STATISTICS IN TRANSITION new series, December 2022
Vol. 23, No. 4, pp. 59–76, DOI 10.2478/stattrans-2022-0042
Received – 21.01.2022; accepted – 24.05.2022

The Weibull lifetime model with randomised failure-free time

Piotr Sulewski[1], **Magdalena Szymkowiak**[2]

ABSTRACT

The paper shows that treating failure-free time in the three-parameter Weibull distribution not a constant, but as a random variable makes the resulting distribution much more flexible at the expense of only one additional parameter.

Key words: Weibull lifetime model, randomised failure-free time, compound Weibull distributions.

1. Introduction

In the reliability domain we routinely treat time to failure of a particular technical product as a random variable. Finding the proper model that fits the reliability data is the main problem of reliability engineers and applied statisticians. The Weibull distribution, which has a particular significance in the reliability theory, is named after the Swedish physicist Waloddi Weibull, who was the first to promote the usefulness of the distribution to model reliability data sets of widely differing character (see, e.g. Weibull, 1951, and Murthy et al., 2004).

Recall that the two-parameter Weibull distribution ($2pW$) has the following cumulative distribution function (cdf)

$$F_{2pW}(t;a,b) = 1 - \exp\left[-\left(\frac{t}{a}\right)^b\right] \quad \text{for} \quad t > 0 \tag{1}$$

and the probability density function (pdf), based on (1), equal to

$$f_{2pW}(t;a,b) = \frac{b}{a}\left(\frac{t}{a}\right)^{b-1} \exp\left[-\left(\frac{t}{a}\right)^b\right] \quad \text{for} \quad t > 0, \tag{2}$$

where $a > 0$ and $b > 0$ are the scale and shape parameters, respectively. The hazard (failure) rate function (hrf) $h(t) = \frac{f(t)}{P(T>t)} = \frac{f(t)}{1-F(t)}$, interpreted as the instantaneous failure rate of a particular product occurring immediately after the time point t, given that the product has survived until the time point t, has for $2pW$, using (1) and (2), the following form

$$h_{2pW}(t;a,b) = ba^{-b}t^{b-1} \quad \text{for} \quad t > 0.$$

[1]Institute of Exact and Technical Sciences, Pomeranian University, Poland.
E-mail: piotr.sulewski@apsl.edu.pl.
[2]Institute of Automatic Control and Robotics, Poznan University of Technology, Poland.
E-mail: magdalena.szymkowiak@put.poznan.pl.
© Piotr Sulewski, Magdalena Szymkowiak. Article available under the CC BY-SA 4.0 licence

It can be increasing, decreasing or constant depending on $b > 1$, $b < 1$ or $b = 1$, respectively. It is easy to note that in the last mentioned case, when $b = 1$, we get the exponential distribution, the most standard distribution in the reliability theory, with a constant hazard rate. It is a well-known fact that the constant hazard rate function $h(t) = \frac{1}{a}$ characterizes the family of exponential distribution with scale parameter a.

Further, let us define another reliability function known as the aging intensity function (aif) in the form

$$L(t) = \frac{h(t)}{\frac{1}{t}\int_0^t h(u)\mathrm{d}u} = \frac{-tf(t)}{[1-F(t)]\ln[1-F(t)]} \quad \text{for} \quad t > 0, \tag{3}$$

being the ratio of the instantaneous hazard rate to its average and expressing the product average aging behaviour (see, e.g., Szymkowiak, 2018a). For $2pW$ it is constant

$$L_{2pW}(t;a,b) = b \quad \text{for} \quad t > 0. \tag{4}$$

This constant aging intensity $L(t) = b$ characterizes the subfamily of the family of $2pW$ with a fixed shape parameter b and varying scale parameter a (Szymkowiak, 2020). Moreover, the aif equal to 1, $L(t) = 1$, characterizes the family of exponential distributions.

However, certain lifetime data (i.a., human mortality, machine life-cycles and some biological studies) require non-monotonic shapes of the hazard rate, e.g., a bathtub shape or a unimodal (upside-down bathtub) shape. Therefore, many researchers have developed various modified forms of the Weibull distribution to achieve non-monotonic shapes of hazard function, i.a., Drapella, 1993, introduced the complementary Weibull distribution ($2pCW$), known also as the inverse Weibull distribution, with the following cdf

$$F_{2pCW}(t;a,b) = \exp\left[-\left(\frac{a}{t}\right)^b\right] \quad \text{for} \quad t > 0.$$

Further extensive literature is also available on modifications of the standard Weibull (see, e.g., Murthy et al., 2004, Almaki and Nadarajah, 2014, Lai, 2014), which in some cases involve one or more additional parameters. For example, the exponentiated Weibull distribution ($2pEW$) with a bathtub hazard rate function has the following cdf

$$F_{2pEW}(t;a,b,d) = \left\{1 - \exp\left[-\left(\frac{t}{a}\right)^b\right]\right\}^d \quad \text{for} \quad t > 0$$

with $d > 0$ being a new shape parameter (Mudholkar and Srivastava, 1993).

To be precise, Waloddi Weibull introduced his distribution as a three-parameter model $3pW$ (known also as the shifted Weibull distribution) with an additional location parameter $\tau \geq 0$ and the following cdf

$$F_{3pW}(t;a,b,\tau) = 1 - \exp\left[-\left(\frac{t-\tau}{a}\right)^b\right] \quad \text{for} \quad t > \tau, \tag{5}$$

pdf, based on (5), equal to

$$f_{3pW}(t;a,b,\tau) = \frac{b}{a}\left(\frac{t-\tau}{a}\right)^{b-1} \exp\left[-\left(\frac{t-\tau}{a}\right)^{b}\right] \quad \text{for} \quad t > \tau, \tag{6}$$

and hrf, based on (5) and (6), equal to

$$h_{3pW}(t;a,b,\tau) = ba^{-b}(t-\tau)^{b-1} \quad \text{for} \quad t > \tau.$$

Note that for a distribution with support $(\tau, +\infty)$ we use a modified version of aif, known as the support dependent aif (see Szymkowiak, 2018b)

$$L^{s}(t) = \frac{h(t)}{\frac{1}{t-\tau}\int_{\tau}^{t} h(u)\mathrm{d}u} = \frac{(\tau - t)f(t)}{[1-F(t)]\ln[1-F(t)]}, \tag{7}$$

which for $\tau = 0$ corresponds to the classical definition (3). Then, the subfamily of $3pW$ distributions with fixed shape parameter b and location parameter τ, and varying scale parameter, a is characterized by constant support dependent aif

$$L^{s}_{3pW}(t;a,b,\tau) = b \quad \text{for} \quad t > \tau,$$

(compare formula (4)).
Figure 1 presents hrf (on the left) and support dependent aif (on the right) of the $3pW(a,b,\tau)$ for the scale parameter $a = 1$ and the location parameter $\tau = 1$, and different values of shape parameter b. As one can note hrf is decreasing for $b < 1$ and increasing if $b > 1$. When $b = 1$, we get the shifted exponential distribution (see, e.g., Szymkowiak, 2020) with the constant hrf. On the other hand, the support dependent aging intensity functions, shown in the figure on the right, are constant, equal to b, for all $b > 0$.

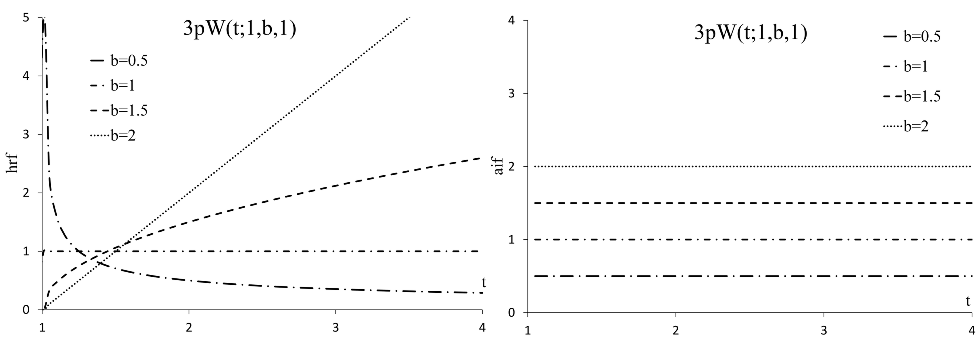

Figure 1: hrf and aif of the 3pW

Under the assumption of $3pW$ distribution, none failure of the analysed product can possibly occur prior to the time τ, therefore the location parameter τ is also referred to as failure-free time or minimum life.

In lifetime data analysis, first we try to find the proper model that best fits the data. Parameter estimation is the second step of our modeling process. In three-parameter Weibull distribution, the determination of a suitable location parameter τ is not a simple task.

The first and simplest τ estimate is $\hat{\tau} = t_{(1)}$, where $t_{(1)}$ is the smallest value in the order data set, or $\hat{\tau} = t_{(1)} - \frac{1}{n}$, where n is a sample size (see, e.g., Murthy et al., 2004). O'Connor, 2012, suggested an alternative procedure to estimate the location parameter given by

$$\hat{\tau} = t_{(2)} - \frac{\left(t_{(3)} - t_{(2)}\right)\left(t_{(2)} - t_{(1)}\right)}{\left(t_{(3)} - t_{(2)}\right) - \left(t_{(2)} - t_{(1)}\right)},$$

where further on $t_{(2)}$ and $t_{(3)}$ are the second and third value in the order data set. Drapella, 1999, improved this method and also Kececioglu, 1991, discussed two other methods to obtain estimates of τ.

In our paper, apart from the fact that we assume that the time to failure follows $3pW$ distribution, we also suggest that the failure-free time (location parameter) can be considered as a new random variable. It allows this parameter to vary and makes the estimation model more complex.

It is well-known fact (see, e.g., Qutb and Rajhi, 2016) that if X is a random variable following the known parametrized distribution with pdf f_X, and one of its parameters θ is considered as a new random variable Y with a specified pdf f_Y then a compound random variable T has a distribution with the following pdf

$$f_T(t) = \int_\theta f_X(t|\theta) f_Y(\theta) \mathrm{d}\theta \quad \text{for} \quad t > 0, \tag{8}$$

where $f_X(t|\theta)$ is a conditional density function depending on the parameter θ.

The compound Weibull distribution with random parameters was introduced earlier, e.g., by Dubey, 1968, and Qutb and Rajhi, 2016, but as far as we know, its location parameter has not yet been considered as being random.

The rest of our paper is organized as follows. In Section 2 the compound Weibull lifetime model with random failure free time is defined. In Section 3 four candidates for being the distribution of the random location parameter are presented. Section 4 contains analysis of three real lifetime data that compares the defined compound Weibull distributions with the standard three-parameter one. The conclusions are presented in Section 5.

2. Failure-free time as random variable

In this section the Weibull lifetime model with random failure-free time denoted as $4pWY$ is defined. Its pdf, according to formula (8), has a form of the convolution integral, namely

$$f_{4pWY}(t; a, b, c, d) = \int_0^t f_{3pW}(t; a, b, \tau) f_Y(\tau; c, d) d\tau \quad \text{for} \quad t > 0, \tag{9}$$

where f_Y is pdf of the failure-free time distribution with parameters c and d. For details related to the above formula please consult any advanced textbook on probability theory, e.g. Rossberg et al., 1985.

Cdf of the $4pWY$, based on (9), is given by

$$F_{4pWY}(t;a,b,c,d) = \int_0^t \left[1 - \exp\left(-\left(\frac{t-\tau}{a}\right)^b\right)\right] F_Y(\tau,c,d)\,d\tau \quad \text{for} \quad t>0, \tag{10}$$

where F_Y is cdf of the failure-free time distribution with parameters c and d.

Hrf of the $4pWY$, using (9) and (10), is obviously defined as

$$h_{4pWY}(t;a,b,c,d) = \frac{f_{4pWY}(t;a,b,c,d)}{1 - F_{4pWY}(t;a,b,c,d)} \quad \text{for} \quad t>0, \tag{11}$$

and aif of the $4pWY$ using (11) is given by

$$L_{4pWY}(t;a,b,c,d) = \frac{h_{4pWY}(t;a,b,c,d)}{\frac{1}{t}\int_0^t h_{4pWY}(u;a,b,c,d)\,du} \quad \text{for} \quad t>0. \tag{12}$$

Regarding the calculations, unfortunately, there will not always be analytical formulas to which (9)-(12) would be transformed. All applications of the $4pWY$ lifetime model will often be numerical. Fortunately, this is not an obstacle these days. Anyone who decides to use $4pWY$ to evaluate reliability must be equipped with a powerful computing environment. Fortunately, we have Excel, Mathcad, Matematica, Matlab, Scilab, and maybe a few other powerful, less known computing environments.

3. Four candidates for the failure-free time model

Now, four compound Weibull random variables with different distributions of the random location parameter will be presented. As the first distribution of the random location parameter we propose the Uniform distribution $U(c,d)$ giving a smooth transformation from $3pW$ to the compound Weibull-Uniform distribution $4pWU$. As the second very natural candidate – we propose the Weibull distribution $W(c,d)$. The next one will be the very popular two-parameter Gamma distribution $G(c,d)$ being the generalization of exponential, Erlang and chi-square distributions. As the last one we use the Normal distribution $N(c,d)$ with possibly positive support.

For all the proposed models, their hrf and aif for different parameters are determined and plotted (using formulas (11) and (12)). Determination of statistical measures of the presented random variables, such as their ordinary and central moments, quantiles, etc., because of their complex distribution forms, is possible only with the numerical calculations.

3.1. Compound Weibull-Uniform distribution

The first candidate for the failure-free time model is a Uniform distribution on interval $[c-d,c+d]$ denoted $U(c-d,c+d)$. For the purposes of the convolution integral (9) we can write pdf of the $U(c-d,c+d)$ using the Heaviside step function H in the form

$$f_U(t;c,d) = \frac{H(t-c+d) - H(t-c-d)}{2d} \quad \text{for} \quad t>0, c \geq d$$

where $c > 0$ and $d > 0$ are the position and scale parameters, respectively, as well as

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 0.5 & \text{for } x = 0 \\ 1 & \text{for } x > 0. \end{cases}$$

The Heaviside step function we can write as $H(x) = \int_{-\infty}^{x} \delta(s)\,ds$, where $\delta(x)$ is the Dirac function. Let us define the Dirac function $\delta(t;\tau)$ as

$$\delta(t;\tau) = \begin{cases} 0 & \text{for } t = \tau \\ \infty & \text{for } t \neq \tau. \end{cases}$$

The Dirac function fulfils, as any probability distribution, Normalization condition $\int_{-\infty}^{\infty} \delta(t;\tau)\,dt = 1$. This function can also be defined as the limit of the sequence of τ - centered Normal distributions, namely

$$\lim_{d \to 0} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-0.5\left(\frac{t-\tau}{\sigma}\right)^2\right\}.$$

Cdf of the $U(c-d, c+d)$ calculated in the Mathematica software is defined as

$$F_U(t;c,d) = I_1 + I_2,$$

where

$$I_1 = \frac{c+d-t-(c+d)H(-c-d)}{2b} H(t-c-d),$$

$$I_2 = \frac{-c+d+t+(c-d)H(d-c)}{2b} H(t-c+d).$$

Let $4pWU(a,b,c,d)$ denotes the four parameters compound Weibull-Uniform distribution. Then pdf of this distribution, based on (9), is given by

$$f_{4pWU}(t;a,b,c,d) = \frac{b}{2a^b d} \int_0^t \frac{H(\tau-c+d)-H(\tau-c-d)}{(t-\tau)^{1-b} \exp\left[\left(\frac{t-\tau}{a}\right)^b\right]} d\tau \quad \text{for} \quad t > 0. \tag{13}$$

The integral in (13) is possible to be calculated, namely

$$f_{4pWU}(t;a,b,c,d) = I_1 + I_2,$$

where

$$I_1 = \frac{b}{2a^b d} \int_0^t \frac{H(\tau-c+d)}{(t-\tau)^{1-b} \exp\left[\left(\frac{t-\tau}{b}\right)^b\right]} d\tau, \tag{14}$$

$$I_2 = \frac{b}{2a^b d} \int_0^t \frac{H(\tau-c-d)}{(t-\tau)^{1-b} \exp\left[\left(\frac{t-\tau}{b}\right)^b\right]} d\tau. \tag{15}$$

Using the Mathematica software by (14) we obtain

$$I_1 = \frac{H(t+d-c)\left[1+\exp\left[-\left(\frac{t+d-c}{a}\right)^b\right][H(d-c)-1]-\exp\left[-\left(\frac{t}{a}\right)^b\right]H(d-c)\right]}{2d}$$

and by (15), we get

$$I_2 = \frac{H(t-d-c)\left[1+\exp\left[-\left(\frac{t-d-c}{a}\right)^b\right][H(-c-d)-1]-\exp\left[-\left(\frac{t}{a}\right)^b\right]H(-c-d)\right]}{2d}.$$

It turns out that we can pass from $3pW$ (6) to $4pWU$ (13) smoothly. If $d \to 0$, then the $4pWU$ tends to the $3pW$ (see Figure 2).

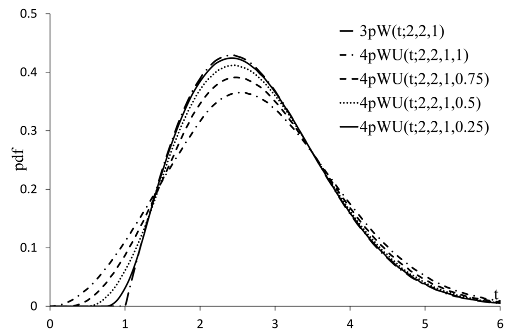

Figure 2: pdf of the $3pW$ and $4pWU$

Cdf of the $4pWU$, based on (13), is given by

$$\begin{aligned} F_{4pWU}(t;a,b,c,d) &= \frac{1}{2d}\int_0^t \left[1-\exp\left[-\left(\frac{t-\tau}{b}\right)^b\right]\right] \\ &\times \quad [H(\tau-c+d)-H(\tau-c-d)]\,d\tau \quad \text{for} \quad t>0. \end{aligned}$$

Figure 3 presents hrf of the $4pWU(a,b,c,d)$ for various parameter values. For $b = 0.5$ (Figure 3, left) hrf increases very quickly and then decreases very slowly. The maximum shifts to the right as d increases. For $d = 0.75$ (Figure 3, right) hrf increases slowly and then decreases very quickly. The maximum shifts to the left as b increases.

Figure 4 presents aif of the $4pWU(a,b,c,d)$ for various parameter values. For $a = 1$, $b = 0.5$, $c = 1$ regardless of the parameter d (Figure 4, left), aif (after some fluctuations for small t) tends to a constant function. For $a = 1$, $c = 1$, $d = 0.75$ regardless of the parameter b (Figure 4, right), aif decreases for small t and then also tends to a constant function.

To generate data that follows a compound Weibull-Uniform distribution we provide that if $T_{4pWU} \sim 4pWU(a,b,c,d)$ and $T_U \sim U(c-d,c+d)$, $T_{2pW} \sim 2pW(a,b)$, $R \sim U(0,1)$ then generator of T_{4pWU} is given by the formula

$$T_{4pWU} = T_{2pW} + T_U = a\left[-\ln(1-R)\right]^{1/b} + (c-d) + 2dR.$$

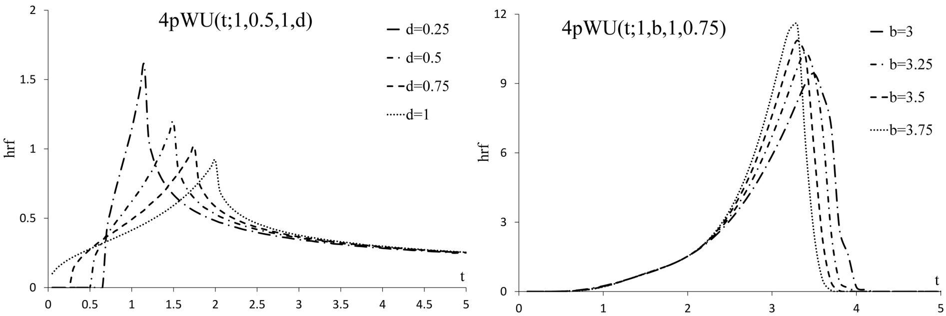

Figure 3: hrf of the $4pWU$

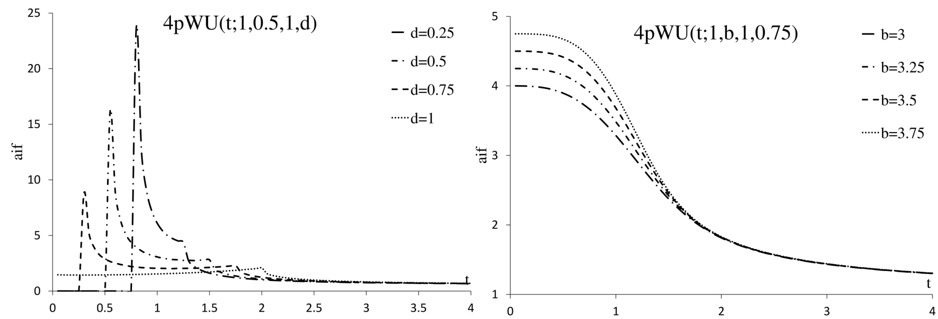

Figure 4: aif of the $4pWU$

3.2. Compound Weibull-Weibull distribution

The second candidate for the failure-free time model is the Weibull distribution with pdf given by formula (2). Let $4pWW(a,b,c,d)$ denotes the four parameters compound Weibull-Weibull distribution. Pdf of this distribution, based on (9), is given by

$$f_{4pWW}(t,a,b,c,d) = \frac{bd\int_0^t \left(\frac{t-\tau}{a}\right)^{b-1}\left(\frac{\tau}{c}\right)^{d-1}\exp\left[-\left(\frac{t-\tau}{a}\right)^{b}-\left(\frac{\tau}{c}\right)^{d}\right]d\tau}{ac} \quad t>0. \tag{16}$$

Cdf of the $4pWW$, based on (16), is given by

$$F_{4pWW}(t,a,b,c,d) = \frac{b\int_0^t \left[1-\exp\left[-\left(\frac{\tau}{c}\right)^{d}\right]\right]\left(\frac{t-\tau}{a}\right)^{b-1}\exp\left[-\left(\frac{t-\tau}{a}\right)^{b}\right]d\tau}{a} \quad t>0$$

Figure 5 presents hrf of the $4pWW(a,b,c,d)$ for various parameter values. For $d=0.5$ (see Figure 6, left) the hrf is a decreasing function. In other cases, the hrf increases strongly and then slowly decreases. The larger the d, the higher the maximum. For $b=2$ (see Figure 5, right) we obtain the inverse-bathtub hrf. For $b=1$ the hrf is initially an increasing function and then remains constant.

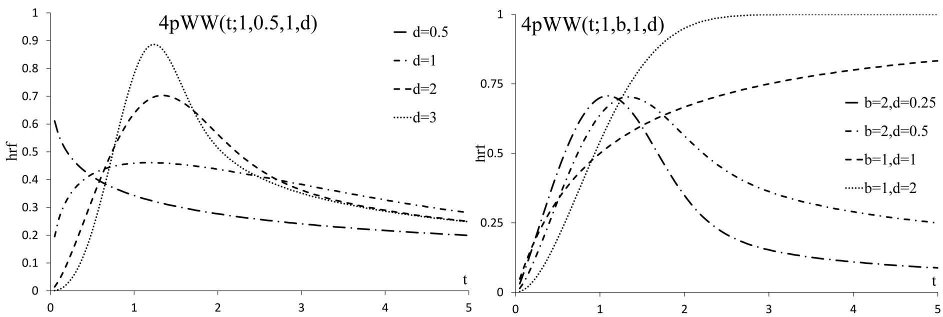

Figure 5: hrf of the $4pWW$

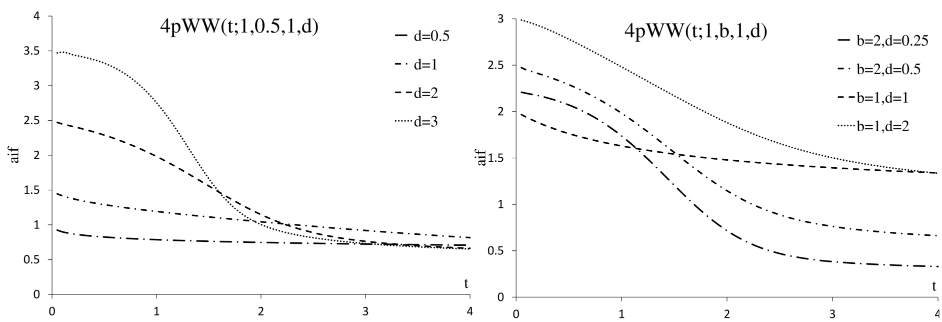

Figure 6: aif of the $4pWW$

Figure 6 presents aif of the $4pWW(a,b,c,d)$ for various parameter values. For $a=1$, $b=0.5$, $c=1$ regardless of the parameter d (Figure 6, left), aif decreases for small t and tends to a constant function. Also, for $a=1$, $b=2$ or $b=1$ and $c=1$, regardless of the parameter d (Figure 6, right), aif decreases for small t and then tends to a constant function.

To generate data that follows a compound Weibull-Weibull distribution we provide that if $T_{4pWW} \sim 4pWW(a,b,c,d)$, $T_{2pW1} \sim 2pW(a,b), T_{2pW2} \sim 2pW(c,d), R \sim U(0,1)$ then the generator of T_{4pWW} is given by the formula

$$T_{4pWW} = T_{2pW1} + T_{2pW2} = a\left[-\ln(1-R)\right]^{1/b} + c\left[-\ln(1-R)\right]^{1/d}.$$

3.3. Compound Weibull-Gamma distribution

The third candidate for the failure-free time model is the Gamma distribution with pdf

$$f_G(t;c,d) = \frac{1}{c^d \Gamma(d)} t^{d-1} \exp\left(-\frac{t}{c}\right) \quad \text{for} \quad t>0,$$

where $c>0, d>0$ are the scale and shape parameters, respectively.

Let $4pWG(a,b,c,d)$ denote the four-parameters compound Weibull-Gamma distribution then pdf of this distribution, based on (9), is given by

$$f_{4pWG}(t;a,b,c,d) = \frac{b\int_0^t (t-\tau)^{b-1}\,\tau^{d-1}\exp\left[-\frac{\tau}{c}-\left(\frac{t-\tau}{a}\right)^b\right]d\tau}{a^b c^d \Gamma(d)} \quad \text{for} \quad t>0. \tag{17}$$

Cdf of the $4pWG$, based on (17), is given by

$$F_{4pWG}(t;a,b,c,d) = \frac{\int_0^t \tau^{d-1}\exp\left[-\frac{\tau}{c}\right]\left[1-\exp\left[-\left(\frac{t-\tau}{a}\right)^b\right]\right]d\tau}{c^d\Gamma(d)} \quad \text{for} \quad t>0.$$

Decades pass, but the Weibull plotting technique remains irreplaceable in failure data analysis. Therefore, Figure 7 shows appropriately transformed cdf of the $2pW$, $3pW$ and $4pWG$ plotted on the Weibull probability paper. Of course, the transformed cdf of $2pW$ appears as straight lines. In contrast, both transformed cdf of $3pW$ and of $4pWG$ appear as curves convex upward. As a rule the transformed cdf of $4pWG$ is less convex then the transformed cdf of $3pWG$.

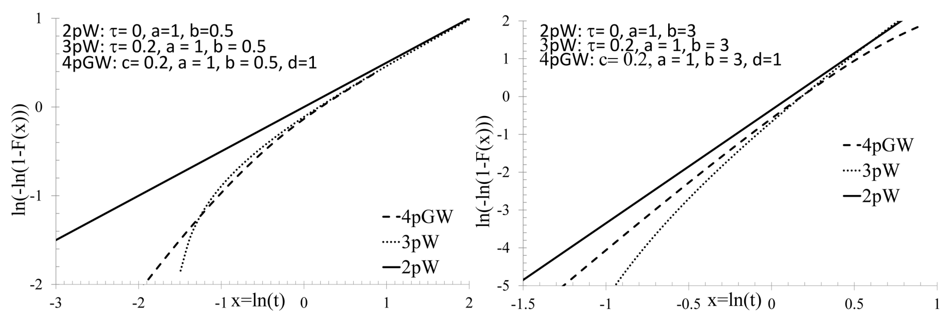

Figure 7: cdf of the $2pW$, $3pW$ and $4pWG$ plotted on the Weibull probability paper

Figure 8 has been prepared to express artificiality the stepwise $3pW$ (denoted as ps0) and compared it with the smooth $4pWG$. Table 1 contains parameter values of the Gamma component of the $4pWG$ used in Figure 8.

The $2pW$ and $3pW$ offer monotonic hazard rate functions, strictly decreasing or increasing ones, always convex downward. In contrast, $4pWG$ offers much more flexible hrf that may be non-monotonic and may have even two points of inflection. The main competitors of the $4pWG$ are the Complementary Weibull distribution (see Rossberg et al., 1985) and LogNormal distribution (see O'Connor, 2012). But they have two-parameter only. Let us look closely at Figure 8. One can pick up something resembling Moivre-Laplace limit process. Let us remember that when $p \to 0$, $n \to \infty$ and pn remains constant, the Binomial distribution tends to the Poisson distribution. By similarity, when $c \to 0$, $d \to \infty$ and cd remains constant, then $4pWG$ tends to $3pW$ (see Table 1).

Table 1. Sets of parameters of the Gamma component of the $4pWG$

Set of parameters	ps0*	ps1	ps2	ps3	ps4	ps5	ps6	ps7
c	0.2	0.2	0.1	0.05	0.033	0.025	0.02	0.01
d	-	1	2	4	6	8	10	20

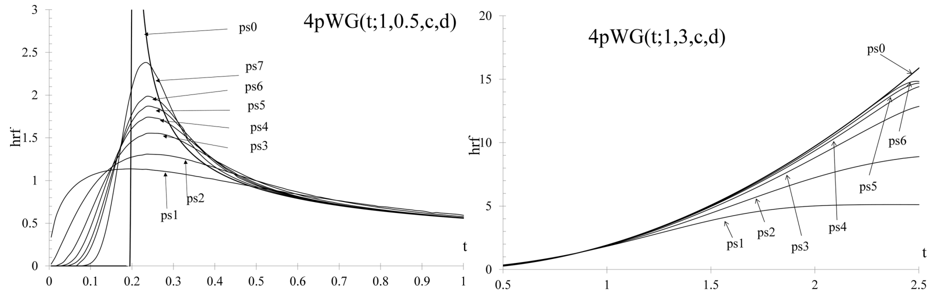

Figure 8: hrf of the $4pWG$

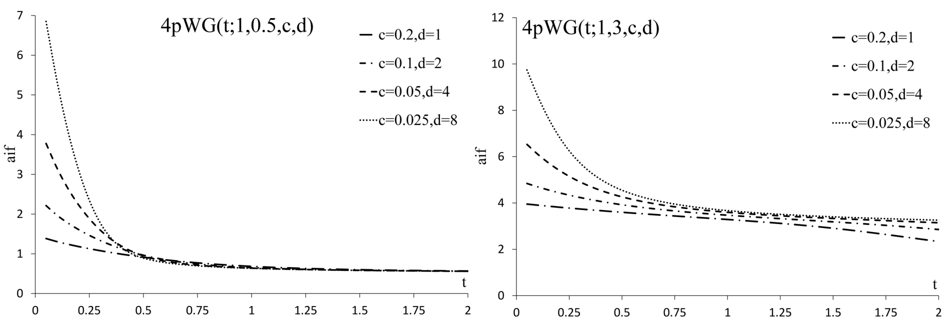

Figure 9: aif of the $4pWG$

Figure 9 presents aif of the $4pWG(a,b,c,d)$ for various parameter values. For $a = 1$, $b = 0.5$ (Figure 9, left) or $a = 1$, $b = 3$ (Figure 9, right), regardless of the parameters c and d, aif decreases for small t and tends to a constant function.

To generate data that follow a compound Weibull-Gamma distribution we provide that if $T_{4pWG} \sim 4pWG(a,b,c,d)$ and $T_G \sim G(c,d), T_{2pW} \sim 2pW(a,b)$ then the generator of T_{4pWG} is given by the formula

$$T_{4pWG} = T_{2pW} + T_G = a\left[-\ln(1-R)\right]^{1/b} + T_G$$

The generator of T_G, implemented in R software, for $d \geqslant 1$ and $0 < d < 1$ is described in Ahrens and Dieter, 1982, and Ahrens and Dieter, 1974, respectively.

3.4. Compound Weibull-Normal distribution

Many researchers may be tempted to replace the Gamma distribution with the Normal distribution with pdf

$$f_N\left(t;c,d\right)=\frac{1}{\sqrt{2\pi}d}\exp\left[-0.5\left(\frac{t-c}{d}\right)^2\right]\quad\text{for}\quad t>0,$$

where $c>0,d>0$ are the position and scale parameters, respectively (to ensure that Normal distribution has positive support – the failure-free time should not be negative – using the three-sigma rule we assume that $c>3.3d$). The argument was that when the shape parameter d increases, then the Gamma distribution tends to the Normal distribution. It directly proceeds from Lindeberg-Levy Limit theorem. Such a replacement of the Gamma distribution with the Normal one, has both advantages and disadvantages.

The advantages are:

- Firstly, one can skip this strong assumption that between-two-portions time interval follows the exponential distribution. Taking the Central Limit Theorem of Lapunov as a base, one can admit that particular intervals follow different distributions.

- Secondly, applying the Normal distribution as that between-two-portions time interval distribution, one makes $4pWN$ more flexible. It is because the mean value $E\left(t\right)$ and variance $D\left(t\right)$ of the Gamma distribution are strongly interrelated because $E\left(t\right)=cd$, $D\left(t\right)=cd^2$.

- Thirdly, the problem of interpretation of non-integer d value disappears.

The disadvantages are:

- Firstly, the moments mentioned above are to some extent interrelated. It is because the value of the coefficient of variation $\gamma_0=D\left(t\right)/E\left(t\right)$ has to be carefully chosen for probability of negative t values to be negligible, for instance γ_0 has to be kept not greater than $1/3$.

- Secondly, it is true that the left-censored Normal distribution can alternatively be applied, and values of $\mathrm{E}(t)$ and $\mathrm{D}(t)$ can be freely set, but then the threshold function, that we wanted to eliminate, returns.

Let $4pWN(a,b,c,d)$ denotes the four-parameters compound Weibull-Normal distribution. Then, pdf of this distribution, based on (9), is given by

$$\begin{aligned}f_{4pWN}\left(t;a,b,c,d\right) &= \frac{b}{\sqrt{2\pi}ad}\int_0^t\left(\frac{t-\tau}{a}\right)^{b-1} \\ &\times \exp\left[-\frac{1}{2}\left(\frac{\tau-c}{d}\right)^2-\left(\frac{t-\tau}{a}\right)^b\right]d\tau\quad\text{for}\quad t>0,\ c>3.3d.\end{aligned}\tag{18}$$

Cdf of the $4pWN$, using (18), is given by

$$F_{4pWN}(t;a,b,c,d) = \frac{b\int_0^t \Phi\left(\frac{\tau-c}{d}\right)\left(\frac{t-\tau}{a}\right)^{b-1}\exp\left[-\left(\frac{t-\tau}{a}\right)^{b}\right]d\tau}{a} \quad \text{for } t>0,\ c>3.3d,$$

where Φ is cdf of the standard Normal distribution.

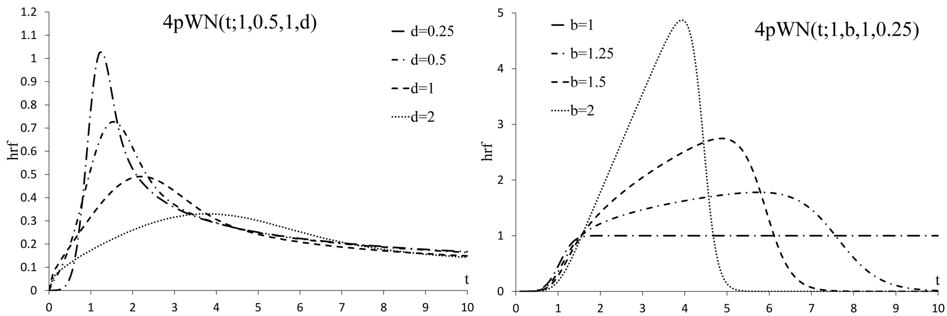

Figure 10: hrf of the 4pWN

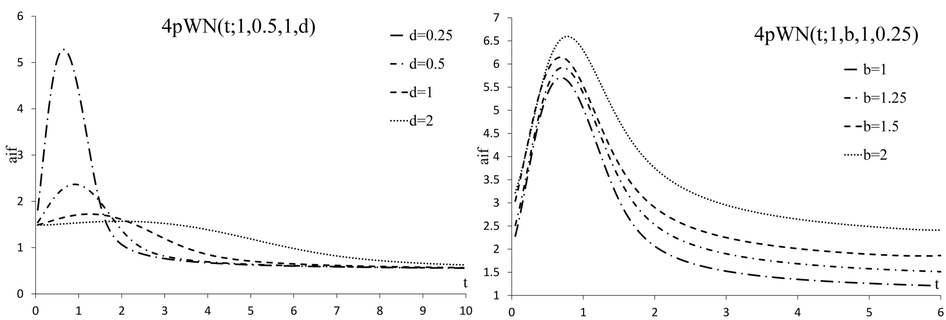

Figure 11: aif of the 4pWN

Figure 10 presents hrf of the $4pWN(a,b,c,d)$ for various parameter values. For $b=0.5$ (see Figure 10, left) hrf increases very quickly and then decreases very slowly. The maximum shifts to the right as d increases. Figure (10,right) presents an upside-down bathtub-shaped hrf except the case $b=1$. Then, initially hrf increases and then remains constant.

Figure 11 presents aif of the $4pWN(a,b,c,d)$ for various parameter values. For $a=1$, $b=0.5$, $c=1$, regardless of the parameter d (Figure 11, left), aif (after some fluctuations for small t) tends to the constant function. For $a=1$, $c=1$, $d=0.25$, regardless of the parameter b (Figure 11, right), aif decreases for small t and tends to the constant function.

To generate data that follows a compound Weibull-Normal distribution we provide that if $T_{4pWN} \sim 4pWN(a,b,c,d)$ and $T_N \sim N(c,d), T_{2pW} \sim 2pW(a,b), R \sim U(0,1)$ then the generator of T_{4pWN} is given by the formula

$$T_{4pWN} = T_{2pW} + T_N = a\left[-\ln(1-R)\right]^{1/b} + \Phi^{-1}(R,c,d),$$

where Φ^{-1} is the inverse cdf of the Normal distribution implemented in R software (see Wichura, 1988).

At the end of this section, let us mention that any infinitely divisible (stable) probability distribution (see Seidel, 2010) with the positive support would be a good candidate for the failure-free time model. The only problem is to determine basic statistical functions for such a candidate.

4. Real lifetime data analysis

To demonstrate the flexibility and applicability of the new compound models in lifetime data analysis we consider three real data examples. Among the known random variables, the standard $3pW$ distribution seems to be the most natural choice for comparing the goodness-of-fit approach of the proposed models.

Example 1 is devoted to 33 operating times (in hours) of the first phase of the construction machine: 0.25, 9.25, 17.75, 22, 22.25, 22.5, 22.5, 23, 23.25, 30.25, 35.75, 41.75, 42.5, 48.75, 49.75, 51, 58.75, 63, 68.25, 72.5, 75.5, 127.5, 138.5, 140.5, 141, 146.5, 173, 193.5, 218.25, 237.5, 257.75, 312, 352 (see, e.g., Mahmood, 2021, Saffawy and Algmal, 2006).

Example 2 concerns leukaemia free survival times (in months) of 51 autologous transplant patients: 0.658, 0.822, 1.414, 2.500, 3.322, 3.816, 4.737, 4.836, 4.934, 5.033, 5.757, 5.855, 5.987, 6.151, 6.217, 6.447, 8.651, 8.717, 9.441, 10.329, 11.480, 12.007, 12.007, 12.237, 12.401, 13.059, 14.474, 15.000, 15.461, 15.757, 16.480, 16.711, 17.204, 17.237, 17.303, 17.644, 18.092, 18.092, 18.750, 20.625, 23.158, 27.730, 31.184, 32.434, 35.921, 42.237, 44.638, 46.480, 47.467, 48.322, 56.086 (see, e.g., LaiXie, 2006).

Example 3 refers to survival times (in days from diagnosis) of 43 patients suffering chronic granulocytic leukaemia: 7, 47, 58, 74, 177, 232, 273, 285, 317, 429, 440, 445, 455, 468, 495, 497, 532, 571, 579, 581, 650, 702, 715, 779, 881, 900, 930, 968, 1077, 1109, 1314, 1334, 1367, 1534, 1712, 1784, 1877, 1886, 2045, 2056, 2260, 2429, 2509 (see, e.g., Lai and Xie, 2006).

To estimate parameters of the considered models, in addition to commonly known estimation tools such as the maximum likelihood (not quite adequate in the case of three-parameter Weibull distribution, (see e.g. Murthy et al., 2004, Lam, 2010, Ramakrishnan, 2017, Park, 2018) or least squares methods, also goodness-of-fit tests can be used. For example, Kendall and Stuart, 1961 presented the minimum chi-square test statistic method in parameter estimation. Moreover, Weber, 2006, used the minimum Kolmogorov–Smirnov test statistic method to estimate the distribution parameters. In our analysis we apply the latter tool.

To avoid local maxima, the optimization routine was run with several different starting values that are widely scattered in the parameter space. The p-values for a given model were calculated as follows. Let Θ be the vector of model parameters. Having estimated parameters vector $\widehat{\Theta}$ for a given sample of size n, we calculate test statistics $T\left(\widehat{\Theta}, n\right)$. Next, we generate 10^4 samples of size n for the given model with the estimated parameters vector $\widehat{\Theta}$. For each obtained sample s, we calculated the $T_i^s\left(\widehat{\Theta},n\right)$. Finally, the p-value is given by (see e.g. Balakrishnan and Ristic, 2016)

$$p \approx \#\left\{i : T_i^s\left(\widehat{\Theta},n\right) > T\left(\widehat{\Theta},n\right)\right\} 10^{-4}$$

Tables 2–4 present the estimated parameters of the analysed models, test statistics and p-values (in parentheses) calculated by the Kolmogorov-Smirnow (KS), Anderson-Darling (AD) and Cramer von Mises (CVM) tests. The lowest statistics values (the highest p-values) are noted in bold. The determined values show that for all the exemplary lifetime data sets, there are some compound Weibull distributions that fit better to the data then the standard three-parameter Weibull distribution $3pW$. For the first data set, two tests point to the compound Weibull-Gamma distribution, $4pWG$, as the distribution that best fits the data, and one test points to the compound Weibull-Weibull distribution, $4pWW$, as the best one (see Table 2). Further, for the second data set, all the tests point to the compound Weibull-Weibull distribution, $4pWW$, as the distribution that best fits the data (see Table 3). Finally, for the third data set, two tests point to the compound Weibull-Uniform distribution, $4pWU$, and one test points at the compound Weibull-Weibull distribution, $4pWW$, as the best models (see Table 4).

Table 2. Goodness-of-fit tests. Example 1

Model	Estimated parameters	KS	AD	CVM
$3pW$	$\widehat{a}=96.91, \widehat{b}=1.045, \widehat{\tau}=0.06$	0.1(0.864)	0.543(0.698)	0.082(0.673)
$4pWU$	$\widehat{a}=90.48, \widehat{b}=0.92, \widehat{c}=5.76, \widehat{d}=5.81$	0.094(0.910)	0.513(0.727)	0.062(0.810)
$4pWW$	$\widehat{a}=90.90, \widehat{b}=0.93, \widehat{c}=4.71, \widehat{d}=0.79$	0.095(0.901)	0.496(0.747)	0.067(0.764)
$4pWG$	$\widehat{a}=80.90, \widehat{b}=0.69, \widehat{c}=4.99, \widehat{d}=2.87$	0.083(0.962)	0.660(0.586)	0.045(0.901)
$4pWN$	$\widehat{a}=91.15, \widehat{b}=0.93, \widehat{c}=5.55, \widehat{d}=0.92$	0.094(0.913)	0.994(0.359)	0.064(0.793)

Table 3. Goodness-of-fit tests. Example 2

Model	Estimated parameters	KS	AD	CVM
$3pW$	$\widehat{a}=15.92, \widehat{b}=1.20, \widehat{\tau}=0.64$	0.076(0.910)	0.673(0.580)	0.071(0.746)
$4pWU$	$\widehat{a}=15.56, \widehat{b}=1.29, \widehat{c}=0.48, \widehat{d}=0.50$	0.071(0.945)	0.542(0.713)	0.065(0.797)
$4pWW$	$\widehat{a}=13.90, \widehat{b}=1.32, \widehat{c}=0.36, \widehat{d}=0.28$	0.070(0.946)	0.469(0.778)	0.062(0.802)
$4pWG$	$\widehat{a}=14.64, \widehat{b}=1.12, \widehat{c}=1.77, \widehat{d}=0.99$	0.074(0.928)	0.573(0.678)	0.066(0.776)
$4pWN$	$\widehat{a}=15.50, \widehat{b}=1.17, \widehat{c}=1.03, \widehat{d}=0.21$	0.075(0.921)	0.836(0.459)	0.070(0.759)

Table 4. Goodness-of-fit tests. Example 3

$3pW$	$\widehat{a}=993.19, \widehat{b}=1.18, \widehat{\tau}=4.37$	0.08(0.928)	0.435(0.812)	0.048(0.891)
$4pWU$	$\widehat{a}=984.62, \widehat{b}=1.18, \widehat{c}=4.85, \widehat{d}=4.5$	0.077(0.949)	0.431(0.834)	0.047(0.908)
$4pWW$	$\widehat{a}=993.61, \widehat{b}=1.18, \widehat{c}=0.96, \widehat{d}=1.5$	0.077(0.943)	0.392(0.866)	0.046(0.901)
$4pWG$	$\widehat{a}=957.54, \widehat{b}=1.18, \widehat{c}=3.55, \widehat{d}=0.4$	0.088(0.863)	0.403(0.847)	0.047(0.892)
$4pWN$	$\widehat{a}=991.98, \widehat{b}=1.17, \widehat{c}=4.94, \widehat{d}=0.6$	0.077(0.941)	0.428(0.817)	0.046(0.904)

5. Conclusions

In the paper, the failure-free time, being the location parameter of the shifted Weibull distribution, was proposed to be treated as a random variable. We defined four compound Weibull distributions with the location parameter having Uniform, Weibull, Gamma and Normal distribution, respectively. Using these proposed models the analysis of three real lifetime data sets were performed. The received results showed that the new models fit better the data under consideration that the standard three-parameter Weibull distribution.

However, anyone who will decide to use any of the proposed compound Weibull distributions in data analysis has to be equipped with a powerful computational environment. Luckily, nowadays it is not a problem since we have Excel, Mathcad, Mathematica, Matlab, Scilab and maybe some other not so widely known computational tools.

Acknowledgements

The authors thank Prof. Antoni Drapella, former Director of Institute of Mathematics at Pomeranian University in Słupsk, for providing assistance in formulating basics of the method presented. The second author was partially supported by PUT under grant 0211/SBAD/0121.

Conflict of interest. The authors declare that they have no conflict of interest.

References

Ahrens, J. H., Dieter, U. (1982). Generating gamma variates by a modified rejection technique. *Communications of the ACM*, 25, pp. 47–54.

Ahrens, J. H., Dieter, U. (1974). Computer methods for sampling from gamma, beta, poisson and binomial distributions. *Computing*, 12, pp. 223–246.

Alamlki, S. J.,Nadarajah, S., (2014). ,Modifications of the Weibull distribution: a review.*Reliability Engineering and System Safety*, 124, pp. 32–55.

Balakrishnan, N., Ristic, M. M., (2016). Multivariate families of gamma-generated distributions with finite or infinite support above or below the diagonal, *Journal of Multivariate Analysis*, 143, pp. 194–207.

Drapella, A., (1993). Complementary Weibull distribution: Unknown or just forgotten. *Quality and Reliability Engineering International*, 9, pp. 383–385.

Drapella, A., (1999). An improved failure-free time estimation method. *Quality and Reliability Engineering International*, 15, pp. 235–238.

Dubey, S. D. (1968). A compound Weibull distribution. *Naval Research Logistics Quarterly*, 15, pp. 179–188.

Gertsbakh, I. B., Kordonskiy, K. H. B., (1969). *Models of failure*, Verlag: Springer.

Kao, J. H. K. (1966). *Lifetime models with applications*, In: Reliability Handbook. W.G. Ireson Editor-in-chief. McGraw-Hill Company.

Kao, J. H. K. (1960). A summary of some new techniques on failure analysis, *Proc. Sixth Natl. Symp. on Reliability and Quality Control*.

Kececioglu, D. (1991). *Reliability Engineering Handbook*, New York: Prentice Hall, Eaglewood Cliffs.

Kendall, M. G., Stuart, A. (1961). *The advanced theory of statistics, Vol. 2*, Charles Griffin and Company.

Lai, C. D. (2014). *Generalized Weibull Distributions*, New York: Springer.

Lai, C. D., Xie, M. (2006). *Stochastic ageing and dependence for reliability*, Springer Science and Business Media.

Lam, S. W., Halim, T., Muthusamy, K. (2010). Models with failure-free life—Applied review and extensions, *IEEE Transactions on Device and Materials Reliability*, 10(2), pp. 263–270.

Mahmood, S. W., Algamal, Z. Y. (2021). Reliability Estimation of Three Parameters Gamma Distribution via Particle Swarm Optimization, *Thailand Statisticia*, 19(2), pp. 308–316.

Mudholkar, G. S., Srivastava, D. K., (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data *IEEE Transactions on Reliability*, 42, pp. 299–302.

Murthy, D. N. P., Xie, M., Jiang, R., (2004). *Weibull models*, Hoboken: Wiley.

O'Connor, P. D. T. (1985). *Practical reliability engineering*, New York: Wiley.

Park, C. (2018). A Note on the Existence of the Location Parameter Estimate of the Three-Parameter Weibull Model Using the Weibull Plot, *Mathematical Problems in Engineering*, 10(2), pp. 1–6.

Qutb, N., Rajhi, E., (2016). Estimation of the Parameters of Compound Weibull Distribution, *IOSR Journal of Mathematics*, 12, pp. 11–18.

Ramakrishnan, M., Viswanathan, N., (2017). Comparing the methods of estimation of three-parameter Weibull distribution, *IOSR Journal of Mathematics*, 13(1), pp. 42–45.

Rossberg, H. J., Jesiak, B., Siegel, G., (1985). *Analytic Methods of Probability Theory*, Berlin: Academie Verlag.

Saffawy, S. Y., Algmal, Z. Y (2006). The Use of Maximum Likelihood and Kaplan-Meir method to Estimate the Reliability Function An Application on Babylon Tires Factory, *Tanmiyat Al-Rafidain*, 82(28), pp. 9–20.

Seidel, W. (2010). *Mixture model*, In Lovric, M., International Encyclopedia of Statistical Science, Heidelberg: Springer.

Stacy, E.W., (1962). A generalization of the gamma distribution, *The Annals of Mathematical Statistics*, pp. 1187–1192.

Szymkowiak, M. (2018a). Characterizations of distributions through aging intensity, *IEEE Transactions on Reliability*, 67(2), pp. 446–296.

Szymkowiak, M. (2018b). Generalized aging intensity functions, *Reliability Engineering and System Safety*, 178, pp. 198–208.

Szymkowiak, M., (2020). *Lifetime analysis by aging intensity functions*, Cham: Springer.

Weber, M. D., Leemis, L. M., Kincaid, R. (2006). Minimum Kolmogorov–Smirnov test statistic parameter estimates, *Journal of Statistical Computation and Simulation*, 76(3), pp. 195–206.

Weibull, W. (1951). A statistical distribution function of wide applicability, *Journal of Applied Mechanics*, 18, pp. 29–296.

Wichura, M. J., (1988). Algorithm AS 241: The percentage points of the normal distribution, *Applied Statistics*, 37, pp. 477–484.