



**HAL**  
open science

# Estimating the parameters of parametric lifetime distributions through an efficient acceptance–rejection sampler

Anis Ben Abdessalem

► **To cite this version:**

Anis Ben Abdessalem. Estimating the parameters of parametric lifetime distributions through an efficient acceptance–rejection sampler. *Engineering Applications of Artificial Intelligence*, 2021, 106, pp.104457. 10.1016/j.engappai.2021.104457 . hal-04395977

**HAL Id: hal-04395977**

**<https://univ-angers.hal.science/hal-04395977>**

Submitted on 22 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Estimating the parameters of parametric lifetime distributions through an efficient acceptance-rejection sampler

Anis Ben Abdesslem<sup>a,\*</sup>

<sup>a</sup>Univ Angers, LARIS, SFR MATHSTIC, F-49000 Angers, France

---

## Abstract

The three-parameter (3-p) Weibull distribution is an extremely important distribution to characterise the statistical behaviour of a large number of real world phenomena. It is also useful as a failure model in analysing the reliability of different types of mechanical and electrical components/systems. Successful applications of the distribution rely on an accurate estimation of its three parameters because it directly affects the reliability and lifetime analysis. Due to the intricate system of nonlinear equations and the complexity of the likelihood function, derivative-based optimisation methods may fail to converge. Thus, an efficient and effective method for estimating the parameters of the model is important from the practical viewpoint. In this paper, an optimisation scheme based on an acceptance-rejection (AR) mechanism coupled with an elegant *nested sampling* (NS) technique is proposed to tackle this problem. The idea is to gradually approach the region of optimal solutions through an efficient sampling technique and a reweighting scheme. The AR-NS algorithm allows a good exploration of the parameter space and converges towards higher likelihood regions by decreasing progressively a pre-specified tolerance threshold. The proposed approach gives the entire distributions of the optimal estimates rather than a single point estimates. To demonstrate the practicality and the efficiency of the proposed approach, numerous numerical examples using simulated data and real-world engineering cases will be given. The obtained results show that the AR-NS algorithm is a suitable method for estimating the parameters of lifetime distributions using different distances.

**Keywords:** Weibull distribution, Likelihood function, Wasserstein distance, Acceptance-rejection sampler, Parameter estimation, Nested sampling, Failure times

---

## 1. Introduction

Weibull family distributions contains a large number of models with different degrees of complexity which have been broadly applied in reliability engineering and lifetime studies [1–12]. The three-parameter (3-p) Weibull distribution is a member of this family which is extremely flexible and can fit very well an extremely wide range of empirical observations. Also, it is especially useful as a failure model in analysing the reliability of different types

---

\*Corresponding author

Email address: [anis.ben-abdesslem@univ-angers.fr](mailto:anis.ben-abdesslem@univ-angers.fr) (Anis Ben Abdesslem)

of components/systems. It exhibits a wide range of shapes for the density and hazard functions, which are suitable to model complex failure data sets. A successful application of the 3-p Weibull distribution depends on having acceptable statistical estimates of the parameters of the model because it directly affects the accuracy of the reliability analysis. However, estimating the parameters of the 3-p Weibull distribution is intrinsically a very difficult task as well  
10 reported in the literature [13–18]. There are several different methods for estimating the model parameters. These include the maximum likelihood (ML), method of moments (MM), modified maximum likelihood (MML), least squares (LS), and other methods; see for example, Kantar and Senoglu [19] and the references therein. Among the estimation methods, the ML is arguably the most widely-used because it has several desirable properties often not enjoyed by alternative methods of estimation. The ML method does not provide explicit estimators of the parameters in certain  
15 cases, therefore, we turn to numerical methods. However, numerical methods may have the following problems: (i) non-convergence of iterations, (ii) slow convergence, and (iii) convergence to the wrong root [20–24].

Due to the practical importance of this distribution, the development of an efficient and effective optimisation techniques has been an important area of research in reliability engineering and lifetime studies to overcome the in-  
20 efficiency of the numerical methods. Nagatsuka et al. [3] proposed a method based on a data transformation, which avoids the problem of unbounded likelihood. Through Monte Carlo simulations, the authors showed that the proposed method performs better than some existing methods in terms of bias and root mean squared error. Abbasi et al. [13] applied a simulated annealing algorithm to maximise the likelihood function to estimate the parameters of a 3-p Weibull distribution and then they employed an artificial neural network in estimating Weibull's parameters [25]. Hasan et al. [26] focused on likelihood method and used differential evolution algorithm to maximise the likelihood function.  
25 Nosal and Nosal [27] used Monte Carlo simulation and array processing language to investigate the performance of the gradient random search minimisation procedure for fitting a Weibull distribution to a given data set using minimum Kolmogorov-Smirnov distance approach. In Acitas et al. [28], the maximum likelihood (ML) estimation method has been considered using particle swarm optimisation (PSO) with a new adaptive search space using confidence intervals  
30 based on a modified maximum likelihood estimators. Örkücü et al. [29] used a population based heuristic optimisation technique based on particle swarm optimisation to estimate the parameters of the 3-p Weibull distribution. They demonstrated the efficiency of the PSO using three numerical examples under different control parameters. Babak et al. [30] proposed a new hybrid methodology based on a variable neighborhood search and a simulated annealing approach to maximise the likelihood function. The efficiency of their method has been demonstrated through a wide  
35 range of numerical examples. Luus and Jammer [31] showed that MLE gives the most reliable parameter estimation in comparison with the errors-in-variables and least-squares methods. Cousineau [32] reviewed estimation methods for 3-p Weibull distribution. Yang and Yue [33] proposed a new method based on the kernel density estimation using the neural network and genetic algorithm to estimate the parameters of 3-p Weibull distribution. Nagatsuka and Balakrishnan [34] proposed a method for the estimation of parameters and quantiles of the 3-p Weibull distribution. Markovich  
40 and Jukic [35] examined moments method for Weibull distribution. Moeni et al. [36] proposed a Cross Entropy (CE)

method in the context of maximum likelihood estimation (MLE) of a 3-p Weibull distribution. Bartolucci et al. [37], Bartkute and Sakalauskas [38], Jukic et al. [39], Jukic and Markovich [40], and Markovich and Jukic [41] examined moments method for Weibull distribution.

4.5 In this paper, a new approach based on the concept of the Approximate Bayesian Computation (ABC) method is considered for solving the optimisation problem and estimating the parameters of the 3-p Weibull distribution. The ABC method, which corresponds to a class of sequential particle filter methods, has proved to be quite effective in a number of Bayesian model inference problems [42–48]. In this study, the inference problem is translated into an optimisation problem using the same concept of the ABC algorithm. In short, the optimisation scheme is based on an  
5.0 acceptance-rejection (AR) mechanism coupled with an efficient *nested sampling* (NS) technique. It has been named AR-NS. The AR-NS algorithm approaches gradually the optimal solution by propagating “good” particles and lowering progressively a pre-specified tolerance threshold. The proposed algorithm considers the optimisation process by generating several populations and each population contains a fixed number of particles. The algorithm accepts/rejects the particles by comparing the objective function value associated to a given particle to a tolerance threshold which  
5.5 itself evolves over the iterations. Through the algorithm a scoring rule is defined to reward good particles by assigning a high weight. The proposed scheme pushes the best particles towards the region of minimum cost function values. In order to show the high performance of the proposed AR-NS algorithm, an extensive Monte-Carlo simulation study is conducted. In addition, the proposed scheme is applied to real life data in order to show the practical side of the method.

6.0 The paper is organised as follows. Section 2 introduces briefly the mathematical structure of the 3-p Weibull model and its specifications. The mathematical expression of the likelihood function and the formulation of the optimisation problem are given in the same Section. Section 3 details the principle of the AR-NS algorithm, and the way it can be applied to the present problem. Several aspects related to the implementation of the proposed optimisation algorithm are addressed in the same section. Four numerical examples based on simulated data sets are provided in Sections 4 and  
6.5 5 to demonstrate the efficiency and the practicality of the AR-NS algorithm. Section 6 demonstrates the efficiency of the algorithm using real data sets. Section 7 closes the paper with some conclusions and future research efforts.

## 2. The 3-p Weibull distribution specifications

### 2.1. Model structure

The cumulative distribution function (CDF) of the 3-p Weibull distribution is given by:

$$F(t; \alpha, \beta, \gamma) = \begin{cases} 1 - \exp\left\{-\left(\frac{t-\gamma}{\alpha}\right)^\beta\right\}, & t > \gamma \\ 0, & t \leq \gamma \end{cases} \quad (1)$$

where  $\gamma \geq 0$  is a location parameter,  $\alpha > 0$  is the scale parameter, and  $\beta > 0$  is the shape parameter and  $t$  is the time to failure. The corresponding probability density function (PDF) is given by:

$$f(t; \alpha, \beta, \gamma) = \begin{cases} \frac{\beta}{\alpha} \left(\frac{t-\gamma}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{t-\gamma}{\alpha}\right)^\beta\right\}, & t > \gamma \\ 0, & t \leq \gamma \end{cases} \quad (2)$$

If  $\gamma = 0$ , the resulting distribution is called the two-parameter (2-p) Weibull distribution.

The reliability function  $R(t)$  and the failure rate function  $h(t)$  (also known as the hazard rate) for the 3-p Weibull distribution are given in Eqs. (3) and (4), respectively:

$$R(t; \alpha, \beta, \gamma) = \exp\left\{-\left(\frac{t-\gamma}{\alpha}\right)^\beta\right\} \quad (3)$$

and

$$h(t; \alpha, \beta, \gamma) = \frac{f(t; \alpha, \beta, \gamma)}{1 - F(t; \alpha, \beta, \gamma)} = \frac{\beta}{\alpha} \left(\frac{t-\gamma}{\alpha}\right)^{\beta-1} \quad (4)$$

The 3-p Weibull distribution can allow for decreasing, constant, and increasing hazard rates. This is one of the beautiful properties that made the Weibull distribution so applicable.

## 2.2. Maximum likelihood estimator (MLE)

Maximum likelihood estimation (MLE) is a commonly used technique for parameter estimation for a large class of distributions. It is widely known that maximum likelihood estimators are asymptotically unbiased with the minimum variance. In this section, we use the method of maximum likelihood to discuss the estimation of model parameters based on a complete sample. Let  $(t_1, t_2, \dots, t_n)$  be a random sample of size  $n$  drawn from a 3-p Weibull probability density function,  $f(t_i; \vartheta)$ , of unknown parameters,  $\vartheta$ . The likelihood function (LF) is formulated as follows:

$$\mathcal{L} = \prod_{i=1}^n f(t_i; \vartheta) \quad (5)$$

where  $\vartheta$  is a parameter vector representing the unknown parameters, i.e.,

$$\vartheta = (\beta, \alpha, \gamma) \quad (6)$$

The goal, here is to find a vector  $\vartheta$ , that maximises the so-called likelihood function. To maximise  $\mathcal{L}$ , we may equivalently use its logarithm, say  $\ln(\mathcal{L})$ . Estimates are obtained through solving the following equation set:

$$\frac{\partial \ln(\mathcal{L})}{\partial \vartheta} = 0 \quad (7)$$

Equation (7) is hard to solve when applied to Weibull distribution, and many others. In this regard, we may be forced to use numerical procedures.

The LF for Weibull distribution is as follows:

$$\mathcal{L}(t_i, \vartheta) = \prod_{i=1}^n \left[ \frac{\beta}{\alpha} \left(\frac{t_i-\gamma}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{t_i-\gamma}{\alpha}\right)^\beta\right\} \right] \quad (8)$$

Its logarithm will be as follows:

$$\ell_{\text{WE}}(t_i, \vartheta) = n \ln \left( \frac{\beta}{\alpha} \right) + \sum_{i=1}^n \left[ - \left( \frac{t_i - \gamma}{\alpha} \right)^\beta + (\beta - 1) \ln \left( \frac{t_i - \gamma}{\alpha} \right) \right] \quad (9)$$

It is very difficult to maximise  $\mathcal{L}$  (or  $\ell_{\text{WE}}$ ), using ordinary optimisation techniques. Gradient method which is dependent on the partial derivations of the objective function is not a good method to use here, because it is very hard to evaluate the gradient terms and the objective function itself at different points the algorithm needs. Also worth noting, that it is quite challenging to derive the gradient of this complicated objective function. It is well known that the regularity conditions are not satisfied for the maximum likelihood (ML) estimation of the three parameter Weibull distribution because the support of the PDF depends on the unknown parameter when the threshold parameter  $\gamma$  is unknown. Then, the maximum likelihood estimators (MLEs) may not exist, and may not have the classical asymptotic properties even if they exist (see Refs. [49–52] for details). For this reason, various alternative methods have been sought in the literature. In this paper, a gradient-free optimisation algorithm is proposed offering to the statistical and reliability analyst another alternative to deal with parameter estimation.

### 3. Optimisation strategy

#### 3.1. Formulation of the optimisation problem

In the framework of this study, the optimisation problem is formulated as follows:

$$\begin{cases} \text{Minimise}_{\vartheta} & \rho(\vartheta) = -\ell_{\text{WE}}(\vartheta|t_i) \\ \text{subject to :} & \beta^l \leq \beta \leq \beta^u \\ & \alpha^l \leq \alpha \leq \alpha^u \\ & \gamma^l \leq \gamma \leq \gamma^u \end{cases} \quad (10)$$

Here,  $\rho(\vartheta)$  is the objective function (the distance/metric measuring the discrepancies in the ABC jargon),  $-\ell_{\text{WE}}(\vartheta|t_i)$  is the minus log-likelihood function.  $(\beta^l, \beta^u)$ ,  $(\alpha^l, \alpha^u)$ , and  $(\gamma^l, \gamma^u)$  are the upper and the lower boundaries of  $\gamma$ ,  $\alpha$ ,  $\beta$ , respectively. To solve the optimisation problem, the proposed AR-NS algorithm and its implementation are detailed in the next section.

#### 3.2. AR-NS sampler

In this section, the iterative process of the AR-NS algorithm given in Algorithm 1 is detailed and each of the steps will be discussed. The algorithm starts by sampling  $N$  particles from the parameter space satisfying the constraint  $\rho(\cdot) \leq \varepsilon_1$ . Here,  $\rho(\cdot)$  is the objective function to be minimised and  $\varepsilon_1$  is a pre-specified tolerance level. The accepted particles are then weighted following the formula given in Step 9 and the next tolerance threshold is computed based on the objective function values ranked in descending order (highest on top, see, Step 11) as the  $(\varphi_0 N)^{\text{th}}$  value where  $\varphi_0$  is the proportion of discarded particles defined by the user. Then, a weight of zero is assigned to the dropped

particles. After that, we normalise the weights of the remaining particles (see, Step 13). From the remaining particles, one selects in a probabilistic way  $\lambda_0 N$  particles based on the updated weight values, where  $\lambda_0$  is the proportion of particles, so-called “surviving” particles (see Step 14). The objective of this step is to ensure a good exploration of the parameter space by giving the possibility to select “bad” particles and to avoid the trapping attraction of local minima. The alive particles are then enclosed in an ellipsoid in which the center  $\mu$  and the covariance matrix  $\mathcal{C}$  are estimated based on the remaining particles; one denotes this ellipsoid by  $\mathcal{E} = (\mu, \mathcal{C})$ . To guarantee that the particles on the borders are inside the ellipsoid, one may enlarge the volume by a factor  $f_0$ . This technique was firstly proposed in [53] to improve the efficiency of the NS algorithm which has been widely used in different domains, mainly in cosmology [54]. Finally, the population is populated by resampling  $(1 - \lambda_0)N$  particles inside the enlarged ellipsoid (see Step 20) following the same scheme and a re-weighting step is carried out (Step 28) [55]. The entire process is repeated in the subsequent iterations until some stopping criterion is met. It should be noted that employing a naive approach that draws blindly from the parameter space would result in a steady decrease in the acceptance rate of new samples with increasing likelihood. To overcome this issue, the algorithm uses an elliptical bound containing the current alive particles at each stage of the process to restrict the region around the optimal solutions from which new samples are drawn. Ellipsoids are intended to reduce the effective initial volume of the parameter space where the drawing has to take place. This considerably improves the speed of the AR-NS algorithm because sampling from uninteresting regions of the parameter space is avoided.

As in other population-based optimisation methods, the AR-NS algorithm requires the selection of some tuning parameters. In this study, the parameters associated to the AR-NS algorithm are defined as follows: the number of particles is set to  $N = 1000$ ;  $\varphi_0$ ,  $\lambda_0$ , and  $f_0$  are set to 0.3, 0.6, and 1.1, respectively. The initial tolerance threshold  $\varepsilon_1$  is selected with the goal to guarantee an acceptance rate roughly equal or higher than 40% at the first iteration. A detailed discussion concerning the effects of these settings can be found in [46].

130

### 3.3. Definition of the parameter space

In order to use the AR-NS algorithm, upper and lower bounds on the parameters being optimised should be determined. To get a crude estimates of the model parameters  $\vartheta = (\alpha_0, \beta_0, \gamma_0)$ , a very simple method in concept is used. A precise description of the applied method is given in the sequel. Throughout this section:

$$0 < t_1 < t_2 < \dots < t_n$$

denotes the data generated from the three-parameter Weibull distribution.

To estimate the model parameter, the Weibull probability plot is used which is based on the following transformation of the Weibull distribution  $F(t; \alpha, \beta, \gamma)$ :

$$y = \ln \left\{ \ln \left[ \frac{1}{1 - F(t; \alpha, \beta, \gamma)} \right] \right\} = \beta \ln(t - \gamma) - \beta \ln \alpha \quad (11)$$

---

**Algorithm 1** AR-NS FOR PARAMETER ESTIMATION

---

**Require:** Data  $t_i$ , Initial tolerance  $\varepsilon_1$ , Objective function  $\rho(\cdot)$ , Tuning parameters  $(\varphi_0, \lambda_0, f_0)$ , Precision level  $\xi$

- 1: Iteration  $t = 1$ : ▷ First iteration
  - 2: **for**  $i = 1$  to  $N$  **do**
  - 3:     **repeat**
  - 4:         Sample a parameter vector,  $\vartheta^*$ , from the parameter space
  - 5:         Evaluate the minus log-likelihood function  $\rho(\cdot) = -\ell_{\text{WE}}(\vartheta^*|t_i)$ .
  - 6:     **until**  $\rho(\cdot) \leq \varepsilon_1$
  - 7:     Store the pair  $(\vartheta^*, \rho(\cdot))$  in population  $\mathcal{P}_1$  ▷  $e_i = \rho(\cdot)$
  - 8: **end for**
  - 9: Evaluate weights for all particles:  $\omega_i = \frac{1}{\varepsilon_1} \left( 1 - \left( \frac{e_i}{\varepsilon_1} \right)^2 \right)$
  - 10: Sort  $e_i$  in descending order and store them in  $e^t$ .
  - 11: Compute the next tolerance threshold:  $\varepsilon_2 = e^t(\varphi_0 N) \implies \varphi_0 = 0.3$
  - 12: Drop particles with  $\rho(\cdot) > \varepsilon_2$ ,  $\omega_{i=1:\varphi_0 N} = 0$
  - 13: Normalise the weights of the remaining particles such that  $\sum_{i=1}^{(1-\varphi_0)N} \omega_i = 1$
  - 14: Select  $\mathcal{A}_t = \lambda_0 N$  particles from the remaining based on the assigned weights  $\implies \lambda_0 = 0.6$
  - 15: Define the ellipsoid by its centre of the mass and covariance matrix  $\mathcal{E}_1 = \{\mu_1, \mathcal{C}_1\}$
  - 16: Enlarge the ellipsoid  $\mathcal{E}_1$  by  $f_0 = 1.1$
  - 17: Iteration  $t = 2$ : while the stop criterion is not reached do ▷ Subsequent iterations
  - 18: **while**  $|\varepsilon_t - \varepsilon_{t-1}| \leq \xi$  **do**
  - 19:     **for**  $j = 1$  to  $(1 - \lambda_0)N$  **do**
  - 20:         **repeat**
  - 21:             Sample a parameter vector  $\vartheta^*$  inside the ellipsoid  $\mathcal{E}_{t-1}$
  - 22:             Evaluate the objective function  $\rho(\cdot) = -\ell_{\text{WE}}(\vartheta^*|t_i)$
  - 23:             **until**  $\rho(\cdot) \leq \varepsilon_t$
  - 24:             Store the pair  $(\vartheta^*, \rho(\cdot))$  in  $\mathcal{S}_t$ ,  $e_j = \rho(\cdot)$
  - 25:         **end for**
  - 26:         Get the new population,  $\mathcal{P}_t = \{\mathcal{A}_{t-1}\} \cup \{\mathcal{S}_t\}$  with their correspondent distance values  $e^t$  ▷  $\mathcal{P}_t$ : is the  $t$ -th replenished population
  - 27:         Sort  $e^t$  and define  $\varepsilon_{t+1} = e^t(\varphi_0 N)$
  - 28:         Compute new weights for all particles as in Step (9) and normalise them
  - 29:         Define the new set of selected particles  $\mathcal{A}_t$  as in Step (14)
  - 30:         Update the centre of the mass and covariance matrix of the ellipsoid ▷ The same enlargement factor is used for all the simulations
  - 31: **end while**
- Ensure:** The entire distribution of the optimal estimates
-



The initial approximation of  $\gamma$  is given by  $\gamma_0 = t_1 - 1/n$  [17]. For the scale and shape parameters, the initial approxi-  
 135 mations are then obtained by using the following steps:

- Compute the empirical CDF  $\hat{F}$  corresponding to the sample data,  $t_1, \dots, t_n$ . The mean rank estimator for the empirical cumulative distribution function is given by Eq. (12):

$$\hat{F}(t_i) = \frac{i-c}{n+1-2c} \quad 0 \leq c \leq 1 \quad (12)$$

There are many different ways of computing empirical distribution function  $\hat{F}$  corresponding to the sample data  $t_1, \dots, t_n$ . They all involve arranging the data in ascending order so that  $t_1 < t_2 < \dots < t_n$ . The estimator used in this study is expressed in the following form (i.e.,  $c = 0.3$ ):

$$\hat{F}(t_i) = \frac{i-0.3}{n+0.4} \quad (13)$$

- Compute

$$y_i = \ln[-\ln(1 - \hat{F}(t_i))] \quad \text{and} \quad x_i = \ln(t_i - \gamma_0) \quad \text{for} \quad 1 \leq i \leq n$$

- Fit a straight line:  $y = ax + b$  to the data  $\{(x_i, y_i)\}_{i=1}^n$ , by using the least square method.
- Let  $y = a^*x + b^*$  be the fitting line.
- Let  $\beta_0 = a^*$  and  $\alpha_0 = \exp(-b^*/\beta_0)$  the initial estimates of the shape and scale parameters, respectively.

Based on the crude estimates, one may define the parameter space. In this study, the shape and scale parameters  
 140 were given uniform priors spanning a range one order of magnitude above and below the crude estimates while the location parameter is varying between zero and the minus of the data as illustrated in Table 9.

Parameter	Type of distribution	Lower bound	Upper bound
$\beta$	Uniform	$0.1 \times \beta_0$	$10 \times \beta_0$
$\alpha$	Uniform	$0.1 \times \alpha_0$	$10 \times \alpha_0$
$\gamma$	Uniform	0	$\min(t)$

Table 1: Parameter bounds for the considered 3-p Weibull model.

#### 4. Numerical examples

To demonstrate the performance of the proposed AR-NS, I shall present some numerical examples using simulated  
 145 data with different parameter settings and sample sizes. It should be noted that no restriction has been imposed on the maximum number of iterations and convergence is assumed when the absolute difference between successive tolerance threshold drops below  $10^{-6}$ .

4.1. Example # 1:  $\vartheta = (2, 2, 2)$

The first example is to estimate the parameters of a 3-p Weibull distribution with  $\vartheta = (2, 2, 2)$  as previously performed by [13, 26]. Furthermore, one investigates the effects of the sample size on the precision estimates. The considered sample sizes are  $n = 100, 500, 1000,$  and  $2500$ . Table 2 displays the obtained results for the considered sample sizes. It should be noted that each optimal vector is being estimated as the mean values of the particles obtained at the last population. It can be seen that as the sample size increases, the more will be better the estimation. It is observed that for  $n = 500, 1000,$  and  $2500$  the estimated parameters are very close to the true values. Additionally, it can be noted that the absolute relative difference between the log-likelihood functions evaluated at the real and optimal estimates is becoming smaller as we increase the sample size. The same finding was found in [13].

Weibull parameters		$\vartheta = (2, 2, 2)$			
Sample size		100	500	1000	2500
Estimated parameters	$\hat{\beta}$	1.7021	1.9250	2.0178	1.9843
(mean values)	$\hat{\alpha}$	1.7639	1.9860	2.0005	1.9892
	$\hat{\gamma}$	2.1013	2.0704	1.9841	2.0047
	$\mathcal{L}(\hat{\vartheta} t_i)$	-128.2156	-655.6756	-1281.9445	-3239.8829
	$\mathcal{L}(\vartheta t_i)$	-129.2754	-657.4319	-1282.2499	-3240.1048
	$\left  \frac{(\mathcal{L}(\hat{\vartheta} t_i) - \mathcal{L}(\vartheta t_i))}{\mathcal{L}(\vartheta t_i)} \right $	0.0082	0.0027	$2.3818 \times 10^{-4}$	$6.8485 \times 10^{-5}$

Table 2: Optimal estimates of the 3-p Weibull parameters using the AR-NS algorithm.

Figs. 1-4 show how the algorithm converges progressively to the optimal estimates considering all the sample sizes. The blue lines represent the mean value obtained at each population while the orange lines delimit the 95% credible intervals over the populations (the dashed line shows the exact value). It is obvious that the bigger the sample size the better the estimation. Figs. 5-8 show trace plots and corresponding histograms of the model parameters at the last population for all the sample sizes. It can be seen that the histograms are too peaked and all the parameters are recovered with small uncertainty.

Fig. 9 shows the evolution of the particles over some intermediate populations. One can see that the prior volume shrinks down and particles move to approach the region of optimal solutions (i.e., high likelihood region) by decreasing progressively the tolerance threshold over the iterations. At the last iteration, the algorithm gives the entire distributions of the optimal parameters rather than simply producing point estimates. From the last population, one may get a point

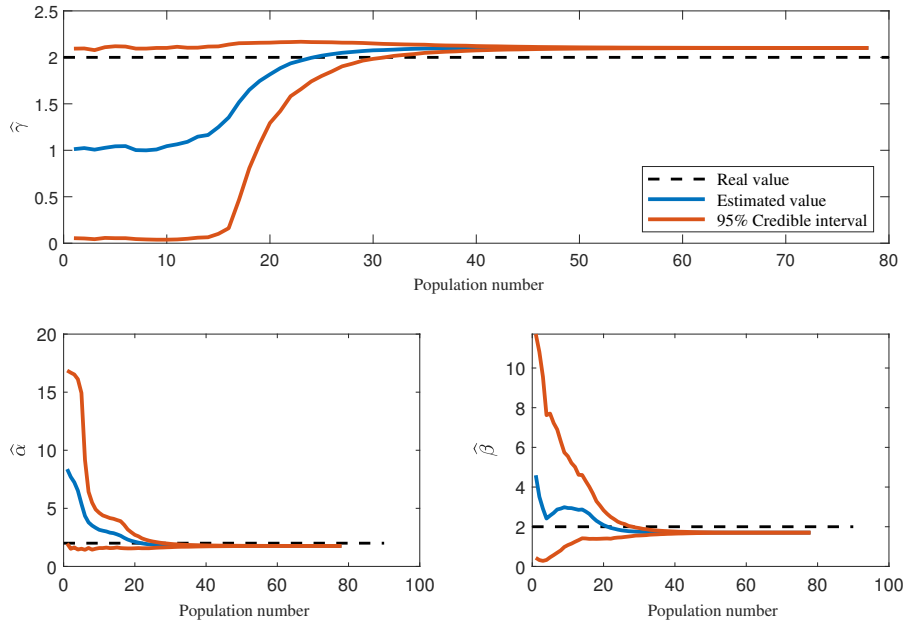


Figure 1: Evolution of the model parameters over the populations for  $n = 100$ .

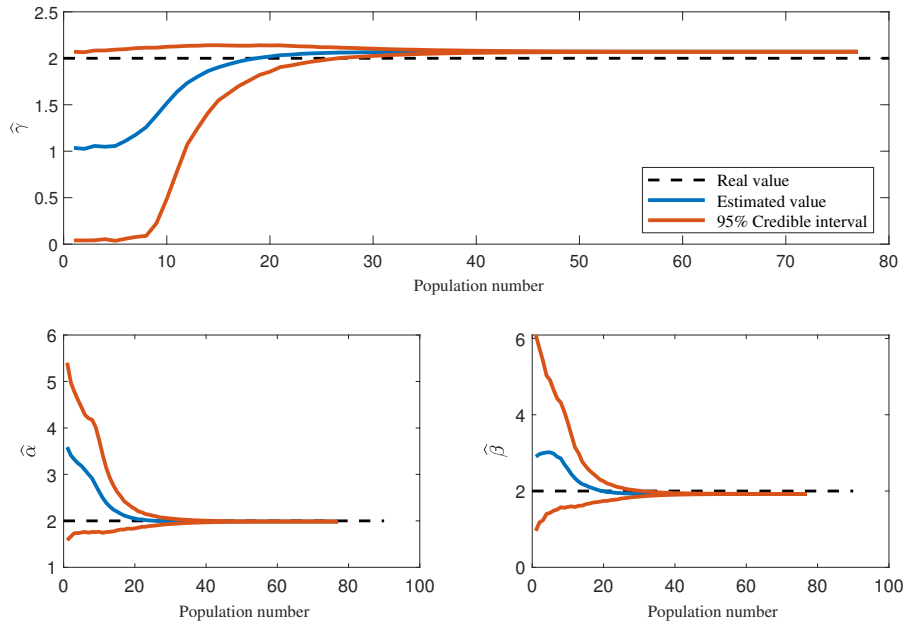


Figure 2: Evolution of the model parameters over the populations for  $n = 500$ .

estimates by taking the mean values.

Now, I shall estimate the parameters of a Weibull distribution with  $\vartheta = (2, 2, 2)$  using 10,000 replications for each

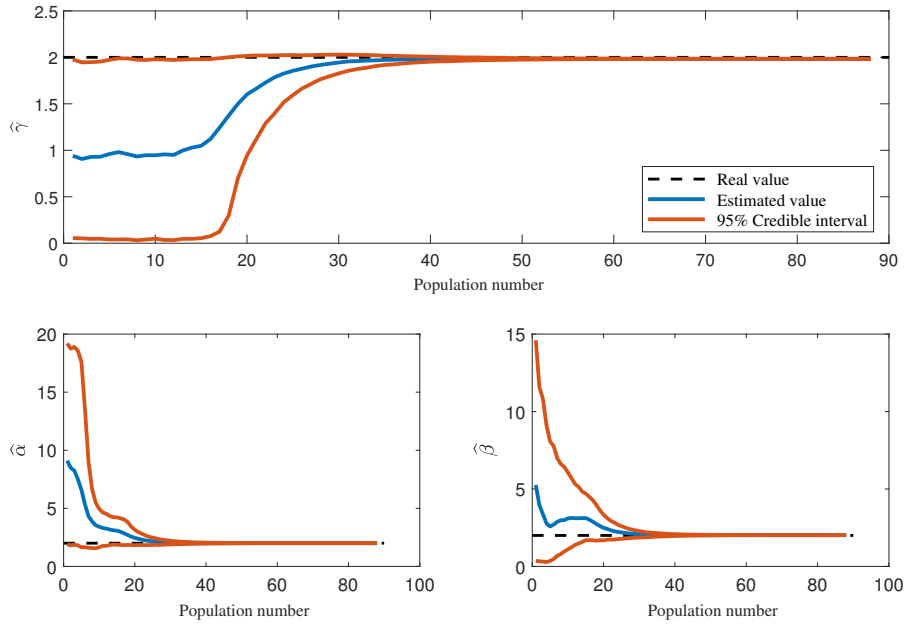


Figure 3: Evolution of the model parameters over the populations for  $n = 1000$ .

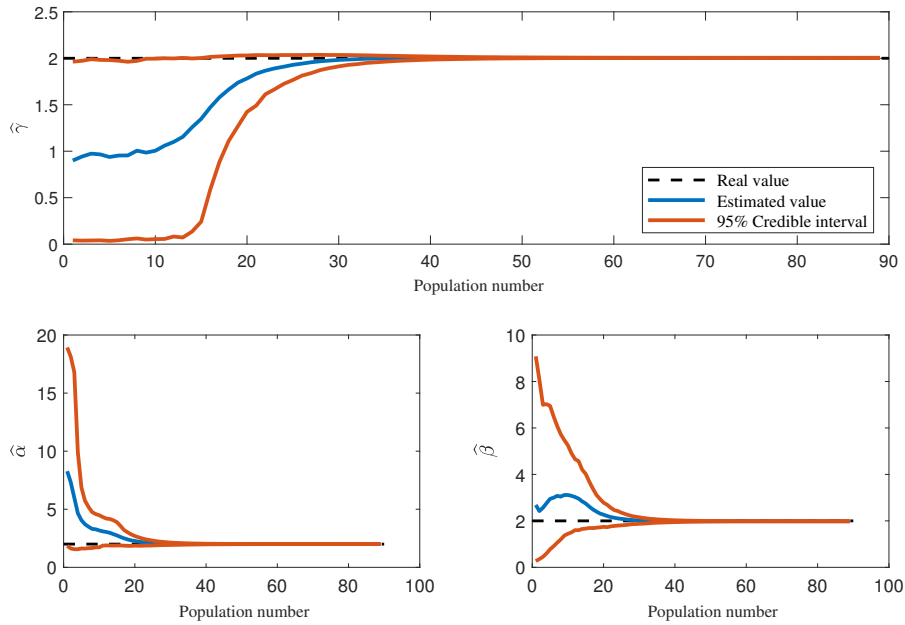


Figure 4: Evolution of the model parameters over the populations for  $n = 2500$ .

sample size. The results of the simulation runs are reported in Table 3 for all the sample sizes. It can be seen that as the sample size increases, not only the estimates get closer to the real values of the parameters, but also more precise

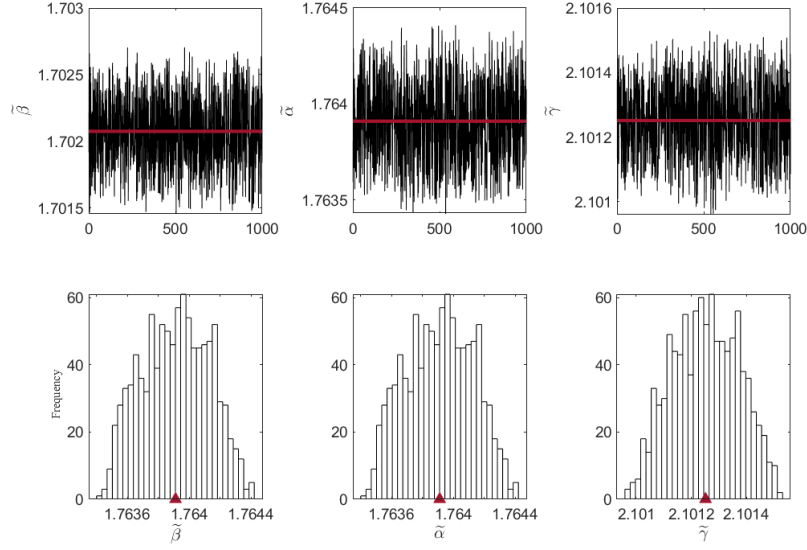


Figure 5: Trace plots (the dashed line represents the mean value) and corresponding histograms of the model parameters (the triangles represent the mean values) at the last population for  $n = 100$ .

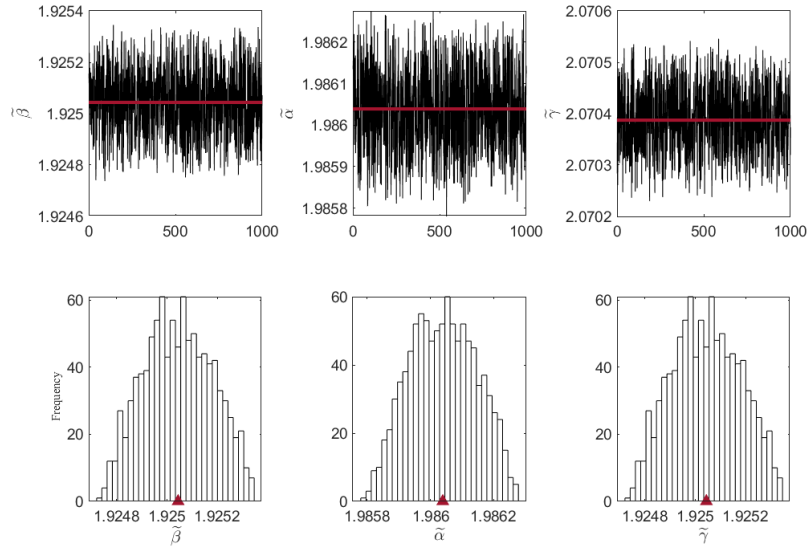


Figure 6: Trace plots (the dashed line represents the posterior mean value) and corresponding histograms of the model parameters (the triangles represent the mean values) at the last population for  $n = 500$ .

170 confidence intervals are obtained. The obtained results prove that the proposed AR-NS performs quite well. Fig. 10 shows the histograms of the model parameters for the different sample sizes (at the end of each simulation, one takes the mean values to get a point estimates). One can see that the histograms are too peaked around the true values mainly

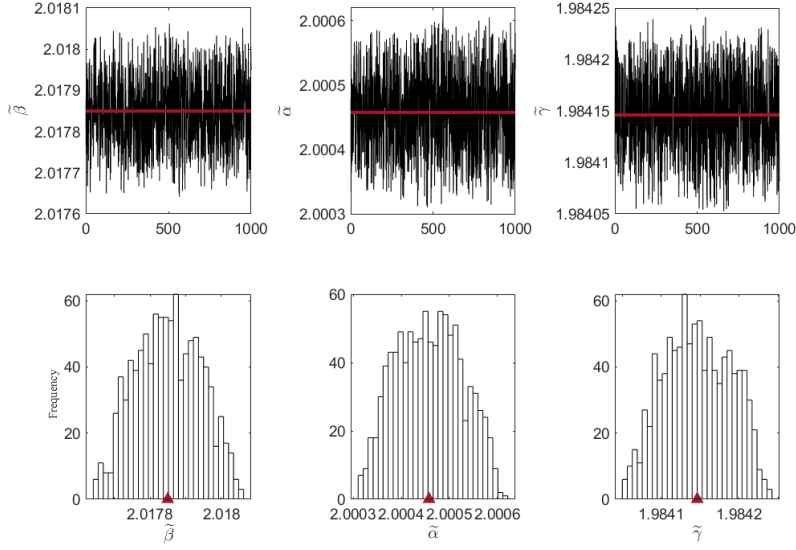


Figure 7: Trace plots (the dashed line represents the mean value) and corresponding histograms of the model parameters (the triangles represent the mean values) at the last population for  $n = 1000$ .

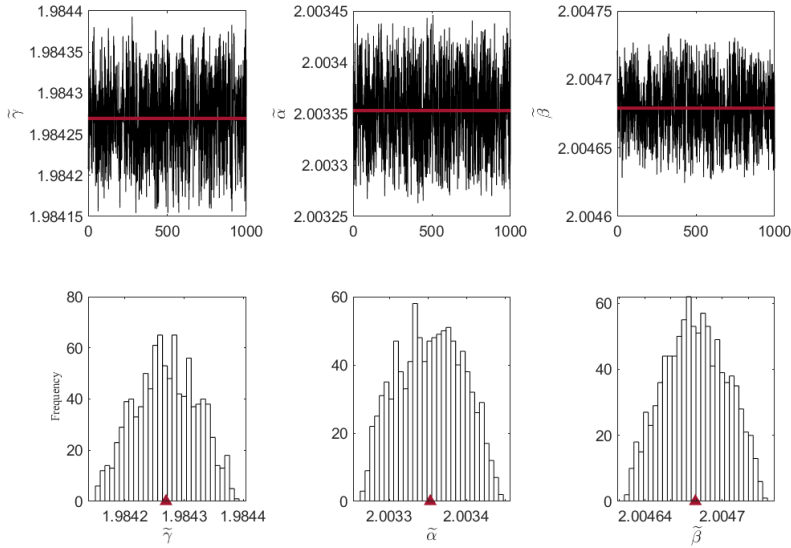


Figure 8: Trace plots (the dashed line represents the mean value) and corresponding histograms of the model parameters (the triangles represent the mean values) at the last population for  $n = 2500$ .

when the sample size increases.

Fig. 11 shows the kernel density function of the minus log-likelihood function values computed at the true and estimated parameters and for all the sample sizes. Overall, we can visually see a good agreement between both kernels

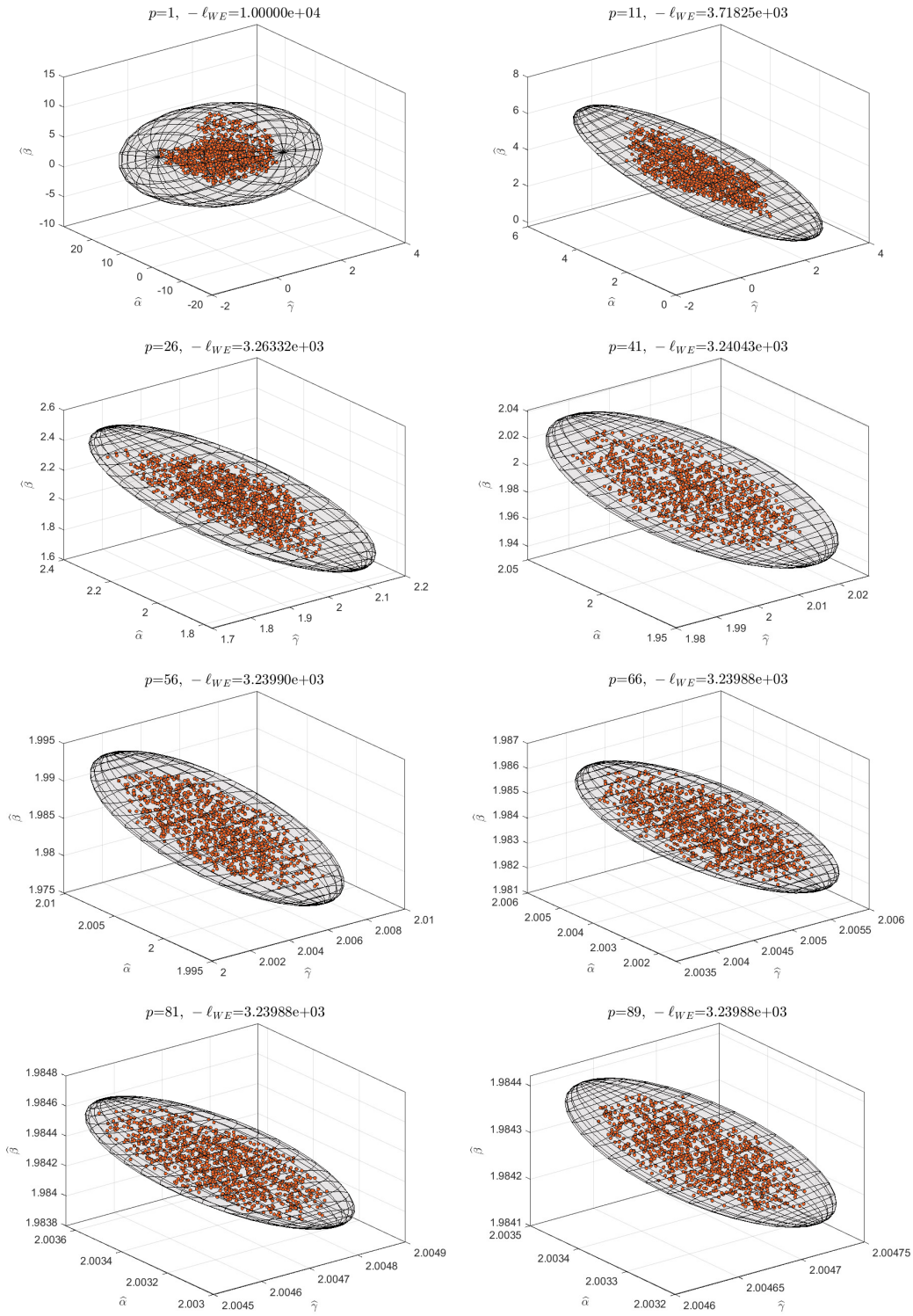


Figure 9: Evolution of the particles over some intermediate populations.

True values	$n$	Estimated values	Mean	Standard deviation	C.I. (95%)		
$\beta = 2$	100	$\hat{\beta}$	1.9529	0.2584	1.5167	2.5239	
$\alpha = 2$		$\hat{\alpha}$	1.9304	0.1910	1.5645	2.3239	
$\gamma = 2$		$\hat{\gamma}$	2.0545	0.1332	1.7846	2.3185	
		$\mathcal{L}(\hat{\vartheta} t_i)$	-127.2106	-	-	-	
		$\mathcal{L}(\vartheta t_i)$	-128.7654	-	-	-	
		$\hat{\beta}$	1.9804	0.0982	1.7971	2.1793	
	500	$\hat{\alpha}$	1.9776	0.0728	1.8359	2.1203	
		$\hat{\gamma}$	2.0180	0.0467	1.9319	2.1188	
		$\mathcal{L}(\hat{\vartheta} t_i)$	-642.6323	-	-	-	
		$\mathcal{L}(\vartheta t_i)$	-644.1584	-	-	-	
		1000	$\hat{\beta}$	1.9882	0.0659	1.8620	2.1228
			$\hat{\alpha}$	1.9874	0.0489	1.8892	2.0827
	$\hat{\gamma}$		2.0108	0.0302	1.8620	2.1228	
	$\mathcal{L}(\hat{\vartheta} t_i)$		-1287.1290	-	-	-	
	$\mathcal{L}(\vartheta t_i)$		-1288.6113	-	-	-	
	2500		$\hat{\beta}$	1.9932	0.0403	1.9155	2.0715
		$\hat{\alpha}$	1.9930	0.0295	1.9338	2.0500	
		$\hat{\gamma}$	2.0058	0.0176	1.9740	2.0440	
		$\mathcal{L}(\hat{\vartheta} t_i)$	-3219.5967	-	-	-	
		$\mathcal{L}(\vartheta t_i)$	-3221.0721	-	-	-	

Table 3: Results obtained by implementation of the AR-NS algorithm in 10,000 replications for  $n = 100, 500, 1000, 2500$ .



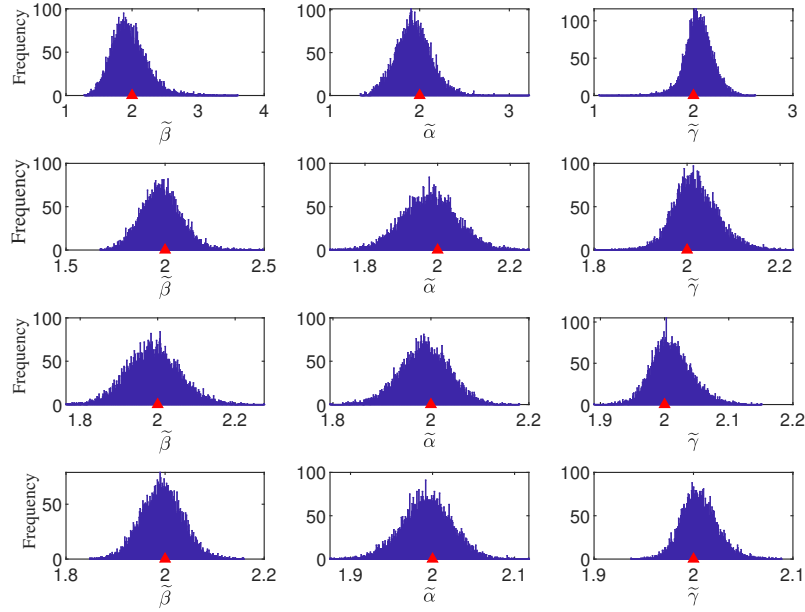


Figure 10: Histograms of the optimal estimates using 10,000 replications considering sample sizes of  $n = 100, 500, 1000, 2500$  from the top to the bottom.

except for  $n = 100$ , where the fit is reasonably well only in the central part of the distribution and poorly in the tails. For more objective analysis, the quality of the estimation method is studied by considering the bias and mean square error (MSE) of the MLEs given by:

$$\text{Bias}(\hat{\vartheta}) = \bar{\vartheta} - \vartheta \quad (14)$$

such that:

$$\bar{\vartheta} = \frac{1}{10,000} \sum_{i=1}^{10,000} \hat{\vartheta}_i \quad (15)$$

$$\text{MSE}(\hat{\vartheta}) = \frac{1}{10,000} \sum_{i=1}^{10,000} (\hat{\vartheta}_i - \vartheta)^2 \quad (16)$$

175 Table 4 gives the bias and mean squared error estimates of the 3-p Weibull model for the different sample sizes. Bias and MSE were calculated for the estimators based on the above formula ( $\vartheta$  can be replaced by  $\beta$ ,  $\alpha$ , and  $\gamma$ ). It can be seen that the MSE and the bias becomes smaller for all the parameters as the sample size increases.

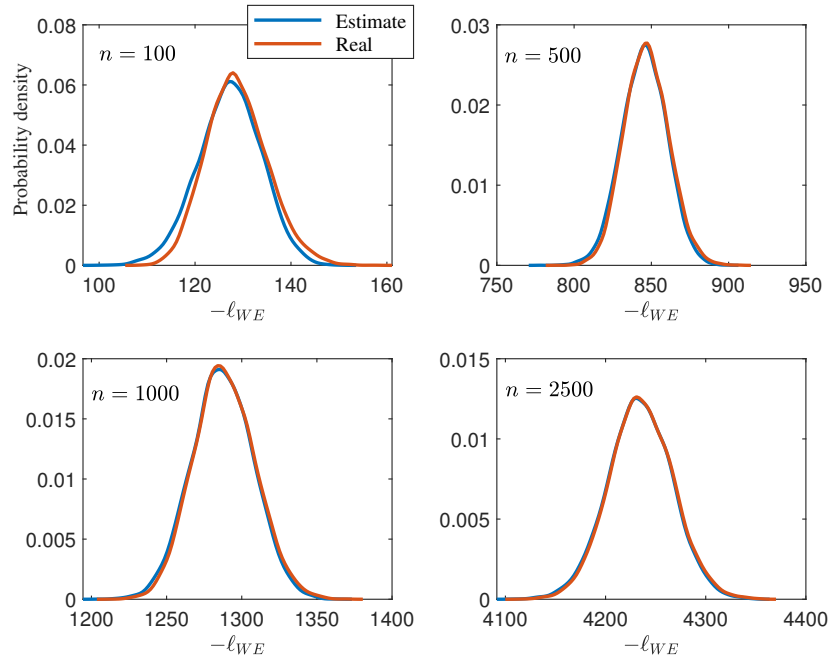


Figure 11: Comparison between the kernel density functions of the negative log-likelihood function estimated at the true and estimated parameters.

Sample size	$\hat{\beta}$		$\hat{\alpha}$		$\hat{\gamma}$	
	Bias	MSE	Bias	MSE	Bias	MSE
100	-0.0471	0.0690	-0.0696	0.0413	0.0545	0.0207
500	-0.0196	0.0100	-0.0224	0.0058	0.0180	0.0025
1000	-0.0118	0.0045	-0.0126	0.0025	0.0108	0.0010
2500	-0.0068	0.0017	-0.0070	0.0009	0.0058	0.0003

Table 4: Bias and mean squared error for 3-p Weibull distribution with  $\vartheta = (2, 2, 2)$ .

#### 4.2. Example # 2: $\vartheta = (2, 3, 4)$

180 This example considers a Weibull distribution with  $\vartheta = (2, 3, 4)$ . For the AR-NS algorithm, we use the same tuning parameters and we run the algorithm for the different sample sizes. The obtained results are displayed in Table 5 and Fig. 12. One can notice that the estimates get closer to the true values as we increase the sample size, but also narrower confidence intervals are obtained. From Fig. 12, it can be seen that the estimated values are all too peaked on the true values.

True values	$n$	Estimated values	Mean	Standard deviation	C.I. (95%)	
$\beta = 2$	100	$\hat{\beta}$	1.9487	0.2536	1.5164	2.5206
$\alpha = 3$		$\hat{\alpha}$	2.8956	0.2811	2.3564	3.4742
$\gamma = 4$		$\hat{\gamma}$	4.0834	0.1963	3.6899	4.4790
		$\mathcal{L}(\hat{\vartheta} t_i)$	-167.9083	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-169.4719	-	-	-
		500	$\hat{\beta}$	1.9811	0.0996	1.7964
		$\hat{\alpha}$	2.9675	0.1093	2.7517	3.1813
		$\hat{\gamma}$	4.0270	0.0687	3.9022	4.1740
		$\mathcal{L}(\hat{\vartheta} t_i)$	-845.4235	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-846.9379	-	-	-
	1000	$\hat{\beta}$	1.9869	0.0671	1.8620	2.1213
		$\hat{\alpha}$	2.9802	0.0742	2.8299	3.1198
		$\hat{\gamma}$	4.0165	0.0459	3.9359	4.1148
		$\mathcal{L}(\hat{\vartheta} t_i)$	-1692.7590	-	-	-
		$\mathcal{L}(\vartheta t)$	-1694.2608	-	-	-
	2500	$\hat{\beta}$	1.9920	0.0404	1.9147	2.0715
		$\hat{\alpha}$	2.9895	0.0445	2.8998	3.0767
		$\hat{\gamma}$	4.0086	0.0261	3.9620	4.0656
		$\mathcal{L}(\hat{\vartheta} t_i)$	-4234.3051	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-4235.7830	-	-	-

Table 5: Results obtained by implementation of the AR-NS algorithm in 10,000 replications for  $n = 100, 500, 1000, 2500$ .

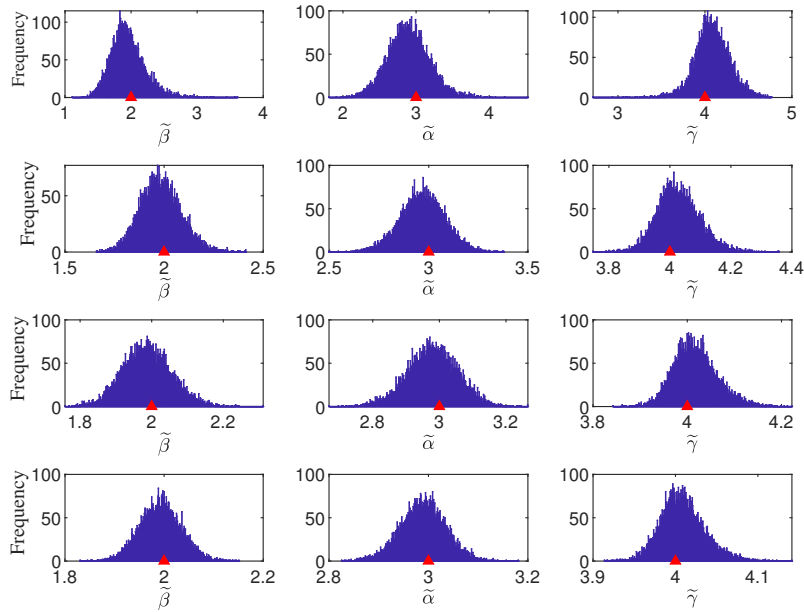


Figure 12: Histograms of the optimal estimates using 10,000 replications considering different sample sizes  $n = 100, 500, 1000, 2500$  from the top to the bottom.

185

Table 6 gives the bias and the mean squared error estimates for the various sample sizes and for all the model parameters. The estimated parameters have considerably smaller biases and mean squared errors as we increase the sample sizes as previously noticed. Fig. 13 displays the kernel density functions of the minus log-likelihood evaluated at the true and estimated values for all the sample sizes. Except for  $n = 100$ , the kernels were in very good agreement.

Sample size	$\hat{\beta}$		$\hat{\alpha}$		$\hat{\gamma}$	
	Bias	MSE	Bias	MSE	Bias	MSE
100	-0.0513	0.0670	-0.1044	0.0899	0.0834	0.0455
500	-0.0189	0.0103	-0.0325	0.0130	0.0270	0.0054
1000	-0.0131	0.0047	-0.0198	0.0059	0.0165	0.0024
2500	-0.0080	0.0017	-0.0105	0.0021	0.0086	0.0008

Table 6: Bias and mean squared error for 3-p Weibull distribution for  $\vartheta = (2, 3, 4)$ .

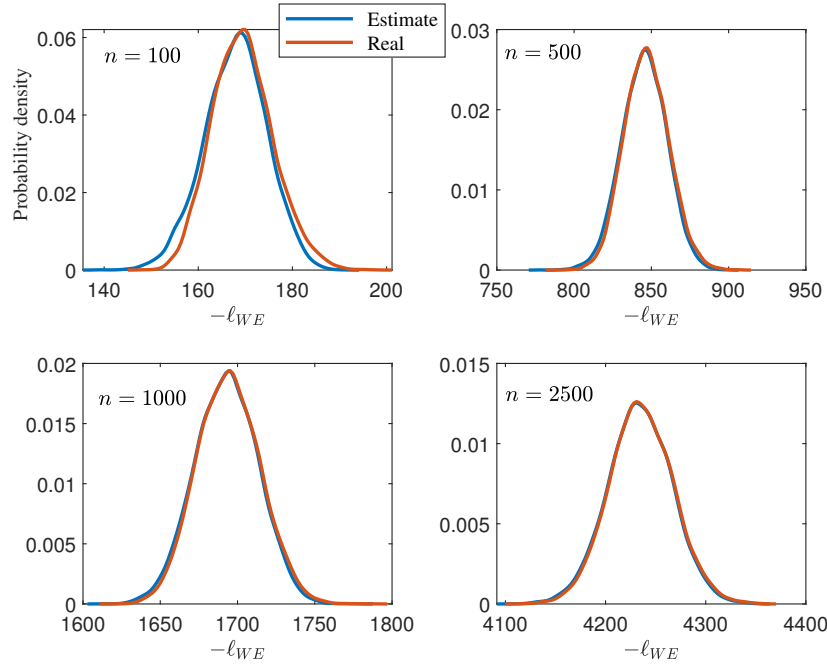


Figure 13: Comparison between the kernel density functions of the negative log-likelihood function estimated at the true and estimated parameters.

190 4.3. Example # 3,  $\vartheta = (3, 2, 5)$

To further evaluate the accuracy and efficiency of the AR-NS algorithm, we consider in this example a Weibull distribution with  $\vartheta = (3, 2, 5)$ . The obtained results are tabulated and depicted in Table 7 and Fig. 14. It can be seen that the AR-NS estimates are very close to the true parameter values. As expected, the optimal estimates become better when the sample size increases and the confidence intervals become narrower as well. One can see from Fig. 14 that the histograms are too peaked on the true parameters. It may be concluded that the AR-NS algorithm may be considered as an effective parameter estimation method for the 3-p Weibull distribution. Table 8 gives the bias and the mean squared error estimates for the various sample sizes and for all the model parameters. It can be seen that both the biases and mean squared errors were reduced as the sample size increases as expected. As for previous examples, we superimpose the minus log-likelihood functions values obtained from the true and estimated parameters for the considered sample sizes. From Fig. 21, one can see that the kernels are very close to each other except in the tails and only for  $n = 100$  as previously noticed.

Throughout the whole numerical examples, it has to be noted that the estimates are quite stable and, more importantly, are close to the true values for the considered sample sizes and for all the setting parameters. Additionally, as the sample size increases, the bias and the MSE decrease. To conclude and as it has been shown via numerical

True values	$n$	Estimated values	Mean	Standard deviation	C.I. (95%)	
$\beta = 3$	100	$\hat{\beta}$	2.9835	0.6309	2.0695	4.5087
$\alpha = 2$		$\hat{\alpha}$	1.9549	0.3284	1.4270	2.7086
$\gamma = 5$		$\hat{\gamma}$	5.0362	0.3014	4.3500	5.5161
	500	$\mathcal{L}(\hat{\vartheta} t_i)$	-96.3374	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-97.8588	-	-	-
		$\hat{\beta}$	2.9797	0.2260	2.5735	3.4545
	1000	$\hat{\alpha}$	1.9814	0.1202	1.7534	2.2259
		$\hat{\gamma}$	5.0163	0.1080	4.7975	5.2235
		$\mathcal{L}(\hat{\vartheta} t_i)$	-488.1443	-	-	-
	2500	$\mathcal{L}(\vartheta t_i)$	-489.6400	-	-	-
		$\hat{\beta}$	2.9883	0.1564	2.6946	3.3116
		$\hat{\alpha}$	1.9900	0.0841	1.8262	2.1559
	5000	$\hat{\gamma}$	5.0090	0.0754	4.8622	5.1543
		$\mathcal{L}(\hat{\vartheta} t_i)$	-977.9615	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-979.4857	-	-	-
	10000	$\hat{\beta}$	2.9931	0.0945	2.8097	3.1793
		$\hat{\alpha}$	1.9941	0.0500	1.8953	2.0898
		$\hat{\gamma}$	5.0054	0.0447	4.9193	5.0940
	20000	$\mathcal{L}(\hat{\vartheta} t_i)$	-2446.2312	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-2447.7345	-	-	-

Table 7: Results obtained by implementation of the AR-NS algorithm in 10,000 replications.

205 experiments, the algorithm presents beautiful properties required for a good estimator such as consistency, asymptotic efficiency, normality, and invariance. In addition, the algorithm is less sensitive to the control parameters, compared with the method proposed in [29].

Sample size	$\hat{\beta}$		$\hat{\alpha}$		$\hat{\gamma}$	
	Bias	MSE	Bias	MSE	Bias	MSE
100	-0.0165	0.3983	-0.0451	0.1099	0.0362	0.0921
500	-0.0203	0.0515	-0.0186	0.0148	0.0163	0.0119
1000	-0.0117	0.0246	-0.0100	0.0072	0.0090	0.0058
2500	-0.0069	0.0090	-0.0059	0.0025	0.0054	0.0020

Table 8: Bias and MSE values for the model  $\vartheta = (3, 2, 5)$ .

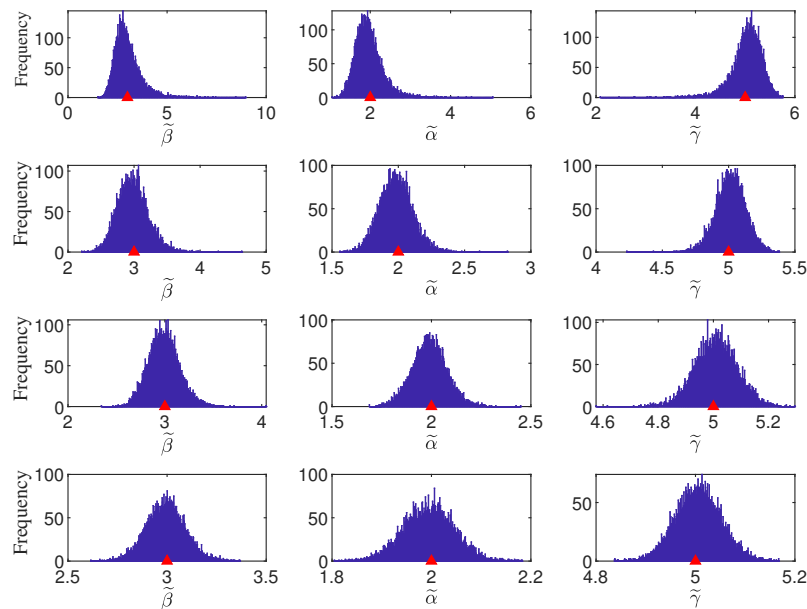


Figure 14: Histograms of the optimum mean values using 10,000 replications considering different sample sizes.

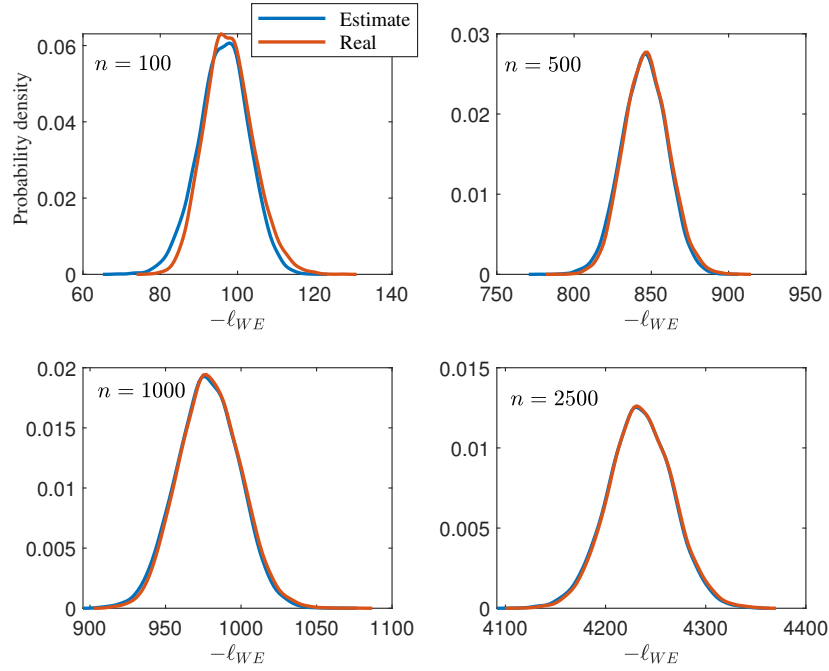


Figure 15: Comparison between the kernel density functions of the negative log-likelihood function estimated at the true and estimated parameters.

## 5. The CE2G distribution specifications

### 210 5.1. Model structure

To demonstrate and validate the efficiency of the AR-NS algorithm, we apply it to the Complementary Exponentiated Exponential Geometric (CE2G) distribution [56]. The PDF and the reliability function are given by:

$$f(t; \alpha, \theta, \lambda) = \frac{\alpha \lambda \theta e^{-\lambda t} (1 - e^{-\lambda t})^{\alpha-1}}{[1 - (1 - \theta)(1 - e^{-\lambda t})^\alpha]^2}, \quad t > 0, \quad (17)$$

$$R(t; \alpha, \theta, \lambda) = \frac{1 - (1 - e^{-\lambda t})^\alpha}{1 - (1 - \theta)(1 - e^{-\lambda t})^\alpha}, \quad t > 0, \quad (18)$$

where  $\lambda$  is a scale parameter of the distribution, and  $\alpha$  and  $\theta$  are shape parameters, ( $\alpha > 0$ ,  $\theta \in (0, 1)$ , and  $\lambda > 0$ ).

Assuming the lifetimes are independently distributed, the MLEs of the parameters are obtained by direct maximization of the log-likelihood function given by:

$$\ell_{\text{CE2G}}(t_i, \vartheta) = n \ln(\alpha \theta \lambda) - \lambda \sum_{i=1}^n t_i + (\alpha - 1) \sum_{i=1}^n \ln(1 - e^{-\lambda t_i}) - 2 \sum_{i=1}^n \ln(1 - (1 - \theta)(1 - e^{-\lambda t_i})^\alpha) \quad (19)$$

### 5.2. Example # 4, $\vartheta = (3, 0.5, 1)$

A simulation study was performed based on generated data sets from the CE2G distribution for different sample sizes,  $n = 100, 500, 1000, \text{ and } 2500$ . The true parameter values were taken as  $\alpha = 3$ ,  $\theta = 0.5$  and  $\lambda = 1$ . All of the



215 model parameters were given uniform priors (see Table 9).

Parameter	True value	Prior range
$\alpha$	3	$\mathcal{U}(0, 10)$
$\theta$	0.5	$\mathcal{U}(0, 1)$
$\lambda$	1	$\mathcal{U}(0, 10)$

Table 9: Parameter bounds for the considered CE2G model.

The results are given in Table 10, which shows the mean values of the MLE, together with the standard deviation, and the 95% confidence intervals for parameters of the CE2G model. The results suggest that the ML estimates have performed adequately. One can notice that the estimates get closer to the true values as we increase the sample size, but also narrower confidence intervals are obtained. Once again, the maximum likelihood estimator is robust against various sample sizes. From Table 11, we can see that the Bias and the MSE of all estimators tend to zero for large  $n$ , i.e., the estimates are asymptotically unbiased for the parameters.

## 6. Real data analysis

### 225 6.1. Example # 5: Failure data of ceramic material

In this example, we shall consider the real data set presented by [57], which represents the failure data of ceramic material shown in Table 12. Table 13 provides some descriptive summary statistics of the data. We note a positive skewness, in consequence, the 3p-Weibull model may be suitable for describing the asymmetry detected in these data. For illustrative purposes, we consider the 2-p Weibull distribution as a competing model which has been widely used for ceramic material.

The histogram of the data and the Weibull plotting plot (WPP) using the 2-p Weibull model are shown in Figs. 16a and 16b, respectively. The 2-p Weibull model may be an initial choice because of its positively skewed density shape and its simplicity. However, one can clearly see from Fig. 16b that the 2-p Weibull distribution could not provide a good fit because the data points do not scatter on a straight line. The 3-p Weibull model would be an appropriate model to fit the data. We will prove this by using the AR-NS algorithm and through a number of goodness-of-fit (GOF) statistics and information criteria (IC).

The AR-NS algorithm is now employed to estimate the parameters of the 2-p and 3-p Weibull models. It should be noted that we keep the same tuning parameters as for the numerical examples. Table 14 displays the optimal estimates

True values	$n$	Estimated values	Mean	Standard deviation	C.I. (95%)	
$\alpha = 3$	100	$\hat{\alpha}$	3.0949	0.8017	1.7008	4.8613
		$\hat{\theta}$	0.5586	0.2841	0.1254	1.000
$\theta = 0.5$	100	$\hat{\lambda}$	1.0241	0.1539	0.7668	1.3624
		$\mathcal{L}(\hat{\vartheta} t_i)$	-156.3107	-	-	-
$\lambda = 1$	100	$\mathcal{L}(\vartheta t_i)$	-157.7530	-	-	-
		$\hat{\alpha}$	3.0278	0.3645	2.3366	3.7804
	500	$\hat{\theta}$	0.5267	0.1641	0.2712	0.9356
		$\hat{\lambda}$	1.0029	0.0759	0.8596	1.1539
	500	$\mathcal{L}(\hat{\vartheta} t_i)$	-788.9552	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-788.4600	-	-	-
	1000	$\hat{\alpha}$	3.0094	0.2546	2.5262	3.5212
		$\hat{\theta}$	0.5129	0.1153	0.3264	0.7817
	1000	$\hat{\lambda}$	1.0012	0.0544	0.8942	1.1074
		$\mathcal{L}(\hat{\vartheta} t_i)$	-1576.3156	-	-	-
	2500	$\mathcal{L}(\vartheta t_i)$	-1577.8084	-	-	-
		$\hat{\alpha}$	3.0047	0.1627	2.6960	3.3336
	2500	$\hat{\theta}$	0.5059	0.0724	0.3802	0.6654
		$\hat{\lambda}$	1.0003	0.0345	0.9324	1.0685
	2500	$\mathcal{L}(\hat{\vartheta} t_i)$	-3942.9761	-	-	-
		$\mathcal{L}(\vartheta t_i)$	-3944.4847	-	-	-

Table 10: Results obtained by implementation of the AR-NS algorithm in 10,000 replications (CE2G distribution).

Sample size	$\hat{\alpha}$		$\hat{\theta}$		$\hat{\lambda}$	
	Bias	MSE	Bias	MSE	Bias	MSE
100	0.0949	0.6516	0.0586	0.0841	0.0241	0.0243
500	0.0278	0.1336	0.0267	0.0277	0.0029	0.0058
1000	0.0094	0.0649	0.0129	0.0135	0.0012	0.0030
2500	0.0047	0.0132	0.0059	0.0026	0.0003	0.0006

Table 11: Bias and MSE values for the CE2G model,  $\vartheta = (3, 0.5, 1)$ .

307	308	322	328	328	329	331	332	335
337	343	345	347	350	352	353	355	356
357	364	371	373	374	375	376	376	381
385	388	395	402	411	413	415	456	

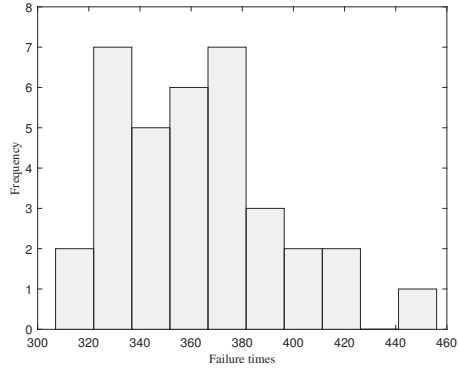
Table 12: Failure data of ceramic material in Megapascal [MPa].

Mean	Median	Standard deviation	Skewness	Kurtosis	$n$
362	356	33.1130	0.6279	3.2812	35

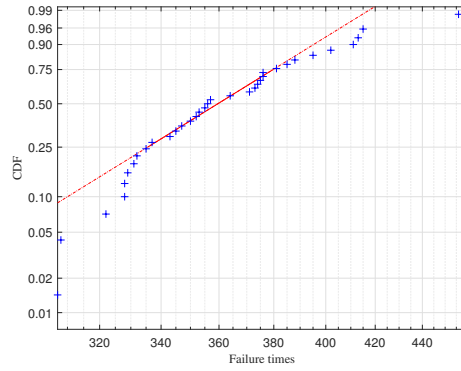
Table 13: Descriptive statistics of the failure data of ceramic material.

for the considered models and the log-likelihood value associated to each model. To illustrate how appropriate the  
 240 considered models describe failure data of ceramic material, we display in Fig. 17a the fitted PDF curves with the optimal estimates. We can visually see that 3-p Weibull model produces a better fit than the 2-p Weibull model. Fig. 17b shows the failure times and the fitted CDFs, one can see that the 3-p Weibull distribution provides a better fit to the data compared with the 2-p Weibull model.

245 The verification of the fitting performance for the fitted distributions is carried out now by comparing the values of their Kolmogorov-Smirnov, Anderson-Darling, Cramer-Von-Mises, AIC, BIC, and the AICc goodness-of-fit statistics.



(a)

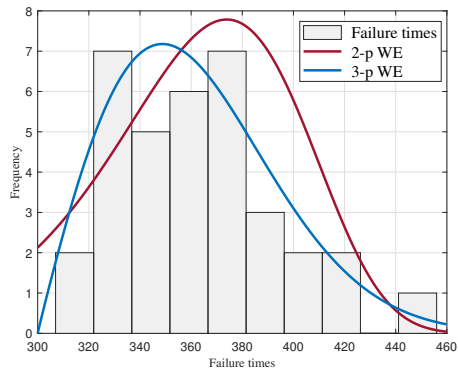


(b)

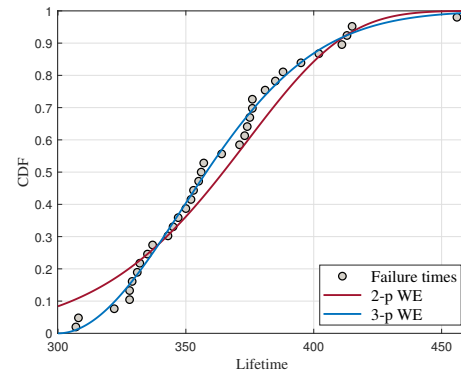
Figure 16: (a) Histogram of the failure data of ceramic material, (b) Weibull probability plot.

Distribution	Estimates			$-\mathcal{L}(\hat{\vartheta} t_i)$
	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\gamma}$	
2-p Weibull	10.6020	377.4461	-	175.4064
3-p Weibull	123	69.8395	300.0082	169.9322

Table 14: Optimal estimates and the minus log-likelihood function values for the 2-p and the 3-p Weibull models.



(a)



(b)

Figure 17: (a) Histogram of the data and the fitted PDFs, (b) failure times and the fitted CDFs.

These statistics have been widely used in the literature to discriminate between competing models when more than one model could fit the data. In general, the distribution with the smallest goodness-of-fit statistics is the best. The analytical expressions for the goodness-of-fit measures are given below:

- **Kolmogorov-Smirnov (KS) criterion** [58]:

$$KS = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - z_i, z_i - \frac{i-1}{n} \right\} \quad (20)$$

- **The Cramer-von-Mises (CVM) distance** [59]:

$$CVM = \frac{1}{12n} + \sum_{i=1}^n \left[ z_i - \frac{2i-1}{2n} \right]^2 \quad (21)$$

- **The Anderson-Darling (AD) distance** [60]:

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln(z_i) + \ln(1 - z_{n+1-i})] \quad (22)$$

- **Akaike information criterion (AIC) due to Akaike (1974)** [61]:

$$AIC = -2\widehat{\mathcal{L}} + 2k \quad (23)$$

- **Bayes information criterion (BIC) due to Schwarz (1978)** [62]:

$$BIC = -2\widehat{\mathcal{L}} + k \log(n) \quad (24)$$

250

- **AIC with a correction (AICc) due to Hurvich and Tsai (1989)** [63]:

$$AICc = -2\widehat{\mathcal{L}} + \frac{2k(k+1)}{n-k-1} \quad (25)$$

where,  $\widehat{\mathcal{L}}$ ,  $k$ ,  $n$  and  $z_i = \widehat{F}(\cdot)$  correspond to the estimate of the model maximised log-likelihood function, number of parameters in the distribution, the sample size of the fitted data, and the estimated distribution function under the ordered data, respectively.

255

Results from the model fittings are reported in Table 15. One can see that the 3-p Weibull distribution performs better due to lower information criterions and GOF statistics values. As a result, the superiority of the 3p-Weibull model is obvious for this data set.

260

Fig. 21 shows the histograms of the optimal estimates for the 2-p and the 3-p Weibull models. It can be seen that the five parameters have been identified with small uncertainty. Figs. 20a-h display for the selected model, the evolution of the particles inside the ellipsoids over some intermediate populations and how the prior parameter space

IC and GOF statistics						
Model	AIC	BIC	AICc	KS	CVM	AD
2-p Weibull	354.81	357.92	351.18	0.1257	0.8753	0.3534
3-p Weibull	345.86	350.53	340.64	0.0726	0.0225	0.1913

Table 15: IC and goodness-of-fit statistics of the fitted models for the failure data of ceramic material.

shrinks down by gradually reducing the tolerance threshold values. Finally, Fig. 19 displays the acceptance rates over the populations for the considered models. It is measured by dividing the number of particles required to replenish a population by the total number of simulations at each population. It can be seen that the AR-NS algorithm maintains  
 265 after a few populations a high acceptance rates around 50 ~ 70% until convergence.

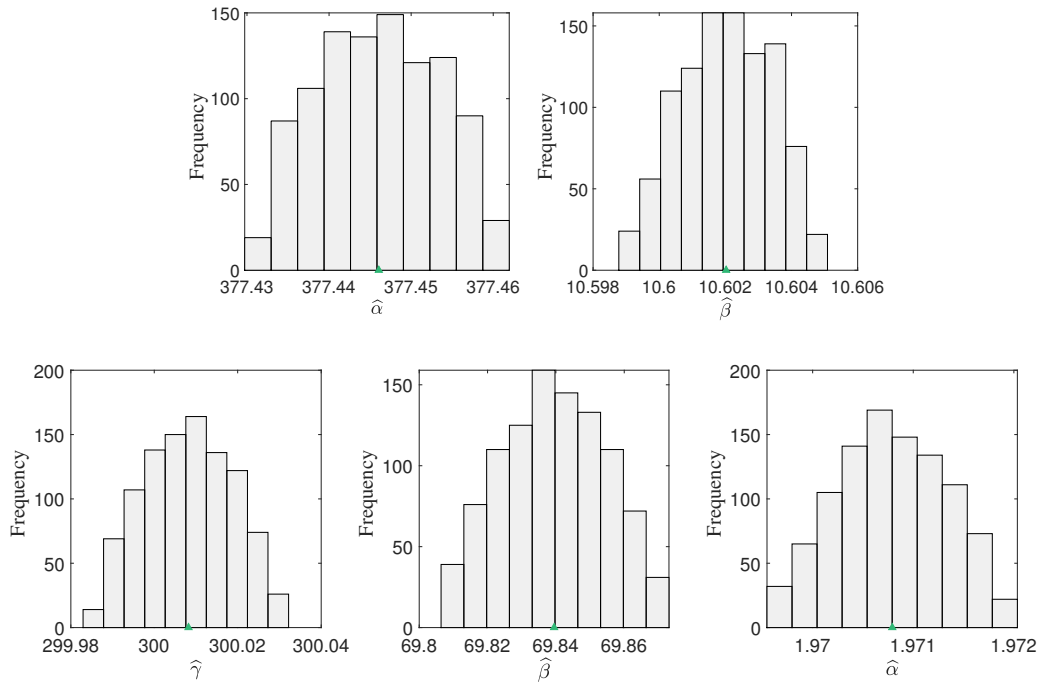


Figure 18: Histograms of the optimal estimates obtained for the 2-p (first row) and 3-p (second row) Weibull distributions.

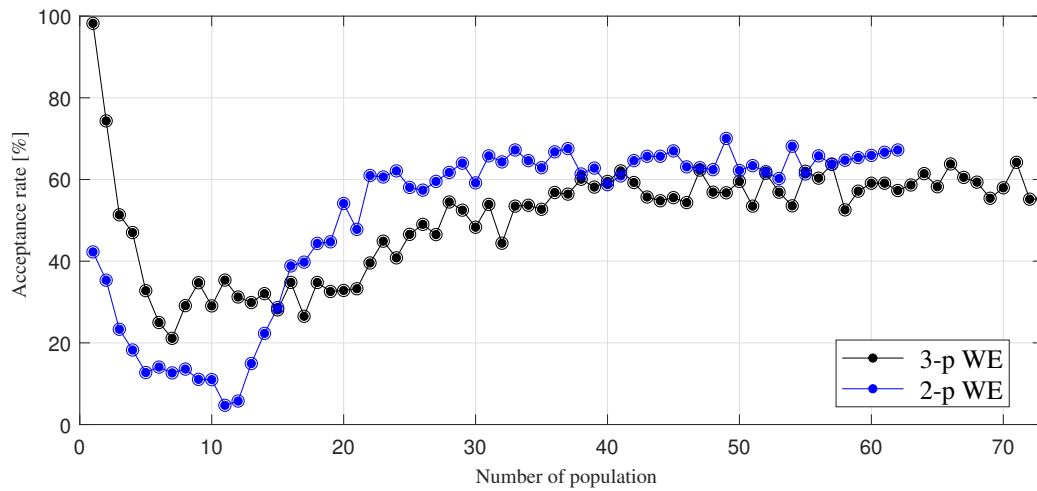


Figure 19: Acceptance rates over the populations for the 2-p and 3-p Weibull distributions.

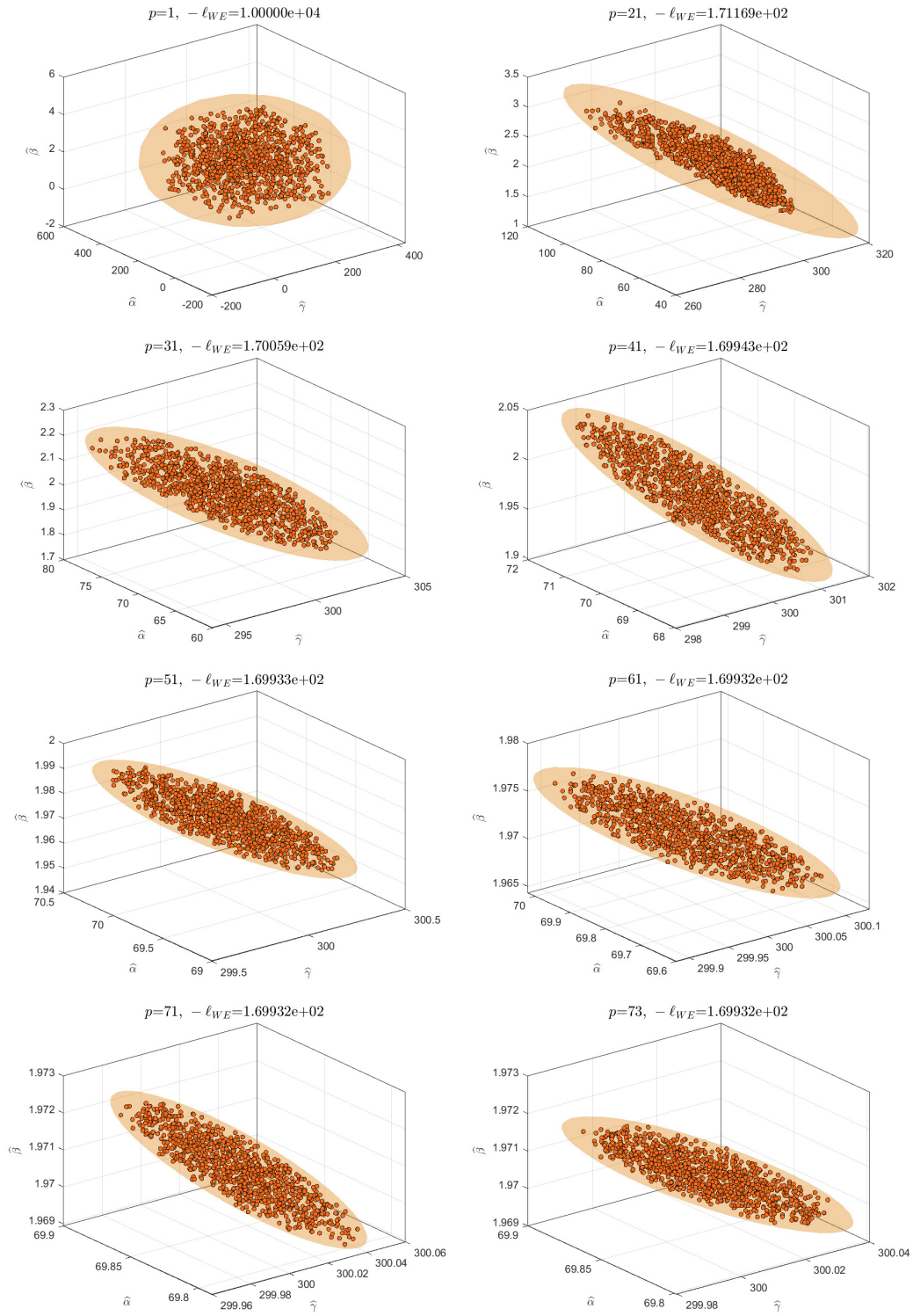


Figure 20: Evolution of the particles for the 3-p Weibull model over some intermediate populations.



## 6.2. Example # 6: Fatigue lives data set

The second data set is bearings' fatigue life data initially reported by McCool [64] and later reanalysed by a number of authors; see for example, Cohen et al. [65]; Cohen and Whitten [66] and Upadhyay and Mukherjee [67]. The data consist of fatigue life in hours of 10 bearings of a certain type given in Table 16.

152.7	172.0	172.5	173.3	193.0
204.7	216.5	234.9	262.6	422.6

Table 16: Fatigue lives in hours of 10 bearings from McCool [64].

Let consider the problem of parameter estimation. For this data set, when the maximum likelihood method is used to estimate the parameters of the 3p-Weibull distribution, there is no convergent solution [3]. To overcome this issue, the AR-NS algorithm is used by using the Wasserstein distance. In the next section, we first review the definitions of the Wasserstein distance.

### 6.2.1. Wasserstein distance

The Wasserstein distance is a powerful metric based on the theory of optimal mass transport [68]. It gives a natural measure of the distance between two distributions with a wide range of applications [69, 70]. Let  $\mathcal{P}(\mathbb{R}^d)$  be the set of Borel probability measures on  $\mathbb{R}^d$  and let  $\mathcal{P}_p(\mathbb{R}^d)$  be the subset of such measures with a finite moment of order  $p \in [1, \infty)$ . For  $P, Q \in \mathcal{P}(\mathbb{R}^d)$ , let  $\Gamma(P, Q)$  be the set of probability measures  $\gamma$  on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $P$  and  $Q$ , i.e., such that  $\gamma(B \times \mathbb{R}^d) = P(B)$  and  $\gamma(\mathbb{R}^d \times B) = Q(B)$  for Borel sets  $B \subseteq \mathbb{R}^d$ . The p-Wasserstein distance between  $P, Q \in \mathcal{P}_p(\mathbb{R}^d)$  is defined by:

$$W_p(P, Q) = \left( \inf_{\gamma \in \Gamma(P, Q)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^p d\gamma(x, y) \right)^{1/p} \quad (26)$$

$\|\cdot\|$  denotes the Euclidean norm. In terms of random variables  $X$  and  $Y$  with laws  $P$  and  $Q$ , respectively, the p-Wasserstein distance is the smallest value of  $\{E(\|X - Y\|^p)\}^{1/p}$  over all possible joint distributions  $\gamma \in \Gamma(P, Q)$  of  $(X, Y)$ .

The p-Wasserstein distance  $W_p$  defines a metric on  $\mathcal{P}_p(\mathbb{R}^d)$ , which thereby becomes a complete separable metric space [68]. Convergence in the  $W_p$  metric is equivalent to weak convergence plus convergence of moments of order  $p$ ; see for instance [71] and [68].

For univariate distributions  $P$  and  $Q$  with distributions functions  $F$  and  $G$ , the p-Wasserstein distance is defined as:

$$W_p(P, Q) = \left( \int_0^1 |F^{-1}(u) - G^{-1}(u)|^p du \right)^{1/p} \quad (27)$$

where  $F^{-1}$  and  $G^{-1}$  are the quantile functions of  $P$  and  $Q$ , respectively.

The optimal estimates given by minimising the Wasserstein distance are given in Table 17. Fig. 21 shows a comparison of the observed cumulative probability and the expected cumulative distribution curves obtained by the AR-NS algorithm using the Wasserstein distance. One can clearly see that overall the optimal estimates fit very well the data.

Distribution	Estimates			Distance measure
	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\gamma}$	
3-p Weibull	1.3190	73.5351	142.9422	0.0272

Table 17: Optimal estimates of the 3-p Weibull model using the Wasserstein distance.

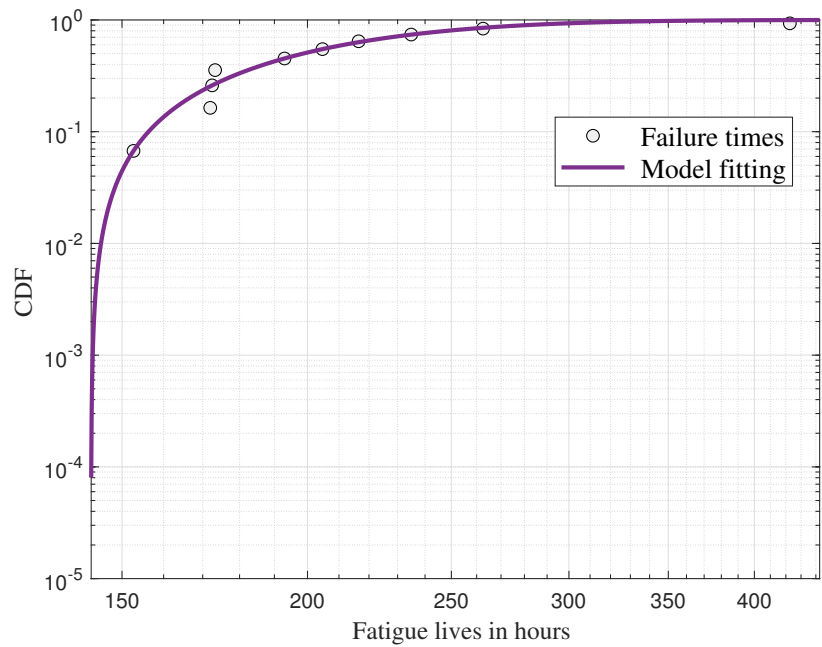


Figure 21: Empirical data and fitted CDF using the 3-p Weibull model.

## 7. Conclusion

295 In this paper, an efficient acceptance-rejection algorithm coupled with an elegant nested sampling technique for estimating the 3-p Weibull distribution is proposed. The algorithm approaches gradually the optimal solutions in an efficient way by reducing progressively a pre-specified tolerance threshold and through a proper reweighting scheme to

gain in efficiency. We conducted simulation studies to demonstrate the efficiency and the robustness of the algorithm. It has shown through various numerical examples and real-world engineering cases that the AR-NS algorithm yields quite efficient results. Consequently, the AR-NS algorithm which is rooted on a well-developed and widely used ABC method can be considered as an effective and an efficient population-based algorithm for parameter estimation available for the reliability and statistical analyst. In addition to being simple, stable, and easily implemented, the algorithm gives the entire distribution of the optimal solutions rather than a single point estimates.

In the framework of this study, although, we have only considered only two parametric distributions, the proposed AR-NS algorithm can be applied to any other family of lifetime distributions having the same or larger number of parameters. More importantly, the algorithm is fairly flexible as it can be implemented with other minimum distance estimators (Cramér-von Mises, Anderson-Darling, etc.) to estimate the model parameters. This property is of great interest mainly when the derivation of the likelihood function is not trivial, intractable, or when one has to deal with small/moderate sample sizes which is often the case in practical applications and where the use of likelihood-based methods are questionable. Moreover, the proposed optimisation scheme, which is based on an acceptance/rejection mechanism, is suitable for a parallel implementation in a high-performance computing environment. Work in these directions is currently under way.

## References

- [1] Murthy DNP, Bulmer M, Eccleston JA. Weibull Model Selection for Reliability Modelling. *Reliability Engineering and System Safety* 86 (3): (2004) 257-267.
- [2] Murthy, DNP, Xie, M, Jiang, R, 2004. *Weibull Models*. Wiley, New York.
- [3] H. Nagatsuka, T. Kamakurab, N. Balakrishnan. A consistent method of estimation for the three-parameter Weibull distribution. *Computational Statistics and Data Analysis* 58 (2013) 210-226.
- [4] Ling J, Pan J. A new method for selection of population distribution and parameter estimation. *Reliability Engineering and System Safety* 60 (1998) 247-255.
- [5] Lawless, J.F., 1982. *Statistical Models and Methods for Lifetime Data*. Wiley, New York.
- [6] Nelson, W., 1982. *Applied Life Data Analysis*. Wiley, New York.
- [7] S.A. Akdag, A. Dinler. A new method to estimate Weibull parameters for wind energy applications, *Energy Convers. Manage.* 50 (7) (2009) 1761–1766.
- [8] R. Bedi, R. Chandra. Fatigue-life distributions and failure probability for glass–fiber reinforced polymeric composites. *Compos. Sci. Technol.* 69 (9) (2009) 1381–1387.

- [9] A. Khalatbari, K. Jenab, A. Varvani-Farahani. Lifetime estimation of composite bone joint screws. *Int. J. Comput. Theory Eng.* 1 (5) (2009) 474–477.
- 330 [10] J.F. Castet, H.S. Joseph. Satellite and satellite subsystems reliability: statistical data analysis and modeling. *Reliability Engineering and System Safety* 94 (11) (2009) 1718–1728.
- [11] Z. Tan. A new approach to MLE of Weibull distribution with interval data. *Reliability Engineering and System Safety* 94 (2) (2009) 394–403.
- [12] K. Jenab, A. Kazemini, DS Liu. A computational intelligent approach to estimate the Weibull parameters. *Int. J. Ind. Syst. Eng.* 6 (1) (2010) 62–78.
- 335 [13] B. Abbasi, A.H. Eshragh-Jahromi, J. Arkat, M. Hosseinkouchack, Estimating the parameters of Weibull distribution using simulated annealing algorithm. *Applied Mathematics and Computation* 183 (1) (2006) 85–93.
- [14] KP. Soman, KB. Misra. A least square estimation of three parameters of a Weibull distribution. *Microelectron. Reliab.*, Vol. 32, No. 3, pp. 303-305, 1992.
- 340 [15] D. Jukić, M. Benšić, R. Scitovski. On the existence of the nonlinear weighted least squares estimate for a three-parameter Weibull distribution. *Computational Statistics and Data Analysis* 52 (2008) 4502–4511.
- [16] H. Rockette, C. Antle, and L. A. Klimko. Maximum likelihood estimation with the Weibull model. *J. American Statistical Association*, Vol. 69, pp. 246-249, 1974.
- [17] D. Markovic, D. Jukic, M. Benšić. Nonlinear weighted least squares estimation of a three-parameter Weibull density with a nonparametric start. *J. Comput. Appl. Math.* 228 (1) (2009) 304–312.
- 345 [18] Hideki Nagatsuka, Toshinari Kamakura, N. Balakrishnan. A consistent method of estimation for the three-parameter Weibull distribution. *Computational Statistics and Data Analysis* 58 (2013) 210–226.
- [19] Kantar YM, Senoglu B. A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter. *Comput Geosci* 2008;34(12):1900–9.
- 350 [20] Puthenpura S, Sinha NK. Modified maximum likelihood method for the robust estimation of system parameters from very noisy data. *Automatica* 1986;22(2):231–5.
- [21] Vaughan DC. On the Tiku-Suresh method of estimation. *Commun Stat-Theory Methods* 1992;21(2):451–69.
- [22] Vaughan DC. The generalized secant hyperbolic distribution and its properties. *Commun Stat-Theory Methods* 2002;31(2):219–38.
- 355 [23] Acitas S, Kasap P, Senoglu B, Arslan O. Robust estimation with the skew  $t_2$  distribution. *Pak J Stat* 2013;29(8):409–30.

- [24] Acitas S, Kasap P, Senoglu B, Arslan O. One-step M-estimators: Jones and Faddy's skewed t-distribution. *J Appl Stat* 2013;40(3):1545–60.
- [25] B. Abbasi, L. Rabelo, M. Hosseinkouchack. Estimating parameters of the three-parameter Weibull distribution using neural network. *Eur. J. Ind. Eng.* 2(4) (2008) 428–445.
- 360
- [26] H. Hasan Örkücü, Ertuğrul Aksoy, Mustafa İsa Doğan. Estimating the parameters of 3-p Weibull distribution through differential evolution. *Applied Mathematics and Computation* 251 (2015) 211–224.
- [27] Nosal, M., & Nosal, E. M. (2003). Three-parameter Weibull generator for replacing missing observations. In *Proceedings of the Hawaii International conference on statistics and related fields, Honolulu, Hawaii, June 5–8.*
- [28] Sukru Acitas, Cagdas Hakan Aladag, Birdal Senoglu. A new approach for estimating the parameters of Weibull distribution via particle swarm optimization: An application to the strengths of glass fibre data. *Reliability Engineering and System Safety* 183 (2019) 116–127.
- 365
- [29] H. Hasan Örkücü, VS. Özsoy, E. Aksoy, M. İsa Dogan. Estimating the parameters of 3-p Weibull distribution using particle swarm optimization: A comprehensive experimental comparison. *Applied Mathematics and Computation* 268 (2015): 201-226.
- 370
- [30] B. Abbasi, S. Taghi Akhavan Niaki, M. Abdi Khalife, Y. Faize. A hybrid variable neighborhood search and simulated annealing algorithm to estimate the three parameters of the Weibull distribution. *Expert Systems with Applications* 38 (2011) 700–708.
- [31] R. Luus, M. Jammer, Estimation of parameters in 3-parameter Weibull probability distribution functions, *Hung. J. Ind. Chem.* 33(1-2)(2005)69–73.
- 375
- [32] D. Cousineau, Fitting the three-parameter Weibull distribution: review and evaluation of existing and new methods, *IEEE Trans. Dielectr. Electr. Insul.* 16(1)(2009) 281–288.
- [33] F. Yang, Z. Yue, Kernel density estimation of three-parameter Weibull distribution with neural network and genetic algorithm, *Appl. Math. Comput.* 247(2014)803–814.
- [34] H. Nagatsuka, N. Balakrishnan, An efficient method of parameter and quantile estimation for the three parameter Weibull distribution based on statistics invariant to unknown location parameter, *Commun. Stat. Simul. Comput.* 44(2)(2015)295–318.
- 380
- [35] D. Markovich, D. Jukic, On non linear total least squares parameter estimation problem for the three-parameter Weibull density, *Appl. Math. Modell.* 34(7)(2010)1839–1848.
- [36] A. Moeini, K. Jenab, M. Mohammadi, M. Foumani, Fitting the three-parameter Weibull distribution with Cross Entropy, *Appl. Math. Modell.* 37(9)(2013)6354–6363.
- 385

- [37] A. Bartolucci, K. Singh, A. Bartolucci, S. Bae, Applying medical survival data to estimate the three-parameter Weibull distribution by the method of probability-weighted moments, *Math. Comput. Simul.* 48 (1999) 385–392.
- [38] V. Bartkute, L. Sakalauskas, The method of three-parameter Weibull distribution estimation, *Acta et Commentationes Universitatis Tartuensis de Mathematica* 12 (2008) 65–78.
- [39] D. Jukic, M. Bencic, R. Scitovski, On the existence of the nonlinear weighted least squares estimate for a three-parameter Weibull distribution, *Comput. Stat. Data Anal.* 52 (2008) 4502–4511.
- [40] D. Jukic, D. Markovich, On nonlinear weighted errors-in-variables parameter estimation problem in the three-parameter Weibull model, *Appl. Math. Comput.* 215 (2010) 3599–3609.
- [41] D. Markovich, D. Jukic, On nonlinear total least squares parameter estimation problem for the three-parameter Weibull density, *Appl. Math. Model.* 34 (2010) 1839–1848.
- [42] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, M.P.H. Stumpf, Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems, *J. Roy. Soc. Interface* 6 (2009) 187–202.
- [43] A. Ben Abdesslem, N. Dervilis, D. Wagg, K. Worden, Identification of nonlinear dynamical systems using approximate Bayesian computation based on a sequential Monte Carlo sampler, in: *International Conference on Noise and Vibration Engineering*, September 19–21, 2016, Leuven, Belgium.
- [44] Theodore Kypraios, Peter Neal, Dennis Prangle. A tutorial introduction to Bayesian inference for stochastic epidemic models using Approximate Bayesian Computation. *Mathematical Biosciences* 287 (2017) 42–53.
- [45] A. Ben Abdesslem, N. Dervilis, D. Wagg, K. Worden. Model selection and parameter estimation in structural dynamics using approximate Bayesian computation, *Mechanical Systems and Signal Processing.* 99 (2018) 306–325.
- [46] A. Ben Abdesslem, N. Dervilis, D. Wagg, K. Worden. Model selection and parameter estimation of dynamical systems using a novel variant of approximate Bayesian computation, *Mechanical Systems and Signal Processing.* 122 (2019) 364–386.
- [47] E. Jennings, M. Madigan. astroABC: An Approximate Bayesian Computation Sequential Monte Carlo sampler for cosmological parameter estimation. *Astronomy and Computing*, 19(2017)16–22.
- [48] Tanaka, MM, Francis, AR, Luciani, F. Sisson, S.A., 2006. Using approximate Bayesian computation to estimate tuberculosis transmission parameters from genotype data. *Genetics* 173, 1511-1520.
- [49] Cheng, B.C.H., Amin, N.A.K., 1983. Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society, Series B* 45, 394–403.

- [50] Smith, R.L., 1985. Maximum likelihood estimation in a class of nonregular cases. *Biometrika* 72, 67–90.
- [51] Griffiths, D.A., 1980. Interval estimation for the three-parameter lognormal distribution via the likelihood function. *Applied Statistics* 29, 58–68.
- [52] Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. *Continuous Univariate Distributions*, Vol. 1, second ed. John Wiley & Sons, New York.
- 420 [53] P. Mukherjee, D. Parkinson, AR. Liddle, A nested sampling algorithm for cosmological model selection, *Astrophysical Journal*. 638 (2006) L51–L54.
- [54] F. Feroz, MP. Hobson, M. Bridges, MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics, *Mon Not R Astron Soc*. 398 (4):(2009) 1601–1614.
- 425 [55] JR. Shaw, M. Bidges, MP Hobson, Efficient Bayesian inference for multimodal problems in cosmology. *Mon. Not. R. Astron. Soc*. 000, 1–7 (2006).
- [56] Francisco Louzada, Vitor Marchi, James Carpenter. The Complementary Exponentiated Exponential Geometric Lifetime Distribution. *Journal of Probability and Statistics*. vol. 2013, Article ID 502159, 12 pages, 2013.
- [57] S. F. Duffy, J. L. Palko, and J. P. Gyekenyesi. Structural reliability analysis of laminated CMC components. *Journal of Engineering for Gas Turbines and Power*, 115, (1), 103–108, 1993.
- 430 [58] Kolmogorov, A., 1933. Sulla determinazione empirica di una lgge di distribuzione. *Inst. Ital. Attuari, Giorn*.4, 83–91.
- [59] Anderson TW (1962). On the Distribution of the Two-Sample Cramer-von Mises Criterion. *The Annals of Mathematical Statistics*, 33(3), 1148–1159.
- 435 [60] Anderson TW, Darling DA. A test of goodness of fit. *J Am Stat Assoc* 1954;49:765–9.
- [61] Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716–723.
- [62] Schwarz, G.E., 1978. Estimating the dimension of a model. *Ann. Stat.*6, 461–464.
- [63] Hurvich, C.M., Tsai, C.L., 1989. Regression and time series model selection in small samples. *Biometrika* 76, 297–307.
- 440 [64] McCool, J.I., 1974. Inferential techniques for Weibull populations. Technical Report TR 74-0180, Wright Patterson Air Force Base, Ohio.
- [65] Cohen, A.C., Whitten, B.J., Ding, Y., 1984. Modified moment estimation for the three-parameter Weibull distribution. *Journal of Quality Technology* 16,159–167.

- [66] Cohen, A.C., Whitten, B.J., 1988. *Parameter Estimation in Reliability and Life Span Models*. Marcel Dekker, New York.
- [67] Upadhyay, S.K., Mukherjee, B., 2008. Assessing the value of the threshold parameter in the Weibull distribution using Bayes paradigm. *IEEE Transactions on Reliability* 36, 489–497.
- [68] Villani, C. (2009). *Optimal Transport: Old and New*. Springer-Verlag, Berlin. von Mises, R. E. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Julius Springer, Berlin.
- [69] M. Halin, G. Mordant, J. Segers, 2020. *Multivariate Goodness-of-Fit Tests Based on Wasserstein Distance*. Working Papers ECARES 2020-06, ULB – Universite Libre de Bruxelles.
- [70] Bernton, E., Jacob, P.E., Gerber, M., & Robert, C. P. (2019). Approximate Bayesian computation with the Wasserstein distance. *Journal of the Royal Statistical Society: Series B*, 81(2), 235-269.
- [71] Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics* 9 1196–1217.