مجله‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **16** (2019): 319–342

دوره‌ی ۱۶، شماره‌ی ۲، پاییز و زمستان ۱۳۹۸، صص ۳۱۹–۳۴۲

DOI: 10.52547/jsri.16.2.319

DOR: 20.1001.1.25385771.2020.16.2.4.4

Analysis of Wind Speed Data Based on the New Upper Truncated Inverse Weibull Distribution; A Case Study: Ardabil, Iran

Gholamhossein Yari* and Zahra Karimi Ezmareh

Iran University of Science and Technology

Received: 2020/23/08 Approved: 2021/08/02

Abstract. In this paper, for the first time, the upper truncated inverse Weibull (UTIW) distribution is proposed for modeling wind speed data. Since there is a upper limit for empirical wind speed data, this data can be represented by using the UTIW distribution. In this study, the UTIW distribution is introduced and some of its statistical properties are studied. Then, the parameters of this distribution are estimated by using different methods. Simulation studies for these estimators are presented. In addition, the mentioned distribution performance is tested on real wind speed data of Ardabil province in Iran. Based on the results of the analysis, it is found that the presented distribution in this study for modeling wind speed data is more appropriate than recently introduced distributions. Finally, this distribution can be used as an alternative model for evaluating wind speed data.

Keywords. Inverse Weibull distribution; upper truncated inverse Weibull distribution; wind speed; parameters estimation; Monte-Carlo simulation; model selection criteria.

MSC 2010: 62Exx, 62Pxx, 97Kxx.

* Corresponding author

1 Introduction

Wind energy is one of the most important renewable energies. Today, the use of wind energy is very much considered because of low cost and friendly behavior with nature. The prediction of wind speed with its random nature can be studied as a probabilistic model. Determining the exact probability distribution of wind speed is a guarantee of the optimum use of wind energy and its importance among other energy sources. For more details see Resources Keyhani et al. (2010), Fyrippis et al. (2010), Ilkiliç and Aydin (2015), Wang et al. (2015), Köse (2004), Kaplan (2015) and Huang and McElroy (2015).

The Weibull distribution is considered as a widely used distribution in fitting to wind speed data, for simplicity, flexibility and the fact that its parameters are easily estimated (Ouammi et al. (2010), Weisser (2003), Ayodele et al. (2012), Celik (2003), Dabbaghiyan et al. (2016), Yip et al. (2016), Akdağ and Güler (2009) and Ozgur et al. (2009)). However, this distribution is unable to model all the wind structures encountered in nature (Carta et al. (2008) and Ouarda et al. (2015)). For this reason, alternative distributions have been proposed for the modeling of wind speed data. Some of the proposed distributions for modeling wind speed data are: Weibull and Rayleigh (Safari and Gasore (2010) and Morgan (1995)), inverse Gaussian, Pearson type V, and Burr (Brano et al. (2011)), Kappa and Wakeby (Morgan et al. (2011)), Johnson SB (Soukissian (2013)) and mixture distributions (Akpinar and Akpinar (2009) and Qin et al. (2012)). Also, Ramírez and Carta (2006), Akpinar and Akpinar (2007), Shamilov et al. (2008) and Kantar and Usta (2008), respectively propose some distributions derived from the maximum entropy principle and the minimum cross-entropy principle for modeling wind speed data. Recently, new models have been proposed for fitting on wind speed data. Kantar and Usta (2015) considers upper-truncated Weibull distribution, limited by an upper bound. Arslan et al. (2017) propose to use Generalized Lindley and Power Lindley distributions for modeling the wind speed data. Kantar et al. (2018) introduce the Extended Generalized Lindley distribution (EGL) as an alternative wind speed distribution. Akgül et al. (2016) suggest the inverse Weibull (IW) distribution for wind speed. Ouarda and Charron (2018) have studied the mixture of wind speed distribution.

Due to the Weibull distribution constraints, various types of it have been introduced. One of the most important of which is the Inverse Weibull (IW) distribution. This distribution is applicable in various fields. Since it has high flexibility for modeling the long tailed right skewed data, it is considered in this paper. Interested readers, see sources Drapella (1993), Keller et al. (1982), Mudholkar et al. (1996), Keller and Kamath (1982), Murthy et al. (2004), De Gusmao et al. (2011), Pasari and Dikshit (2014), Khan et al. (2008), Aydin (2018) and Jiang et al. (2001).

Truncated distributions are used in situations where occurrences are limited to a maximum or a minimum value, or within a specific range. If occurrences are limited to values which lie below a given threshold or above a given threshold, the lower (left) truncated distribution and the upper (right) truncated distribution are obtained respectively (Zhang and Xie (2011)).

Identifying a statistical model that has the best fit on real data is one of the main topics in the statistics. Since the wind speed data are limited to a maximum value, this paper suggests the UTIW distribution as an alternative to widely-used wind speed distribution. Then, evaluating this new model using various criteria shows that it has the best performance relative to the Weibull (W), Gamma (G) and Extended Generalized Lindley (EGL) distributions. Estimation of parameters is one of the main issues in determining wind speed distribution. In this study, four estimation methods are presented and these methods are compared using the Monte Carlo (MC) simulation.

The rest of this article is organized as follows: In section 2, the UTIW distribution is introduced. In section 3 some of the statistical properties of this distribution are studied. In section 4, the parameters of the UTIW distribution are estimated using four methods of the moment (MME), maximum likelihood (MLE), least square (LS) and Bayes and these methods are compared by using the MC simulation. Then, in section 5, some of the required criteria are presented to evaluate the UTIW distribution. Finally, in section 6, mentioned distribution performance is tested on real wind speed data in Ardabil province that measured from 2013 to 2016.

2 New Model

When the random variable is limited from above to an unknown cut-off point, upper truncated distribution is used. This point is called the truncation point. Since wind speed data has a upper bound, in this research, the UTIW distribution is proposed for modeling wind speed data.

Definition 1. Upper truncated cumulative distribution function (cdf) is defined as follows

$$G(x) = \frac{F(x)}{F(T)}, \quad 0 \leq x \leq T, \tag{1}$$

where T is truncation point and F(x) is the cdf of X random variable (Zhang and Xie (2011)).

If IW distribution cdf, $F(x) = e^{-(\frac{x}{\beta})^{-\alpha}}$, is replaced in 1 then, the cdf and the probability density function (pdf) of the UTIW distribution are given by

$$G(x) = \frac{e^{-(\frac{x}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}}, \quad 0 \leq x \leq T, \tag{2}$$

$$g(x) = \frac{\frac{\alpha}{\beta}(\frac{x}{\beta})^{-(\alpha+1)} e^{-(\frac{x}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}}, \quad 0 \leq x \leq T. \tag{3}$$

Where $\alpha, \beta > 0$, β is the scale parameter, α is the shape parameter and the truncation point T is generally assumed to be known. The reliability (R) and hazard rate function (H) for the UTIW distribution are

$$R(x;\alpha,\beta) = 1 - G(x) = \frac{e^{-(\frac{T}{\beta})^{-\alpha}} - e^{-(\frac{x}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}}, \quad 0 \leq x \leq T, \tag{4}$$

$$H(x;\alpha,\beta) = \frac{g(x)}{1 - G(x)} = \frac{\frac{\alpha}{\beta}(\frac{x}{\beta})^{-(\alpha+1)} e^{-(\frac{x}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}} - e^{-(\frac{x}{\beta})^{-\alpha}}}, \quad 0 \leq x \leq T\cdot \tag{5}$$

Where $\alpha, \beta > 0$. Figure 1 shows the pdf, cdf, reliability and hazard rate function plots of the UTIW distribution for different values of α and β.

3 Statistical Properties of the New Distribution

In this section, some properties of the introduced distribution such as mode, rth moment, quantile function, moment generating function, characteristic function and order statistics are obtained.

3.1 The rth Moment

Since moments are used in computing measures of central tendency, dispersion, shapes and estimation parameters, It is imperative to derive the

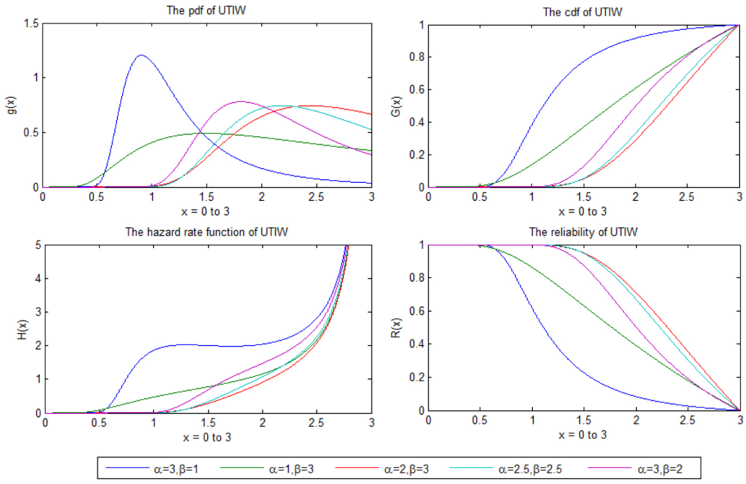


Figure 1. The pdf, cdf, reliability and hazard rate function plots of the UTIW distribution for different values of the α and β.

moments when a new distribution is proposed.

***Theorem* 1.** *The rth moment of the UTIW distribution is*

$$E(X^r) = \frac{-\beta^r \Gamma\big(1-\frac{r}{\alpha}, (\frac{T}{\beta})^{-\alpha}\big)}{e^{-(\frac{T}{\beta})^{-\alpha}}}. \qquad (6)$$

Where Γ *is incomplete gamma function* $(\Gamma(c,x) = \int_0^x t^{c-1}e^{-t}dt)$.

Proof. To prove this, it is enough to use the $u = (\frac{x}{\beta})^{-\alpha}$ transform

$$\begin{aligned} E(X^r) &= \frac{1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \int_0^T x^r \frac{\alpha}{\beta}(\frac{x}{\beta})^{-(\alpha+1)} e^{-(\frac{x}{\beta})^{-\alpha}} dx, \\ &= \frac{-1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \int_0^{(\frac{T}{\beta})^{-\alpha}} (\beta u^{-\frac{1}{\alpha}})^r e^{-u} du, \\ &= \frac{-\beta^r \Gamma\big(1-\frac{r}{\alpha}, (\frac{T}{\beta})^{-\alpha}\big)}{e^{-(\frac{T}{\beta})^{-\alpha}}}. \end{aligned}$$

□

By using 6, the mean and variance of the UTIW distribution with a known truncation point are respectively

$$\mu_{UTIW} = \frac{-\beta\Gamma\big(1-\frac{1}{\alpha},(\frac{T}{\beta})^{-\alpha}\big)}{e^{-(\frac{T}{\beta})^{-\alpha}}},$$
$$Var_{UTIW} = \frac{\beta^2\Gamma^2\big(1-\frac{1}{\alpha},(\frac{T}{\beta})^{-\alpha}\big)}{\big[e^{-(\frac{T}{\beta})^{-\alpha}}\big]^2} - \frac{\beta^2\Gamma\big(1-\frac{2}{\alpha},(\frac{T}{\beta})^{-\alpha}\big)}{e^{-(\frac{T}{\beta})^{-\alpha}}}.$$

The skewness (S) and kurtosis (K) of the proposed distribution are obtained by using the following two relations:

$$S(x;\alpha,\beta) = \frac{E(X^3) - 3E(X)E(X^2) + 2E^3(X)}{Var^{\frac{3}{2}}(X)},$$
$$K(x;\alpha,\beta) = \frac{E(X^4) - 4E(X)E(X^3) + 6E^2(X)E(X^2) - 4E^4(X)}{Var^2(X)} - 3.$$

3.2 Quantile Function

The quantile function is useful in generating sample and in computing the median, kurtosis and skewness of a distribution. To obtain a quantile function, it is suffices to solve equation $G(x) = U$. Then the quantile function of the UTIW distribution is

$$Q(u) = \beta\big[(\frac{T}{\beta})^{-\alpha} - \log u\big]^{-\frac{1}{\alpha}}. \tag{7}$$

Where $U \sim Uniform(0,1)$ (U is a random variable that has a Uniform distribution).

Table 1 shows the mean, variance, median, skewness, kurtosis, quartiles, cdf, pdf, reliability and hazard function values for $\alpha = 3, \beta = 1.1$ and different values of x and T.

3.3 Moment Generating Function and Characteristics Function

In the following two theorems, the moment generating function and characteristics function of the UTIW distribution are calculated.

Table 1. The mean, variance, median, quartiles, kurtosis, skewness, cdf, pdf, hazard rate function and reliability values for $\alpha = 3, \beta = 1.1$ and different values of x and T.

Statistics	$T = 2, x = 1.5$	$T = 4, x = 3$	$T = 5, x = 2$	$T = 8, x = 5$	$T = 10, x = 8$
Mean	1.7592	1.5208	1.5055	1.4934	1.4915
Variance	0.7336	0.9967	1.0098	1.0197	1.0212
Median	1.1569	1.2308	1.2366	1.2414	1.2421
Q_1	0.9499	0.9816	0.9840	0.9859	0.9862
Q_3	1.4312	1.6280	1.6462	1.6613	1.6637
Kurtosis	1.1955	1.3873	1.4139	1.4374	1.4413
Skewness	2.5439	2.1333	2.1044	2.0815	2.0778
G(x)	0.7961	0.9719	0.8558	0.9920	0.9987
g(x)	0.6279	0.0479	0.2136	0.0063	0.0010
H(x)	3.0801	1.7052	1.4810	0.7906	0.7680
R(x)	0.2039	0.0281	0.1442	0.0080	0.0013

Theorem 2. *The moment generating function of the UTIW distribution is*

$$M_X(t) = \frac{-1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \sum_{i=0}^{\infty} \frac{(t\beta)^i}{i!} \Gamma\big(1 - \frac{i}{\alpha}, (\frac{T}{\beta})^{-\alpha}\big). \tag{8}$$

Where Γ *is incomplete gamma function.*

Proof. By using transform $u = (\frac{x}{\beta})^{-\alpha}$ we have

$$\begin{aligned} M_X(t) &= E(e^{tx}), \\ &= \frac{1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \int_0^T e^{tx} \frac{\alpha}{\beta} (\frac{x}{\beta})^{-(\alpha+1)} e^{-(\frac{x}{\beta})^{-\alpha}} dx, \\ &= \frac{1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \int_0^T \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} \frac{\alpha}{\beta} (\frac{x}{\beta})^{-(\alpha+1)} e^{-(\frac{x}{\beta})^{-\alpha}} dx, \\ &= \frac{-1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \sum_{i=0}^{\infty} \frac{t^i}{i!} \int_0^{(\frac{T}{\beta})^{-\alpha}} (\beta u^{-\frac{1}{\alpha}})^i e^{-u} du, \\ &= \frac{-1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \sum_{i=0}^{\infty} \frac{(t\beta)^i}{i!} \Gamma(1 - \frac{i}{\alpha}, (\frac{T}{\beta})^{-\alpha}). \end{aligned}$$

□

Theorem 3. *The characteristics function of the UTIW distribution is*

$$M_X(ti) = \frac{-1}{e^{-(\frac{T}{\beta})^{-\alpha}}} \sum_{k=0}^{\infty} \frac{(it\beta)^k}{k!} \Gamma(1 - \frac{k}{\alpha}, (\frac{T}{\beta})^{-\alpha}). \tag{9}$$

Proof. Similar to the proof of Theorem 2, the characteristics function of the UTIW distribution is obtained. □

3.4 Mode

For obtaining the mode of a distribution, it is enough to derive from the density function respect to x and put the derivative equal to zero, then the mode of the UTIW distribution is $x = (\frac{\alpha\beta^\alpha}{\alpha+1})^{\frac{1}{\alpha}}$.

3.5 Order Statistics

The density of rth order statistics of a distribution is calculated by

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} g(x)[G(x)]^{r-1}[1 - G(x)]^{n-r}. \tag{10}$$

Where $G(x)$ and $g(x)$ are the cdf and the pdf of desired distribution respectively.
By using expansion $[1 - G(x)]^{n-r} = \sum_{i=0}^{n-r} \binom{n-r}{i} [-G(x)]^i$ and 10, one can obtain

$$f_{r:n}(x) = \sum_{i=0}^{n-r} \frac{n!(-1)^i}{i!(r-1)!(n-r-1)!} g(x)[G(x)]^{r+i-1} \tag{11}$$

By substituting 2 and 3 in 11, the density of rth order statistics for UTIW distribution is obtained.

4 Parameter Estimation Methods

In this section, with four methods of MME, MLE, LS and Bayse the unknown parameters of the UTIW distribution are estimated. In addition, the performance of these methods is compared using the MC simulation.

4.1 MME Method

For estimating the UTIW distribution parameters by the MME method, it is enough to equate the population moments and the sample moments. Thus, the MME estimators of the parameters is obtained by solving equations 12 and 13.

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{-\beta\Gamma\big(1-\frac{1}{\alpha},(\frac{T}{\beta})^{-\alpha}\big)}{e^{-(\frac{T}{\beta})^{-\alpha}}}, \tag{12}$$

$$\bar{x^2} = \sum_{i=1}^{n} \frac{x_i^2}{n} = \frac{-\beta^2\Gamma\big(1-\frac{2}{\alpha},(\frac{T}{\beta})^{-\alpha}\big)}{e^{-(\frac{T}{\beta})^{-\alpha}}}. \tag{13}$$

4.2 MLE Method

The MLE method is one of the most widely used methods for estimating unknown parameters. This method has asymptomatic statistical characteristics such as consistency, efficiency and normality. The MLE estimator is the value that maximizes the log-likelihood function.

For this purpose, let $X_1, X_2, ..., X_n \sim UTIW(\alpha, \beta)$ denote random samples with parameter vector $\Theta = (\alpha, \beta)$, then the likelihood function is given by

$$l(\Theta|x) = \prod_{i=1}^{n} \frac{\frac{\alpha}{\beta}(\frac{x_i}{\beta})^{-(\alpha+1)}e^{-(\frac{x_i}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}} = \frac{(\frac{\alpha}{\beta})^n \prod_{i=1}^{n}(\frac{x_i}{\beta})^{-(\alpha+1)}e^{-\sum_{i=1}^{n}(\frac{x_i}{\beta})^{-\alpha}}}{\big(e^{-(\frac{T}{\beta})^{-\alpha}}\big)^n}. \tag{14}$$

The log-likelihood function is

$$L = \log l(\Theta|x) = n\log\alpha - n\log\beta + n(\frac{T}{\beta})^{-\alpha} - (\alpha+1)\sum_{i=1}^{n}\log(\frac{x_i}{\beta}) - \sum_{i=1}^{n}(\frac{x_i}{\beta})^{-\alpha}. \tag{15}$$

Where $x_i \in [0, T], i = 1, 2, ..., n$. If the truncation point (T) is taken as the largest sample observation ($T \to x_{max}$) the log-likelihood function reaches the maximum value. Now, by deriving from the above function respect to the α and β parameters and equating to zero of these partial derivatives, the following equations are obtained:

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} - n(\frac{T}{\beta})^{-\alpha}\log(\frac{T}{\beta}) + \sum_{i=1}^{n}\log(\frac{x_i}{\beta})\{(\frac{x_i}{\beta})^{-\alpha} - 1\} = 0,$$
$$\frac{\partial L}{\partial \beta} = \frac{n\alpha}{\beta}(\frac{T}{\beta})^{-\alpha} + \frac{n\alpha}{\beta} - \frac{\alpha}{\beta}\sum_{i=1}^{n}(\frac{x_i}{\beta})^{-\alpha} = 0.$$

The above equations are not solved by analytical methods. Therefore, "Newton Raphson method" is used to solve these equations. By obtaining the second derivative of the log-likelihood function, the elements of the matrix of information are obtained as follows

$$\frac{\partial^2 L}{\partial \alpha^2} = -\frac{n}{\alpha^2} + n(\frac{T}{\beta})^{-\alpha}[\log(\frac{T}{\beta})]^2 - \sum_{i=1}^{n}(\frac{x_i}{\beta})^{-\alpha}[\log(\frac{x_i}{\beta})]^2,$$
$$\frac{\partial^2 L}{\partial \beta^2} = \frac{n\alpha(\alpha-1)}{\beta^2}(\frac{T}{\beta})^{-\alpha} - \frac{n\alpha}{\beta^2} - \frac{\alpha(\alpha-1)}{\beta^2}\sum_{i=1}^{n}(\frac{x_i}{\beta})^{-\alpha},$$
$$\frac{\partial^2 L}{\partial \beta \partial \alpha} = \frac{n}{\beta} - \frac{1}{\beta}\sum_{i=1}^{n}(\frac{x_i}{\beta})^{-\alpha}\big[1 - \alpha\log(\frac{x_i}{\beta})\big] + \frac{n}{\beta}(\frac{T}{\beta})^{-\alpha}\big[1 - \alpha\log(\frac{T}{\beta})\big].$$

Information matrix is given by

$$I^{-1}(\Theta) = \begin{pmatrix} \frac{\partial^2 L}{\partial \alpha^2} & \frac{\partial^2 L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 L}{\partial \beta \partial \alpha} & \frac{\partial^2 L}{\partial \beta^2} \end{pmatrix}. \quad (16)$$

Where $\sqrt{n}[(\alpha - \hat{\alpha}), (\beta - \hat{\beta})] \longrightarrow N(0, I^{-1}(\alpha, \beta))$. The approximate $(1 - \xi)100\%$ confidence intervals for the parameters α and β are determined, respectively as

$$\hat{\alpha} \pm Z_{\frac{\xi}{2}}\sqrt{\frac{Var(\hat{\alpha})}{n}}, \quad \hat{\beta} \pm Z_{\frac{\xi}{2}}\sqrt{\frac{Var(\hat{\beta})}{n}}.$$

Where $Var(\hat{\alpha})$ and $Var(\hat{\beta})$ are the diagonal elements of $I^{-1}(\Theta)$ and $Z_{\frac{\xi}{2}}$ is the upper $\frac{\xi}{2}$ percentile of the standard normal distribution.

4.3 LS Method

let $X_1, ..., X_n$ are random samples that have UTIW distribution. Then, the LS estimator of parameters α and β are obtained by minimizing the following function

$$LS = \sum_{i=1}^{n} [G(x_{(i)}) - u_i]^2,$$

where $G(x_{(i)}) = \frac{e^{-(\frac{x_{(i)}}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}}$ is the distribution function of the ordered random variables and $u_i = \frac{i}{n+1}$. Then

$$\begin{aligned}\frac{\partial LS}{\partial \alpha} =& \frac{2}{\left[e^{-(\frac{T}{\beta})^{-\alpha}}\right]^2} \sum_{i=1}^{n} e^{-(\frac{x_{(i)}}{\beta})^{-\alpha}} e^{-(\frac{T}{\beta})^{-\alpha}} \left\{(\frac{x_{(i)}}{\beta})^{-\alpha} \log(\frac{x_{(i)}}{\beta}) - (\frac{T}{\beta})^{-\alpha} \log(\frac{T}{\beta})\right\} \\ & \left\{\frac{e^{-(\frac{x_{(i)}}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}} - u_i\right\} = 0, \qquad (17) \\ \frac{\partial LS}{\partial \beta} =& \frac{2\alpha}{\beta\left[e^{-(\frac{T}{\beta})^{-\alpha}}\right]^2} \sum_{i=1}^{n} e^{-(\frac{x_{(i)}}{\beta})^{-\alpha}} e^{-(\frac{T}{\beta})^{-\alpha}} \left\{ - (\frac{x_{(i)}}{\beta})^{-\alpha} + (\frac{T}{\beta})^{-\alpha}\right\} \\ & \times \left\{\frac{e^{-(\frac{x_{(i)}}{\beta})^{-\alpha}}}{e^{-(\frac{T}{\beta})^{-\alpha}}} - u_i\right\} = 0. \qquad (18)\end{aligned}$$

By numerical solution of equations 17 and 18, estimation of unknown parameters are obtained.

4.4 Bayes Method

In this method, let the previous information of α and β are independent of each other, so $\Pi(\alpha, \beta) = \Pi(\alpha)\Pi(\beta)$. The prior distribution Uniform (BU) is used. Since the denominators of the posterior distributions have a two-integral, it is not easy to calculate them, so the "important sampling method" is used (Rubinstein and Kroese, 2016). In the relevant section, the results are presented.

4.5 Simulation Study With Using MC Method

Now, the performance of the four mentioned methods is compared to estimate the UTIW distribution parameters using the MC simulation. At the first, a random sample of the UTIW distribution with parameters $\alpha_{true} = 5$ and $\beta_{true} = 3$ (let $T = 6$) is generated. The required random sample is generated by the method ”inverse transform”, that the algorithm of this method is as follows

1. Generate U from $U(0, 1)$.
2. Return $X = F^{-1}(U)$.

Where $X = \beta\big[(\frac{T}{\beta})^{-\alpha} - \log u\big]^{-\frac{1}{\alpha}}$.

This study is done for sample size $n = 30, 50, 70, 100, 120, 150, 170, 200$, and the number of replicates is 1000. Finally, for comparing these methods, criteria such as bias, variance and mean square error (MSE) are derived. Bias and MSE of the estimators $\hat{\alpha}$ and $\hat{\beta}$ are calculated respectively

$$bias(\hat{\alpha}) = E(\hat{\alpha}) - \alpha_{true}, \quad bias(\hat{\beta}) = E(\hat{\beta}) - \beta_{true},$$
$$MSE(\hat{\alpha}) = E(\hat{\alpha} - \alpha_{true})^2, \quad MSE(\hat{\beta}) = E(\hat{\beta} - \beta_{true})^2.$$

MATLAB software has been used for simulation studies and the results have been summarized in Figures 2 and 3.

According to Figures 2 and 3, we conclude that MLE method have the best performance because this method has the least variance and MSE and the bias is close to zero and the MME method is the most inappropriate. However, among these methods, the estimators obtained by using the MLE approach are very close to real value. Hence, in the remainder of this paper, the MLE estimators of the parameters are used.

5 Goodness-of-Fit Evaluation

To evaluate performance of the UTIW distribution, criteria such as root mean square error (RMSE), coefficient of determination (R^2), maximum value of the log-likelihood function corresponding to the MLE estimates of the parameters (L), Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criteria (CAIC) and Kolmogorov-Smirnov test (KS) are used. The formulas for these criteria have been listed in Table 2.

Where, in Table 2, L is log likelihood function, q is the number of parameters, n is sample size and $x_{(i)}$ is the ordered observation.

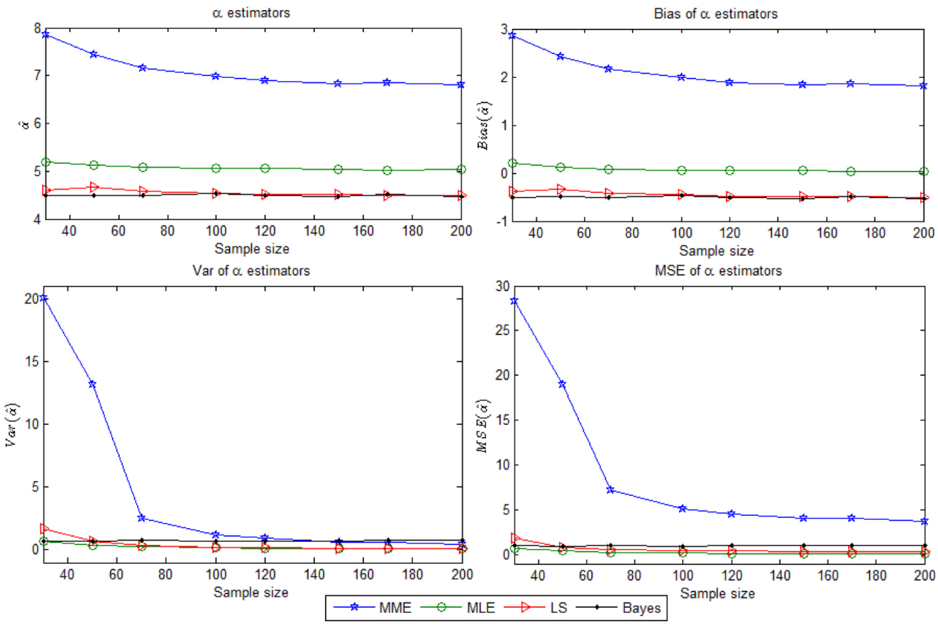


Figure 2. The results of α estimators.

Table 2. Goodness-of-fit criteria

Goodness-of-fit	Equation
Akaike information criterion	$AIC = -2L + 2q$
Bayesian information criterion	$BIC = -2L + q\log(n)$
Consistent Akaike information criteria	$CAIC = -2L + \frac{2qn}{n-q-1}$
Kolmogorov-Smirnov	$KS = \max_{1\leq i\leq n}\{\lvert\frac{i}{n} - F(x_{(i)})\rvert, \lvert F(x_{(i)} - \frac{i-1}{n}\rvert\}$
Root mean square error	$RMSE = \Big[\sum_{i=1}^{n}\frac{\left(\hat{F}(x_{(i)})-\frac{i}{n+1}\right)^2}{n}\Big]^{\frac{1}{2}}$
Coefficient of determination	$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\hat{F}(x_{(i)})-\frac{i}{n+1}\right)^2}{\sum_{i=1}^{n}\left(\hat{F}(x_{(i)})-\bar{\hat{F}}(x_{(i)})\right)^2}$

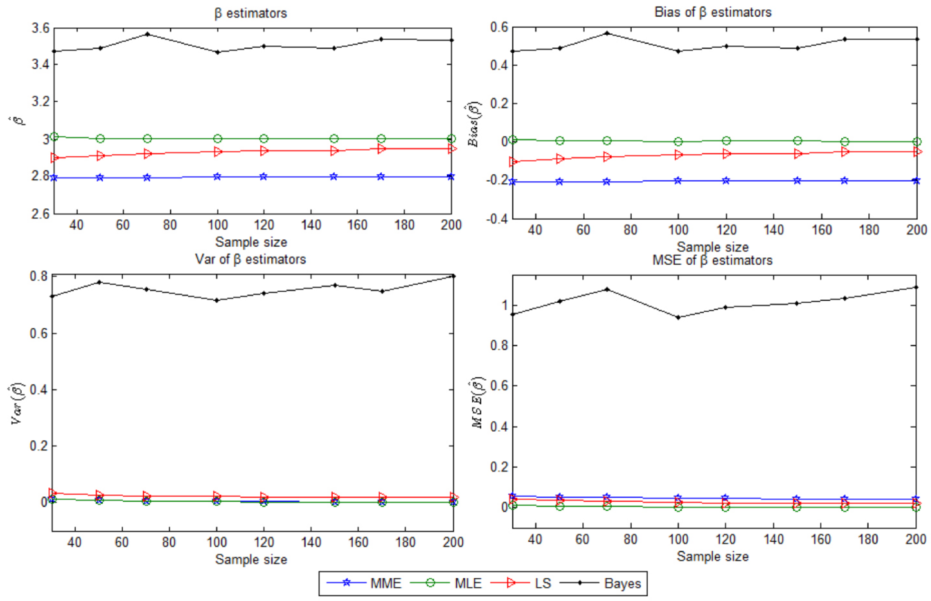


Figure 3. The results of β estimators.

6 Application: Wind Speed Data

The data used in this paper have been taken from the Meteorological Organization of Iran. These data include the monthly average wind speed of Ardebil province from 2013 to 2016, which is given in Table 3. For evaluating the performance of the UTIW distribution in the modeling of wind speed data, this distribution is fitted on these data. Table 5 shows the MLE estimator of unknown parameters and Table 6 compares the effectiveness of the UTIW distribution, for modeling wind speed data versus EGL, W and G distributions, through criteria such as AIC, BIC, CAIC, L, KS, P-value, RMSE and R^2. Table 4 shows the pdf of EGL, W and G distributions.

It is clear from Table 6 that the UTIW distribution has the best performance in fitting on these data and is the most appropriate relative to other distributions based on all the considered criteria. For example in 2013 the UTIW distribution in this research has the best performance because it has the lowest values of AIC, BIC, CAIC, KS and RMSE and the highest values of L, P-value and R^2. The results of these criteria can be summarized as follows

Table 3. Mean wind speed of Ardebil province in Iran from 2013 to 2016.

Year	2013	2014	2015	2016
January	5.7	4.6	4.9	7.4
February	3.2	3.1	3.8	4.9
March	5.1	4.7	3.1	4.6
April	2.6	3.5	4.6	3
May	3.4	3.1	3.3	3.4
June	3.1	3.5	3.8	3.4
July	4	3.4	3.9	3.9
August	3.9	3.5	3.9	3
September	3.7	3.3	2.4	2.7
October	3.7	3.6	3.7	2.8
November	2.8	3.1	4.2	3.6
December	3.6	4.8	4.9	5

Table 4. The density of selected distributions for fitting on wind speed data.

Distribution	Density	Range
EGL	$f_{EGL}(x) = \frac{pk^2 c(1+cx)^{2p-1} e^{k-k(1+cx)^p}}{k+1}$	$x, c, p, k > 0$
Weibull	$f_W(x) = \frac{\alpha}{\beta}(\frac{x}{\beta})^{\alpha-1} e^{-(\frac{x}{\beta})^\alpha}$	$x, \alpha, \beta > 0$
Gamma	$f_G(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)}$	$x, \alpha, \beta > 0$

Table 5. MLE values of the parameters

	Year			
Distribution	2013	2014	2015	2016
UTIW	$\hat{\alpha} = 4.8250$	$\hat{\alpha} = 7.9765$	$\hat{\alpha} = 2.054$	$\hat{\alpha} = 4.4243$
	$\hat{\beta} = 3.3651$	$\hat{\beta} = 3.4285$	$\hat{\beta} = 5.231$	$\hat{\beta} = 3.3788$
EGL	$\hat{c} = 4.4458628$	$\hat{c} = 1.0798254$	$\hat{c} = 0.07895$	$\hat{c} = 1.416e+01$
	$\hat{p} = 3.2576718$	$\hat{p} = 5.1846309$	$\hat{p} = 18.18965$	$\hat{p} = 2.276e+00$
	$\hat{k} = 0.0001451$	$\hat{k} = 0.0003965$	$\hat{k} = 0.01212$	$\hat{k} = 1.727e-04$
Weibull	$\hat{\alpha} = 4.369101$	$\hat{\alpha} = 6.050477$	$\hat{\alpha} = 6.463977$	$\hat{\alpha} = 3.119020$
	$\hat{\beta} = 4.080962$	$\hat{\beta} = 3.954042$	$\hat{\beta} = 4.162333$	$\hat{\beta} = 4.432718$
Gamma	$\hat{\alpha} = 20.880102$	$\hat{\alpha} = 39.52571$	$\hat{\alpha} = 28.22638$	$\hat{\alpha} = 11.603501$
	$\hat{\beta} = 5.592762$	$\hat{\beta} = 10.73087$	$\hat{\beta} = 7.28425$	$\hat{\beta} = 2.919286$

Table 6. Results of goodness-of-fit criteria for wind speed data of Ardabil province.

Year	Distribution	AIC	BIC	CAIC	L	KS	Pvalue	RMSE	R^2
2013	UTIW	30.49572	31.46554	31.82906	−13.24785	0.1802	0.8304	0.05116186	0.9680814
	EGL	36.54219	37.99691	39.54219	−15.27108	0.2267	0.5685	0.08151827	0.9054252
	Weibull	35.51438	36.48419	36.84771	−15.75719	0.2334	0.5305	0.08403572	0.893143
	Gamma	32.8138	33.78361	34.14713	−14.4069	0.1811	0.8263	0.063328	0.9486589
2014	UTIW	20.44323	21.41304	21.77656	−8.221616	0.2168	0.6252	0.1049911	0.888284
	EGL	29.28622	30.74094	32.28622	−11.64311	0.3139	0.1879	0.1313041	0.7882557
	Weibull	27.92791	28.89773	29.26125	−11.96396	0.3173	0.1784	0.1323845	0.7795464
	Gamma	25.02029	25.9901	26.35363	−10.51015	0.2859	0.2803	0.1229977	0.828399
2015	UTIW	27.1373	28.10711	28.47063	−11.56865	0.185	0.8058	0.06710672	0.9441288
	EGL	31.26519	32.71991	34.26519	−12.6326	0.1918	0.7694	0.0662878	0.9480651
	Weibull	29.15501	30.12483	30.48835	−12.57751	0.1853	0.8043	0.06470574	0.9503291
	Gamma	30.1964	31.16622	31.52973	−13.0982	0.1783	0.8399	0.06135646	0.9522002
2016	UTIW	36.99553	37.96534	38.32886	−16.49777	0.1451	0.9622	0.0632527	0.9580506
	EGL	44.89427	46.34899	47.89427	−19.44707	0.1792	0.8355	0.0768337	0.9139807
	Weibull	44.20502	45.17483	45.53835	−20.10251	0.1919	0.769	0.08051086	0.8956057
	Gamma	41.05548	42.02529	42.38881	−18.52774	0.1755	0.8537	0.07416397	0.9316567

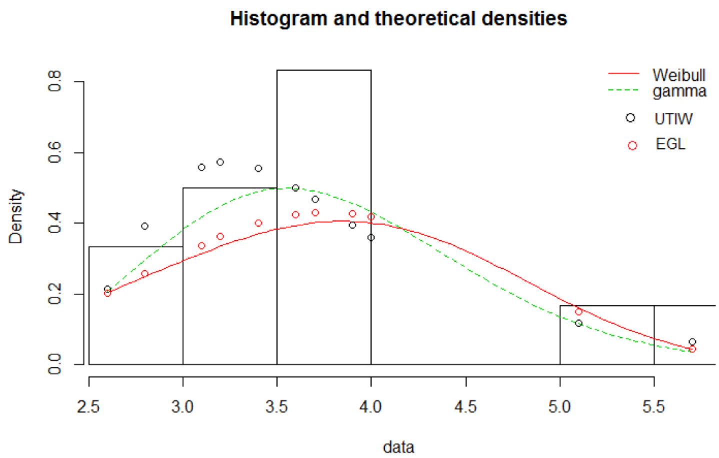
- In all years, the UTIW distribution has the least AIC, BIC and CAIC.
- In all years, except for 2015, the UTIW distribution has the least RMSE and the highest R^2.
- According to the KS and P-value, except for 2015, in all year the least KS and the highest P-value belong to the UTIW distribution.

Figures 4, 5, 6 and 7 represent histogram and theoretical densities and empirical and theoretical cdfs of UTIW, EGL, W and G for wind speed data measured in Ardabil from 2013 to 2016. Based on these Figures, it is seen that the UTIW distribution is the best model.

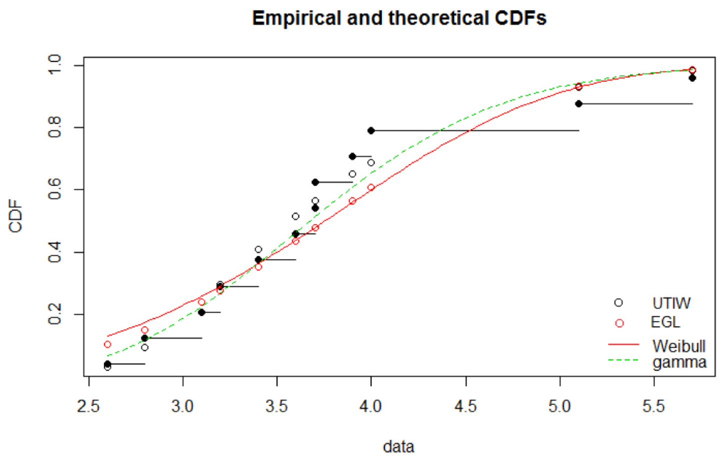
7 Conclusion

The main results of this research can be summarized as follows:

- In this paper for the first time, the UTIW distribution was introduced for modeling wind speed data.
- Four methods for estimating parameters were introduced and these methods were compared by using MC simulation. Finally, the MLE method was chosen as the most appropriate estimation method.

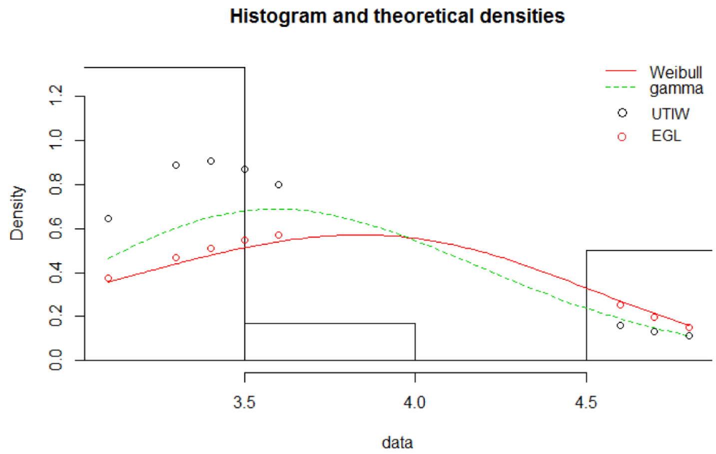


(A) Histogram and theoretical densities for wind speed data in 2013.

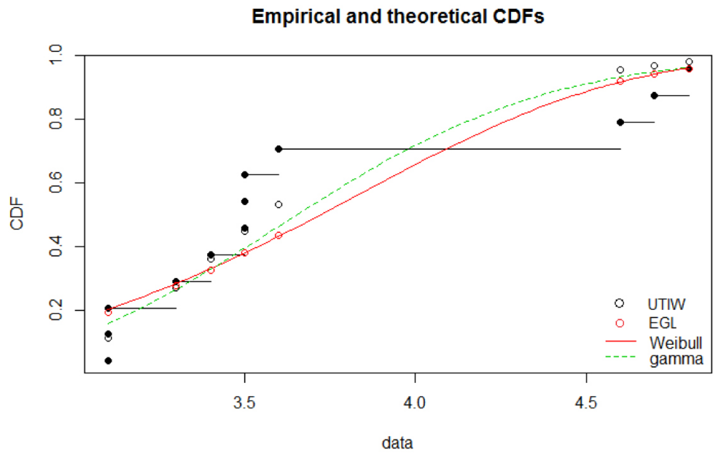


(B) Empirical and theoretical CDFs for wind speed data in 2013.

Figure 4. Histogram, theoretical densities and empirical and theoretical cdfs of the UTIW, EGL, W and G for wind speed data measured in Ardabil in year 2013.

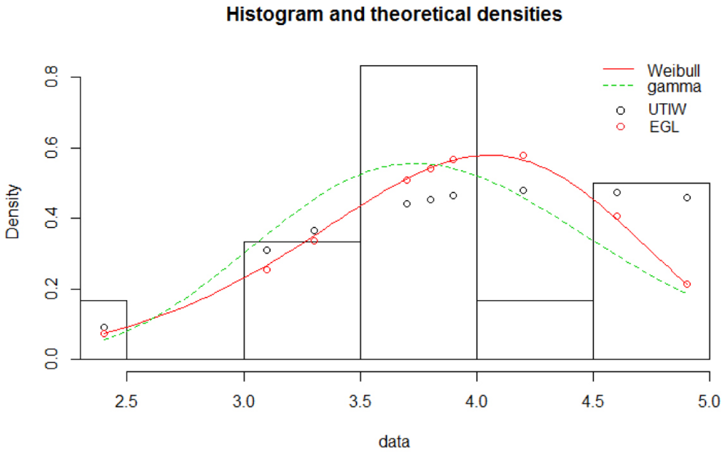


(A) Histogram and theoretical densities for wind speed data in 2014.

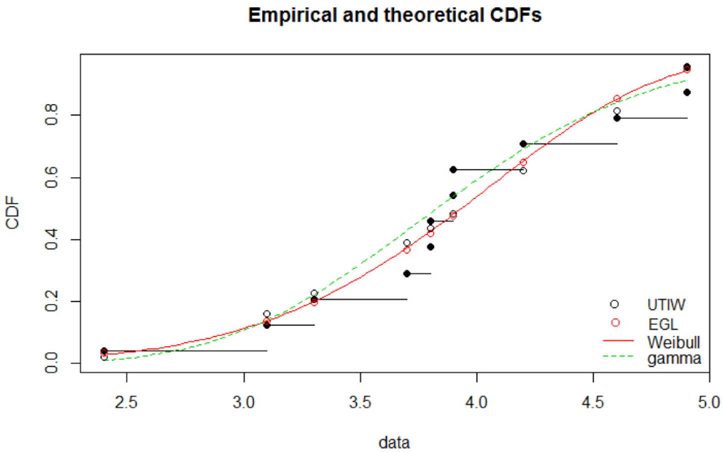


(B) Empirical and theoretical CDFs for wind speed data in 2014.

Figure 5. Histogram, theoretical densities and empirical and theoretical cdfs of the UTIW, EGL, W and G for wind speed data measured in Ardabil in year 2014.

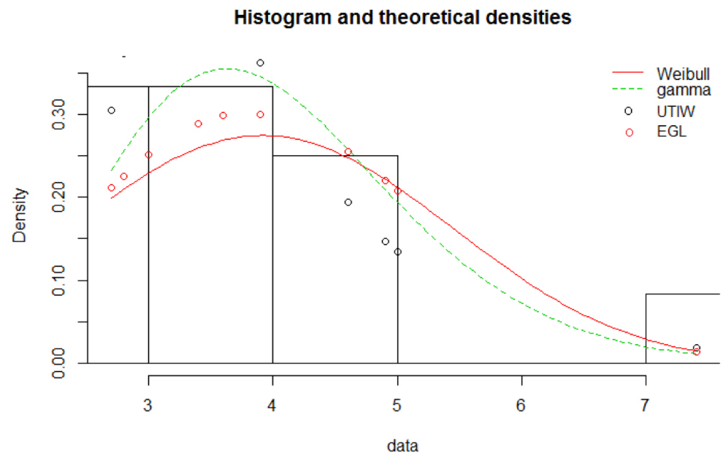


(A) Histogram and theoretical densities for wind speed data in 2015.

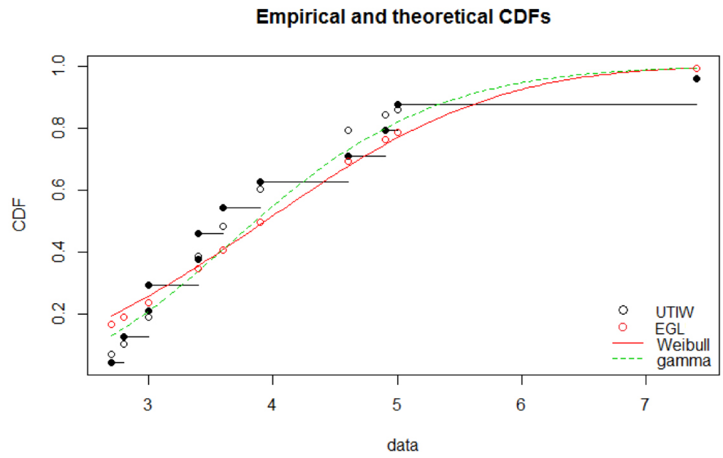


(B) Empirical and theoretical CDFs for wind speed data in 2015.

Figure 6. Histogram, theoretical densities and empirical and theoretical cdfs of the UTIW, EGL, W and G for wind speed data measured in Ardabil in year 2015.



(A) Histogram and theoretical densities for wind speed data in 2016.



(B) Empirical and theoretical CDFs for wind speed data in 2016.

Figure 7. Histogram, theoretical densities and empirical and theoretical cdfs of the UTIW, EGL, W and G for wind speed data measured in Ardabil in year 2016.

- The accuracy of the UTIW distribution versus EGL, W and G distributions was compared by fitting on wind speed data in Ardabil province. Based on criteria such as AIC, BIC, CAIC, L, KS, RMSE and R^2, the superiority of the proposed model was demonstrated in comparison with other models.
- Finally, the UTIW distribution can be used as a new and alternative model for evaluating wind speed data.

References

Akdağ, S., and Güler, Ö. (2009). Calculation of Wind Energy Potential and Economic Analysis by Using Weibull Distribution—a Case Study from Turkey. part 1: Determination of Weibull Parameters. *Energy Sources*, **4**, 1-8.

Akgül, F.G., Şenoğlu, B., and Arslan, T. (2016). An Alternative Distribution to Weibull for Modeling the Wind Speed Data: Inverse Weibull Distribution. *Energy Conversion and Management*, **114**, 234-240.

Akpinar, S., and Akpinar, E.K. (2007). Wind Energy Analysis Based on Maximum Entropy Principle (mep)-type Distribution Function. *Energy Conversion and Management*, **48**, 1140-1149.

Akpinar, S., and Akpinar, E.K. (2009). Estimation of Wind Energy Potential Using Finite Mixture Distribution Models. *Energy Conversion and Management*, **50**, 877-884.

Arslan, T., Acitas, S., and Senoglu, B. (2017). Generalized Lindley and Power Lindley Distributions for Modeling the Wind Speed Data. *Energy Conversion and Management*, **152**, 300-311.

Aydin, and Demet, (2018). The Doubly-truncated Exponentiated Inverse Weibull Distribution Anadolu University of Sciences & Technology-B: Theoretical Sciences **6**,

Ayodele, T., Jimoh, A., Munda, J., and Agee, J. (2012). Wind Distribution and Capacity Factor Estimation for Wind Turbines in the Coastal Region of South Africa. *Energy Conversion and Management*, **64**, 614-625.

Brano, VL., Orioli, A., Ciulla, G., and Culotta, S. (2011). Quality of Wind Speed Fitting Distributions for the Urban Area of Palermo, Italy. *Renewable Energy*, **36**, 1026-1039.

Carta, J.A., Ramírez, P., and Velázquez, S. (2008). Influence of The Level of Fit of a Density Probability Function to Wind-Speed Data on the Wecs Mean Power Output Estimation. *Energy Conversion and Management*, **49**, 2647-2655.

Celik, A.N. (2003). Energy Output Estimation for Small-Scale Wind Power Generators Using Weibull-Representative Wind Data. *Journal of Wind Engineering and Industrial Aerodynamics*, **91**, 693-707.

Dabbaghiyan, A., Fazelpour, F., Abnavi, M.D., and Rosen, M.A. (2016). Evaluation of Wind Energy Potential in Province of Bushehr, Iran. *Renewable and Sustainable Energy Reviews*, **55**, 455-466.

De Gusmao, FR., Ortega, E.M., and Cordeiro, G.M. (2011). The Generalized Inverse Weibull Distribution. *Statistical Papers*, **52**, 591-619.

Drapella, A. (1993). The Complementary Weibull Distribution: Unknown or Just Forgotten? *Quality and Reliability Engineering International*, **9**, 383-385.

Fyrippis, I., Axaopoulos, P.J., and Panayiotou, G. (2010). Wind Energy Potential Assessment in Naxos Island, Greece. *Applied Energy*, **87**, 577-586.

Huang, J., and McElroy, M.B. (2015). A 32-Year Perspective on the Origin of Wind Energy in a Warming Climate. *Renewable Energy*, **77**, 482-492.

Ilkiliç, C., and Aydin, H. (2015). Wind Power Potential and Usage in the Coastal Regions of Turkey. *Renewable and Sustainable Energy Reviews*, **44**, 78-86.

Jiang, R., Murthy, D., and Ji, P. (2001). Models Involving Two Inverse Weibull Distributions. *Reliability Engineering & System Safety*, **73**, 73-81.

Kantar, Y.M., and Usta, I. (2008). Analysis of Wind Speed Distributions: Wind Distribution Function Derived from Minimum Cross Entropy Principles as Better Alternative to Weibull Function. *Energy Conversion and Management*, **49**, 962-973.

Kantar, Y.M., and Usta, I. (2015). Analysis of the Upper-Truncated Weibull Distribution for Wind Speed. *Energy Conversion and Management*, **96**, 81-88.

Kantar, Y.M., Usta, I., Arik, I., and Yenilmez, I. (2018). Wind Speed Analysis Using the Extended Generalized Lindley Distribution. *Renewable Energy*, **118**, 1024-1030.

Kaplan, Y.A. (2015). Overview of Wind Energy in the World and Assessment of Current Wind Energy Policies inTurkey. *Renewable and Sustainable Energy Reviews*, **43**, 562-568.

Keller, A., and Kamath, A. (1982). Alternate Reliability Models for Mechanical Systems. *ESA Reliability and Maintainability*, p 411-415(SEE N 83-20178 10-38)

Keller, A., Kamath, A., and Perera, U. (1982). Reliability Analysis of Cnc Machine Tools. *Reliability Engineering*, **3**, 449-473.

Keyhani, A., Ghasemi-Varnamkhasti, M., Khanali, M., and Abbaszadeh, R. (2010). An Assessment of Wind Energy Potential as a Power Generation Source in the Capital of Iran, Tehran. *Energy*, **35**, 188-201.

Khan, M.S., Pasha, G., and Pasha, A.H. (2008). Theoretical Analysis of Inverse Weibull Distribution. *WSEAS Transactions on Mathematics*, **7**, 30-38.

Köse, R. (2004). An Evaluation of Wind Energy Potential as a Power Generation Source in Kütahya, Turkey. *Energy Conversion and Management*, **45**, 1631-1641.

Morgan, E.C., Lackner, M., Vogel, R.M., and Baise, L.G. (2011). Probability Distributions for Offshore Wind Speeds. *Energy Conversion and Management*, **52**, 15-26.

Morgan, V.T. (1995). Statistical Distributions of Wind Parameters at Sydney, Australia. *Renewable Energy*, **6**, 39-47.

Mudholkar, G.S., Srivastava, D.K., and Kollia, G.D. (1996). A Generalization of the Weibull Distribution with Application to the Analysis of Survival Data. *Journal of the American Statistical Association*, **91**, 1575-1583.

Murthy, D.P., Bulmer, M., and Eccleston, J.A. (2004). Weibull Model Selection for Reliability Modelling. *Reliability Engineering & System Safety*, **86**, 257-267.

Ouammi, A., Dagdougui, H., Sacile, R., and Mimet, A. (2010). Monthly and Seasonal Assessment of Wind Energy Characteristics at Four Monitored Locations in Liguria Region (Italy). *Renewable and Sustainable Energy Reviews*, **14**, 1959-1968.

Ouarda, T.B., and Charron, C. (2018). On the Mixture of Wind Speed Distribution in a Nordic Region. *Energy Conversion and Management*, **174**, 33-44.

Ouarda, T.B., Charron, C., Shin, J.Y., Marpu, P.R., Al-Mandoos, A.H., Al-Tamimi, M.H., Ghedira, H., and Al Hosary, T. (2015). Probability Distributions of Wind Speed in the Uae. *Energy Conversion and Management*, **93**, 414-434.

Ozgur, M., Arslan, O., Kose, R., and Peker, K. (2009). Statistical Evaluation of Wind Characteristics in Kutahya, Turkey. *Energy Sources*, **31**, 1450-1463.

Pasari, S., and Dikshit, O. (2014). Impact of Three-Parameter Weibull Models in Probabilistic Assessment of Earthquake Hazards. *Pure and Applied Geophysics*, **171**, 1251-1281.

Qin, X., Zhang, J.S., and Yan, X.D. (2012). Two Improved Mixture Weibull Models for the Analysis of Wind Speed Data. *Journal of Applied Meteorology and Climatology*, **51**, 1321-1332.

Ramírez, P., and Carta, J.A. (2006). The Use of Wind Probability Distributions Derived from the Maximum Entropy Principle in the Analysis of Wind Energy. a Case Study. *Energy Conversion and Management*, **47**, 2564-2577.

Rubinstein, R.Y., and Kroese, D.P. (2016). Simulation and the Monte Carlo Method, vol 10. John Wiley & Sons

Safari, B., and Gasore, J. (2010). A Statistical Investigation of Wind Characteristics and Wind Energy Potential Based on the Weibull and Rayleigh Models in Rwanda. *Renewable Energy*, **35**, 2874-2880.

Shamilov, A., Kantar, Y.M., and Usta, I. (2008). Use of Minmaxent Distributions Defined on Basis of Maxent Method in Wind Power Study. *Energy Conversion and Management*, **49**, 660-677.

Soukissian, T. (2013). Use of Multi-Parameter Distributions for Offshore Wind Speed Modeling: The Johnson Sb Distribution. *Applied Energy*, **111**, 982-1000.

Wang, J., Qin, S., Jin, S., and Wu, J. (2015). Estimation Methods Review and Analysis of Offshore Extreme Wind Speeds and Wind Energy Resources. *Renewable and Sustainable Energy Reviews*, **42**, 26-42.

Weisser, D. (2003). A Wind Energy Analysis of Grenada: An Estimation Esing the 'Weibull'Density Function. *Renewable Energy*, **28**, 1803-1812.

Yip, C.M.A., Gunturu, U.B., and Stenchikov, G.L. (2016). Wind Resource Characterization in the Arabian Peninsula. *Applied Energy*, **164**, 826-836.

Zhang, T., and Xie, M. (2011). On the Upper Truncated Weibull Distribution and Its Reliability Implications. *Reliability Engineering & System Safety*, **96**, 194-200.

Gholamhossein Yari
Department of Statistics,
Iran University of Science and Technology,
Tehran, Iran.
email: *yari@iust.ac.ir*

Zahra Karimi Ezmareh
Department of Statistics,
Iran University of Science and Technology,
Tehran, Iran.
email: *z_ karimi@mathdep.iust.ac.ir*