

## Research Article

# EM Algorithm for Estimating the Parameters of Weibull Competing Risk Model

Mohamed Kayid 

Department of Statistics and Operations Research, College of Science, King Saud University, Riyadh, Saudi Arabia

Correspondence should be addressed to Mohamed Kayid; drkayid@ksu.edu.sa

Received 14 August 2021; Revised 2 October 2021; Accepted 5 October 2021; Published 21 October 2021

Academic Editor: Domenico Speranza

Copyright © 2021 Mohamed Kayid. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the most commonly used models in survival analysis is the additive Weibull model and its generalizations. They are well suited for modeling bathtub-shaped hazard rates that are a natural form of the hazard rate. Although they have some advantages, the maximum likelihood and the least square estimators are biased and have poor performance when the data set contains a large number of parameters. As an alternative, the expectation-maximization (EM) algorithm was applied to estimate the parameters of the additive Weibull model. The accuracy of the parameter estimates and the simulation study confirmed the advantages of the EM algorithm.

## 1. Introduction

There are many situations in which the hazard rate function shows a bathtub shape (BTS) with three life periods, early decreasing hazard rate period, useful period (where the hazard rate is approximately constant), and eventually increasing hazard rate period. Lai et al. [1] and Nadarajah [2] presented a list of distributions with BTS hazard rate. Many authors assumed a BTS hazard rate model in their research, and among them, we can point to Block et al. [3], Glaser [4], Leemis and Beneke [5], Mi [6, 7], Mitra and Basu [8], and Noughabi et al. [9].

Xie and Lai [10] and Jiang and Murthy [11, 12] studied some aspects of the additive Weibull model with the hazard rate function:

$$h(x) = \alpha_1 \lambda_1^{\alpha_1} x^{\alpha_1 - 1} + \alpha_2 \lambda_2^{\alpha_2} x^{\alpha_2 - 1}, x \geq 0, \quad (1)$$

as a good candidate for describing the bathtub-shaped failure rate function. They urged that the statistical inference about this model is complex due to the number of the parameters, and as a remedy, they suggested the reduced model by considering  $\lambda_1 = \lambda_2$  and  $\alpha_2 = 1/\alpha_1$  which still accommodates the bathtub-shaped hazard rate.

Lai et al. [13] added one constant magnitude to the additive Weibull hazard rate (1) to provide more realistic model:

$$h(x) = \alpha_1 \lambda_1^{\alpha_1} x^{\alpha_1 - 1} + \alpha_2 \lambda_2^{\alpha_2} x^{\alpha_2 - 1} + \lambda_3, x \geq 0. \quad (2)$$

In addition, Bebbington et al. [14] considered the additive Weibull models (1) and (2) in their research to express the concept of the useful period of life of a bathtub-shaped hazard rate distribution. They concluded that the additive Weibull model is sufficiently flexible to describe a bathtub-shaped hazard rate. The common estimator of the parameters of these models is the maximum likelihood estimator (MLE).

The EM algorithm is an iterative algorithm for estimating parameters of models involving latent variables, e.g., when data is derived from a mixture or competing risk model. It was used by Dempster et al. [15], Balakrishnan et al. [16], Davies et al. [17], Yang et al. [18], and Okamura and Dohi [19] to estimate the parameters in their models. In this paper, we use the EM algorithm to estimate the parameters of the additive model (2) and show by a simulation study that this algorithm gives a better estimate than the MLE and the least square estimator (LSE).

The paper has been organized as the following. In Section 2, we present a short representation of the MLE and

LSE of the parameters. Then, the EM algorithm has been discussed. Section 3 provides a simulation study for comparing the results related to MLE, LSE, and EM estimator. In Section 4, a data set has been analyzed to show the applicability of the proposed estimators.

## 2. The MLE, Least Square, and EM

Let  $x_1, x_2, \dots, x_n$  represents a realization of an iid random sample of the additive Weibull hazard rate distribution with the hazard rate (2). The log-likelihood function has five parameters and is of the form

$$l(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \lambda_3) = \sum_{i=1}^n \ln \left( \alpha_1 \lambda_1^{\alpha_1} x_i^{\alpha_1-1} + \alpha_2 \lambda_2^{\alpha_2} x_i^{\alpha_2-1} + \lambda_3 \right) - \sum_{i=1}^n ((\lambda_1 x_i)^{\alpha_1} + (\lambda_2 x_i)^{\alpha_2} + \lambda_3 x_i). \quad (3)$$

To find the MLE, we should find admissible values of  $(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \lambda_3)$  which maximize the log-likelihood function.

To find the LSE of the parameters, we apply the reliability function related to the hazard rate model (2) which is

$$R(x) = e^{-(\lambda_1 x)^{\alpha_1} - (\lambda_2 x)^{\alpha_2} - \lambda_3 x}, \quad x \geq 0. \quad (4)$$

The empirical reliability function is defined to be

$$\tilde{R}(t) = \frac{1}{n} \sum_{i=1}^n I(t < x_i), \quad (5)$$

in which the indicator function  $I(t < x_i)$  equals 1 when  $t < x_i$  and otherwise is 0. So, the LSE of the parameters can be computed by minimizing the following sum of squares of errors in terms of  $(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \lambda_3)$ .

$$S^2 = \sum_{i=1}^n \left( \tilde{R}(x_i) - e^{-(\lambda_1 x_i)^{\alpha_1} - (\lambda_2 x_i)^{\alpha_2} - \lambda_3 x_i} \right)^2. \quad (6)$$

**2.1. The EM Algorithm.** Let  $X_{1i}, X_{2i}$ , and  $X_{3i}$ ,  $i = 1, 2, \dots, n$  follows from the Weibull distributions with parameters  $(\alpha_1, \lambda_1)$  and  $(\alpha_2, \lambda_2)$  and the exponential distribution with mean  $1/\lambda_3$ , respectively. Assume that in a lifetime experiment, the observations are realizations of the competing risk random variable  $X_i = \min \{X_{1i}, X_{2i}, X_{3i}\}$ . This means that the lifetime event may be due to one of three competing causes. Let the latent random variable  $Z_i$  with the support  $\{1, 2, 3\}$  such that

$$Z_i = \begin{cases} 1 & \text{when } X_{1i} \leq \min \{X_{2i}, X_{3i}\}, \\ 2 & \text{when } X_{2i} \leq \min \{X_{1i}, X_{3i}\}, \\ 3 & \text{when } X_{3i} \leq \min \{X_{1i}, X_{2i}\}. \end{cases} \quad (7)$$

With these notations, the likelihood function is

$$\begin{aligned} L(\theta; x, z) &= \prod_{i=1}^n \left( \prod_{j=1}^3 f(Z_i = j | X_i = x_i, \theta)^{I(Z_i=j)} \right) f(X_i = x_i | \theta) \\ &= \prod_{i=1}^n \frac{\prod_{j=1}^3 \lambda_j(x_i | \theta)^{I(Z_i=j)}}{\lambda_1(x_i | \theta) + \lambda_2(x_i | \theta) + \lambda_3(x_i | \theta)} \\ &\quad \times (f_1(x_i | \theta)R_2(x_i | \theta)R_3(x_i | \theta) \\ &\quad + f_2(x_i | \theta)R_1(x_i | \theta)R_3(x_i | \theta) \\ &\quad + f_3(x_i | \theta)R_1(x_i | \theta)R_2(x_i | \theta)) \\ &= \prod_{i=1}^n \left( \prod_{j=1}^3 \lambda_j(x_i | \theta)^{I(Z_i=j)} \right) \\ &\quad \cdot R_1(x_i | \theta)R_2(x_i | \theta)R_3(x_i | \theta), \end{aligned} \quad (8)$$

where  $R_j(x | \theta)$ ,  $f_j(x | \theta)$ , and  $\lambda_j(x | \theta)$  show the corresponding reliability function, the density function, and the hazard rate function of  $X_{ji}$ ,  $j = 1, 2, 3$ . Then, the log-likelihood function is

$$\begin{aligned} l(\theta; x, z) = \ln L(\theta; x, z) &= \sum_{i=1}^n \sum_{j=1}^3 I(Z_i = j) \ln \lambda_j(x_i | \theta) \\ &\quad + \sum_{i=1}^n \sum_{j=1}^3 \ln R_j(x_i | \theta). \end{aligned} \quad (9)$$

The EM algorithm is an iterative algorithm, and every iteration of it consists of two consecutive steps, namely, the E step and the M step. In the E step, the expectation of the log-likelihood with respect to the estimate of the conditional probabilities of the latent variables has been constructed. Then, in the M step, the constructed expectation of the E step is maximized to compute the estimate of the parameters in the current iteration.

**2.2. The E Step.** Suppose that the estimate of  $\theta$  at iteration  $t$  be denoted by  $\theta_t$ , then the conditional distribution of  $Z_i$  is

$$p_{i1,t} = P(Z_i = 1 | X_i = x_i, \theta_t) = \frac{f_1(x_i | \theta_t)R_2(x_i | \theta_t)R_3(x_i | \theta_t)}{f(x_i | \theta_t)}, \quad (10)$$

$$p_{i2,t} = P(Z_i = 2 | X_i = x_i, \theta_t) = \frac{f_2(x_i | \theta_t)R_1(x_i | \theta_t)R_3(x_i | \theta_t)}{f(x_i | \theta_t)}, \quad (11)$$

and  $p_{i3,t} = 1 - p_{i1,t} - p_{i2,t}$  where

$$\begin{aligned} f(x_i | \theta_t) &= f_1(x_i | \theta_t)R_2(x_i | \theta_t)R_3(x_i | \theta_t) \\ &\quad + f_2(x_i | \theta_t)R_1(x_i | \theta_t)R_3(x_i | \theta_t) \\ &\quad + f_3(x_i | \theta_t)R_1(x_i | \theta_t)R_2(x_i | \theta_t). \end{aligned} \quad (12)$$

The probabilities  $p_{i1,t}$ ,  $p_{i2,t}$ , and  $p_{i3,t}$  are called the membership probabilities.

Now we define the expectation of the log-likelihood function (8) with respect to the conditional distribution of  $Z_i$ .

$$\begin{aligned}
Q(\theta | \theta_t) &= E_{Z_i | X, \theta_t}(l(\theta; x, Z)) = \sum_{i=1}^n p_{1i,t} \left( \ln \left( \alpha_1 \lambda_1^{\alpha_1} x_i^{\alpha_1 - 1} \right) \right. \\
&\quad \left. - (\lambda_1 x_i)^{\alpha_1} - (\lambda_2 x_i)^{\alpha_2} - \lambda_3 x_i \right) \\
&\quad + \sum_{i=1}^n p_{2i,t} \left( \ln \left( \alpha_2 \lambda_2^{\alpha_2} x_i^{\alpha_2 - 1} \right) - (\lambda_1 x_i)^{\alpha_1} \right. \\
&\quad \left. - (\lambda_2 x_i)^{\alpha_2} - \lambda_3 x_i \right) + \sum_{i=1}^n p_{3i,t} \left( \ln \lambda_3 - (\lambda_1 x_i)^{\alpha_1} \right. \\
&\quad \left. - (\lambda_2 x_i)^{\alpha_2} - \lambda_3 x_i \right) = \sum_{i=1}^n p_{1i,t} \ln \alpha_1 \lambda_1^{\alpha_1} x_i^{\alpha_1 - 1} \\
&\quad - \sum_{i=1}^n (\lambda_1 x_i)^{\alpha_1} + \sum_{i=1}^n p_{2i,t} \ln \alpha_2 \lambda_2^{\alpha_2} x_i^{\alpha_2 - 1} \\
&\quad - \sum_{i=1}^n (\lambda_2 x_i)^{\alpha_2} + \sum_{i=1}^n p_{3i,t} \ln \lambda_3 - \sum_{i=1}^n \lambda_3 x_i.
\end{aligned} \tag{13}$$

So the  $Q(\theta | \theta_t)$  can be written as sum of three distinct expressions  $Q_1(\theta | \theta_t)$ ,  $Q_2(\theta | \theta_t)$ , and  $Q_3(\theta | \theta_t)$ , where

$$Q_1(\theta | \theta_t) = \sum_{i=1}^n p_{1i,t} \ln \alpha_1 \lambda_1^{\alpha_1} x_i^{\alpha_1 - 1} - \sum_{i=1}^n (\lambda_1 x_i)^{\alpha_1}, \tag{14}$$

$$Q_2(\theta | \theta_t) = \sum_{i=1}^n p_{2i,t} \ln \alpha_2 \lambda_2^{\alpha_2} x_i^{\alpha_2 - 1} - \sum_{i=1}^n (\lambda_2 x_i)^{\alpha_2}, \tag{15}$$

$$Q_3(\theta | \theta_t) = \sum_{i=1}^n p_{3i,t} \ln \lambda_3 - \sum_{i=1}^n \lambda_3 x_i. \tag{16}$$

These statements will be applied in the M step, to compute the estimates of the parameters.

2.3. *The M Step.* The estimate of the parameters at iteration  $t + 1$  can be obtained by

$$\hat{\alpha}_{1,t+1}, \hat{\lambda}_{1,t+1} = \arg \max_{\alpha, \lambda} Q_1(\theta | \theta_t), \tag{17}$$

$$\hat{\alpha}_{2,t+1}, \hat{\lambda}_{2,t+1} = \arg \max_{\alpha, \lambda} Q_2(\theta | \theta_t), \tag{18}$$

$$\hat{\lambda}_{3,t+1} = \arg \max_{\lambda} Q_3(\theta | \theta_t). \tag{19}$$

We should optimize  $Q_1(\theta | \theta_t)$  and  $Q_2(\theta | \theta_t)$  numerically since they have not closed form for their critical points. But, the point which maximizes  $Q_3(\theta | \theta_t)$  has a closed form, and by solving the equation  $(\partial/\partial \lambda_3)Q_3 = 0$ , we have

$$\hat{\lambda}_{3,t+1} = \frac{\sum_{i=1}^n p_{3i,t}}{\sum_{i=1}^n x_i}. \tag{20}$$

The iterative process can be concluded if  $Q(\theta_{t+1} | \theta_{t+1}) < Q(\theta_t | \theta_t) + \varepsilon$  for some small predefined  $\varepsilon$ .

### 3. Simulation Study

To provide a random instance of the competing risk model with hazard rate (2), we simulate one random instance of Weibull with parameters  $(\alpha_1, \lambda_1)$ , namely,  $X_{11}$ , one random instance of Weibull with parameters  $(\alpha_2, \lambda_2)$ , namely,  $X_{12}$ , and one random instance of the exponential distribution with parameter  $\lambda_3$ , namely,  $X_{13}$ . Then, the random variable  $X_1 = \min \{X_{11}, X_{12}, X_{13}\}$  follows from the desired competing risk model.

In every run of the simulation study, we drive  $r = 500$  replicates of samples of sizes  $n = 50$  and  $100$ . Then, for each sample, the parameters have been estimated applying the MLE, LSE, or EM algorithm (see Supplementary Materials for all R codes used for simulation study). Every cell of Table 1 shows the results of one run. The results contain the bias ( $B$ ), the absolute bias ( $AB$ ), and the mean squared error ( $MSE$ ) which, for example for  $\alpha_1$ , have been computed by the following relations.

$$B_{\alpha_1} = \frac{1}{r} \sum_{i=1}^r (\hat{\alpha}_{1i} - \alpha_1), \tag{21}$$

$$AB_{\alpha_1} = \frac{1}{r} \sum_{i=1}^r |\hat{\alpha}_{1i} - \alpha_1|, \tag{22}$$

$$MSE_{\alpha_1} = \frac{1}{r} \sum_{i=1}^r (\alpha \hat{\alpha}_{1i} - \alpha_1)^2, \tag{23}$$

where  $\hat{\alpha}_{1i}$  is the estimate of  $\alpha_1$  based on the  $i$ th replication. Some important observations of the simulation results have been pointed out in the following.

- (i) As sample size increases, the  $AB$  and  $MSE$  decrease in all approaches
- (ii) The EM estimator outperforms the LSE and MLE in terms of  $AB$  and  $MSE$  in all cases
- (iii) The LSE outperforms the MLE in terms of  $AB$  and  $MSE$  in all cases

### 4. Applications

Lawless [20] analyzed failure times of some electrical appliances. The scaled TTT transform plot drawn in Figure 1 shows a bathtub shape for the hazard rate function. This gives us some nonparametric information indicating that the data come from a BT hazard rate model. So we tried to fit some distributions accommodating bathtub-shaped hazard rate to this data set. We use the MLE and the EM algorithm to fit the five parameters

TABLE 1: Every cell consists of the bias, the absolute bias, and the mean squared error for five parameters  $\alpha_1, \lambda_1, \alpha_2, \lambda_2,$  and  $\lambda_3$  from top to bottom, respectively.

	$n$	The EM algorithm			The LSE			The MLE			
		B	AB	MSE	B	AB	MSE	B	AB	MSE	
$\alpha_1 = 0.8$ $\lambda_1 = 0.01$ $\alpha_2 = 1.3$ $\lambda_2 = 0.02$ $\lambda_3 = 0.04$	50	0.114788	0.209461	0.073327	0.064984	0.266447	0.295491	0.150792	0.278228	0.475591	
		-0.001410	0.003386	0.000017	0.015999	0.020659	0.000707	0.026929	0.030571	0.001399	
		0.354883	0.414357	0.325645	0.450476	0.596059	1.216075	1.710262	1.829587	30.45458	
	100	0.001644	0.004211	0.000030	0.0008324	0.015219	0.000326	0.003705	0.013934	0.000315	
		0.006223	0.007680	0.000101	-0.018570	0.025329	0.000851	-0.025577	0.031822	0.001172	
		0.090420	0.161060	0.041728	0.075634	0.208668	0.090426	0.074398	0.221743	0.092066	
	$\alpha_1 = 1.1$ $\lambda_1 = 0.1$ $\alpha_2 = 0.9$ $\lambda_2 = 0.2$ $\lambda_3 = 0.3$	50	-0.001857	0.002883	0.000011	0.014107	0.018573	0.000558	0.022671	0.027087	0.001132
			0.269857	0.300074	0.151175	0.351361	0.551910	0.787756	1.266889	1.407221	38.77269
			0.000711	0.002802	0.000012	0.007800	0.013041	0.000264	0.000171	0.012493	0.000243
100		0.006117	0.006801	0.000072	-0.017343	0.024008	0.000788	-0.018373	0.027122	0.000936	
		0.087247	0.189355	0.067184	0.388006	0.476704	0.656696	1.686055	1.736100	193.3874	
		0.010944	0.026164	0.001247	0.131134	0.184852	0.053799	0.057770	0.124714	0.025303	
$\alpha_1 = 1.1$ $\lambda_1 = 0.1$ $\alpha_2 = 0.9$ $\lambda_2 = 0.2$ $\lambda_3 = 0.3$	50	0.049117	0.165828	0.047906	-0.011756	0.306156	0.273852	0.019276	0.239581	0.241092	
		0.011833	0.033445	0.001913	0.082536	0.215583	0.070223	0.161446	0.250880	0.084662	
		0.010969	0.033325	0.001867	-0.149815	0.212298	0.056467	-0.165021	0.245053	0.068212	
	100	0.057342	0.128094	0.027325	0.332300	0.424207	0.583780	0.919810	0.977701	8.022046	
		0.005099	0.018084	0.000525	0.104450	0.160100	0.040268	0.030251	0.106644	0.017581	
		0.035447	0.107349	0.018491	-0.011164	0.227723	0.132404	0.008467	0.179967	0.311876	
$\alpha_1 = 1.1$ $\lambda_1 = 0.1$ $\alpha_2 = 0.9$ $\lambda_2 = 0.2$ $\lambda_3 = 0.3$	100	0.004282	0.023096	0.000883	0.087848	0.195858	0.057835	0.170213	0.248756	0.082494	
		0.003564	0.025512	0.001046	-0.151875	0.206708	0.053532	-0.158807	0.236292	0.064980	

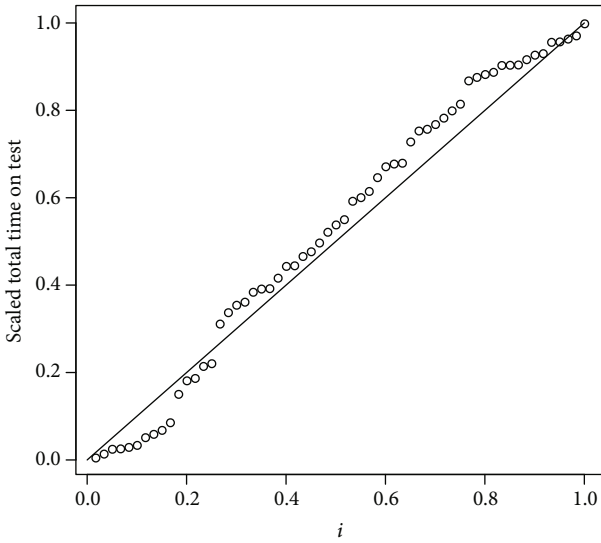
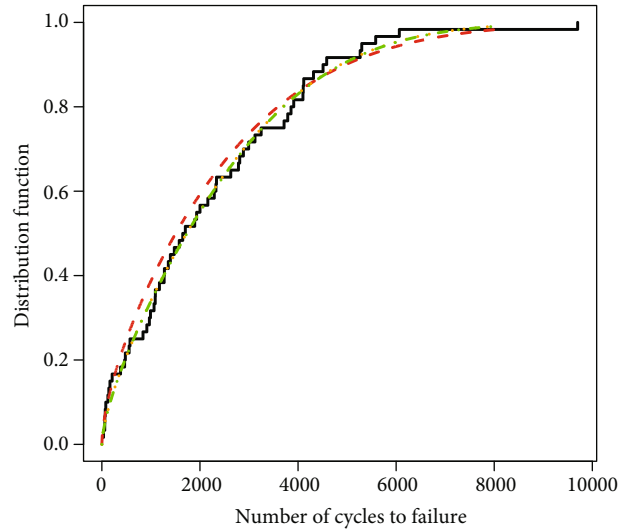


FIGURE 1: The scaled TTT transform plot for data sets of Table 2.

competing risk model (2). Also, the reduced model with the reliability function

$$R(x) = \exp \left\{ -(\lambda x)^\alpha - (\lambda x)^{1/\alpha} \right\}, \quad (24)$$

has been fitted by computing the MLE of the parameters. In Figure 2, the cumulative distribution function (CDF) for fitted models has been drawn. There is a significant distance



— The empirical distribution function  
 - - - The reduced model with two parameters (MLE)  
 . . . The five parameters model (MLE)  
 - . - The five parameters model (EM)

FIGURE 2: The empirical CDF and some fitted models to data set of Table 2.

between the fitted model (14) and the empirical CDF which shows that the reduced model may be improper in some examples.

Moreover, some results of fit have been abstracted in Table 3. Based on the Kolmogorov-Smirnov (K-S) statistics

TABLE 2: Failure time of electrical appliances in terms of 1000 s cycles.

34	59	61	69	80	123	142	165	210	381
479	556	574	839	917	969	991	1064	1088	1091
1270	1275	1355	1397	1477	1578	1649	1702	1893	1932
2161	2292	2326	2337	2628	2785	2811	2886	2993	3122
3715	3790	3857	3912	4100	4106	4116	4315	4510	4584
5299	5583	6065	9701						

TABLE 3: The results of fitting some suitable models to the data set.

The model	The method	Estimations	K-S statistics	K-S $p$ value	AIC
Model (2)	EM	$\hat{\alpha}_1 = 0.62125, \hat{\lambda}_1 = 0.00011$ $\hat{\alpha}_2 = 1.98826, \hat{\lambda}_2 = 0.00021$ $\hat{\lambda}_3 = 0.00012$	0.05246	0.9936	-
Model (2)	MLE	$\hat{\alpha}_1 = 0.62445, \hat{\lambda}_1 = 0.00011$ $\hat{\alpha}_2 = 1.99617, \hat{\lambda}_2 = 0.00021$ $\hat{\lambda}_3 = 0.00012$	0.05359	0.9917	1049.014
Model (20)	MLE	$\hat{\alpha} = 0.00023, \hat{\lambda} = 0.62891$	0.09987	0.5539	1044.176

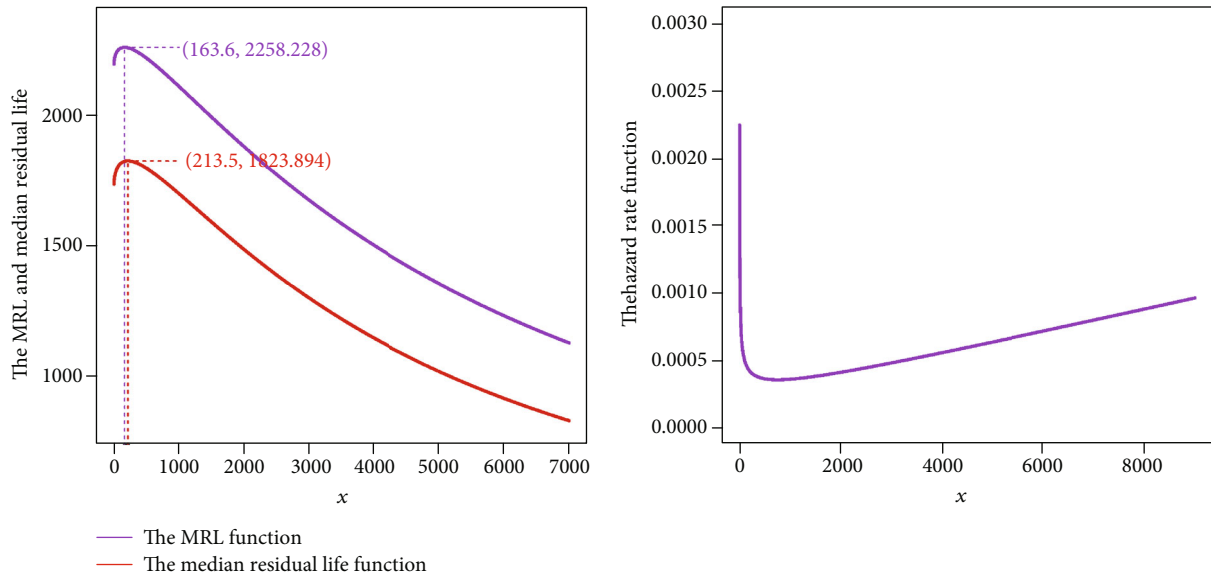


FIGURE 3: (a) The mean residual life and the median residual life of the fitted model. (b) The hazard rate function of this fitted model.

and  $p$  value, the competing risk model (2) which has been fitted by the EM algorithm gives the best description of the data. The MLE has also provided good results, but it is worthy to denote that we applied the estimates of the EM algorithm as the initial values in the likelihood maximization process. All of the fitted models confirm a BT hazard rate model which was firstly recognized by the TTT transform plot. So it may be interesting to investigate the point which maximizes the mean residual life and/or the median residual life functions. These points are referred to burn-in points and show the time at which the component is in its most

reliable condition. The left side of Figure 3 draws the mean residual life function along with the median residual life function related to the best fitted model. Also, the burn-in points related to both functions have been determined in the figure. The right side of Figure 2 draws the hazard rate function of this model and shows a BT hazard rate model.

### 5. Conclusion

The competing risk model of the baseline distribution Weibull plays a vital role in describing nonmonotone hazard rate

models. One drawback of this model is that it has large number of the parameters which causes the estimation problem harder. Some authors suggested reduced versions to overcome this problem. But, there are many examples showing that the reduced model may not be proper. So, we implemented the EM algorithm for estimating the parameters. The simulation results confirm that this algorithm is better than MLE and LSE. As future works, such EM algorithm may be constructed for similar competing risk models or mixture models, for example, the gamma competing risk model with the following reliability function may be a good candidate:

$$R(x) = \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \Gamma(\alpha_1, \lambda_1 x) \Gamma(\alpha_2, \lambda_2 x), x \geq 0, \quad (25)$$

in which  $\Gamma(\alpha, t) = \int_t^\infty y^{\alpha-1} e^{-y} dy$  is the upper incomplete gamma function and  $\alpha_1 > 0$ ,  $\lambda_1 > 0$ ,  $\alpha_2 > 0$ , and  $\lambda_2 > 0$ .

### Data Availability

The lifetime data used to support the findings of this study are included within the article.

### Conflicts of Interest

There is no any conflict of interest.

### Acknowledgments

The author thanks the two anonymous reviewers for their comments and suggestions. This work is supported by Researchers Supporting Project number RSP-2021/392, King Saud University, Riyadh, Saudi Arabia.

### Supplementary Materials

The R codes used for simulation study and applications are presented in a supplementary file. (*Supplementary materials*)

### References

- [1] C. D. Lai, M. Xie, and D. N. P. Murthy, "Bathtub-Shaped Failure Rate Life Distributions. Handbook of Statistics Vol. 20 - Advances in Reliability," in *Handbook of Statistics*, N. Balakrishnan and C. R. Rao, Eds., Netherlands: Elsevier Science B.V.69-104, 2001.
- [2] S. Nadarajah, "Bathtub-shaped failure rate functions," *Quality & Quantity*, vol. 43, no. 5, pp. 855–863, 2009.
- [3] H. W. Block, T. H. Savits, and H. Singh, "A Criterion for burn-in that balances mean residual life and residual variance," *Operations Research*, vol. 50, no. 2, pp. 290–296, 2002.
- [4] R. E. Glaser, "Bathtub and related failure rate characterizations," *Journal of the American Statistical Association*, vol. 75, pp. 667–672, 1980.
- [5] L. M. Leemis and M. Beneke, "Burn-in models and methods: a review," *IIE Transactions*, vol. 22, no. 2, pp. 172–180, 1990.
- [6] J. Mi, "Maximization of a survival probability and its application," *Journal of Applied Probability*, vol. 31, no. 4, pp. 1026–1033, 1994.
- [7] J. Mi, "Bathtub failure rate and upside-down bathtub mean residual life," *IEEE Transactions on Reliability*, vol. 44, no. 3, pp. 388–396, 1995.
- [8] M. Mitra and S. K. Basu, "On some properties of the bathtub failure rate family of life distributions," *Microelectronics Reliability*, vol. 36, no. 5, pp. 679–684, 1996.
- [9] M. Noughabi, G. Borzadaran, and A. Roknabadi, "On the reliability properties of some weighted models of bathtub shaped hazard rate distributions," *Probability in the Engineering and Informational Sciences*, vol. 27, no. 1, pp. 125–140, 2013.
- [10] M. Xie and C. D. Lai, "Reliability analysis using additive Weibull model with bathtub-shaped failure rate function," *Reliability Engineering and System Safety*, vol. 52, pp. 87–93, 1996.
- [11] R. Jiang and D. N. P. Murthy, "Two sectional models involving three Weibull distributions," *Quality and Reliability Engineering International*, vol. 13, pp. 83–96, 1997.
- [12] R. Jiang and D. N. P. Murthy, "Parametric study of competing risk model involving two Weibull distributions," *International Journal of Reliability, Quality and Safety Engineering*, vol. 4, no. 1, pp. 17–34, 1997.
- [13] C. D. Lai, L. Y. Zhang, and M. Xie, "Mean residual life and other properties of Weibull related bathtub shaped failure rate distributions," *International Journal of Reliability, Quality and Safety Engineering*, vol. 11, pp. 113–132, 2004.
- [14] M. Bebbington, C. D. Lai, and R. Zitikis, "Useful periods for lifetime distributions with bathtub shaped hazard rate functions," *IEEE Transactions on Reliability*, vol. 55, no. 2, pp. 245–251, 2006.
- [15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, pp. 1–22, 1977.
- [16] N. Balakrishnan, H. Y. So, and M. H. Ling, "EM algorithm for one-shot device testing with competing risks under Weibull distribution," *IEEE Transactions on Reliability*, vol. 65, no. 2, pp. 973–991, 2016.
- [17] K. Davies, S. Pal, and J. A. Siddiqua, "Stochastic EM algorithm for generalized exponential cure rate model and an empirical study," *Journal of Applied Statistics*, vol. 48, no. 12, pp. 2112–2135, 2021.
- [18] J. Yang, J. Chen, and X. Wang, "EM algorithm for estimating reliability of multi-release open source software based on general masked data," *IEEE Access*, vol. 9, pp. 18890–18903, 2021.
- [19] H. Okamura and T. Dohi, "Application of EM algorithm to NHPP-based software reliability assessment with generalized failure count data," *Mathematics*, vol. 9, no. 9, p. 985, 2021.
- [20] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, Wiley, New York, 2nd ed edition, 2003.