

Wind speed analysis with the upper truncated quasi Lindley distribution

Emrah Altun^{1,*} and Gamze Ozel²

Abstract

We propose a new distribution with extra shape parameter named a truncated quasi Lindley distribution which is more flexible than many well-known distributions. Mathematical and statistical properties of truncated quasi Lindley distribution are given only for upper truncated version of quasi Lindley distribution to reduce the mathematical complexity. An important property of the upper truncated quasi Lindley distribution is that it can have bathtub-shaped failure rate function. We present some of its mathematical properties including ordinary moments, quantile and moment generating functions, order statistics. The method of maximum likelihood to estimate the model parameters is discussed and the behavior of maximum likelihood estimator is studied. The importance of the new distribution is illustrated by means of the wind speed data and the capability in modeling wind speed is evaluated. The results indicate that the truncated version of quasi Lindley distribution can provide better fits than Exponential,

¹ Department of Statistics, Hacettepe University, 06800, Ankara, Turkey.
E-mail: emrahaltun@hacettepe.edu.tr (E. Altun).

² Department of Statistics, Hacettepe University, 06800, Ankara, Turkey.

* Corresponding author.

Lindley, quasi Lindley and Weibull distributions in estimating wind speed distribution. Therefore, truncated quasi Lindley can be an alternative for use in the assessment of wind energy potential.

Mathematics Subject Classification: 60E05; 60E10; 62G30; 62P30

Keywords: Truncated Quasi Lindley Distribution; Maximum Likelihood Estimation; Quantile Function, Order Statistics; Wind Speed

1 Introduction

The Lindley was introduced by Lindley [1] as a new distribution useful to analyze failure time data especially in applications modeling stress-strength reliability. The motivation of the Lindley distribution arises from its ability to model failure time data with increasing, decreasing, unimodal and bathtub shaped hazard rates. The distribution represents a good alternative to the exponential failure time distribution that suffer from not exhibiting unimodal and bathtub shaped failure rates [2]. Ghitany et al. [3] obtained some properties of the Lindley distribution and showed that the Lindley distribution provides a better modeling for some applications than the exponential distribution. Mazucheli and Achcar [4] also found that many of the mathematical properties are more flexible than those of the exponential distribution and proposed the Lindley distribution as a possible alternative to the exponential distribution. The probability distribution function (pdf) and cumulative distribution function (cdf) of the Lindley distribution are given by, respectively,

$$f_L(x) = \frac{\theta^2}{1+\theta} (1+x)e^{-\theta x}, \quad x > 0, \theta > 0 \quad (1)$$

$$F_L(x) = 1 - \frac{\theta+1+\theta x}{\theta+1} e^{-\theta x}, \quad x > 0, \theta > 0$$

The Lindley distribution is a mixture of exponential (θ) and gamma ($2, \theta$) distributions with their mixing proportions are $(1/(1+\theta))$ and $(\theta/(1+\theta))$. Let us point out that some researchers proposed new classes of distributions based on modifications of the Lindley distribution, including also their properties. The generalized Lindley distribution was proposed in Zakerzadeh and Dolati [5]. The Poisson-Lindley distribution was introduced by Sankaran [6]. The zero-truncated Poisson-Lindley distribution and generalized Poisson-Lindley distribution were considered in Ghianty et al. [7]. Recently, Shanker and Mishra [8] have introduced a two parameter quasi Lindley (QL) distribution which is the particular case of the Lindley distribution. They showed that the QL distribution provides better fits than the Lindley distribution based on different data sets. The QL distribution is defined by its pdf and cdf as, respectively,

$$f_{QL}(x) = \frac{\theta(\alpha + x\theta)}{\alpha + 1} e^{-\theta x}, \quad x > 0, \theta, \alpha > 0 \quad (2)$$

$$F_{QL}(x) = 1 - \frac{1 + \alpha + \theta x}{\alpha + 1} e^{-\theta x}, \quad x > 0, \theta, \alpha > 0$$

In Eq. (2), the QL distribution reduces to the Lindley and gamma($2, \theta$) distributions for $\alpha = \theta$ and $\alpha = 0$, respectively.

A truncated distribution is defined as a conditional distribution that results from restricting the domain of the statistical distribution. Hence, truncated distributions are used in cases where occurrences are limited to values which lie above or below a given threshold or within a specified range. If occurrences are limited to values which lie below a given threshold, the lower (left) truncated distribution is obtained. Similarly, if occurrences are limited to values which lie above a given threshold, the upper (right) truncated distribution arises [9, 10]. Truncated versions of the well-known statistical distributions are proposed by many researchers to model the truncated data in various fields. Zhang and Xie [11] studied the characteristics of the truncated Weibull distribution and illustrated the applicability of this distribution to modeling lifetime data. Ahmed et al. [12]

proposed the truncated version of the Birnbaum-Saunders (BS) distribution and showed that truncated BS distribution is more appropriate than the classical BS model for describing the financial loss data from a commercial bank. Recently, Singh et al. [13] have introduced the truncated version of the Lindley distribution and discussed statistical properties of proposed distribution and showed that truncated version of the Lindley distribution provides a better modeling than Weibull, Lindley and exponential distributions based on a real data. As far as we know, truncated version of the QL distribution has not been investigated. Therefore, the aim of this study is to obtain a truncated QL distribution which is more flexible than the exponential, Lindley, quasi Lindley, and lower truncated quasi distributions.

Accurately modeling wind speed is critical in estimating the wind energy potential of a certain region. Several statistical distributions have been studied in order to model wind speed data smoothly. The Erlang, inverse normal and Gumbel distributions presented as wind speed distributions in [27], while a generalized extreme value distribution was used in [14]. Chang [15] introduced mixture truncated normal distributions, while Usta and Kantar [16] proposed certain flexible families of distributions as an alternative to the exponential distribution in estimating wind speed distribution. Philippopoulos et al. [17] compares various distributions and shows that the gamma distribution could be an efficient alternative to the exponential distribution. Consequently, the mentioned studies [18–21] emphasize that exponential distribution does not present good performance in the modeling of wind speed data in comparison for all wind types encountered in nature, such as low or high, skewed or kurtotic, or skewed and kurtotic wind speed. Thus, in order to minimize errors in wind speed estimation, it is necessary to select the most appropriate distribution for the description of wind speed measured for a specific area. An upper truncated distribution is applicable to this situation where the range of random variable is bounded from above by an unknown cut-off point, called a truncation point. In other words, if the values of

random variables are observed in the interval $[0, T]$, then the upper truncated distribution can be used. The wind speed measurements are generally observed in the range of $[0, V]$, and therefore, the upper truncated distribution can be applied to model the wind speed data. Therefore, this paper also proposes, for the first time, the use of the upper truncated QL (UTQL) distribution, in modeling wind speed data. In addition, a comparison is made between UTQL distribution and well known distributions using wind speed data.

The paper is organized as follows: Section 2 introduces the probability density, cumulative density, survival and hazard functions of UTQL distribution. Then the plots of the proposed distribution for several values of parameters are presented. In Section 3, some statistical properties of the UTQL distribution, such as ordinary moments, moment generating and quantile functions, skewness and kurtosis, order statistics are derived and mean and variance for the several values of the parameters are obtained. Section 4 provides the maximum likelihood estimation of the model parameters. The wind speed data is used to evaluate the performance of truncated versions of the UTQL distribution in Section 5. Section 6 concludes the study with the obtained results and certain suggestions for further research.

2 Upper Truncated Quasi Lindley Distribution

2.1. The probability density and cumulative density functions

The cdf of a double truncated distribution is given by

$$G(x) = \frac{F(x) - F(v)}{F(\zeta) - F(v)}, \quad v < x < \zeta, \quad -\infty < v < \zeta < \infty \quad (3)$$

where v and ζ are the intervals of the truncated distribution. The pdf of the truncated distribution is defined as

$$g(x) = \frac{f(x)}{F(\zeta) - F(v)}, \quad v < x < \zeta, \quad -\infty < v < \zeta < \infty \quad (4)$$

where $f(x)$ and $F(x)$ are the pdf and the cdf of the baseline distribution, respectively. It is easily seen from Eq. (4) that the distribution reduces the baseline distribution for $v = 0$ and $\zeta \rightarrow \infty$.

Eq. (4) is called as the upper truncated distribution for any baseline distribution for $v = 0$. If $f_{QL}(x)$ in Eq. (2) is considered as a baseline distribution and using the Eqs. (3) and (4), we obtain the pdf and the cdf of the UTQL distribution respectively as follows:

$$g_{UTQL}(x) = \frac{\theta(\alpha + \theta x)e^{-\theta x}}{(1 + \alpha)(1 - e^{-\theta \zeta}) + \zeta \theta e^{-\zeta \theta}}, \quad x > 0, \quad \theta, \alpha > 0, \quad 0 < x < \zeta \quad (5)$$

$$G_{UTQL}(x) = \frac{e^{\theta(\zeta - x)} \left((1 + \alpha)(1 - e^{-\theta x}) + \theta x \right)}{(1 + \alpha)(1 - e^{-\zeta \theta}) + \zeta \theta}, \quad x > 0, \quad \theta, \alpha > 0, \quad 0 < x < \zeta \quad (6)$$

The density function in Eq. (5) is much more flexible than the Lindley and quasi Lindley density functions. Thus it can allow for greater flexibility of the tails. It can exhibit different behavior depending on the parameter values. Note that the probability density in Eq. (5) is reduced the quasi Lindley distribution for $\zeta = 0$.

To obtain the mode of the UTQL distribution, we give the first derivate of (5) as

$$\frac{\partial g_{UTQL}(x)}{\partial x} = \frac{e^{-\theta x} \theta^2 (\alpha + \theta x - 1)}{(\alpha + 1)(e^{-\zeta \theta} - 1) + \zeta \theta e^{-\zeta \theta}} \quad (7)$$

Solving the Eq. (7) for zero point, gives $x = \frac{1 - \alpha}{\theta}$. When $|\alpha| < 1$, $x = \frac{1 - \alpha}{\theta}$ is

the maximum point of the distribution. Then, the mode of the UTQL distribution can be defined as

$$\text{Mode} = \begin{cases} \frac{1 - \alpha}{\theta}, & |\alpha| < 1 \\ 0, & \text{o/w} \end{cases} \quad (8)$$

Figure 1 illustrates some of the possible shapes of the pdf of the UTQL distribution for selected values of the parameters.

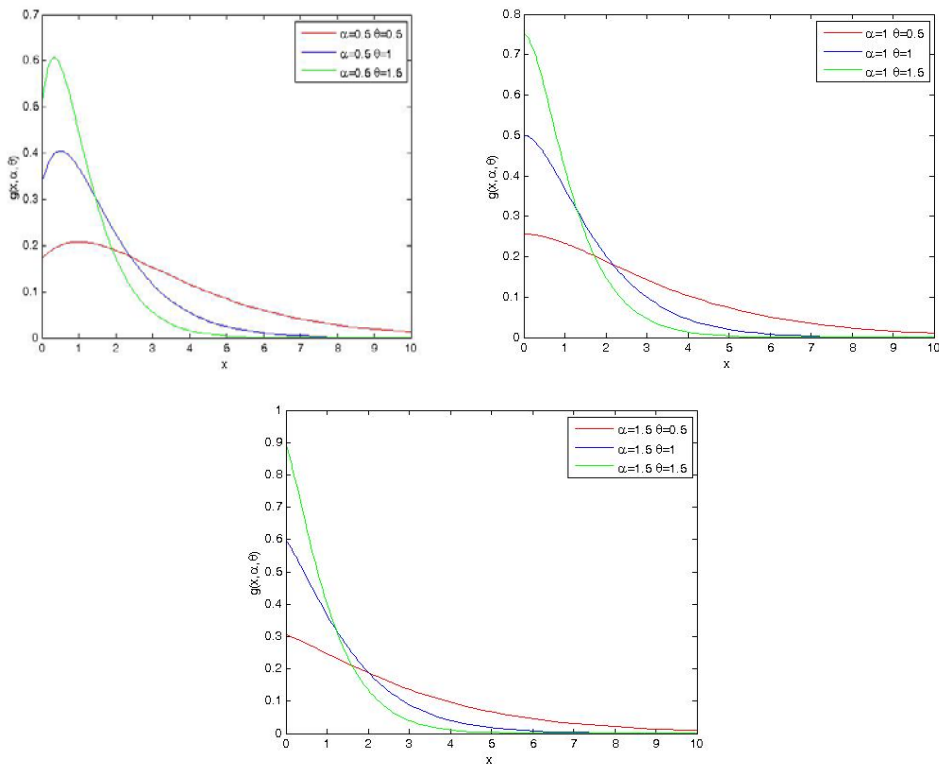


Figure 1: Plots of the UTQL density functions for given $\theta = 0.5, 1, 1.5$ $\alpha = 0.5, 1, 1.5$ and $\zeta=10$

As seen from Figure 1, decreasing θ parameter yields to fat-tail structure and θ parameter specifies the shape of the distribution. In fact, plots in Figure 1 reveals that the mode of the pdf increases as θ . It is evident that the UTQL distribution is much more flexible than the Lindley distribution, i.e. the additional shape parameter θ allow for a high degree of flexibility of the UTQL distribution. Hence, the new model can be very useful in many practical situations for modeling positive real data sets such as wind speed data.

2.2. Survival and Hazard Rate Functions

Central role is playing in the reliability theory by the ratio of the probability

density and survival functions. The survival function of the UTQL distribution is obtained from Eq. (6) as

$$S(x) = 1 - G_{\text{UTQL}}(x) = 1 - \left\{ \frac{e^{\theta(\zeta-x)} \left((1+\alpha)(1-e^{\theta x}) + \theta x \right)}{(1+\alpha)(1-e^{\zeta\theta}) + \zeta\theta} \right\} \quad (9)$$

Figure 2 represents plots of the survival function for the UTQL distribution for several parameter values α , θ and ζ .

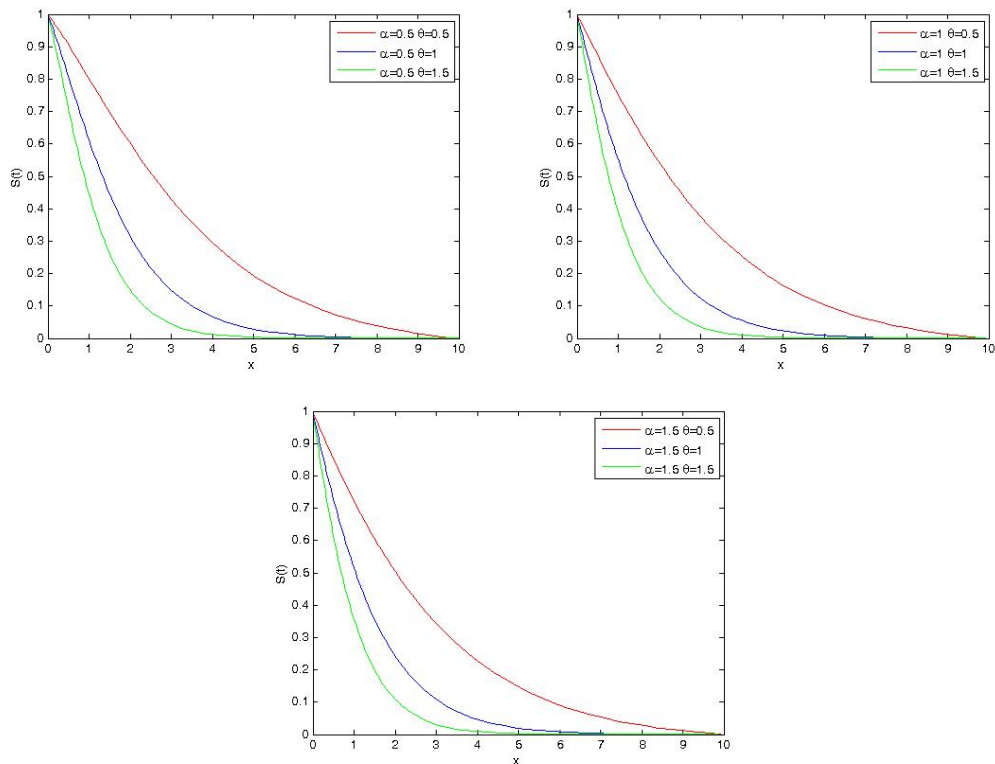


Figure 2: Plots for the survival functions of UTQL distribution for given $\theta = 0.5, 1, 1.5$ $\alpha = 0.5, 1, 1.5$ and $\zeta = 10$

The other characteristic of interest of a random variable is the hazard rate function. It can be loosely interpreted as the conditional probability of failure, given it has survived to time t . We obtain the hazard function of the UTQL distribution as

$$h(x) = \frac{g_{UTQL}(x)}{1 - G_{UTQL}(x)} = \frac{\theta(\alpha + \theta x)e^{-\theta x}}{(1 + \alpha)(1 - e^{-\theta\zeta}) + \zeta\theta e^{-\zeta\theta}} \left[\frac{e^{\theta(\zeta-x)}((1 + \alpha)(1 - e^{\theta x}) + \theta x)}{(1 + \alpha)(1 - e^{\theta\zeta}) + \zeta\theta} \right]$$

Figure 3 illustrates the hazard rate function of the UTQL distribution for different values of the parameters α , θ and ζ .

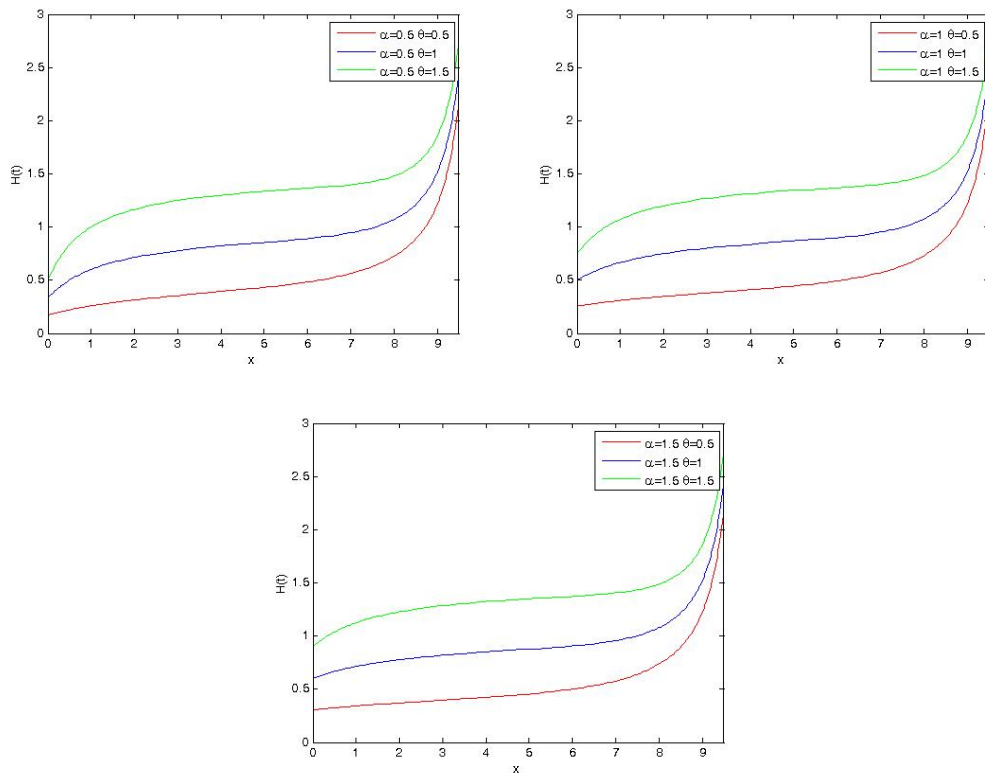


Figure 3: Plots for the hazard functions of the UTQL distribution for given $\theta = 0.5, 1, 1.5$ $\alpha = 0.5, 1, 1.5$ and $\zeta = 10$

Figure 3 shows that the hazard rate function of the UTQL distribution is flexible for several values of parameters. It is seen in Figure 3 that the hazard rate function are increasing.

3 Main Properties

We derive computational sum-representations and explicit expressions for the ordinary and central moments, skewness, kurtosis, generating and quantile functions, order statistics of X . These expressions can be evaluated analytically or numerically using packages such as Mathematica, Matlab and Maple.

3.1. Moments

Some key features of a distribution such as skewness and kurtosis can be studied through its moments. We derive closed-form expressions for the ordinary moments, generating function, skewness and kurtosis of X . We obtain the ordinary moments of the UTQL distribution as

$$E(X^r) = \frac{\theta}{(\alpha+1)F(\zeta)} \int_0^{\zeta} x^r (\alpha + \theta x) e^{-\theta x} dx \quad (10)$$

$$= \frac{\Gamma(r+1) - \Gamma(r+1, \zeta\theta)}{\theta^r (\alpha+1)} \frac{1}{F(\zeta)}$$

where $\Gamma(\cdot)$ and $\Gamma(\cdot, \cdot)$ are gamma and upper incomplete gamma functions, respectively, and defined as follows:

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx, \quad (11)$$

$$\Gamma(s, t) = \int_t^{\infty} x^{s-1} e^{-x} dx.$$

From Eq. (10), the mean and variance of the UTQL distribution are obtained using the first and the second ordinary moments as

$$E(X) = \mu = \frac{\alpha - 2e^{\zeta\theta} + 2\zeta\theta - \alpha e^{\zeta\theta} + \zeta^2\theta^2 + \alpha\zeta\theta + 2}{\theta(\alpha - e^{\zeta\theta} + \zeta\theta - \alpha e^{\zeta\theta} + 1)}, \quad (12)$$

$$E(X^2) = \frac{2\alpha - 6e^{\zeta\theta} + 6\zeta\theta - 2\alpha e^{\zeta\theta} + 3\zeta^2\theta^2 + \zeta^3\theta^3 + \alpha\zeta^2\theta^2 + 2\alpha\zeta\theta + 6}{\theta^2(\alpha - e^{\zeta\theta} + \zeta\theta - \alpha e^{\zeta\theta} + 1)}$$

The mean and variance (σ^2) of the UTQL distribution are presented in Table 1 for different values of parameters α , θ and ζ . Note that variance $V(X) = \sigma^2 = E(X^2) - \{E(X)\}^2$ is found from Eq. (12).

Table 1: Mean and variance of the UTQL distribution for several values of the parameters

θ	α	$\xi = 5$		$\xi = 10$		$\xi = 15$	
		μ	σ^2	μ	σ^2	μ	σ^2
0.1	0.25	2.7219	1.9798	5.3677	7.6760	7.7444	16.9309
	0.75	2.5026	2.0655	4.8594	8.1311	6.9664	17.8787
	1	2.4599	2.0711	4.7396	8.1631	6.7633	17.9269
0.25	0.25	2.6373	1.8994	4.6299	7.1384	5.8916	14.1595
	0.75	2.3789	2.0116	4.1306	7.3616	5.2129	14.1105
	1	2.3144	2.0188	3.9864	7.3333	5.0068	13.9133
0.5	0.25	2.3149	1.7846	3.2960	5.1894	3.5470	7.1269
	0.75	2.0653	1.8404	2.8986	5.0279	3.1016	6.6901
	1	1.9932	1.8333	2.7757	4.9140	2.9626	6.4725
1	0.25	1.6480	1.2973	1.7962	1.9250	1.7999	1.9592
	0.75	1.4493	1.2570	1.5685	1.7889	1.5714	1.8157
	1	1.3879	1.2285	1.4974	1.7251	1.5000	1.7495

Table 1 shows that the mean and variance decrease when θ increases for a fixed α . Besides, for a fixed θ , while mean decreases, variance increases when the α increases. Assuming α and θ is kept fix, the mean and variance increase when ξ increases.

Further, the central moments (μ_r) and cumulants (κ_r), $r = 1, 2, \dots$, of the UTQL distribution can be obtained from

$$\mu_r = \sum_{k=0}^r (-1)^k \binom{r}{k} \mu_1^r \mu'_{r-k} \quad \text{and} \quad \kappa_r = \mu'_r - \sum_{k=0}^r (-1)^k \binom{r}{k} \mu_1^r \mu'_{n-r} \quad (13)$$

Here $\kappa_1 = \mu'_1$, $\kappa_2 = \mu'_2 - \mu_1'^2$, $\kappa_3 = \mu'_3 - 3\mu'_2\mu_1' + 2\mu_1'^3$, $\kappa_4 = \mu'_4 - 4\mu'_3\mu_1' - 3\mu_2'^2 + 12\mu'_2\mu_1'^2 - 6\mu_1'^4$ etc. The skewness $\gamma_1 = \kappa_3 / \kappa_2^{3/2}$ and kurtosis $\gamma_2 = \kappa_4 / \kappa_2^2$ are also computed from the second, third and fourth cumulants.

3.2. Moment Generating Function

The moment generating function (mgf) is widely used as an alternative way to analytical results compared with working directly with pdf and cdf. Here, we give a formula for the mgf $M(t) = E(e^{tx})$ of X as

$$M_x(t) = \frac{\alpha\theta(e^{\xi t - \xi\theta} - 1)}{(\theta - t)((\alpha + 1)(e^{-\xi\theta} - 1) + \xi\theta e^{-\xi\theta})} - \frac{\theta^2(1/(\theta - t)^2 - (e^{\xi t - \xi\theta}(\xi\theta - \xi t + 1))/(\theta - t)^2)}{(\alpha + 1)(e^{-\xi\theta} - 1) + \xi\theta e^{-\xi\theta}} \quad (14)$$

Using the Eq. (14), the ordinary moments of the UTQL distribution can be also calculated.

3.3. Quantile Function

The quantile function of a probability distribution is the inverse of the cdf and defined as follows:

$$\begin{aligned} Q(p) &= \inf\{x \in \mathbb{R} : p \leq F(x)\} \\ &= F^{-1}(x) \end{aligned} \quad (15)$$

We obtain the quantile function of the UTQL distribution from Eqs. (6) and

(15) as

$$Q_{\text{UTQL}}(p) = \left(\frac{1 + \alpha + W\left((1 + \alpha)\left[e^{\xi\theta}(1 - p) + p\right]e^{-(1 + \alpha + \xi\theta)}\right)}{2\theta} \right) \quad (16)$$

where $W(\cdot)$ is a Lambert-W function which is a multi-valued complex function defined as the solution of the equation

$$W(z)\exp[W(z)] = z \quad (17)$$

Here, z is a complex number [12].

Simulating UTQL distribution random variable is straightforward. Let p be a uniform variable on the unit interval $(0, 1)$. Thus, by means of the inverse transformation method, the random variable X given by

$$X = \left(\frac{1 + \alpha + W\left((1 + \alpha)\left[e^{\xi\theta}(1 - p) + p\right]e^{-(1 + \alpha + \xi\theta)}\right)}{2\theta} \right) \quad (18)$$

3.4. Skewness and Kurtosis

The effects of the parameters on the skewness and kurtosis of X can be based on quantile function in Eq. (16). There are many heavy tailed distributions for which this measure is infinite. So, it becomes uninformative precisely when it needs to be. The Bowley's skewness is based on quartiles:

$$S = -\frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)}$$

and the Moors' kurtosis is based on octiles:

$$K = -\frac{Q(7/8) - Q(5/8) - Q(3/8) + Q(1/8)}{Q(6/8) - Q(2/8)},$$

where $Q(\cdot)$ represents the quantile function of X . These measures are less sensitive to outliers and they exist even for distributions without moments. Skewness measures the degree of the long tail and kurtosis is a measure of the

degree of tail heaviness. When the distribution is symmetric, $S = 0$ and the when the distribution is right (or left) skewed, $S > 0$ (or $S < 0$). As K increases, the tail of the distribution becomes heavier. From Eq. (16), skewness and kurtosis of the UTQL distribution are obtained and presented in Table 2.

Table 2: Skewness and kurtosis of the UTQL distribution for several values of the parameters

θ	α	$\xi = 5$		$\xi = 10$		$\xi = 15$	
		Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
0.1	0.25	0.0207	0.0414	0.0354	0.0712	0.0453	0.0914
	0.75	0.0199	0.0398	0.0318	0.0639	0.0406	0.0817
	1	0.0174	0.0348	0.0294	0.059	0.0374	0.0752
0.25	0.25	0.0409	0.0823	0.0558	0.1131	0.0605	0.1227
	0.75	0.0391	0.0787	0.0498	0.1005	0.0538	0.1089
	1	0.0338	0.068	0.0457	0.0922	0.0494	0.0997
0.5	0.25	0.0558	0.1131	0.0618	0.1256	0.0624	0.1266
	0.75	0.0533	0.1078	0.055	0.1113	0.0554	0.1122
	1	0.0457	0.0922	0.0504	0.1019	0.0508	0.1027
1	0.25	0.0618	0.1256	0.0624	0.1267	0.0624	0.1267
	0.75	0.055	0.1113	0.0555	0.1123	0.0555	0.1123
	1	0.0504	0.1019	0.0509	0.1027	0.0509	0.1028

The values in Table 2 indicate a narrow range for the skewness of X , similarly, the kurtosis does not vary much.

3.5. Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose X_1, X_2, \dots, X_n is a random sample from the UTQL distribution. Let $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ be a ordered statistics and $g_{s:n}(t)$ represents the pdf of the $X_{s:n}$

$$g_{s:n}(t) = \frac{1}{B(s, n-s+1)} \sum_{i=1}^{n-s} (-1)^i \binom{n-s}{i} \left(\frac{G_{UTQL}(t, \theta)}{G_{UTQL}(\xi, \theta)} \right)^{s+i} \frac{g_{UTQL}(t, \theta)}{G_{UTQL}(t, \theta)}$$

where $B(s, n-s+1)$ is beta function, $g_{UTQL}(t)$ and $G_{UTQL}(t)$ are the pdf and cdf of the UTQL distribution, respectively.

For $s=1$, the pdf of the first order statistics is defined as

$$g_{1:n}(t) = n \sum_{i=1}^{n-s} (-1)^i \frac{\left(\frac{\theta(\alpha + \theta t)e^{-\theta t}}{(1+\alpha)(1-e^{-\theta \zeta}) + \zeta \theta e^{-\zeta \theta}} \right) \left(\frac{e^{\zeta \theta} e^{-\theta t} ((1+\alpha)(1-e^{\theta t}) + \theta t)}{(1+\alpha)(1-e^{\zeta \theta}) + \zeta \theta} \right)^i}{B(n-i, i+1) \left(\frac{e^{\zeta \theta} e^{-\theta \gamma} ((1+\alpha)(1-e^{\theta \gamma}) + \theta \gamma)}{(1+\alpha)(1-e^{\zeta \theta}) + \zeta \theta} \right)^{i+1}} \quad (19)$$

Similarly, the pdf of $X_{n:n}$ can be obtained by taking $s=n$.

4 Parameter Estimation

Several approaches for parameter point estimation were proposed in the literature but the maximum likelihood estimation (MLE) method is the most commonly employed. The MLEs enjoy desirable properties and can be used when constructing confidence intervals and regions and also in test statistics. In this section, the parameter estimation of the UTQL distribution is obtained using the maximum likelihood estimation (MLE) procedure. Let x_1, x_2, \dots, x_n be a random sample from the UTQL distribution. The likelihood function of the UTQL

distribution is given as

$$L(\alpha, \theta, \zeta | \mathbf{x}) = \left[\frac{\theta}{(1+\alpha)(1-e^{\theta\zeta}) - \theta\zeta e^{-\theta\zeta}} \right]^n \prod_{i=1}^n (\alpha + x_i \theta) e^{-\theta \sum_{i=1}^n x_i} \quad (20)$$

and the log-likelihood function is obtained as

$$\ln L(\alpha, \theta, \zeta | \mathbf{x}) = n \ln(\theta) - \ln \left[(1+\alpha)(1-e^{\theta\zeta}) - \theta\zeta e^{-\theta\zeta} \right] + \sum_{i=1}^n \ln(\alpha + x_i \theta) - \theta \sum_{i=1}^n x_i \quad (21)$$

The log-likelihood function can be maximized either directly by using the SAS (PROCNLMIXED) or the Ox program (sub-routine MaxBFGS) or by solving the nonlinear likelihood equations obtained by differentiating Eq. (21). Here, the log-likelihood function is maximized by solving nonlinear equations obtained by differentiating the log-likelihood function. The first derivatives of the log-likelihood function with respect to the parameters are

$$\frac{\partial \ln L(\alpha, \theta, \zeta | \mathbf{x})}{\partial \theta} = \frac{n}{\theta} + n \left[\frac{-\xi e^{\xi\theta} (1+\alpha) - \xi e^{-\xi\theta} (1+\xi\theta)}{(1+\alpha)(1-e^{\xi\theta}) - \xi\theta e^{\xi\theta}} \right] + \sum_{i=1}^n \frac{x_i}{\alpha + x_i \theta} - \sum_{i=1}^n x_i \quad (22)$$

$$\frac{\partial \ln L(\alpha, \theta, \zeta | \mathbf{x})}{\partial \alpha} = -n \left[\frac{1-e^{\theta\zeta}}{(1+\alpha)(1-e^{\theta\zeta}) - \theta\zeta e^{-\theta\zeta}} \right] + \sum_{i=1}^n \frac{1}{\alpha + x_i \theta} \quad (23)$$

The given equations can be solved using iterative methods such as Newton Raphson method. The MLE of ξ is taken as $\hat{\xi} = \max \{x_1, x_2, \dots, x_n\}$, $\hat{\xi}$ is the largest value in sample. Let θ_0 and α_0 be the initial values of θ and α , respectively. Then, we have

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \log L}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2} \end{bmatrix} \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \alpha} \end{bmatrix} \quad (24)$$

where

$$\begin{aligned} \frac{\partial^2 \ln L(\alpha, \theta, \zeta | \mathbf{x})}{\partial \theta^2} &= \frac{n}{\theta^2} + \frac{n(\zeta e^{-\zeta \theta} + \zeta e^{\zeta \theta} (\alpha + 1) - \zeta^2 \theta e^{-\zeta \theta})^2}{((\alpha + 1)(e^{\zeta \theta} - 1) + \zeta \theta e^{-\zeta \theta})^2} \\ &\quad - \frac{n(\zeta^3 \theta e^{-\zeta \theta} - 2\zeta^2 e^{-\zeta \theta} + \zeta^2 e^{\zeta \theta} (\alpha + 1))}{((\alpha + 1)(e^{\zeta \theta} - 1) + \zeta \theta e^{-\zeta \theta})} - \sum_{i=1}^n \frac{x_i^2}{(\alpha + \theta x_i)^2}, \end{aligned} \quad (25)$$

$$\frac{\partial^2 \ln L(\alpha, \theta, \zeta | \mathbf{x})}{\partial \alpha^2} = \frac{n(e^{\zeta \theta} - 1)^2}{((\alpha + 1)(e^{\zeta \theta} - 1) + \zeta \theta e^{-\zeta \theta})^2} - \sum_{i=1}^n \frac{1}{(\alpha + \theta x_i)^2}, \quad (26)$$

$$\begin{aligned} \frac{\partial^2 \ln L(\alpha, \theta, \zeta | \mathbf{x})}{\partial \theta \partial \alpha} &= -n \left[\frac{-e^{\zeta \theta} \zeta [(1 + \alpha)(1 - e^{\zeta \theta}) - \zeta \theta e^{-\zeta \theta}] - (1 - e^{\zeta \theta}) [\zeta^2 \theta e^{-\zeta \theta} - e^{\zeta \theta} (\zeta + \alpha)]}{[(1 + \alpha)(1 - e^{\zeta \theta}) - \zeta \theta e^{-\zeta \theta}]^2} \right] \\ &\quad + \sum_{i=1}^n \frac{x_i}{(\alpha + \theta x_i)^2}. \end{aligned} \quad (27)$$

5 Application

In this section, we provide applications to a real data set to demonstrate the potentiality of the truncated versions of quasi Lindley distribution. A set of real test data representing wind speed reported by [22] is used to evaluate the performance of the truncated quasi Lindley distribution. It consists of the observations listed in Table 3.

As seen from Table 3, each data point represents the average wind speed over some time period. Then, the descriptive statistics for the wind speed data are given in Table 4.

Table 4 shows that the wind speed data is positively skewed. . According to kurtosis value, wind speed data distribution is flatter than a normal distribution with a wider peak. This means that the probability for extreme values of wind speed is less than for a normal distribution, and the values are wider spread around the mean. Here, we have fitted the data with the exponential, Lindley, QL, upper, lower and double truncated Quasi Lindley distributions. We obtain the pdf of the lower truncated Quasi Lindley (LTQL) and double truncated Quasi Lindley

(DTQL) distribution as

$$g_{LTQL}(x) = \frac{\theta e^{\theta(v-x)}(\alpha + \theta x)}{(\alpha + \theta v + 1)}; \quad x > v; \quad \theta, \alpha > 0 \quad (28)$$

$$g_{DTQLD} = \frac{-\theta e^{-\theta x}(\alpha + \theta x)}{e^{-\zeta\theta}(\alpha + \zeta\theta + 1) - e^{-v\theta}(\alpha + v\theta + 1)}; \quad v, \zeta, \theta, \alpha > 0 \quad (29)$$

Table 3: Wind speed data (m/s)

3.3	5.7	4	4.5	4.7	6.3
3.8	8.5	4	5.8	4.5	9.4
4.2	8.9	2	4.8	4.2	7.7
3.3	9.3	2.7	4.8	5.7	6
2.8	6.5	2.7	5.5	2.7	8.9
3	4.2	3.3	5.7	4.3	7.7
4	4.3	2.7	5	4.3	6.2
2.7	3.7	2.7	4.3	4.5	5.7
5.2	4	5.8	4	4.5	5.7
6.7	2.8	5.7	3.5	6	7.5
6.8	3.7	6.2	5	10.4	7.5
6.8	3.3	6.5	3.7	6.7	5.3

Table 4: Descriptive statistics for the wind speed data

n	Minimum	Median	Mean	Maximum	Variance	Skewness	Kurtosis
72	2	4.75	5.12	10.4	3.498	0.722	0.041

The log-likelihood function of UTQL and DTQL can be obtained using the Eq. (28) and (29) to estimate the parameters of UTQLD and DTQL distributions.

We estimate the unknown parameters of the distributions by the maximum likelihood. Table 5 represents the fitting summary of these distributions including the estimates of parameters, log-likelihood, Akaike information criterion (AIC), Corrected Akaike information criterion (AICC), and Bayesian information criterion (BIC) which are calculated by given equations,

$$\begin{aligned} \text{AIC} &= -2\log(L) + 2k, \\ \text{AICC} &= \text{AIC} + \frac{2k(k+1)}{(n-k-1)} \\ \text{BIC} &= -2\log(L) + k \log(n) \end{aligned} \quad (30)$$

where k is the number of parameters and n is the sample size.

According to the AIC, AICC and BIC statistics values, the DTQL distribution gives the best fitting to the wind speed data. Figure 4 represents the fitting performance of the DTQL distribution graphically, including the quantile-quantile (Q-Q) plot, probability-probability (P-P) plot, empirical and theoretical densities and empirical and theoretical cumulative distribution functions.

Table 5: Maximum likelihood estimates, AIC, AICC, BIC statistics values under considered distributions based on wind speed data

Distribution	Parameter Estimations	LogL	AIC	AICC	BIC
$Exp(\lambda)$	$(\hat{\lambda}) = 0.1952$	189,6184	381,2368	381,29	381,0941
$W(k, \lambda)$	$(\hat{k}, \hat{\lambda}) = (2.93, 5.75)$	146,496	296,9912	297,17	296,7059
$L(\theta)$	$(\hat{\theta}) = 0.341$	174,552	351,104	351,16	350,9613
$QL(\theta, \alpha)$	$(\hat{\theta}, \hat{\alpha}) = (0.390, 0.001)$	166,462	336,924	337,1	336,6387
$UTQL(\theta, \alpha, \zeta)$	$(\hat{\theta}, \hat{\alpha}, \hat{\zeta}) = (0.269, 0.013, 10.4)$	156,366	318,732	319,08	318,304
$LTQL(\theta, \alpha, v)$	$(\hat{\theta}, \hat{\alpha}, \hat{v}) = (0.390, 0.0002, 2)$	166,462	338,924	339,28	338,496
$DTQL(\theta, \alpha, v, \zeta)$	$(\hat{\theta}, \hat{\alpha}, \hat{v}, \hat{\zeta}) = (0.387, 0.014, 2, 10.4)$	143,643	295,286	295,88	294,7153

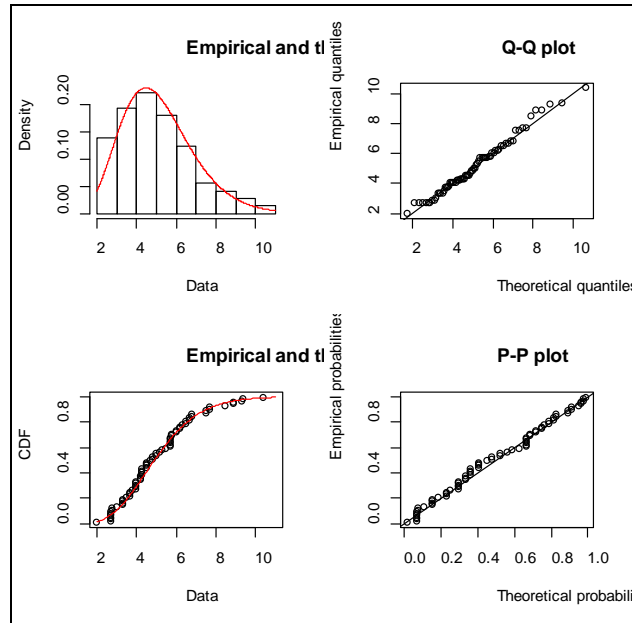


Figure 4: Fitting performance of the DTQL distribution based on wind speed data

Figure 4 shows that the wind speed data fits very well to the DTQL distribution.

6 Conclusion

In this study we propose a new model, the so-called the distribution which extends the Lindley distribution in the analysis of data with real support. An obvious reason for obtaining a standard distribution is because the generalized form provides larger flexibility in modeling real data. We derive expansions for the moments, quantile function, order statistics, survival function, hazard function, moments and for the moment generating function. The estimation of parameters is approached by the method of maximum likelihood, also the information matrix is derived. The wind speed data is modeled by the lower, upper and double truncated quasi Lindley distributions, exponential, Lindley, quasi Lindley and Weibull distributions to evaluate the performance of the truncated versions of quasi Lindley distribution. Fitting performance of these distributions are compared

according to AIC, AICC and BIC statistics values and clearly the wind speed data is best modeled by DTQL distribution. The present study might provide different and useful insights to scientists dealing with wind energy by firstly introducing a truncation parameter into QL distribution.

References

- [1] Lindley, D.V., Fiducial distributions and Bayes' theorem, *J. Royal Stat. Soc. Series B*, (1958), **20**, 102-107.
- [2] Bakouch, H.S., Bander, M.A., Al-Shaomrani, A.A. Marchi, V.A.A. and Louzada, F., An extended Lindley distribution, *Journal of The Korean Statistical Society*, **41**, (2012), 75-85.
- [3] Ghitany, M.A., Atieh, B. and Nadarajah S., Lindley distribution and its Applications, *Math. Cmput. Simul.*, **81**(11), (2008), 1405-1416.
- [4] Mazucheli, J. and Achcar, J.A., The Lindley distribution applied to competing risks lifetime data, *Computation Methods Programs*, **104**, (2011), 188-192.
- [5] Zakerzadeh, H. and Dolati, A., Generalized Lindley distribution, *Journal of Mathematical Extension*, **3**(2), (2009), 13-25.
- [6] Sankaran, M., The discrete Poisson–Lindley distribution, *Biometrics*, **26**, (1970), 145–149.
- [7] Ghitany, M.E., Al-Mutairi, D.K. and Nadarajah, S., Zero-truncated Poisson–Lindley distribution and its application, *Mathematics and Computers in Simulation*, **79**, (2008), 279–287.
- [8] Shanker, R. and Mishra, A., A quasi Lindley distribution, *African Journal of Mathematics and Computer Science Research*, **6**(4), (2013), 67-71.
- [9] Dusit C. and Cohen A.C., Estimation in the singly truncated Weibull distribution with an unknown truncation point, *Commun. Statist-Theory Meth.*, **13**(7), (1984), 843–857.

- [10]Zhang T. and Xie M., On the upper truncated Weibull distribution and its reliability implications, *Reliability Engineering and System Safety*, **96**, (2011), 94–200.
- [11]Ahmed, S.E., Castro-Kuriss, C., Leiva, V. and Sanhueza, A., A truncated version of the Birnbaum-Saunders distribution with an application in financial risk, *Pakistan Journal of Statistics*, **26**(1), (2010), 293-311.
- [12]Sing, S. K., Sing, U. and Sharma, V. K., The Truncated Lindley Distribution: Inference and Application, *Journal of Statistics Applications & Probability*, **3**(2), (2014), 219-228.
- [13]Zhou J., Erdem E., Li G. and Shi J., Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites, *Energy Convers Manage*, **51**(7), (2010), 1449–1458.
- [14]Bauer E., Characteristic frequency distributions of remotely sensed in situ and modeled wind speeds, *Int. J. Climatol*, **16**, (1996), 1087–1102.
- [15]Chang T.P., Estimation of wind energy potential using different probability density functions, *Appl. Energy*, **88**(5), (2011), 1848–1856.
- [16]Usta I. and Kantar Y.M., Analysis of some flexible families of distributions for estimation of wind speed distributions, *Appl. Energy*, **89**(1), (2012), 355–367.
- [17]Philippopoulos K., Deligiorgi D. and Karvounis G., Wind speed distribution modeling in the Greater Area of Chania, Greece, *Int. J. Green Energy*, **9**(2), (2012), 174–193.
- [18]Akdag S.A., Bagiorgas H.S. and Mihalakakou G., Use of two-component Weibull mixtures in the analysis of wind speed in the Eastern Mediterranean, *Appl. Energy*, **87**(8), (2010), 2566–2573.
- [19]Carta J.A. and Ramirez P., Use of finite mixture distribution models in the analysis of wind energy in the Canarian Archipelago, *Energy Convers Manage*, **48**, (2007), 281–291.

- [20]Soukissian T., Use of multi-parameter distributions for offshore wind speed modeling: The Johnson SB distribution, *Appl. Energy*, **111**, (2013), 982–1000.
- [21]Morgan V.T., Statistical distributions of wind parameters at Sydney, Australia, *Renew Energy*, **6**(1), (1995), 39–47.
- [22]Seguro, J.V. and Lambert, T.W., Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis, *Journal of Engineering and Industrial Aerodynamics*, **85**, (2000), 75-84.