Simulation study for the regression model

We study the performance of the MLEs in the GOLLMax regression model based on 1,000 replications from the true parameters $\beta_{10} = 2.45$, $\beta_{11} = -0.35$, $\beta_{20} = 0.15$, $\beta_{21} = 0.50$, $\beta_{30} = -0.55$ and $\beta_{31} = 0.20$ for different sample sizes ($n = 100, 350, 850$) using `optim` package in `R` software by BFGS method. We consider three different scenarios for the systematic components:

–scenario 1: $\log(\mu_i) = \beta_{10} + \beta_{11}v_{1i}$, $\log(\sigma_i) = \beta_{20} + \beta_{21}v_{1i}$, $\log(\nu_i) = \beta_{30} + \beta_{31}v_{1i}$.

–scenario 2: $\log(\mu_i) = \beta_{10} + \beta_{11}v_{2i}$, $\log(\sigma_i) = \beta_{20} + \beta_{21}v_{2i}$, $\log(\nu_i) = \beta_{30} + \beta_{31}v_{2i}$.

–scenario 3: $\log(\mu_i) = \beta_{10} + \beta_{11}v_{1i}$, $\log(\sigma_i) = \beta_{20} + \beta_{21}v_{2i}$, $\log(\nu_i) = \beta_{30} + \beta_{31}v_{1i}$,

where $v_{1i} \sim \text{Binomial}(1, 0.5)$ by considering two groups (0 and 1) and $v_{2i} \sim \text{Normal}(0, 0.5)$.

The response variables $x_1, \cdots, x_n$ are generated from the GOLLMax regression model (8) as follow:

1. Generate v_{1i} and v_{2i}.
2. Estimate μ_i, σ_i and ν_i for the fixed scenario.
3. Generate $u_i \sim$ U(0,1).
4. Use the steps i., ii. and iii. to calculate the observations x_i's from (6).

Tables 1, 2 and 3 give the average estimates (AEs), biases, mean squared errors (MSEs) of the MLEs, their average lengths (ALs) and the empirical coverage probabilities (CPs), say C(ψ), corresponding to the 95% confidence intervals calculated from the simulations for the parameters ψ=(β_{10}, β_{11}, β_{20}, β_{21}, β_{30}, β_{31})^T. We verify that the AEs tend to be closer to the true parameter values and the MSEs, as well as biases of the sample estimates decay toward zero when the sample size n increases as expected under first-order asymptotic theory. The figures in Tables 1, 2 and 3 indicate that the CPs are close to the nominal level and ALs decrease substantially when n increases, respectively.

Table 1: AEs, Biases, MSEs, ALs and CPs of the parameters for the GOLLMax regression model for scenario 1.

n	ψ	AEs	Biases	MSEs	ALs	C(ψ)
	β_{10}	2.437	-0.013	0.174	1.495	0.898
	β_{11}	-0.314	0.036	0.373	2.334	0.981
100	β_{20}	0.353	0.203	0.847	1.355	0.891
	β_{21}	0.484	-0.016	1.855	5.146	0.918
	β_{30}	-0.636	-0.086	0.413	2.354	0.909
	β_{31}	0.245	0.045	0.817	3.417	0.947
	β_{10}	2.479	0.029	0.067	0.883	0.922
	β_{11}	-0.333	0.017	0.153	1.272	0.970
350	β_{20}	0.140	-0.009	0.320	2.049	0.910
	β_{21}	0.472	-0.028	0.790	3.071	0.913
	β_{30}	-0.525	0.025	0.167	1.454	0.911
	β_{31}	0.224	0.024	0.355	2.035	0.928
	β_{10}	2.452	0.002	0.021	0.545	0.937
	β_{11}	-0.344	0.005	0.044	0.740	0.967
850	β_{20}	0.168	0.018	0.126	1.355	0.935
	β_{21}	0.490	-0.009	0.293	1.995	0.949
	β_{30}	-0.557	-0.007	0.062	0.949	0.930
	β_{31}	0.211	0.011	0.124	1.297	0.952

Due to the difficulty of working analytically with the proposed model, the regularity conditions are verified on the basis of the qq-plots of the sample estimates. In Appendix 2 the Figures 11-16, for n= 850, are presented to better visualize and understand the behavior of the asymptotic distribution of the MLEs. These plots reveal empirically that the asymptotic distributions of the MLEs tend to the normal distribution (as expected) when the sample size increases. This fact asserts that the asymptotic normal distribution provides an adequate approximation to the finite sample distribution of the estimates.

Table 2: AEs, Biases, MSEs, ALs and CPs of the parameters for the GOLLMax regression model for scenario 2.

n	ψ	AEs	Biases	MSEs	ALs	C(ψ)
	β_{10}	2.450	0.000	0.111	1.042	0.880
	β_{11}	-0.297	0.053	0.156	1.205	0.784
100	β_{20}	0.300	0.150	0.532	2.387	0.833
	β_{21}	0.410	-0.090	0.952	2.997	0.769
	β_{30}	-0.604	-0.054	0.258	1.686	0.853
	β_{31}	0.301	0.101	0.458	2.084	0.778
	β_{10}	2.460	0.010	0.030	0.601	0.919
	β_{11}	-0.322	0.028	0.058	0.793	0.861
350	β_{20}	0.164	0.014	0.169	1.459	0.913
	β_{21}	0.453	-0.047	0.357	1.991	0.843
	β_{30}	-0.547	0.003	0.082	1.024	0.916
	β_{31}	0.254	0.054	0.176	1.381	0.848
	β_{10}	2.453	0.003	0.011	0.387	0.935
	β_{11}	-0.326	0.024	0.028	0.569	0.888
850	β_{20}	0.157	0.007	0.068	0.971	0.929
	β_{21}	0.451	-0.049	0.184	1.475	0.879
	β_{30}	-0.550	0.000	0.033	0.677	0.931
	β_{31}	0.241	0.041	0.088	1.013	0.876

Table 3: AEs, Biases, MSEs, ALs and CPs of the parameters for the GOLLMax regression model for scenario 3.

n	ψ	AEs	Biases	MSEs	ALs	C(ψ)
	β_{10}	2.431	-0.019	0.120	1.148	0.881
	β_{11}	-0.337	0.013	0.014	0.430	0.923
100	β_{20}	0.316	0.166	0.614	2.624	0.855
	β_{21}	0.546	0.046	0.102	1.107	0.906
	β_{30}	-0.610	-0.060	0.302	1.878	0.871
	β_{31}	0.215	0.015	0.047	0.757	0.910
	β_{10}	2.436	-0.014	0.027	0.607	0.921
	β_{11}	-0.347	0.003	0.004	0.229	0.947
350	β_{20}	0.213	0.063	0.162	1.493	0.914
	β_{21}	0.509	0.009	0.025	0.571	0.923
	β_{30}	-0.581	-0.031	0.082	1.057	0.915
	β_{31}	0.204	0.004	0.012	0.397	0.937
	β_{10}	2.446	-0.004	0.010	0.387	0.933
	β_{11}	-0.350	0.000	0.001	0.147	0.947
850	β_{20}	0.173	0.023	0.061	0.973	0.936
	β_{21}	0.506	0.006	0.009	0.363	0.924
	β_{30}	-0.559	-0.009	0.032	0.685	0.931
	β_{31}	0.199	-0.001	0.004	0.254	0.951

4 Diagnostic and residual analysis

We adopt diagnostic measures based on case deletion (global influence) to detect influential observations in the proposed regression model. The case-deletion model with systematic structures (8) is

$$g_1(\mu_l) = \mathbf{v}_l^T \beta_1, \quad g_2(\sigma_l) = \mathbf{v}_l^T \beta_2, \quad g_3(\nu_l) = \mathbf{v}_l^T \beta_3 \qquad l = 1, \ldots, n, \quad l \neq i. \tag{10}$$

In the following, a quantity with subscript "(i)" means the original quantity with the ith observation deleted. For model (10), the log-likelihood function for ψ is denoted by $l_{(i)}(\psi)$. Let $\hat{\psi}_{(i)} = (\hat{\beta}_{1(i)}^T, \hat{\beta}_{2(i)}^T, \hat{\beta}_{3(i)}^T)^T$ be the MLE of ψ from $l_{(i)}(\psi)$. To assess the influence of the ith observation on the MLEs $\hat{\psi} = (\hat{\beta}_1^T, \hat{\beta}_2^T, \hat{\beta}_3^T)^T$, we can compare the difference between $\hat{\psi}_{(i)}$ and $\hat{\psi}$. If deletion of an observation seriously influences the estimates, more attention should be paid to that observation. Hence, if $\hat{\psi}_{(i)}$ is far from $\hat{\psi}$, the ith observation can be regarded as influential. A first measure of the global influence is defined as the standardized norm of $\hat{\psi}_{(i)} - \hat{\psi}$ (generalized Cook distance), namely

$$GD_i = (\hat{\psi}_{(i)} - \hat{\psi})^T \left[\ddot{\mathbf{L}}(\hat{\psi})\right] (\hat{\psi}_{(i)} - \hat{\psi}). \tag{11}$$

Another popular measure of the difference between $\hat{\psi}_{(i)}$ and $\hat{\psi}$ is the likelihood distance given by

$$LD_i = 2\left\{ l(\hat{\psi}) - l(\hat{\psi}_{(i)}) \right\}. \tag{12}$$

We can study departures from the error assumption as well as the presence of outliers for various residuals introduced in the literature but we consider the *quantile residuals* (qr's). For the new regression model, they are defined by

$$\hat{qr}_i = \Phi^{-1} \left\{ \frac{\gamma_1^{\hat{\sigma}_i \hat{\nu}_i}(3/2, x_i^2/\hat{\mu}_i^2)}{\gamma_1^{\hat{\sigma}_i \hat{\nu}_i}(3/2, x_i^2/\hat{\mu}_i^2) + \left[1 - \gamma_1^{\hat{\sigma}_i}(3/2, x_i^2/\hat{\mu}_i^2)\right]^{\hat{\nu}_i}} \right\}, \tag{13}$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative standard normal distribution.

We built envelopes to allow better interpretation of the probability normal plot of the residuals. These envelopes are simulated confidence bands described by [16] and contain the residuals, such that if the model is well-fitted, the majority of points will be randomly distributed within these bands.

5 Applications

In this section, we perform three applications in the engineering area and analyze three real data sets. In the first and second applications, we fit the Maxwell, EMax, OLLGMax and GOLLMax (nested models). In addition, we analyze the data sets using non-nested alternative models. We consider the cdf of the gamma-Maxwell (GMax) distribution [7] given by

$$F(x) = \frac{\gamma\left(\delta, -\log\left(1 - \frac{2}{\pi}\gamma\left(\frac{3}{2}, \mu x^2\right)\right)\right)}{\Gamma(\delta)}, \quad x > 0, \tag{14}$$

where $\mu > 0$ is a scale parameter and $\delta > 0$ is a shape parameter.

The cdf of the three-parameter transmuted exponentiated Maxwell (TEMax) distribution [8] is given by

$$F(x) = (1 - \lambda)\, G(x)^{\delta} - \lambda\, G(x)^{2\delta}, \quad x > 0, \tag{15}$$

where $\mu > 0$ is a scale parameter, $\delta > 0$ is a shape parameter, $|\lambda| \leq 1$ is a parameter that makes the asymmetry more flexible and $G(x) = G(x;\mu)$ is the Maxwell cdf given by (1).

The beta-Weibull (BW) distribution (with four positive parameters) $\alpha > 0$, $\lambda > 0$, $a > 0$ and $b > 0$ [17] is (for $x > 0$)

$$F(x) = I_{G_{\alpha,\lambda}(x)}(a,b) = \frac{1}{B(a,b)} \int_0^{G_{\alpha,\lambda}(x)} w^{a-1}(1-w)^{b-1}dw, \quad (16)$$

where $B(a,b) = \Gamma(a+b)/[\Gamma(a)\Gamma(b)]$ is the beta function, $\Gamma(a) = \int_0^\infty z^{a-1}\mathrm{e}^{-z}\mathrm{d}z$ is the gamma function and $I_y(a,b) = B(a,b)^{-1}\int_0^y z^{a-1}(1-z)^{b-1}dz$ is the incomplete beta function ratio.

Finally, the fourth generalization is the Kumaraswamy-Weibull (KwW) distribution (with four positive parameters) defined from [18] family by the cdf

$$F(x) = 1 - \left\{1 - G_{\alpha,\lambda}(x)^a\right\}^b, \quad x > 0. \quad (17)$$

For both BW and KwW distributions,the base cdf $G_{\alpha,\lambda}(x)$ is given by Weibull distribution as

$$G_{\alpha,\lambda}(x) = 1 - \exp\{-(\lambda x)^\alpha\}, \quad \lambda,\alpha,x > 0. \quad (18)$$

In the applications, we define the MLEs and their estimated standard errors (SEs) (given in parentheses) of the model parameters and the values of the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Cramér-von Mises (W^*), Anderson Darling (A^*) and Kolmogorov-Smirnov (KS) statistics for the fitted models. The lower the values of these measures, the better the fit. For all distributions, the parameters are estimated by maximum likelihood. We adopt the `AdequacyModel` script BFGS and CG algorithms for the first two applications, whereas the `gamlss` function by RS method described by [19] in the `R` software is used for the third application.

5.1 Application 1: Strength data

The data set with 49 observations presented by [20] was obtained from a process of manufacturing a plastic laminate whose resistance must exceed a few pounds per square inch (psi). We calculate the MLEs of the model parameters and the above-mentioned statistics for each model fitted to these data. The results are presented in Tables 4 and 5. The five statistical measures are favorable to the GOLLMax distribution, which can be chosen as the best distribution to explain the current data.

Formal tests to verify the inclusion or not of the additional parameters σ and ν in the proposed distribution can be done based on the LR statistics as listed in Table 6. We reject the null hypotheses in the three tests in favor of the GOLLMax distribution. The rejection is significant and provides clear evidence of the flexibility of the shape parameters σ and ν when modeling real data with bimodal characteristics. More information is provided by a visual comparison of the data histogram and fitted density functions. The plots of the fitted GOLLMax, TEMax and Weibull densities are displayed in Figure 4a. The estimated GOLLMax density provides the closest fit to the histogram of the data.

To assess if the model is appropriate, the plots of the fitted GOLLMax, TEMax and Weibull cumulative distributions and the empirical cdf are displayed in Figure 4b. They also indicate that the wider distribution provides a good fit to these data.

Table 4: MLEs of the model parameters, their SEs (given in parentheses) for the strength data.

Model	μ	σ	ν	
GOLLMax	21.416	18.819	0.311	
	(0.001)	(0.002)	(0.037)	
OLLMax	40.603	1	1.216	
	(2.097)	(-)	(0.153	
EMax	36.332	1.561	1	
	(2.768)	(0.351)	(-)	
Maxwell	40.988	1	1	
	(2.390)	(-)	(-)	
Model	μ	δ	λ	
GMax	34.578	1.485		
	(3.502)	(0.305)		
TEMax	38.676	1.626	0.409	
	(0.001)	(0.281)	(0.293)	
Model	λ	α	a	b
BW	0.263	0.673	615.67	0.477
	(0.023)	(0.0005)	(0.013)	(0.001)
KwW	0.106	1.432	92.091	0.164
	(0.001)	(0.002)	(0.003)	(0.002)
Weibull	52.945	2.890	1	1
	(2.777)	(0.304)	(-)	(-)

Table 5: The statistics AIC, BIC, W^*, A^* and KS (p-value associated in parentheses) for the strength data.

Model	AIC	BIC	W^*	A^*	KS
GOLLMax	411.9	417.6	0.038	0.255	0.085
					(0.836)
OLLMax	417.5	421.3	0.120	0.7397	0.094
					(0.732)
EMax	416.2	420.0	0.114	0.700	0.137
					(0.289)
Maxwell	418.0	419.9	0.1214	0.744	0.133
					(0.317)
Model	AIC	BIC	W^*	A^*	KS
GMax	416.6	420.4	0.119	0.733	0.133
					(0.314)
TEMax	417.3	422.9	0.100	0.618	0.131
					(0.334)
Model	AIC	BIC	W^*	A^*	KS
BW	416.5	424.1	0.053	0.357	0.107
					(0.581)
KwW	414.8	422.4	0.040	0.275	0.072
					(0.940)
Weibull	420.1	423.9	0.158	0.968	0.114
					(0.504)

Table 6: LR tests for strength data.

Models	Hypotheses	Statistic w	p-value
GOLLMax vs OLLMax	$H_0 : \sigma = 1$ vs $H_1 : H_0$ is false	7.5	0.006
GOLLMax vs EMax	$H_0 : \nu = 1$ vs $H_1 : H_0$ is false	6.3	0.012
GOLLMax vs Maxwell	$H_0 : \sigma = \nu = 1$ vs $H_1 : H_0$ is false	10.0	0.006

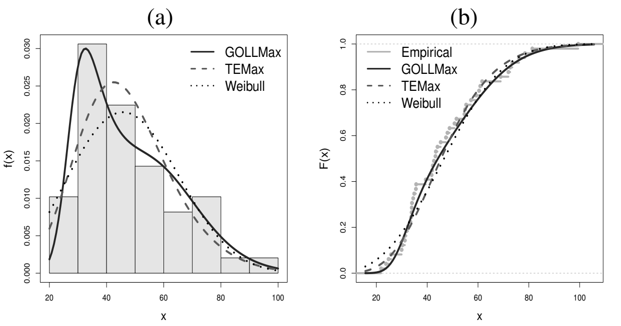

Fig. 4: (a) Estimated densities of the GOLLMax, TEMax and Weibull models for strength data. (b) Estimated cumulative functions of the GOLLMax, TEMax and Weibull models for strength data.

5.2 Application 2: Image data

The data set was extracted from an image of Foulum (Denmark) obtained by the EMISAR sensor Lee and Pottier [21]) jointly built by the Electro Magnetics Institute (EMI), the Technical University of Denmark (TUD), and its Danish Centre for Remote Sensing (DCRS), operated at C- and L-bands (though not simultaneously) with quad-polarizations. The data are retrieved from http://earth.eo.esa.int/polsarpro/datasets.html by means of the PolSARpro software. For each geographic position, each element consists in norm squared of a complex number, which represents the information of the polarization channel resulting of a pulse transmitted and recorded in horizontal direction. A scenario of this data set is presented in [22].

We calculate the MLEs of the model parameters and the above-mentioned statistics for the fitted models to these data. The results are reported in Tables 7 and 8. The five statistics agree with the suitability of the proposed model. In fact, the lowest values of them indicate that the GOLLMax distribution could be chosen as the best model to these data.

Formal tests to verify the inclusion or not of the additional parameters σ and ν in the proposed distribution can be performed based on the LR statistics given in Table 9. We reject the null hypotheses in the three tests in favor of the GOLLMax distribution. The rejection is significant and provides clear evidence of the flexibility of the shape parameters σ and ν for modeling real data with bimodal characteristics. More information is provided by a visual comparison of the data histogram and adjusted densities. Figure 5a displays the plots of the estimated GOLLMax, OLLMax and Weibull densities.

Table 7: MLEs of the model parameters, their SEs (given in parentheses) for the image data.

Model	μ	σ	ν	
GOLLMax	0.065	4.687	0.184	
	(0.013)	(2.205)	(0.075)	
OLLMax	0.110	1	0.517 2	
	(0.006)	(-)	(0.048)	
EMax	0.170	0.422	1	
	(0.0122)	(0.050)	(-)	
Maxwell	0.129	1	1	
	(0.005)	(-)	(-)	
Model	μ	δ	λ	
GMax	0.187	0.422		
	(0.015)	(0.051)		
TEMax	0.174	0.442	0.158	
	(0.017)	(0.061)	(0.304)	
Model	λ	α	a	b
BW	1.479	2.062	0.623	61.586
	(0.001)	(0.0001)	(0.005)	(0.615)
KwW	1.438	5.210	0.284	29.078
	(0.001)	(0.029)	(0.0002)	(0.035)
Weibull	0.145	1.505	1	1
	(0.478)	(0.119)	(-)	(-)

Table 8: The statistics AIC, BIC, W^*, A^* and KS (p-value associated in parentheses) for the image data.

Model	AIC	BIC	W^*	A^*	KS
GOLLMax	-243.0	-235.1	0.064	0.437	0.057
					(0.887)
OLLMax	-233.4	-228.2	0.208	1.176	0.116
					(0.130)
EMax	-226.9	-221.7	0.287	1.614	0.128
					(0.071)
Maxwell	-165.4	-162.8	0.297	1.671	0.307
					(<0.001)
Model	AIC	BIC	W^*	A^*	KS
GMax	-226.4	-221.2	0.293	1.650	0.129
					(0.068)
TEMax	-225.2	-217.3	0.285	1.598	0.123
					(0.093)
Model	AIC	BIC	W^*	A^*	KS
BW	-222.5	-212.1	0.292	1.644	0.128
					(0.072)
KwW	-222.8	-212.35	0.290	1.622	0.119
					(0.111)
Weibull	-226.3	-221.1	0.294	1.648	0.118
					(0.115)

Table 9: LR tests for the image data.

Models	Hypotheses	Statistic w	p-value
GOLLMax vs OLLMax	$H_0: \sigma = 1$ vs H_1 : H_0 is false	11.6	0.001
GOLLMax vs EMax	$H_0: \nu = 1$ vs H_1 : H_0 is false	18.2	<0.001
GOLLMax vs Maxwell	$H_0: \sigma = \nu = 1$ vs $H_1 : H_0$ is false	81.6	<0.001

The estimated GOLLMax density provides the closest fit to the histogram of these data.

To assess if the model is appropriate, the plots of the estimated cdfs of the GOLLMax, OLLMax and Weibull distributions and the empirical cdf are displayed in Figure 5b. They also support that the GOLLMax distribution provides a good fit to these data.

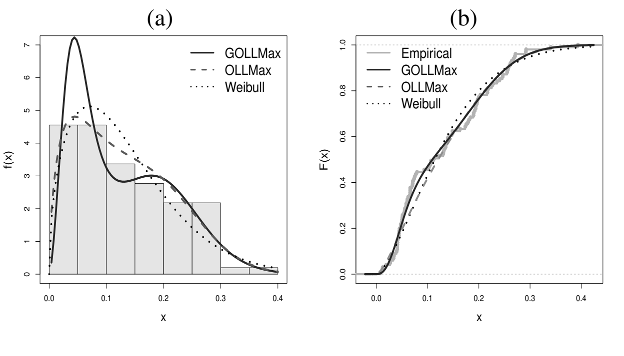

Fig. 5: (a) Estimated densities of the GOLLMax, OLLMax and Weibull models to the image data. (b) Estimated cumulative functions of the GOLLMax, OLLMax and Weibull models to the image data.

5.3 Application 3: Brittle materials

Basu et al. [23] presented a detailed analysis of the data on resistance measurements (in MPa) for different components of ceramic materials and a glass material. In this way, the data set is divided into four subsets with measurements of materials with different chemical compositions, such as, ZrO_2, $ZrO_2 - TiB_2$, Si_3N_4 and Glass (with unknown composition). The interest is to verify the strength properties of solid materials, extremely brittle as glass and the most resistant as the materials to be made of ZrO_2 and Si_3N_4.

The Weibull distribution is an alternative in applications involving resistance studies of brittle materials, for example, in [24]. Basu et al. [23] proposed the Weibull distribution for an alternative modeling and compared it with other two-parameter distributions. However, in the application, the analyses are performed considering the compositions separately. Here, we propose a joint analysis by considering a regression model with three systematic components given by (8).

The explanatory variables are:

– x_i -observed value of the strength of the material (until rupture or crack occurs);
– v_i - chemical compounds of materials with four levels (ZrO_2, $ZrO_2 - TiB_2$, Si_3N_4 and Glass) are defined by dummy variables: ZrO_2 ($v_{i1} = 0$, $v_{i2} = 0$ and $v_{i3} = 0$), $ZrO_2 - TiB_2$ ($v_{i1} = 1$, $v_{i2} = 0$ and $v_{i3} = 0$), Si_3N_4 ($v_{i1} = 0$, $v_{i2} = 1$ and $v_{i3} = 0$) and Glass Si_3N_4 ($v_{i1} = 0$, $v_{i2} = 0$ and $v_{i3} = 1$).

First, we perform an exploratory analysis for these data. We can verify by means of Figure 6 that the different groups of chemical compounds have bimodality and asymmetry. This behavior indicates that a more flexible model, for example, the GOLLMax distribution can be more adequate than the most popular gamma, exponential, log-normal and Weibull distributions.

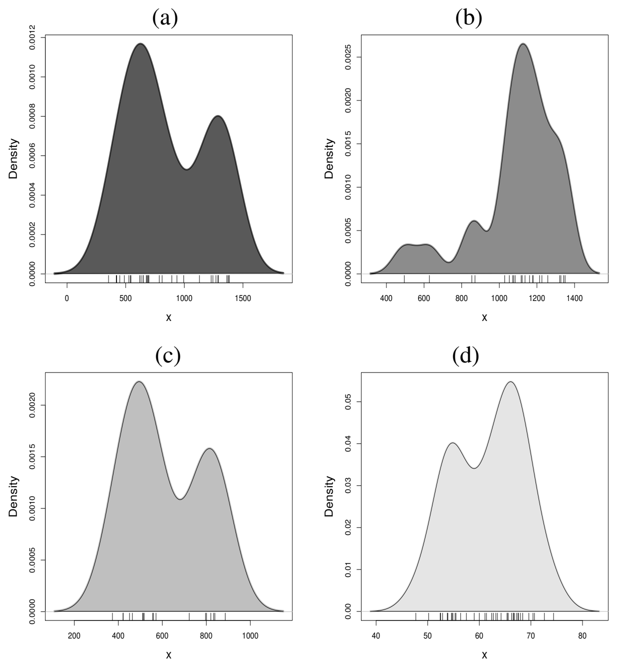

Fig. 6: Plots for brittle materials data. (a) ZrO2. (b) ZrO2-TiB2. (c) Si3N4. (d) Glass.

Considering only the response variable x_i, we verify the suitability of the proposed model and compare it with the Weibull distribution. Tables 10 and 11 give the MLEs and their SEs (in parentheses) and the AIC, BIC, W^*, A^* and KS statistics.

Table 10: MLEs of the model parameters for brittle materials data, their SEs (given in parentheses) and the AIC, BIC, W^*, A^* and KS statistics.

Model	μ	σ	ν
GOLLMax	230.073	3.679	0.099
	(0.029)	(0.017)	(0.007)
	λ	α	
Weibull	591.012	1.032	
	(56.592)	(0.082)	

Table 11: The statistics AIC, BIC, W^*, A^* and KS for brittle materials data

Model	AIC	BIC	W^*	A^*	KS
GOLLMax	1623.6	1631.8	0.541	4.077	0.152
					(0.010)
	AIC	BIC	W^*	A^*	KS
Weibull	1669.5	1674.9	1.080	7.202	0.242
					(<0.001)

Figure 7a displays the estimated GOLLMax and Weibull densities and the histogram to verify which model is more appropriate. As an alternative to check the quality fit, Figure 7b displays the hrfs of the GOLLMax and Weibull models and the empirical hazard function. We conclude that the GOLLMax distribution provides a better fit to these data.

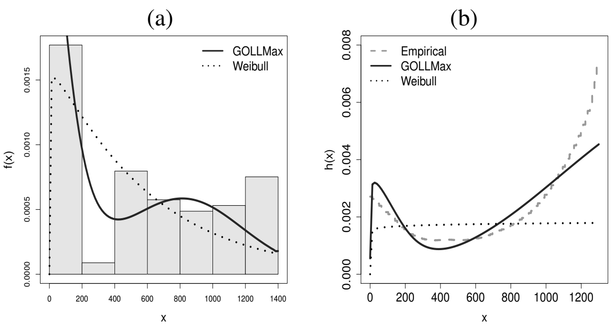

Fig. 7: (a) Estimated densities of the GOLLMax and Weibull models for brittle materials data. (b) Estimated cumulative functions of the GOLLMax and Weibull models for brittle materials data.

We consider the following systematic structures:

$$\mu_i = \exp(\beta_{10} + \beta_{11}v_{i1} + \beta_{12}v_{i2} + \beta_{13}v_{i3}),$$

$$\sigma_i = \exp(\beta_{20} + \beta_{21}v_{i1} + \beta_{22}v_{i2} + \beta_{23}v_{i3})$$

and

$$\nu_i = \exp(\beta_{10} + \beta_{11}v_{i1} + \beta_{12}v_{i2} + \beta_{13}v_{i3}),$$

Table 12: MLEs, SEs and p-values for the fitted GOLLMax regression model to the brittle materials data.

Parameters	Estimates	SEs	p-values
β_{10}	5.9882	0.0673	<0.0001
β_{11}	1.0437	0.0816	<0.0001
β_{12}	-0.4956	0.0857	<0.0001
β_{13}	-2.1062	0.0695	<0.0001
β_{20}	2.6242	0.2220	<0.0001
β_{21}	-2.8636	0.2362	<0.0001
β_{22}	1.4901	0.3797	0.0001
β_{23}	-2.1661	0.2286	<0.0001
β_{30}	-1.2495	0.1784	<0.0001
β_{31}	2.3455	0.2241	<0.0001
β_{32}	0.0570	0.3064	0.8530
β_{33}	2.4431	0.2597	<0.0001

for $i = 1, \ldots, 113$.

In addition, we present the estimates of the parameters, SEs and the associated p-values of the MLEs in Table 12. The figures in this table give evidence that the presence of the covariate v_{i2} (Si_3N_4) is insignificant at a significance level of 5% in the regression structure for the parameter τ_i in relation to the ZrO_2 level. This fact confirms the exploratory analysis shown in Figure 6 in which the groups ZrO_2 in Figure 6a and Si_3N_4 in Figure 6c present similar bimodal forms, which contribute to the non-significance.

Model checking

The next step is to detect possible influential points in the GOLLMax regression model. The generalized Cook distance and likelihood distance are displayed in Figure 8. These plots reveal that the cases 29 and 56 are possible influential observations.

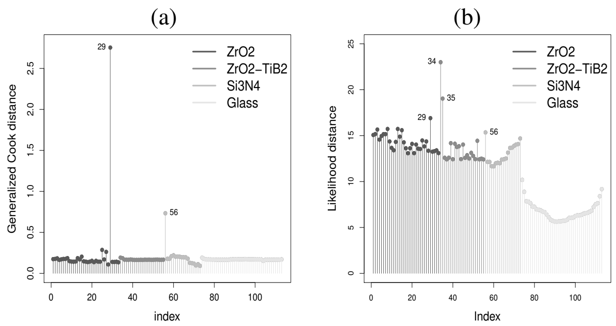

Fig. 8: (a) Generalized Cook distance for the GOLLMax regression model to the brittle materials data. (b) Likelihood distance for the GOLLMax regression model to the brittle materials data.

Furthermore, we verify the quality of the adjustment of the GOLLMax regression by constructing the normal probability plot for the qr's for the waste diversion with simulated envelope. There is evidence of a good fit of the GOLLMax regression model as illustrated in Figures 9a and 9b.

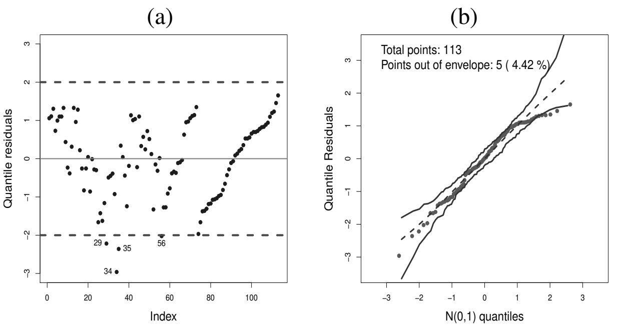

Fig. 9: (a) Index plot of the qr's from the fitted GOLLMax regression model to brittle materials data. (b) Normal probability plot for the qr's with envelopes.

In this analysis, we make an analogy with the methodology of survival analysis. We consider that the event of interest is the breaking or rupture of the material.

The survival function corresponding to (5) is

$$S(x) = 1 - \frac{\gamma_1^{\sigma\nu}(3/2, x^2/\mu^2)}{\gamma_1^{\sigma\nu}(3/2, x^2/\mu^2) + \left[1 - \gamma_1^{\sigma}(3/2, x^2/\mu^2)\right]^{\nu}} \quad (19)$$

Table 12 suggests that the materials $ZrO_2 - TiB_2$, Si_3N_4 and glass are statistically different from the ZrO_2 material in all structures, except as mentioned above. This fact reveals the modeling ability of the proposed structure to model the scale of the data using the parameter μ_i and asymmetry and bimodality through the parameters σ_i and ν_i, respectively.

In Figure 10a, we display the Kaplan-Meier empirical curves and estimated survival functions defined from (8). Note that glass is the weakest material with least resistance. For a strength of 800 MPa the probabilities that in the materials ZrO_2, $ZrO_2 - TiB_2$ and Si_3N_4 do not occur the event of interest are 0.4912, 0.9446 and 0.1633, respectively. In Figure 10b, we give the estimated hrf for each material and verify the presence of different forms for this function. Based on this figure, we note that both models show satisfactory fits. However, the GOLLMax regression model presents a better fit to the current data.

6 Concluding Remarks

We proposed a three-parameter model called the *generalized odd log-logistic Maxwell* (GOLLMax) distribution, which includes as special cases the odd log-logistic Maxwell (OLLMax), exponentiated Maxwell (EMax) and Maxwell distributions. We provide some of its mathematical properties. We define a GOLLMax regression model with three systematic components based on the new distribution. The proposed regression serves as an important extension to several existing regression models and could be a valuable addition to the literature. The maximum likelihood method is described for estimating the model parameters. Some simulations are performed for different parameter settings and sample sizes to evaluate the precision of the maximum likelihood estimates. Diagnostic analysis is presented to assess global influences. We also discuss the sensitivity of the estimates from the fitted model via quantile residuals. The utility of the introduced models was discussed by means of three real data sets.

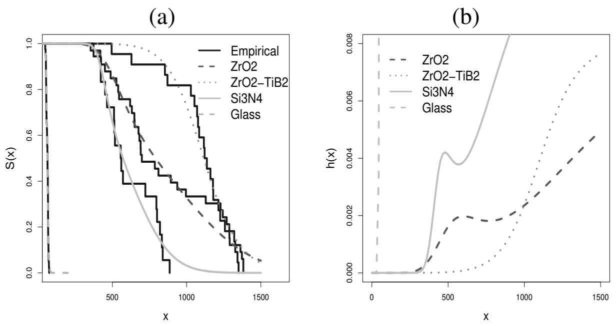

Fig. 10: The GOLLMax regression model. (a) Estimated survival functions and the empirical survival for the brittle materials data. (b) Estimated hrf for the brittle materials data.

Acknowledgment

This paper was supported by CNPq and CAPES, Brazil.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article

Appendix 1: Structural properties

We derive here some mathematical properties of $X \sim \text{GOLLMax}(\mu, \sigma, \nu)$. First, we introduce some notation. Let $\Pi(x; \mu, k, \beta) = \gamma_1\left(k, [\frac{x}{\mu}]^{\beta}\right)$ (for $x > 0$) be the cdf of the *generalized gamma* (GG) distribution [25] with shape parameters $k > 0$ and $\beta > 0$ and scale parameter $\mu > 0$. The corresponding density function, say $\pi(x;\mu,k,\beta)$, is

$$\pi(x,\mu,k,\beta) = \frac{\beta}{\mu\,\Gamma(k)}\left(\frac{x}{\mu}\right)^{k\beta-1}\exp\left[-\left(\frac{x}{\mu}\right)^{\beta}\right]. \quad (20)$$

Clearly, the Maxwell density (2) follows from (20) as $g(x;\mu) = \pi(x;\mu,3/2,2)$.

Theorem 1: We can express the density $f(x;\mu,\sigma,\nu)$ of X as a linear combination of GG densities

$$f(x;\mu,\sigma,\nu) = \sum_{m,i=0}^{\infty} p_{m,i}\,\pi(x;\mu,k^*,2), \quad (21)$$

where the GG densities have common parameters μ and 2 and varying shape parameter $k^* = k^*(m,i) = 3(m+1)/2 + i$, and the coefficients $p_{m,i} = p_{m,i}(\mu,\sigma,\nu)$ are functions of the model quantities (defined below) given by

$$p_{m,i} = \frac{2(m+1)\,\mu^{3m+2i}\,w_{m+1}\,c_{m,i}}{\sqrt{\pi}}\,\Gamma\left(\frac{3m}{2}+i\right).$$

Proof Theorem 1:

For any real non-integer $\lambda > 0$, the generalized binomial theorem

$$\gamma_1(3/2,x^2/\mu^2)^{\lambda} = \left[1-\left\{1-\gamma_1(3/2,x^2/\mu^2)\right\}\right]^{\lambda} = \sum_{j=0}^{\infty}(-1)^j\binom{\lambda}{j}\left\{1-\gamma_1(3/2,x^2/\mu^2)\right\}^j,$$

holds, and it is always convergent since $0 < \gamma_1(3/2,x^2/\mu^2) < 1$. Hence,

$$\gamma_1(3/2,x^2/\mu^2)^{\lambda} = \sum_{j=0}^{\infty}\sum_{m=0}^{j}(-1)^{j+m}\binom{\lambda}{j}\binom{j}{m}\gamma_1(3/2,x^2/\mu^2)^m.$$

Substituting $\sum_{j=0}^{\infty}\sum_{m=0}^{j}$ for $\sum_{m=0}^{\infty}\sum_{j=m}^{\infty}$ and after some algebra, we obtain

$$\gamma_1(3/2,x^2/\mu^2)^{\lambda} = \sum_{m=0}^{\infty} s_m(\lambda)\,\gamma_1(3/2,x^2/\mu^2)^m, \quad (22)$$

where (for $m \geq 0$)

$$s_m(\lambda) = \sum_{j=m}^{\infty}(-1)^{r+m}\binom{\lambda}{j}\binom{j}{m}.$$

Using (22), the numerator of (4) can be expanded as

$$\gamma_1^{\sigma\nu}(3/2,x^2/\mu^2) = \sum_{m=0}^{\infty} s_m(\sigma\nu)\,\gamma_1(3/2,x^2/\mu^2)^m \quad (23)$$

where $s_m(\sigma\nu)$ comes from a previous quantity. Using the generalized binomial theorem and (22), one part of the denominator (4) can be written as

$$\left[1-\gamma_1^{\sigma}(3/2,x^2/\mu^2)\right]^{\nu} = 1+\sum_{j=1}^{\infty}(-1)^j\binom{\nu}{j}\gamma_1^{j\sigma}(3/2,x^2/\mu^2)$$
$$= 1+\sum_{m=0}^{\infty} t_m(\sigma,\nu)\,\gamma_1(3/2,x^2/\mu^2)^m,$$

where $t_m(\sigma,\nu) = \sum_{j=1}^{\infty}(-1)^j\binom{\nu}{j}s_m(j\sigma)$.

The denominator of (4) can be defined from (23) and the last power series as

$$\gamma_1^{\sigma\nu}(3/2,x^2/\mu^2)+\left[1-\gamma_1^{\sigma}(3/2,x^2/\mu^2)\right]^{\nu} = \sum_{m=0}^{\infty} v_m(\sigma\nu)\,\gamma_1(3/2,x^2/\mu^2)^m, \quad (24)$$

where $v_0(\sigma,\nu) = 1+s_0(\sigma\nu)+t_0(\sigma,\nu)$ and $v_m(\sigma,\nu) = s_m(\sigma\nu)+t_m(\sigma,\nu)$ for $m \geq 1$.

Combining (23) and (24), we can express (4) as

$$F(x;\mu,\sigma,\nu) = \frac{\sum_{m=0}^{\infty} s_m(\sigma\nu)\,\gamma_1(3/2,x^2/\mu^2)^m}{\sum_{m=0}^{\infty} v_m(\sigma,\nu)\,\gamma_1(3/2,x^2/\mu^2)^m}.$$

The ratio of the two power series in the last equation reduces to

$$F(x;\mu,\sigma,\nu) = \sum_{m=0}^{\infty} w_m\,\gamma_1(3/2,x^2/\mu^2)^m, \quad (25)$$

where the coefficients w_m's (for $m \geq 1$) are determined from the recurrence equation

$$w_m = w_m(\sigma,\nu) = v_0(\sigma,\nu)^{-1}\left[s_m(\sigma\nu)-\sum_{r=1}^{m} v_r(\sigma,\nu)\,w_{m-r}(\sigma,\nu)\right]$$

and $w_0 = w_0(\sigma,\nu) = s_0(\sigma\nu)/v_0(\sigma,\nu)$.

Differentiating (25), we can rewrite (5) as

$$f(x;\mu,\sigma,\nu) = \sum_{m=0}^{\infty} w_{m+1}\,h_{m+1}(x;\mu), \quad (26)$$

where $h_{m+1}(x;\mu) = (m+1)g(x;\mu)\,\gamma_1(3/2,x^2/\mu^2)^m$ is the EMax density with power parameter $(m+1)$ (for $m \geq 0$). Then, we have

$$h_{m+1}(x;\mu) = \frac{4\,(m+1)}{\sqrt{\pi}\,\mu^3}\,x^2\exp\left(-\frac{x^2}{\mu^2}\right)\gamma_1(3/2,x^2/\mu^2)^m.$$

In addition, we adopt the power series for the incomplete gamma function ratio

$$\gamma_1(3/2,x^2/\mu^2) = \sum_{i=0}^{\infty} a_i\,x^{2i+3}, \quad (27)$$

where $a_i = a_i(\mu) = \frac{(-1)^i}{(3/2+i)\,\mu^{2i+3}\,\Gamma(3/2)\,i!}$ (for $i \geq 0$).

By application of an equation in Section 0.314 of [26] for power series raised to powers, we can write (for any m positive integer)

$$\left(\sum_{i=0}^{\infty} a_i\,z^i\right)^m = \sum_{i=0}^{\infty} c_{m,i}\,z^i, \quad (28)$$

where the coefficients $c_{m,i} = c_{m,i}(\mu)$ are determined by $c_{m,0} = a_0^m$ and, for $i = 1,2,\ldots$, from the recurrence relation

$$c_{m,i} = (i\,a_0)^{-1}\sum_{r=1}^{i}\left[(m+1)r-i\right]a_r\,c_{m,i-r}.$$

Thus, the coefficient $c_{m,i}$ can follow from $c_{m,0},\cdots,c_{m,i-1}$ and then from $a_0,\cdots,a_j$.

Furthermore, we have from equation (28)

$$\gamma_1(3/2,x^2/\mu^2)^m = \sum_{i=0}^{\infty} c_{m,i}\, x^{3m+2i}. \quad (29)$$

Combining (22) and (29), the EMax density can be expanded as

$$h_{m+1}(x;\mu) = \frac{4}{\sqrt{\pi}\,\mu^3} \sum_{i=0}^{\infty} (m+1)\, c_{m,i}\, x^{3m+2i+2} \exp\left(-\frac{x^2}{\mu^2}\right).$$

Inserting this expression in Equation (26) gives

$$f(x;\mu,\sigma,\nu)=\frac{4}{\sqrt{\pi}\mu^3}\sum_{m,i=0}^{\infty}(m+1)\, w_{m+1}\, c_{m,i}\, x^{3m+2i+2} \exp\left(-\frac{x^2}{\mu^2}\right).$$

Finally, we can rewrite $f(x;\mu,\sigma,\nu)$ as a linear combination of GG densities with two common parameters 2 and μ and the third parameter $k^* = k^*(m,i) = 3(m+1)/2+i$ as in (21), and the coefficients $p_{m,i} = p_{m,i}(\mu,\sigma,\nu)$ are functions of previous quantities given by

$$p_{m,i} = \frac{2(m+1)\,\mu^{3m+2i}}{\sqrt{\pi}}\,\Gamma(3m/2+i)\, w_{m+1}\, c_{m,i}. \quad \blacksquare$$

The linear representation (21) becomes very useful in deriving some mathematical properties for the GOLLMax distribution using well-known GG properties. We can adopt at most ten terms in (21) to provide accurate results in most analytical platforms.

Corollary 1 The nth moment of X takes the form

$$\mu_n' = E(X^n) = \mu^n \sum_{m,i=0}^{\infty} p_{m,i}\, \frac{\Gamma(3[m+1]/2+i+n/2)}{\Gamma(3[m+1]/2+i)}. \quad (30)$$

Proof Corollary 1:

The nth ordinary moment of the GG pdf $\pi(x;\mu,k,\beta)$ is known to be $\delta'_{n,GG} = \mu^n\,\Gamma(k+n/\beta)/\Gamma(k)$. Then, the nth moment of X can be defined from (21) as

$$\mu_n' = E(X^n) = \mu^n \sum_{m,i=0}^{\infty} p_{m,i}\, \frac{\Gamma(3[m+1]/2+i+n/2)}{\Gamma(3[m+1]/2+i)}. \quad \blacksquare \quad (31)$$

Some of the most important features and characteristics of a distribution can be investigated through moments (e.g. tendency, dispersion, skewness and kurtosis). The central moments and cumulants of X can be determined from the ordinary moments in (30) using well-known formulae.

Corollary 2 The nth incomplete of X can be expressed as

$$M_n(s) = \int_0^s x^n f(x;\mu,\sigma,\nu)\, dx = \sum_{m,i=0}^{\infty} p_{m,i}\, I_n(s;\mu,k^*,2).$$

Proof Corollary 2:

The nth incomplete moment $I_n(s;\mu,k,2) = \int_0^s x^n\,\pi(x;\mu,k,2)dx$ of $\pi(x;\mu,k,2)$ is easily found by transforming variables $z = (t/\mu)^2$ as

$$I_n(s;\mu,k,2) = \frac{\mu^n}{\Gamma(k)}\,\gamma(n/2+k,(s/\mu)^2).$$

Hence, the nth incomplete of X follows from Equation (21) as

$$M_n(s) = \int_0^s x^n f(x;\mu,\sigma,\nu)\, dx = \sum_{m,i=0}^{\infty} p_{m,i}\, I_n(s;\mu,k^*,2). \quad \blacksquare$$

The first incomplete moment $M_1(s)$ plays an important role in measuring inequality such as the mean deviations and Lorenz and Bonferroni curves. First, the mean deviations about the mean $\tau_1' = E(X)$ and the median m of X are determined from $\delta_1 = 2\tau_1' F(\tau_1') - 2M_1(\tau_1')$ and $\delta_2 = \tau_1' - 2M_1(m)$, where $F(\tau_1')$ and $F(m)$ are easily calculated from (4).

Another application of $M_1(s)$ refers to the Bonferroni and Lorenz curves of X. These curves are very useful in economics, reliability, demography, insurance and medicine. For a given probability π, the Bonferroni and Lorenz curves are given by $B(\pi) = M_1(p)/(p\tau_1')$ and $L(p) = M_1(p)/\tau_1'$, where $p = Q(\pi) = F^{-1}(\pi)$ can be computed from (6).

Corollary 3 The moment generating function (mgf) of X can be reduced to

$$M(s) = \sum_{m,i=0}^{\infty} p_{m,i}\, M_{\mu,k^*,2}(s),$$

where $M_{\mu,k^*,2}(s)$ is the mgf of $\pi(x;\mu,k^*,2)$.

Proof Corollary 3:

The mgf of $\pi(x;\mu,k^*,2)$ follows from [20] as

$$M_{\mu,k^*,2}(s) = \frac{1}{\Gamma(k^*)}\, {}_1\Psi_0\left[\begin{matrix}(1,1/2)\\ -\end{matrix};\mu\, s\right], \quad (32)$$

where ${}_1\Psi_0$ is the Wright generalized hypergeometric function defined by

$${}_p\Psi_q\left[\begin{matrix}(\mu_1,A_1),\cdots,(\mu_p,A_p)\\ (\beta_1,B_1),\cdots,(\beta_q,B_q)\end{matrix};x\right]=\sum_{m=0}^{\infty}\frac{\prod_{j=1}^{p}\Gamma(\mu_j+A_j m)}{\prod_{j=1}^{q}\Gamma(\beta_j+B_j m)}\frac{x^m}{m!}.$$

Hence, the mgf of X can be determined from (21) and (32) as

$$M(s)=E(e^{sX})={}_1\Psi_0\left[\begin{matrix}(1,1/2)\\ -\end{matrix};\mu\, s\right]\sum_{m,i=0}^{\infty}\frac{p_{m,i}}{\Gamma\left(3[m+1]/2+i\right)}. \quad \blacksquare$$

Appendix 2: Plots for checking the regularity conditions

References

[1] I. Prigogine and E. Xhrouet. On the perturbation of Maxwell distribution function by chemical reactions in gases, *Physica*, **15**, 913-932, (1949).

[2] T. Pöschel and N.V. Brilliantov. *Granular gas dynamics*, Springer Science & Business Media, (2003).

[3] H. Krishna and M. Malik. Reliability estimation in Maxwell distribution with progressively type-II censored data, *Journal of Statistical Computation and Simulation*, **82**, 623-641, (2012).

[4] S. Kazmi, M. Aslam and S. Ali. On the Bayesian estimation for two component mixture of Maxwell distribution, assuming type I censored data, *SOURCE International Journal of Applied Science & Technology*, **2**, 197-218, (2012).

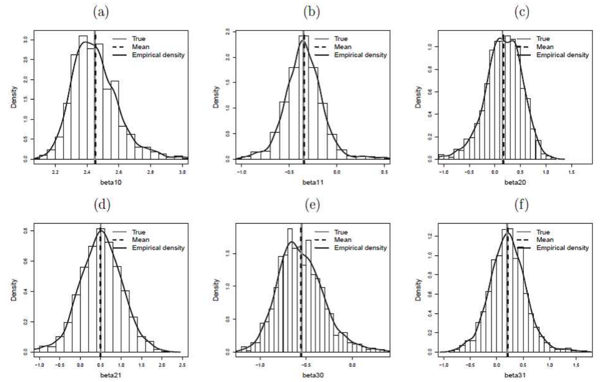

Fig. 11: Plots of the empirical distributions of the 1,000 parameter estimates for $n = 850$ for scenario 1. (a) For $\hat{\beta}_{10}$. (b) For $\hat{\beta}_{11}$. (c) For $\hat{\beta}_{20}$. (d) For $\hat{\beta}_{21}$. (e) For $\hat{\beta}_{30}$. (f) For $\hat{\beta}_{31}$.

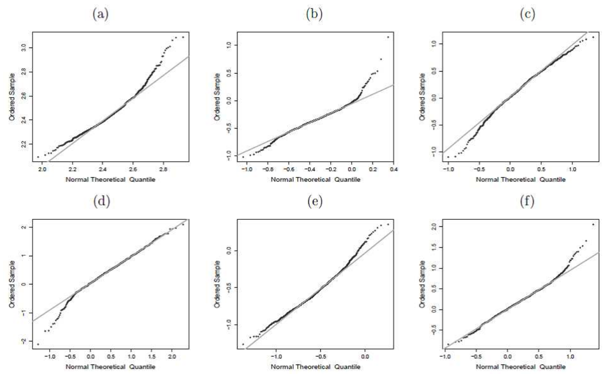

Fig. 12: Normal probability plots when $n = 850$ for scenario 1. (a) For $\hat{\beta}_{10}$. (b) For $\hat{\beta}_{11}$. (c) For $\hat{\beta}_{20}$. (d) For $\hat{\beta}_{21}$. (e) For $\hat{\beta}_{30}$. (f) For $\hat{\beta}_{31}$.

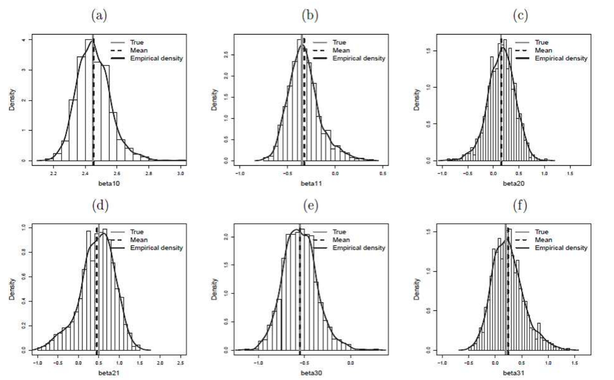

Fig. 13: Plots of the empirical distributions of the 1,000 parameter estimates for $n = 850$ for scenario 1. (a) For $\hat{\beta}_{10}$. (b) For $\hat{\beta}_{11}$. (c) For $\hat{\beta}_{20}$. (d) For $\hat{\beta}_{21}$. (e) For $\hat{\beta}_{30}$. (f) For $\hat{\beta}_{31}$.

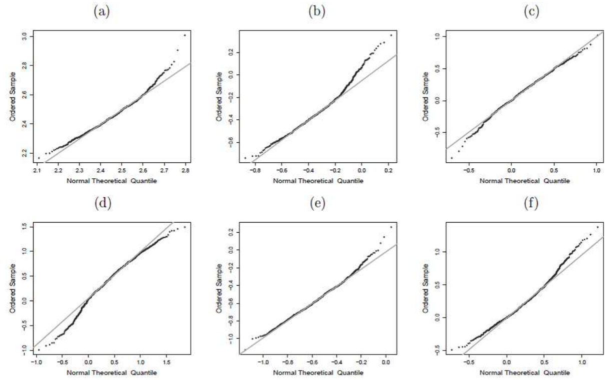

Fig. 14: Normal probability plots when $n = 850$ for scenario 1. (a) For $\hat{\beta}_{10}$. (b) For $\hat{\beta}_{11}$. (c) For $\hat{\beta}_{20}$. (d) For $\hat{\beta}_{21}$. (e) For $\hat{\beta}_{30}$. (f) For $\hat{\beta}_{31}$.

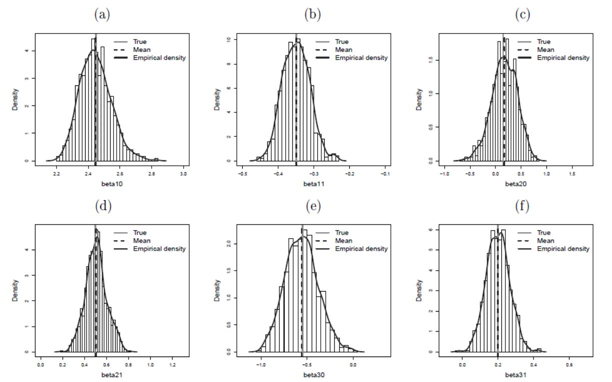

Fig. 15: Plots of the empirical distributions of the 1,000 parameter estimates for $n = 850$ for scenario 1. (a) For $\hat{\beta}_{10}$. (b) For $\hat{\beta}_{11}$. (c) For $\hat{\beta}_{20}$. (d) For $\hat{\beta}_{21}$. (e) For $\hat{\beta}_{30}$. (f) For $\hat{\beta}_{31}$.

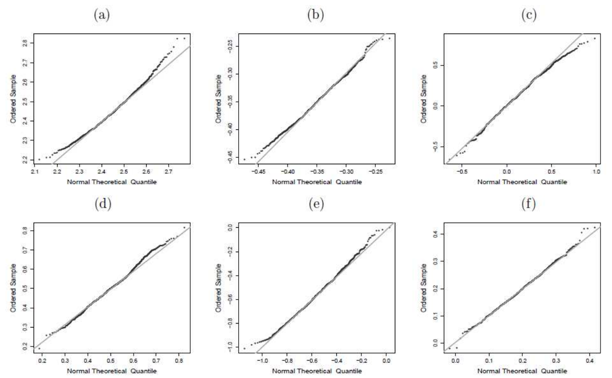

Fig. 16: Normal probability plots when $n = 850$ for scenario 1. (a) For $\hat{\beta}_{10}$. (b) For $\hat{\beta}_{11}$. (c) For $\hat{\beta}_{20}$. (d) For $\hat{\beta}_{21}$. (e) For $\hat{\beta}_{30}$. (f) For $\hat{\beta}_{31}$.

[5] S.K. Tomer and M.S. Panwar. Estimation procedures for Maxwell distribution under type-I progressive hybrid censoring scheme, *Journal of Statistical Computation and Simulation*, **85**, 339-356, (2015).
[6] S. Dey, T. Dey, S. Ali and M.S. Mulekar. Two-parameter Maxwell distribution: Properties and different methods of estimation, *Journal of Statistical Theory and Practice*, **10**, 291-310, (2016).
[7] Y.A. Iriarte, J.M. Astorga, H. Bolfarine and H.W. Gómez. Gamma-Maxwell distribution, *Communications in Statistics-Theory and Methods*, **46**, 4264-4274, (2017).
[8] O. Venegas, Y.A. Iriarte, J.M. Astorga, A. Borger, H. Bolfarine and H.W. Gómez. A New Generalization of the Maxwell Distribution, *Applied Mathematics & Information Sciences*, **11**, 867-876, (2017).
[9] G.M. Cordeiro, M. Alizadeh, G. Ozel, B. Hosseini, E.M.M. Ortega and E. Altun. The generalized odd log-logistic family of distributions: properties, regression models and applications, *Journal of Statistical Computation and Simulation*, **87**, 908-932, (2017).
[10] A. Qoshja and M. Muça. A new modified generalized odd log-logistic distribution with three parameters, *Mathematical Theory and Modeling*, **8**, 49-60, (2018).
[11] F. Prataviera, E.M.M., Ortega, G.M. Cordeiro, R.R. Pescim and B.A.W. Verssani. A new generalized odd log-logistic flexible Weibull regression model with applications in repairable systems, *Reliability Engineering & System Safety*, **176**, 13-26, (2019).
[12] A.Z. Afify, A.K. Suzuki, C. Zhang and M. Nassar. On three-parameter exponential distribution: properties, Bayesian and non-Bayesian estimation based on complete and censored samples, *Communication in Statitics - Simulation and Computation*, in Proc. 10.1080/03610918.2019.1636995, 1-21, (2019).
[13] F. Prataviera, S.M.C. Loibel, K.F. Grego, E.M.M. Ortega and G.M. Cordeiro. Modelling non-proportional hazard for survival data with different systematic components, *Environmental and Ecological Statistics*, in Proc. 10.1007/s10651-020-00453-5, 1-23, (2020).
[14] J.C.S. Vasconcelos, G.M. Cordeiro, E.M.M. Ortega and M.A. Biaggioni. The parametric and additive partial linear regressions based on the generalized odd loglogistic log-normal distribution, *Communications in Statistics - Theory and Methods*, in Poc. 10.1080/03610926.2020.1795681, 1-28, (2020).
[15] J.U. Gleaton and J.D. Lynch. Properties of generalized log-logistic families of lifetime distributions, *Journal of Probability and Statistical Science*, **4**, 51-64, (2006).
[16] A.C. Atkinson. *Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis*, Oxford University Press, New York, (1985).
[17] C. Lee, F. Famoye and O. Olumolade. Beta-Weibull distribution: some properties and applications to censored data, *Journal of Modern Applied Statistical Methods*, **6**, 17, (2007).
[18] G.M. Cordeiro, E.M.M. Ortega and S. Nadarajah. The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*, **347**, 1399-1429 (2010).
[19] R.A. Rigby and D.M. Stasinopoulos. Generalized additive models for location, scale and shape, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**, 507-554, (2005).
[20] C. Chen. Tests of fit for the three-parameter log-normal distribution, *Computational Statistics and Data Analysis*, **50**, 1418-1440, (2006).
[21] J.S. Lee and E. Pottier. *Polarimetric radar imaging: from basics to applications*, CRC press, (2009).
[22] M. Alizadeh, G.M. Cordeiro, A.D. Nascimento, M.D.C.S. Lima and E.M.M. Ortega. Odd Burr generalized family of distributions with some applications, *Journal of Statistical Computation and Simulation*, **87**, 367-389, (2017).
[23] B. Basu, D. Tiwari, D. Kundu and R. Prasad. Is Weibull distribution the most appropriate statistical strength distribution for brittle materials?, *Ceramics International*, **35**, 237-246, (2009).
[24] Y. Xu, L. Cheng, L. Zhang, D. Yan and C. You. Optimization of sample number for Weibull function of brittle materials strength, *Ceramics International*, **27**, 239-241, (2001).
[25] E.W. Stacy. A generalization of the gamma distribution, *Annals of Mathematical Statistics*, **33**, 1187-1192, (1962).
[26] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series and Products*, Academic Press, San Diego, (2000).

Fábio Prataviera a is PhD in Statistic and Agronomic Experimentation from University of São Paulo, ESALQ-USP, Brazil (2020). His research interests are parametric and semiparametric regression models, distribution theory, engineering reliability and survival analysis.

Edwin M. M. Ortega is PhD in Statistics from, University of São Paulo (2002). He has a postdoctoral degree from the Federal University of Pernambuco (2010-2011). He is currently a full professor at the Department of Exact Sciences at the University of São Paulo, ESALQ-USP, Brazil. He has published 205 papers in international statistics journals with peer review systems and has supervised master and doctoral students. He works on regression models, semiparametric regression models, distribution theory and survival analysis. He has worked as a reviewer for several scientific journals since 2005.

Gauss M. Cordeiro is PhD in Statistics from Imperial College, University of London (1982). He has a postdoctoral degree from the University of London (1986-1987) and the Institute of Pure and Applied Mathematics (1990-1992). He is currently a full professor at the Department of Statistics at the Federal University of Pernambuco. He has published 433 articles in international statistics journals with peer review systems and has supervised 59 master and doctoral students. He works on distribution theory and asymptotic theory. He has worked as a reviewer for several scientific journals since 1983. He was President of the Brazilian Statistical Association (ABE) from 2000 to 2002 and member of the Mathematical Advisory Committee of the National Council for Scientific and Technological Development in four different periods. He was editor of the Brazilian Journal of Probability and Statistics (1995-2000) and received the ABE-2008 Prize for outstanding services to Statistics and, in 2010, the degree of Commander of the Order of Scientific Merit.