



**HAL**  
open science

## The Beta Exponentiated Nadarajah-Haghighi Distribution: Theory and Applications

Abdus M Saboor, Muhammad Nauman Khan, Gauss M Cordeiro, Ibrahim M  
Elbatal, Rodrigo R Pescim

► **To cite this version:**

Abdus M Saboor, Muhammad Nauman Khan, Gauss M Cordeiro, Ibrahim M Elbatal, Rodrigo R Pescim. The Beta Exponentiated Nadarajah-Haghighi Distribution: Theory and Applications. *Mathematica Slovaca*, In press. hal-01570564

**HAL Id: hal-01570564**

**<https://hal.science/hal-01570564>**

Submitted on 31 Jul 2017

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

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

# The Beta Exponentiated Nadarajah-Haghighi Distribution: Theory and Applications

Abdus Saboor<sup>a</sup>, Ibrahim Elbatal<sup>b</sup>, Muhammad Nauman Khan<sup>a,\*</sup>, Gauss M. Cordeiro<sup>c</sup>, Rodrigo R. Pescim<sup>d</sup>

<sup>a</sup>Department of Mathematics, Kohat University of Science & Technology, Kohat, Pakistan 26000

<sup>b</sup>Department of Mathematics and Statistics, College of Science Al Imam Mohammad Ibn Saud Islamic University, Saudi Arabia

<sup>c</sup>Departamento de Estatística, Universidade Federal de Pernambuco, 50740-540, Brazil

<sup>d</sup>Departamento de Estatística, Universidade Estadual de Londrina, 86057-970, Brazil

---

## Abstract

We introduce and study the beta exponentiated Nadarajah-Haghighi model, which has increasing, decreasing, upside-down bathtub and bathtub shaped hazard functions. Some of its mathematical properties are determined including a power series for the quantile function. The model parameters are estimated by the method of maximum likelihood. The new distribution and some other competitive models are fitted to a meteorological data set. We prove empirically that it has a superior performance compared with other distributions based on some goodness-of-fit statistics. As a result, some parameters such as return level and mean deviation about the return level are estimated under the best model. We define a new parametric model based on the introduced distribution.

Keywords: Beta distribution, Moment generating function, Maximum likelihood estimation, Monte Carlo simulation, Return level, Wind data.

2000 MSC: 60E05, 62E15, 62E20

---

## 1. Introduction

The exponential distribution was the first lifetime model for which statistical methods were extensively developed in the life testing literature. In many applied sciences such as medicine, engineering and finance, among others, modeling and analyzing lifetime data are crucial. Several lifetime distributions have been used to model these types of data, including the exponential, Weibull, gamma and Rayleigh distributions and their generalizations (see, e.g., [1, 2]). Each distribution has its own characteristics due specifically to the shapes of the hazard rate function (hrf), which can be monotonically decreasing or increasing, bathtub and unimodal.

Nadarajah and Haghighi [3] introduced an extension of the exponential distribution as an alternative to the gamma, Weibull and exponentiated exponential (EE) distributions. The cumulative distribution function (cdf) of the Nadarajah-Haghighi (NH) distribution is given by

$$G(x) = 1 - e^{1-(1+\lambda x)^\alpha}, \quad x > 0, \quad (1)$$

---

<sup>\*</sup>Corresponding author

Email addresses: saboorhangu@gmail.com (Abdus Saboor), iielbatal@imamu.edu.sa (Ibrahim Elbatal), zaybasdf@gmail.com (Muhammad Nauman Khan), gausscordeiro@gmail.com (Gauss M. Cordeiro), rrpescim@gmail.com (Rodrigo R. Pescim)

where  $\lambda > 0$  and  $\alpha > 0$  are scale and shape parameters, respectively. The corresponding probability density function (pdf) and hrf are given by

$$g(x) = \alpha \lambda (1 + \lambda x)^{\alpha-1} e^{1-(1+\lambda x)^\alpha} \quad (2)$$

and  $h(x) = \alpha \lambda (1 + \lambda x)^{\alpha-1}$ .

Equation (2) has two parameters just like the gamma, Weibull and EE distributions. Note also that the NH model has closed-form survival and hrfs just like the Weibull and EE distributions. For  $\alpha = 1$ , it reduces to the exponential distribution. For general properties about the NH model, the reader is referred to [3]. The NH distribution has its mode at zero and allows for increasing, decreasing and constant hrfs.

Lemonte [4] proposed a three-parameter generalization of the NH distribution called the exponentiated NH (ENH) model, whose cdf is given by

$$G(x, \alpha, \lambda, \theta) = \left\{ 1 - e^{1-(1+\lambda x)^\alpha} \right\}^\theta, \quad (3)$$

where the parameters  $\alpha > 0$  and  $\theta > 0$  control the shapes of the distribution and  $\lambda > 0$  is the scale parameter. The ENH density function is given by

$$g(x, \alpha, \lambda, \theta) = \alpha \lambda \theta (1 + \lambda x)^{\alpha-1} e^{1-(1+\lambda x)^\alpha} \left\{ 1 - e^{1-(1+\lambda x)^\alpha} \right\}^{\theta-1}. \quad (4)$$

The beta generalized family pioneered by Eugene *et al.* [6] includes nearly all of well-known models as special or limiting cases such as those exponentiated distributions. It is a rich class of generalized distributions, which allows for greater flexibility of its tails and can be widely applied in many areas such as engineering, biology and medicine, among others. One major benefit of this family is its ability of fitting skewed data that cannot be properly fitted by existing distributions.

Several models have been investigated in the beta family in the last ten years. In fact, this class has been studied in the literature for some special baselines, among them, we cite the beta normal (BN) [6], beta exponential (BE) [2], beta generalized half-normal (BGHN) [9], beta Weibull geometric (BWG) [8], beta generalized exponential (BGE) [10], beta generalized Weibull (BGW) [5], beta gamma (BG) [12] and beta exponentiated Weibull (BEW) [13] distributions.

The cdf of the beta-G class is given by

$$F(x) = \frac{1}{B(a, b)} \int_0^{G(x)} w^{a-1} (1-w)^{b-1} dw = \frac{B(G(x); a, b)}{B(a, b)}, \quad (5)$$

where  $G(x)$  denotes the baseline cdf,  $a > 0$  and  $b > 0$  are two additional parameters,  $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$  is the beta function and  $B(z; a, b) = \int_0^z t^{a-1} (1-t)^{b-1} dt$  is the incomplete beta function. The role of these parameters is to control skewness and vary tail weights of the generated model.

The pdf corresponding to (5) is given by

$$f(x) = \frac{1}{B(a, b)} G(x)^{a-1} \{1 - G(x)\}^{b-1} g(x), \quad (6)$$

where  $g(x) = \frac{d}{dx}G(x)$ . The density  $f(x)$  will be most tractable when  $G(x)$  and  $g(x)$  have simple forms.

The rest of the paper is organized as follows. In Section 2, we define the *beta exponentiated Nadarajah-Haghighi* (BENH) distribution by using the NH model as baseline in the beta family. We show the flexibility of its hrf and present some special models. In Section 3, we study some of its structural properties, which include ordinary and conditional moments, mean deviations and Bonferroni and Lorenz curves. We derive a useful power series for the quantile function (qf) in Section 4. A Monte Carlo simulation study is performed in section 5. In Section 6, we obtain the maximum likelihood estimates (MLEs) of the unknown parameters. In Section 9.1, we provide an application from meteorology data to prove empirically that the new model can yield better fits than some other known lifetime models. Finally, Section 10 offers some conclusions.

## 2. The BENH distribution

In this section, we introduce the five-parameter BENH distribution by taking  $G(x)$  in Equation (3) to be the cdf of the ENH distribution. Let  $\varphi = (\lambda, \theta, \alpha, a, b)$  be the parameters of the proposed model, where  $\lambda > 0, \alpha > 0, \theta > 0, a > 0$  and  $b > 0$ . By using (3) in Equation (5), the cdf of the BENH distribution reduces to

$$F(x; \varphi) = \frac{B\left(\left\{1 - e^{1-(1+\lambda x)^\alpha}\right\}^\theta; a, b\right)}{B(a, b)}. \quad (7)$$

Henceforth, if  $X$  is a random variable with cdf (7), we use the notation  $X \sim BENH_{a,b}(\lambda, \theta, \alpha)$ . The pdf of  $X$  is given by

$$f(x; \varphi) = \frac{\alpha\theta\lambda}{B(a, b)} (1 + \lambda x)^{\alpha-1} e^{1-(1+\lambda x)^\alpha} \left\{1 - e^{1-(1+\lambda x)^\alpha}\right\}^{\theta a-1} \left[1 - \left\{1 - e^{1-(1+\lambda x)^\alpha}\right\}^\theta\right]^{b-1}. \quad (8)$$

Clearly, the sequential double construction **beta class**→**exponentiated class**→**NH model** yields the BENH density function (8). The derivations of some properties of the BENH model can be facilitated by this construction. We eliminate sometimes the dependence of the parameters and write  $F(x) = F(x; \varphi)$ ,  $f(x) = f(x; \varphi)$ , etc. The hazard rate function of  $X$  follow from equations (7) and (8) as  $h(x) = f(x)/[1 - F(x)]$ .

### 2.1. Shapes of the density and hazard functions

The BENH distribution given by (8) is much more flexible than the ENH distribution since it allows for greater flexibility of the tails. In fact, it can approach different distributions when its parameters change. Figures 1 and 2 reveal that the BENH density function can exhibit different behaviors depending on its parameter values. The density function (8) provides more flexible forms than the ENH model and other extended forms of the exponential and NH distributions. Further, Figure 3 and 4 display increasing, decreasing, upside-down bathtub and bathtub shaped forms for the hrf of  $X$ , respectively.

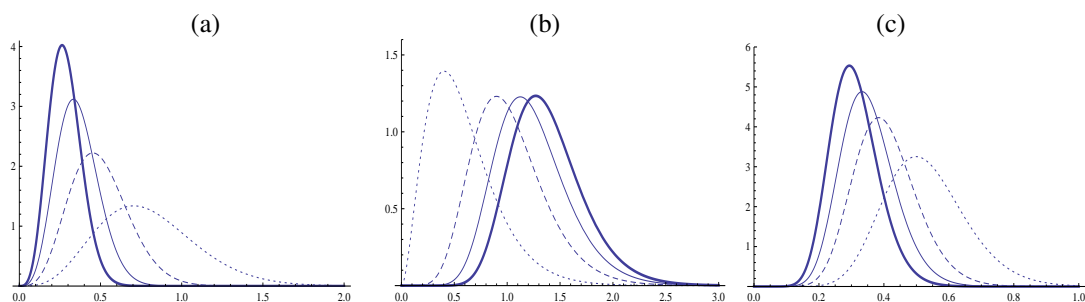


Figure 1: Plots of the BENH density function. (a)  $a = 1.9$ ,  $b = 3.9$ ,  $\lambda = 0.5$ ,  $\theta = 2.3$  and  $\alpha = 2$  (dotted line),  $\alpha = 3$  (dashed line),  $\alpha = 4$  (solid line),  $\alpha = 5$  (thick line). (b)  $a = 2.9$ ,  $b = 1.9$ ,  $\lambda = 1.3$ ,  $\alpha = 1.3$  and  $\theta = 1$  (dotted line),  $\theta = 3$  (dashed line),  $\theta = 5$  (solid line),  $\theta = 7$  (thick line). (c)  $a = 9$ ,  $b = 1.5$ ,  $\alpha = 3$ ,  $\theta = 1.5$  and  $\lambda = 1$  (dotted line),  $\lambda = 1.3$  (dashed line),  $\lambda = 1.5$  (solid line),  $\lambda = 1.7$  (thick line).

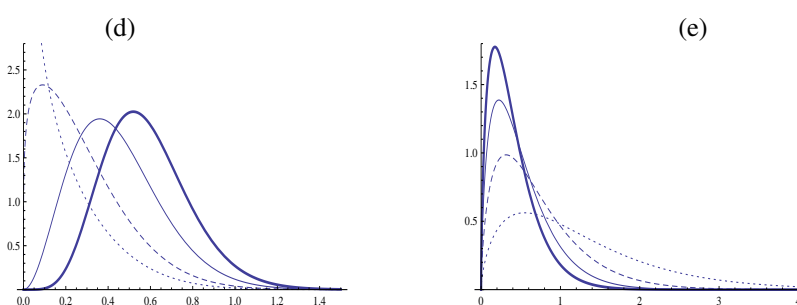


Figure 2: Plots of the BENH density function. (d)  $b = 1.3$ ,  $\lambda = 1$ ,  $\alpha = 2.4$ ,  $\theta = 1.2$  and  $a = 0.5$  (dotted line),  $a = 1$  (dashed line),  $a = 2.5$  (solid line),  $a = 5$  (thick line). (e)  $a = 1.5$ ,  $\lambda = 0.8$ ,  $\alpha = 1.2$ ,  $\theta = 1.1$  and  $b = 1$  (dotted line),  $b = 2$  (dashed line),  $b = 3$  (solid line),  $b = 4$  (thick line).

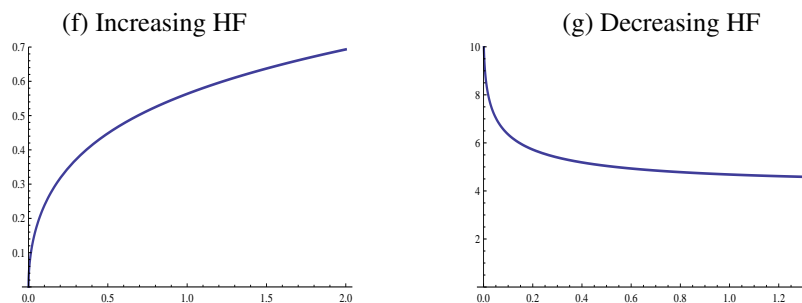


Figure 3: Plots of the BENH hazard function. (f)  $a = 2.5$ ,  $b = 1.5$ ,  $\alpha = 1.4$ ,  $\lambda = 0.3$ ,  $\theta = 0.6$  and (g)  $a = 2$ ,  $b = 4$ ,  $\alpha = 1$ ,  $\lambda = 1.1$ ,  $\theta = 0.5$ .

## 2.2. Special models of the BENH distribution

The  $BENH_{a,b}(\lambda, \theta, \alpha)$  model contains as special cases some well-known distributions. The  $BENH_{1,1}(\lambda, \theta, \alpha)$  model is the ENH [4],  $BENH_{a,b}(\lambda, \theta, 1)$  is the beta generalized exponential (BGE) distribution [10],  $BENH_{a,b}(\lambda, 1, 1)$  is the beta exponential (BE) distribution pioneered by [2],  $BENH_{1,1}(\lambda, 1, \alpha)$  is identical to the NH model [3] and  $BENH_{1,1}(\lambda, \theta, 1)$  coincides with the generalized exponential (GE) distribution [11].

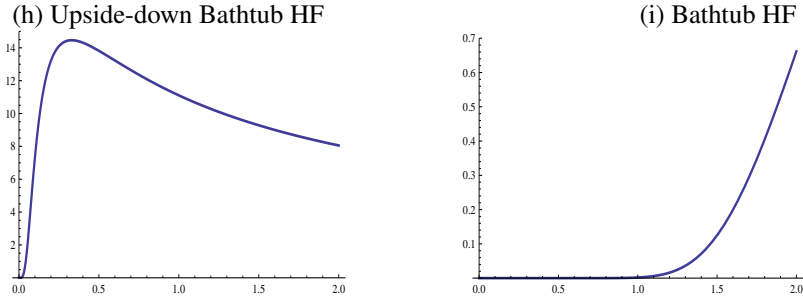


Figure 4: Plots of the BENH hazard function. (h)  $a = 4$ ,  $b = 11$ ,  $\alpha = 0.4$ ,  $\lambda = 15$ ,  $\theta = 1.7$  and (i)  $a = 3.2$ ,  $b = 1.5$ ,  $\alpha = 1.3$ ,  $\lambda = 0.8$ ,  $\theta = 8$ .

### 2.3. Simulation

The BEHN distribution is easily simulated from (7) as follows: if  $v$  is a generated beta variate with shape parameters  $a$  and  $b$ , then

$$x = \lambda^{-1} \left\{ \left[ 1 - \log(1 - v^{1/\theta}) \right]^{1/\alpha} - 1 \right\} \quad (9)$$

is a generated variate having density (8).

### 2.4. Linear representation

Equations (7) and (8) are straightforward to be evaluated using any software with algebraic facilities. In this section, we present a linear representation for the BENH density in terms of EHN densities.

If  $|z| < 1$  and  $b > 0$  is real non-integer, the power series holds

$$(1 - z)^{b-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j!} z^j, \quad (10)$$

where  $\Gamma(\cdot)$  is the gamma function. By using (10), the pdf (8) can be rewritten as

$$f(x; \varphi) = \frac{\alpha \lambda \theta}{B(a, b)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j!} e^{1-(1+\lambda x)^\alpha} (1 + \lambda x)^{\alpha-1} \left\{ 1 - e^{1-(1+\lambda x)^\alpha} \right\}^{\theta(a+j)-1}$$

and then

$$f(x; \varphi) = \sum_{j=0}^{\infty} v_j f_{ENH}(x; \alpha, \lambda, (a+j)\theta), \quad (11)$$

where  $f_{ENH}(x; \alpha, \lambda, (a+j)\theta)$  denotes the ENH density with parameters  $\alpha$ ,  $\lambda$  and  $(a+j)\theta$  and the coefficient  $v_j$  (for  $j \geq 0$ ) is given by

$$v_j = \frac{(-1)^j \Gamma(b)}{(a+j)j! \Gamma(b-j) B(a, b)}.$$

Equation (11) is the main result of this section. It reveals that the BENH density is an infinite mixture of ENH densities with parameters  $\alpha$ ,  $\lambda$  and  $(a+j)\theta$  (for  $j \geq 0$ ) (See [4]).

### 3. Mathematical properties

In this section, we study some structural properties of the BENH distribution, specifically moments, cumulants, mean deviations and Bonferroni and Lorenz curves. Established algebraic expansions to determine some structural properties of this distribution can be more efficient than computing those directly by numerical integration of the density function (8), which can be prone to rounding off errors among others. The formulae derived throughout the paper can be easily handled in softwares such as **Mathematica** and **Maple**.

**Theorem 1.** *The  $r$ th ordinary moment of  $X \sim \text{BENH}_{a,b}(\lambda, \theta, \alpha)$  is given by*

$$\mu'_r = E(X^r) = \frac{\theta}{\lambda^r} \sum_{j,k=0}^{\infty} \sum_{i=0}^r \frac{(-1)^{r+k-i} e^{k+1} v_j}{(k+1)^{\frac{i}{\alpha}+1}} \binom{(a+j)\theta-1}{k} \binom{r}{i} \Gamma\left(\frac{i}{\alpha} + 1, k+1\right). \quad (12)$$

*Proof.* We can write

$$\mu'_r = E(X^r) = \int_0^{\infty} x^r f(x; \varphi) dx = \sum_{j,k=0}^{\infty} (-1)^k e^{k+1} v_j \binom{(a+j)\theta-1}{k} \int_0^{\infty} x^r (1+\lambda x)^{\alpha-1} e^{-(k+1)(1+\lambda x)^\alpha} dx. \quad (13)$$

It follows, since  $r > 0$  is an integer, that

$$\int_0^{\infty} x^r (1+\lambda x)^{\alpha-1} e^{-(k+1)(1+\lambda x)^\alpha} dx = \frac{(-1)^{r-i}}{\alpha \lambda^{r+1} (k+1)^{\frac{i}{\alpha}+1}} \sum_{i=0}^r \binom{r}{i} \Gamma\left(\frac{i}{\alpha} + 1, k+1\right),$$

where  $\Gamma(a, x) = \int_x^{\infty} w^{a-1} e^{-w} dw$  denotes the complementary incomplete gamma function. Then, we have

$$\mu'_r = \frac{\theta}{\lambda^r} \sum_{j,k=0}^{\infty} \sum_{i=0}^r \frac{(-1)^{r+k-i} e^{k+1} v_j}{(k+1)^{\frac{i}{\alpha}+1}} \binom{(a+j)\theta-1}{k} \binom{r}{i} \Gamma\left(\frac{i}{\alpha} + 1, k+1\right), \quad (14)$$

which completes the proof.  $\square$

For lifetime models, it is also of interest to find the conditional moments and the mean residual lifetime function.

**Remark 1.** *The conditional moments of  $X$ , say  $E(X^s | X > t) = \int_t^{\infty} x^s f(x; \varphi) dx$ , are given by*

$$\begin{aligned} E(X^s | X > t) &= \alpha \lambda \theta \sum_{j,k=0}^{\infty} (-1)^k e^{k+1} v_j \binom{(a+j)\theta-1}{k} \int_t^{\infty} x^s (1+\lambda x)^{\alpha-1} e^{-(k+1)(1+\lambda x)^\alpha} dx \\ &= \lambda^{-s} \theta \sum_{j,k=0}^{\infty} \sum_{i=0}^s \frac{(-1)^{s+k-i} e^{k+1} v_j}{(k+1)^{\frac{i}{\alpha}+1}} \binom{(a+j)\theta-1}{k} \binom{s}{i} \Gamma\left(\frac{i}{\alpha} + 1, (k+1)(1+\lambda t)^\alpha\right). \end{aligned}$$

The mean residual lifetime function  $\mu(t) = E(X | X > t) - t$  is given by

$$\mu(t) = \frac{\theta}{\lambda} \sum_{j,k=0}^{\infty} \sum_{i=0}^1 (-1)^{1+k-i} e^{k+1} v_j \binom{(a+j)\theta-1}{k} \binom{1}{i} \Gamma\left(\frac{i}{\alpha} + 1, (k+1)(1+\lambda t)^\alpha\right) - t.$$

Generally, there has been a great interest in obtaining the first incomplete moment of a distribution. Based on this quantity, we can obtain, for example, the mean deviations that provide important information about characteristics of a population. Indeed, the amount of dispersion in a population may be measured to some extent

by all the deviations from the mean and median. One way to measure the amount of scatter in a population is the totality of deviations about the mean and the median. The mean deviations of  $X$  about the mean  $\mu'_1 = E(X)$  and about the median  $M$  can be expressed as  $\delta_1 = 2\mu'_1 F(\mu'_1) - 2m_1(\mu'_1)$  and  $\delta_2 = \mu - 2m_1(M)$ , respectively, where  $F(\mu'_1)$  is easily evaluated from (7), the median  $M$  follows from (7) as the solution of  $F(M; \varphi) = 0.5$  and  $m_1(z) = \int_0^z x f(x) dx$ . We can write

$$\begin{aligned} m_1(z) &= \alpha \lambda \theta \sum_{j,k=0}^{\infty} (-1)^k e^{k+1} v_j \binom{(a+j)\theta - 1}{k} \int_0^z x (1 + \lambda x)^{\alpha-1} e^{-(k+1)(1+\lambda x)^\alpha} dx \\ &= \frac{\theta}{\lambda} \sum_{j,k=0}^{\infty} \sum_{i=0}^1 (-1)^{1+k-i} e^{k+1} v_j \binom{(a+j)\theta - 1}{k} \binom{1}{i} \Gamma\left(\frac{i}{\alpha} + 1, k + 1\right) \left\{1 - (1 + \lambda z)^\alpha\right\}. \end{aligned} \quad (15)$$

#### 4. Quantile function

In this section, we derive a power series for the qf of the BENH distribution obtained by inverting (7) as  $Q(u) = F^{-1}(u)$ . First, we can expand (9) using **Mathematica** as

$$x = \sum_{i=1}^{\infty} s_i v^{i/\theta}, \quad (16)$$

where  $s_1 = 1/(\lambda\alpha)$ ,  $s_2 = 1/(2\lambda\alpha^2)$ ,  $s_3 = (\alpha^2 + 1)/(6\lambda\alpha^3)$ ,  $s_4 = (\alpha^3 + 4\alpha^2 + 1)/(24\lambda\alpha^4)$ , etc.

An expansion for the inverse of the incomplete beta function  $B(v; a, b) = u$  can be found in Wolfram website<sup>1</sup> as

$$v = B^{-1}(u; a, b) = \sum_{j=1}^{\infty} q_j y^j, \quad (17)$$

where  $y = y(u) = [aB(a, b)u]^{1/a}$  and  $q_1 = 1$ ,  $q_2 = (b-1)/(a+1)$ , etc.

Combining (16) and (17), the BENH qf can be expressed as

$$x = Q(u) = \sum_{i=1}^{\infty} s_i \left( \sum_{j=1}^{\infty} t_j u^{j/a} \right)^{i/\theta}, \quad (18)$$

where  $t_j = q_j a^j B(a, b)^j$  and the  $q'_j$ s are given above.

By expanding  $\left(\sum_{j=1}^{\infty} t_j u^{j/a}\right)^{i/\theta}$  in Taylor series, we can rewrite (16) as

$$\left( \sum_{j=1}^{\infty} t_j u^{j/a} \right)^{i/\theta} = \sum_{k=0}^{\infty} f_k(i\theta^{-1}) \left( \sum_{j=1}^{\infty} t_j u^{j/a} \right)^k, \quad (19)$$

where  $f_k(p) = \sum_{m=k}^{\infty} (-1)^{m-k} \binom{m}{k} (p)_m / m!$  (for  $k \geq 0$ ) and  $(p)_m = p(p-1)\dots(p-m+1)$  is the descending factorial.

<sup>1</sup><http://functions.wolfram.com/06.23.06.0004.01>.



We now use an equation of Gradshteyn and Ryzhik [7] for a power series raised to a positive integer  $k$

$$\left( \sum_{j=0}^{\infty} a_j z^j \right)^k = \sum_{j=0}^{\infty} c_{k,j} z^j, \quad (20)$$

where the coefficients  $c_{k,j}$  (for  $k, j = 1, 2, \dots$ ) are obtained from the recurrence equation

$$c_{k,j} = (k a_0)^{-1} \sum_{m=1}^k [m(k+1) - j] a_m c_{k,j-m} \quad (21)$$

and, for  $k \geq 0$ ,  $c_{k,0} = a_0^k$ . The coefficient  $c_{k,j}$  can be determined from  $c_{k,0}, \dots, c_{k,j-1}$  and then from the quantities  $a_0, \dots, a_k$  listed above.

We can write using (20) and (21)

$$\left( u^{1/a} \sum_{j=0}^{\infty} p_j u^{j/a} \right)^k = u^{k/a} \sum_{j=0}^{\infty} d_{k,j} u^{j/a}$$

where  $p_j = t_{j+1}$  for  $j \geq 0$ , and, for  $k \geq 0$ ,  $d_{k,0} = p_0^k$  and, for  $j \geq 1$ ,  $d_{k,j} = (k p_0)^{-1} \sum_{m=1}^k [m(k+1) - j] p_m d_{k,j-m}$ .

Inserting the last equation in (19), we obtain

$$\left( \sum_{j=1}^{\infty} t_j u^{j/a} \right)^{i/\theta} = \sum_{k,j=0}^{\infty} f_k(i\theta^{-1}) d_{k,j} u^{(j+k)/a}$$

and then  $Q(u)$  in (16) can be expressed as

$$Q(u) = \sum_{k,j=0}^{\infty} e_{k,j} u^{(j+k)/a},$$

where  $e_{k,j} = \sum_{i=1}^{\infty} s_i f_k(i\theta^{-1}) d_{k,j}$  (for  $j, k \geq 0$ ).

Finally, we can write

$$Q(u) = \sum_{r=0}^{\infty} h_r u^{r/a}, \quad (22)$$

where  $h_r = \sum_{\substack{k,j=0 \\ k+j=r}}^{\infty} e_{k,j}$ .

Equation (22) reveals that the qf of the BEHN distribution is just a power series. Then, several mathematical quantities of  $X$  can be expressed in terms of integrals over  $(0, 1)$ . In practical terms, the upper limit  $\infty$  in the above sums can be substituted by a number like twenty or thirty. In fact, if  $W(\cdot)$  be any integrable function in the real line, we can write

$$\int_0^{\infty} W(x) f(x) dx = \int_0^1 W \left( \sum_{r=0}^{\infty} h_r u^{r/a} \right) du. \quad (23)$$

Equations (22) and (23) are the main results of this section since we can obtain from them numerically various BENH structural quantities. They can be determined by using the integral on the right-hand side for special  $W(\cdot)$  functions, which can be simpler than if they are evaluated from the left-hand integral.

## 5. A simulation study

We perform a Monte Carlo simulation study to assess the finite sample behavior of the MLEs of  $\lambda$ ,  $\theta$ ,  $\alpha$ ,  $a$  and  $b$ . The results were obtained from 2,000 Monte Carlo replications and the simulations were carried out using the R statistical software. In each replication, a random sample of size  $n$  is drawn from the  $BENH_{a,b}(\lambda, \theta, \alpha)$  distribution and the parameters were estimated by the maximum likelihood method. The random variable  $X$  was generated using the inversion method. We considered two setups with the following values for the parameters of the model:  $\lambda = 1.5$ ,  $\theta = 10.0$ ,  $\alpha = 2.0$ ,  $a = 4.0$ , and  $b = 2.0$  (setup 1) and  $\lambda = 3.0$ ,  $\theta = 5.0$ ,  $\alpha = 1.5$ ,  $a = 2.5$ , and  $b = 3.0$  (setup 2). The mean estimates of the five model parameters and the corresponding root mean squared errors (RMSEs) for the sample sizes  $n = 50$ , 100 and 200 are listed in Tables 1 and 2, respectively. In both setups, we note that the biases and RMSEs of the MLEs of  $\lambda$ ,  $\theta$ ,  $\alpha$ ,  $a$  and  $b$  decay toward zero when the sample size increases, as expected. There is a small sample bias in the estimation of the model parameters. Future research should be conducted to obtain bias corrections for these estimators.

Table 1: Monte Carlo simulation results for Setup 1: Mean estimates and RMSEs of  $\lambda$ ,  $\theta$ ,  $\alpha$ ,  $a$  and  $b$ .

$n$	Parameter	Mean	RMSE
50	$\lambda$	1.1910	1.2007
	$\theta$	13.3681	7.8019
	$\alpha$	3.6020	2.7661
	$a$	7.6681	8.0093
	$b$	3.9407	5.1279
100	$\lambda$	1.3289	1.0321
	$\theta$	11.6246	5.0460
	$\alpha$	2.8785	1.7866
	$a$	5.9771	5.3011
	$b$	3.1445	3.3964
200	$\lambda$	1.4131	0.7452
	$\theta$	10.7861	3.5364
	$\alpha$	2.4810	1.2138
	$a$	4.8404	3.0189
	$b$	2.5862	2.2210

Table 2: Monte Carlo simulation results for Setup 2: Mean estimates and RMSEs of  $\lambda$ ,  $\theta$ ,  $\alpha$ ,  $a$  and  $b$ .

$n$	Parameter	Mean	RMSE
50	$\lambda$	2.7536	2.5369
	$\theta$	8.8775	10.0651
	$\alpha$	2.6233	1.9267
	$a$	4.6692	4.0526
	$b$	4.8437	5.1057
100	$\lambda$	2.9982	2.2718
	$\theta$	7.4226	7.5222
	$\alpha$	2.2102	1.3976
	$a$	3.9831	3.0587
	$b$	4.0726	3.5745
200	$\lambda$	2.9956	2.0220
	$\theta$	6.8338	6.2232
	$\alpha$	2.0991	1.2690
	$a$	3.7933	2.7461
	$b$	3.6886	2.8409

## 6. Estimation and inference

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The MLEs enjoy desirable properties and can be used when constructing confidence intervals for the parameters and also in test statistics. Given the observed values  $x_1, \dots, x_n$ , the MLEs of the model parameters of the BENH distribution are determined by maximization the log-likelihood function given by

$$\begin{aligned} \ell(\varphi) = \ell(x_i, \alpha, \lambda, \theta, a, b) = & n [\log \alpha + \log \theta + \log \lambda - \log \{B(a, b)\}] + (\alpha - 1) \sum_{i=1}^n \log(1 + \lambda x_i) \\ & + \sum_{i=1}^n \{1 - (1 + \lambda x_i)^\alpha\} + (\theta a - 1) \sum_{i=1}^n \log \left\{1 - e^{1-(1+\lambda x_i)^\alpha}\right\} + (b - 1) \sum_{i=1}^n \log \left[1 - \left\{1 - e^{1-(1+\lambda x_i)^\alpha}\right\}^\theta\right]. \end{aligned} \quad (24)$$

The above log-likelihood can be maximized numerically using by using the R (`optim` function), SAS (`PROC NLMIXED`), Ox program (sub-routine `MaxBFGS`), `Nmaximize` command in **Mathematica**, among others.

To check the goodness of fit of all statistical models, some goodness-of-fit statistics shall be used are computed using computational package **Mathematica**. The MLEs are evaluated using `Nmaximize` as well as the goodness-of-fit statistics including the Anderson-Darling ( $A^*$ ), Cramér-von Mises ( $W^*$ ), Liao-Shimokawa (L-S) and Kolmogrov-Simnorov (K-S) statistics (with their p-values). These statistics are used as goodness-of-fit measures and to evaluate how closely a specific density fits the corresponding histogram of a given data set. Usually, the distribution with smallest statistic values will yield a better fit. These statistics are given by

$$\begin{aligned} A^* = \left(\frac{2.25}{n^2} + \frac{0.75}{n} + 1\right) \left[-n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \log(z_i(1 - z_{n-i+1}))\right], \quad L - S = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\text{Max}\left[\frac{i}{n} - z_i, z_i - \frac{i-1}{n}\right]}{\sqrt{z_i(1 - z_i)}}, \\ W^* = \left(\frac{0.5}{n} + 1\right) \left[\sum_{i=1}^n \left(z_i - \frac{2i-1}{2n}\right)^2 + \frac{1}{12n}\right], \quad K - S = \text{Max}\left[\frac{i}{n} - z_i, z_i - \frac{i-1}{n}\right], \end{aligned}$$

where  $z_i = \text{cdf}(y_i)$  and the  $y_i$ 's are the *ordered observations*.

## 7. The log-beta exponentiated Nadarajah-Haghighi distribution

Let  $X$  be a random variable having the BENH density function (8). The random variable  $Y = \log(X)$  defines the *log-beta exponentiated Nadarajah-Haghighi* (LBENH) distribution. The density function of  $Y$  reparameterized in terms of  $\lambda = e^{-\mu}$  can be expressed as

$$f(y) = \frac{\alpha \theta}{B(a, b)} e^{y-\mu} (1 + e^{y-\mu})^{\alpha-1} e^{1-(1+e^{y-\mu})^\alpha} \left\{1 - e^{1-(1+e^{y-\mu})^\alpha}\right\}^{\theta a-1} \left[1 - \left\{1 - e^{1-(1+e^{y-\mu})^\alpha}\right\}^\theta\right]^{b-1}, \quad (25)$$

where  $y, \mu \in \mathbb{R}$ ,  $\theta > 0$ ,  $\alpha > 0$ ,  $a > 0$  and  $b > 0$ .

If  $Y$  is a random variable having density function (25), we can write  $Y \sim \text{LBENH}(\mu, \theta, \alpha, a, b)$ . Thus, if  $X \sim \text{BENH}(\lambda, \theta, \alpha, a, b)$  then  $Y = \log(X) \sim \text{LBENH}(\mu, \theta, \alpha, a, b)$ . The LBENH distribution contains well-known distributions as special models. It simplifies to the log-exponentiated Nadarajah-Haghighi (LENH)

distribution when  $a = b = 1$ . If  $a = b = \theta = 1$ , it reduces to the log-Nadarajah-Haghighi (LNH) distribution. Further, if  $\theta = 1$ , it becomes the log-beta exponential (LBE) distribution. Plots of the density function (25) for selected parameter values are displayed in Figure 5. These plots indicate that the density function (25) is very flexible for different values of the shape parameters  $\theta$ ,  $a$  and  $b$  and then can be used in many practical situations.

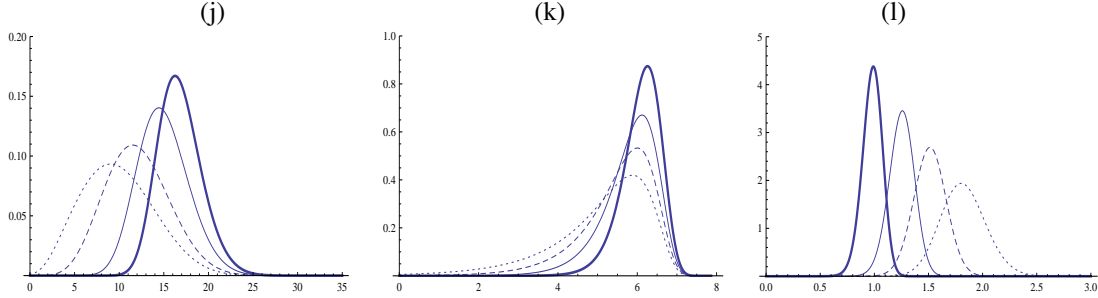


Figure 5: Plots of the LBENH density function. (j)  $\mu = 0.5$ ,  $\alpha = 0.1$ ,  $a = 0.5$ ,  $b = 0.9$  and  $\theta = 5$  (dotted line),  $\theta = 10$  (dashed line),  $\theta = 30$  (solid line),  $\theta = 70$  (thick line). (k)  $\mu = 7$ ,  $\theta = 0.8$ ,  $\alpha = 3$ ,  $b = 0.8$  and  $a = 1$  (dotted line),  $a = 1.5$  (dashed line),  $a = 2.3$  (solid line),  $a = 4$  (thick line). (l)  $\mu = 0.8$ ,  $\theta = 11$ ,  $\alpha = 1.2$ ,  $a = 5$  and  $b = 1.5$  (dotted line)  $b = 5$ , (dashed line),  $b = 15$  (solid line),  $b = 55$  (thick line).

The survival function corresponding to (25) is

$$S(y) = 1 - \frac{B\left[\left\{1 - e^{-(1+e^{y-\mu})^\alpha}\right\}^\theta; a, b\right]}{B(a, b)}.$$

We define the standardized random variable  $Z = (Y - \mu)$  having density function given by

$$f(z) = \frac{\alpha \theta}{B(a, b)} e^z (1 + e^z)^{\alpha-1} e^{1-(1+e^z)^\alpha} \left\{1 - e^{1-(1+e^z)^\alpha}\right\}^{\theta a-1} \left[1 - \left\{1 - e^{1-(1+e^z)^\alpha}\right\}^\theta\right]^{b-1}, \quad z \in \mathbb{R} \quad (26)$$

## 8. The LBENH regression model for censored data

A parametric regression model that provides a good fit to lifetime data tends to yield more precise estimates of the quantities of interest. Next, we construct a linear regression model for the response variable  $y_i$  and the explanatory variable vector  $\mathbf{v}_i^T = (v_{i1}, \dots, v_{ip})$  based on the LBENH distribution given by

$$y_i = \mathbf{v}_i^T \boldsymbol{\beta} + z_i, \quad i = 1, \dots, n. \quad (27)$$

Here, the random error  $z_i$  is assumed to have the density function (26) with unknown parameters  $\mu_i \in \mathbb{R}$ ,  $\alpha > 0$ ,  $\theta > 0$ ,  $a > 0$  and  $b > 0$ . The parameter  $\mu_i = \mathbf{v}_i^T \boldsymbol{\beta}$  is the location of  $y_i$ . The vector  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$  of location parameters is represented by a linear model  $\boldsymbol{\mu} = \mathbf{V}\boldsymbol{\beta}$ , where  $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)^T$  is a known model matrix. The LBENH model (27) opens new possibilities for fitting many different types of data.

Consider a sample  $(y_1, \mathbf{v}_1), \dots, (y_n, \mathbf{v}_n)$  of  $n$  independent observations, where the random response is defined by  $y_i = \min\{\log(x_i), \log(c_i)\}$ . We assume non-informative censoring such that the observed lifetimes and

censoring times are independent. Let  $F$  and  $C$  be the sets of individuals for which  $y_i$  is the log-lifetime and log-censoring, respectively. The log-likelihood function for the vector of parameters  $\boldsymbol{\eta} = (\boldsymbol{\beta}^T, \theta, \alpha, a, b)^T$  from model (27) is given by  $l(\boldsymbol{\eta}) = \sum_{i \in F} \log[f(y_i)] + \sum_{i \in C} \log[S(y_i)]$ , where  $f(y_i)$  is the density function (25) and  $S(y_i)$  is the survival function (26) of  $Y_i$ . The log-likelihood function for  $\boldsymbol{\eta}$  reduces to

$$\begin{aligned}
l(\boldsymbol{\eta}) = & r \left\{ 1 + \log \left( \frac{\alpha \theta}{\sigma B(a, b)} \right) \right\} + \sum_{i \in F} y_i - \mathbf{v}_i^T \boldsymbol{\beta} + (\alpha - 1) \sum_{i \in F} \log \left( 1 + e^{y_i - \mathbf{v}_i^T \boldsymbol{\beta}} \right) \\
& - \sum_{i \in F} \left( 1 + e^{y_i - \mathbf{v}_i^T \boldsymbol{\beta}} \right) + (\theta a - 1) \sum_{i \in F} \log \left\{ 1 - e^{-\left( 1 + e^{y_i - \mathbf{v}_i^T \boldsymbol{\beta}} \right)^\alpha} \right\} + (b - 1) \sum_{i \in F} \log \left[ 1 - \left\{ 1 - e^{-\left( 1 + e^{y_i - \mathbf{v}_i^T \boldsymbol{\beta}} \right)^\alpha} \right\}^\theta \right] \\
& + \sum_{i \in C} \log \left\{ 1 - \frac{B \left[ \left\{ 1 - e^{-\left( 1 + e^{y_i - \mathbf{v}_i^T \boldsymbol{\beta}} \right)^\alpha} \right\}^\theta ; a, b \right]}{B(a, b)} \right\}, \tag{28}
\end{aligned}$$

where  $r$  is the number of uncensored observations (failures).

The MLE  $\hat{\boldsymbol{\eta}}$  of  $\boldsymbol{\eta}$  can be obtained by maximizing the log-likelihood function (28). We use the statistical R software to compute the estimates. From the fitted model (27), the survival function for  $y_i$  can be estimated by

$$S(y_i; \hat{\boldsymbol{\beta}}^T, \hat{\alpha}, \hat{\theta}, \hat{a}, \hat{b}) = 1 - \frac{B \left[ \left\{ 1 - e^{-\left( 1 + e^{y_i - \mathbf{v}_i^T \hat{\boldsymbol{\beta}}^T} \right)^{\hat{\alpha}}} \right\}^{\hat{\theta}} ; \hat{a}, \hat{b} \right]}{B(\hat{a}, \hat{b})}. \tag{29}$$

Under general regularity conditions, the asymptotic distribution of  $(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})$  can be approximated by the multivariate normal  $N_{p+4}(0, J(\boldsymbol{\theta})^{-1})$ , where  $J(\boldsymbol{\theta})$  is the observed information matrix. Then, the inference on the parameter vector  $\boldsymbol{\eta}$  can be based on this approximation distribution for  $\hat{\boldsymbol{\eta}}$ . The likelihood ratio (LR) statistic can be used to discriminate between the LBENH and LENH regression models since they are nested models.

## 9. Applications

In this section, we provide two applications to show the usefulness of the BENH distribution and the new LBENH regression model.

### 9.1. Meteorological application

We fit some well-known statistical distributions and the BENH distribution to a meteorology data set and determine the best fitted distribution among the fitted models. We also provide some meteorological parameters for these data under the best model.

#### 9.1.1. Model fitting

We consider the following models as competitive for the proposed distribution:

- The Gumbel model with density function

$$f(x) = \frac{e^{-e^{-\frac{x-\mu}{\sigma}} - \frac{x-\mu}{\sigma}}}{\sigma}, \quad x, \mu \in R, \sigma > 0,$$

- The gamma Weibull (GW) [14] model with density function

$$f(x) = \frac{k \lambda^{-k-\xi} x^{\xi+k-1} e^{-\lambda^{-k} x^k}}{\Gamma(1 + \xi/k)}, \quad x > 0, \xi + k, \lambda > 0,$$

- The BE [2] model with density function

$$f(x) = \frac{\lambda e^{-(b\lambda x)} (1 - e^{-\lambda x})^{a-1}}{B(a, b)}, \quad x > 0, a, b, \lambda > 0,$$

- The ENH [4] model with density function

$$f(x) = \alpha \lambda \theta (1 + \lambda x)^{\alpha-1} e^{1-(1+\lambda x)\theta} \left\{ 1 - e^{1-(1+\lambda x)\theta} \right\}^{\theta-1}, \quad x > 0, \alpha, \theta, \lambda > 0,$$

- The BG [12] model with density function

$$f(x) = \frac{x^{\rho-1} e^{-x/\lambda}}{\lambda^\rho B(\alpha, \beta) \Gamma(\rho)} \left\{ \frac{\gamma(\rho, x/\lambda)}{\Gamma(\rho)} \right\}^{\alpha-1} \left\{ 1 - \frac{\gamma(\rho, x/\lambda)}{\Gamma(\rho)} \right\}^{\beta-1}, \quad x > 0, \alpha, \beta, \rho, \lambda > 0,$$

- The BEW [13] model with density function

$$f(x) = \frac{\alpha c \lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{\alpha a-1} (1 - (1 - e^{-(\lambda x)^c})^\alpha)^{b-1}}{B(a, b)}, \quad x > 0, a, b, c, \lambda, \beta > 0.$$

The application refers to the “yearly maxima wind” data, which represents the yearly maxima wind speed (miles per hour) registered at a certain location over a period of 50 years as reported in Table 1.1 by [15]. These data may help in the design of structures and buildings.

The maximum likelihood method is used to obtain the parameter estimates and their standard errors (computed by inverting the observed information matrix) given in Table 3 for the current data. Further, all mentioned goodness-of-fit statistics are determined for each fitted distribution and listed in Table 4.

Based on the figures in Table 4, we note that the BENH distribution has the smallest statistics. Therefore, we can conclude that this distribution gives the best fit to these data among the other distributions. Also, this fact is confirmed by observing the plots of the estimated cdfs of these distributions and the empirical cdf given in the Appendix. In fact, the estimated BENH cdf and the empirical cdf are quite close. Further, the LR statistic can be used to compare a distribution having extra parameters with some of its sub-models. Then, we apply the LR statistic for comparing the BENH and ENH distributions. The LR statistic  $w$  for testing the following hypotheses:  $H_0: a = b = 1$  (ENH) versus  $H_1: a \neq b \neq 1$  (BENH) yields the value  $w = 12.3273$  (p-value=0.0021). Thus, the BENH model provides a better fit to the yearly maxima wind data when compared with the ENH model.

The 95% and 99% confidence intervals for these parameters are listed in Table 5.

Table 3: MLEs of the parameters (standard errors in parentheses) for yearly maxima wind data

Distributions	Estimates				
Gumbel( $\mu, \sigma$ )	29.448375 (0.986481)	6.728401 (0.811886)			
GW( $\xi, k, \lambda$ )	16.37274 (3.81043)	0.636736 (0.102927)	0.191498 (0.250299)		
BE( $\lambda, a, b$ )	0.174056 (0.033464)	99.999996 (63.3027)	0.667858 (0.214419)		
ENH( $\alpha, \lambda, \theta$ )	1.090938 (0.230356)	0.098860 (0.0479538)	29.99999 (14.0858)		
BG( $\alpha, \beta, \rho, \lambda$ )	82.54150 (89.3679)	75.99302 (98.5805)	0.285808 (0.280189)	595.603 (873.244)	
BEW( $\alpha, \lambda, a, b, c$ )	89.96635 (98.776)	0.299499 (0.358327)	14.99929 (16.1477)	0.365674 (0.351981)	0.992785 (0.57515)
BENH( $\lambda, \theta, \alpha, a, b$ )	0.440040 (0.24795)	41.48587 (68.1182)	0.899999 (0.197686)	88.218035 (144.411)	0.364810 (0.194824)

Table 4: Goodness-of-fit statistics for yearly maxima wind data

Distributions	A*	W*	L-S	K-S	p-value
Gumbel( $\mu, \sigma$ )	1.0083	0.149555	5.73492	0.130391	0.363071
GW( $\xi, k, \lambda$ )	1.81995	0.276761	6.2735	0.130799	0.359295
BE( $\lambda, a, b$ )	0.995606	0.160182	3.41307	0.12506	0.414754
ENH( $\alpha, \lambda, \theta$ )	1.32567	0.204332	3.84876	0.126957	0.395883
BG( $\alpha, \beta, \rho, \lambda$ )	1.83913	0.282842	7.01132	0.134266	0.328222
BEW( $\alpha, \lambda, a, b, c$ )	0.644849	0.107817	2.89811	0.112035	0.556868
BENH( $\lambda, \theta, \alpha, a, b$ )	<b>0.561666</b>	<b>0.094480</b>	<b>2.75311</b>	<b>0.107788</b>	<b>0.606717</b>

### 9.1.2. Meteorological parameters

We prove empirically that the BENH distribution is the best model for the yearly maxima wind data. So, it is of interest to obtain the following meteorological parameters for these data.

### 9.1.3. Return level

A return period is an estimate of the likelihood of an event, like a flood or wind speed to occur and it is the inverse of the probability that the event will be exceeded in any one year (or the expected waiting time or the mean number of years that will be taken for an exceeding event to occur). The probability, return period and return level (high quantile of the distribution) of the event can be estimated by  $P(x_T) = 1 - F(x_T)$ ,  $T = 1/P(x_T)$  and  $x_T = F^{-1}\left(1 - \frac{1}{T}\right)$ , respectively, where  $F^{-1}(\cdot)$  is the qf and  $P(x_T)$  is called the exceedance probability (Nadarajah and Choi [18]).

The return level  $x_T$  under the BENH distribution is obtained by solving

$$B\left(1 - e^{-(\lambda x_T + \beta x_T^k)}; \alpha, \rho\right) = B(\alpha, \rho) \left(1 - \frac{1}{T}\right).$$

where  $x_T > 0$  and  $T \geq 1$ . Table 6 provides the estimates of the return level  $x_T$  for the yearly maxima wind speed data in Table 3 by replacing the model parameters by their MLEs.

Table 5: Confidence intervals for the model parameters

CI	$\lambda$	$\beta$	k	$\alpha$	$\rho$
95%	[0 0.926]	[0 174.9]	[0.512 1.287]	[0 371.2]	[0 0.746]
99%	[0 1.079]	[0 217.2]	[0.389 1.410]	[0 460.7]	[0 0.867]

Table 6: Return level estimates  $\hat{x}_T$  for the T-year of the maxima wind speed data

T	Return level estimate of wind data (mm)
2	31.4687
5	40.1367
10	46.6312
20	53.1969
50	61.9929
100	68.7284
200	75.5285

#### 9.1.4. Mean deviation about the return level

The mean deviation about the return level is the mean of the distances of each value from their return level. Analogously to the mean deviation about the mean given in Section 4, the mean deviation about return level is given by

$$\eta = \int_0^{\infty} |x - x_T| f(x) dx = 2x_T F(x_T) - x_T + \mu - 2m(x_T),$$

where  $m(x_T)$  is given by (15) and  $f(x)$  is the density of the BENH distribution.

#### 9.2. Diabetic Retinopathy Study

We consider a data set analyzed by Huster *et al.*[16]. Patients with diabetic retinopathy in both eyes and 20/100 or better visual acuity for both eyes were eligible for the study. The patients were followed for two consecutively completed 4 month follow-ups and the endpoint was the occurrence of visual acuity less than 5/200. We choose only the treatment time. A 50% sample of the high-risk patients defined by diabetic retinopathy criteria was taken for the data set ( $n = 197$ ) and the percentage of censored observations was 72.4%. The variables involved in the study are:  $t_i$  - failure time for the treatment (in min); censoring indicator (0=censoring, 1=lifetime observed);  $x_{i1}$  - age (0 = patient is an adult diabetic, 1 = patient is a juvenile diabetic). We adopt the model  $y_i = \beta_0 + \beta_1 x_{i1} + z_i$ , where the random variable  $Y_i$  follows the LBENH distribution (25) for  $i = 1, \dots, 197$ . The MLEs of the model parameters are determined using the statistical software R. In order to estimate  $\alpha$ ,  $\theta$ ,  $a$ ,  $b$  and  $\beta$ , we perform iterative maximization of the logarithm of the likelihood function (28) starting with initial values for  $\alpha$  and  $\beta$  taken from the fitted log-NH regression model (with  $a = b = \theta = 1$ ). The MLEs (approximate standard errors and p-values in parentheses) are:  $\hat{\alpha} = 0.4536 (0.1887)$ ,  $\hat{\theta} = 10.7832 (0.1247)$ ,  $\hat{a} = 2.5057 (2.5872)$ ,  $\hat{b} = 0.0219 (0.0327)$ ,  $\hat{\beta}_0 = -3.6222 (1.1698) (0.4994)$  and  $\hat{\beta}_1 = 0.6399 (0.3984) (0.0274)$ . The explanatory variable  $x_1$  is marginally significant for the model at the significance level of 5%.

In order to assess if the model is appropriate, the empirical survival function and the estimated survival function (29) from the fitted LBENH regression model are plotted in Figure 6. In fact, the LBENH regression model provides a good fit for these data.



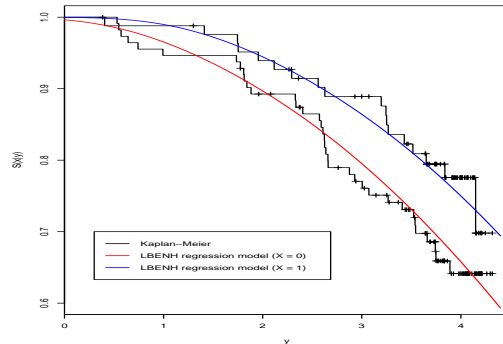


Figure 6: Estimated survival function by fitting the LBENH regression model and the empirical survival for each level of the diabetic retinopathy study.

## 10. Conclusions

We introduce a five-parameter lifetime model called the beta exponentiated Nadarajah-Haghighi (BENH) distribution, which extends several distributions proposed and widely used in the literature. The new distribution is useful to model lifetime data with increasing, decreasing, upside-down bathtub and bathtub shaped hazard functions. It can be much more flexible than the exponentiated Nadarajah-Haghighi, beta exponentiated Weibull, beta gamma, beta exponential, gamma Weibull and Gumbel distributions. We provide some closed-form expressions for the ordinary, incomplete and conditional moments, mean deviations and quantile and generating functions. By using nine goodness-of-fit statistics we prove empirically that the new model can be superior to some distributions generated from other families in terms of model fitting by means of a meteorological application.

## Acknowledgement

The research of Abdus Saboor has been supported in part by the Higher Education Commission of Pakistan under NRP project No. 3104. The research of Gauss M. Cordeiro has been supported by CNPq (Brazil).

## References

- [1] Gupta, R.D. and Kundu, D. Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, (1999) 41, 173–188.
- [2] Nadarajah, S., Kotz, S. The beta exponential distribution, *Reliability Engineering and System Safety*, (2006), 91, 689–697.
- [3] Nadarajah, S. and Haghighi, F. An extension of the exponential distribution. *Statistics*, (2011), 45, 543–558.
- [4] Lemonte, A.J. A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. *Computational Statistics and Data Analysis*, (2013), 62, 149–170.
- [5] Singla, N, Jain, K and Sharma, S.K. The Beta Generalized Weibull distribution: Properties and applications. *Reliability Engineering and System Safety*, (2012), 102, 5–15.
- [6] Eugene, N., Lee, C. and Famoye, F. Beta-normal distribution and its applications. *Communications in Statistics: Theory and Methods*, (2002), 31, 497–512.
- [7] Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of integrals, series, and products*. Academic Press. San Diego
- [8] Cordeiro, G.M., Gomes, A. E, da-Silva, C. Q and Ortega, E.M.M. Cordeiro, G.M., Silva, G.O. and Ortega, E.M.M. The beta-Weibull geometric distribution. *Statistics*. (2011). DOI: 10.1080/02331888.2011.577897.

- [9] Pescim, R.R., Demétrio, C.G.B., Cordeiro, G.M., and Ortega, E.M.M. and Urbano, M.R. The beta generalized half-normal distribution. *Computational Statistics and Data Analysis*, (2010), 54, 945–957.
- [10] Souza, W. B, Alessandro H.S. Santos and Gauss, G.M. The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, (2010), 80, 159–172.
- [11] Gupta, R.D. and Kundu, D. Discriminating between the Weibull and the GE distributions. *Computational Statistics and Data Analysis*, (2003), 43, 179–196.
- [12] Kong, L., Lee, C and Sepanski, J.H. On the Properties of Beta–Gamma Distribution, *Journal of Modern Applied Statistical Methods*, (2007), 6, 187–211.
- [13] Cordeiro, G.M., Gomes. A. E, da-Silva. C. Q and Ortega, E.M.M. The beta exponentiated Weibull distribution. *Journal of Statistical Computation and Simulation*, (2013) 83, 141–138.
- [14] Provost, S.B., Saboor, A. and Ahmad, M. The gamma–Weibull distribution. *Pakistan Journal of Statistics*, (2011), 27, 111–113.
- [15] Castillo, E., Hadi, A.S., Balakrishnan, N. and Sarabia, J.M. *Extreme Value and Related Models with Applications in Engineering and Science*. John Wiley (2005).
- [16] Huster, W.J., Brookmeyer, R., Self, S.G. Modelling paired survival data with covariates. *Biometrics*, (1989), 45, 145-156.
- [17] Bonferroni C.E. *Elementi di statistica generale*. Libreria Seber, Firenze, (1930).
- [18] Nadarajah, S. and choi, D. Maximum daily rainfall in South Korea. *Journal of Earth System Science*, (2007), 116, 311-320.

## Appendix

The symbol  $\psi^{(1)}(\cdot)$  denotes the trigamma function, and the polygamma function of order  $m$  is denoted by:

$$\psi^{(m)}(z) = \frac{d^{m+1}}{dz^{m+1}} \log[\Gamma(z)].$$

Figure 7 displays the estimated cdfs for some models and the empirical cdf for yearly maxima wind data.

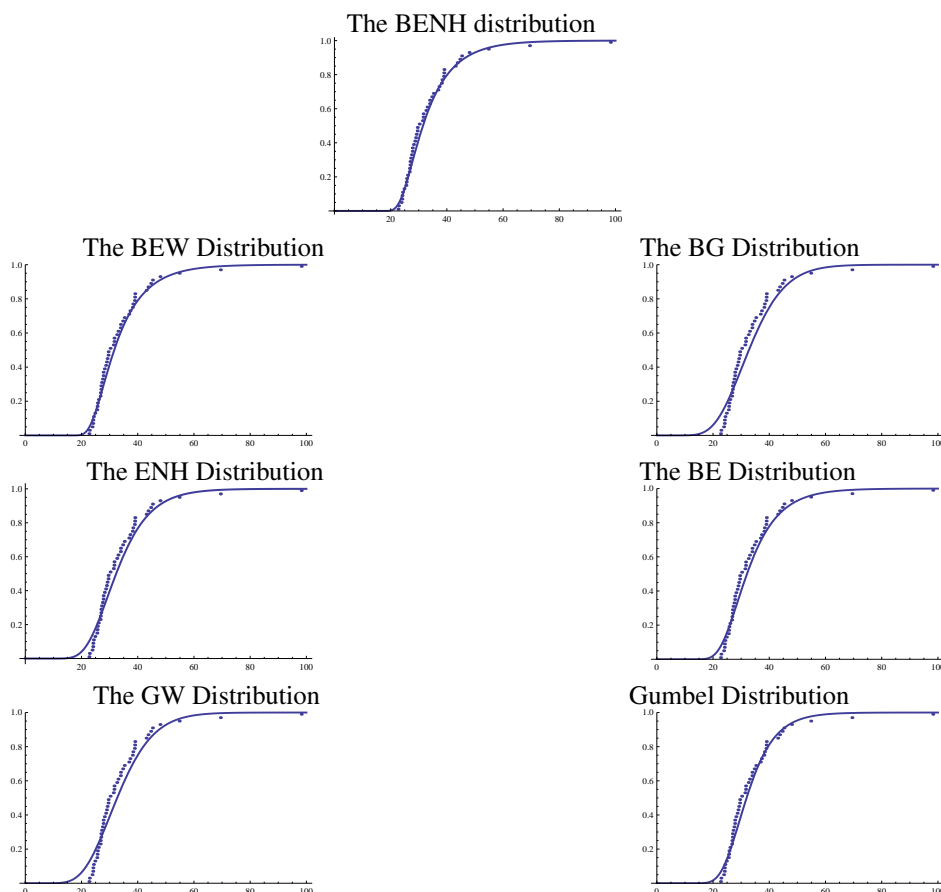


Figure 7: The estimated cdfs for some models and the empirical cdf for yearly maxima wind data.