JOURNAL OF APPLIED STATISTICS
2022, VOL. 49, NO. 11, 2825–2844
https://doi.org/10.1080/02664763.2021.1924638

Taylor & Francis
Taylor & Francis Group

Improved maximum likelihood estimation of the shape-scale family based on the generalized progressive hybrid censoring scheme

M. Maswadah

Department of Mathematics, Faculty of Science, Aswan University, Aswan, Egypt

ABSTRACT

In parametric estimates, the maximum likelihood estimation method is the most popular method widely used in the social sciences and psychology, although it is biased in situations where sample sizes are small or the data are heavily censored. Therefore, the main objective of this research is to improve this estimation method using the Runge–Kutta technique. The improved method was applied to derive the estimators of the shape scale family parameters and compare them with Bayesian estimators based on the informative and kernel priors, via Monte Carlo simulation. The simulation results showed that the improved maximum likelihood estimation method is highly efficient and outperforms the Bayesian method for different sample sizes. Finally, from a future perspective, the proposed model could be important for analyzing real data sets including data on COVID-19 deaths in Egypt, for potential comparative studies with other countries.

ARTICLE HISTORY
Received 4 November 2018
Accepted 25 April 2021

KEYWORDS
Bayesian estimations; COVID-19 epidemic; informative prior; Kernel prior; ordinary differential equation; Runge–Kutta method

1. Introduction

In statistical inference, it is known that the maximum likelihood estimation method is biased when sample sizes are small or when the data are heavily censored and ineffective as the Bayesian method. These biases can mislead subsequent inferences and in some distributions, contain nonlinear equations that require numerical techniques. Therefore, the challenge in this paper is to improve the maximum likelihood estimation method, using the Runge–Kutta technique. The simulations and real data set results indicated that the improved method was more efficient than Bayesian method, even using informative and kernel priors. Thus, the statistical significance of this method is its efficiency compared to most estimation methods, and it is reliable and easy to apply, especially for researchers in social sciences and psychology. To illustrate this, we applied the proposed method to a general lifetime distribution which contains some of the lifetime distributions most commonly used in reliability and survival analyzes such as Weibull and Weibull extension models.

CONTACT M. Maswadah maswadah@hotmail.com Department of Mathematics, Faculty of Science, Aswan University, Aswan, Egypt

Supplemental data for this article can be accessed here. https://doi.org/10.1080/02664763.2021.1924638

These distributions have flexibility in describing the lifetime variables of constant hazard rate as well as non-constant hazard rate and are useful for modeling and analyzing lifetime data in medical, biological, and engineering sciences. Recently, we applied these distributions to analyze actual real data sets of current significance including COVID-19 deaths observed in Egypt from December 22, 2020, to February 16, 2021.

Several authors have discussed the estimation of the Weibull model parameters, including [31,37] who derived confidence intervals using some pivotal quantities based on progressively censored samples [2]. Derived the estimates of the parameters of the Weibull model based on the classical and Bayesian approaches [20] presented reliability and quantile analyzes of the Weibull distribution [5] derived the maximum likelihood estimates for the Weibull model parameters based on complete and censored data [22] presented some methods for estimating the parameters of the Weibull model [31,38] derived the MLEs for the Weibull model parameters based on progressive type-II censored samples, and [26] derived the empirical Bayes estimates of the Weibull model parameters. For further discussion of the Weibull distribution, see Zhang *et al.* [39,40].

The Weibull extension model proposed as a new lifetime distribution with a bathtub-shaped hazard rate function with the ability to describe bathtub-shaped lifetime data, see [9,13,32] the Weibull extension model presented as a practical model for numerous applications in reliability and lifetime modeling of mechanical and electro-mechanical products with a bathtub-shaped failure rate function [34] the statistical analysis of this model as a bathtub-shaped hazard rate function were discussed [35,36] derived the exact confidence intervals for the shape parameter based on censored samples [15] the Quasi-likelihood estimates of the Weibull extension parameters were derived [3] the estimates of the Weibull extension model parameters were derived based on the generalized order statistics [33] the Bayesian modeling of bathtub-shaped hazard rate function using various forms of this model and some issues related to model choices were discussed [27] the confidence intervals of the parameters of the Weibull extension model were derived using the conditional inference based on the generalized order statistics. The aim of this paper is to continue these efforts by deriving point estimates for the general lifetime model parameters that contain the Weibull and the Weibull extension models as members of this family, based on a generalized progressive hybrid censoring scheme (GPHCS).

A random variable X is said to have the general lifetime model, which belongs to the shape-scale family if its probability density function (PDF) is given by

$$f(x) = \alpha\beta g^{\alpha-1}(x)g'(x)\exp(-\beta g^{\alpha}(x)), \quad x > 0, \alpha, \beta > 0, \tag{1}$$

and its cumulative distribution function (CDF) is given by

$$F(x) = 1 - \exp(-\beta g^{\alpha}(x)), \quad x > 0, \alpha, \beta > 0, \tag{2}$$

α and β are the shape and scale parameters respectively. For convenience, we assume $g(x)$ to be differentiable as well as strictly increasing function of x such that, $g(0^+) = 0$ and $g(x) \to \infty$ as $x \to \infty$.

This family includes the most common lifetime distributions such as the Weibull extension, modified Weibull, Weibull, Pareto, Burr-type-XII, Lomax, and generalized Pareto distributions according to the values of $g^{\alpha}(x)$. Some of the important members of this family are given in Table 1.

Table 1. Some important distributions as special cases from the shape-scale family.

No.	$g^{\alpha}(x)$	$F(x)$	Distributions
1	$\exp(x^{\alpha})-1$	$1-\exp(-\beta(\exp(x^{\alpha})-1))$	Weibull Extension
2	$x^{\alpha}\exp(\lambda x)$	$1-\exp(-x^{\alpha}\beta\exp(\lambda x))$	Modified Weibull
3	x^{α}	$1-\exp(-\beta x^{\alpha})$	Weibull
4	$\ln(1+x^{\alpha})$	$1-(1+x^{\alpha})^{-\beta}$	Burr-type XII
5	$\ln(1+x/\alpha)$	$1-(1+x/\alpha)^{-\beta}$	Lomax
6	$-\ln(1-x/\alpha)$	$1-(1-x/\alpha)^{\beta}$	Generalized Pareto
7	$\ln(x/\alpha)$	$1-(x/\alpha)^{-\beta}$	Pareto-type I

In reliability analysis, experiments often terminate before all test units fail due to cost and time considerations or may be lost or removed from testing prior to failure. Hence, progressive censoring sampling schemes appear in these life test experiments. The general scheme for studying such experiments is the progressive censoring scheme, which is the most popular and useful scheme for both industrial life testing and clinical trial applications [4,6] presented comprehensive studies on the topic of the progressive censored scheme and its applications. The progressive censoring scheme allows for some of the experimental units remaining at various stages to be removed before the end of the test, although the trial time may be quite long due to the presence of some very reliable units. Thus, recently [21] proposed a progressive hybrid censoring scheme, which is a mixture of progressive type-II and hybrid censoring scheme, see [12,14,16,19]. However, the disadvantage of the progressive hybrid censored scheme is those very few failures may occur before the time point T. In order to provide a guarantee of the number of failures observed as well as the time required to complete the test [11,12] suggested a generalized progressive hybrid censoring scheme. This scheme modifies the progressive hybrid censoring scheme by allowing the experiment to continue beyond the time point T if the number of failures is less than m and this allows the experimenter to at least monitor k failures. This scheme can be described as follows:

Consider a life-testing experiment in which n identical units $X_1, X_2, \ldots, X_n$ placed on the test. For $T \in (0,\infty)$ and the integers k and m are pre-fixed such that $k < m$ with the random removal units $R_1, R_2, \ldots, R_m$, which are fixed at the beginning of the experiment, where $n = m + \sum_{i=1}^{m} R_i$. Generally, at the time of the i_th failure, R_i units are randomly removed from the remaining surviving units $S_i = n - i - \sum_{j=1}^{i-1} R_j$, $i \in [1, m]$. Continue this process until the terminated time $T^* = \max\{X_{k:m}, \min\{X_{m:n}, T\}\}$, which is the time for removing all the remaining surviving units from the experiment according to the following cases. Let J denote the number of observed failures up to the time T. Thus, we have one of the following types of observations:

Case I: $X_{1:m:n} \le \ldots \le X_{J:m:n} < X_{J+1:m:n} < \ldots < X_{k:m:n}$ If $T < X_{k:m:n} < X_{m:m:n}$.

Case II: $X_{1:m:n} \le \ldots \le X_{k:m:n} < X_{k+1:m:n} < \ldots < X_{J:m:n}$ If $X_{K:m:n} < T < X_{m:m:n}$.

Case III: $X_{1:m:n} \le \ldots \le X_{k:m:n} < X_{k+1:m:n} < \ldots < X_{m:m:n}$ If $X_{k:m:n} < X_{m:m:n} < T$.

Note that for Case I, $T < X_{k:m:n} < X_{m:m:n}$. and $X_{k+1:m:n}, \ldots, X_{m:m:n}$ are not observed. For Case II, $X_{J:m:n} < T < X_{J+1:m:n}$ and $X_{J+1:m:n}, \ldots, X_{m:m:n}$ are not observed.

Thus, given a generalized progressive hybrid censored sample, the likelihood function for the three different cases can be written in a unified form as follows:

$$L(\underline{x};\theta) = C\prod_{i=1}^{D} f(x_{i,m,n})[1-F(x_{i,m,n})]^{R_i}[1-F(T)]^{R_T^*\delta}. \tag{3}$$

$$D = \begin{cases} k, \delta = 0 \; if \; T \le X_k < X_m \\ J, \delta = 1 \; if \; X_k < T \le X_m \\ m, \delta = 0 \; if \; X_k < X_m < T \end{cases}, \quad \underline{X} = \begin{cases} (X_1, X_2, \ldots, X_J, X_{J+1} \ldots, X_k) \; if \\ \quad T < X_k < X_m \\ (X_1, X_2, \ldots, X_J) \; if \; X_k \le T < X_m \\ (X_1, X_2, \ldots, X_m) \; if \; X_k < X_m \le T \end{cases}$$

and

$$R = \begin{cases} (R_1, R_2, \ldots, R_J, 0, 0, \ldots 0, R_k), R_k = D - k - \sum_{i=1}^{k-1} R_i \; if \; T < X_k < X_m \\ (R_1, R_2, \ldots, R_J, R_T^*), R_T^* = D - J - \sum_{i=1}^{J} R_i \; if \; X_k \le T < X_m \\ (R_1, R_2, \ldots, R_m), R_m = D - m - \sum_{i=1}^{m-1} R_i \; if \; X_k < X_m \le T \end{cases},$$

where $R_T{}^*$ is the number of surviving units that are removed at the stopping time T.

Recently, GPHCS was applied to some distributions such as Weibull distribution, see [12], inverse Weibull distribution, see [28,29], exponential distribution, see [8,11], Rayleigh distribution, see [10].

2. An improved MLE method

It is known that the likelihood function $L(\theta;x)$ contains all the information in the sample and depends on the unknown parameters $\theta = (\alpha, \beta)$ and the data $X = (x_1, x_2, \ldots, x_n)$. Thus, the MLE $\hat{\theta} = \hat{\theta}(x)$ of θ is the solution of the stationary equation $(\partial H(\theta;X)/\partial\theta)|_{\theta=\hat{\theta}} = 0$, which is a function of $\hat{\theta}(x)$ and X, where $H(\theta;X)$ is the log-likelihood function.

Applying the implicit function theorem to the stationary equation with considering all partial derivatives, as well as the total derivatives, are assumed to be evaluated at some known value of $\hat{\theta}(x) = \theta_0$, say. We might say that for any x in $\underline{X}$ there is a value $\hat{\theta}(x)$ satisfying the stationary equation. Taking the x-derivative for the stationary equation, see Ramsay *et al.* [30], we obtain:

$$\frac{d}{dx}\left(\frac{\partial H(\theta;X)}{\partial\theta}\right)|_{\theta=\hat{\theta}} = \frac{\partial^2 H(\theta;X)}{\partial\theta\partial x}|_{\theta=\hat{\theta}} + \frac{\partial^2 H(\theta;X)}{\partial\theta^2}|_{\theta=\hat{\theta}}\frac{d\hat{\theta}(x)}{dx} = 0. \tag{4}$$

Solving (4), we obtain the first derivative for $\hat{\theta}$ with respect to x at $\theta = \hat{\theta}$ as:

$$\frac{d\hat{\theta}(x)}{dx} = -\left(\frac{\partial^2 H(\theta;X)}{\partial\theta^2}|_{\theta=\hat{\theta}}\right)^{-1}\frac{\partial^2 H(\theta;X)}{\partial\theta\partial x}|_{\theta=\hat{\theta}}. \tag{5}$$

Thus, we can write (5) as a first-order ordinary differential equation for the maximum likelihood estimator $\hat{\theta}(x)$ as:

$$d\hat{\theta}(x)/dx = f(x,\hat{\theta}), \text{with the initial condition } \hat{\theta}(x_0) = \theta_0, \tag{6}$$

It is clear that $f(x,\hat{\theta})$ and $df(x,\hat{\theta})/d\tilde{\theta}$ are defined and continuous functions in a rectangular region containing the point (x_0,θ_0), which ensures the existence of a unique solution for (6) in the neighborhood of the point (x_0,θ_0). Using any numerical technique such as the fourth-order Runge–Kutta method, we can find the approximate solution for $\hat{\theta}(x)$. If the initial conditions are unavailable, then they should be appended to the parameter $\hat{\theta}$ as quantities with respect to which the fit is optimized. This method was applied to the inverse Weibull model parameters based on the generalized progressive hybrid censoring scheme, see [28].

For the general lifetime model (1) and the likelihood function (3), the log-likelihood function and its derivatives can be derived as follows:

$$H(\alpha,\beta|X) = D\ln(\alpha\beta) + (\alpha-1)\sum_{i=1}^{D}\ln(g(x_i)) - \beta\left[\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i) + \delta R_T^* g^{\alpha}(T)\right].$$

$$\frac{\partial H}{\partial \alpha} = D/\alpha + \sum_{i=1}^{D}\ln(g(x_i)) - \beta\left[\sum_{i=1}^{D}(R_I+1)g^{\alpha}(x_i)\ln g(x_i) + \delta R_T^* g^{\alpha}(T)\ln g(T)\right],$$

$$\frac{\partial^2 H}{\partial \alpha^2} = -D/\alpha^2 - \beta\left[\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i)(\ln(g(x_i))^2 + \delta R_T^* g^{\alpha}(T)(\ln(g(T))^2\right],$$

$$\frac{\partial^2 H}{\partial x\partial \alpha} = \sum_{i=1}^{D}\frac{g'(x_i)}{g(x_i)} - \beta\left[\alpha\sum_{i=1}^{D}(R_i+1)g^{\alpha-1}(x_i)g'(x_i)\ln(g(x_i)) + g^{\alpha-1}(x_i)g'(x_i)\right],$$

$$\frac{\partial H}{\partial \beta} = D/\beta - \left[\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i) + \delta R_T^* g^{\alpha}(T)\right], \frac{\partial^2 H}{\partial \beta^2} = -D/\beta^2,$$

$$\frac{\partial^2 H}{\partial x\partial \beta} = -\alpha\sum_{i=1}^{D}(R_i+1)g^{\alpha-1}(x_i)g'(x_i).$$

Thus, using (6) with the corresponding derivatives above, we can find the point estimates for each α and β, using the Runge–Kutta method. Programs for the simulation with Fortran codes are provided in the 'Supplemental Material'.

3. Bayesian estimation based on the informative prior

We suggest using independent priors for each of the parameters α and β such as gamma distributions. Hence, the joint prior density is given by

$$g(\alpha,\beta) \propto \alpha^{a-1}\beta^{c-1}e^{-b\alpha-d\beta}, a,b \geq 0, c,d \geq 0. \tag{7}$$

Using the informative prior (7) and the likelihood function (3) based on (1) and (2), the joint posterior density can be derived as follows:

$$f(\alpha,\beta|X) = K\alpha^{D+a-1}\beta^{D+c-1}\exp\left[-\alpha\left(b-\sum_{i=1}^{D}\ln(g(x_i)\right)-\sum_{i=1}^{D}\ln(g(x_i)\right]$$
$$\times\exp\left[-\beta(d+\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i)+\delta R_T^{*}g^{\alpha}(T))\right]$$

The marginal posterior densities of the parameters α and β can be derived as

$$g_1(\alpha|X) = K\Gamma(D+c)\alpha^{D+a-1}\left[d+\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i)+\delta R_T^{*}g^{\alpha}(T)\right]^{-(n+c)}$$
$$\times\exp\left[-\alpha\left(b-\sum_{i=1}^{D}\ln(g(x_i))-\sum_{i=1}^{D}\ln(g(x_i))\right)\right]$$
$$g_2(\beta|X) = K\int_0^{\infty}\alpha^{D+a-1}\beta^{D+c-1}\exp\left[-\alpha\left(b-\sum_{i=1}^{D}\ln(g(x_i))-\sum_{i=1}^{D}\ln(g(x_i)\right)\right]$$
$$\times\exp\left[-\beta\left(d+\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i)+\delta R_T^{*}g^{\alpha}(T)\right)\right]d\alpha$$

K is the normalization constant that can be derived as:

$$K^{-1} = \Gamma(D+c)\int_0^{\infty}\alpha^{D+a-1}\left[d+\sum_{i=1}^{D}(R_i+1)g^{\alpha}(x_i)+\delta R_T^{*}g^{\alpha}(T)\right]^{-(D+c)}$$
$$\times\exp\left[-\alpha\left(b-\sum_{i=1}^{D}\ln(g(x_i))-\sum_{i=1}^{D}\ln(g(x_i))\right)\right]d\alpha.$$

4. Bayesian estimation based on the Kernel prior

For deriving the kernel prior, we present the bivariate kernel density estimator for the unknown probability density function $g(\alpha,\beta)$ with support on $(0,\infty)$, which is defined as

$$\hat{g}(\alpha,\beta) = \frac{1}{Dh_1h_2}\sum_{i=1}^{D}K\left(\frac{\alpha-\hat{\alpha}_i}{h_1},\frac{\beta-\hat{\beta}_i}{h_2}\right), \tag{8}$$

$h_i, i=1,2$ are called the bandwidths or smoothing parameters, which chosen such that $h_i\to 0$ and $Dh_i\to\infty$ as $D\to\infty$, where D is the sample size. The influence of the smoothing parameter h is critical because it determines the amount of smoothing. A small value of h leads to the estimator having insignificant details while a large value of h causes exceeders of the information contained in the sample, which in consequence, may mask some of the important characteristics. Hence, a certain compromise is needed. However, the

optimal choice for h, which minimizes the mean squared errors is $h_i = \gamma \widehat{\sigma}_i D^{-0.2}$ where $0.3 \leq \gamma \leq 1.06$ and the estimated value of the population standard deviations $\widehat{\sigma}_i$ could be used as S_i the sample standard deviation. The optimal choice for the kernel function $K(.,.)$ can be used as the bivariate standard normal distribution. The basic elements associated with the kernel density estimation function have been extensively studied in [17,18]. Based on the properties of the MLEs of the parameters, which converge in probability to the original parameters, the kernel prior estimate can be derived using the following algorithm:

(i) Generate a random sample $X = (x_1, x_2, x_3, \ldots, x_n)$ from the parent distribution $f(x; \alpha, \beta)$ with given specified values for the unknown parameters α and β.
(ii) Bootstrapping with replacement n samples $x_1^*, x_2^*, x_3^*, \ldots, x_n^*$, with size n each, where $x_i^* = (x_{i1}^*, x_{i2}^*, \ldots, x_{in}^*)$ for $i \in [1, n]$ from the given random sample in step (i).
(iii) For each sample in step (ii) calculate the MLEs for the parameters α and β, we get the random variables $Y = (\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_n)$ and $Z = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_n)$ for their MLEs.
(iv) Finally, based on the random variables Y and Z, the kernel density estimation (8) can be used to derive the kernel prior density estimator $\hat{g}(\alpha, \beta)$.

Using the kernel prior (8) and (3) the joint posterior density for α and β is given by

$$f(\alpha, \beta | X) = K\hat{g}(\alpha, \beta)(\alpha\beta)^D \exp\left[(\alpha - 1)\sum_{i=1}^{D} \ln(g(x_i))\right] \times \exp\left[-\beta\left(\sum_{i=1}^{D}(R_i + 1)g^{\alpha}(x_i) + \delta R_T^* g^{\alpha}(T)\right)\right].$$

It is worthwhile to mention that, this kernel prior has been used for some distributions, see [1,23–26].

5. Simulation study

We study the performance of the IMLE and Bayes methods, based on the root mean squared errors (RMSEs). The RMSEs were computed by generating 1000 replications for sample sizes, n= 20,40, and 60 to represent small, moderate, and large sizes from Weibull and Weibull extension models with $\alpha = (0.59, 1.11)$ and $\beta = (1.11, 1.97)$. In the simulation study, Bayesian estimates based on the squared error loss function were driving using the informative gamma prior and the kernel prior distributions using different combinations for the hyper-parameter of the gamma prior for α and β as $a = (2, 4)$, $b = (8, 7)$, $c = (4, 7)$, and $d = (7, 6)$.

Generation of the generalized progressive hybrid censored order statistics can be performed according to the following procedure:

Let $X = \{X_1, X_2, \ldots, X_n\}$ be a random sample with size n from the parent distribution. Thus, based on the random sample the generalized progressive hybrid censored sample with size $m(< n)$: $X_{1:m:n}, X_{2:m:n}, \ldots, X_{m:m:n}$, can be generated as follows:

(i) Let $T \in (0, \infty)$ and the integers k and m are pre-fixed such that $k < m$ with $\Re = (R_1, R_2, \ldots, R_m)$ be the predetermined number of uniform random removal observations, which can be generated as $R_i = An\text{int}[2 * (mm - \sum_{j=1}^{i-1} R_j + 1) * U/mm]$, $i = 1(1)(m-1)$, where $U \approx Uniform(0,1)$ and $R_m = mm - \sum_{i=1}^{m-1} R_i$, $mm = n - m$.

(ii) From $i = 1$, choose the minimum observation of the random sample say $X_{i,m:n}$, which is the i_th observation that can be selected for the generalized progressive hybrid censored random sample.

(iii) If $i <= m$, remove R_i uniform random observations that is $R_i = \{r_{1,i}, r_{2,i}, \ldots r_{c,i}\}$, $i = 1(1)m$ and c is the number of censored observations. The removed units without replacement from the subset $X_i^* = X \backslash \{A_i \bigcup B_i\}$ are $r_{j,i} = An\text{ int}\left[(n - \sum_{l=1}^{i-1} R_l + 1) * U\right]$, $j = 1(1)c$ is the subscript of the removed units where $A_i = \bigcup_{j=1}^{i} X_{j:m:n}$ and $B_i = \bigcup_{j=1}^{i-1} R_j$ with noting that $r_{0,i} = 0$ it means $R_i = 0$.

(iv) If $T < X_{k:m:n} < X_{m:m:n}$, let $R_K = D - k - \sum_{j=1}^{k-1} R_j$, $m = k$ and stop.

(v) If $X_{k:m:n} < T < X_{m:m:n}$, let $R_T^* = D - J - \sum_{j=1}^{J} R_j$, $m = J$ and stop.

(vi) If $X_{k:m:n} < X_{m:m:n} < T$, let $R_m = D - m - \sum_{j=1}^{m-1} R_j$, and stop.

(vii) If $i < m$, set $i = i + 1$ and go to step 2 or else stop.

From the simulation results in Tables 2–5, some points are quite clear based on these estimates, and others are summarized in the following key points:

(i) Obviously, in general for the parameters α and β the IMLE method has smaller RMSEs compared to those based on the Bayesian method.

(ii) The Bayesian estimates for both models based on the kernel prior are more efficient than the ones based on the gamma prior and they are relatively close to IMLE estimates.

(iii) The estimated RMSEs values increase as the value of α increases and decrease as the value of β increases.

(iv) The estimated RMSEs values decrease as the scale parameter increases and the shape parameter decreases for the prior's hyper-parameter for α and vice versa for β.

(v) The estimated RMSE values decrease with increasing the values of k, m, and n as expected.

(vi) In general, the estimated RMSE values for small values of n increase with increasing the termination time T and vice versa for large values of the sample size n as expected for both models.

6. Real data analysis

In this section, we studied three real datasets to study the performance of the proposed methods on Weibull and Weibull extension models, which are the most desirable and widely used lifetime distributions. These distributions have been used in many applications in various fields and in new areas such as biomedical sciences and survival analysis to describe the lifetime of specific mortality and failure rates. Hence, we have fitted these datasets using some goodness of fit tests such as the Kolmogorov–Smirnov (K-S),

Table 2. The root mean square errors (RMSEs) for the Weibull parameter α using the IMLE and Bayes methods at $T = 0.75$ and $T = 1.5$ with $m = (n/2$ and $3n/4)$ and $k = (m/2$ and $3\,m/4)$.

					IMLE		Gamma prior		Kernel prior	
n	m	k	α	β	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$
20	10	5	0.59	1.11	0.0642	0.0553	0.1773	0.1677	0.1322	0.1372
				1.97	0.0676	0.0707	0.1613	0.1643	0.1068	0.1104
			1.11	1.11	0.0799	0.0715	0.3452	0.3237	0.1481	0.1515
				1.97	0.0844	0.0759	0.3149	0.3105	0.1410	0.1386
		8	0.59	1.11	0.0501	0.0472	0.1510	0.1439	0.1312	0.1338
				1.97	0.0548	0.0552	0.1529	0.1521	0.1015	0.1058
			1.11	1.11	0.0663	0.0667	0.2884	0.2960	0.1479	0.1444
				1.97	0.0630	0.0661	0.2863	0.2913	0.1370	0.1355
	15	8	0.59	1.11	0.0518	0.0557	0.1518	0.1532	0.1339	0.1317
				1.97	0.0566	0.0555	0.1535	0.1502	0.1051	0.1023
			1.11	1.11	0.0657	0.0623	0.2872	0.2877	0.1438	0.1396
				1.97	0.0656	0.0622	0.2878	0.2846	0.1389	0.1374
		11	0.59	1.11	0.0430	0.0422	0.1268	0.1283	0.1287	0.1227
				1.97	0.0478	0.0459	0.1311	0.1279	0.1016	0.1043
			1.11	1.11	0.0538	0.0554	0.2361	0.2395	0.1415	0.1432
				1.97	0.0549	0.0544	0.2512	0.2454	0.1334	0.1364
40	20	10	0.59	1.11	0.0452	0.0420	0.1298	0.1135	0.0972	0.0966
				1.97	0.0544	0.0487	0.1283	0.1188	0.0766	0.0776
			1.11	1.11	0.0572	0.0470	0.2721	0.2213	0.1321	0.1274
				1.97	0.0690	0.0521	0.2610	0.2287	0.1201	0.1134
		15	0.59	1.11	0.0418	0.0421	0.1152	0.1142	0.0979	0.0962
				1.97	0.0488	0.0482	0.1190	0.1187	0.0785	0.0763
			1.11	1.11	0.0494	0.0491	0.2298	0.2263	0.1248	0.1268
				1.97	0.0511	0.0538	0.2262	0.2306	0.1140	0.1127
	30	15	0.59	1.11	0.0414	0.0413	0.1134	0.1098	0.0981	0.0993
				1.97	0.0482	0.0468	0.1197	0.1134	0.0749	0.0773
			1.11	1.11	0.0492	0.0464	0.2264	0.2174	0.1235	0.1236
				1.97	0.0544	0.0523	0.2349	0.2275	0.1095	0.1103
		23	0.59	1.11	0.0275	0.0269	0.0901	0.0927	0.0930	0.0971
				1.97	0.0390	0.0401	0.0901	0.0931	0.0766	0.0753
			1.11	1.11	0.0375	0.0401	0.1634	0.1676	0.1171	0.1223
				1.97	0.0417	0.0386	0.1775	0.1730	0.1116	0.1087
60	30	15	0.59	1.11	0.0412	0.0382	0.1044	0.0948	0.0841	0.0834
				1.97	0.0484	0.0456	0.1076	0.1009	0.0632	0.0664
			1.11	1.11	0.0513	0.0386	0.2408	0.1765	0.1123	0.1140
				1.97	0.0621	0.0474	0.2309	0.1889	0.1011	0.0993
		23	0.59	1.11	0.0391	0.0380	0.0965	0.0932	0.0854	0.0820
				1.97	0.0448	0.0436	0.0991	0.0954	0.0635	0.0681
			1.11	1.11	0.0420	0.0372	0.1918	0.1727	0.1117	0.1069
				1.97	0.0471	0.0472	0.1908	0.1872	0.1007	0.0973
	45	23	0.59	1.11	0.0399	0.0358	0.0979	0.0874	0.0853	0.0814
				1.97	0.0447	0.0438	0.0986	0.0938	0.0628	0.0632
			1.11	1.11	0.0393	0.0378	0.1812	0.1754	0.1099	0.1115
				1.97	0.0504	0.0475	0.2007	0.1881	0.1038	0.1012
		34	0.59	1.11	0.0239	0.0233	0.0768	0.0763	0.0787	0.0779
				1.97	0.0373	0.0369	0.0762	0.0774	0.0622	0.063
			1.11	1.11	0.0318	0.0320	0.1405	0.1432	0.1052	0.1079
				1.97	0.0373	0.0351	0.1523	0.1434	0.0955	0.0916

Anderson–darling (A-D), and Chi-Square (CH2) tests for a significance level equals to 0.05 [13,14] provided a comprehensive study of these tests.

6.1. Vinyl chloride data application

Vinyl chloride is a known human carcinogen, exposure to this compound should be avoided as much as possible, and its level should be kept as low as technically possible.

Table 3. The root mean square errors (RMSEs) for the Weibull parameter β using the IMLE and Bayes methods at $T = 0.75$ and $T = 1.5$ with $m = (n/2$ and $3n/4)$ and $k = (m/2$ and $3\,m/4)$.

					IMLE		Gamma prior		Kernel prior	
n	m	k	α	β	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$
20	10	5	0.59	1.11	0.2584	0.2490	0.4092	0.4013	0.1124	0.1191
				1.97	0.6268	0.6288	0.7699	0.7590	0.2072	0.1981
			1.11	1.11	0.2533	0.2382	0.4497	0.4212	0.1233	0.1265
				1.97	0.5719	0.5335	0.8083	0.7646	0.1573	0.1424
		8	0.59	1.11	0.2201	0.2151	0.3481	0.3318	0.1239	0.1296
				1.97	0.5437	0.5442	0.6871	0.6798	0.1750	0.1658
			1.11	1.11	0.2129	0.2129	0.3507	0.3545	0.1310	0.1326
				1.97	0.4764	0.4763	0.6873	0.6914	0.1381	0.1417
	15	8	0.59	1.11	0.2198	0.2198	0.3470	0.3436	0.1304	0.1269
				1.97	0.5442	0.5443	0.6831	0.6782	0.1694	0.1634
			1.11	1.11	0.2129	0.2129	0.3535	0.3515	0.1310	0.1265
				1.97	0.4766	0.4767	0.6982	0.6968	0.1478	0.1434
		11	0.59	1.11	0.1896	0.1890	0.2601	0.2737	0.1380	0.1362
				1.97	0.4596	0.4594	0.5709	0.5622	0.1477	0.1426
			1.11	1.11	0.1894	0.1894	0.2747	0.2741	0.1320	0.1373
				1.97	0.4193	0.4194	0.5835	0.5756	0.1465	0.1364
40	20	10	0.59	1.11	0.2510	0.2253	0.3126	0.2630	0.1165	0.1193
				1.97	0.6301	0.5647	0.6557	0.5768	0.1981	0.1723
			1.11	1.11	0.2523	0.2184	0.3697	0.2773	0.1189	0.1174
				1.97	0.5723	0.4892	0.7252	0.5992	0.1485	0.1416
		15	0.59	1.11	0.2296	0.2228	0.2673	0.2617	0.1260	0.1222
				1.97	0.5647	0.5647	0.5764	0.5797	0.1699	0.1722
			1.11	1.11	0.2183	0.2184	0.2817	0.2842	0.1279	0.1208
				1.97	0.4891	0.4888	0.5950	0.5835	0.1402	0.1335
	30	15	0.59	1.11	0.2293	0.2210	0.2701	0.2527	0.1217	0.1213
				1.97	0.5645	0.5465	0.5843	0.5482	0.1748	0.1633
			1.11	1.11	0.2184	0.2130	0.2825	0.2643	0.1205	0.1205
				1.97	0.4889	0.4763	0.6002	0.5665	0.1426	0.1355
		23	0.59	1.11	0.1882	0.1879	0.1935	0.2022	0.1244	0.1322
				1.97	0.4547	0.4557	0.4032	0.3913	0.1405	0.1331
			1.11	1.11	0.1864	0.1864	0.1904	0.1972	0.1243	0.1287
				1.97	0.4117	0.4119	0.4067	0.4223	0.1283	0.1368
60	30	15	0.59	1.11	0.2424	0.2184	0.2476	0.2091	0.1139	0.1189
				1.97	0.6134	0.5464	0.5695	0.4815	0.1906	0.1661
			1.11	1.11	0.2521	0.2132	0.3225	0.2229	0.1103	0.1133
				1.97	0.5725	0.4765	0.6653	0.4813	0.1457	0.1292
		23	0.59	1.11	0.2270	0.2158	0.2198	0.2046	0.1163	0.1153
				1.97	0.5582	0.5347	0.4917	0.4570	0.1659	0.1577
			1.11	1.11	0.2166	0.2098	0.2328	0.2083	0.1134	0.1106
				1.97	0.4849	0.4687	0.5067	0.4750	0.1326	0.1325
	45	23	0.59	1.11	0.2266	0.2164	0.2193	0.1985	0.1160	0.1188
				1.97	0.5583	0.5353	0.4889	0.4543	0.1648	0.1559
			1.11	1.11	0.2167	0.2098	0.2313	0.2094	0.1148	0.1159
				1.97	0.4847	0.4688	0.5128	0.4737	0.1381	0.1326
		34	0.59	1.11	0.1890	0.1892	0.1717	0.1657	0.1189	0.1151
				1.97	0.4580	0.4577	0.3456	0.3467	0.1390	0.1382
			1.11	1.11	0.1875	0.1875	0.1680	0.1628	0.1155	0.1117
				1.97	0.4142	0.4143	0.3498	0.3445	0.1258	0.1232

It is known that the concentration of vinyl chloride in drinking water of 0.5 mg/liter is being associated with an increased risk of liver and Brain tumors for exposure beginning at adulthood and will double cancer risk for continuous exposure from birth. Therefore, we consider the dataset used by [7] which represents 34 data points in mg/L from the vinyl chloride obtained from clean upgrade monitoring wells as:

Table 4. The root mean square errors (RMSEs) for the Weibull extension parameter α using the IMLE and Bayes methods at $T = 0.75$ and $T = 1.5$ with $m = (n/2$ and $3n/4)$ and $k = (m/2$ and $3\,m/4)$.

					IMLE		Gamma prior		Kernel prior	
n	m	K	α	β	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$
20	10	5	0.59	1.11	0.0597	0.0545	0.1726	0.1576	0.1148	0.1146
				1.97	0.0866	0.0675	0.1649	0.1573	0.0964	0.0901
			1.11	1.11	0.0763	0.0607	0.3469	0.2914	0.1331	0.1302
				1.97	0.0888	0.0688	0.3151	0.2948	0.1369	0.1317
		8	0.59	1.11	0.0621	0.0519	0.1759	0.1549	0.1130	0.1064
				1.97	0.0753	0.0635	0.1622	0.1526	0.0935	0.0894
			1.11	1.11	0.0678	0.0654	0.3336	0.3023	0.1295	0.1333
				1.97	0.0743	0.0643	0.3013	0.2877	0.1304	0.1244
	15	8	0.59	1.11	0.0586	0.0474	0.1721	0.1388	0.1118	0.1072
				1.97	0.0739	0.0592	0.1598	0.1438	0.0947	0.0931
			1.11	1.11	0.0697	0.0576	0.3282	0.2632	0.1342	0.1255
				1.97	0.0736	0.0645	0.3065	0.2789	0.1311	0.1292
		11	0.59	1.11	0.0520	0.0496	0.1482	0.1412	0.1137	0.1111
				1.97	0.0587	0.0579	0.1453	0.1449	0.0931	0.0937
			1.11	1.11	0.0579	0.0563	0.2829	0.2661	0.1266	0.1273
				1.97	0.0623	0.0608	0.2804	0.2720	0.1256	0.1228
40	20	10	0.59	1.11	0.0490	0.0448	0.1301	0.1140	0.0857	0.0866
				1.97	0.0626	0.0554	0.1245	0.1149	0.0697	0.0689
			1.11	1.11	0.0555	0.0465	0.2730	0.2226	0.1141	0.1177
				1.97	0.0733	0.0560	0.2608	0.2234	0.1140	0.1088
		15	0.59	1.11	0.0491	0.0445	0.1305	0.1133	0.0844	0.0881
				1.97	0.0634	0.0558	0.1262	0.1159	0.0702	0.0735
			1.11	1.11	0.0521	0.0502	0.2520	0.2232	0.1146	0.1166
				1.97	0.0675	0.0555	0.2545	0.2211	0.1130	0.1066
	30	15	0.59	1.11	0.0488	0.0416	0.1283	0.1056	0.0874	0.0901
				1.97	0.0653	0.0536	0.1302	0.1103	0.0712	0.0692
			1.11	1.11	0.0516	0.0454	0.2545	0.2117	0.1153	0.1145
				1.97	0.0662	0.0561	0.2525	0.2203	0.1118	0.1089
		23	0.59	1.11	0.0406	0.0387	0.1033	0.1005	0.0848	0.0844
				1.97	0.0523	0.0513	0.1078	0.1050	0.0702	0.0687
			1.11	1.11	0.0402	0.0420	0.1904	0.1962	0.1077	0.1121
				1.97	0.0534	0.0512	0.2092	0.2045	0.1114	0.1081
60	30	15	0.59	1.11	0.0450	0.0431	0.1055	0.0983	0.0750	0.0733
				1.97	0.0592	0.0541	0.1096	0.0987	0.0623	0.0598
			1.11	1.11	0.0479	0.0410	0.2175	0.1825	0.1055	0.1017
				1.97	0.0626	0.0549	0.2188	0.1952	0.1031	0.1012
		23	0.59	1.11	0.0478	0.0421	0.1126	0.0964	0.0742	0.0744
				1.97	0.0619	0.0536	0.1111	0.0984	0.0632	0.0614
			1.11	1.11	0.0471	0.0414	0.2137	0.1846	0.1056	0.1047
				1.97	0.0580	0.0512	0.2087	0.1845	0.0990	0.0958
	45	23	0.59	1.11	0.0452	0.0357	0.1046	0.0851	0.0745	0.0731
				1.97	0.0603	0.0493	0.1090	0.0868	0.0634	0.0580
			1.11	1.11	0.0478	0.0347	0.2131	0.1629	0.1084	0.1036
				1.97	0.0625	0.0472	0.2194	0.1695	0.1023	0.0950
		34	0.59	1.11	0.0355	0.0358	0.0806	0.0872	0.0708	0.0735
				1.97	0.0510	0.0495	0.0914	0.0886	0.0610	0.0601
			1.11	1.11	0.0329	0.0323	0.1548	0.1591	0.0964	0.1007
				1.97	0.0497	0.0492	0.1770	0.1735	0.0992	0.1005

5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8.0, 0.8, 0.4, 0.6, 0.9, 0.4, 2.0, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1.0, 0.2, 0.1, 0.1, 1.8, 0.9, 2.0, 4.0, 6.8, 1.2, 0.4, 0.2.

We found the Weibull extension and Weibull models are a good fit for this dataset as shown in Table 6 and Figures 1(a) and 2(a) respectively. To study the concentration of the vinyl chloride in the water of these wells based on this data set, we find the estimates of the

Table 5. The root mean square errors (RMSEs) for the Weibull Extension parameter β using the IMLE and Bayes methods at $T = 0.75$ and $T = 1.5$ with $m = (n/2$ and $3n/4)$ and $k = (m/2$ and $3\,m/4)$.

					IMLE		Gamma prior		Kernel prior	
n	m	k	α	β	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$	$T = 0.75$	$T = 1.5$
20	10	5	0.59	1.11	0.2665	0.2424	0.4013	0.3534	0.1181	0.1183
				1.97	0.6555	0.5664	0.7834	0.6820	0.2839	0.2362
			1.11	1.11	0.2643	0.2389	0.4417	0.3543	0.1234	0.1242
				1.97	0.5631	0.5033	0.7938	0.7020	0.2366	0.2198
		8	0.59	1.11	0.2626	0.2425	0.3987	0.3476	0.1219	0.1137
				1.97	0.6218	0.5661	0.7486	0.6825	0.2618	0.2353
			1.11	1.11	0.2535	0.2386	0.4080	0.3569	0.1212	0.1243
				1.97	0.5406	0.5032	0.7617	0.6938	0.2285	0.2141
	15	8	0.59	1.11	0.2646	0.2233	0.4010	0.2986	0.1174	0.1257
				1.97	0.6214	0.5230	0.7487	0.6123	0.2633	0.2161
			1.11	1.11	0.2543	0.2277	0.4124	0.3083	0.1295	0.1261
				1.97	0.5408	0.4755	0.7690	0.6354	0.2320	0.2113
		11	0.59	1.11	0.2326	0.2243	0.3298	0.2970	0.1235	0.1233
				1.97	0.5426	0.5230	0.6531	0.6226	0.2275	0.2231
			1.11	1.11	0.2331	0.2279	0.3354	0.3117	0.1324	0.1336
				1.97	0.4890	0.4765	0.6665	0.6394	0.2141	0.2142
40	20	10	0.59	1.11	0.2653	0.2444	0.3090	0.2587	0.1180	0.1198
				1.97	0.6236	0.5689	0.6320	0.5517	0.2571	0.2313
			1.11	1.11	0.2658	0.2388	0.3633	0.2634	0.1239	0.1166
				1.97	0.5760	0.5038	0.7101	0.5694	0.2269	0.2118
		15	0.59	1.11	0.2672	0.2444	0.3099	0.2549	0.1173	0.1143
				1.97	0.6237	0.5688	0.6368	0.5539	0.2597	0.2324
			1.11	1.11	0.2561	0.2384	0.3289	0.2612	0.1154	0.1165
				1.97	0.5512	0.5033	0.6786	0.5619	0.2266	0.2076
	30	15	0.59	1.11	0.2653	0.2366	0.3025	0.2301	0.1177	0.1129
				1.97	0.6240	0.5574	0.6390	0.5238	0.2613	0.2227
			1.11	1.11	0.2565	0.2358	0.3327	0.2522	0.1213	0.1191
				1.97	0.5509	0.4957	0.6737	0.5475	0.2229	0.2089
		23	0.59	1.11	0.2308	0.2293	0.2222	0.2170	0.1159	0.1146
				1.97	0.5365	0.5364	0.4938	0.4868	0.2181	0.2144
			1.11	1.11	0.2301	0.2303	0.2251	0.2309	0.1161	0.1208
				1.97	0.4816	0.4830	0.4987	0.5122	0.1999	0.2081
60	30	15	0.59	1.11	0.2580	0.2445	0.2456	0.2159	0.1131	0.1092
				1.97	0.6039	0.5690	0.5430	0.4759	0.2501	0.2265
			1.11	1.11	0.2567	0.2388	0.2827	0.2202	0.1168	0.1111
				1.97	0.5469	0.5031	0.5909	0.4940	0.2148	0.2049
		23	0.59	1.11	0.2674	0.2444	0.2692	0.2165	0.1160	0.1133
				1.97	0.6238	0.5689	0.5624	0.4702	0.2541	0.2236
			1.11	1.11	0.2544	0.2389	0.2766	0.2240	0.1136	0.1139
				1.97	0.5407	0.5037	0.5796	0.4850	0.2150	0.2015
	45	23	0.59	1.11	0.2612	0.2274	0.2475	0.1778	0.1138	0.1075
				1.97	0.6135	0.5330	0.5465	0.4075	0.2490	0.2063
			1.11	1.11	0.2530	0.2291	0.2726	0.1838	0.1121	0.1064
				1.97	0.5469	0.4801	0.5942	0.4203	0.2166	0.1921
		34	0.59	1.11	0.2319	0.2274	0.1866	0.1850	0.1087	0.1127
				1.97	0.5396	0.5331	0.4180	0.4060	0.2074	0.2057
			1.11	1.11	0.2313	0.2294	0.1933	0.1871	0.1105	0.1059
				1.97	0.4841	0.4794	0.4336	0.4118	0.1953	0.1890

parameters, which represent the scale and shape of the concentration using both models, to determine the average concentration in the water. We observed that the estimates based on the IMLE and Bayes methods for α to both models fall in the interval [0.47, 1.2], indicating that the above dataset is moderately right-skewed and this means that the concentration decreases with increasing time, see Figures 1(b) and 2(b). Also, the IMLE and

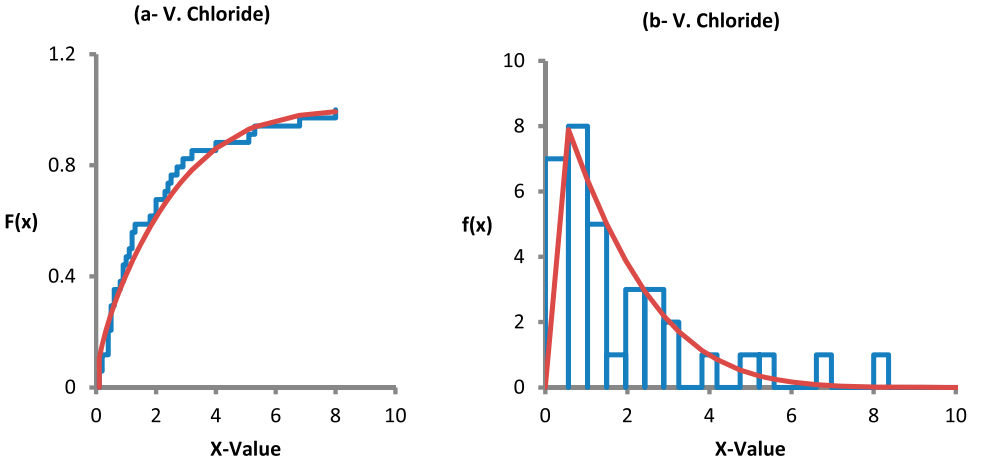

Figure 1. (a) The Empirical CDF and the CDF for the Weibull extension model based on the Vinyl Chloride Data. (b) The Histogram and the PDF for the Weibull extension model based on the Vinyl Chloride Data.

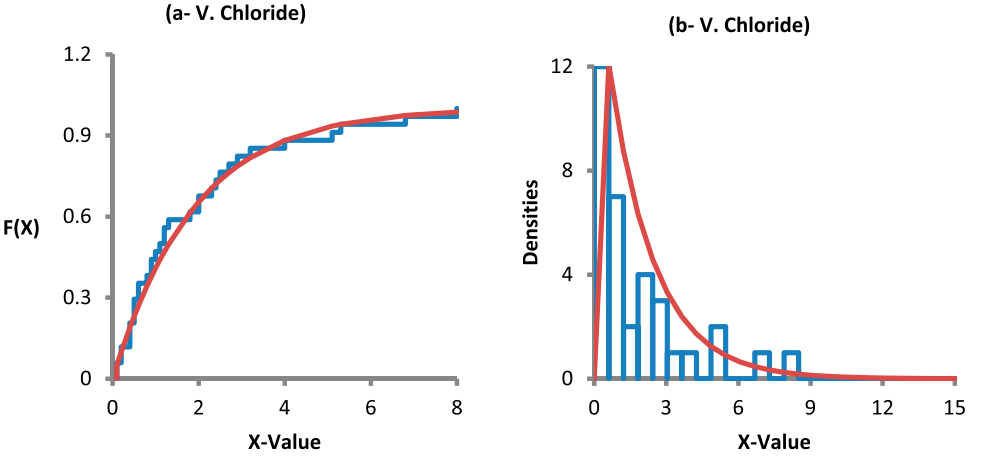

Figure 2. (a) The Empirical and the CDF for the Weibull model based on the Vinyl Chloride Data. (b) The Histogram and the PDF for the Weibull model based on the Vinyl Chloride Data.

Bayes estimates for β to both models fall in the interval [0.28, 0.6], which ensures that the dataset is right-skewed and the vinyl chloride concentration will decrease with increasing time, therefore monitoring these wells is very important.

6.2. Leukemia data application

In the health care field, Leukemia affects the blood status that can be detected with the Blood Cell Count (CBC), and mostly leukemia patients undergo chemotherapy. Therefore, we study the effect of this treatment on leukemia patients based on a dataset collected and used by the Ministry of Health Hospital in the Kingdom of Saudi Arabia [2], which indicated the lifetimes in days for forty-three patients with leukemia after given them the chemotherapy treatment:

Table 6. The critical and calculated values for the K-S, A-D and CH2 tests and their powers (*p*-values) for the models.

Models	Data	The tests	Critical value	Calculated value	The *P*-values	$\hat{\alpha}$	$\hat{\beta}$
Weibull **Model**	Chloride $N = 34$	K-S	0.8624	0.5355	0.6525	1.0102	0.5263
		A-D	0.7504	0.2826	0.6708		
		CH2	15.428	4.9912	0.4474		
	Leukemia $N = 43$	K-S	0.8699	0.7285	0.1915	2.5533	1.04E-08
		A-D	0.7598	0.9159	0.0206		
		CH2	15.399	12.409	0.0528		
	Covid-19 $N = 57$	K-S	0.8705	0.8314	0.7660	11.6306	6.4E-21
		A-D	0.7563	0.5064	0.2030		
		CH2	15.416	3.6758	0.5500		
Weibull Extension **Model**	Chloride $N = 34$	K-S	0.8621	0.6731	0.2629	0.5056	0.3029
		A-D	0.7732	0.6886	0.0814		
		CH2	8.6539	7.5479	0.2409		
	Leukemia $N = 43$	K-S	0.8689	0.5681	0.5572	0.3086	9.2E-05
		A-D	0.7800	0.5051	0.2086		
		CH2	10.119	10.783	0.3776		
	COVID-19 $N = 57$	K-S	0.8708	0.7739	0.1290	0.7031	5.9E-08
		A-D	0.7798	0.4779	0.2337		
		CH2	8.6616	2.8097	0.6774		

The MLE's for the parameters for these datasets have been calculated.

115, 181, 255, 418, 441, 461, 516, 739, 743, 789, 807, 865, 924, 983, 1025, 1062, 1063, 1165, 1191, 1222, 1222, 1251, 1277, 1290, 1357, 1369, 1408, 1455, 1478, 1549, 1578, 1578, 1599, 1603, 1605, 1696, 1735, 1799, 1815, 1852, 1899, 1925, 1965.

We found the Weibull extension model is more fitting for this dataset than the Weibull model as shown in Table 6 and Figures 3(a) and 4(a) respectively. To study the effect of chemotherapy on patients based on this dataset, we find that the IML and Bayes estimates for α to both models fall in the interval [0.25, 2.4] and for β lying in the interval [1E−03, 1E−010], which is approximately zero. This means that the curves that represent this dataset are approximately symmetric, see Figures 3(b) and 4(b). Also, the Weibull extension model is more fitting than the Weibull model, where the parameter estimate for α is less than one ($0.25 \leq \alpha* \leq 0.3$), but the estimate for α to the Weibull model is greater than one ($1.2 \leq \alpha* \leq 2.4$), indicating that the curve will be right oblique, see Figures 3(b) and 4(b). As a result, the hazard rate will decrease with increasing time for patients and this means that they are more likely to reach their maximum normal lifespan. So in general, based on the two models, this dataset indicates that the patient's lifespan is more stable and lives longer due to the chemotherapy dose, and is highly effective in giving patients more hope of survival.

6.3. COVID-19 data application

Here, we propose a concrete application with an actual dataset to assess interest in Weibull and Weibull extension models. The considered data set is the deaths from COVID-19 in Egypt, which are related to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Unfortunately, this epidemic spread rapidly at the beginning of the year

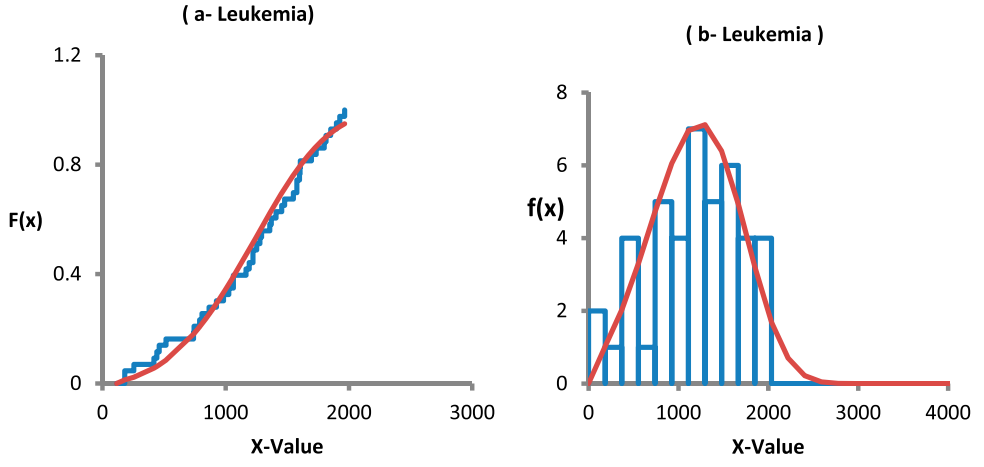

Figure 3. (a) The Empirical CDF and the CDF for the Weibull extension based on the Leukemia Data. (b) The Histogram and the PDF for the for the Weibull extension based on the Leukemia Data.

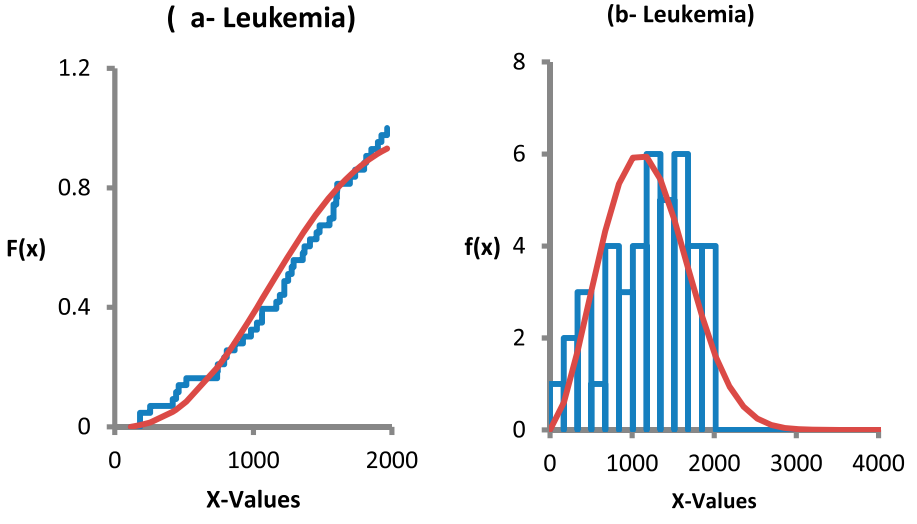

Figure 4. (a) The Empirical and the CDF for the Weibull model based on the Leukemia Data. (b) The Histogram and the PDF for the Weibull model based on the Leukemia Data.

2020, killing thousands of victims, and forcing governments to take exceptional measures to protect their people. Naturally, the general understanding of the COVID-19 pandemic is a challenge for all scientists, but it is essential for the sake of future generations. In this section, we contribute modestly to the subject by applying these models to analyze the daily data set of confirmed deaths for COVID-19 in Egypt from December 22, 2020, to February 16, 2021, as shown below, to provide an estimate for some important measures such as the average cases, the standard deviation of cases, and the probability to have a certain number of cases in the near future to make more efforts to confront these epidemics. This dataset obtained from the following email address: http://covid.gov.Eg/Coronatracker.com/Country/Egypt.

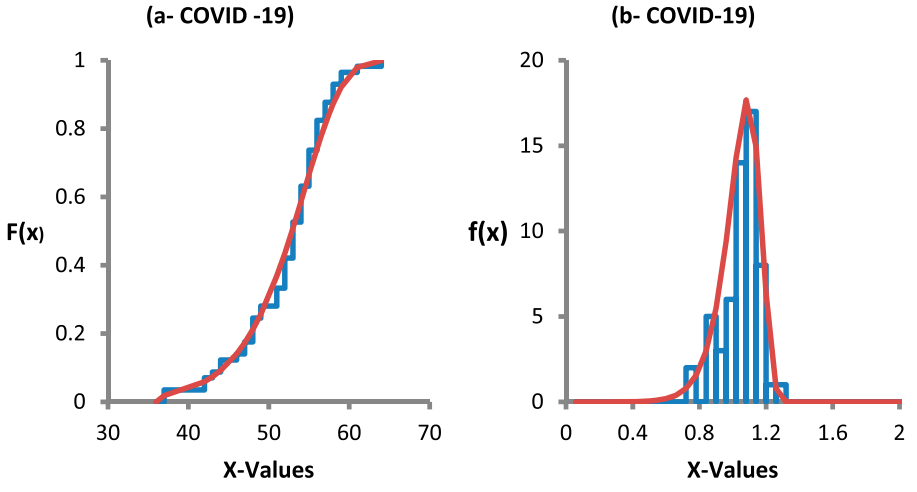

Figure 5. (a) The Empirical CDF and the CDF for the Weibull extension based on the COVID-19 Death Data. (b) The Histogram and the PDF for the for the Weibull extension based on the COVID-19 Death Data.

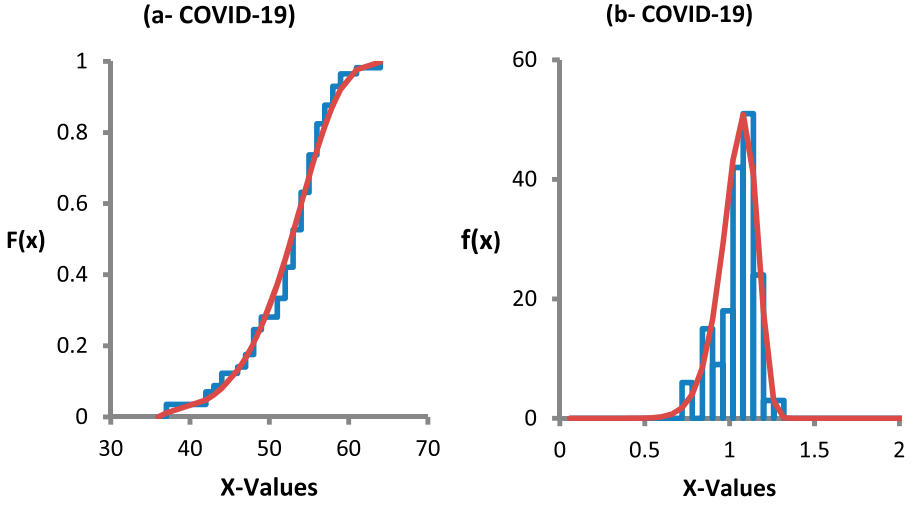

Figure 6. (a) The Empirical and the CDF for the Weibull model based on the COVID-19 Death Data. (b) The Histogram and the PDF for the Weibull model based on the COVID-19 Death Data.

It is given as follows:

37, 42, 51, 49, 43, 53, 61, 54, 56, 55, 56, 54, 64, 58, 55, 57, 54, 56, 57, 55, 52, 55, 58, 59, 52, 54, 56, 55, 58, 51, 54, 52, 49, 57, 53, 55, 48, 54, 48, 46, 53, 44, 47, 53, 52, 48, 44, 47, 48, 52, 53, 53, 42, 36, 59, 56, 51.

For the COVID-19 data set, the results of both distributions are a good fit for this dataset as shown in Figures 5(a) and 6(a). From the information in Table 2, the MLEs for the model parameters are derived to estimate the PDFs for the Weibull and Weibull extension models

Table 7. The estimates and the (RMSEs) in parentheses for Weibull extension and Weibull parameters α and β based on the IMLE and Bayes method at the hyper parameters (A = 2, B = 3, C = 4, D = 2): for $m = n/2, k = m/2$.

			Weibull extension model			Weibull model		
				Bayes			Bayes	
Data	T	Par.	IMLE	Gamma	Kernel	IMLE	Gamma	Kernel
Vinyl. data $N = 34$	Complete sample	α	0.4817	0.4791	0.5019	0.9494	0.9426	1.1971
			(0.0238)	(0.0264)	(0.0037)	(0.0609)	(0.0678)	(0.1869)
		β	0.2817	0.3564	0.3086	0.4891	0.6126	0.3011
			(0.0212)	(0.0535)	(0.0057)	(0.0371)	(0.0864)	(0.2251)
	0.75	α	1.0238	0.7741	0.9339	1.3448	1.1102	1.2873
			(0.0280)	(0.2217)	(0.0619)	(0.0119)	(0.2268)	(0.0455)
		β	0.1297	0.0122	0.2017	0.2328	0.3579	0.3741
			(0.0137)	(0.1313)	(0.0582)	(0.0251)	(0.0999)	(0.1162)
	1.5	α	0.7757	1.5575	0.7530	1.1071	0.9899	1.1377
			(0.0209)	(0.7608)	(0.0437)	(0.0805)	(0.1977)	(0.0499)
		β	0.12066	0.0892	0.1225	0.2158	0.1865	0.2097
			(0.0092)	(0.0407)	(0.0074)	(0.0212)	(0.0505)	(0.0273)
	3.5	α	0.5930	0.6706	0.5659	1.1001	0.9709	1.1643
			(0.0342)	(0.0433)	(0.0613)	(0.0251)	(0.1264)	(0.0391)
		β	0.1084	0.0013	0.0501	0.2059	0.3315	0.3857
			(0.0121)	(0.1193)	(0.0705)	(0.0158)	(0.1091)	(0.1640)
Leuk. data $N = 43$	Complete sample	α	0.2934	0.2532	0.3001	2.4087	1.8020	1.2941
			(0.0154)	(0.0556)	(0.0088)	(0.1268)	(0.7335)	(0.9501)
		β	8.6E−05	2.64E−03	1.0E−04	1.1E−08	2.1E−03	1.0E−04
			(4.5E−06)	(2.55E−03)	(9.2E−06)	(5.9E−10)	(2.1E−04)	(9.9E−05)
	850	α	0.2741	0.1980	0.3001	1.5374	1.2896	1.2792
			(0.0144)	(0.0905)	(0.0116)	(0.1708)	(0.4186)	(0.4290)
		β	2.8E−04	0.0125	1.0E−04	2.8E−06	0.0293	1.0E−04
			(1.5E−05)	(0.0122)	(1.9E−04)	(3.1E−07)	(0.0293)	(9.7E−05)
	1250	α	0.2571	0.1812	0.2999	1.4147	1.1118	1.2498
			(0.0135)	(0.0894)	(0.0293)	(0.1572)	(0.4601)	(0.3221)
		β	5.4E−04	0.0187	1.8E−04	6.6E−06	0.0266	1.0E−04
			2.8E−05	(0.0181)	(3.8E−04)	(7.3E−07)	(0.0266)	(9.3E−05)
	1700	α	0.2739	0.2527	0.3001	1.7539	1.2083	1.3603
			(0.0144)	(0.0356)	(0.0118)	(0.1949)	(0.7405)	(0.5885)
		β	2.5E−04	0.0031	1.0E−04	5.7E−07	0.0069	1.0E−04
			(1.3E−05)	(0.0028)	(1.9E−04)	(6.3E−08)	(0.0069)	(9.9E−05)
Covid-19 data $N = 57$	Complete sample	α	0.6682	0.5896	0.5501	5.9183	5.5325	5.2679
			(0.0352)	(0.1137)	(0.1533)	(0.1566)	(0.2435)	(0.2899)
		β	5.5E−08	5.0E−05	1.0E−04	0.89231	0.9542	0.7671
			(2.9E−09)	(5.0E−05)	(9.9E−05)	(0.1523)	(0.2421)	(0.0146)
	40	α	0.5706	0.3530	0.5001	5.9178	6.1374	5.5213
			(0.0301)	(0.2476)	(0.1005)	(0.3123)	(0.0931)	(0.5402)
		β	3.9E−06	7.0E−03	1.0E−04	3.7E−12	9.3E−08	1.0−E04
			(2.1E−07)	(7.0E−03)	(9.6E−05)	(2.0E−13)	(9.3E−08)	(9.9E−05)
	50	α	0.5803	0.3604	0.5101	6.2578	6.4811	6.3450
			(0.0306)	(0.2506)	(0.1108)	(0.3306)	(0.1107)	(0.4251)
		β	2.8E−06	6.5E−03	1.0E−04	1.0E−12	4.4E−08	1.0−04
			(1.4E−07)	(6.5E−03)	(9.7E−05)	(5.4E−14)	(4.4E08)	(1.0E−04)
	55	α	0.5842	0.3618	0.5002	6.3911	6.6143	6.4210
			(0.0308)	(0.2533)	(0.1149)	(0.3377)	(0.1145)	(0.3552)
		β	2.3E−06	6.4E−03	1.0E−04	5.9E−13	3.4E−08	4.2E−10
			(1.3E−07)	(6.4E−03)	(9.8E−05)	(3.2E−14)	(3.4E−08)	(3.2E−10)

as given respectively by:

$$\hat{f}_1(x) = \hat{f}(x;\hat{\alpha},\hat{\beta}) = \hat{\alpha}\hat{\beta}x^{\hat{\alpha}-1}e^{-\hat{\beta}x^{\hat{\alpha}}},$$

and $\hat{f}_2(x) = \hat{f}(x;\hat{\alpha},\hat{\beta}) = \hat{\alpha}\hat{\beta}x^{\hat{\alpha}-1}e^{-x^{\hat{\alpha}}}exp(-\hat{\beta}(e^{x^{\hat{\alpha}}}-1))$.

Thus, $\hat{f}_1(x)$ and $\hat{f}_2(x)$ are the estimated functions of the unobservable underlying PDFs of the number of COVID-19 deaths in Egypt. Using these functions, one can estimate some interesting measures. By denoting X the random variable modeling the daily COVID-19 confirmed death cases in Egypt during the epidemic, the probability that X belongs to a chosen interval, say [a, b], can be estimated by $\hat{P}_{A,B} = P(a < x < b) = \int_a^b \hat{f}_i(x)dx, \quad i = 1, 2$. More generally, the estimation of the mean of a certain function of X, say T(X) can be estimated as $\hat{\mu} = E(T(x)) = \int_0^\infty T(x)\hat{f}_i(x)dx, \quad i = 1, 2$.

For instance, the average number of COVID-19 deaths in Egypt can be approximated with precision by taking T(x) = x, and so on. Thus, based on this dataset, the average number of COVID-19 deaths in Egypt based on the Weibull model is 52 confirmed deaths with a standard deviation of 6 death, while the average number of deaths based on the Weibull extension model is 52 with a standard deviation of 5 deaths. From Table 7, the IML and Bayes estimates for β to both models are nearly zero, ensuring that the standard deviation is small and the distributions are bell-shaped with a thin tail as shown in Figures 5(b) and 6(b) which indicate the COVID-19 deaths will decrease rapidly in the next few months. Thus, these results indicate that both models are very efficient for modeling the COVID-19 datasets.

Finally, the results in Table 6, based on the Weibull extension and Weibull models, are a good fit for the vinyl chloride data where the power of the tests is greater than the significance level of the tests, but the Weibull extension model is more fitting for the Leukemia data than the Weibull model as shown in Figures 3(a) and 4(a). The results in Table 7 for these datasets indicate that the estimated RMSEs values based on the IMLE method are smaller than those based on the Bayesian method for large values of T with considering the MLEs are the true values of the parameters. Hence, the results of these datasets ensure the simulation results

7. Conclusions

We conclude that the improved MLE method is more efficient than the Bayesian method using the informative and kernel priors, based on the generalized progressive hybrid censored scheme. However, the estimates based on the kernel prior are more efficient than the ones based on the informative prior and are relatively close to the IML estimates. Thus, the IMLE method is a viable estimation method for effectively any lifetime model and is reliable and easy to apply especially for medical, biological, and engineering researchers. Moreover, we applied the proposed methods to analyze real data applications including the COVID-19 pandemic, which concluded that the number of COVID-19 deaths in Egypt is declining in the next few months. Hence, the proposed models provide a better understanding of the COVID-19 epidemic and may provide insights for researchers and potential users with models that can be broadly applicable to real-life situations.

Acknowledgements

The author is very grateful to the Academy of Scientific Research and Technology (ASRT) in Egypt, which supported this work with a grant from Science UP, Faculty of Science, with Project No. 6380. The author is very grateful to the anonymous referees and the Editor for their constructive efforts and Comments that improved this work.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This research was supported by the Academy of Scientific Research and Technology in Egypt, with a grant from Science UP, Faculty of Science, with Project No. 6380.

References

[1] M. Ahsanullah, M. Maswadah, and M. Ali Seham, *Kernel inference on the generalized gamma distribution based on generalized order statistics*. J. Stat. Theory Appl. 12 (2013), pp. 152–172.
[2] A.M. Abd-Elrahman, and K.S. Sultan, *Reliability estimation based on general progressive censored data from the Weibull model: comparison between Bayesian and classical approaches*. Int. J. Stat. LX (2007), pp. 239–257.
[3] S. Abu El Fotouh, and M.A. Nassar, *Estimation for the parameters of the Weibull extension model based on generalized order statistics*. Int. J. Contemp. Math. Sci. 6 (2011), pp. 1749–1760.
[4] N. Balakrishnan, and R. Aggarwala, *Progressive Censoring: Theory, Methods and Applications*, Birkhäuser Publishers, Boston, 2000.
[5] N. Balakrishnan, and M. Kateri, *On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data*. Stat. Probab. Lett. J. 78 (2008), pp. 2971–2975.
[6] N. Balakrishnan, and E. Cramer, *The art of Progressive Censoring: Applications to Reliability and Quality Statistics for Industry and Technology*, Springer, New York, 2014.
[7] D.K. Bhaumik, K. Kapur, and R.D. Gibbons, *Testing parameters of a gamma distribution for small samples*. Technometrics. 51 (2009), pp. 326–334.
[8] P. Chan, H. Ng, and F. Su, *Exact likelihood inference for the two-parameter exponential distribution under Type-II progressive hybrid censoring*. Metrika 78 (2015), pp. 747–770.
[9] Z. Chen, *A new two-parameter lifetime distribution with bathtub-shape or increasing failure rate function*. Stat. Probab. Lett. 49 (2000), pp. 155–161.
[10] Y. Cho, H. Sun, and K. Lee, *An estimation of the entropy for a Rayleigh distribution based on doubly generalized Type-II hybrid censored samples*. Entropy 16 (2014), pp. 3655–3669.
[11] Y. Cho, H. Sun, and K. Lee, *Exact likelihood inference for an exponential parameter under generalized progressive hybrid censoring scheme*. Stat. Method 23 (2015a), pp. 18–34.
[12] Y. Cho, H. Sun, and K. Lee, *Estimating the entropy of a Weibull distribution under generalized progressive hybrid censoring*. Entropy 17 (2015b), pp. 102–122.
[13] G.M. Cordeiro, J.M.F. Carrasco, and E.M.M. Ortega, *A generalized modified Weibull distribution for lifetime modeling*. Comput. Stat. Data Anal. 53 (2008), pp. 450–462.
[14] G.M. Cordeiro, and A. Lemonte, *The -Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling*. Comput. Stat. Data Anal. 55 (2011), pp. 1445–1461. doi:10.1016/j.csda.2010.10.007.
[15] M.A.T. Elshahat, *Quasi-likelihood estimation of the parameters of the Weibull extension model*. Aust. J. Basic Appl. Sci. 1 (2007), pp. 271–278.
[16] J. Gorny, and E. Cramer, *Exact likelihood inference for exponential distribution under generalized progressive hybrid censoring schemes*. Stat. Methodol. 29 (2016), pp. 70–94.
[17] A. Guillamon, J. Navarro, and J.M. Ruiz, *Kernel density estimation using weighted data*. Commun. Statist.-Theory Meth. 27 (1998), pp. 2123–2135.

[18] A. Guillamon, J. Navarro, and J.M. Ruiz, *A note on kernel estimators for positive valued random variables*. Sankhya A 6 (1999), pp. 276–281.
[19] F. Hemmati, and E. Khorram, *Statistical analysis of log-normal distribution under Type-II progressive hybrid censoring schemes*. Commun. Stat.- Simul. Comput. 42 (2013), pp. 52–75.
[20] M.S. Khan, G.R. Pasha, and A.H. Pasha, *Reliability and quantile analysis of the Weibull distribution*. J. Stat. 14 (2007), pp. 32–52.
[21] D. Kundu, and A. Joarder, *Analysis of Type-II progressive hybrid censored data*. Comput. Stat. Data Anal. 50 (2006), pp. 2509–2528.
[22] Y. Lei, *Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine*. J. For. Sci. 54 (2008), pp. 566–571.
[23] M. Maswadah, *Kernel inference on the inverse Weibull distribution*. Korean Commun. Stat. 13 (2006), pp. 503–512.
[24] M. Maswadah, *Kernel inference on the Weibull distribution*. Proc. of the Third National Statistical Conference. Lahore, Pakistan 14 (2007), pp. 77–86.
[25] M. Maswadah, *Kernel Inference on the Type-II extreme Value Distribution*. Proceedings of the Tenth Islamic Countries Conference on Statistical Sciences (ICCS-X), Lahore, Pakistan II (2010), pp. 870–880.
[26] M. Maswadah, *Empirical Bayes inference for the Weibull Model*. Comput. Stat. J. 28 (2013). doi:10.1007/s00180-013-0438-1.
[27] M. Maswadah, and A.A. Faheem, *Conditional inference for the Weibull extension model based on the generalized order statistics*. Pak. J. Stat. Oper. Res. XIV (2018), pp. 119–214.
[28] M. Maswadah, *An optimal point estimation method for the inverse Weibull model parameters using the Runge-Kutta method*. Aligarh J. Stat. (2021). (Accepted).
[29] M.M. Mohie El-Din, and M. Nagy, *Estimation of the inverse Weibull distribution under generalized progressive hybrid censoring scheme*. J. Stat. Appl. Pro. Lett. 4 (2017), pp. 97–107.
[30] J.O. Ramsay, G. Hooker, D. Campbell, and J. Cao, *Parameter estimation for differential equations: a generalized smoothing approach*. J. R. Statist. Soc. B 69 (2007), pp. 741–796.
[31] A.M. Sarhan, and A. Al-Ruzaizza, *Statistical inference in connection with the Weibull model using Type-II progressive censored data with random scheme*. Pak. J. Stat. 26 (2010), pp. 267–279.
[32] Y. Tang, M. Xie, and T.N. Goh, *Statistical analysis of a Weibull extension model*. Commun. Statist. Theor. Meth. 32 (2003), pp. 913–928.
[33] S.K. Upadhyay, A. Gupta, and D.K. Day, *Bayesian modelling of bathtub-shaped hazard rate using various Weibull extensions and related issues of model selection*. Sankhya B 74 (2012), pp. 15–43.
[34] R. Wang, N. Sha, B. Gu, and X. Xu, *Statistical analysis of a Weibull extension with bathtub-shaped failure rate function*. Adv. Stat. 2 (2014). doi:10.1155/2014/304724
[35] J.W. Wu, H. Lu, C. Chen, and C. Wu, *Statistical inference about shape parameter of the new two-parameter Bathtub-shaped lifetime distribution*. Qual. Reliab. Eng. Int. 20 (2004), pp. 607–616.
[36] J.W. Wu, C.C. Wu, and M.H. Tsai, *Optimal parameter estimation of the two-parameter bathtub-shaped lifetime distribution based on a Type-II right censored sample*. Appl. Math. Comput. 2 (2005), pp. 807–819.
[37] S.J. Wu, *Estimation of the parameters of the Weibull distribution with progressively censored data*. J. Japan Statist. Soc. 32 (2002), pp. 155–163.
[38] M. Xie, Y. Tang, and T.N. Goh, *A modified Weibull extension with bathtub-shaped failure rate function*. Reliab. Eng. Syst. Saf. 76 (2002), pp. 279–285.
[39] L.F. Zhang, M. Xie, and L.C. Tang, *Bias correction for the least squares estimator of Weibull shape parameter with complete and censored data*. Reliab. Eng. Syst. Saf. 91 (2006), pp. 930–939.
[40] L.F. Zhang, M. Xie, and L.C. Tang, *A study of two estimation approaches for parameters of Weibull distribution based on WPP*. Reliab. Eng. Syst. Saf. 92 (2007), pp. 360–368.