



Application of approximate Bayesian computation for estimation of modified weibull distribution parameters for natural fiber strength with high uncertainty

M. Ravandi^{1,*}  and P. Hajizadeh¹

¹Department of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran

Received: 28 September 2021

Accepted: 27 December 2021

Published online:
13 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

ABSTRACT

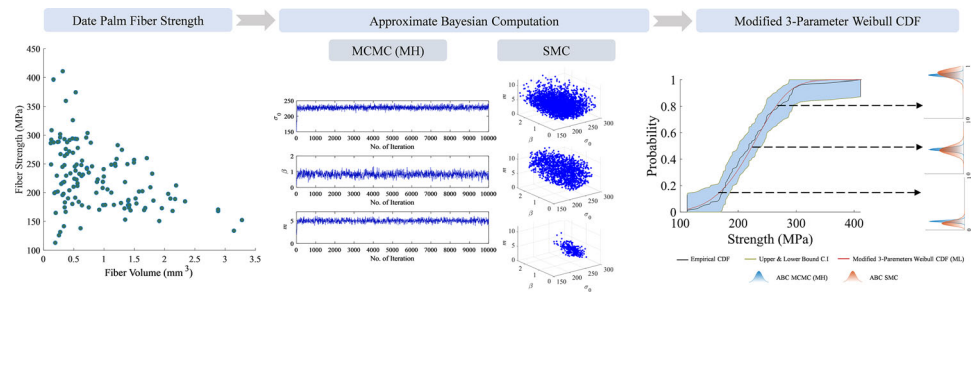
Despite the unique advantages of natural fibers as a reinforcement in polymer composites, they have high natural variability in their mechanical properties, resulting in significant uncertainties in the properties of natural fiber composites. This study aims to propose a multilevel framework based on the Approximate Bayesian Computation (ABC) to analyze the uncertainty of fitting the Weibull distribution to the strength data of date palm fibers. Two computationally efficient algorithms of the ABC, namely the Metropolis–Hasting as a family of Markov Chain Monte Carlo and the Sequential Monte Carlo (SMC), are employed for estimating the highest density interval of the fitting parameters of the modified 3-parameter Weibull distribution, and their performances are evaluated. Moreover, appropriate probability distributions that best fit the estimated parameters are determined based on the goodness of fit to describe their characteristics. It is found that the SMC algorithm leads to a higher scatter in the posterior predictive distribution of the fitting parameters. The results suggest that the uncertainty of the fitting parameters should be considered to have a reliable model for the probability of natural fiber failure.

Handling Editor: Gregory Rutledge.

Address correspondence to E-mail: m.ravandi@u.nus.edu

<https://doi.org/10.1007/s10853-021-06850-w>

GRAPHICAL ABSTRACT



Introduction

Natural (plant) fibers are promising reinforcements to replace synthetic glass fibers in many nonstructural and semi-structural applications thanks to their desirable properties such as high specific strength and stiffness, good acoustic insulation, vibration damping, and lower environmental impacts [1–4]. As a result, the use of these natural reinforcements in polymer composite materials has recently gained substantial attention. However, unlike synthetic fibers, the properties of natural fibers have a relatively large uncertainty that arises from their natural variabilities, extraction, and processing. These uncertainty sources are generally hard or impossible to control due to their stochastic nature. Therefore, it is necessary to consider the variation of properties, such as fiber strength, when modeling or predicting the behavior of natural fiber-reinforced composites.

Weibull statistics is a popular data processing tool for interpreting life data such as time to failure and strength. Weibull distribution parameters can be adjusted to fit many life distributions that are compatible with the weakest link theory or have a sudden failure characteristic. For example, the Weibull distribution has been widely used to process the strength data of brittle materials, such as glass and carbon fibers, in which the most severe flaw controls the strength [2, 5, 6]. It has also been successfully applied to a variety of natural fibers such as hemp, flax, sisal, agave, and bamboo fibers [7–14], as natural fiber failure follows the “brittleness” and “weakest link” assumptions of the Weibull distribution [15].

Although the basic version of the Weibull distribution is the 2-parameter model, it usually is not capable of accurately describing the strength distribution of technical natural fibers [8, 15]. Many efforts have been made to improve the accuracy of the Weibull distribution by considering the influential factors of the failure mechanism in the distribution model [4, 16–18]. The modified 3-parameter Weibull distribution is an improved version in which the third parameter is introduced such that the effect of defect density is considered as a function of fiber geometry. For instance, E. Trujillo et al. [15] applied the 3-parameter Weibull model to the strength data of bamboo fibers tested at different gauge lengths and showed that the model can predict the strength of fibers against their length with an acceptable degree of accuracy. According to the modified 3-parameter Weibull distribution, the probability of failure of a fiber with a volume V at stress less than or equal to σ is given by [15]:

$$F(\sigma|\sigma_0, \beta, m) = 1 - \exp\left(-\left(\frac{V}{V_0}\right)^\beta \left(\frac{\sigma}{\sigma_0}\right)^m\right) \quad (1)$$

where m is the shape parameter, σ_0 is the scale parameter, and β is the geometry sensitivity. In Eq. (1), V is the fiber volume, and V_0 is the reference volume, which is an arbitrary parameter used to normalize the effect of the fiber volume. Interpretations of values of the distribution parameters are available in [14, 15]. There are several classical methods for estimating the Weibull distribution parameters from the experimentally measured dataset, such as Linear Regression (LR), Maximum

Likelihood (ML), and Method of Moments (MM). However, depending on the scatter of the experimental data and the corresponding confidence interval, these methods may result in different values for the distribution parameters. This divergence in the estimated parameters becomes more pronounced for the natural fibers due to the high variation in their strength. Therefore, the classical methods to estimate the Weibull parameters are not reliable approaches for finding the best fitting distribution to the observed strength data.

Unlike the classical methods, the Bayesian approach is a comprehensive alternative for parameter estimation and model calibration, which has gained a lot of attention in recent years. The Bayesian approach is not only able to avoid over-fitting but also provides a credible interval for the fitting parameters, which is a valuable quantity for evaluating the uncertainty of the unknown parameters [19]. The latter is the main advantage of the Bayesian estimator when the scatter of the experimental data is high. In fact, the Bayesian estimator can describe the uncertainty of the estimated parameters through the posterior predictive distribution. Using the Bayesian inference, the posterior distribution is obtained based on our prior beliefs about the parameters and the assessment of their fit to the experimental data [20, 21].

Amongy et al. [22] recently brought up a Bayesian approach to estimate the credible intervals of generalized Weibull distribution parameters in the case of type-II censored data. They applied their model to carbon fiber strength data. Chako et al. [23] have used a Bayesian estimator to obtain the parameters of the Weibull distribution for analyzing competing risk data with binomial removals, which is applicable to specific failure data sets. Ducros et al. [24] proposed a Bayesian restoration maximization approach for a set of heterogeneous data to fit a mixture of 2-parameter Weibull distributions.

Using the Bayesian approach, we are interested in the posterior distribution of the estimated parameters. The posterior probability according to the Bayesian inference is given by:

$$p(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int f(\mathbf{x}|\theta)\pi(\theta)} \quad (2)$$

where \mathbf{x} is the observed data and $\theta(\theta \subseteq R^q, q \geq 1)$ is the parameters vector, $\pi(\theta)$ denotes the prior distribution, and $f(\mathbf{x}|\theta)$ is the likelihood function that is derived from a statistical model for the observed

data. The prior probability is subjectively determined in advance; however, computing the likelihood function in closed-form is not possible in most cases [24–27]. The term $\int f(\mathbf{x}|\theta)\pi(\theta)$ is the normalization constant. As a result, the posterior distribution is always directly proportional to the product of the likelihood function and the prior distribution.

Approximation Bayesian Computation (ABC) is a computational method based on the Bayesian inference, seeking to estimate the posterior distributions of model parameters computationally rather than analytically. The ABC approach is one of the popular methods for assessing statistical models, particularly for analyzing complex problems. This approach facilitates the estimation of the posterior distribution by bypassing the evaluation of the likelihood function of the conventional Bayesian inference. As a result, various methods have been introduced under this scheme over the last years [28–32]. However, it has been shown that the application of the direct Monte Carlo to implement the ABC is not computationally efficient, particularly for models with high dimensions. Hence, various algorithms have been proposed to enhance the efficiency of the ABC [33]. One way is to take advantage of the Markov Chain Monte Carlo (MCMC) sampling. Another way is to apply sequential importance sampling with some variations called Sequential Monte Carlo (SMC) [34].

In this study, for the first time, we applied the ABC approach to fit a modified 3-parameter Weibull distribution to the experimentally measured strengths of date palm fibers and determined the highest posterior density intervals for the non-deterministic distribution parameters. The Metropolis–Hasting (MH), as a class of MCMC sampling, and an SMC algorithm were employed to obtain the posterior distributions of the parameters. This study provides a reliable method for predicting the fiber strength, considering fiber dimensions.

Estimation of modified weibull distribution using ABC

In order to estimate the modified 3-parameter Weibull distribution parameters, the ABC algorithm is applied to the strength data of date palm fibers measured at seven different gauge lengths. Here, the uncertainty of the distribution parameters (i.e., posterior distribution) is modeled by fitting the most

appropriate statistical distribution based on the chi-square goodness of fit test. The general idea of the ABC algorithm can be described as follows:

- Generate a family of random parameters from prior distribution (previous belief)
- Simulate data according to generative model using generated parameters
- Compare the simulated data and data observation (update our belief)
- Repeat this process till the posterior parameters converge to stationary distributions

The prior distributions are chosen to be exponential with parameters $\lambda_1, \lambda_2, \lambda_3$ for σ_0, β , and m , respectively. The likelihood of the modified Weibull distribution, considering the model independence, can be written as:

$$f(\mathbf{x}|\boldsymbol{\theta}) = f(x_1|\boldsymbol{\theta})f(x_2|\boldsymbol{\theta})\dots f(x_n|\boldsymbol{\theta}) \quad (3)$$

thus

$$f(\mathbf{x}|\boldsymbol{\theta}) = \prod_{i=1}^n \left[\frac{m}{\sigma_0^m} \left(\frac{V_i}{V_0} \right)^\beta (\sigma_i)^{m-1} \right] \exp \left(- \sum_{i=1}^n \left(\frac{V_i}{V_0} \right)^\beta \left(\frac{\sigma_i}{\sigma_0} \right)^m \right) \quad (4)$$

where n represents the number of the observed data and vector $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is our data observation. In order to compare the simulated and real data, we calculate the norm of the difference vector and compare it with a threshold ϵ . The threshold value depends on the level of accuracy. Although various methods based on quantiles have been suggested in the literature to determine optimal ϵ , they do not apply to all sorts of models [35, 36]. The comparison process is carried out as follows:

Algorithm: Comparison in ABC
 Generate $\mathbf{y} \sim f(\cdot|\boldsymbol{\theta})$
 if $\|\Delta(\mathbf{y}, \mathbf{x})\| \leq \epsilon$ then
 | accept parameters vector $\boldsymbol{\theta}$
 else
 | reject parameters vector $\boldsymbol{\theta}$
 end

where $\|\cdot\|$ is the norm of the vector. This process is repeated until N particles are accepted, and the posterior distribution is converged to a stationary distribution. In the following, we explore different

approaches, such as MCMC and SMC, to improve the computational efficiency of the process.

Approximation bayesian computation using markov chain monte carlo (ABC MCMC)

MCMC sampling is a promising method for smartly generating a family of random variables from a population. This method uses the Markov process to produce a new random variable (i.e., new state) corresponding to the previous one (i.e., old state). From the Ergodic theorem, which is the central limit theorem (CLT) in the Markov process, and the rule of big numbers, it can be assured that this process will converge to a stationary distribution [37]. As a result, the MCMC sampling is used extensively in the ABC algorithm to facilitate the process.

Several MCMC-based algorithms have been proposed in the literature that apply to various statistical models. Among them, Metropolis–Hasting (MH), Gibbs Sampling, and Hamiltonian Monte Carlo are the most frequently used algorithms [38–40]. In this study, we aim to use the Metropolis–Hasting algorithm to efficiently generate random numbers in order to reduce computational costs.

The Metropolis–Hasting algorithm comprises two-part: generating random variables and evaluating the generated numbers. In the first part, we need to define a new distribution called proposal distribution. There have been many proposal distributions brought up by previous studies [41–43]. Here, the gamma distribution is used as the proposal distribution due to the positivity of the Weibull parameters. The gamma distribution is generally described by two parameters; the shape parameter a and the scale parameter b . The scale parameter is defined in terms of the precision parameter τ as:

$$b = \frac{1}{\tau} \quad (5)$$

Considering that the mean value of the gamma distribution is obtained by $\mu = ab$, the proposal distribution can be stated as:

$$f(x|\mu, \tau) = \text{gamma} \left(\mu\tau, \frac{1}{\tau} \right) \quad (6)$$

This simple substitution facilitates generating the Weibull parameters by proposing parameters with the precision (τ) and the mean value (μ) instead of the shape and scale parameters.

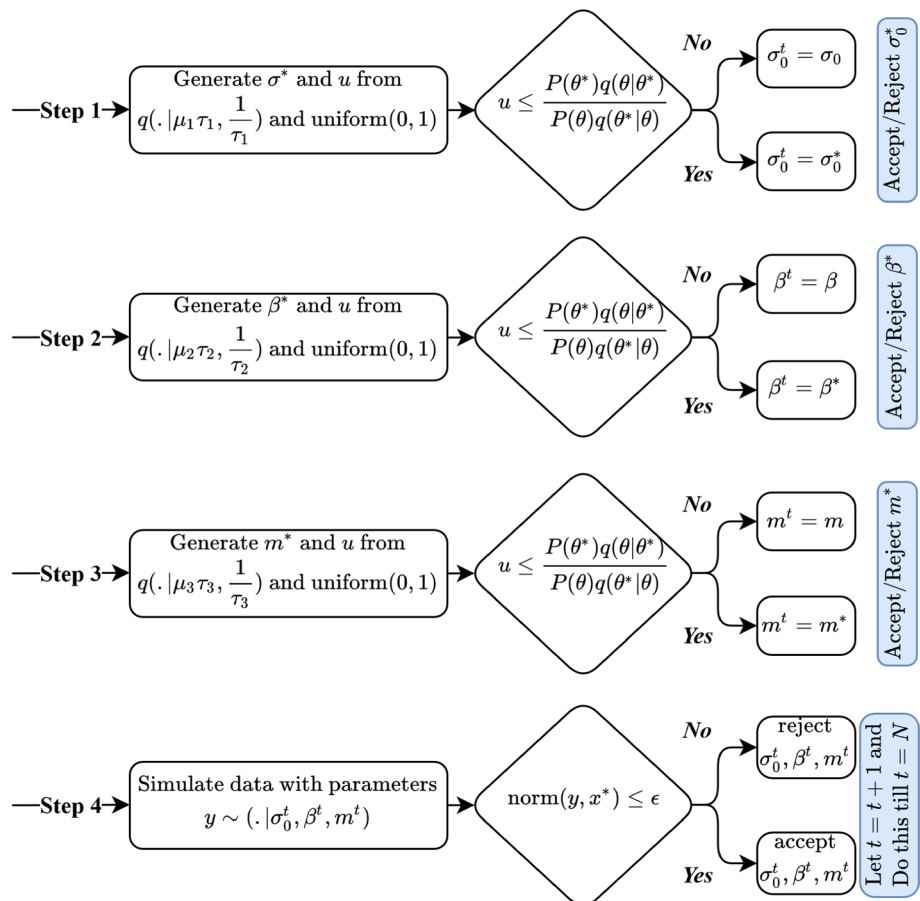
In the second part of the MH algorithm, a judgment is performed on the proposed value. The algorithm is implemented by accepting/rejecting particles component-wisely. In other words, for each parameter, a value is proposed, and then, the judgment is made. For initiating the Markov process, we need to guess an initial value for the parameters and choose a fixed value for the precision τ . This process is repeated till $t < N$, where N is the total number of iterations that the chain is run, and t is the specific iteration that the chain runs through it. The iteration is continued until the required convergence is achieved. The error ellipse and parameters distribution are visualized in the next section for different numbers of iterations. Figure 1 shows the entire process of the ABC MCMC-MH algorithm, including the judgment process.

Approximation bayesian computation using sequential monte carlo (ABC SMC)

The idea of the SMC algorithm is based on the population Monte Carlo and the importance sampling. The application of different types of SMC samplers in the ABC framework has been investigated by several researchers [44, 45]. The ABC SMC algorithm produces random particles within some measurable common spaces using particle distribution $\{P_{\theta^t}(\theta|x)\}_{1 \leq t \leq T}$. The ABC algorithm is applied to the particles with decreasing thresholds $\{\epsilon_t\}_{1 \leq t \leq T}$ in each time step t .

For $t \geq 2$, the parameters are sampled from the set of accepted particles at the previous sequence and perturbed according to a suitable perturbation kernel instead of sampling from the prior distribution. Each of these particles is associated with a weight $\{\theta^{(i,t-1)}, w^{(i,t-1)}\}_{1 \leq i \leq N}$ at each time, which is propagated corresponding to a kernel $\{K_t(\theta^{(i,t)}|\theta^{(j,t-1)})\}_{1 \leq t \leq T}$ for any particles $\theta^{(i,t)}$ and

Figure 1 Flowchart of the ABC MCMC using Metropolis–Hasting algorithm.



$\theta^{(j,t-1)}$. Note that only the first sequence particles are sampled from the prior distribution. Thus, the choice of the kernel is so influential on the acceptance rate and the efficiency of the ABC SMC.

Here, the uniform kernel is employed for perturbing the particles [34]. In this kernel, each component $\theta_j, j \in \{1, 2, \dots, d\}$ of the parameter vector $\theta = (\theta_1, \dots, \theta_d)$ is perturbed independently within $[\theta_j - \sigma_j^{(t)}; \theta_j + \sigma_j^{(t)}]$ with a density of $\frac{1}{2\sigma_j^{(t)}}$, in which $\sigma_j^{(t)}$ is the kernel width and defined as:

$$\sigma_j^{(t)} = \frac{1}{2} \left(\max_{1 \leq k \leq N} \{ \theta_j^{(k,t-1)} \} - \min_{1 \leq k \leq N} \{ \theta_j^{(k,t-1)} \} \right) \quad (7)$$

This sequence is repeated until $t = T$, when the measurable common space of particles converges to a stationary distribution. Another influential factor in ABC SMC efficiency is the sequence of decreasing thresholds, $\{\epsilon_t\}_{1 \leq t \leq T}$. Various methods have been proposed in the literature for determining the thresholds. However, they have limitations that make them not applicable to all models [35, 46]. In this study, the thresholds are determined by trial and error. Figure 2 illustrates the implementation flowchart of the ABC SMC algorithm.

These two methods are implemented using MATLAB and applied to the experimentally measured strength of the date palm fibers to determine the fitting parameters of the modified 3-parameter Weibull distribution. The performance of these methods and their computational efficiency are compared in the next section.

Results and discussion

Fiber strength data

The tensile strength of the technical date palm fibers, extracted from leaf sheath, was measured through the single fiber tensile test as per ASTM C1557-20. The fibers were tested at seven different gauge lengths; 10 mm, 15 mm, 20 mm, 25 mm, 30 mm, 40 mm, and 50 mm. The mean diameter (d) of an individual fiber was obtained by measuring the width of the fiber in six different locations along the fiber length using optical microscopy. Then, the corresponding cross-sectional area (A) of the fibers was calculated, assuming a circular cross section for fibers. Moreover, the volume of fibers was calculated

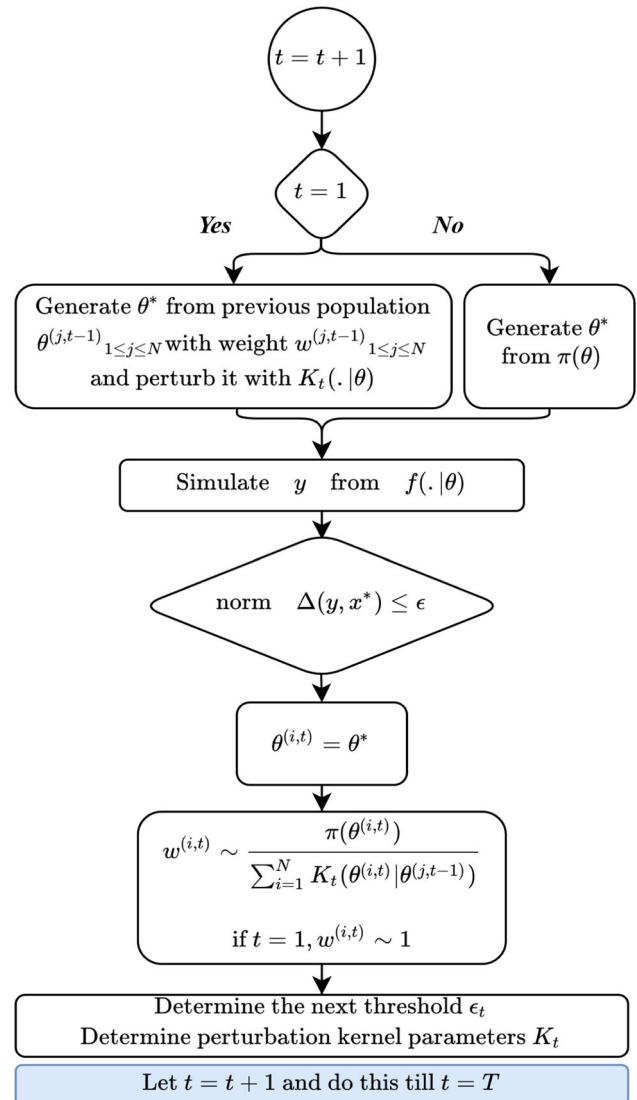


Figure 2 Flowchart of the ABC SMC algorithm.

by multiplying the cross-sectional area by the fiber length. At least twenty fibers were tested at each gauge length, resulting in at least 140 data points in total.

Figure 3 shows the distributions of fiber diameters of all tested fibers and their corresponding strengths. The range of fiber diameters was from 0.107 to 0.361 mm. To assure an unbiased comparison, a one-way ANOVA was carried out at a confidence level of 95% ($\alpha = 0.05$). Table 1 reports the mean value and standard deviation of the fiber diameters for different gauge lengths and the p value of the ANOVA test. It can be seen that the calculated P value is greater than 0.05, meaning there is no statistically significant

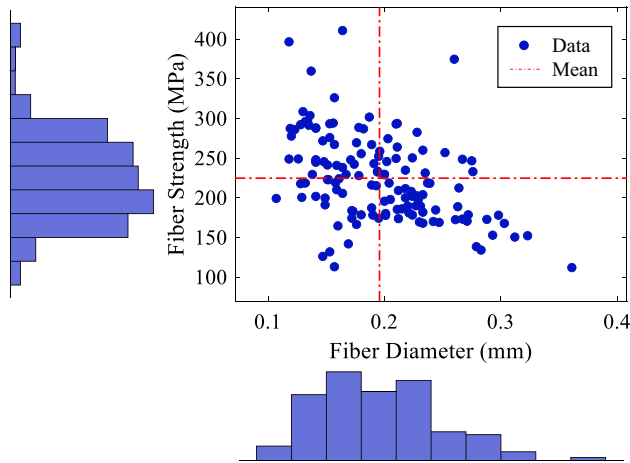


Figure 3 Scatterplot of fiber diameter versus strength with their corresponding marginal distributions.

difference in fiber diameters between the groups of different gauge lengths.

Estimation of modified 3-parameter weibull distribution parameters

The ABC MCMC and ABC SMC algorithms were applied to the strength data of the date palm fibers to estimate the fitting parameters of the modified Weibull distribution of Eq. (1), and results were presented. Their performances were also compared and evaluated.

Figure 4 illustrates the error ellipses of estimated parameters for both the ABC MCMC and ABC SMC methods with 10,000 iterations. Error ellipses are a useful graphical tool to show the pair-wise correlation between the computed values. Here, we used a 95% confidence interval to plot the error ellipses, meaning 95% of the population of each parameter falls within the ellipse. Comparing the orientation of the error ellipses, both the methods resulted in a relatively similar pair-wise correlation between the

parameters. In addition, it can be seen that all the error ellipses of the ABC SMC method are larger than that of the ABC MCMC method. The larger error ellipse is an indication of the higher scatter of the posterior distribution of the parameters. Therefore, we can conclude that the use of the ABC SMC algorithm results in higher standard deviations in the estimated parameters.

The traces of the ABC MCMC estimations of the parameters σ_0 , β , and m against the number of iterations are given in Fig. 5. The results show that, after a few iterations, the generated parameters fall within a stable range, which results in a stationary distribution.

Figure 6 visualizes the population of the parameters of the modified Weibull distribution, $\theta = (\sigma_0, \beta, m)$, at different sequences of the ABC SMC simulation. We can clearly see that the parameter space rapidly shrinks as the sequence advances, meaning that the simulation converges to the posterior distribution of the parameters.

To provide a better understanding of each algorithm’s efficiency, the ABC was implemented using the direct Monte Carlo to estimate the model parameters. Figure 7 compares the acceptance rate of the ABC MCMC, ABC SMC, and ABC (direct Monte Carlo). The results indicate that the acceptance rate of the ABC and ABC MCMC have the lowest and highest acceptance rate, respectively. Moreover, the acceptance rate of the ABC is significantly lower than the two other algorithms. Besides, it can be seen that the ABC MCMC has a higher acceptance rate compared to the ABC SMC. This result is consistent with the error ellipses of these two models shown in Fig. 4.

The histograms of the marginal posterior distribution of the estimated fitting parameters obtained from the ABC (direct Monte Carlo), ABC MCMC, and the ABC SMC after reaching the convergence are

Table 1 ANOVA analysis for different gauge lengths of fibers

Group no.	Gauge length (mm)	Mean fiber diameter (mm)	Standard deviation (mm)	<i>P</i> value
1	10	0.191	0.043	0.393
2	15	0.185	0.039	
3	20	0.203	0.046	
4	25	0.184	0.046	
5	30	0.209	0.049	
6	40	0.210	0.033	
7	50	0.197	0.046	

Figure 4 Error ellipses of the estimated parameters with respect to each other for **a** ABC MCMC (MH) and **b** ABC SMC algorithms.

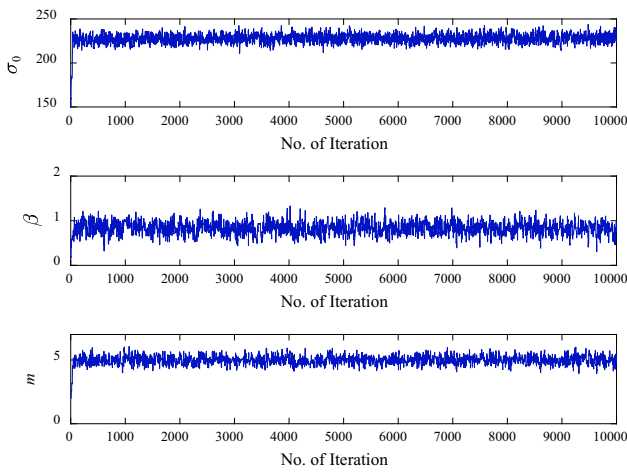
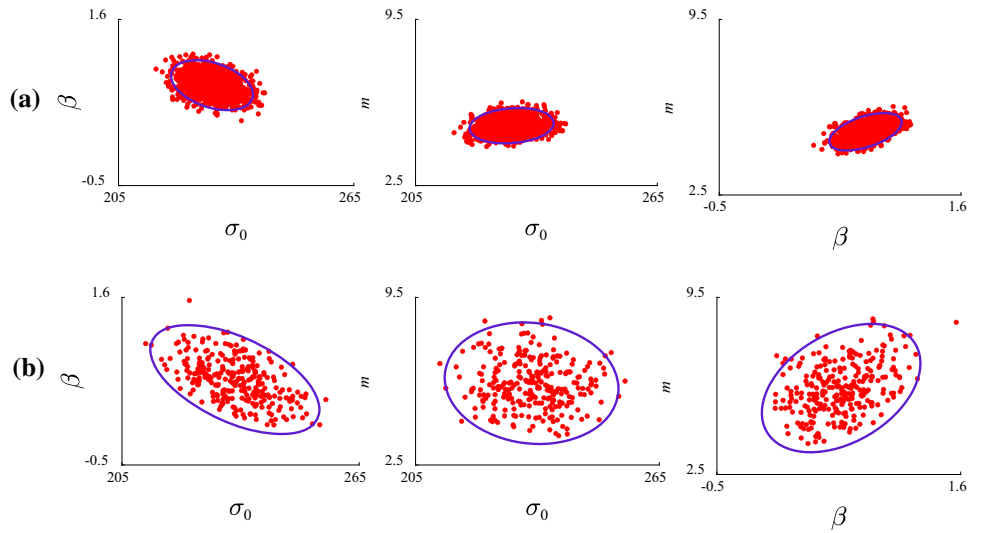


Figure 5 Trace of the generated fitting parameters $\sigma_0, \beta,$ and m against the iteration of the ABC MCMC algorithm.

compared in Fig. 8. The ABC MCMC and ABC SMC were found to converge to stationary distributions for all the parameters after 50,000 iterations, whereas the ABC required at least 2 million iterations to converge.

As can be seen, all the posterior distributions of the parameters have the central tendency and are unimodal. This enables us to model the estimated parameters characterization by finding the best fitting

Figure 6 Population of the parameters of the modified Weibull distributions in different sequences of the ABC SMC simulation: **a** sequence 2, **b** sequence 6, **c** sequence 12.

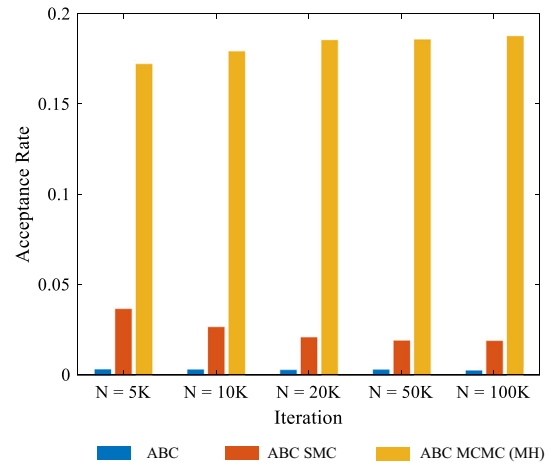
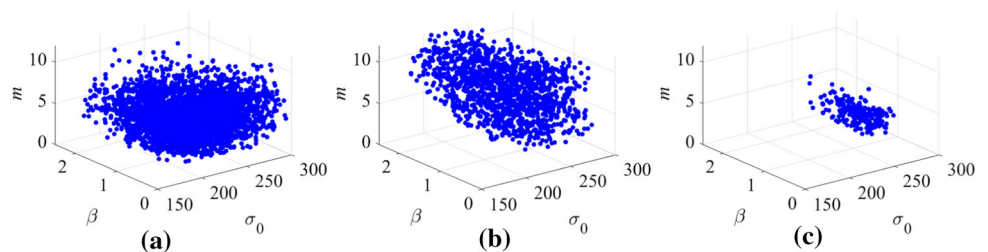


Figure 7 Acceptance rate of ABC MCMC, ABC SMC, and ABC-direct Monte Carlo.

probability distributions. However, there are some differences in the spread and mode of the estimated distribution of the parameters between the algorithms. The results obtained from the ABC SMC are very close to the ABC. Both are more scattered and have different peaks compared to those of the ABC MCMC. The larger scatter of the ABC SMC results is consistent with its bigger error ellipses shown in

Figure 8 Histograms of the fitting parameters σ_0 , β , and m .

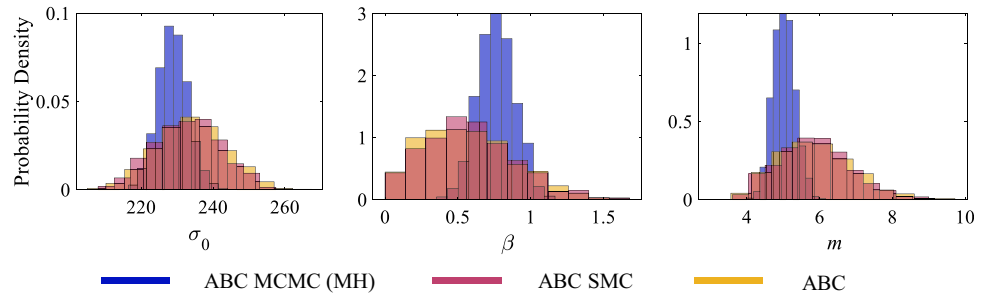


Fig. 4. This difference is associated with the sampling algorithm of the ABC SMC, where instead of the prior distribution, the parameters are sampled from the set of accepted particles at the previous stage and perturbed according to the uniform kernel. The higher scatter in the estimated parameters means higher uncertainty in the fiber strength predicted by the modified Weibull distribution. This is discussed later in the next section.

Parameters distribution

In order to characterize the uncertainty of the fitting parameters of the modified Weibull distribution, we fit probability distributions to the posterior data of the parameters σ_0 , β , and m . The chi-squared (χ^2) goodness of fit test was used to find the distributions that best fit the estimated parameters. The test is based on the null hypothesis (H_0) that the estimated data follow a specific distribution. The p value greater than 0.05 indicates that the null hypothesis is not rejected at the 5% significance level. The summary of the distribution fitting the posterior data of the Weibull parameters estimated by the ABC (direct Monte Carlo), ABC SMC, and ABC MCMC, as well as their associated p values with respect to different

numbers of iterations, are given in Tables 2, 3, and 4, respectively. This information is useful because it can be used to describe and compare the uncertainty of the estimated parameters of the modified Weibull distribution.

The estimated highest density interval (HDI) of the Weibull distribution parameters with a 95% credible interval indicated on their probability density functions is shown in Fig. 9. When comparing the HDIs of the corresponding parameters from both algorithms, it is clearly seen that the HDIs of the ABC SMC consist of a wider range compared to the ABC MCMC. This difference in the estimated HDIs is associated with the higher scattering of the posterior distributions estimated by the ABC SMC algorithm (Fig. 8).

Therefore, it can be said that a higher level of uncertainty is predicted when the ABC SMC algorithm is employed to fit the Weibull distribution to the experimental observations. This is illustrated in Fig. 10, in which the uncertainty in the Cumulative Distribution Function (CDF) of the fiber strength was compared with the empirical CDF and the modified Weibull distribution fitted by the ML method. Moreover, the upper and lower bounds of the empirical CDF are given for comparison purposes.

Table 2 Statistical distribution fitting to estimated parameters by ABC (direct Monte Carlo) with corresponding chi-squared goodness of fit

No. of iterations	σ_0			β			m		
	Lognormal			Weibull			Lognormal		
	χ^2	μ	σ	χ^2	a	b	χ^2	μ	σ
	p value			p value			p value		
200 K [†]	0.296	5.151	0.042	0.037	0.689	1.971	0.008	1.958	0.187
500 K	0.905	5.450	0.037	0.080	0.614	1.736	0.651	1.861	0.171
1 M	0.916	5.453	0.036	0.051	0.625	1.744	0.103	1.759	0.164
2 M	0.159	5.453	0.037	0.123	0.621	1.745	0.392	1.763	0.169

[†]‘k’ stands for thousand and ‘M’ for million

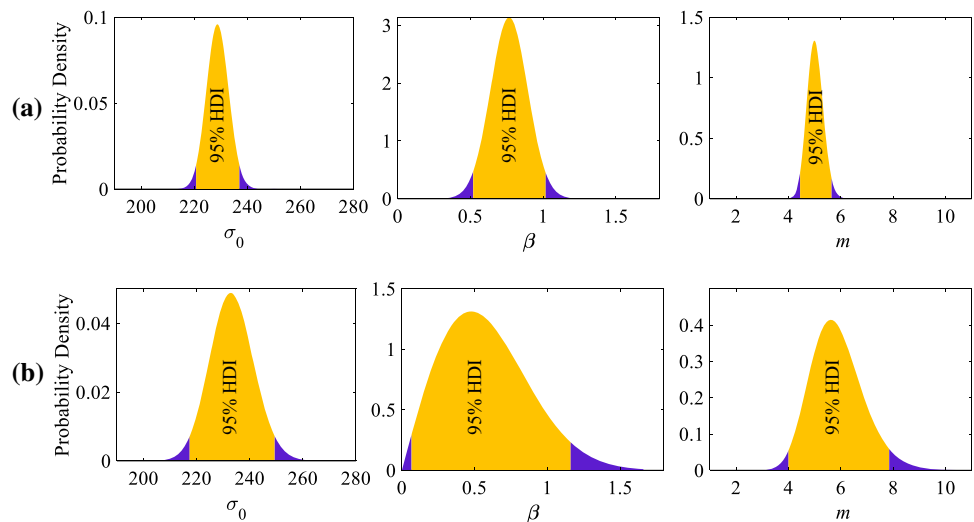
Table 3 Statistical distribution fitting to estimated parameters by ABC SMC with corresponding chi-squared goodness of fit

No. of iterations	σ_0			β			m		
	Lognormal			Weibull			Lognormal		
	χ^2 p value	μ	σ	χ^2 p value	a	b	χ^2 p value	μ	σ
5 K	0.199	5.451	0.037	0.199	0.608	2.206	0.872	1.719	0.151
10 K	0.347	5.453	0.036	0.440	0.656	2.141	0.225	1.736	0.178
20 K	0.643	5.451	0.038	0.054	0.649	1.903	0.142	1.738	0.161
50 K	0.370	5.452	0.038	0.126	0.637	1.860	0.158	1.754	0.172

Table 4 Statistical distribution fitting to estimated parameters by ABC MCMC with corresponding chi-squared goodness of fit

No. of iterations	σ_0			β			m		
	Lognormal			Normal			Lognormal		
	χ^2 p value	μ	σ	χ^2 p value	μ	σ	χ^2 p value	μ	σ
5 K	0.501	5.432	0.019	0.151	0.767	0.131	0.149	1.608	0.063
10 K	0.061	5.433	0.019	0.176	0.770	0.128	0.002	1.613	0.061
20 K	0.252	5.433	0.018	0.105	0.765	0.127	0.183	1.612	0.061
50 K	0.347	5.433	0.018	0.202	0.768	0.125	0.205	1.613	0.060

Figure 9 Highest density interval (HDI) for parameters σ_0, β, m estimated by **a** ABC MCMC (MH) and **b** ABC SMC algorithms.



The advantage of this illustration is that we can evaluate the uncertainty of the corresponding failure probability using a probability distribution for any value of the fiber strength. Also, the difference between the deterministic estimation of fitting parameters and the ABC method can be visualized.

Despite the differences in the Bayesian estimation algorithms, both led to conservative results compared

to the ML. For example, the probability of failure at $\sigma = 170MPa$ estimated by the modified Weibull distribution with deterministic parameters is 0.157, whereas the mean probability value by the ABC MCMC and ABC SMC is 0.161 and 0.131, respectively. As a result, neglecting the uncertainty effects of the parameters can lead to an overestimation of the strength value of the fibers.

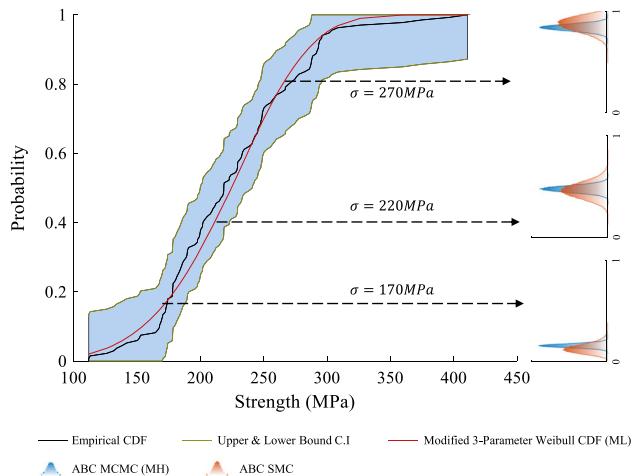


Figure 10 The cumulative distribution functions of the date palm fiber strength where the uncertainty in the failure probability estimated by ML are shown.

Conclusion

In this paper, the application of the ABC framework was brought up for estimating the parameters of the modified 3-parameter Weibull distribution to model the strength of the date palm fibers. Two computationally efficient methods, namely MCMC and SMC, were employed, and their performances in fitting the distribution to the strength data were investigated. The Metropolis–Hastings algorithm was implemented for sampling from the probability distribution in the ABC MCMC method. Conversely, the sampling in the ABC SMC was conducted from the set of previously accepted particles and perturbed according to the uniform perturbation kernel. The marginal posterior distributions of the fitting parameters were modeled by the best-fit probability distributions, and their corresponding HDIs were with a 95% credible interval were computed.

It is found that the posterior distributions of the parameters obtained by the ABC SMC algorithm were more spread compared to that of the ABC MCMC. Therefore, the latter resulted in smaller HDIs for the fitting parameters. The results of this study suggest that the uncertainties in the estimation of the Weibull distribution parameters should be assessed when there is high variability in the experimental data, like natural fiber strength. Consequently, the consideration of these uncertainties is significant to have a reliable and accurate predictive model for the failure probability of date palm fibers.

Availability of data and material

The data that support the findings of this study are available from the corresponding author, M.R., upon reasonable request.

Declarations

Conflict of interest The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- [1] Gurvich MR, Dibenedetto AT, Ranade SV (1997) A new statistical distribution for characterizing the random strength of brittle materials. *J Mater Sci* 32:2559–2564. <https://doi.org/10.1023/A:1018594215963>
- [2] Zhang Y, Wang X, Pan N, Postle R (2002) Weibull analysis of the tensile behavior of fibers with geometrical irregularities. *J Mater Sci* 37:1401–1406. <https://doi.org/10.1023/A:1014580814803>
- [3] Roman RE, Cranford SW (2019) Defect sensitivity and weibull strength analysis of monolayer silicene. *Mech Mater* 133:13–25. <https://doi.org/10.1016/j.mechmat.2019.01.014>
- [4] Acitas S, Aladag CH, Senoglu B (2019) A new approach for estimating the parameters of Weibull distribution via particle swarm optimization: an application to the strengths of glass fibre data. *Reliab Eng Syst Saf* 183:116–127. <https://doi.org/10.1016/j.res.2018.07.024>
- [5] Naik DL, Fronk TH (2016) Weibull distribution analysis of the tensile strength of the kenaf bast fiber. *Fibers Polym* 17:1696–1701. <https://doi.org/10.1007/s12221-016-6176-6>
- [6] Bourahli MEH (2018) Uni- and bimodal Weibull distribution for analyzing the tensile strength of diss fibers. *J Nat Fibers* 15:843–852. <https://doi.org/10.1080/15440478.2017.1371094>
- [7] Wang W, Zhang X, Chouw N et al (2018) Strain rate effect on the dynamic tensile behaviour of flax fibre reinforced polymer. *Compos Struct* 200:135–143. <https://doi.org/10.1016/j.compstruct.2018.05.109>
- [8] Monteiro SN, Margem FM, de Oliveira Braga F et al (2017) Weibull analysis of the tensile strength dependence with fiber diameter of giant bamboo. *J Mater Res Technol* 6:317–322. <https://doi.org/10.1016/j.jmrt.2017.07.001>
- [9] Fuentes CA, Willekens P, Petit J et al (2017) Effect of the middle lamella biochemical composition on the non-linear behaviour of technical fibres of hemp under tensile loading using strain mapping. *Compos A Appl Sci Manuf*

- 101:529–542. <https://doi.org/10.1016/j.compositesa.2017.07.017>
- [10] Guo M, Zhang TH, Chen BW, Cheng L (2014) Tensile strength analysis of palm leaf sheath fiber with Weibull distribution. *Compos A Appl Sci Manuf* 62:45–51. <https://doi.org/10.1016/j.compositesa.2014.03.018>
- [11] Belaadi A, Bezazi A, Bouchak M et al (2014) Thermochemical and statistical mechanical properties of natural sisal fibres. *Compos B Eng* 67:481–489. <https://doi.org/10.1016/j.compositesb.2014.07.029>
- [12] da Costa LL, Loiola RL, Monteiro SN (2010) Diameter dependence of tensile strength by Weibull analysis: part I bamboo fiber. *Matéria (Rio J)* 15:110–116. <https://doi.org/10.1590/S1517-70762010000200004>
- [13] Zafeiropoulos NE, Baillie CA (2007) A study of the effect of surface treatments on the tensile strength of flax fibres: part II. Application of Weibull statistics. *Compos A Appl Sci Manuf* 38:629–638. <https://doi.org/10.1016/j.compositesa.2006.02.005>
- [14] Langhorst A, Ravandi M, Mielewski D, Banu M (2021) Technical agave fiber tensile performance: the effects of fiber heat-treatment. *Ind Crops Prod* 171:113832. <https://doi.org/10.1016/j.indcrop.2021.113832>
- [15] Trujillo E, Moesen M, Osorio L et al (2014) Bamboo fibres for reinforcement in composite materials: strength Weibull analysis. *Compos A Appl Sci Manuf* 61:115–125. <https://doi.org/10.1016/j.compositesa.2014.02.003>
- [16] Korabel'nikov YUG, Tamuzh VP, Siluyanov OF et al (1984) Scale effect of the strength of fibers and properties of unidirectional composites based on them. *Mech Compos Mater* 20:129–134. <https://doi.org/10.1007/BF00610351>
- [17] Watson AS, Smith RL (1985) An examination of statistical theories for fibrous materials in the light of experimental data. *J Mater Sci* 20:3260–3270. <https://doi.org/10.1007/BF00545193>
- [18] Huang D, Zhao X (2019) Novel modified distribution functions of fiber length in fiber reinforced thermoplastics. *Compos Sci Technol* 182:107749. <https://doi.org/10.1016/j.compscitech.2019.107749>
- [19] Canavos GC, Taokas CP (1973) Bayesian estimation of life parameters in the weibull distribution. *Oper Res* 21:755–763. <https://doi.org/10.1287/opre.21.3.755>
- [20] Green EJ, Roesch FA, Smith AFM, Strawderman WE (1994) Bayesian estimation for the three-parameter weibull distribution with tree diameter data. *Biometrics* 50:254–269. <https://doi.org/10.2307/2533217>
- [21] Guure CB, Ibrahim NA, Ahmed AOM (2012) Bayesian estimation of two-parameter weibull distribution using extension of jeffreys' prior information with three loss functions. *Math Probl Eng* 2012:e589640. <https://doi.org/10.1155/2012/589640>
- [22] Almongy HM, Almetwally EM, Alharbi R et al (2021) The weibull generalized exponential distribution with censored sample: estimation and application on real data. *Complexity* 2021:e6653534. <https://doi.org/10.1155/2021/6653534>
- [23] Chacko M, Mohan R (2019) Bayesian analysis of Weibull distribution based on progressive type-II censored competing risks data with binomial removals. *Comput Stat* 34:233–252. <https://doi.org/10.1007/s00180-018-0847-2>
- [24] Ducros F, Pamphile P (2018) Bayesian estimation of Weibull mixture in heavily censored data setting. *Reliab Eng Syst Saf* 180:453–462. <https://doi.org/10.1016/j.res.2018.08.008>
- [25] Minter A, Retkute R (2019) Approximate bayesian computation for infectious disease modelling. *Epidemics* 29:100368. <https://doi.org/10.1016/j.epidem.2019.100368>
- [26] Krit M, Gaudoin O, Xie M, Remy E (2016) Simplified likelihood based goodness-of-fit tests for the weibull distribution. *Commun Stat-Simul Comput* 45:920–951. <https://doi.org/10.1080/03610918.2013.879889>
- [27] Medjoudj R, Mazighi I (2020) Estimation of photovoltaic energy conversion using mixed weibull distribution. *JESA* 53:385–391. <https://doi.org/10.18280/jesa.530309>
- [28] Edwards W, Lindman H, Savage LJ (1963) Bayesian statistical inference for psychological research. *Psychol Rev* 70:193–242. <https://doi.org/10.1037/h0044139>
- [29] Liepe J, Kirk P, Filippi S et al (2014) A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nat Protoc* 9:439–456. <https://doi.org/10.1038/nprot.2014.025>
- [30] Liu C-C (1997) A comparison between the weibull and lognormal models used to analyse reliability data. University of Nottingham, Ph.D.
- [31] Datsiou KC, Overend M (2018) Weibull parameter estimation and goodness-of-fit for glass strength data. *Struct Saf* 73:29–41. <https://doi.org/10.1016/j.strusafe.2018.02.002>
- [32] Vandebroek M, Belis J, Louter C, Van Tendeloo G (2012) Experimental validation of edge strength model for glass with polished and cut edge finishing. *Eng Fract Mech* 96:480–489. <https://doi.org/10.1016/j.engfracmech.2012.08.019>
- [33] Gupta PK, Singh AK (2017) Classical and bayesian estimation of Weibull distribution in presence of outliers. *Cogent Math*. <https://doi.org/10.1080/23311835.2017.1300975>
- [34] Filippi S, Barnes CP, Cornebise J, Stumpf MPH (2013) On optimality of kernels for approximate Bayesian computation using sequential Monte Carlo. *Stat Appl Genet Mol Biol* 12:87–107. <https://doi.org/10.1515/sagmb-2012-0069>

- [35] Beaumont MA, Cornuet J-M, Marin J-M, Robert CP (2009) Adaptive approximate Bayesian computation. *Biometrika* 96:983–990. <https://doi.org/10.1093/biomet/asp052>
- [36] Drovandi CC, Pettitt AN (2011) Estimation of parameters for macroparasite population evolution using approximate bayesian computation. *Biometrics* 67:225–233. <https://doi.org/10.1111/j.1541-0420.2010.01410.x>
- [37] Maruyama G, Tanaka H (1959) Ergodic property of n-dimensional recurrent markov processes. *Mem Faculty Sci Kyushu Univ Series A Math* 13:157–172. <https://doi.org/10.2206/kyushumfs.13.157>
- [38] Thach TT, Bris R (2020) Improved new modified Weibull distribution: a Bayes study using Hamiltonian Monte Carlo simulation. *Proc Inst Mech Eng Part O J Risk Reliability* 234:496–511. <https://doi.org/10.1177/1748006X19896740>
- [39] Fox J-P, Glas CAW (2001) Bayesian estimation of a multi-level IRT model using gibbs sampling. *Psychometrika* 66:271–288. <https://doi.org/10.1007/BF02294839>
- [40] Liu F, Li X, Zhu G (2020) Using the contact network model and Metropolis-Hastings sampling to reconstruct the COVID-19 spread on the “Diamond Princess.” *Sci Bull (Beijing)* 65:1297–1305. <https://doi.org/10.1016/j.scib.2020.04.043>
- [41] Shao W, Guo G, Meng F, Jia S (2013) An efficient proposal distribution for Metropolis-Hastings using a B-splines technique. *Comput Stat Data Anal* 57:465–478. <https://doi.org/10.1016/j.csda.2012.07.014>
- [42] Christensen OF, Roberts GO, Rosenthal JS (2005) Scaling limits for the transient phase of local metropolis-hastings algorithms. *J Royal Stat Soc Series B (Stat Methodology)* 67:253–268
- [43] Abbasi-Yadkori Y (2016) Fast mixing random walks and regularity of incompressible vector fields. [arXiv:1611.09252](https://arxiv.org/abs/1611.09252)
- [44] Sisson SA, Fan Y, Tanaka MM (2007) Sequential monte carlo without likelihoods. *PNAS* 104:1760–1765. <https://doi.org/10.1073/pnas.0607208104>
- [45] Del Moral P, Doucet A, Jasra A (2012) An adaptive sequential monte carlo method for approximate bayesian computation. *Stat Comput* 22:1009–1020. <https://doi.org/10.1007/s11222-011-9271-y>
- [46] Drovandi CC, Pettitt AN (2011) Likelihood-free bayesian estimation of multivariate quantile distributions. *Comput Stat Data Anal* 55:2541–2556. <https://doi.org/10.1016/j.csda.2011.03.019>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.