The generalized weighted Lindley distribution: Properties, estimation, and applications

P.L. Ramos and F. Louzada

Abstract: In this study, a three-parameter lifetime distribution namely generalized weighted Lindley (GLW) distribution is proposed. The GLW distribution is an useful generalization of the weighted Lindley distribution which accommodates increasing, decreasing, decreasing-increasing-decreasing, bathtub, and unimodal hazard rate making it a flexible model for reliable data. A significant account of mathematical properties for this distribution is presented. Different estimation procedures are discussed such as maximum likelihood estimators, method of moments, ordinary and weighted least-squares, percentile, maximum product of spacings, and minimum distance estimators. The estimators are compared by extensive numerical simulations. Finally, two data-sets are analyzed for illustrative purposes proving that the GLW outperforms several other three-parameter lifetime distributions.

Subjects: Science; Mathematics & Statistics; Statistics & Probability; Statistics; Statistical Computing; Statistical Theory & Methods

Keywords: generalized weighted Lindley distribution; different estimation methods; maximum product of spacings

1. Introduction

In recent years, several new extensions of the exponential distribution have been introduced in the literature for describing real problems. Ghitany, Atieh, and Nadarajah (2008) investigated different properties of the Lindley distribution and outlined that in many cases the Lindley distribution
outperforms exponential distribution. Since then, many generalizations of the Lindley distribution have been introduced such as generalized Lindley (Zakerzadeh & Dolati, 2009), extended Lindley (Bakouch, Al-Zahrani, Al-Shomrani, Marchi, & Louzada, 2012), exponential Poisson Lindley (Barreto-Souza & Bakouch, 2013), and Power Lindley (Ghitany, Al-Mutairi, Balakrishnan, & Al-Enezi, 2013) distribution.

Ghitany, Alqallaf, Al-Mutairi, and Husain (2011) introduced a new class of weighted Lindley (WL) distribution adding more flexibility to the Lindley distribution. Let \( T \) be a random variable with a WL distribution. Then probability density function (p.d.f) is given by

\[
f(t; \lambda, \phi) = \frac{\lambda^{\phi+1}}{(\lambda + \phi) \Gamma(\phi)} t^{\phi-1} (1 + t) e^{-\lambda t},
\]

for all \( t > 0, \phi > 0 \) and \( \lambda > 0 \) and \( \Gamma(\phi) = \int_{0}^{\infty} e^{-x} x^{\phi-1} dx \) is the gamma function. One of its peculiarities is that the hazard function can have an increasing \((\phi \geq 1)\) or bathtub \((0 < \phi < 1)\) shape. Different properties and estimation methods for this model were presented by Mazucheli, Louzada, and Ghitany (2013), Ali (2015), Wang and Wang (in press), Al-Mutairi, Ghitany, and Kundu (2015).

In this study, a new lifetime distribution family is proposed which is a direct generalization of the WL distribution. The p.d.f is given by

\[
f(t; \phi, \alpha, \lambda) = \frac{\alpha^\phi}{(\lambda + \phi) \Gamma(\phi)} t^{\phi-1} (\lambda + (\lambda t)^\alpha) e^{-(\lambda t)^\alpha},
\]

for all \( t > 0, \phi > 0, \lambda > 0 \) and \( \alpha > 0 \). Important probability distributions can be obtained from the GWL distribution as the WL distribution \((\alpha = 1)\), Power Lindley distribution \((\phi = 1)\) and the Lindley distribution \((\phi = 1\) and \(\alpha = 1)\). Due to this relationship, such model could also be named as weighted power Lindley or generalized power Lindley distribution.

Torabi, Falahati-Naeini, and Montazeri (2014) discussed a class of distribution with four parameters which is a generalization of the proposed model. Such distribution includes the generalized WL, generalized gamma (GG) distribution, gamma and Weibull, among others. The main difference of this study lies in the fact that the proposed three-parameter distribution has a simple structure with less computational issues. In this way, the behavior of the p.d.f and the hazard function can be studied. This model has different forms of hazard function such as increasing, decreasing, bathtub, unimodal, or decreasing-increasing-decreasing shape making the GWL distribution a flexible model for reliable data. Moreover, a significant account of mathematical properties for the new distribution is provided.

The inferential procedures for the parameters of GLW distribution are presented considering different methods such as maximum likelihood estimators (MLE), methods of moments (ME), ordinary least-squares estimation (OLSE), weighted least-squares estimation (WLSE), maximum product of spacings (MPS), Cramer-von Mises type minimum distance (CME), Anderson–Darling (ADE) and right-tail Anderson–Darling (RADE). The performance of these estimation procedures are compared using extensive numerical simulations. Finally, two data-sets are analyzed for illustrative purposes proving that the GWL outperforms several usual three-parameter lifetime distributions such as the GG distribution (Stacy, 1962), the generalized Weibull (GW) distribution (Mudholkar, Srivastava, & Kollia, 1996), the generalized exponential-Poisson (GEP) distribution (Barreto-Souza & Cribari-Neto, 2009), and the exponentiated Weibull (EW) distribution (Mudholkar, Srivastava, & Freimer, 1995).

The results of this paper are organized as follows. Section 2 provides a significant account of mathematical properties for the new distribution. Section 3 presents the eight estimation methods which are considered. In the Section 4, a simulation study is presented in order to identify the most efficient procedure. Section 5 illustrates the proposed methodology in two real data-sets. Section 6 summarizes the present work.
Generalized weighted Lindley distribution

The generalized WL distribution (2) can be expressed as a two-component mixture

\[ f(t|\phi, \lambda, \alpha) = pf_j(t|\phi, \lambda, \alpha) + (1 - p) f_j(t|\phi, \lambda, \alpha) \]

where \( p = \lambda/(\lambda + \phi) \) and \( T_j \sim \text{GG}(\phi + j - 1, \lambda, \alpha) \), for \( j = 1, 2 \), i.e. \( f_j(t|\lambda, \phi) \) has GG distribution, given by

\[ f_j(t|\phi, \lambda, \alpha) = \frac{\alpha}{\Gamma(\phi + j - 1)} t^{\phi + j - 1} e^{-(\lambda + \phi)t}. \]  

(3)

The behavior of the p.d.f. (2) when \( t \to 0 \) and \( t \to \infty \) are, respectively, given by

\[ f(0) = \begin{cases} \infty, & \text{if } \alpha \phi < 1 \\ \frac{\alpha \lambda^2}{(\phi + \lambda) \Gamma(\phi)}, & \text{if } \alpha \phi = 1 \\ 0, & \text{if } \alpha \phi > 1 \end{cases} \]

Figure 1 gives examples of the shapes of the density function for different values of \( \phi, \lambda \) and \( \alpha \).

The cumulative distribution function from the GWL distribution is given by

\[ F(t|\phi, \lambda, \alpha) = \frac{\gamma[\phi, (\lambda t)^{\alpha}](\lambda + \phi) - (\lambda t)^{\alpha}e^{-(\lambda + \phi)t}}{(\lambda + \phi) \Gamma(\phi)}. \]  

(4)

where \( \gamma[y, x] = \int_0^y w^{x-1}e^{-w}dw \) is the lower incomplete gamma function.

2.1. Moments

Many important features and properties of a distribution can be obtained through its moments such as mean, variance, kurtosis, and skewness. In this section, important moment functions such as the moment-generating function, \( r \)-th moment, \( r \)-th central moment, among others are presented.

**Theorem 2.1** For the random variable \( T \) with GWL distribution, the moment-generating function is given by

\[ M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{\lambda^r r!} \frac{\Gamma(\frac{\phi + r}{\alpha} + \lambda)}{(\lambda + \phi) \Gamma(\phi)}. \]  

(5)
Proof  Note that, the moment-generating function from GG distribution (3) is given by

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\Gamma(\frac{1}{\alpha} + \phi + j - 1)}{\lambda^r \Gamma(\phi + j - 1)}.$$  

Since the GWL distribution (2) can be expressed as a two-component mixture, we have

$$M_p(x) = E[e^{tx}] = \int_0^\infty e^{tx} f(x) \, dx = p M_{X_1}(t) + (1-p) M_{X_2}(t) \tag{6}$$

$$= \frac{\lambda}{(\lambda + \phi)} \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\phi}{\lambda^r \Gamma(\phi + 1)} \tag{7}$$

$$= \frac{1}{(\lambda + \phi)} \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\Gamma(\frac{1}{\alpha} + \phi + 1)}{\lambda^r \Gamma(\phi + 1)} \tag{8}$$

$$= \sum_{r=0}^{\infty} \frac{t^r}{\lambda^r r!} \frac{\Gamma(\frac{1}{\alpha} + \phi + 1)}{(\lambda + \phi) \Gamma(\phi)}. \tag{9}$$

COROLLARY 2.2  For the random variable $T$ with GWL distribution, the $r$-th moment is given by

$$\mu_r = E[T^r] = \frac{\Gamma(\frac{1}{\alpha} + \phi + \lambda)}{(\lambda + \phi) \lambda \Gamma(\phi)}. \tag{10}$$

Proof  Note that, $\mu_r = M_p^{(r)}(0) = \frac{d^r M_p}{dt^r}$ and the result follows. \hfill \Box

COROLLARY 2.3  For the random variable $T$ with GWL distribution, the $r$-th central moment is given by

$$M_r = E[T - \mu] = \sum_{i=0}^{r} \binom{r}{i} (-\mu)^{r-i} E[T^i] \tag{11}$$

$$= \sum_{i=0}^{r} \binom{r}{i} \frac{\Gamma(\frac{1}{\alpha} + \phi + \lambda) \Gamma(\frac{1}{\alpha} + \phi)}{\lambda(\lambda + \phi) \lambda \Gamma(\phi)} \frac{\lambda^i (\lambda + \phi)^i}{(\lambda + \phi) \lambda \Gamma(\phi)}. \tag{12}$$

COROLLARY 2.4  A random variable $T$ with GWL distribution has the mean and variance respectively given by

$$\mu = \frac{\Gamma(\frac{1}{\alpha} + \phi + \lambda)}{\lambda(\lambda + \phi) \lambda \Gamma(\phi)}, \tag{13}$$

$$\sigma^2 = \frac{\lambda(\lambda + \phi) \Gamma(\frac{1}{\alpha} + \phi + \lambda) - \Gamma(\frac{1}{\alpha} + \phi + \lambda)^2 \Gamma(\frac{1}{\alpha} + \phi)}{\lambda^2 (\lambda + \phi)^2 \lambda \Gamma(\phi)^2}. \tag{14}$$

Proof  From (6) and considering $r = 1$ we have $\mu_1 = \mu$. The second result follows from (7) considering $r = 2$ with some algebraic manipulations. \hfill \Box

Another moment function that can be easily achieved for GWL distribution and plays an important role in information theory is given by

$$E[\log(T)] = \frac{\psi(\phi) - \alpha \log(\lambda + \phi)^{-1}}{\alpha}. \tag{15}$$
2.2. Survival properties

In this section, we present the survival, hazard, and mean residual life (MRL) function for the GWL distribution. The survival function of $T$ is given by

$$S(t|\phi, \lambda, \alpha) = \frac{\Gamma[\phi, (\lambda t)^{\alpha}] (\lambda t)^{\phi} e^{-(\lambda t)^{\alpha}}}{(\lambda t)^{\phi} \Gamma(\phi)}$$  \hspace{1cm} (11)

where $\Gamma(x, y) = \int_{y}^{\infty} w^{x-1} e^{-w} \, dw$ is called upper incomplete gamma. The hazard function is given as

$$h(t|\phi, \lambda, \alpha) = \frac{\frac{\lambda^{\alpha} t^{\alpha-1} (\lambda t)^{\phi} e^{-(\lambda t)^{\alpha}}}{\Gamma[\phi, (\lambda t)^{\alpha}]}}{\frac{\lambda^{\alpha} t^{\alpha-1} (\lambda t)^{\phi} e^{-(\lambda t)^{\alpha}}}{\Gamma[\phi, (\lambda t)^{\alpha}]}},$$  \hspace{1cm} (12)

The behavior of the hazard function (12) when $t \rightarrow 0$ and $t \rightarrow \infty$ are, respectively, given by

$$h(0) = \begin{cases} \infty, & \text{if } \alpha \phi < 1 \\ \frac{\alpha \lambda^{\alpha} t^{\alpha-1} (\lambda t)^{\phi} e^{-(\lambda t)^{\alpha}}}{\Gamma(\phi)}, & \text{if } \alpha \phi = 1 \\ 0, & \text{if } \alpha \phi > 1 \end{cases}$$

and $h(\infty) = \begin{cases} 0, & \text{if } \alpha \phi < 1 \\ \lambda, & \text{if } \alpha \phi = 1 \\ \infty, & \text{if } \alpha \phi > 1. \end{cases}$

**Theorem 2.5** The hazard rate function $h(t)$ of the GWL distribution is increasing, decreasing, bathtub, unimodal, or decreasing-increasing-decreasing shaped.

**Proof** The theorem proposed by Glaser (1980) is not easily applied in the GLW distribution. Since the hazard rate function (12) is complex, we considered the following cases:

1. Let $\alpha = 1$, then GWL distribution reduces to the WL distribution. Ghitany et al. (2011) proved that the hazard function is bathtub-shaped (increasing) if $0 < \phi < 1$ and considering different values of $\phi$ and $\lambda$.
2. Let $\phi = 1$, then GWL distribution reduces to the PL distribution. Considering $\beta = \lambda^{\alpha}$, Ghitany et al. (2013) proved that the hazard function is increasing when $\{0 < \alpha \leq 1, \beta > 0\}$; decreasing when $\{0 < \alpha \leq \frac{1}{2}, \beta > 0\}$ or $\{\frac{1}{2} < \alpha < 1, \beta \geq (2a - 1)^{2}(4a(1-a))^{-1}\}$; decreasing-increasing-decreasing if $\{\frac{1}{2} < \alpha < 1, 0 < \beta < (2a - 1)^{2}(4a(1-a))^{-1}\}$.
3. Let $\alpha = 2$ and $\lambda = 1$, from Glaser’s theorem (Glaser, 1980), the hazard rate function is decreasing shaped (unimodal) for $0 < \phi < 1$.

These properties make the GWL distribution a flexible model for reliable data. Figure 2 gives examples of the shapes of the hazard function for different values of $\phi$, $\lambda$, and $\alpha$.

Figure 2. Hazard function shapes for GWL distribution and considering different values of $\phi$, $\lambda$, and $\alpha$. 
The MRL has been widely used in survival analysis and represents the expected additional lifetime given that a component has survived until time $t$.

**Proposition 2.6** The MRL function $r(t|\phi, \lambda, \alpha)$ of the GWL distribution is given by

$$r(t|\phi, \lambda, \alpha) = \frac{\left( \phi + \frac{1}{e} + \lambda \right) \Gamma \left( \phi + \frac{1}{e}, (\lambda t)^{\alpha} \right) - \lambda t (\lambda + \phi) \Gamma \left( \phi, (\lambda t)^{\alpha} \right)}{\lambda (\lambda + \phi) \Gamma(\phi, (\lambda t)^{\alpha}) + (\lambda t)^{\alpha \phi} e^{-\lambda t^\alpha}} .$$

(13)

**Proof** Note that

$$r(t|\phi, \lambda, \alpha) = \frac{1}{S(t)} \int_{t}^{\infty} y f(y|\lambda, \phi) \, dy - t$$

$$= \frac{1}{S(t)} \left[ \int_{t}^{\infty} y f(y|\lambda, \phi) \, dy + (1 - p) \int_{x}^{\infty} y f(y|\lambda, \phi) \, dy \right] - t$$

$$= \frac{\left( \phi + \frac{1}{e} + \lambda \right) \Gamma \left( \phi + \frac{1}{e}, (\lambda t)^{\alpha} \right) - \lambda t (\lambda + \phi) \Gamma \left( \phi, (\lambda t)^{\alpha} \right)}{\lambda (\lambda + \phi) \Gamma(\phi, (\lambda t)^{\alpha}) + (\lambda t)^{\alpha \phi} e^{-\lambda t^\alpha}} .$$

The behavior of the MRL function when $t \to 0$ and $t \to \infty$ are, respectively, given by

$$r(0) = \frac{1}{\lambda ((\lambda + \phi) \Gamma(\phi))} \quad \text{and} \quad r(\infty) = \begin{cases} \infty, & \text{if } \alpha < 1 \\ \frac{1}{\lambda}, & \text{if } \alpha = 1 \\ 0, & \text{if } \alpha > 1. \end{cases}$$

\[2.3.\text{Entropy}\]

In information theory, entropy has played a central role as a measure of uncertainty associated with a random variable. Shannon's entropy is one of the most important metrics in information theory.

For the GWL distribution, Shannon's entropy can be obtained by solving

$$H_{S}(\phi, \lambda, \alpha) = - \int_{0}^{\infty} \log \left( \frac{a \lambda^\phi t^{\phi-1} (\lambda + (\lambda t)^{\alpha}) e^{-\lambda t^\alpha}}{(\lambda + \phi) \Gamma(\phi)} \right) f(t|\phi, \lambda, \alpha) \, dt .$$

(14)

**Proposition 2.7** A random variable $T$ with GWL distribution has Shannon's entropy given by

$$H_{S}(\phi, \lambda, \alpha) = \log(\lambda + \phi) + \log \Gamma(\phi) - \log \alpha - \log \lambda - \frac{\phi(1 + \phi + \lambda)}{(\lambda + \phi)}$$

$$- \frac{\psi(\phi)(\alpha \phi - 1)}{\alpha} - \frac{(\alpha \phi - 1)}{\alpha (\lambda + \phi)} - \frac{\eta(\phi, \lambda)}{(\lambda + \phi) \Gamma(\phi)}$$

(15)

where

$$\eta(\phi, \lambda) = \int_{0}^{\infty} (\lambda + y) \log(\lambda + y) y^{\phi-1} e^{-y} \, dy = \int_{0}^{\infty} (\lambda - \log u) \log(\lambda - \log u)(-\log u)^{\phi-1} \, du .$$

**Proof** From the Equation (14), we have

$$H_{S}(\phi, \lambda, \alpha) = - \log \alpha - \alpha \phi \log \lambda + \log(\lambda + \phi) + \log(\Gamma(\phi)) + \lambda \alpha E[T^\alpha]$$

$$- (\alpha \phi - 1) E[T] - E[\log(\lambda + (\lambda t)^{\alpha})] .$$

(16)

Note that

$$E[\log(\lambda + (\lambda t)^{\alpha})] = \int_{0}^{\infty} \log(\lambda + (\lambda t)^{\alpha}) \frac{a \lambda^\phi t^{\phi-1} (\lambda + (\lambda t)^{\alpha}) e^{-\lambda t^\alpha}}{(\lambda + \phi) \Gamma(\phi)} \, dt ,$$

using the change of variable $y = (\lambda t)^{\alpha}$ and after some algebra.


\[
E[\log(\lambda + (\lambda T)^\rho)] = \frac{1}{\lambda + \phi \Gamma(\phi)} \int_0^\infty (\lambda + y) \log(\lambda + y) y^{\lambda-1} e^{-y} dy
\]

\[
= \eta(\phi, \lambda) \lambda + \phi \Gamma(\phi).
\]

From Equations (6) and (10), we can easily find the solution of \(E[T^\rho]\) and \(E[\log T]\) and the result as follows.

Another popular entropy measure is proposed by Renyi (1961). Some recent applications of the Renyi entropy can be seen in Popescu and Aiordachioaie (2013). If \(T\) has the probability density function (1) then Renyi entropy is defined by

\[
\frac{1}{1 - \rho} \log \int_0^\infty f^\rho(x) \, dx.
\]

(17)

Proposition 2.8 A random variable \(T\) with GWL distribution, has the Renyi entropy given by

\[
H_\rho(\rho) = \frac{(\rho - 1)(\log \alpha + \log \lambda) - \rho(\log(\lambda + \phi) + \log \Gamma(\phi)) - \log(\delta(\rho, \phi, \lambda, \alpha))}{1 - \rho}
\]

(18)

where \(\delta(\rho, \phi, \lambda, \alpha) = \int_0^\infty y^{\alpha - \rho + 1 - \rho} (\lambda + y)^\rho e^{-y} dy\).

Proof The Renyi entropy is given by

\[
H_\rho(\rho) = \frac{1}{1 - \rho} \log \left( \frac{\alpha^{\rho} \lambda^\rho}{(\lambda + \phi)\Gamma(\phi)} \int_0^\infty (\lambda t)^{\rho - 1} (\lambda + (\lambda t)^\rho) e^{-\rho(\lambda t)^\rho} dt \right)
\]

\[
= \frac{1}{1 - \rho} \log \left( \frac{\alpha^{\rho} \lambda^\rho}{(\lambda + \phi)\Gamma(\phi)} \int_0^\infty y^{\rho - \rho + 1 - \rho} (\lambda + y)^\rho e^{-\rho y} dy \right)
\]

\[
= \frac{1}{1 - \rho} \log \left( \frac{\alpha^{\rho} \lambda^\rho}{(\lambda + \phi)\Gamma(\phi)} \delta(\rho, \phi, \lambda, \alpha) \right)
\]

and with some algebra the proof is completed.

2.4. Lorenz curves

The Lorenz curve (Bonferroni, 1930) is a well-known measure used in reliability, income inequality, life testing and renewal theory. The Lorenz curve for a non-negative \(T\) random variable is given through the consecutive plot of

\[
L(F(t)) = \frac{\int_0^t x f(x) \, dx}{\int_0^\infty x f(x) \, dx} = \frac{1}{\mu} \int_0^t x f(x) \, dx.
\]

Proposition 2.9 The Lorenz curve for the GWL distribution is

\[
L(p) = \frac{\left( \frac{1}{\alpha} + \phi + \lambda \right) y \left( \phi + \frac{1}{\alpha} \right) \left( \lambda F^{-1}(p) \right)^\rho - \left( \lambda F^{-1}(p) \right)^{\alpha\lambda + \rho - 1} e^{-\left( \lambda F^{-1}(p) \right)^\rho}}{\left( \frac{1}{\alpha} + \phi + \lambda \right) \Gamma \left( \frac{1}{\alpha} + \phi \right)}
\]

where \(F^{-1}(p) = t_p\).
3. Methods of estimation

In this section, we present eight different estimation methods for the parameters $\phi, \lambda$ and $\alpha$ of the GWL distribution.

### 3.1. Maximum likelihood estimation

The maximum likelihood method has been widely used due to its better asymptotic properties. The estimates are obtained by maximizing the likelihood function. Let $T_1, \ldots, T_n$ be a random sample where $T \sim GWL(\phi, \lambda, \alpha)$, the likelihood function is given by

$$
L(\phi, \lambda, \alpha; t) = \frac{a^n \lambda^{n\phi}}{(\lambda + \phi)^\Gamma(\phi)^n} \left\{ \prod_{i=1}^{n} \left( t_i^{\phi-1} \right) \right\} \left\{ \prod_{i=1}^{n} (\lambda + (\lambda t_i)^\alpha) \exp \left\{ -\lambda^\alpha \sum_{i=1}^{n} t_i^\alpha \right\} \right\}. \quad (19)
$$

The log-likelihood function $l(\phi, \lambda, \alpha; t) = \log L(\phi, \lambda, \alpha; t)$ is given by

$$
l(\phi, \lambda, \alpha; t) = n \log \alpha + n \phi \log \lambda - n \log(\lambda + \phi) - n \log \Gamma(\phi) + (\alpha \phi - 1) \sum_{i=1}^{n} \log(t_i) \quad + \sum_{i=1}^{n} \log(\lambda + (\lambda t_i)^\alpha) - \lambda^\alpha \sum_{i=1}^{n} t_i^\alpha. \quad (20)
$$

From the expressions $\frac{\partial}{\partial \phi} l(\phi, \lambda, \alpha; t) = 0$, $\frac{\partial}{\partial \lambda} l(\phi, \lambda, \alpha; t) = 0$, and $\frac{\partial}{\partial \alpha} l(\phi, \lambda, \alpha; t) = 0$, the likelihood equations are

$$
n \alpha \log(\lambda) + \lambda \sum_{i=1}^{n} \log(t_i) = \frac{n}{\lambda + \phi} + n \psi(\phi) \quad (21)
$$

$$
n \alpha \phi \frac{1}{\lambda} + \sum_{i=1}^{n} \frac{1 + \phi}{\lambda + (t_i)^\alpha} \lambda^{\phi-1} \sum_{i=1}^{n} t_i^\phi \quad + \frac{n}{\lambda + \phi} \quad (22)
$$

and

$$
n \alpha \phi \log(\lambda) + \phi \sum_{i=1}^{n} \log(t_i) + \sum_{i=1}^{n} \frac{1}{\lambda + (t_i)^\alpha} \lambda^{\phi-1} \sum_{i=1}^{n} t_i^\phi \log(\lambda t_i) = \lambda^\alpha \sum_{i=1}^{n} t_i^\alpha \log(\lambda t_i), \quad (23)
$$

where $\psi(k) = \frac{\partial}{\partial \phi} \log \Gamma(k) = \frac{\Gamma(k)}{\Gamma(k) \Gamma(\phi)}$. Numerical methods such as Newton-Rapshon are required to find the solution of the nonlinear system. Note that from (21) and (23) and after some algebra we have

$$
\hat{\phi}_{\text{MLE}} = \left( \frac{1}{n \log(\lambda) + \sum_{i=1}^{n} \log(t_i)} \right) \left( \frac{n}{\lambda + \phi} + n \psi(\phi) \right) \quad (24)
$$

$$
\hat{\phi}_{\text{MLE}} = \left( \frac{\lambda^\alpha \sum_{i=1}^{n} t_i^\alpha \log(\lambda t_i) - \sum_{i=1}^{n} (t_i)^\alpha \log(\lambda t_i)}{n \alpha} \right) \quad (25)
$$

Under mild conditions, the maximum likelihood estimates (MLEs) are asymptotically normal distributed with a joint multivariate normal distribution given by

$$
(\hat{\phi}_{\text{MLE}}, \hat{\lambda}_{\text{MLE}}, \hat{\alpha}_{\text{MLE}}) \sim N_3 \left( (\phi, \lambda, \alpha), I^{-1} ((\phi, \lambda, \alpha)) \right) \text{ as } n \to \infty.
$$

where $I((\phi, \lambda, \alpha))$ is the Fisher information matrix is given as

$$
I((\phi, \lambda, \alpha)) = \begin{bmatrix}
I_{\phi,\phi}(\phi, \lambda, \alpha) & I_{\phi,\lambda}(\phi, \lambda, \alpha) & I_{\phi,\alpha}(\phi, \lambda, \alpha) \\
I_{\lambda,\phi}(\phi, \lambda, \alpha) & I_{\lambda,\lambda}(\phi, \lambda, \alpha) & I_{\lambda,\alpha}(\phi, \lambda, \alpha) \\
I_{\alpha,\phi}(\phi, \lambda, \alpha) & I_{\alpha,\lambda}(\phi, \lambda, \alpha) & I_{\alpha,\alpha}(\phi, \lambda, \alpha)
\end{bmatrix}, \quad (26)
$$

and the elements of the matrix are given in Appendix 2.
3.2. Moments estimators

The method of moments is one of the oldest methods used for estimating parameters in statistical models. The moments estimators (MEs) of the GLW distribution can be obtained by equating the first three sample moments \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} t_i \), \( \frac{1}{n} \sum_{i=1}^{n} t_i^2 \) and \( \frac{1}{n} \sum_{i=1}^{n} t_i^3 \) with the theoretical moments:

\[
\frac{1}{n} \sum_{i=1}^{n} t_i = \left( \frac{\lambda + \phi + x}{\phi} \right) \Gamma\left( \frac{\lambda + \phi}{\phi} \right) \quad \frac{1}{n} \sum_{i=1}^{n} t_i^2 = \left( \frac{\lambda + \phi + x}{\phi} \right) \Gamma\left( \frac{\lambda + \phi}{\phi} \right) \left( \frac{\lambda}{\phi^2} \right) \\
and \frac{1}{n} \sum_{i=1}^{n} t_i^3 = \left( \frac{\lambda + \phi + x}{\phi} \right) \Gamma\left( \frac{\lambda + \phi}{\phi} \right) \left( \frac{\lambda}{\phi^2} \right)^2.
\]

Therefore, the ME \( \hat{\phi}_{ME}, \hat{\lambda}_{ME} \) and \( \hat{\alpha}_{ME} \) can be obtained by solving the non-linear equations

\[
\left( \frac{\lambda + \phi + x}{\phi} \right) \Gamma\left( \frac{\lambda + \phi}{\phi} \right) - \frac{1}{n} \sum_{i=1}^{n} t_i^j = 0, \quad j = 1, 2, 3.
\]

3.3. Ordinary and weighted least-square estimate

Let \( t_{(1)}, t_{(2)}, \ldots, t_{(n)} \) be the order statistics (the same notation is assumed for the next subsections) of the random sample of size \( n \) from \( F(t|\phi, \lambda, \alpha) \). The least square estimators \( \hat{\phi}_{LSE}, \hat{\lambda}_{LSE} \) and \( \hat{\alpha}_{LSE} \) can be obtained by minimizing

\[
V(\phi, \lambda, \alpha) = \sum_{i=1}^{n} \left[ F(t_{(i)}|\phi, \lambda, \alpha) - \frac{i}{n+1} \right]^2
\]

with respect to \( \phi, \lambda \) and \( \alpha \). Equivalently, the estimates can be obtained by solving the non-linear equations

\[
\sum_{i=1}^{n} \left[ F(t_{(i)}|\phi, \lambda, \alpha) - \frac{i}{n+1} \right] \Delta_j(t_{(i)}|\phi, \lambda, \alpha) = 0, \quad j = 1, 2, 3
\]

where

\[
\Delta_1(t_{(i)}|\phi, \lambda, \alpha) = \frac{\partial}{\partial \phi} F(t_{(i)}|\phi, \lambda, \alpha), \quad \Delta_2(t_{(i)}|\phi, \lambda, \alpha) = \frac{\partial}{\partial \lambda} F(t_{(i)}|\phi, \lambda, \alpha)
\]

and

\[
\Delta_3(t_{(i)}|\phi, \lambda, \alpha) = \frac{\partial}{\partial \alpha} F(t_{(i)}|\phi, \lambda, \alpha).
\]

Note that the solution of \( \Delta_i \) for \( i = 1, 2, 3 \) involves partial derivatives of the lower incomplete gamma function. However, this can be easily achieved numerically with high precision.

The weighted least-squares estimates (WLSEs), \( \hat{\phi}_{WLSE}, \hat{\lambda}_{WLSE} \) and \( \hat{\alpha}_{WLSE} \) can be obtained by minimizing

\[
W(\phi, \lambda, \alpha) = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[ F(t_{(i)}|\phi, \lambda, \alpha) - \frac{i}{n+1} \right]^2
\]

These estimates can also be obtained by solving the non-linear equations

\[
\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[ F(t_{(i)}|\phi, \lambda, \alpha) - \frac{i}{n+1} \right] \Delta_j(t_{(i)}|\phi, \lambda, \alpha) = 0, \quad j = 1, 2, 3,
\]

where \( \Delta_1(\cdot|\phi, \lambda, \alpha), \Delta_2(\cdot|\phi, \lambda, \alpha) \) and \( \Delta_3(\cdot|\phi, \lambda, \alpha) \) are given in (27).
3.4. Method of maximum product of spacings

The MPS method is a powerful alternative to MLE for the estimation of unknown parameters of continuous univariate distributions. Proposed by Cheng and Amin (1979, 1983), this method was also independently developed by Ranneby (1984) as an approximation to the Kullback–Leibler information measure. Cheng and Amin (1983) proved desirable properties of the MPS such as asymptotic efficiency, invariance, and more importantly, the consistency of maximum product of spacing estimators holds under more general conditions than for MLEs.

Let \( D_i(\phi, \lambda, \alpha) = F(t_{i1}(\phi, \lambda, \alpha)) - F(t_{i-1}(\phi, \lambda, \alpha)) \) for \( i = 1, 2, \ldots, n + 1 \), be the uniform spacings of a random sample from the GWL distribution, where \( F(t_{i0}(\phi, \lambda, \alpha)) = 0 \) and \( F(t_{i+1}(\phi, \lambda, \alpha)) = 1 \). Clearly \( \sum_{i=1}^{n+1} D_i(\phi, \lambda, \alpha) = 1 \). The MPS estimates \( \hat{\phi}_{MPS}, \hat{\lambda}_{MPS} \) and \( \hat{\alpha}_{MPS} \) are obtained by maximizing the geometric mean of the spacings

\[
G(\phi, \lambda, \alpha) = \left[ \prod_{i=1}^{n+1} D_i(\phi, \lambda, \alpha) \right]^{\frac{1}{n+1}}
\]

with respect to \( \phi, \lambda \) and \( \alpha \), or, equivalently, by maximizing the logarithm of the geometric mean of sample spacings

\[
H(\phi, \lambda, \alpha) = \frac{1}{n + 1} \sum_{i=1}^{n+1} \log D_i(\phi, \lambda, \alpha).
\]

The estimates \( \hat{\phi}_{MPS}, \hat{\lambda}_{MPS} \) and \( \hat{\alpha}_{MPS} \) of the parameters \( \phi, \lambda \) and \( \alpha \) can be obtained by solving the nonlinear equations

\[
\frac{1}{n + 1} \sum_{i=1}^{n+1} \frac{1}{D_i(\phi, \lambda, \alpha)} \left[ \Delta_j(t_{i1}(\phi, \lambda, \alpha)) - \Delta_j(t_{i-1}(\phi, \lambda, \alpha)) \right] = 0, \quad j = 1, 2, 3,
\]

where \( \Delta_1(\cdot|\phi, \lambda, \alpha), \Delta_2(\cdot|\phi, \lambda, \alpha) \) and \( \Delta_3(\cdot|\phi, \lambda, \alpha) \) are given respectively in (27). Note that if \( t_{i+k} = t_{i+k-1} = \cdots = t_{i1} \) then \( D_i(\phi, \lambda, \alpha) = D_{i+k-1}(\phi, \lambda, \alpha) = \cdots = D_i(\phi, \lambda, \alpha) = 0 \). Therefore, the MPS estimators are sensitive to closely spaced observations, especially ties. When the ties are due to multiple observations, \( D_j(\phi, \lambda, \alpha) \) should be replaced by the corresponding likelihood \( f(t_{i1}, \phi, \lambda, \alpha) \) since \( t_{i1} = t_{i-1} \).

Under mild conditions for the GWL distribution, the MPS estimators are asymptotically normal distributed with a joint trivariate normal distribution given by

\[
(\hat{\phi}_{MPS}, \hat{\lambda}_{MPS}, \hat{\alpha}_{MPS}) \sim N_3\left( (\phi, \lambda, \alpha), \Gamma^{-1}(\phi, \lambda, \alpha) \right) \quad \text{as } n \to \infty.
\]

3.5. The Cramer-von Mises minimum distance estimators

The Cramer-von Mises estimator is a type of minimum distance estimators (also called maximum goodness-of-fit estimators) and is based on the difference between the estimate of the cumulative distribution function and the empirical distribution function (Luceño, 2006).

Macdonald (1971) motivated the choice of the CME estimators providing empirical evidence that the bias of the estimator is smaller than the other minimum distance estimators. The Cramer-von Mises estimates \( \hat{\phi}_{CME}, \hat{\lambda}_{CME} \) and \( \hat{\alpha}_{CME} \) of the parameters \( \phi, \lambda \) and \( \alpha \) are obtained by minimizing

\[
C(\phi, \lambda, \alpha) = \frac{1}{12n} + \sum_{i=1}^{n} \left( F(t_{i1}(\phi, \lambda, \alpha)) - \frac{2i - 1}{2n} \right)^2,
\]

with respect to \( \phi, \lambda \) and \( \alpha \). These estimates can also be obtained by solving the nonlinear equations:
\[
\sum_{i=1}^{n} \left( F(t_{(i)}|\phi, \lambda, \alpha) - \frac{2i-1}{2n} \right) \Delta_j(t_{(i)}|\phi, \lambda, \alpha) = 0, \quad j = 1, 2, 3,
\]
where \(\Delta_1(\cdot|\phi, \lambda, \alpha), \Delta_2(\cdot|\phi, \lambda, \alpha)\) and \(\Delta_3(\cdot|\phi, \lambda, \alpha)\) are given respectively in (27).

### 3.6. The Anderson–Darling and Right-tail Anderson–Darling estimators

Another type of minimum distance estimator is based on ADE statistic and is known as ADE estimator. The ADE estimates \(\hat{\phi}_{\text{ADE}}, \hat{\lambda}_{\text{ADE}}\) and \(\hat{\alpha}_{\text{ADE}}\) of the parameters \(\phi, \lambda\) and \(\alpha\) are obtained by minimizing, with respect to \(\phi, \lambda\) and \(\alpha\), the function

\[
A(\phi, \lambda, \alpha) = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left( \log F(t_{(i)}|\phi, \lambda, \alpha) + \log S(t_{(n+1-i)}|\phi, \lambda, \alpha) \right). \tag{32}
\]

These estimates can also be obtained by solving the nonlinear equations

\[
\sum_{i=1}^{n} (2i-1) \left[ \frac{\Delta_j(t_{(i)}|\phi, \lambda, \alpha)}{F(t_{(i)}|\phi, \lambda, \alpha)} - \frac{\Delta_j(t_{(n+1-i)}|\phi, \lambda, \alpha)}{S(t_{(n+1-i)}|\phi, \lambda, \alpha)} \right] = 0, \quad j = 1, 2, 3.
\]

The Right-tail ADE estimates \(\hat{\phi}_{\text{RADE}}, \hat{\lambda}_{\text{RADE}}\) and \(\hat{\alpha}_{\text{RADE}}\) of the parameters \(\phi, \lambda\) and \(\alpha\) are obtained by minimizing

\[
R(\phi, \lambda, \alpha) = \frac{n}{2} - 2 \sum_{i=1}^{n} F(t_{(i)}|\phi, \lambda, \alpha) - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log S(t_{(n+1-i)}|\phi, \lambda, \alpha). \tag{33}
\]

with respect to \(\phi, \lambda\) and \(\alpha\). These estimates can also be obtained by solving the nonlinear equations:

\[
-2 \sum_{i=1}^{n} \Delta_j(t_{(i)}|\phi, \lambda, \alpha) + \frac{1}{n} \sum_{i=1}^{n} (2i-1) \frac{\Delta_j(t_{(n+1-i)}|\phi, \lambda, \alpha)}{S(t_{(n+1-i)}|\phi, \lambda, \alpha)} = 0, \quad j = 1, 2, 3.
\]

where \(\Delta_1(\cdot|\phi, \lambda, \alpha), \Delta_2(\cdot|\phi, \lambda, \alpha)\) and \(\Delta_3(\cdot|\phi, \lambda, \alpha)\) are given respectively in (27).

### 4. Simulation study

In this section, an intensive simulation study is presented to compare the efficiency of the estimation procedures for parameters of the GWL distribution. The following procedure was adopted:

1. Generate pseudo-random values from the GWL(\(\phi, \lambda, \alpha\)) with size \(n\).
2. Using the values obtained in step 1, calculate \(\hat{\phi}, \hat{\lambda}\) and \(\hat{\alpha}\) via 1-MLE, 2-MPS, 3-ADE, 4-RTADE, 5-LSE, 6-WLSE, 7-ME, 8-CME.
3. Repeat the steps 1 and 2 \(N\) times.
4. Using \(\hat{\Theta} = (\hat{\phi}, \hat{\lambda}, \hat{\alpha})\) and \(\Theta = (\phi, \lambda, \alpha)\), compute the mean relative estimates (MRE) \(\sum_{j=1}^{N} \frac{\hat{\theta}_j/\theta_j}{N}\) and the mean square errors (MSE) \(\sum_{j=1}^{N} \frac{\hat{\theta}_j - \theta_j}{N}\), for \(i = 1, 2, 3\).

Considering this approach, the most efficient estimation method will have MREs closer to one and MSEs closer to zero. The results were computed using the software R using the seed 2015 to generate the pseudo-random values. The initial values considered were the same values used to generate the random samples. The chosen values to perform this procedure were \(N = 10,000\) and \(n = (50, 60, \ldots, 250)\). For reasons of space, we have presented the results only for \(\Theta = (2, 0.5, 0.1)\). However, the following results are similar for other choices of \(\Theta\).
For this comparison to be meaningful, the estimation procedures need to be performed under same conditions. However, for some particular samples and estimation methods, the numerical techniques do not work well in finding the parameter estimates. Therefore, a rate study is presented to verify the frequency of convergence of the numerical solutions. This procedure is carried out by counting the number of times each estimation fails in finding the numerical solution. In Figure 3 we present the proportion of failure from each method.

From Figure 3, the MLE, LSE, WLSE, ME, and the CME estimators fail in finding the parameter estimates for a significant number of samples. Therefore, such methods are not recommended for estimation of the GLW parameters. Hereafter, we consider the MPS, ADE, RADE estimators due to their better computational stability. The MLE is considered only for illustrative purposes since it is the most used estimation method. Figure 4 presents the MREs, MSEs for the estimates of $\phi, \lambda$ and $\alpha$ using the MLE, MPS, ADE, RADE with $N$ simulated samples and different values of $\theta = (2, 0.5, 0.1)$ and $n$. The horizontal lines in both figures correspond to MREs and MSEs being one and zero, respectively.

From these results, the MSE of the MLE, MPS, ADE, and RADE estimators tend to zero for large $n$ and also, as expected, the values of MREs tend to one, i.e. the estimates are consistent and asymptotically unbiased for the parameters. For small sample sizes, the MLE has the largest MSEs. The MPS has smaller MSEs with MREs closer to one for almost all values of $n$. Additionally, the MPS, ADE, and RADE estimators were the only methods that were able to find $\hat{\phi}, \hat{\lambda}$ and $\hat{\alpha}$ for all the $2 \times 10^6$ generated samples. Therefore, combining all results with the good properties of the MPS method such as consistency, asymptotic efficiency, normality and invariance, we conclude that the MPS estimators are a highly competitive method compared to the maximum likelihood for estimating the parameters of the GWL distribution.

Figure 3. Proportion of failure from $N$ simulated samples, considering different values of $n$ using the following estimation method 1-MLE, 2-MPS, 3-ADE, 4-RADE, 5-LSE, 6-WLSE, 7-ME, 8-CME.
5. Application

In this section, we compare the GWL distribution with other three-parameter lifetime distributions considering two data-sets, the first with bathtub hazard rate and the other with the increasing hazard function. The following lifetime distributions were considered. The GG distribution with p.d.f given by

\[ f(t) = \alpha \Gamma(\phi)^{-1} \beta^\alpha t^{\alpha - 1} e^{\beta t} \]

where \( \beta > 0, \phi > 0 \) and \( \alpha > 0 \). The GW distribution where the p.d.f is

\[ f(t) = (\alpha \phi)^{-1} (t/\phi)^{1/\alpha - 1} (1 - \lambda (t/\phi)^{1/\alpha})^{1/\lambda - 1} \]

where \( \lambda \in \mathbb{R}, \phi > 0 \) and \( \alpha > 0 \). The GEP distribution with p.d.f given by

\[ f(t) = \left( a \beta \phi / (1 - e^{-\phi})^a \right) e^{-\phi - \beta t + \phi \exp(-\beta t)} \left( 1 - e^{-\phi + \phi \exp(-\beta t)} \right)^{a-1} \]

where \( \beta > 0, \phi > 0 \) and \( \alpha > 0 \).

The EW distribution with p.d.f

\[ f(t) = \frac{1}{u \Gamma(\phi)} t^{\phi - 1} e^{-t/u} \frac{1}{u^\phi} \]

where \( u > 0, \phi > 0 \) and \( \phi > 0 \).

Figure 4. MREs, MSEs related from the estimates of \( \phi = 0.5, \lambda = 0.7 \) and \( \alpha = 1.5 \) for \( N \) simulated samples, considering different values of \( n \) obtained using the following estimation method 1-MLE, 2-MPS, 3-ADE, 4-RTADE.
where $\beta > 0$, $\phi > 0$ and $\alpha > 0$.

The TTT-plot (total time on test) is considered in order to verify the behavior of the empirical hazard function (Barlow & Campo, 1975). The TTT-plot is obtained through the plot of $\frac{r}{n}, G\left(\frac{r}{n}\right)$ where $G(r/n)$ is the statistical order. If the curve is concave (convex), the hazard function is increasing (decreasing). On the other hand, when it starts convex and then becomes concave (concave and then convex) the hazard function has bathtub (inverse bathtub) shape.

The goodness of fit is checked considering the Kolmogorov–Smirnov (KS) test. This procedure is based on the KS statistic $D_n = \sup_t |F_n(t) - F(t; \lambda, \mu)|$, where $F$ is the empirical distribution function and $F(t; \lambda, \mu)$ is c.d.f. A hypothesis test is conducted at the 5\% level of significance to test whether or not the data come from $F(t; \lambda, \mu)$. In this case, the null hypothesis is rejected if the returned p-value is smaller than 0.05.

To carry out the model selection, the following discrimination criterion methods are adopted: AIC (Akaike information criteria) and AICc (Corrected Akaike information criterion) computed, respectively, by $AIC = -2\ell(\hat{\theta}; \mathbf{t}) + 2k$ and $AICc = AIC + 2k(k + 1)(n - k - 1)^{-1}$, where $k$ is the number of parameters to be fitted and $\hat{\theta}$ is estimation of $\theta$. For a set of candidate models for $\mathbf{t}$, the best one provides the minimum values.

5.1. Lifetimes data

Aarset (1987) presents the data-set (see Table 1) related to the lifetime in hours of 50 devices on test.

Figure 5 shows (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution.
Table 2 presents the AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the Aarset dataset.

| Criteria | Gen. WL | Gen. Gamma | Gen. Weibull | Exp. Weibull | Gen. EP |
|----------|---------|------------|--------------|--------------|---------|
| AIC      | 418.031 | 448.294    | 430.055      | 463.674      | 486.255 |
| AICc     | 412.552 | 442.816    | 424.576      | 458.196      | 480.777 |
| KS       | 0.5787  | 0.0115     | 0.0453       | 0.0222       | 0.0302  |

Comparing the empirical survival function with the adjusted distributions, it can be observed that the GWL distribution is as a better fit. This result is also confirmed from the AIC and AICC (see Table 3) since GWL distribution has the minimum values and also the p-values returned from the KS test are greater than 0.05. It should be emphasized that considering a significance level of 5%, the others models are not able to fit the proposed data. Table 3 displays the MPS estimates, standard errors, and the confidence intervals (CI) for \( \phi \), \( \lambda \) and \( \alpha \) of the GWL distribution.

### 5.2. Average flows data

The study of average flows has been proved to be of high importance to protect and maintain aquatic resources in streams and rivers (Reiser, Wesche, & Estes, 1989). In this section, we consider a real data-set related to the average flows (m\(^3\)/s) of the Cantareira system during January at São Paulo city in Brazil. It is worth mentioning that the Cantareira system provides water to 9 million people in the São Paulo metropolitan area. The data-set available in Table 4 was obtained from the National Water Agency from 1930 to 2012.

Table 3. MPS estimates, standard-error and 95\% CI for \( \phi \), \( \lambda \) and \( \alpha \)

| \( \theta \) | \( \hat{\theta}_{MPS} \) | S.E(\( \hat{\theta} \)) | CI\(_{95\%}\)(\( \hat{\theta} \)) |
|------------|-----------------|-----------------|-----------------|
| \( \phi \)  | 0.0057          | 0.00091         | (0.0039; 0.0075)|
| \( \lambda \) | 0.0118          | 0.00003         | (0.0117; 0.0118)|
| \( \alpha \) | 110.4964        | 6.58144         | (97.5971; 123.3958)|

Table 2 presents the AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the Aarset dataset.
In this section, we consider the ML estimator showing that both MPS or MLE could be used successfully in applications. Figure 6 shows (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution. Table 5 presents the AIC and AICc criteria and the $p$-value from the KS test for all fitted distributions considering the data-set related to the January average flows (m$^3$/s) of the Cantareira system.

From the empirical survival function and the adjusted distributions, it can be observed that the GWL distribution is better. This result is also confirmed from AIC and AICC since GWL distribution has the minimum values and the $p$-values returned from the KS test are greater than 0.05. Table 6 displays the ML estimates, standard errors, and the CI for $\phi$, $\lambda$ and $\alpha$ of the GWL distribution.

### 6. Concluding remarks

To summarize, we have proposed a three-parameter lifetime distribution. The GLW distribution is a straightforward generalization of the WL distribution proposed by Ghitany et al. (2011), which accommodates increasing, decreasing, decreasing-increasing-decreasing, bathtub, and unimodal hazard rate making the GWL distribution a flexible model for reliable data. The mathematical properties of this distribution are also discussed.

The estimation procedures for the parameters of GWL distribution are also derived considering eight estimation methods. Since it is not feasible to compare these methods theoretically, we have presented an extensive simulation study in order to identify the most efficient procedure. We
observed that the MLE, ME, LSE, WLSE, and the CME estimators fail in finding the parameter estimates for a significant number of samples. The simulations showed that the MPS (maximum product of spacing) is the most efficient method for estimating the parameters of the GWL distribution in comparison to its competitors. Finally, two data-sets were analyzed for illustrative purposes proving that the GWL distribution outperforms several usual three parameter lifetime distributions.

Acknowledgements
We are grateful to the Editorial Board and the reviewers for their valuable comments and suggestions which has improved the manuscript.

Funding
The research was partially supported by CNPq, FAPESP, and CAPES of Brazil.

Author details
P.L. Ramos1
E-mail: pedro.l.ramos@usp.br
F. Louzada1
E-mail: louzada@icmc.usp.br
1 Institute of Mathematical Science and Computing, Universidade de São Paulo, São Carlos-SP, Brazil.

Citation information
Cite this article as: The generalized weighted Lindley distribution: Properties, estimation, and applications, P.L. Ramos & F. Louzada, Cogent Mathematics (2016), 3:1256022.

References
Aarset, M. V. (1987). How to identify a bathtub hazard rate. IEEE Transactions on Reliability, 36, 106–108.
Ali, S. (2015). On the bayesian estimation of the weighted lindley distribution. Journal of Statistical Computation and Simulation, 85, 855–880.
Al-Mutairi, D., Ghitany, M., & Kundu, D. (2013). Inferences on stress-strength reliability from weighted lindley distributions. Communications in Statistics-Theory and Methods, 44, 4096–4113.
Bakouch, H. S., Al-Zahrani, B. M., Al-Shomrani, A. A., Marchi, V. A., & Louzada, F. (2012). An extended lindley distribution. Journal of the Korean Statistical Society, 41, 75–85.
Barlow, R. E., & Campo, R. A. (1975). Total time on test processes and applications to failure data analysis (Technical report). Berkeley, CA: DTIC Document.
Barroto-Souza, W., & Bokouch, H. S. (2013). A new lifetime model with decreasing failure rate. Statistics, 47, 465–476.
Barreto-Souza, W., & Cribari-Neto, F. (2009). A generalization of the exponential-poisson distribution. Statistics & Probability Letters, 79, 2493–2500.
Bonferroni, C. (1930). Elementi di statistica generale. Firenze: Seeber.
Cheng, R., & Amin, N. (1979). Maximum product of spacings estimation with application to the lognormal distribution (Mathematical Report 79-1). Cardiff: University of Wales IST.
Cheng, R., & Amin, N. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. Journal of the Royal Statistical Society. Series B (Methodological), 45, 394–403.
Ghitany, M., Al-Mutairi, D., Balokrishnan, N., & Al-Enezi, L. (2013). Power lindley distribution and associated inference. Computational Statistics & Data Analysis, 64, 20–33.
Ghitany, M., Alqallaf, F., Al-Mutairi, D., & Husain, H. (2011). A two-parameter weighted lindley distribution and its applications to survival data. Mathematics and Computers in Simulation, 81, 1190–1201.
Ghitany, M., Alteh, B., & Nababarajah, S. (2008). Lindley distribution and its application. Mathematics and Computers in Simulation, 78, 493–506.
Gleser, R. E. (1980). Bathtub and related failure rate characterizations. Journal of the American Statistical Association, 75, 667–672.
Luceño, A. (2006). Fitting the generalized paretos distribution to data using maximum goodness-of-fit estimators. Computational Statistics & Data Analysis, 51, 904–917.
Macdonald, P. (1971). An estimation procedure for mixtures of distribution. Journal of the Royal Statistical Society. Series B (Methodological), 33, 326–329.
Mazucheli, J., Louzada, F., & Ghitany, M. (2013). Comparison of estimation methods for the parameters of the weighted lindley distribution. Applied Mathematics and Computation, 220, 463–471.
Mudholkar, G. S., Srivastava, D. K., & Freimer, M. (1995). The exponentiated weibull family: A reanalysis of the bus-motor-failure data. Technometrics, 37, 436-445.
Mudholkar, G. S., Srivastava, D. K., & 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.
Pope, T. D., & Aiorachioia, D. (2013). Signal segmentation in time-frequency plane using renyi entropy-application in seismic signal processing. In 2013 Conference on Control and Fault-Tolerant Systems (SysTol) (pp. 312–317). Nice: IEEE.
Ranneby, B. (1984). The maximum spacing method. An estimation method related to the maximum likelihood method. Scandinavian Journal of Statistics, 11, 93–112.
Reiser, D. W., Wescot, T. A., & Estes, C. (1989). Status of instream flow legislation and practices in north america. Fisheries, 14, 22–29.
Renyi, A. (1961). On measures of entropy and information. In Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1, 47-561.
Stacy, E. W. (1962). A generalization of the gamma distribution. The Annals of Mathematical Statistics, 33, 1187–1192.
Torabi, H., Falahati-Noei, M., & Montazeri, N. (2014). An extended generalized lindley distribution and its applications to lifetime data. Journal of Statistical Research of Iran, 11, 203–222.
Wang, M., & Wang, W. (in press). Bias-corrected maximum likelihood estimation of the parameters of the weighted lindley distribution. Communications in Statistics-Simulation and Computation, 46, 530–545.
Zakerzadeh, H., & Dolati, A. (2009). Generalized lindley distribution. Journal of Mathematical Extension, 3(2), 1–17.
Appendix 1

The elements of the Fisher information matrix are

\[ I_{\phi,\phi} = -E \left[ \frac{\partial^2 \ell(\theta, t)}{\partial \phi^2} \right] = -\frac{1}{(\lambda + \phi)^2} + \psi'(\theta) \]

\[ I_{\phi,\lambda} = -E \left[ \frac{\partial \ell(\theta, t)}{\partial \phi \partial \lambda} \right] = -\frac{\alpha}{\lambda} + \frac{1}{(\lambda + \phi)^2} \]

\[ I_{\phi,\alpha} = -E \left[ \frac{\partial \ell(\theta, t)}{\partial \phi \partial \alpha} \right] = -\alpha \log(\lambda) - \psi(\phi) + \alpha \log(\lambda) - (\lambda + \phi)^{-1} \frac{1}{\alpha} \]

\[ I_{\lambda,\lambda} = -E \left[ \frac{\partial^2 \ell(\theta, t)}{\partial \lambda^2} \right] = \frac{\alpha \phi}{\lambda^2} + (\alpha - 1)\lambda^{-2} \psi(\phi) - \alpha \log(\lambda) + (\lambda + \phi)^{-1} \]

\[ + E \left[ \frac{a T^a \lambda^{-2} (\alpha - 2\lambda - (\lambda T)^a)}{(\lambda + (\lambda T)^a)} \right] - \frac{1}{(\lambda + \phi)^2} \]

\[ I_{\alpha,\alpha} = -E \left[ \frac{\partial^2 \ell(\theta, t)}{\partial \alpha^2} \right] = \frac{\phi(\lambda + \phi + 1)\left(\psi(\phi)^2 + \psi(\phi)\right)}{\alpha^2(\lambda + \phi)} + \frac{1}{\alpha^2} \]

\[ + 2(\lambda + 2\phi + 1)\psi(\phi) + 2 \left[ \frac{\lambda(\lambda T)^a \log(\lambda T)^a}{(\lambda + (\lambda T)^a)} \right] \]

\[ I_{\lambda,\alpha} = -E \left[ \frac{\partial \ell(\theta, t)}{\partial \alpha \partial \lambda} \right] = -\frac{\phi}{\lambda} + \frac{\lambda(1 + \phi \psi(\phi)) + \phi(1 + (\phi + 1)\psi(\phi) + 1)}{\lambda(\lambda + \phi)} \]

\[ - E \left[ \frac{(1 + a \lambda^{-1} T^a)(\lambda T)^a \log(\lambda T)^a}{(\lambda + (\lambda T)^a)} \right] + \frac{(\phi + \lambda + 1 - \frac{1}{\alpha}) \Gamma \left( \frac{\phi + 1 - \frac{1}{\alpha}}{(\lambda + \phi) \Gamma(\phi)} \right)}{(\lambda + (\lambda T)^a)^2} \]

\[ - E \left[ \frac{(a \lambda^{-1} T^a \log(\lambda T)^a + (\lambda T)^{a-1})}{(\lambda + (\lambda T)^a)} \right] \]