IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 15, 2021, accepted November 1, 2021, date of publication November 4, 2021, date of current version November 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125456

One Generalized Mixture Pareto Distribution and Estimation of the Parameters by the EM Algorithm for Complete and Right-Censored Data

MOHAMED KAYID
Department of Statistics and Operations Research, College of Science, King Saud University, Riyadh 11451, Saudi Arabia
e-mail: drkayid@ksu.edu.sa

This work was supported by King Saud University, Riyadh, Saudi Arabia, under Researchers Supporting Project RSP-2021/392.

ABSTRACT A new mixture generalized Pareto distribution is introduced. Then, some of its attributes are explored. The maximum likelihood method and expectation maximization (EM) algorithm have been applied to estimate the parameters for complete and right-censored data. In a simulation study, the bias, absolute bias and mean squared error of the maximum likelihood estimator are compared with those related to the EM estimator. The results show that the absolute bias and mean squared error of the EM estimator are smaller than the related values for the maximum likelihood estimator. Finally, to illustrate its usefulness, the model has been applied to describe real data sets.

INDEX TERMS Generalized Pareto distribution, mixture model, maximum likelihood estimator, EM algorithm.

I. INTRODUCTION

The Pareto distribution is a power-law distribution that was originally used to describe wealth distribution. Additionally, it may be useful in describing observations from scientific, geophysical, actuarial, social, and quality control events and many other fields. The Pareto model may be applied to situations, where an equilibrium is found in the distribution of "small" to "large" values. In many cases, we may point to: the sizes of files transferred on the internet network by TCP/IP, which consists of many smaller files and few larger ones; the hard disk drive error rates, which consist of many small error rates and few large ones; the sizes of human settlements, which consist of many small values related to hamlets/villages and few large values related to cities; the oil reserves volumes in oil fields, which consist of many small fields and few large fields and many other examples. Among many studies in this field, Burroughs and Tebbens [1] analyzed observations from earthquakes and forest fire areas, and Schroeder *et al.* [2] fitted the data from disk-driven sector errors to this distribution.

The associate editor coordinating the review of this manuscript and approving it for publication was Jiajie Fan.

To provide more flexible distributions, various generalizations of the Pareto distribution have been proposed. Among them, Akinsete *et al.* [3] studied the beta Pareto distribution, Nassar and Nada [4] and Mahmoudi [5] considered the beta generalized Pareto, Zea *et al.* [6] presented the beta exponentiated Pareto distribution, Alzaatreh *et al.* [7] introduced the gamma Pareto distribution, ElbataL [8] investigated the Kumaraswamy exponentiated Pareto distribution, Bourguignon *et al.* [9] introduced the Kumaraswamy Pareto distribution, Mead [10] defined one generalized beta exponentiated Pareto distribution, Clifton *et al.* [11] applied one extended version of the Pareto distribution for novelty detection, and Papastathopoulos and Tawn [12] applied the extended generalized Pareto models to the tail estimation problem. Recently, Jayakumar *et al.* [13], Tahir *et al.* [14], Korkmaz *et al.* [15], Elbatal and Aryal [16], and Chananet and Phaphan [17] proposed new distributions based on the Pareto model.

One generalized Pareto distribution, i.e., the $GP(\alpha, \beta, \theta)$ distribution, can be defined by the probability density function (pdf)

$$f(x) = \frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)} \frac{1}{\beta} \frac{(\frac{x}{\beta})^{\theta-1}}{(1+\frac{x}{\beta})^{\alpha+\theta}}, \quad x > 0, \tag{1}$$

where $\alpha > 0, \beta > 0, \theta > 0$, see Beirlant *et al.* [18] and Wiborg [19]. When $\alpha \to \infty$ and $\lambda = \frac{\theta\beta}{\alpha}$ are fixed constants, this model tends to the gamma distribution with pdf

$$f(x) = \frac{1}{\Gamma(\theta)}(\frac{\theta}{\lambda})^{\theta}x^{\theta-1}e^{-\frac{\theta}{\lambda}x}, \quad x > 0. \tag{2}$$

In this paper, we propose a new model based on a mixture of the GP distribution (1). The proposed model extends mixtures of the exponential, gamma, or Pareto distributions and has sufficient flexibility to describe many real situations.

The paper is organized as follows. In Section 2, the new model is defined, and some of its statistical and reliability attributes are studied. In Section 3, the parameters of the model are estimated for the complete and right-censored data by the maximum likelihood estimator (MLE) and EM algorithm. The behavior of the estimators of the parameters has been investigated by a simulation study in Section 4. In Section 5, the proposed model is fitted to real data sets to show its applicability.

II. THE NEW MODEL

We propose a generalized Pareto mixture $GPM(\alpha, \beta, \gamma)$ by the pdf

$$f(x) = \frac{\alpha}{\beta(\gamma+1)}(1+\frac{x}{\beta})^{-\alpha-1}[\gamma + (\alpha+1)\frac{x}{\beta+x}], \quad x \geq 0, \quad \alpha > 0, \beta > 0, \gamma > 0. \tag{3}$$

which is a mixture of $GP(\alpha, \beta, 1)$ and $GP(\alpha, \beta, 2)$ with weights $\frac{\gamma}{\gamma+1}$ and $\frac{1}{\gamma+1}$, respectively.

Since the GP model tends to the gamma distribution, a special limiting case of the GPM is the quasi Lindley distribution studied by Shanker [20].

For a random lifetime X, the most important function in reliability engineering and survival analysis is the reliability function $R(x) = P(X \geq x)$ and gives the probability that an object works (survives) at least until a specified time x. The reliability function of GPM is

$$R(x) = \frac{1}{\gamma+1}(1+\frac{x}{\beta})^{-\alpha}\Big(\gamma+\alpha+1-\frac{\alpha\beta}{x+\beta}\Big), \quad x \geq 0, \quad \alpha > 0, \beta > 0, \gamma > 0. \tag{4}$$

Proposition 1: When $\alpha > k$, the k^{th} moments of $GPM(\alpha, \beta, \gamma)$ are finite and equal to

$$E(X^k) = \frac{\alpha\beta^k}{\gamma+1}\Big[\frac{\alpha+1}{\alpha-k} + \sum_{j=0}^{k}(-1)^{k-j}(\binom{k}{j}\frac{\gamma}{\alpha-j} - \binom{k+1}{j}\frac{\alpha+1}{\alpha-j+1})\Big]. \tag{5}$$

For $\alpha \leq k$, it is infinite.

Proof: Because $GPM(\alpha, \beta, \gamma)$ is a mixture of $X_1 \sim GP(\alpha, \beta, 1)$ and $X_2 \sim GP(\alpha, \beta, 2)$, we have

$$E(X^k) = \frac{\gamma}{\gamma+1}E(X_1^k) + \frac{1}{\gamma+1}E(X_2^k). \tag{6}$$

But

$$E(X_1^k) = \int_0^{\infty}\frac{\alpha}{\beta}\frac{x^k}{(1+\frac{x}{\beta})^{\alpha+1}}dx = \alpha\beta^k\int_1^{\infty}\frac{(y-1)^k}{y^{\alpha+1}}dy = \alpha\beta^k\sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}\int_1^{\infty}y^{j-\alpha-1}dy. \tag{7}$$

It is easy to check that for $\alpha \leq k$, (7) is infinite, but for $\alpha > k$, it can be simplified as follows:

$$E(X_1^k) = \alpha\beta^k\sum_{j=0}^{k}\binom{k}{j}\frac{(-1)^{k-j}}{\alpha-j}. \tag{8}$$

Using a similar approach, we can check that for $\alpha \leq k$, $E(X_2^k)$ is infinite, and for $\alpha > k$, we have

$$E(X_2^k) = \alpha(\alpha+1)\beta^k\sum_{j=0}^{k+1}\binom{k+1}{j}\frac{(-1)^{k+1-j}}{\alpha-j+1}. \tag{9}$$

Then, the result follows by (7), (8) and (9). □

The quantile function $q(p) = F^{-1}(p)$ for $GPM(\alpha, \beta, \gamma)$ has no closed form and can be numerically computed by solving the following equation:

$$\frac{(1-p)(\gamma+1)}{\beta^{\alpha}} = \frac{1}{(\beta+q(p))^{\alpha}}(\gamma+\alpha+1-\frac{\alpha\beta}{\beta+q(p)}). \tag{10}$$

When the corresponding moments exit, the Pearson's moment coefficient of skewness of a random variable X is defined to be

$$B = E\Big(\frac{X-\mu}{\sigma}\Big)^3 = \frac{E(X^3) - 3\mu\sigma^2 - \mu^3}{\sigma^3}, \tag{11}$$

and the kurtosis of X is

$$K = E\Big(\frac{X-\mu}{\sigma}\Big)^4 = \frac{E(X^4) - 4\mu E(X^3) + 6\mu^2E(X^2) - 3\mu^4}{\sigma^4}, \tag{12}$$

where $\mu = E(X)$, and σ^2 is the variance of X.

Moreover, the skewness and kurtosis of a distribution can be described in terms of the quantile function. MacGillivray [21] suggested the following relation for skewness:

$$B = \frac{q(1-u) + q(u) - 2q(0.50)}{q(1-u) - q(u)},$$

where $u \in (0, 0.5)$. When $u = 0.25$, it reduces to Bowley's measure of skewness, Bowley [22]. Moreover, Moor [23] introduced the kurtosis in terms of the quantile function by

$$K = \frac{q(\frac{7}{8}) - q(\frac{5}{8}) + q(\frac{3}{8}) - q(\frac{1}{8})}{q(\frac{6}{8}) - q(\frac{2}{8})}.$$

In the economics literature, for a cumulative distribution function (cdf) F, the Lorenz curve is:

$$L(p) = \frac{p\mu_p}{\mu} = \frac{\int_0^{q(p)}xdF(x)}{\int_0^{q(1)}xdF(x)}.$$

and provides a graphical representation of wealth inequality. In fact, $L(p)$ shows the proportion of overall income or wealth of the $100 \times p$ percent of people with lower income or wealth. This plot will be a convex plot joining two points (0, 0) and (1, 1). For an ideal society where every person has identical income, the Lorenz curve is a straight line that joins these points. For $GPM(\alpha, \beta, \gamma)$, μ is not finite for $\alpha \leq 1$, but μ_p is finite, so $L(p)$ is zero and does not describe the wealth distribution. Fortunately, Prendergast and Staudte [24] have proposed alternatives in terms of the quantile function, which replace μ by the median of the distribution and μ_p by $q(\frac{p}{2})$. Precisely

$$L_1(p) = p\frac{q(\frac{p}{2})}{q(0.5)},$$
$$L_2(p) = p\frac{q(\frac{p}{2})}{q(1-\frac{p}{2})},$$

and

$$L_3(p) = 2p\frac{q(\frac{p}{2})}{q(\frac{p}{2}) + q(1-\frac{p}{2})}.$$

A. RELIABILITY MEASURES

The failure rate (FR), mean residual life (MRL) and p-quantile residual life (p-QRL) concepts play important roles in describing the dynamic attributes of lifetime variables and have been studied from different perspectives in the reliability and survival analysis literature. The FR function of the $GPM(\alpha, \beta, \gamma)$ is

$$h(x) = \frac{f(x)}{R(x)} = \frac{\alpha\gamma(\beta + x) + \alpha(\alpha+1)x}{(\gamma+\alpha+1)(\beta+x)^2 - \alpha\beta(\beta+x)}.$$

For a distribution with reliability function R, the MRL function is:

$$m(x) = \frac{1}{R(x)}\int_x^{\infty} R(t)dt.$$

The following proposition gives the form of the MRL function for the proposed model.

Proposition 2: When $\alpha > 1$, the MRL function of $GPM(\alpha, \beta, \gamma)$ is finite and of the form

$$m(x) = \frac{(\beta+x)(\gamma+\alpha-\alpha\beta+\beta+1)}{(\alpha-1)(\beta+x)(\gamma+\alpha+1) - \alpha\beta(\alpha-1)}.$$

When $\alpha \leq 1$, the MRL is infinite.

Proof: It is straightforward to check that

$$\int_x^{\infty} R(t)dt = \Big(\frac{\gamma}{\gamma+1} + \frac{\alpha+1}{\gamma+1}\Big)\int_x^{\infty}(1+\frac{t}{\beta})^{-\alpha}dt - \frac{\alpha}{\gamma+1}\int_x^{\infty}(1+\frac{t}{\beta})^{-\alpha-1}dt. \quad (13)$$

The first integral of the right side of (13) is equal to $\frac{\beta}{(\alpha-1)}(1+\frac{x}{\beta})^{-\alpha+1}$ when $\alpha > 1$ and infinite when $\alpha \leq 1$. The second integral is equal to $\frac{\beta}{\alpha}(1+\frac{x}{\beta})^{-\alpha}$. Then, the result follows by some algebra. □

The p-QRL function of a distribution with reliability function R is defined to be

$$q_p(x) = R^{-1}(\bar{p}R(x)) - x = q(1-\bar{p}R(x)) - x, \quad x > 0, \quad (14)$$

where $\bar{p} = 1-p$, and $q()$ is the quantile function and defined by (10). Similar to the quantile function, the p-QRL function of GPM has no closed form and should be numerically computed.

III. ESTIMATION OF THE PARAMETERS

A. THE MLE FOR COMPLETE DATA

Suppose that $x_1, x_2, \ldots, x_n$ represent realizations from $GPM(\alpha, \beta, \gamma)$. The log-likelihood function of the parameters is

$$l(\alpha, \beta, \gamma; x) = n\ln\alpha - n\ln\beta - n\ln(\gamma+1) - (\alpha+1)\sum_{i=1}^{n}\ln(1+\frac{x_i}{\beta}) + \sum_{i=1}^{n}\ln[\gamma + (\alpha+1)\frac{x_i}{x_i+\beta}], \quad (15)$$

and the log-likelihood equations are

$$\frac{\partial}{\partial\alpha}l(\alpha, \beta, \gamma; x) = \frac{n}{\alpha} - \sum_{i=1}^{n}\ln(1+\frac{x_i}{\beta}) + \sum_{i=1}^{n}\frac{x_i}{(\alpha+\gamma+1)x_i + \gamma\beta} = 0,$$

$$\frac{\partial}{\partial\beta}l(\alpha, \beta, \gamma; x) = -\frac{n}{\beta} + (\alpha+1)\sum_{i=1}^{n}\frac{x_i}{\beta^2+\beta x_i} - (\alpha+1)\sum_{i=1}^{n}\frac{x_i}{\gamma(x_i+\beta)^2 + (\alpha+1)x_i(x_i+\beta)} = 0,$$

and

$$\frac{\partial}{\partial\gamma}l(\alpha, \beta, \gamma; x) = -\frac{n}{\gamma+1} + \sum_{i=1}^{n}\frac{x_i+\beta}{(\alpha+\gamma+1)x_i+\gamma\beta} = 0.$$

We calculate the MLE by directly maximizing the log-likelihood function (15) or solving the likelihood equations.

Let $l = \ln f(X)$. Then, the Fisher information matrix can be applied to obtain the variance of the MLE. Here, it is defined as follows:

$$K = \begin{bmatrix} E(-\frac{\partial^2 l}{\partial\alpha^2}) & E(-\frac{\partial^2 l}{\partial\alpha\partial\beta}) & E(-\frac{\partial^2 l}{\partial\alpha\partial\gamma}) \\ E(-\frac{\partial^2 l}{\partial\beta\partial\alpha}) & E(-\frac{\partial^2 l}{\partial\beta^2}) & E(-\frac{\partial^2 l}{\partial\beta\partial\gamma}) \\ E(-\frac{\partial^2 l}{\partial\gamma\partial\alpha}) & E(-\frac{\partial^2 l}{\partial\gamma\partial\beta}) & E(-\frac{\partial^2 l}{\partial\gamma^2}) \end{bmatrix}. \quad (16)$$

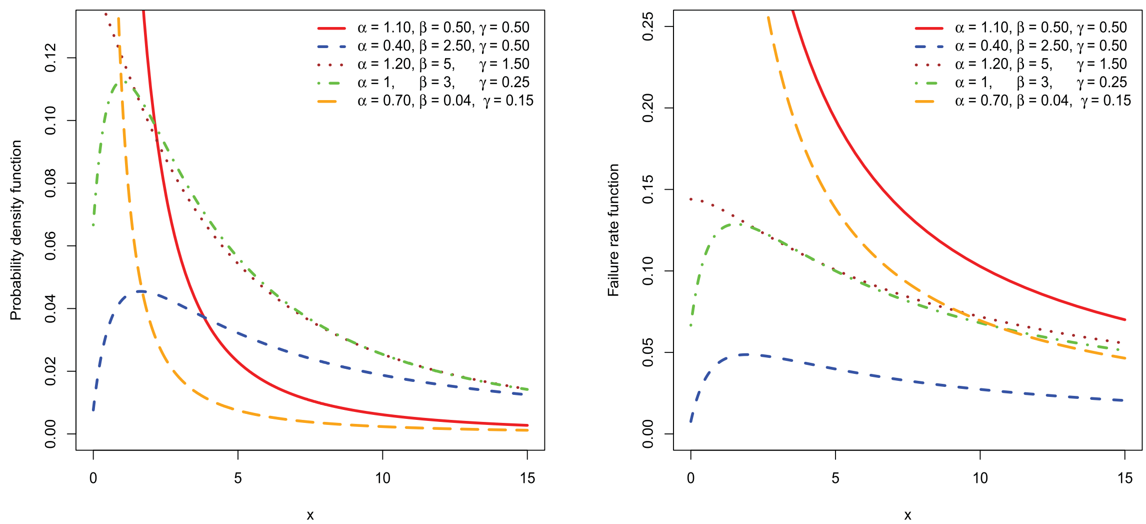

FIGURE 1. **The pdf (left) and FR function (right) of the GPM for some values of the parameters.**

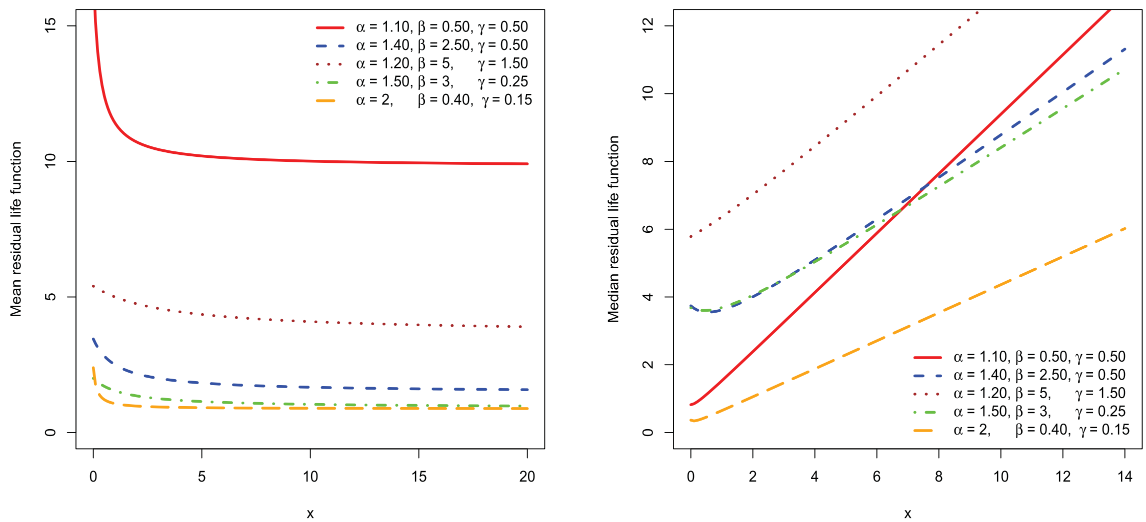

FIGURE 2. **The MRL (left) and median residual life (right) of the GPM for some values of the parameters.**

The elements of this matrix are complicated expressions and should be numerically computed. For the iid random sample X_i, $i = 1, 2, \ldots, n$ from $GPM(\alpha_0, \beta_0, \gamma_0)$, the MLE $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ weakly converges to the multivariate normal $N((\alpha_0, \beta_0, \gamma_0), n^{-1}K^{-1})$, where K^{-1} is the inverse of the information matrix.

B. EM ALGORITHM FOR COMPLETE DATA

Assume that X_i, $i = 1, 2, \ldots, n$ shows an iid random sample of $GPM(\alpha, \beta, \gamma)$. Every X_i arises from the GPM distribution. In the EM algorithm, one indicator latent random variable Z_i is considered, which determines that X_i has been taken from $GP(\alpha, \beta, 1)$ or $GP(\alpha, \beta, 2)$. Precisely, $X_i|Z_i = 1 \sim GP(\alpha, \beta, 1)$, $X_i|Z_i = 2 \sim GP(\alpha, \beta, 2)$ and

$$P(Z_i = j) = \frac{1}{\gamma + 1}\gamma^{I(j=1)}, \quad j = 1, 2. \tag{17}$$

Let $\Theta = (\alpha, \beta, \gamma)$ for a brief representation. The likelihood function can be written in the following form.

$$L(\Theta; \mathbf{x}, \mathbf{z}) = \prod_{i=1}^{n}\prod_{j=1}^{2}\Big[f_j(x_i; \alpha, \beta)P(Z_i = j)\Big]^{I(z_i=j)}, \tag{18}$$

where indicator $I(z_i = j)$ is equal to 1 when $z_i = j$ and otherwise equal to 0, and f_j shows the pdf of $GP(\alpha, \beta, j)$, i.e.,

$$f_j(x_i; \alpha, \beta) = \frac{\alpha}{\beta}(1 + \frac{x_i}{\beta})^{-\alpha-1}\Big[(\alpha + 1)\frac{x_i}{\beta + x_i}\Big]^{I(j=2)}, \quad j = 1, 2. \tag{19}$$

The log-likelihood function is $l(\Theta; \mathbf{x}, \mathbf{z}) = \ln L(\Theta; \mathbf{x}, \mathbf{z})$, which is simplified in (32), see Appendix A.

In the sequel, one specific implementation of the EM algorithm for estimating the model parameters is presented. It involves two steps: Expectation (E) and Maximization (M).

TABLE 1. Simulation results for the MLE and EM estimators of the parameters of the GPM distribution for uncensored data. In every cell, the first, second and third lines are related to α, β and γ, respectively.

n	α, β, γ	MLE B	MLE AB	MLE MSE	EM B	EM AB	EM MSE
80	0.7, 0.5,0.5	-0.03492 0.17192 6.78662	0.11637 0.31603 7.18736	0.02406 0.21845 195.5139	0.02741 -0.00456 -0.15203	0.10378 0.14279 0.25741	0.02051 0.03572 0.08905
	1.2,0.2,0.3	0.19721 0.10909 3.47297	0.38371 0.16394 3.70657	0.94833 0.18364 98.6911	0.08766 0.01572 -0.06096	0.24584 0.07015 0.16820	0.12508 0.01009 0.04073
	1,0.05,0.7	0.08056 0.01862 7.75307	0.22707 0.03515 8.27545	0.10473 0.00297 228.2012	0.03739 -0.00049 -0.20123	0.18307 0.01600 0.32371	0.06338 0.00045 0.14411
	0.8, 2, 0.3	0.05179 0.82140 5.08492	0.15407 1.38307 5.34278	0.05193 5.71420 146.2850	0.03329 0.08495 -0.06897	0.12621 0.63135 0.17288	0.03033 0.83074 0.04298
100	0.7, 0.5,0.5	0.02231 0.12514 5.79579	0.10068 0.26544 6.19919	0.01744 0.15814 168.1027	0.00704 -0.02188 -0.15839	0.09279 0.13403 0.23175	0.01380 0.02979 0.07457
	1.2,0.2,0.3	0.11896 0.07237 3.00987	0.26948 0.12039 3.23867	0.16718 0.04334 85.6446	0.03174 -0.00330 -0.09155	0.19863 0.05532 0.15410	0.06876 0.00505 0.03305
	1,0.05,0.7	0.07070 0.01858 7.49583	0.18715 0.03139 7.96049	0.07003 0.00209 216.9558	-0.02938 -0.00048 -0.20585	0.15915 0.01408 0.30586	0.04929 0.00035 0.12870
	0.8, 2, 0.3	0.04446 0.64949 3.84656	0.12836 1.11213 4.09132	0.03415 3.90560 109.7019	0.02768 0.06441 -0.08576	0.10948 0.51717 0.15784	0.02180 0.53767 0.03439

At iteration t, in the E step, the expectation of the log-likelihood in terms of the current estimate of the conditional latent variable, $Q(\Theta|\Theta_t) = E_{Z|X,\Theta_t}[l(\Theta; \mathbf{x}, \mathbf{Z})]$, is prepared. Then in the M step, the parameters are estimated in the current iteration by maximizing the likelihood expectation $Q(\Theta|\Theta_t)$, prepared in the E step. The iterative process continues until the iteration does not noticeably improve the expectation. The simulation results show that this implementation of the EM algorithm gives a better estimator than the MLE, see Tables 1 and 2.

1) THE E STEP

Suppose that the estimate of the parameters Θ_t is known at iteration t. Then, the conditional distribution of Z_i can be computed by the well-known Bayes theorem as follows. It is simplified in Appendix A, (33). After the simplification, we have

$$p_{i1,t} = \frac{\gamma_t(\beta_t + x_i)}{\gamma_t(\beta_t + x_i) + (\alpha_t + 1)x_i}, \quad i = 1, 2, \ldots, n,$$

and $p_{i2,t} = 1 - p_{i1,t}$. These probabilities are known as membership probabilities at iteration t and are used to construct the expectation function $Q(\Theta|\Theta_t) = E_{Z|X,\Theta_t}[l(\Theta; \mathbf{x}, \mathbf{Z})]$. The simplified expression in Appendix A, (34), shows that the expectation is a sum of two expressions, one of which depends on α and β, and the other depends on γ. We can thus write

$$Q(\Theta|\Theta_t) = Q_1(\alpha, \beta|\Theta_t) + Q_2(\gamma|\Theta_t), \tag{20}$$

where

$$Q_1(\alpha, \beta|\Theta_t) = n \ln \frac{\alpha}{\beta} - \sum_{i=1}^{n} (\alpha + 1) \ln(1 + \frac{x_i}{\beta}) + \sum_{i=1}^{n} p_{i2,t} \big[\ln(\alpha + 1) + \ln \frac{x_i}{\beta + x_i} \big], \tag{21}$$

and

$$Q_2(\gamma|\Theta_t) = \sum_{i=1}^{n} p_{i1,t} \ln \gamma - n \ln(\gamma + 1). \tag{22}$$

TABLE 2. Simulation results for the MLE and EM estimators of the parameters of the GPM distribution when the censorship rate is 20 percent. In every cell, the first, second and third lines are related to α, β and γ, respectively.

n	α, β, γ	MLE B	MLE AB	MLE MSE	EM B	EM AB	EM MSE
80	0.7, 0.5,0.5	0.05682 0.19560 5.95804	0.17180 0.37706 6.38670	0.06061 0.34730 171.9882	-0.00933 -0.01936 -0.15251	0.13277 0.16777 0.23555	0.02971 0.04473 0.07630
	1.2,0.2,0.3	0.78733 0.37793 3.61086	1.03488 0.44331 3.85548	13.47012 2.67611 99.04898	0.10901 0.01741 -0.08889	0.35045 0.08802 0.16944	0.28269 0.01881 0.03970
	1,0.05,0.7	0.25037 0.04462 6.87167	0.48243 0.06573 7.47129	3.84762 0.08483 201.6821	-0.05258 -0.00038 -0.23615	0.27952 0.02206 0.32885	0.17328 0.00108 0.14982
	0.8,2,0.3	0.11562 1.26735 4.62239	0.24517 1.88498 4.88856	0.18100 14.37903 131.7597	0.04134 -0.12257 -0.07956	0.18238 0.78199 0.17278	0.07003 1.62294 0.04225
100	0.7, 0.5,0.5	0.04973 0.20270 5.25818	0.16121 0.36651 6.16985	0.05001 0.35795 165.6812	0.01461 -0.01850 -0.17395	0.12647 0.16854 0.24806	0.02888 0.04617 0.08129
	1.2,0.2,0.3	0.41433 0.21964 3.00607	0.64580 0.28243 3.25485	6.16477 1.61510 84.9875	0.05642 -0.00942 -0.08119	0.29300 0.07685 0.15752	0.17374 0.01292 0.03446
	1,0.05,0.7	0.11682 0.02325 6.25681	0.32925 0.04175 6.80138	0.59276 0.02537 182.5726	-0.00913 -0.00327 -0.25698	0.22640 0.01835 0.32232	0.09748 0.00059 0.14114
	0.8,2,0.3	0.08435 0.89854 3.78973	0.18718 1.40920 4.01250	0.07648 6.67489 103.7488	-0.01617 0.00274 -0.08853	0.15508 0.65027 0.15401	0.04225 0.80846 0.03262

TABLE 3. Number of successive failure for the air conditioning system.

50	130	487	57	102	15	14	10	57	320	261	51	44
9	254	493	33	18	209	41	58	60	48	56	87	11
102	12	5	14	14	29	37	186	29	104	7	4	72
270	283	7	61	100	61	502	220	120	141	22	603	35
98	54	100	11	181	65	49	12	239	14	18	39	3
12	5	32	9	438	43	134	184	20	386	182	71	80
188	230	152	5	36	79	59	33	246	1	79	3	27
201	84	27	156	21	16	88	130	14	118	44	15	42
106	46	230	26	59	153	104	20	206	5	66	34	29
26	35	5	82	31	118	326	12	54	36	34	18	25
120	31	22	18	216	139	67	310	3	46	210	57	76
14	111	97	62	39	30	7	44	11	63	23	22	23
14	18	13	34	16	18	130	90	163	208	1	24	70
16	101	52	208	95	62	11	191	14	71			

2) THE M STEP

In this step, the parameters at iteration $t + 1$ are estimated by maximizing $Q(\Theta|\Theta_t)$ in terms of Θ. Specifically,

$$\Theta_{t+1} = \arg\max_{\Theta} Q(\Theta|\Theta_t),$$

By (20), this problem can be reduced to two separate maximization problems as follows:

$$(\alpha_{t+1}, \beta_{t+1}) = \arg\max_{\alpha,\beta} Q_1(\alpha, \beta|\Theta_t), \tag{23}$$

TABLE 4. Remission times of the bladder cancer.

0.08	0.20	0.40	0.50	0.51	0.81	0.90	1.05	1.19	1.26	1.35	1.40	1.46
1.76	2.02	2.02	2.07	2.09	2.23	2.26	2.46	2.54	2.62	2.64	2.69	2.69
2.83	2.87	3.02	3.25	3.31	3.36	3.36	3.48	3.52	3.57	3.64	3.70	3.82
3.88	4.18	4.23	4.26	4.33	4.34	4.40	4.50	4.51	4.87	4.98	5.06	5.09
5.17	5.32	5.32	5.34	5.41	5.41	5.49	5.62	5.71	5.85	6.25	6.54	6.76
6.93	6.94	6.97	7.09	7.26	7.28	7.32	7.39	7.59	7.62	7.63	7.66	7.87
7.93	8.26	8.37	8.53	8.65	8.66	9.02	9.22	9.47	9.74	10.06	10.34	10.66
10.75	11.25	11.64	11.79	11.98	12.02	12.03	12.07	12.63	13.11	13.29	13.80	14.24
14.76	14.77	14.83	15.96	16.62	17.12	17.14	17.36	18.10	19.13	20.28	21.73	22.69
23.63	25.74	25.82	26.31	32.15	34.26	36.66	43.01	46.12	79.05	2.75		

TABLE 5. Results of fitting the data sets to some models.

Data set	Model	Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	D_n	$p-value$	AIC
Table 3	GPM	MLE	4.8842	341.9168	47.0243	0.0456	0.8505	1965.446
	GPM	EM	4.2001	233.7010	3.2691	0.0446	0.8677	
	GuLx	MLE	21.9694	995.0593	1.9508	0.0481	0.8030	1961.591
	WP	MLE	3.74224	0.23231	1	0.0433	0.8943	1965.856
Table 4	GPM	MLE	4.0242	15.0751	0.13901	0.03605	0.9963	825.2818
	GPM	EM	3.9345	14.50030	0.11655	0.03409	0.9984	
	GuLx	MLE	44.5793	236.7911	1.7973	0.1050	0.1185	841.8991
	WP	MLE	4.9035	0.21297	0.08	0.0450	0.9591	840.0858

and

$$\gamma_{t+1} = \arg\max_{\gamma} Q_2(\gamma|\Theta_t), \tag{24}$$

The maximization problem (23) has no analytical solution and should be numerically solved. However, by solving the equation $\frac{\partial}{\partial\gamma}Q_2(\gamma|\Theta_t) = 0$, we obtain

$$\gamma_{t+1} = \frac{\sum_{i=1}^{n} p_{i1,t}}{\sum_{i=1}^{n} p_{i2,t}}.$$

The iterative process can be terminated if for a predefined small $\epsilon > 0$, $Q(\Theta_{t+1}|\Theta_{t+1}) < Q(\Theta_t|\Theta_t) + \epsilon$.

C. THE MLE FOR RIGHT-CENSORED DATA

Now, suppose that we have a right-censored iid random sample from $GPM(\alpha, \beta, \gamma)$. The random variable X_i is said to be censored from right by a censorship random variable C_i, when $X_i > C_i$, so we only know that the event time is greater than C_i. Thus, the observations are $T_i = \min(X_i, C_i)$ and d_i, where $d_i = 1$; if the event is not censored, $X_i \leq C_i$, and $d_i = 0$; if it is censored, $X_i > C_i$. Let (t_i, d_i), $i = 1, 2, \ldots, n$ show a right-censored sample; then, the log-likelihood function is:

$$l(\alpha, \beta, \gamma; t, d) = \sum_{i=1}^{n} d_i \ln f(t_i) + \sum_{i=1}^{n} (1-d_i)\ln R(t_i).$$

where f and R are the density and reliability functions of the GPM distribution, respectively. The log-likelihood function reduces to

$$\begin{aligned} &l(\alpha, \beta, \gamma; t, d) \\ &= \sum_{i=1}^{n} d_i\Big[\ln\alpha - \ln\beta - \ln(\gamma+1)\Big] \\ &- \sum_{i=1}^{n} d_i(\alpha+1)\ln(1+\frac{t_i}{\beta}) + \ln[\gamma + (\alpha+1)\frac{t_i}{t_i+\beta})] \\ &+ \sum_{i=1}^{n}(1-d_i)\Big[-\ln(1+\gamma) - \alpha\ln(1+\frac{t_i}{\beta}) \\ &+ \ln(\gamma+\alpha+1-\frac{\alpha\beta}{t_i+\beta})\Big]. \end{aligned}$$

D. EM ALGORITHM FOR RIGHT-CENSORED DATA

In the presence of latent variable Z_i, which is defined in the previous sections, the likelihood function for the right-censored data is:

$$\begin{aligned} &L(\Theta; \mathbf{t}, \mathbf{d}, \mathbf{z}) \\ &= \prod_{i=1}^{n}\prod_{j=1}^{2}\Big[f_j(t_i; \alpha, \beta)P(Z_i = j)\Big]^{I(z_i=j)d_i} \\ &\times \prod_{i=1}^{n}\prod_{j=1}^{2}\Big[R_j(t_i; \alpha, \beta)P(Z_i = j)\Big]^{I(z_i=j)(1-d_i)}, \end{aligned} \tag{25}$$

where $P(Z_i = j)$ and f_j are defined by (17) and (19), respectively, and R_j shows the reliability function of f_j. The log-likelihood function $\ln L(\Theta; \mathbf{t}, \mathbf{d}, \mathbf{z})$ has been simplified in Appendix A, (35).

1) THE E STEP

Suppose that the estimate of the parameters at iteration t, Θ_t, is known; then, the conditional distribution of Z_i is equal to $p_{ij,t} = d_iP(Z_i = j|X_i = t_i, \Theta_t) + (1-d_i)P(Z_i = j|X_i > t_i, \Theta_t)$ and can be computed by the well-known Bayes theorem (see Appendix A, (36)).

Specifically, taking $j = 1$,

$$p_{i1,t} = d_i \frac{\gamma_t(\beta_t + t_i)}{\gamma_t(\beta_t + t_i) + (\alpha_t + 1)t_i} + (1 - d_i)\frac{\gamma_t}{\gamma_t + \alpha_t + 1 - \frac{\alpha\beta}{\beta_t + t_i}}, \quad i = 1, 2, \ldots, n, \tag{26}$$

and $p_{i2,t} = 1 - p_{i1,t}$. Then, using (35), the expectation function at iteration t is $Q(\Theta|\Theta_t) = E_{Z|t,d,\Theta_t}[l(\Theta; \mathbf{t}, \mathbf{d}, \mathbf{z})]$. In Appendix A, (37) and (38) show that $Q(\Theta|\Theta_t)$ is a sum of two expressions, one of which solely depends on α and β, and the other depends on γ. In other words,

$$Q(\Theta|\Theta_t) = Q_1(\alpha, \beta|\Theta_t) + Q_2(\gamma|\Theta_t), \tag{27}$$

where

$$\begin{aligned} Q_1(\alpha, \beta|\Theta_t) = &\sum_{i=1}^{n} d_i\Big[\ln\frac{\alpha}{\beta} - (\alpha + 1)\ln(1 + \frac{t_i}{\beta}) \\ &+ p_{i2,t}\ln(\alpha + 1) + p_{i2,t}\ln\frac{t_i}{t_i + \beta}\Big] \\ &+ \sum_{i=1}^{n}(1 - d_i)\Big[-\alpha\ln(1 + \frac{t_i}{\beta}) \\ &+ p_{i2,t}\ln(\alpha + 1 - \frac{\alpha\beta}{t_i + \beta})\Big], \end{aligned} \tag{28}$$

and

$$Q_2(\gamma|\Theta_t) = -n\ln(\gamma + 1) + \ln\gamma\sum_{i=1}^{n} p_{i1,t}, \tag{29}$$

2) THE M STEP

Similar to the uncensored case, the estimation of the parameters at iteration $t + 1$ will be computed by maximizing $Q(\Theta|\Theta_t)$ in terms of Θ. By (27), $Q(\Theta|\Theta_t)$ is a sum of two separate statements, so it follows that

$$(\alpha_{t+1}, \beta_{t+1}) = \arg\max_{\alpha,\beta} Q_1(\alpha, \beta|\Theta_t), \tag{30}$$

and

$$\gamma_{t+1} = \arg\max_{\gamma} Q_2(\gamma|\Theta_t). \tag{31}$$

The maximization problem (30) cannot be analytically solved and should be numerically solved. However, it is easy to check that the solution of (31) is

$$\gamma_{t+1} = \frac{\sum_{i=1}^{n} p_{i1,t}}{\sum_{i=1}^{n} p_{i2,t}}.$$

IV. SIMULATION STUDY

Because the GPM is a mixture model, we can apply the following steps to generate a random sample with size n from it:

1. Generate one random instance of binomial distribution with parameters n and $\frac{\gamma}{\gamma+1}$. Let the generated instance be n_1.
2. Generate one random sample with size n_1 from $GP(\alpha, \beta, 1)$ and one random sample with size $n_2 = n - n_1$ from $GP(\alpha, \beta, 2)$. Then, we combine these two samples to provide one random sample of GPM.
3. To generate a right-censored sample with censorship p, we take the random censorship variable C_i to follow the degenerate distribution with mean t_*. Then, t_* can be computed by solving equation $t_* = q(\bar{p})$, where $q()$ is the quantile function and defined in (10).

The results of a simulation study to estimate the parameters of GPM have been abstracted in Tables 1 and 2 for uncensored data and right-censored data with censorship rate $p = 0.2$, respectively. In every run, $r = 500$ replicates of samples with sizes $n = 80$ and 100 were generated. Then, the parameters were estimated by the maximum likelihood method or EM algorithm. The EM algorithm is an iterative method to find maximum likelihood and can be applied when the model depends on some latent variables, e.g. the mixture or competing risk models. For MLE, we use the built-in ''optim'' function of R with the default optimization method ''Nelder-Mead''. In both MLE and EM approaches, for real parameters α, β and γ, the initial points have been generated from the uniform distribution on the intervals $(0, \alpha)$, $(0, \beta)$ and $(0, \gamma)$ respectively. For real data, we should guess some different initial values for parameters and then compare the results by AIC criterion and Kolmogorov–Smirnov (K-S) statistic. Unfortunately, checking the conditions for terminating the EM process in every EM iteration, makes the simulation runs very slow and time consuming. So, we tested the EM algorithm for many times to find a big constant that is sufficiently large for EM iterations to converge. In this way, we found that 20 iterations is sufficient.

For each parameter, the bias (B), absolute bias (AB) and mean squared error (MSE) were computed. Every cell of these tables shows the A, AB and MSE for α, β and γ from top to bottom. The results related to the maximum likelihood method and EM algorithm are presented on the left and right sides, respectively. The results show smaller values of AB and MSE for the EM algorithm, which indicates that the EM algorithm outperforms the maximum likelihood method.

V. APPLICATIONS

Table 3 shows the number of successive failures for the air conditioning system of 13 Boeing 720 jet airplanes that were analyzed by Tahir *et al.* [25] and many others. We fit the GPM to this data set by the maximum likelihood method and EM algorithm. Additionally, the Gumbel-Lomax (GuLx) and Weibull-Pareto (WP) distributions, which exhibit decreasing and upside-down bathtub-shaped FR functions, were fit to this data set. Table 5 shows the estimates of the parameters, K-S statistic, $D_n = \sup_x\{F_n(x) - F(x)\}$ where $F_n(x)$ is the empirical distribution function, and its corresponding p-value. Akaike information criterion (AIC) is also included in this table. Large values of the K-S statistic and/or the AIC clearly indicate poor fit, so we can compare the fitted models using these measurements. The K-S statistic calculated for

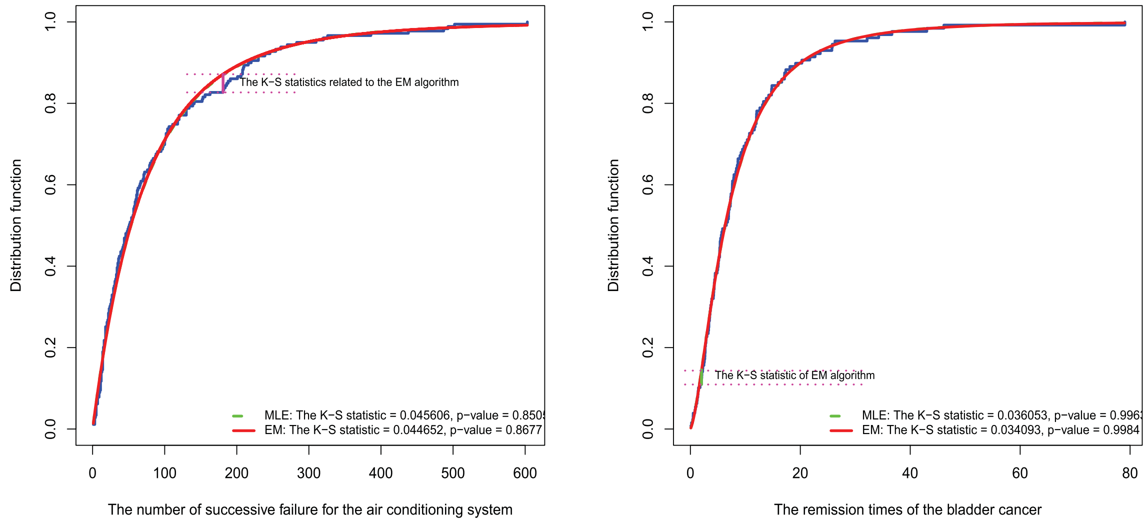

FIGURE 3. **Empirical distribution and fitted GPM distribution for data sets of Table 3 (left) and Table 4 (right).**

GPM (in particular by the EM algorithm) is smaller than others. Based on this criterion, GPM performs better than the others. However, based on the AIC, GuLx shows a better fit. The empirical and fitted CDFs are shown in Figure 3, left panel, visually confirming that the estimated GPM distribution is very close to the empirical distribution function.

Table 4 shows a dataset corresponding to the remission times (in months) of a random sample of 128 bladder cancer patients, referring to Tahir *et al.* [26]. The results of fitting GPM, GuLx and WP to this dataset are also shown in Table 5. The K-S statistics and the AIC for GPM show smaller values than GuLx and WP, indicating that GPM is a good candidate to describe this dataset. Figure 3, right panel, which draws the empirical and fitted CDFs, confirms a very good fit.

VI. CONCLUSION

A new generalized Pareto model was introduced and some of the statistical and reliability properties were investigated. The proposed model can be applied in a variety of real-world situations, such as reliability and survival data where equilibrium is found in the distribution from "small" to "large" values, etc. The parameters of the proposed model were estimated using the MLE and EM algorithms. Simulation results confirm that EM performs better than MLE.

APPENDIX A
SOME DETAILED INFORMATION

The log-likelihood function related to the EM algorithm for complete data can be simplified as follows.

$$
\begin{aligned}
&l(\Theta;\mathbf{x},\mathbf{z})\\
&\quad=\ln L(\Theta;\mathbf{x},\mathbf{z})=\sum_{i=1}^{n}\sum_{j=1}^{2}I(z_i=j)\\
&\qquad\times\Big\{\ln\frac{\alpha}{\beta}-(\alpha+1)\ln(1+\frac{x_i}{\beta})\\
&\qquad+I(j=2)\Big[\ln(\alpha+1)+\ln(\frac{x_i}{\beta+x_i})\Big]\\
&\qquad+I(j=1)\ln\gamma-\ln(\gamma+1)\Big\}\\
&\quad=\sum_{i=1}^{n}I(Z_i=1)\Big[\ln\frac{\alpha}{\beta}-(\alpha+1)\ln(1+\frac{x_i}{\beta})+\ln\gamma\\
&\qquad-\ln(1+\gamma)\Big]+I(Z_i=2)\Big[\ln\frac{\alpha}{\beta}-(\alpha+1)\ln(1+\frac{x_i}{\beta})\\
&\qquad+\ln(\alpha+1)=\ln\frac{x_i}{x_i+\beta}-\ln(1+\gamma)\Big].
\end{aligned}
\tag{32}
$$

Membership probabilities of the EM algorithm for complete data are simplified as follows.

$$
\begin{aligned}
&p_{ij,t}\\
&\quad=P(Z_i=j|X_i=x_i,\Theta_t)\\
&\quad=\frac{f(X_i=x_i|Z_i=j,\Theta_t)P(Z_i=j|\Theta_t)}{f(X_i=x_i|\Theta_t)}\\
&\quad=\frac{\frac{\alpha_t}{\beta_t}(1+\frac{x_i}{\beta_t})^{-\alpha_t-1}\Big[(\alpha_t+1)\frac{x_i}{\beta_t+x_i}\Big]^{I(j=2)}\gamma_t^{I(j=1)}\frac{1}{\gamma_t+1}}{\sum_{j=1}^{2}\frac{\alpha_t}{\beta_t}(1+\frac{x_i}{\beta_t})^{-\alpha_t-1}\Big[(\alpha_t+1)\frac{x_i}{\beta_t+x_i}\Big]^{I(j=2)}\gamma_t^{I(j=1)}\frac{1}{\gamma_t+1}}\\
&\quad=\frac{\gamma_t(\beta_t+x_i)I(j=1)+(\alpha_t+1)x_iI(j=2)}{\gamma_t(\beta_t+x_i)+(\alpha_t+1)x_i},\\
&\qquad i=1,2,\ldots,n,\ j=1,2.
\end{aligned}
\tag{33}
$$

Expectation function of the EM algorithm for complete data can be written as follows.

$$
\begin{aligned}
&Q(\Theta|\Theta_t)\\
&\quad=E_{Z|X,\Theta_t}[l(\Theta;\mathbf{x},\mathbf{Z})]\\
&\quad=\sum_{i=1}^{n}E_{Z_i|X_i,\Theta_t}\sum_{j=1}^{2}I(z_i=j)\Big[\ln\frac{\alpha}{\beta}-(\alpha+1)\ln(1+\frac{x_i}{\beta})\\
&\qquad+I(j=2)(\ln(\alpha+1)+\ln\frac{x_i}{x_i+\beta})+I(j=1)
\end{aligned}
$$

$$\times \ln\gamma - \ln(\gamma+1)\Big] = \sum_{i=1}^{n} P(Z_i = 1|X_i, \Theta_t)$$
$$\times\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)\ln(1+\frac{x_i}{\beta}) - \ln\gamma - \ln(\gamma+1)\Big]$$
$$+P(Z_i = 2|X_i, \Theta_t)\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)\ln(1+\frac{x_i}{\beta})$$
$$+\ln(\alpha+1) + \ln\frac{x_i}{\beta+x_i} - \ln(\gamma+1)\Big]$$
$$= n\ln\frac{\alpha}{\beta} - \sum_{i=1}^{n}(\alpha+1)\ln(1+\frac{x_i}{\beta})$$
$$+\sum_{i=1}^{n} p_{i2,t}\big(\ln(\alpha+1) + \ln\frac{x_i}{\beta+x_i}\big)$$
$$+\sum_{i=1}^{n} p_{i1,t}\ln\gamma - n\ln(\gamma+1). \quad (34)$$

The log-likelihood function related to the EM algorithm for right-censored data has been simplified aa follows.

$$l(\Theta; \mathbf{t}, \mathbf{d}, \mathbf{z})$$
$$= \ln L(\Theta; \mathbf{t}, \mathbf{d}, \mathbf{z})$$
$$= \sum_{i=1}^{n} d_i \sum_{j=1}^{2} I(z_i = j)\ln\Big[f_j(t_i; \alpha, \beta)P(Z_i = j)\Big]$$
$$+\sum_{i=1}^{n}(1-d_i)\sum_{j=1}^{2} I(z_i = j)\ln\Big[R_j(t_i; \alpha, \beta)P(Z_i = j)\Big]$$
$$= \sum_{i=1}^{n} d_i I(Z_i = 1)\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)\ln(1+\frac{t_i}{\beta}) + \ln\gamma$$
$$-\ln(\gamma+1)\Big] + \sum_{i=1}^{n} d_i I(Z_i = 2)\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)$$
$$\times\ln(1+\frac{t_i}{\beta}) + \ln(\alpha+1) + \ln\frac{t_i}{t_i+\beta} - \ln(\gamma+1)\Big]$$
$$+\sum_{i=1}^{n}(1-d_i)I(Z_i = 1)\ln\Big[(1+\frac{t_i}{\beta})^{-\alpha}\frac{\gamma}{\gamma+1}\Big]$$
$$+\sum_{i=1}^{n}(1-d_i)I(Z_i = 2)\ln\Big[(1+\frac{t_i}{\beta})^{-\alpha}$$
$$\times(\alpha+1-\frac{\alpha\beta}{t_i+\beta})\frac{1}{\gamma+1}\Big] \quad (35)$$

Membership probabilities of the EM algorithm for right-censored data.

$$p_{ij,t}$$
$$= d_i P(Z_i = j|X_i = t_i, \Theta_t) + (1-d_i)P(Z_i = j|X_i > t_i, \Theta_t)$$
$$= d_i\frac{f(X_i = t_i|Z_i = j, \Theta_t)P(Z_i = j|\Theta_t)}{f(X_i = t_i|\Theta_t)}$$
$$+(1-d_i)\frac{P(X_i > t_i|Z_i = j, \Theta_t)P(Z_i = j|\Theta_t)}{f(X_i > t_i|\Theta_t)}$$
$$= d_i\frac{\frac{\alpha_t}{\beta_t}(1+\frac{t_i}{\beta_t})^{-\alpha_t-1}\Big[(\alpha_t+1)\frac{t_i}{\beta_t+t_i}\Big]^{I(j=2)}\gamma_t^{I(j=1)}\frac{1}{\gamma_t+1}}{\sum_{j=1}^{2}\frac{\alpha_t}{\beta_t}(1+\frac{t_i}{\beta_t})^{-\alpha_t-1}\Big[(\alpha_t+1)\frac{t_i}{\beta_t+t_i}\Big]^{I(j=2)}\gamma_t^{I(j=1)}\frac{1}{\gamma_t+1}}$$
$$+(1-d_i)$$
$$\times\frac{(1+\frac{t_i}{\beta_t})^{-\alpha_t}\Big[\alpha_t+1-\frac{\alpha\beta}{\beta_t+t_i}\Big]^{I(j=2)}\gamma_t^{I(j=1)}\frac{1}{\gamma_t+1}}{\sum_{j=1}^{2}(1+\frac{t_i}{\beta_t})^{-\alpha_t}\Big[\alpha_t+1-\frac{\alpha\beta}{\beta_t+t_i}\Big]^{I(j=2)}\gamma_t^{I(j=1)}\frac{1}{\gamma_t+1}}$$
$$= d_i\frac{\gamma_t(\beta_t+t_i)I(j=1) + (\alpha_t+1)t_i I(j=2)}{\gamma_t(\beta_t+t_i) + (\alpha_t+1)t_i}$$
$$+(1-d_i)$$
$$\times\frac{\gamma_t I(j=1) + (\alpha_t+1-\frac{\alpha\beta}{\beta_t+t_i})I(j=2)}{\gamma_t+\alpha_t+1-\frac{\alpha\beta}{\beta_t+t_i}},$$
$$i = 1, 2, \ldots, n,\ j = 1, 2. \quad (36)$$

Expectation function of the EM algorithm for right-censored data is equal to

$$Q(\Theta|\Theta_t)$$
$$= E_{Z|t,d,\Theta_t}[l(\Theta; \mathbf{t}, \mathbf{d}, \mathbf{z})]$$
$$= \sum_{i=1}^{n} d_i P(Z_i = 1|\mathbf{t}, \mathbf{d}, \Theta_t)$$
$$\times\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)\ln(1=\frac{t_i}{\beta}) + \ln\gamma - \ln(\gamma+1)\Big]$$
$$+\sum_{i=1}^{n} d_i P(Z_i = 2|\mathbf{t}, \mathbf{d}, \Theta_t)\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)\ln(1+\frac{t_i}{\beta})$$
$$+\ln(\alpha+1) + \ln\frac{t_i}{t_i+\beta} - \ln(\gamma+1)\Big]$$
$$+\sum_{i=1}^{n}(1-d_i)P(Z_i = 1|\mathbf{t}, \mathbf{d}, \Theta_t)$$
$$\times\Big[-\alpha\ln\frac{t_i}{t_i+\beta} + \ln\gamma - \ln(\gamma+1)\Big]$$
$$+\sum_{i=1}^{n}(1-d_i)P(Z_i = 2|\mathbf{t}, \mathbf{d}, \Theta_t)$$
$$\times\Big[-\alpha\ln\frac{t_i}{t_i+\beta} + \ln(\alpha+1-\frac{\alpha\beta}{t_i+\beta}) - \ln(\gamma+1)\Big], \quad (37)$$

After simplification, it reduces to

$$Q(\Theta|\Theta_t)$$
$$= \sum_{i=1}^{n} d_i\Big[\ln\frac{\alpha}{\beta} - (\alpha+1)\ln(1+\frac{t_i}{\beta})$$
$$+p_{i2,t}\ln(\alpha+1) + p_{i2,t}\ln\frac{t_i}{t_i+\beta}\Big]$$
$$+\sum_{i=1}^{n}(1-d_i)\Big[-\alpha\ln(1+\frac{t_i}{\beta}) + p_{i2,t}\ln(\alpha+1-\frac{\alpha\beta}{t_i+\beta})\Big]$$
$$= -n\ln(\gamma+1) + \ln\gamma\sum_{i=1}^{n} p_{i1,t}. \quad (38)$$

ACKNOWLEDGMENT

The author thanks four anonymous reviewers for their useful comments that lead to this improved version of the paper.

REFERENCES

[1] S. M. Burroughs and S. F. Tebbens, "Upper-truncated power law distributions," *Fractals*, vol. 9, no. 2, pp. 209–222, Jun. 2001.

[2] B. Schroeder, S. Damouras, and P. Gill, "Understanding latent sector errors and how to protect against them," *ACM Trans. Storage*, vol. 6, no. 3, pp. 1–23, Sep. 2010.

[3] A. Akinsete, F. Famoye, and C. Lee, "The beta-Pareto distribution," *Statistics*, vol. 42, no. 6, pp. 547–563, Dec. 2008.

[4] M. M. Nassar and N. K. Nada, "The beta generalized Pareto distribution," *J. Statist., Adv. Theory Appl.*, vol. 6, no. 1/2, pp. 1–17, 2011.

[5] E. Mahmoudi, "The beta generalized Pareto distribution with application to lifetime data," *Math. Comput. Simul.*, vol. 81, no. 11, pp. 2414–2430, Jul. 2011.

[6] L. M. Zea, R. B. Silva, M. Bourguignon, A. M. Santos, and G. M. Cordeiro, "The beta exponentiated Pareto distribution with application to bladder cancer susceptibility," *Int. J. Statist. Probab.*, vol. 1, no. 2, pp. 8–19, Oct. 2012.

[7] A. Alzaatreh, F. Famoye, and C. Lee, "Gamma-Pareto distribution and its applications," *J. Mod. Appl. Stat. Methods*, vol. 11, no. 1, pp. 78–94, 2012, doi: 10.22237/jmasm/133584516.

[8] I. Elbatal, "The Kumaraswamy exponentiated Pareto distribution," *Econ. Qual. Control*, vol. 28, no. 1, pp. 1–8, 2013, doi: 10.1515/eqc-2013-0006.

[9] M. Bourguignon, R. B. Silva, L. M. Zea, and G. M. Cordeiro, "The Kumaraswamy Pareto distribution," *J. Stat. Theory Appl.*, vol. 12, no. 2, pp. 129–144, 2013, doi: 10.2991/jsta.2013.12.2.1.

[10] M. Mead, "An extended Pareto distribution," *Pakistan J. Statist. Operation Res.*, vol. 10, pp. 313–329, Oct. 2014.

[11] D. A. Clifton, L. Clifton, S. Hugueny, and L. Tarassenko, "Extending the generalised Pareto distribution for novelty detection in high-dimensional spaces," *J. Signal Process. Syst.*, vol. 74, no. 3, pp. 323–339, Mar. 2014, doi: 10.1007/s11265-013-0835-2.

[12] I. Papastathopoulos and J. A. Tawn, "Extended generalised Pareto models for tail estimation," *J. Stat. Planning Inference*, vol. 143, no. 1, pp. 131–143, Jan. 2013, doi: 10.1016/j.jspi.2012.07.001.

[13] K. Jayakumar, B. Krishnan, and G. G. Hamedani, "On a new generalization of Pareto distribution and its applications," *Commun. Statist. Simul. Comput.*, vol. 49, no. 5, pp. 1264–1284, May 2020.

[14] A. Tahir, A. S. Akhter, and M. A. Haq, "Transmuted new Weibull–Pareto distribution and its applications," *Appl. Appl. Math. Int. J.*, vol. 13, no. 1, pp. 30–46, 2018.

[15] M. Korkmaz, E. Altun, H. Yousof, A. Afify, and S. Nadarajah, "The burr X Pareto distribution: Properties, applications and VaR estimation," *J. Risk Financial Manage.*, vol. 11, no. 1, p. 1, Dec. 2017.

[16] I. Elbatal and G. Aryal, "A new generalization of the exponential Pareto distribution," *J. Inf. Optim. Sci.*, vol. 38, no. 5, pp. 675–697, Jul. 2017.

[17] C. Chananet and W. Phaphan, "On the new weight parameter of the mixture Pareto distribution and its application to real data," *Appl. Sci. Eng. Prog.*, vol. 14, pp. 460–467, Oct. 2020.

[18] J. Beirlant, J. Teugels, and P. Vynckier, *Practical Analysis of Extreme Values*. Leuven, Belgium: Leuven Univ. Press, 1996.

[19] R. Wiborg, "The extended Pareto distribution as default loss model," M.S. thesis, Dept. Math. Natural Sci., Univ. Oslo, Oslo, Norway, 2013.

[20] R. Shanker and A. Mishra, "A quasi Lindley distribution," *Afr. J. Math. Comput. Sci. Res.*, vol. 6, pp. 64–71, Apr. 2013.

[21] H. L. MacGillivray, "Skewness and asymmetry: Measures and orderings," *Ann. Statist.*, vol. 14, no. 3, pp. 994–1011, Sep. 1986.

[22] A. L. Bowley, *Elements of Statistics*. London, U.K.: P.S. King Son, 1901.

[23] J. Moors, "A quantile alternative for kurtosis," *J. Roy. Stat. SoC., D*, vol. 562, no. 37, pp. 25–32, 1988.

[24] L. A. Prendergast and R. G. Staudte, "Quantile versions of the Lorenz curve," *Electron. J. Statist.*, vol. 10, n0. 2, pp. 1896–1926, 2016, doi: 10.1214/16-EJS1154.

[25] M. H. Tahir, M. A. Hussain, G. M. Cordeiro, G. G. Hamedani, M. Mansoor, and M. Zubair, "The Gumbel-Lomax distribution: Properties and applications," *J. Stat. Theory Appl.*, vol. 15, no. 1, pp. 61–79, 2016.

[26] M. H. Tahir, G. M. Cordeiro, A. Alzaatreh, M. Mansoor, and M. Zubair, "A new Weibull–Pareto distribution: Properties and applications," *Commun. Statist. Simul. Comput.*, vol. 45, no. 10, pp. 3548–3567, Nov. 2016, doi: 10.1080/03610918.2014.948190.

MOHAMED KAYID received the B.Sc. degree in mathematics from Minia University, in 1991, the M.Sc. degree in statistics from Al-Azhar University, in 1999, and the Ph.D. degree in statistics from South Valley University, in 2003. He is a Professor of statistics with the Department of Statistics and Operations Research, King Saud University. His research interests include reliability theory, preventive maintenance, stochastic orders, and their applications. He has published in various international journals on these issues. He has also varied research interests in applied statistics. Some of the research methodologies proposed by him and his coauthors have found interesting applications in the context of reliability and engineering. His current research interests can be classified into four broad areas: theory of stochastic ageing and its applications, estimation in restricted parametric space, theory of stochastic orders, ranking and selection problems, and related estimation problems. His work in these areas has received attention among engineers and been referenced in electrical engineering and computer science journals. He is an elected Reviewer of American Mathematical Society.