

# On a Comparative Study on Some Trigonometric Classes of Distributions by the Analysis of Practical Data Sets

Christophe Chesneau<sup>1,†</sup> and Amaury Artault<sup>1</sup>

**Abstract** The aim of this paper is twofold. Firstly, we elaborate and investigate a new trigonometric class of distribution, called the type II Tan-G class. Secondly, we perform a practical comparative evaluation of certain trigonometric classes; the Sin-G, Cos-G, Tan-G classes and the new one, with each other and with their common baseline distribution. More specifically, the usefulness and flexibility of these trigonometric classes are demonstrated through twelve practical data sets, by using the Weibull distribution as baseline. Among the data sets, two of them concern the Covid-19 pandemic in France from March to June 2020. As main results, it is shown that the related trigonometric models can outperform the former Weibull model in various cases and that the proposed type II Tan Weibull model can be, in certain situations, the best of them. The main lines of the code written in the R software language are provided.

**Keywords** Trigonometric classes of distributions, Weibull distribution, Data fitting, Covid-19 pandemic.

**MSC(2010)** 62N01, 62N02, 62E10.

## 1. Introduction

In recent years, there has been great interest in the general classes of trigonometric distributions. This interest lies in the fact that such classes are fairly simple to handle, with well documented and accessible mathematical properties, and they generally offer excellent fits of different types of practical data sets. In this paper, we bring an assessment to these trigonometric classes of distributions. Before going further, let us briefly present them. We begin with the most popular one: the Sin-G class of distributions introduced by [11] and [19] and completed by [20]. The starting point is an univariate baseline distribution, defined by a cumulative distribution function (cdf) denoted by  $G(x)$ , with a probability density function (pdf) denoted by  $g(x)$ . This baseline distribution is very general; it can be of any support and have no, one or more parameters. Then, the Sin-G class is defined by the cdf given as

$$F(x) = \sin \left[ \frac{\pi}{2} G(x) \right], \quad x \in \mathbb{R},$$

---

<sup>†</sup>the corresponding author.

Email address: [christophe.chesneau@gmail.com](mailto:christophe.chesneau@gmail.com)(C. Chesneau), [amaury.artault@gmail.com](mailto:amaury.artault@gmail.com)(A. Artault)

<sup>1</sup>Université de Caen Normandie, LMNO, Campus II, Science 3, 14032, Caen, France

and possesses the pdf specified by

$$f(x) = \frac{\pi}{2}g(x) \cos \left[ \frac{\pi}{2}G(x) \right], \quad x \in \mathbb{R}.$$

The qualities of the Sin-G class are the following ones: (i) It is based on simple functions, (ii) no new parameter is added; the Sin-G distribution only depends on the parameters involved in the baseline distribution, (iii) it satisfies the following first-order stochastic dominance:  $F(x) \geq G(x)$ ,  $x \in \mathbb{R}$ , and (iv) the related models can be easily enriched for some baseline distributions. The qualities of the Sin-G class have motivated the constructions of other trigonometric classes. The two most famous of them are the Cos-G class of distributions provided by [19] and [21], and the Tan-G class given by [19], [22] and [2], which will be treated in our study.

In this paper, we are also going to introduce a new class of trigonometric distributions along side the more classics ones, called the type II Tan-G (TIIT-G) class. With the above notations, it is defined by the cdf given as

$$F(x) = 1 - \tan \left[ \frac{\pi}{4}(1 - G(x)) \right], \quad x \in \mathbb{R}, \quad (1.1)$$

the corresponding pdf being defined by

$$f(x) = \frac{\pi}{4}g(x) \frac{1}{\left\{ \cos \left[ \frac{\pi}{4}(1 - G(x)) \right] \right\}^2}, \quad x \in \mathbb{R}. \quad (1.2)$$

The definition of  $F(x)$  is based on the second type of the T-X transformation by [1] applied to the unit tangent cdf as generator; we can write it as

$$F(x) = 1 - F_o(1 - G(x)),$$

where  $F_o(x) = \tan[(\pi/4)x]$ . The first type of the T-X transformation applied to  $F_o(x)$  providing the cdf of the former Tan-G class defined by  $F_*(x) = F_o(G(x))$ . With this remark in mind, we see that the structural relationship between the TIIT-G and Tan-G classes is the same as that between the Sin-G and Cos-G classes. The definition of the TIIT-G class is therefore legitimate in this regard. As special features, (i) the TIIT-G class is a sub-class of the original T-X class by [1] (as shown later), (ii) the TIIT-G class also corresponds to the class combining the M class by [12] and the Tan-G class (as shown later), and (iii) from the inequality  $\tan(y) \leq (4/\pi)y$  for  $y \in [0, \pi/4]$ , we deduce the following first-order stochastic dominance:  $F(x) \geq G(x) \geq F_*(x)$ ,  $x \in \mathbb{R}$ . Thus, the TIIT-G class can be considered as an alternative to the Tan-G distribution class, while retaining similar flexibility and simplicity. It offers a new modeling option, extending the scope of the existing trigonometric classes of distributions. In the paper, we develop this point by exploring various of its theoretical and applied aspects.

As main part, applications are made to expose the wide coverage of different practical data sets which can be well fitted by trigonometric classes of distributions, by adopting the Weibull distribution as the baseline. These applications also exhibit the usefulness of each of the above classes depending on the data set. Moreover, in regard to the difficult times we are in living with the Covid-19 threat, some of these applications are about this pandemic. More precisely, it concerns the Covid-19 pandemic in France from March to June 2020. The objective is to offer a model that fits as best possible, aiming to be better prepared in the future in the face of

another pandemic with correct forecasts of its expansion. The obtained results are quite satisfying, supported by remarkable values of solid goodness-of-fit criteria and graphics.

We organize the paper as follows. Section 2 completes the presentation of the TIIT-G class. The statistical methodology and applications are given in Section 3. A conclusion is proposed in Section 4. The paper ends with Appendix containing the essential of the code used in this paper, written in the R software language.

## 2. Type II Tan-G class of distributions

This section is devoted to the TIIT-G class, including more relations between him and some existing classes, the description of other related functions and some mathematical properties.

### 2.1. Complements

As already evoked, the TIIT-G class is defined by the cdf given by Equation (1.1), with pdf specified by Equation (1.2). Formerly, it is based on the "second type (type II) of the T-X transformation" by [1] applied to the unit tan cdf as generator. New remarks are developed below. After some developments, we can express the cdf of the TIIT-G class under the following integral form:  $F(x) = \int_0^{G(x)} p(t)dt$ , where  $p(t)$  is the pdf defined by

$$p(t) = \frac{\pi}{2} \frac{1}{1 + \sin(\frac{\pi}{2}t)}, \quad t \in (0, 1),$$

and equal to 0 for  $t \notin (0, 1)$ . As far as we know, this pdf on the unit interval is not listed in the literature. The expression above shows that the TIIT-G class is a subclass of the original T-X class proposed by [1].

A more original expression comes from the scholar trigonometric formula:  $\tan(x+y) = (\tan(x) + \tan(y))/(1 - \tan(x)\tan(y))$ , for  $x, y \in \mathbb{R}$  (such that all the involved functions are well defined); we can express  $F(x)$  as

$$F(x) = 1 - \frac{1 - \tan\left[\frac{\pi}{4}G(x)\right]}{1 + \tan\left[\frac{\pi}{4}G(x)\right]} = \frac{2F_*(x)}{1 + F_*(x)},$$

where  $F_*(x) = \tan[(\pi/4)G(x)]$ , recalling that is the cdf of the Tan-G class. This expression reveals that the TIIT-G class is also a sub-class of the M family proposed by [12], defined with the Tan-G class as baseline. As is known to us all, the TIIT-G class is one of the few classes of distributions to be a sub-class of both "second type of T-X" and M classes defined with the same baseline class, and here comes the Tan-G class.

### 2.2. On other functions of interest

Now, let us present some functions of interest. The survival function (sf) and hazard rate function (hrf) of the TIIT-G class are defined by

$$\bar{F}(x) = 1 - F(x) = \tan\left[\frac{\pi}{4}(1 - G(x))\right]$$

and

$$h(x) = \frac{f(x)}{F(x)} = \frac{\pi}{2} g(x) \frac{1}{\cos \left[ \frac{\pi}{2} G(x) \right]}, \quad x \in \mathbb{R} \quad (2.1)$$

respectively. As for  $f(x)$ , the possible shapes of  $h(x)$  are strong indicators of the capability of the related statistical model to fit data; more these shapes are pliant, the more the model is able to capture the details of a wide variety of data sets. This is particularly true for lifetime data sets.

The quantile function (qf) of the TIIT-G class is derived by inverting  $F(x)$ . After some developments, we get

$$H(y) = Q \left[ 1 - \frac{4}{\pi} \arctan(1 - y) \right], \quad y \in (0, 1), \quad (2.2)$$

where  $Q(y)$  denotes the qf of the baseline distribution. The qf allows us to express the quartiles:  $Q_1 = H(1/4)$ ,  $Q_2 = H(1/2)$  and  $Q_3 = H(3/4)$ , octiles and various measures of skewness and kurtosis, such as the MacGillivray skewness and Moors kurtosis, among others. It is also the main tool to generate values for any distribution of the TIIT-G class, through the technique called inverse transform sampling.

For example, let us consider the Weibull distribution with parameters  $a$  (scale) and  $b$  (shape) as baseline, i.e. with cdf and pdf given as

$$G(x) = 1 - e^{-ax^b}, \quad g(x) = abx^{b-1}e^{-ax^b}, \quad x > 0, \quad (2.3)$$

and both equal to 0 for  $x \leq 0$ . The corresponding TIIT-G distribution is called the type II Tan Weibull (TIITW) distribution. Based on Equations (1.1) and (2.3), it is defined with the cdf given as

$$F(x) = 1 - \tan \left( \frac{\pi}{4} e^{-ax^b} \right), \quad x > 0, \quad (2.4)$$

and equal to 0 for  $x \leq 0$ . Also, based on Equations (1.2), (2.1) and (2.3), we can express the corresponding pdf and hrf as

$$f(x) = \frac{\pi}{4} abx^{b-1} e^{-ax^b} \frac{1}{\left[ \cos \left( \frac{\pi}{4} e^{-ax^b} \right) \right]^2}$$

and

$$h(x) = \frac{\pi}{2} abx^{b-1} e^{-ax^b} \frac{1}{\sin \left( \frac{\pi}{2} e^{-ax^b} \right)}, \quad x > 0,$$

and both equal to 0 for  $x \leq 0$  respectively. Graphical analysis reveals a plethora of different shapes for these functions, attesting the flexibility of the TIITW distribution. Finally, based on Equation (2.2) and the qf of the Weibull distribution given as  $Q(y) = [-\log(1 - y)/a]^{1/b}$ ,  $y \in (0, 1)$ , the qf of the TIITW distribution is specified by

$$H(y) = \left\{ -\frac{1}{a} \log \left[ \frac{4}{\pi} \arctan(1 - y) \right] \right\}^{1/b}, \quad y \in (0, 1).$$

Hence, the median can be approximated as  $Q_2 \approx (0.527/a)^{1/b}$  among others.

### 2.3. Some properties

Some basic properties of the TIIT-G class are now discussed. Firstly, let us investigate the equivalences of the main functions when  $G(x)$  tends to its limit values. When  $G(x)$  tends to 0, we have

$$F(x) \sim \frac{\pi}{2}G(x), \quad f(x) \sim \frac{\pi}{2}g(x), \quad h(x) \sim \frac{\pi}{2}g(x)$$

and, when  $G(x)$  tends to 1, we get

$$F(x) \sim 1 - \frac{\pi}{4}(1 - G(x)), \quad f(x) \sim \frac{\pi}{4}g(x), \quad h(x) \sim \frac{g(x)}{1 - G(x)}.$$

These results reveal that the asymptotic behaviors of  $F(x)$ ,  $f(x)$  and  $h(x)$  are closed to the corresponding function of the baseline distribution.

The rest of this part is devoted to the expressions of the moments of a random variable  $X$  having the cdf of the TIIT-G class, as well as those of some related measures and functions. Following the spirit of the modern classes of distributions, one can investigate a series expansion for  $f(x)$ . The motivations and details of this approach can be found in [7]. Then, for any  $x \in (-\pi/2, \pi/2)$ , the Taylor series expansion of the tangent function gives  $\tan(x) = \sum_{k=1}^{+\infty} a_k x^{2k-1}$ , where  $a_k = (-1)^{k-1} 2^{2k} (2^{2k} - 1) B_{2k} / [(2k)!]$  and  $B_{2k}$  denotes the Bernoulli number with index  $2k$ . Therefore, one can expand  $F(x)$  as

$$F(x) = 1 - \sum_{k=1}^{+\infty} a_k^* \bar{G}(x)^{2k-1}, \quad x \in \mathbb{R},$$

where  $a_k^* = (\pi/4)^{2k-1} a_k$  and  $\bar{G}(x) = 1 - G(x)$ , being the sf of the baseline distribution. By differentiation with respect to  $x$  (almost surely), we obtain

$$f(x) = \sum_{k=1}^{+\infty} a_k^* \left[ (2k - 1)g(x)\bar{G}(x)^{2(k-1)} \right], \quad x \in \mathbb{R}. \tag{2.5}$$

From this representation, one can express or approximate several key measures or functions of the TIIT-G class in a tractable manner. Indeed, from  $X$  and the related expectation denoted by  $E$ , most of them can be written as (or depend on quantities that can be written as)  $E[\Pi(X)]$  where  $\Pi(x)$  refers to a certain function. Some of them are listed in Table 1.

**Table 1.** Examples of standard mathematical objects of the TIIT-G class that can be expressed as  $E[\Pi(X)]$ 

$\Pi(x)$	Name of $E[\Pi(X)]$	Notation
$I\{x \leq t\}$	cdf ( $t$ )	$F(t)$
$x$	mean	$m$
$x^2 - m^2$	variance	$\sigma^2$
$x^r$	$r^{\text{th}}$ (ordinary) moment	$m_r$
$\left(\frac{x-m}{\sigma}\right)^r$	$r^{\text{th}}$ general coefficient	$C_r$
$x^r I\{x \leq t\}$	$r^{\text{th}}$ incomplete moment ( $t$ )	$m_r(t)$
$e^{tx}$	moment generating function ( $t$ )	$M(t)$

Thus, thanks to Equation (2.5), we have

$$E[\Pi(X)] = \int_{-\infty}^{+\infty} \Pi(x)f(x)dx = \sum_{k=1}^{+\infty} a_k^* \mathcal{I}_k^{[\Pi]} \approx \sum_{k=1}^K a_k^* \mathcal{I}_k^{[\Pi]},$$

where  $\mathcal{I}_k^{[\Pi]} = \int_{-\infty}^{+\infty} \Pi(x) [(2k-1)g(x)\bar{G}(x)^{2(k-1)}] dx$  and  $K$  denotes a great integer. The advantage of this expression is that  $\mathcal{I}_k^{[\Pi]}$  is often manageable for a myriad of baseline distributions, making the computational treatment of  $E[\Pi(X)]$  straightforward.

For instance, if  $X$  follows the TIITW distribution as defined in Equation (2.4), the  $r^{\text{th}}$  ordinary moment of  $X$  is given as

$$m_r = E(X^r) = \sum_{k=1}^{+\infty} a_k^* \mathcal{I}_k^{[r]} \approx \sum_{k=1}^K a_k^* \mathcal{I}_k^{[r]},$$

where, after some developments,

$$\mathcal{I}_k^{[r]} = \int_0^{+\infty} x^r [(2k-1)abx^{b-1}e^{-a(2k-1)x^b}] dx = \frac{1}{a^{r/b}(2k-1)^{r/b}} \Gamma\left(\frac{r}{b} + 1\right),$$

$\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$  being the standard gamma function. Other measures and functions can be expressed in a similar manner, exactly as those developed in [7].

### 3. Applications

In this section, we describe our statistical methodology and applied it to twelve different data sets, for models comparison purpose.

### 3.1. Statistical methodology

The following statistical methodology is adopted. Firstly, twelve data sets with diverse characteristics are considered. Then, for each data set, we proceed as follows.

- The data are presented with adequate reference(s).
- A descriptive summary of the data is provided, indicating the mean, median, standard deviation (standard dev.), variance, skewness, kurtosis, minimum and maximum.
- We fit the data with the following five two-parameter models.
  1. Weibull (W) model with parameters  $a$  and  $b$  as defined in Equation (2.3), i.e. whose cdf and pdf are given by

$$G(x; a, b) = 1 - e^{-ax^b}, \quad g(x; a, b) = abx^{b-1}e^{-ax^b}, \quad x > 0,$$

and both equal to 0 for  $x \leq 0$  respectively. The Weibull model serves as the baseline of the next trigonometric models.

2. Sin Weibull (SW) model with parameters  $a$  and  $b$  whose cdf and pdf are given by

$$F_{SW}(x; a, b) = \sin \left[ \frac{\pi}{2} G(x; a, b) \right]$$

and

$$f_{SW}(x; a, b) = \frac{\pi}{2} g(x; a, b) \cos \left[ \frac{\pi}{2} G(x; a, b) \right]$$

respectively.

3. Cos Weibull (CW) model with parameters  $a$  and  $b$  whose cdf and pdf are given by

$$F_{CW}(x; a, b) = 1 - \cos \left[ \frac{\pi}{2} G(x; a, b) \right]$$

and

$$f_{CW}(x; a, b) = \frac{\pi}{2} g(x; a, b) \sin \left[ \frac{\pi}{2} G(x; a, b) \right]$$

respectively. This model has been the object of a special **R** package in [23].

4. Tan Weibull (TW) model with parameters  $a$  and  $b$  whose cdf and pdf are given by

$$F_{TW}(x; a, b) = \tan \left[ \frac{\pi}{4} G(x; a, b) \right]$$

and

$$f_{TW}(x; a, b) = \frac{\pi}{4} g(x; a, b) \frac{1}{\left\{ \cos \left[ \frac{\pi}{4} G(x; a, b) \right] \right\}^2}$$

respectively.

5. Type II Tan Weibull (TIITW) model with parameters  $a$  and  $b$  as described in Equation (2.4), i.e., whose cdf and pdf are given by

$$F_{TIITW}(x; a, b) = 1 - \tan \left[ \frac{\pi}{4} (1 - G(x; a, b)) \right]$$

and

$$f_{TIITW}(x; a, b) = \frac{\pi}{4} g(x; a, b) \frac{1}{\left\{ \cos \left[ \frac{\pi}{4} (1 - G(x; a, b)) \right] \right\}^2}$$

respectively.

Here, it is supposed that  $a$  and  $b$  are unknown, both of which needed to be estimated from the data.

- We employ the maximum likelihood method outlined by [6], providing the so-called maximum likelihood estimates (MLEs) of  $a$  and  $b$ , denoted by  $\hat{a}$  and  $\hat{b}$  respectively. Let us mention that the estimated pdfs and cdfs of the models are obtained by substituting  $a$  and  $b$  by  $\hat{a}$  and  $\hat{b}$ , which are respectively in their own definitions.
- We compare the models by using the following standard goodness-of-fit statistics:  $-Log$  (corresponding to the minus estimated log-likelihood function), Anderson Darling ( $A^*$ ), Cramer-von Mises ( $W^*$ ), Akaike information criterion (AIC), corrected Akaike information criterion (CAIC), Bayesian information criterion (BIC) and Hannan-Quinn information criterion (HQIC), along with the Kolmogorov-Smirnov (K-S) statistic of the K-S test coupled with its p-value (PV). All these criteria are useful to show the relevance of these models to fit the considered data set; the best model being the one with the lowest value of each of these statistics, excluding the p-value which must be the greatest.
- Plots of the estimated pdfs over the histogram of the data as well as plots of the estimated cdfs over the empirical cdf of the data are shown to support the conclusions.

In practice, we operate with the statistical software R (developed by [17]) and use the `goodness.fit` function from the package `AdequacyModel` elaborated by [14]. The outputs provide all the numerical values of the mentioned MLEs and statistics. The main lines of our R code are given in Appendix.

*Note:* In this study, in accordance with the original spirit of model comparison, we considered only unknown and estimated pdfs and cdfs. For the sake of brevity and to avoid some redundancy in the conclusions, statistical objects related to hazard rates, such as the Total of Time (TTT) plots and estimated hrfs, have been omitted. However, they are of some interest and remain available from the author upon request.

### 3.2. Application

The application of our statistical methodology on twelve data sets is described as follows.

**Data set 1:** This data set displays the waiting times (in minutes) before service of 100 bank customers. It was investigated and interpreted by [10] in order to fit the Lindley distribution. It is given by: 0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5.

Table 2 provides a summary of descriptive statistics of this data set.

**Table 2.** Descriptive statistics of Data set 1

Mean	Median	Standard dev.	Variance	Skewness	Kurtosis	Minimum	Maximum
9.877	8.100	7.236996	52.37411	1.472765	5.540292	0.800	38.500

In Table 2, we see that the data are right-skewed and leptokurtic with a high variance. Table 3 contains the MLEs of the parameters of the W, SW, CW, TW and TIITW models, that is,  $\hat{a}$  and  $\hat{b}$  for each model as well as the  $-Log$ ,  $A^*$ ,  $W^*$ , K-S, PV, AIC, AICC, BIC and HQIC statistics of these models.

**Table 3.** MLEs and goodness-of-fit statistics for Data set 1

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	0.0304	1.4586	318.73	0.40	0.06	0.06	0.89	641.46	641.59	646.67	643.57
SW	0.0210	1.3796	319.06	0.45	0.07	0.06	0.87	642.11	642.24	647.32	644.22
CW	0.1187	1.0620	317.06	0.13	0.02	0.04	1.00	638.11	638.24	643.32	640.22
TW	0.0456	1.3767	319.78	0.57	0.09	0.07	0.76	643.56	643.68	648.77	645.67
TIITW	0.0165	1.6190	318.32	0.33	0.05	0.06	0.91	640.65	640.77	645.86	642.76

From Table 3, since they present the lowest  $-Log$ ,  $A^*$ ,  $W^*$ , K-S, AIC, AICC, BIC and HQIC values along with the greatest PV values, the CW and TIITW models are more suitable fitted models than the competitors, the CW model being the better of the two. Figure 1 displays the estimated pdfs over the histogram of the data.

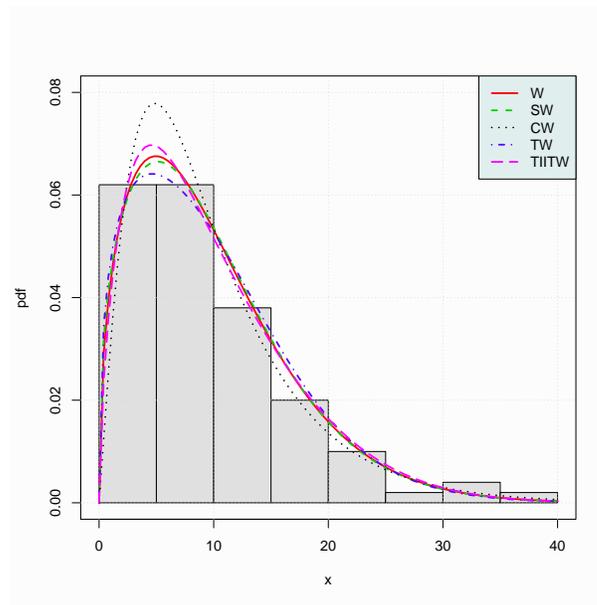


Figure 1. Estimated pdfs for Data set 1

Figure 2 displays the estimated cdfs over the empirical cdf of the data.

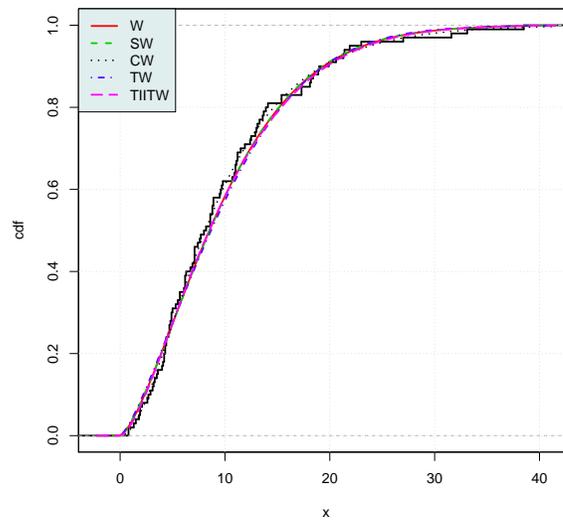


Figure 2. Estimated cdfs for Data set 1

Note: In the next applications, for the sake of conciseness, we will provide less details on the used tools and the reasons why the considered models are the best,

implicitly invoking the same arguments as those developed for the analysis of Data set 1.

**Data set 2:** This data set includes the waiting times between 65 consecutive Kiama Blowhole eruptions. The values had been observed on 12 July, 1998 by Jim Irish and had been lately mentioned by [3]. It is given by: 83, 51, 87, 60, 28, 95, 8, 27, 15, 10, 18, 16, 29, 54, 91, 8, 17, 55, 10, 35, 47, 77, 36, 17, 21, 36, 18, 40, 10, 7, 34, 27, 28, 56, 8, 25, 68, 146, 89, 18, 73, 69, 9, 37, 10, 82, 29, 8, 60, 61, 61, 18, 169, 25, 8, 26, 11, 83, 11, 42, 17, 14, 9, 12.

A summary of descriptive statistics of this data set is provided in Table 4.

**Table 4.** Descriptive statistics of Data set 2

Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
39.83	28.00	33.75051	1139.097	1.546409	5.771077	7.00	169.00

We see in Table 4 that the data are right-skewed and leptokurtic with a very high variance. Table 5 contains the MLEs and goodness-of-fit statistics of the considered models.

**Table 5.** MLEs and goodness-of-fit statistics for Data set 2

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	0.0082	1.2744	296.90	1.01	0.15	0.11	0.41	597.80	598.00	602.12	599.50
SW	0.0061	1.2035	297.28	1.05	0.16	0.11	0.40	598.57	598.76	602.88	600.27
CW	0.0459	0.9291	295.17	0.85	0.12	0.11	0.43	594.35	594.55	598.67	596.05
TW	0.0127	1.2115	297.57	1.09	0.16	0.11	0.38	599.13	599.33	603.45	600.83
TIITW	0.0038	1.4206	296.13	0.91	0.13	0.11	0.43	596.26	596.46	600.58	597.96

According to Table 5, the CW and TIITW models are the best to fit this data set, the CW model being the better of the two. In Figures 3 and 4, the estimated pdfs and cdfs are displayed over the histogram and empirical cdf of the data respectively.

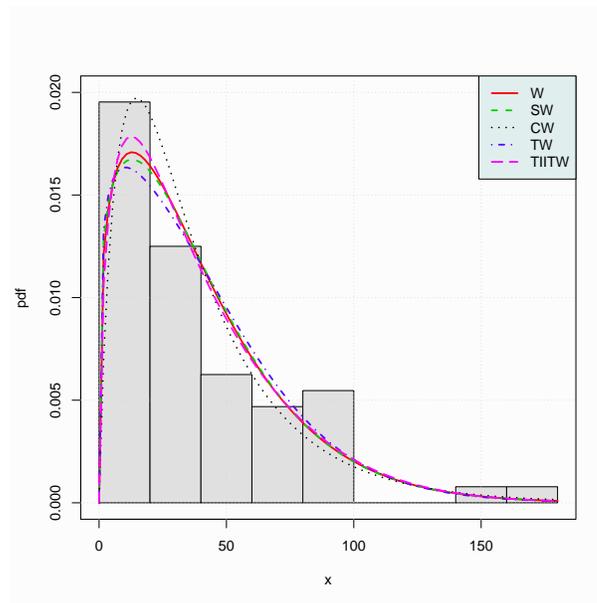


Figure 3. Estimated pdfs for Data set 2

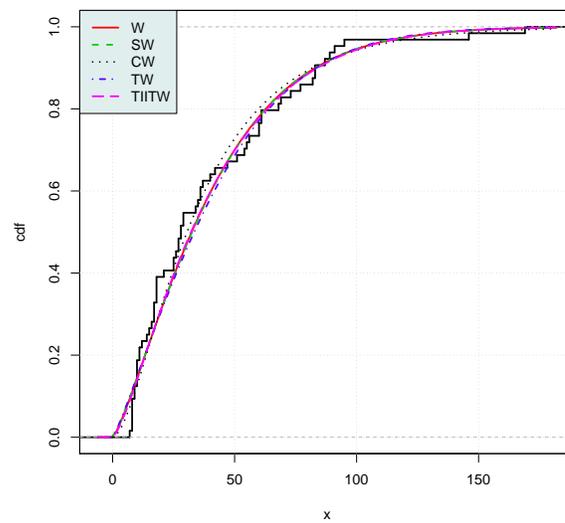


Figure 4. Estimated cdfs for Data set 2

Figures 3 and 4 illustrate the superior adequacy of the CW and TIITW models, and particularly highlights the adjustment of the CW one.

**Data set 3:** This data set consists of 202 athletes’ lean body mass (in kilograms) extracted from physical measurements and blood measurements from high performance athletes at the Australian Institute of Sports and examined by [25]. It is given by: 63.32, 58.55, 55.36, 57.18, 53.20, 53.77, 60.17, 48.33, 54.57, 53.42, 68.53, 61.85, 48.32, 66.24, 57.92, 56.52, 54.78, 56.31, 62.96, 56.68, 62.39, 63.05, 56.05, 53.65, 65.45, 64.62, 60.05, 56.48, 41.54, 52.78, 52.72, 61.29, 59.59, 61.70, 62.46, 53.14, 47.09, 53.44, 48.78, 56.05, 56.45, 53.11, 54.41, 55.97, 51.62, 58.27, 57.28, 57.30, 54.18, 42.96, 54.46, 57.20, 54.38, 57.58, 61.46, 53.46, 54.11, 55.35, 55.39, 52.23, 59.33, 61.63, 63.39, 60.22, 55.73, 48.57, 51.99, 51.17, 57.54, 68.86, 63.04, 63.03, 66.85, 59.89, 72.98, 45.23, 55.06, 46.96, 53.54, 47.57, 54.63, 46.31, 49.13, 53.71, 53.11, 46.12, 53.41, 51.48, 53.20, 56.58, 56.01, 46.52, 51.75, 42.15, 48.76, 41.93, 42.95, 38.30, 34.36, 39.03, 61.00, 69.00, 74.00, 80.00, 78.00, 71.00, 71.00, 78.00, 77.00, 81.00, 66.00, 77.00, 91.00, 78.00, 75.00, 78.00, 87.00, 78.00, 79.00, 79.00, 48.00, 82.00, 82.00, 82.00, 83.00, 88.00, 83.00, 78.00, 85.00, 73.00, 82.00, 79.00, 97.00, 90.00, 90.00, 74.00, 82.00, 72.00, 76.00, 70.00, 57.00, 67.00, 67.00, 70.00, 88.00, 83.00, 74.00, 62.00, 67.00, 70.00, 64.00, 58.00, 57.00, 73.00, 54.00, 67.00, 66.00, 75.00, 78.00, 102.00, 74.00, 78.00, 106.00, 68.00, 77.00, 69.00, 66.00, 62.00, 65.00, 62.00, 66.00, 67.00, 65.00, 63.00, 59.00, 86.00, 87.00, 89.00, 80.00, 68.00, 69.00, 77.00, 68.00, 77.00, 71.00, 72.00, 74.00, 68.00, 85.00, 75.00, 78.00, 86.00, 69.00, 79.00, 80.00, 82.00, 68.00, 82.00, 72.00, 68.00, 63.00, 72.00.

Table 6 proposes a summary of descriptive statistics of this data set.

**Table 6.** Descriptive statistics of Data set 3

Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
64.87	63.03	13.0702	170.8301	0.3585088	2.75988	34.36	106.00

In Table 6, we notice that the data are approximately symmetric and rather mesokurtic with a high variance. Table 7 indicates the MLEs and goodness-of-fit statistics of the considered models.

**Table 7.** MLEs and goodness-of-fit statistics for Data set 3

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	$1.9184 \times 10^{-10}$	5.2618	811.61	2.09	0.37	0.08	0.17	1627.21	1627.27	1633.83	1629.89
SW	$3.7898 \times 10^{-10}$	4.9702	812.60	2.25	0.40	0.08	0.16	1629.19	1629.25	1635.81	1631.87
CW	$9.6118 \times 10^{-8}$	3.8973	804.95	1.26	0.23	0.08	0.19	1613.90	1613.96	1620.52	1616.58
TW	$7.9479 \times 10^{-10}$	4.9753	814.33	2.50	0.44	0.09	0.10	1632.66	1632.72	1639.28	1635.34
TIITW	$1.2190 \times 10^{-11}$	5.8560	810.00	1.79	0.32	0.08	0.18	1623.99	1624.05	1630.61	1626.67

The analysis of Table 7 ensures that the CW and TIITW models are the best, the CW model being the better of the two. In Figures 5 and 6, the estimated pdfs and cdfs are displayed over the histogram and empirical cdf of the data respectively.

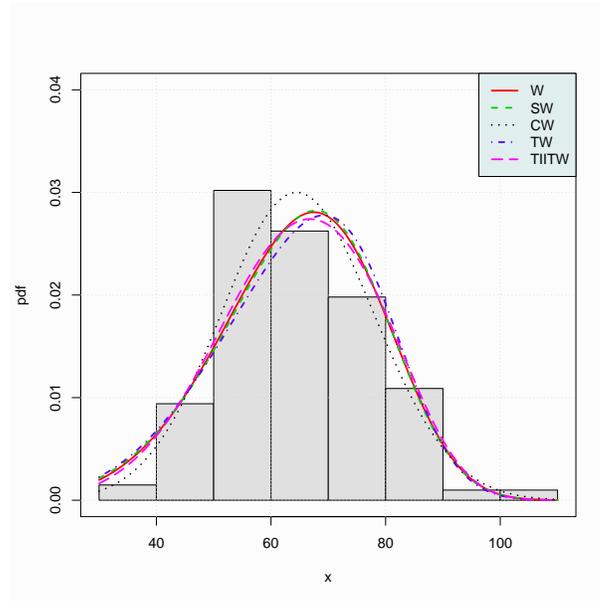


Figure 5. Estimated pdfs for Data set 3

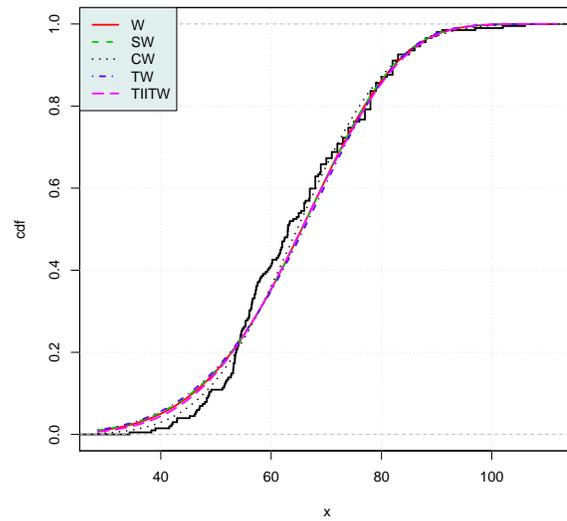


Figure 6. Estimated cdfs for Data set 3

Figures 5 and 6 visually confirm the better adjustment of the CW model.

**Data set 4:** This data set contains the number of successive failures for the air-conditioning system of each member in a fleet of 13 Boeing 720 jet airplanes, given by [16]. It is given by: 1, 1, 3, 3, 3, 4, 5, 5, 5, 5, 5, 7, 7, 7, 9, 9, 10, 11, 11, 11, 11, 12, 12, 12, 12, 13, 14, 14, 14, 14, 14, 14, 14, 14, 14, 15, 15, 16, 16, 16, 18, 18, 18, 18, 18, 18, 20, 20, 21, 22, 22, 22, 23, 23, 23, 24, 25, 26, 26, 27, 27, 29, 29, 29, 30, 31, 31, 32, 33, 33, 34, 34, 34, 35, 35, 36, 36, 37, 39, 39, 41, 42, 43, 44, 44, 44, 46, 46, 48, 49, 50, 50, 51, 52, 54, 54, 55, 56, 57, 57, 57, 58, 59, 59, 60, 61, 61, 62, 62, 63, 65, 66, 67, 70, 71, 71, 72, 74, 76, 79, 79, 80, 82, 84, 87, 88, 90, 90, 95, 97, 97, 98, 100, 100, 101, 102, 102, 104, 104, 106, 111, 118, 118, 120, 120, 130, 130, 130, 134, 139, 141, 152, 153, 156, 163, 181, 182, 184, 186, 188, 191, 194, 201, 206, 208, 208, 209, 210, 216, 220, 230, 230, 239, 246, 254, 261, 270, 283, 310, 320, 326, 359, 386, 413, 438, 487, 493, 502, 603.

Descriptive statistics of this data set can be seen in Table 8.

**Table 8.** Descriptive statistics of Data set 4

Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
92.07	54.00	107.9163	11645.93	2.139207	8.023109	1.00	603.00

Table 8 indicates that the data are right-skewed and leptokurtic with a really high variance. Table 9 contains the MLEs and goodness-of-fit statistics of the considered models.

**Table 9.** MLEs and goodness-of-fit statistics for Data set 4

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	0.0170	0.9107	1036.75	1.00	0.16	0.06	0.57	2077.50	2077.57	2083.98	2080.12
SW	0.0121	0.8608	1037.58	1.12	0.18	0.06	0.53	2079.16	2079.23	2085.64	2081.79
<b>CW</b>	<b>0.0783</b>	<b>0.6614</b>	<b>1033.34</b>	<b>0.36</b>	<b>0.05</b>	<b>0.04</b>	<b>0.85</b>	<b>2070.69</b>	<b>2070.75</b>	<b>2077.16</b>	<b>2073.31</b>
TW	0.0260	0.8605	1038.81	1.36	0.22	0.07	0.34	2081.63	2081.69	2088.10	2084.25
TIITW	0.0085	1.0139	1035.37	0.78	0.12	0.06	0.61	2074.74	2074.80	2081.21	2077.36

Table 9 indicates that the CW and TIITW models are the most relevant fitted models with a preference for the CW model. In Figure 7, the estimated pdfs are displayed over the histogram of the data.

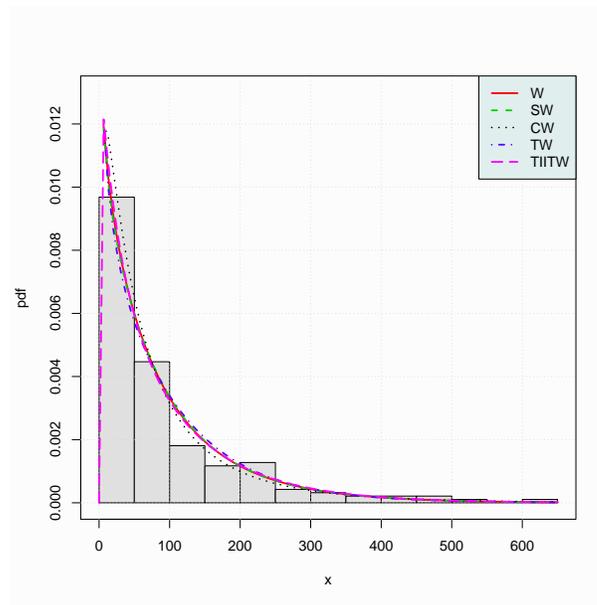


Figure 7. Estimated pdfs for Data set 4

In Figure 8, the estimated cdfs are plotted over the empirical cdf of the data.

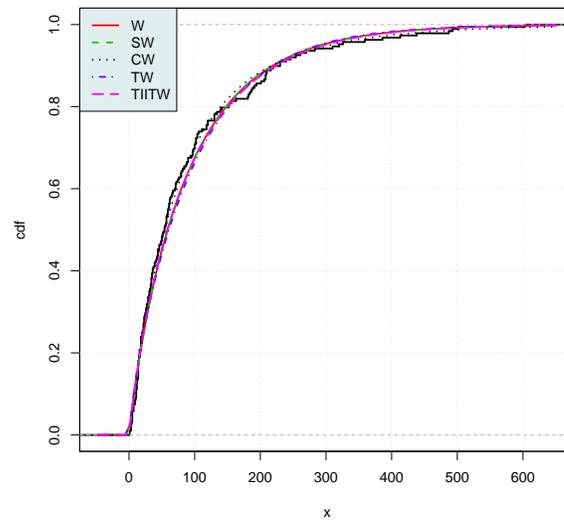


Figure 8. Estimated cdfs for Data set 4.

**Data set 5:** This data set exhibits the vinyl chloride data collected from clean

upgrading, monitoring wells in mg/L which was employed by [4]. It is given by: 5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8.0, 0.8, 0.4, 0.6, 0.9, 0.4, 2.0, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1.0, 0.2, 0.1, 0.1, 1.8, 0.9, 2.0, 4.0, 6.8, 1.2, 0.4, 0.2.

Table 10 provides a summary of descriptive statistics of this data set.

**Table 10.** Descriptive statistics of Data set 5

Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
1.879	1.150	1.952586	3.812594	1.603688	5.005408	0.100	8.000

In Table 10, we see that the data are right-skewed and leptokurtic with a low variance. Table 11 collects the MLEs and goodness-of-fit statistics of the considered models.

**Table 11.** MLEs and goodness-of-fit statistics for Data set 5

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	0.5262	1.0102	55.45	0.30	0.05	0.09	0.94	114.90	115.29	117.95	115.94
SW	0.3105	0.9542	55.61	0.32	0.05	0.09	0.92	115.22	115.61	118.28	116.26
<b>CW</b>	<b>0.9534</b>	<b>0.7281</b>	<b>54.94</b>	<b>0.20</b>	<b>0.03</b>	<b>0.08</b>	<b>0.98</b>	<b>113.88</b>	<b>114.27</b>	<b>116.93</b>	<b>114.92</b>
TW	0.6651	0.9563	55.77	0.36	0.06	0.10	0.86	115.55	115.93	118.60	116.59
TIITW	0.3922	1.1251	55.15	0.26	0.04	0.09	0.95	114.30	114.68	117.35	115.34

According to Table 11, the CW and TIITW models are the best in the fitting work, the CW model being the better of the two. In Figures 9 and 10, the estimated pdfs and cdfs are plotted over the histogram and empirical cdf of the data respectively.

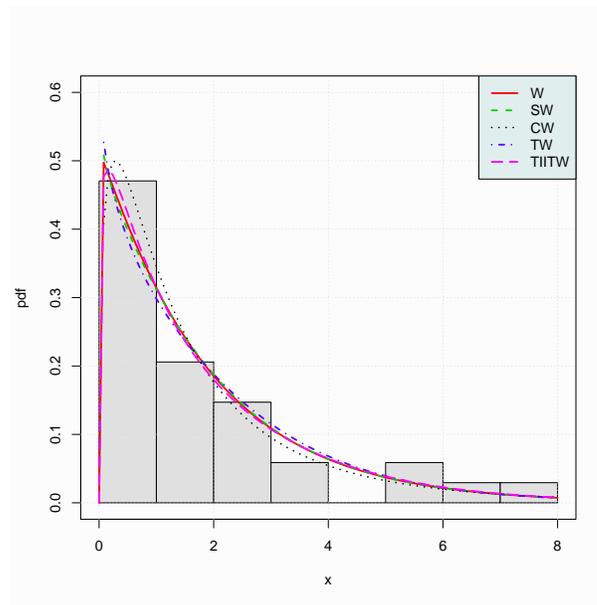


Figure 9. Estimated pdfs for Data set 5

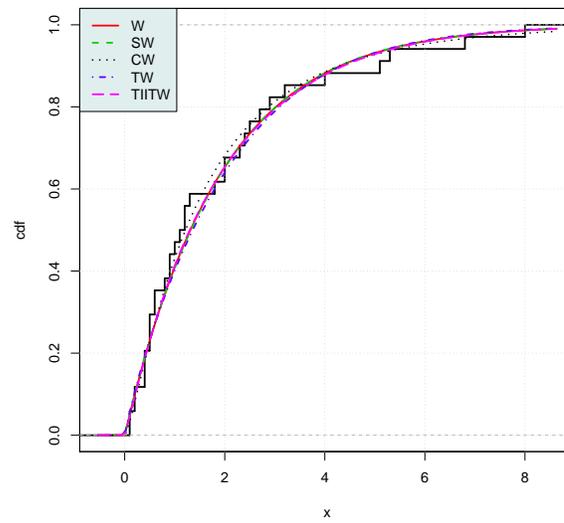


Figure 10. Estimated cdfs for Data set 5

Figures 9 and 10 visually confirms the best settings for the CW and TIITW models, notably the CW one.

**Data set 6:** This data set is based on the survival times (in years) of a group of patients given chemotherapy treatment alone. It is a subset of data presented by [5]. It is given by: 0.047, 0.115, 0.121, 0.132, 0.164, 0.197, 0.203, 0.260, 0.282, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.641, 0.644, 0.696, 0.841, 0.863, 1.099, 1.219, 1.271, 1.326, 1.447, 1.485, 1.553, 1.581, 1.589, 2.178, 2.343, 2.416, 2.444, 2.825, 2.830, 3.578, 3.658, 3.743, 3.978, 4.003, 4.033.

Table 12 shows a summary of descriptive statistics of this data set.

**Table 12.** Descriptive statistics of Data set 6

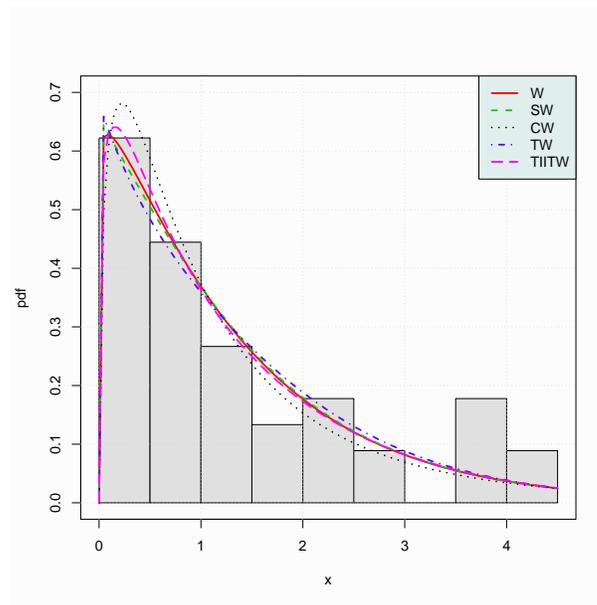
Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
1.341	0.841	1.246601	1.554014	0.9721463	2.663833	0.047	4.033

From Table 12, we see that the data are right-skewed and rather mesokurtic with a high variance. Table 13 presents the MLEs and goodness-of-fit statistics of the considered models.

**Table 13.** MLEs and goodness-of-fit statistics for Data set 6

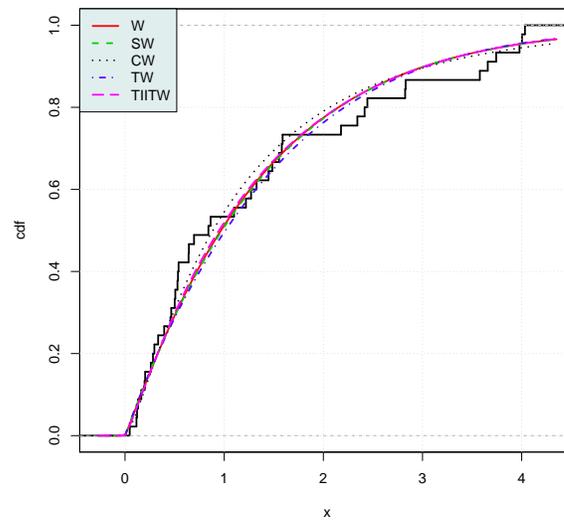
Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	0.7177	1.0532	58.12	0.54	0.08	0.11	0.61	120.25	120.53	123.86	121.59
SW	0.4144	0.9956	58.33	0.58	0.09	0.11	0.59	120.66	120.95	124.28	122.01
CW	1.2010	0.7404	58.09	0.44	0.06	0.09	0.81	120.18	120.47	123.80	121.53
TW	0.8816	1.0054	58.26	0.60	0.09	0.12	0.54	120.52	120.81	124.13	121.87
<b>TIITW</b>	<b>0.5618</b>	<b>1.1746</b>	<b>57.61</b>	<b>0.48</b>	<b>0.07</b>	<b>0.10</b>	<b>0.68</b>	<b>119.21</b>	<b>119.50</b>	<b>122.83</b>	<b>120.56</b>

According to Table 13, the CW and TIITW models are more pertinent fitted models than the competitors for this data set, the TIITW model being the better of the two. Figure 11 shows the estimated pdfs over the histogram of the data.



**Figure 11.** Estimated pdfs for Data set 6

Figure 12 presents the estimated cdfs over the empirical cdf of the data.



**Figure 12.** Estimated cdfs for Data set 6

Figures 11 and 12 illustrate the superior adequacy of the CW and TIITW models, especially the one of the TIITW model.

**Data set 7:** This data set reports the failure times of the air conditioning system of an airplane. It has been published by [13]. It is given by: 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95.

A summary of descriptive statistics of this data set is given in Table 14.

**Table 14.** Descriptive statistics of Data set 7

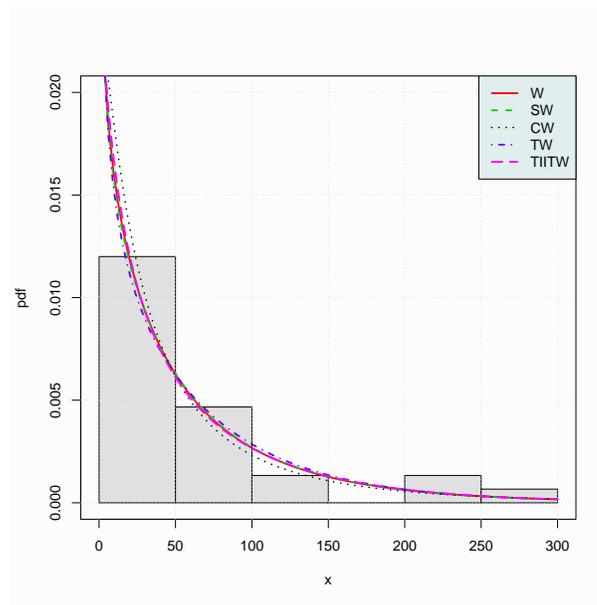
Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
59.6	22.0	71.88477	5167.421	1.693605	4.966655	1.0	261.0

From Table 14, we can say that the data are right-skewed and leptokurtic with a very high variance. Table 15 presents the MLEs and goodness-of-fit statistics of the considered models.

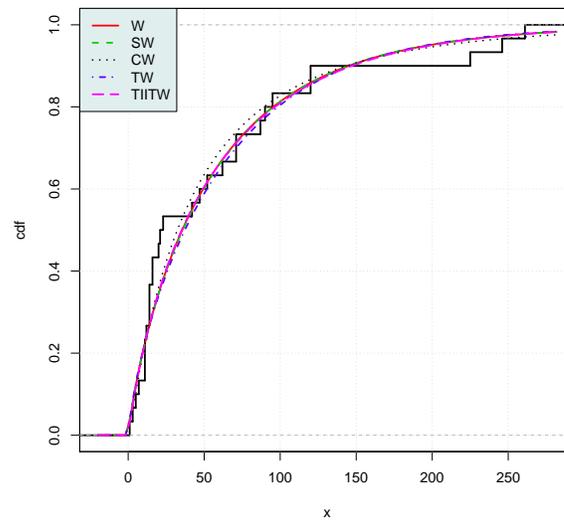
**Table 15.** MLEs and goodness-of-fit statistics for Data set 7

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	0.0329	0.8535	151.94	0.58	0.10	0.15	0.48	307.87	308.32	310.68	308.77
SW	0.0227	0.8052	152.10	0.61	0.11	0.15	0.47	308.21	308.65	311.01	309.10
CW	0.1289	0.6161	151.42	0.47	0.09	0.13	0.65	306.84	307.28	309.64	307.74
TW	0.0477	0.8092	152.24	0.64	0.11	0.16	0.41	308.49	308.93	311.29	309.38
TIITW	0.0176	0.9540	151.60	0.54	0.09	0.15	0.51	307.19	307.64	310.00	308.09

From Table 15, it is clear that the CW and TIITW models are more relevant fitted models than the others, the CW model being the better of the two. In Figures 13 and 14, the estimated pdfs and cdfs are displayed over the histogram and empirical cdf of the data respectively.



**Figure 13.** Estimated pdfs for Data set 7



**Figure 14.** Estimated cdfs for Data set 7

**Data set 8:** This data set, disclosed by [9], involves several measures of the strength of glass of the aircraft window. It is given by: 18.83, 20.8, 21.657, 23.03, 23.23, 24.05, 24.321, 25.5, 25.52, 25.8, 26.69, 26.77, 26.78, 27.05, 27.67, 29.9, 31.11,

33.2, 33.73, 33.76, 33.89, 34.76, 35.75, 35.91, 36.98, 37.08, 37.09, 39.58, 44.045, 45.29, 45.381.

Table 16 provides a summary of descriptive statistics of this data set.

**Table 16.** Descriptive statistics of Data set 8

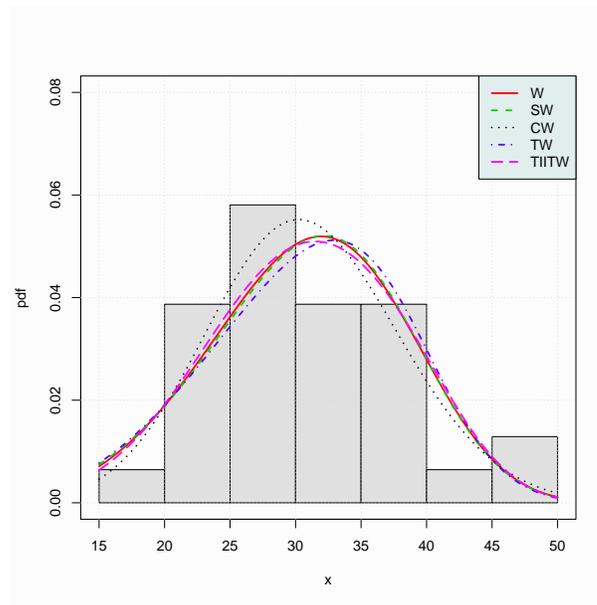
Mean	Median	Standard dev.	Variance	Skewness	Kurtosis	Minimum	Maximum
30.81	29.90	7.253381	52.61154	0.4053965	2.286637	18.830	45.381

We see in Table 16 that the data are approximately symmetric and platykurtic with a high variance. Table 17 contains the MLEs and goodness-of-fit statistics of the considered models.

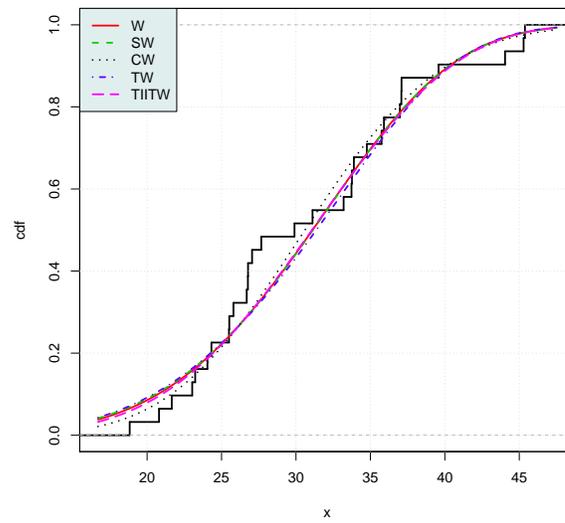
**Table 17.** MLEs and goodness-of-fit statistics for Data set 8

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	$8.3379 \times 10^{-8}$	4.6349	105.49	0.64	0.10	0.15	0.42	214.98	215.41	217.85	215.91
SW	$1.2024 \times 10^{-7}$	4.3712	105.69	0.67	0.10	0.15	0.42	215.38	215.81	218.25	216.31
CW	$1.0279 \times 10^{-5}$	3.3890	104.60	0.50	0.09	0.14	0.51	213.19	213.62	216.06	214.13
TW	$2.4182 \times 10^{-7}$	4.3885	105.89	0.71	0.11	0.16	0.39	215.78	216.21	218.65	216.72
TIITW	$9.9612 \times 10^{-9}$	5.1753	105.11	0.60	0.09	0.15	0.44	214.21	214.64	217.08	215.15

According to Table 17, the CW and TIITW models are more pertinent than the other models to fit these data, the CW model being the better of the two. Figure 15 shows the estimated pdfs over the histogram of the data, whereas Figure 16 shows the estimated cdfs over empirical cdf of the data respectively.



**Figure 15.** Estimated pdfs for Data set 8



**Figure 16.** Estimated cdfs for Data set 8

Figures 15 and 16 illustrate the superior adequacy of the CW model, which has better captured the "top" of the data.

**Data set 9:** This data set disclosed 50 observations with hole diameter of 12 mm and sheet thickness 3.15 mm, documented by [8]. Hole diameter readings are taken into account with respect to one hole selected and fixed as per a predetermined orientation. It is given by: 0.04, 0.02, 0.06, 0.12, 0.14, 0.08, 0.22, 0.12, 0.08, 0.26, 0.24, 0.04, 0.14, 0.16, 0.08, 0.26, 0.32, 0.28, 0.14, 0.16, 0.24, 0.22, 0.12, 0.18, 0.24, 0.32, 0.16, 0.14, 0.08, 0.16, 0.24, 0.16, 0.32, 0.18, 0.24, 0.22, 0.16, 0.12, 0.24, 0.06, 0.02, 0.18, 0.22, 0.14, 0.06, 0.04, 0.14, 0.26, 0.18, 0.16.

A summary of descriptive statistics of this data set is given in Table 18.

**Table 18.** Descriptive statistics of Data set 9

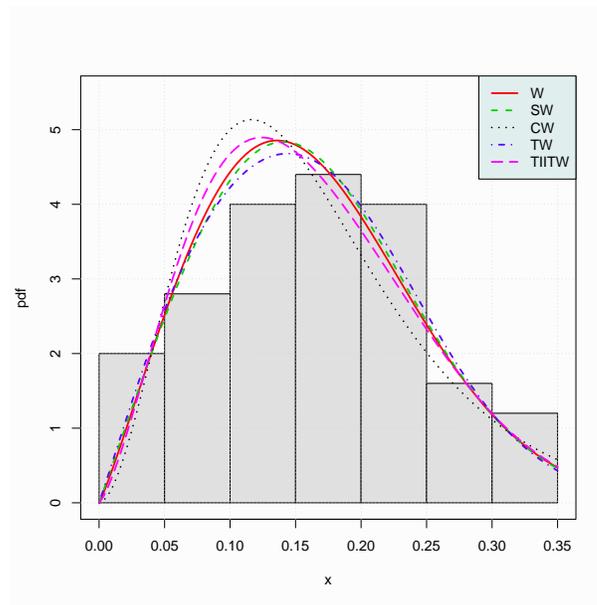
Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
0.1632	0.1600	0.08105025	0.006569143	0.07233688	2.216649	0.02	0.32

Table 18 suggests that the data are approximately symmetric and platykurtic with a variance close to 0. Table 19 collects the MLEs and goodness-of-fit statistics of the considered models.

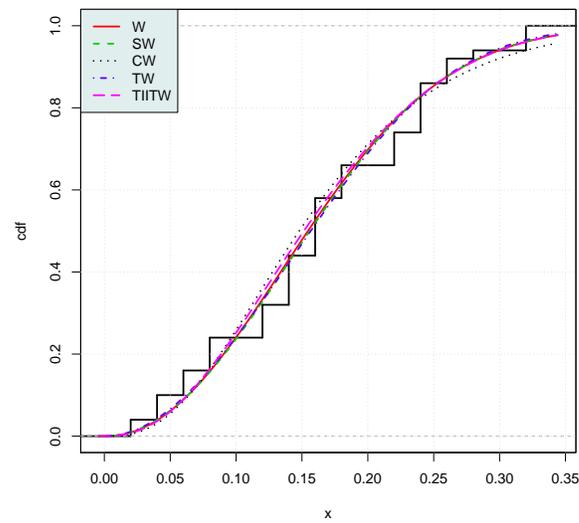
**Table 19.** MLEs and goodness-of-fit statistics for Data set 9

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	36.2758	2.1196	-55.89	0.64	0.11	0.11	0.58	-107.78	-107.53	-103.96	-106.33
SW	17.3726	2.0220	-56.08	0.61	0.10	0.11	0.62	-108.16	-107.90	-104.33	-106.70
CW	16.3020	1.4112	-53.48	1.05	0.17	0.14	0.27	-102.96	-102.70	-99.13	-101.50
<b>TW</b>	<b>38.1814</b>	<b>2.0264</b>	<b>-56.44</b>	<b>0.56</b>	<b>0.09</b>	<b>0.10</b>	<b>0.65</b>	<b>-108.88</b>	<b>-108.63</b>	<b>-105.06</b>	<b>-107.43</b>
TIITW	42.1280	2.3210	-55.54	0.72	0.12	0.13	0.41	-107.08	-106.82	-103.25	-105.62

Table 19 indicates that the SW and TW models are the most adequate fitted models than the competitors concerning this data set, the TW model being the better of the two. In Figures 17 and 18, the estimated pdfs and cdfs are displayed over the histogram and empirical cdf of the data respectively.



**Figure 17.** Estimated pdfs for Data set 9



**Figure 18.** Estimated cdfs for Data set 9

Figures 17 and 18 highlight the better adjustment of the SW and TW models.

**Data set 10:** This data set implicates the time to failure (in  $10^3h$ ) of tur-

bocharger of one type of engine. It is presented by [26]. It is given by: 1.6, 2.0, 2.6, 3.0, 3.5, 3.9, 4.5, 4.6, 4.8, 5.0, 5.1, 5.3, 5.4, 5.6, 5.8, 6.0, 6.0, 6.1, 6.3, 6.5, 6.5, 6.7, 7.0, 7.1, 7.3, 7.3, 7.3, 7.7, 7.7, 7.8, 7.9, 8.0, 8.1, 8.3, 8.4, 8.4, 8.5, 8.7, 8.8, 9.0.

Table 20 provides a summary of descriptive statistics of this data set.

**Table 20.** Descriptive statistics of Data set 10

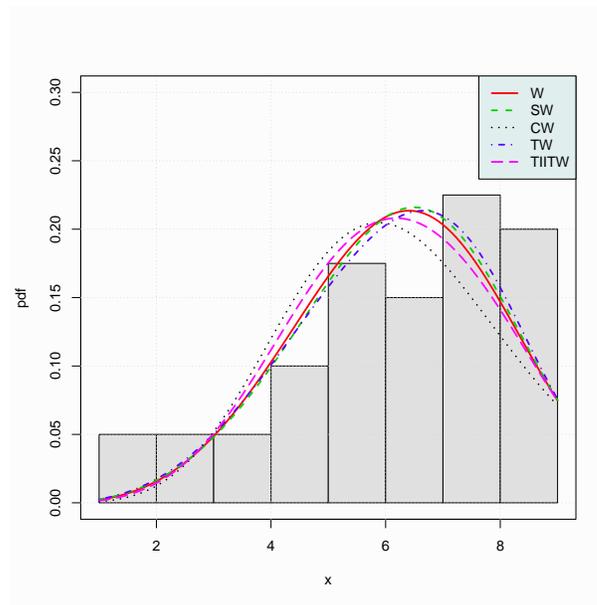
Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
6.253	6.500	1.95553	3.824096	-0.6625574	2.641014	1.60	9.00

We see in Table 20 that the data are left-skewed and approximately mesokurtic with a low variance. Table 21 contains the MLEs and goodness-of-fit statistics of the considered models.

**Table 21.** MLEs and goodness-of-fit statistics for Data set 10

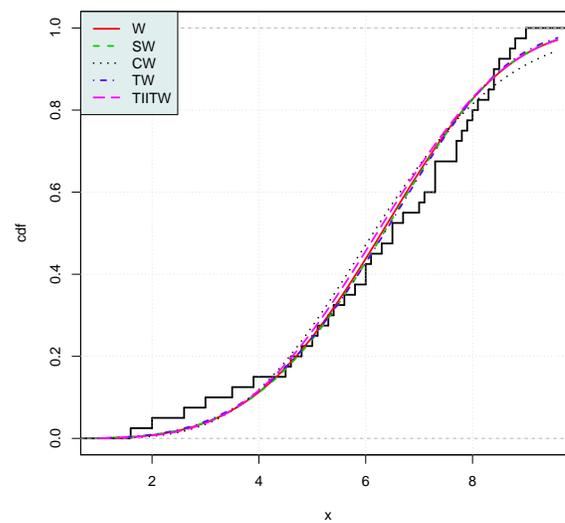
Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	$5.5573 \times 10^{-4}$	3.8731	82.48	0.57	0.08	0.11	0.74	168.95	169.28	172.33	170.17
SW	$4.3338 \times 10^{-4}$	3.7060	82.20	0.53	0.07	0.10	0.78	168.39	168.71	171.77	169.61
CW	$1.1176 \times 10^{-2}$	2.5005	85.34	1.00	0.14	0.12	0.66	174.69	175.01	178.06	175.91
TW	$8.9843 \times 10^{-4}$	3.7339	81.73	0.47	0.06	0.10	0.81	167.46	167.78	170.83	168.68
TIITW	$2.3331 \times 10^{-4}$	4.2294	82.96	0.66	0.09	0.12	0.65	169.92	170.24	173.30	171.14

From Table 21, we see that the SW and TW models are the best fitted models for this data set, the TW model being the better of the two. In Figure 19, the estimated pdfs are plotted over the histogram of the data.



**Figure 19.** Estimated pdfs for Data set 10

To complete, Figure 20 displays the estimated cdfs over the empirical cdf of the data.



**Figure 20.** Estimated cdfs for Data set 10

Figures 19 and 20 spotlight the greater adequacy of the TW models.

## Covid-19 relative data sets

The COVID-19 disease is a mild to severe respiratory illness that is caused by a coronavirus (severe acute respiratory syndrome coronavirus 2 of the genus Beta-coronavirus). Spreading all around the world at a rapid pace in the early of the year 2020, contaminating millions of people and making thousands of victims, it had forced governments to adopt exceptional measures like quarantine in order to limit the pandemic and protect their citizens. An overall comprehension of the COVID-19 disease is a challenge for all scientists, for the sake of the future generations. The analysis of COVID-19 data aims to predict and be better prepared in case of future pandemic.

In this part, we consider COVID-19 data from France from March 2020 to 1 June, 2020. The day of 1 June, 2020 is symbolic for it marks the term of the first stage of the release of lockdown in France. At this point, the pandemic is considered on the decline. Schools and restaurants are reopening, but meetings of more than ten people are still forbidden.

**Data set 11:** This data set is based on the daily amount of new admission in hospital concerning the COVID-19 in France from 19 March, 2020 to 1 June, 2020. The data are collected by [18] and are available on:

<https://www.data.gouv.fr/en/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/>

Explicitly, this data set is given by: 438, 242, 298, 309, 448, 571, 607, 612, 608, 695, 543, 694, 767, 771, 728, 640, 502, 390, 478, 518, 482, 369, 431, 255, 220, 227, 275, 284, 270, 242, 206, 137, 208, 190, 183, 178, 155, 124, 79, 125, 153, 110, 121, 73, 64, 80, 84, 111, 69, 99, 89, 38, 38, 82, 92, 69, 52, 64, 46, 24, 38, 69, 43, 28, 36, 30, 24, 45, 37, 32, 36, 29, 29, 18, 9.

Table 22 provides a summary of descriptive statistics of this data set.

**Table 22.** Descriptive statistics of Data set 11

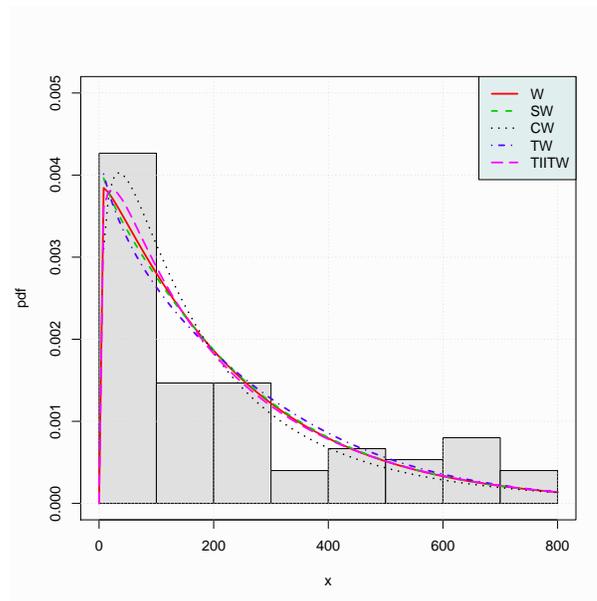
Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
234.1	137.0	222.9707	49715.93	0.9860842	2.699223	9.0	771.0

In Table 22, we see that the data are right-skewed and rather mesokurtic with a really high variance. Table 23 displays the MLEs and goodness-of-fit statistics of the considered models.

**Table 23.** MLEs and goodness-of-fit statistics for Data set 11

Model	$\hat{a}$	$\hat{b}$	$-\text{Log}$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	$3.6581 \times 10^{-3}$	1.0263	484.15	1.22	0.18	0.10	0.47	972.30	972.47	976.94	974.15
SW	$2.8345 \times 10^{-3}$	0.9691	484.59	1.27	0.19	0.10	0.44	973.18	973.35	977.82	975.03
CW	$2.877 \times 10^{-2}$	0.7256	483.58	1.03	0.15	0.09	0.57	971.16	971.33	975.79	973.01
TW	$5.6536 \times 10^{-3}$	0.9809	484.47	1.31	0.20	0.11	0.37	972.94	973.10	977.57	974.79
TIITW	$1.5358 \times 10^{-3}$	1.1472	483.05	1.09	0.16	0.09	0.54	970.10	970.27	974.74	971.95

From Table 23, it is clear that the CW and TIITW models are the most relevant, the best model being the TIITW model. In Figures 21 and 22, the estimated pdfs and cdfs are plotted over the histogram and empirical cdf of the data respectively.

**Figure 21.** Estimated pdfs for Data set 11

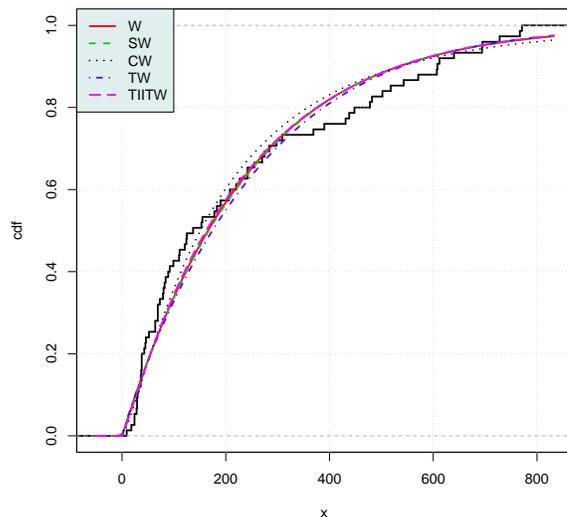


Figure 22. Estimated cdfs for Data set 11

Figures 21 and 22 visually confirm the good adjustment of the TIITW model.

**Data set 12:** This data set displays the amount of people hospitalize on account of the COVID-19 each day in France from 18 March, 2020 to 1 June, 2020. The data are collected by [18] and available on:

<https://www.data.gouv.fr/en/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/>

In expanded form, the data set is given by: 5905, 7961, 10272, 11599, 13675, 17087, 20181, 23992, 27601, 31237, 34975, 38424, 41682, 45115, 48849, 51999, 54328, 55716, 57175, 58813, 59430, 60121, 60897, 61911, 62009, 63009, 63572, 63942, 62946, 62045, 61826, 60741, 60686, 60625, 59673, 58963, 57953, 56844, 55976, 55962, 55640, 54951, 53392, 52311, 51533, 51419, 51395, 50859, 49317, 47737, 46184, 45220, 45001, 44906, 44343, 42971, 41934, 40723, 39522, 38666, 38527, 37833, 36740, 35693, 34985, 34586, 34175, 34190, 33414, 32347, 31186, 30245, 29223, 28594, 28477, 28409.

Descriptive statistics of this data set are described in Table 24.

Table 24. Descriptive statistics of Data set 12

Mean	Median	Standard deviation	Variance	Skewness	Kurtosis	Minimum	Maximum
43978	45168	15136.18	229103947	-0.6479744	2.650809	5905	63942

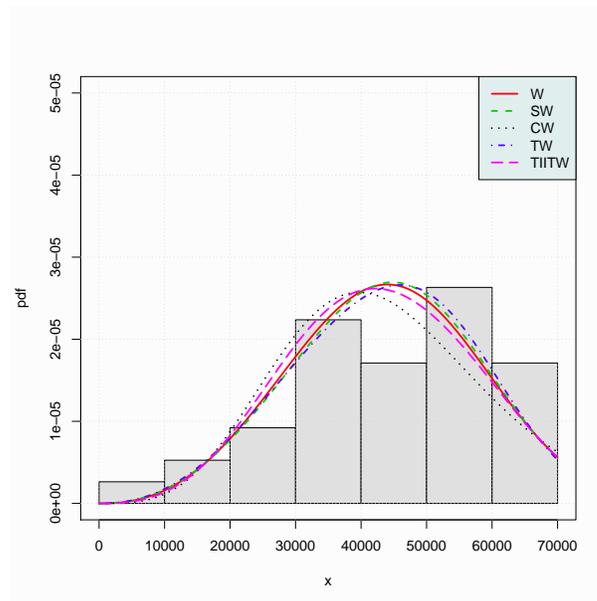
From Table 24, we observe that the data are left-skewed and approximately mesokurtic with a really high variance. Table 25 contains the MLEs and goodness-

of-fit statistics of the considered models.

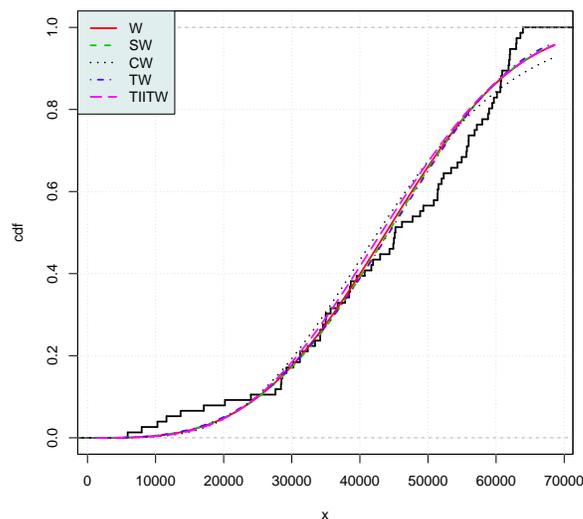
**Table 25.** MLEs and goodness-of-fit statistics for Data set 12.

Model	$\hat{a}$	$\hat{b}$	$-Log$	$A^*$	$W^*$	K-S	PV	AIC	AICC	BIC	HQIC
W	$1.4930 \times 10^{-16}$	3.3750	839.44	1.43	0.18	0.12	0.24	1682.88	1683.04	1687.54	1684.74
SW	$4.0442 \times 10^{-16}$	3.2296	838.81	1.35	0.17	0.11	0.28	1681.61	1681.78	1686.27	1683.48
CW	$1.0689 \times 10^{-10}$	2.1625	845.67	2.22	0.30	0.13	0.17	1695.35	1695.51	1700.01	1697.21
TW	$6.0230 \times 10^{-16}$	3.2650	837.88	1.25	0.16	0.11	0.33	1679.76	1679.93	1684.42	1681.63
TIITW	$4.222 \times 10^{-18}$	3.6897	840.52	1.59	0.21	0.13	0.16	1685.04	1685.20	1689.70	1686.90

According to Table 25, the SW and TW models are the best, the TW model being the better of the two. In Figures 23 and 24, the estimated pdfs and cdfs are displayed over the histogram and empirical cdf of the data respectively.



**Figure 23.** Estimated pdfs for Data set 12.



**Figure 24.** Estimated cdfs for Data set 12

Figures 23 and 24 portray the greater adequacy of the TW model.

## 4. Conclusion

The first contribution of the paper is to offer a new motivated trigonometric class of distributions, which is called the type II Tan-G (TIIT-G) class. We discuss its main features, with some mathematical developments. Then, we perform a complete practical comparative evaluation of the main existing trigonometric classes, defined with the Weibull distribution as baseline. That is, we deal with the Weibull (W), Sin Weibull (SW), Cos Weibull (CW), Tan Weibull (TW) and type II Tan Weibull (TIITW) models. Twelve practical data sets are considered, including two on the COVID-19 pandemic in France from March to June 2020.

Here are some concluding notes. When the data are strongly asymmetric on the right, they can be better fitted by the CW model, mainly, or the TIITW model than the Weibull model. However, when the data are moderately left-skewed, they can be better fitted by the TW model or the SW model than the Weibull model. Furthermore, when the distribution of the data is approximately symmetric, it seems that there is no predominant model. Applications illustrate the usefulness of the considered class of distributions. For certain data sets, the newly introduced TIIT-G class can be the best. In all cases, the results obtained are fairly satisfactory, demonstrating that these trigonometric classes of distributions can be used fairly effectively to analyze a large panel of data sets.

## Acknowledgements

We would like to thank the reviewers for the constructive comments which have helped to improve the paper.

## References

- [1] A. Alzaatreh, F. Famoye and C. Lee, *A new method for generating families of continuous distributions*, *Metron*, 2013, 71, 63–79.
- [2] C. B. Ampadu, *The Tan-G family of distributions with illustration to data in the health sciences*, *Physical Science & Biophysics Journal*, 2019, 3(3), 000125.
- [3] G. R. Aryal, M. O. Edwin, G. G. Hamedani and M. Y. Haitham, *The Topp-Leone generated Weibull distribution: Regression model, characterizations and applications*, *International Journal of Statistics and Probability* 2017, 6, 126–141.
- [4] D. K. Bhaumik, K. Kapur and R. D. Gibbons, *Testing parameters of a gamma distribution for small samples*, *Technometrics*, 2009, 51(3), 326–334.
- [5] A. Bekker, J. J. J. Roux and P. J. Mostert, *A generalization of the compound Rayleigh distribution: using a Bayesian methods on cancer survival times*, *Communication in Statistics: Theory Methods*, 2000, 29(7), 1419–1433.
- [6] G. Casella and R. L. Berger, *Statistical Inference*, Brooks/Cole Publishing Company, California, 1990.
- [7] G. M. Cordeiro, R. B. Silva and A. D. C. Nascimento, *Recent Advances in Lifetime and Reliability Models*, Bentham Books, Sharjah, 2020.  
DOI: 10.2174/97816810834521200101
- [8] R. Dasgupta, *On the distribution of Burr with applications*, *Sankhya*, 2011, 73, 1–19.
- [9] E. J. Fuller, S. Frieman, J. Quinn, G. Quinn and W. Carter, *Fracture mechanics approach to the design of glass aircraft windows: A case study*, *Proceedings of SPIE - The International Society for Optical Engineering*, 1994, 2286, 419–430.
- [10] M. E. Ghitany, B. Atieh and S. Nadarajah, *Lindley distribution and its applications*, *Mathematics Computing and Simulation*, 2008, 78(4), 493–506.
- [11] D. Kumar, U. Singh and S. K. Singh, *A new distribution using sine function- Its application to bladder cancer patients data*, *Journal of Statistics Applications & Probability*, 2015, 4(3), 417–427.
- [12] D. Kumar, U. Singh, S. K. Singh and S. Mukherjee, *The new probability distribution: An aspect to a life time distribution*, *Mathematical Sciences Letters*, 2017, 6(1), 35–42.
- [13] H. Linhart and W. Zucchini, *Model Selection*, John Wiley, New York, 1986.
- [14] P. R. D. Marinho, R. B. Silva, M. Bourguignon, G. M. Cordeiro and S. Nadarajah, *AdequacyModel: An R package for probability distributions and general purpose optimization*, *PLoS ONE*, 2019, 14(8), e0221487.
- [15] D. N. P. Murthy, M. Xie and R. Jiag, *Weibull Models*, John Wiley and Sons, Incorporate Hoboken, New Jersey, 2004.

- [16] F. Proschan, *Theoretical explanation of observed decreasing failure rate*, *Technometrics*, 1963, 5, 375–383.
- [17] R Development Core Team, *R: A language and environment for statistical computing*, *R Foundation for Statistical Computing*, Vienna, 2005.  
<http://www.R-project.org>
- [18] Santé publique France, *Hospital data relative to the COVID-19 epidemic in France*, 1 June, 2020.  
<https://www.data.gouv.fr/en/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/>
- [19] L. Souza, *New trigonometric classes of probabilistic distributions*, Thesis, Universidade Federal Rural de Pernambuco, 2015.
- [20] L. Souza, W. R. O. Junior, C. C. R. de Brito, C. Chesneau, T. A. E. Ferreira and L. Soares, *On the Sin-G class of distributions: theory, model and application*, *Journal of Mathematical Modeling*, 2009, 7(3), 357–379.
- [21] L. Souza, W. R. O. Junior, C. C. R. de Brito, C. Chesneau, T. A. E. Ferreira and L. Soares, *General properties for the Cos-G class of distributions with applications*, *Eurasian Bulletin of Mathematics*, 2019 2(2), 63–79.
- [22] L. Souza, W. R. O. Junior, C. C. R. de Brito, C. Chesneau and T. A. E. Ferreira, *Tan-G class of trigonometric distributions and its applications* (submitted), 2019.
- [23] L. Souza, L. Gallindo and L. Serafim-de-Souza, *CosW: The CosW distribution, R package version 0.1*, 2016.  
<https://CRAN.R-project.org/package=CosW> or by running `install.packages("CosW"); library("CosW"); help("rcosw")` inside R (see [17])
- [24] L. Souza, L. Gallindo and L. Serafim-de-Souza, *SinIW: The SinIW distribution, R package version 0.2*, 2016.  
<https://CRAN.R-project.org/package=SinIW> or by running `install.packages("SinIW"); library("SinIW"); help("rsiniw")` inside R (see [17])
- [25] R. D. Telford and R. B. Cunningham, *Sex, sport, and body-size dependency of hematology in highly trained athletes*, *Medicine and Science in Sports and Exercise*, 1991, 23(7), 788–794.
- [26] K. Xu, M. Xie, L. C. Tang and S. L. Ho, *Application of neural networks in forecasting engine systems reliability*, *Applied Soft Computing*, 2003, 2(4), 255–268.

## Appendix

In this appendix, the code to be used to output all the results for Data set 1 is presented, using the software R. These codes are inspired by those of [23] and [24].

*Technical note:* In the code below, we use the definition of the Weibull distribution behind the following predefined R functions: `pweibull(x, shape = a, scale = b)` and `dweibull(x, shape = a, scale = b)` for the cdf and pdf, respectively. The definitions of these functions do not correspond exactly to those defined in Equation (2.3); the parameters “a” and “b” do not coincide, a re-parametrization

must and will be done. We have voluntarily done that because these R functions allow the activation of useful options in other predefined R functions in specific packages, without alert or error. After extracting the essential, we manipulate the obtained values of the estimated parameters to recover exactly the definition of the Weibull distribution as defined in Equation (2.3); in the code below, the estimates of the former  $a$  and  $b$  are contained into the R vectors named as  $\mathbf{a}$  and  $\mathbf{b}$ , respectively.

```

library(AdequacyModel)
library(moments)
library(fitdistrplus)

## Baseline Weibull ##

cdf.w=function(par,x) {
  a=par[1]
  b=par[2]
  G=pweibull(x, shape = a, scale = b)
  return(G)
}
pdf.w=function(par,x) {
  a=par[1]
  b=par[2]
  g=dweibull(x, shape = a, scale = b)
  return(g)
}

## Distribution ##

# Distribution SW
cdf.sw=function(par,x) {
  G=cdf.w(par,x)
  F=sin((pi/2)*G)
  return(F)
}
pdf.sw=function(par,x) {
  G=cdf.w(par,x)
  g=pdf.w(par,x)
  F= (pi/2)*g*cos((pi/2)*G)
  return(F)
}

# Distribution CW
cdf.cw=function(par,x) {
  G=cdf.w(par,x)
  F=1-cos((pi/2)*G)
  return(F)
}
pdf.cw=function(par,x) {
  G=cdf.w(par,x)
  g=pdf.w(par,x)
  F= (pi/2)*g*sin((pi/2)*G)
  return(F)
}

# Distribution TW
cdf.tw=function(par,x) {
  G=cdf.w(par,x)
  F=tan((pi/4)*G)
  return(F)
}
pdf.tw=function(par,x) {
  G=cdf.w(par,x)
  g=pdf.w(par,x)
  F= (pi/4)*g*(1/cos((pi/4)*G)**2)
  return(F)
}

# Distribution TIITW
cdf.tiitw=function(par,x) {
  G=cdf.w(par,x)

```

```

    F=1-tan((pi/4)*(1-G))
    return(F)
}
pdf_tiiw=function(par,x) {
  G=cdf_w(par,x)
  g=pdf_w(par,x)
  F= (pi/4)*g*(1/cos((pi/4)*(1-G))**2)
  return(F)
}

## Data 1 ##

# Data.

g_p=c( 0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3,
4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1,
7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2,
11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2,
18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0 ,27.0, 31.6, 33.1 ,38.5)

# Initial parameters.

fit.w = fitdist(g_p, "weibull")
a=fit.w$estimate[1]
b=fit.w$estimate[2]

# goodness.fit.

Res_w = goodness.fit(data = g_p, pdf = pdf_w, cdf = cdf_w, starts = c(a,b),
  method = "N", mle = NULL, domain = c(0,Inf))

Res_sw = goodness.fit(data = g_p, pdf = pdf_sw, cdf = cdf_sw, starts = c(a,b),
  method = "N", mle = NULL, domain = c(0,Inf))

Res_cw = goodness.fit(data = g_p, pdf = pdf_cw, cdf = cdf_cw, starts = c(a,b),
  method = "N", mle = NULL, domain = c(0,Inf))

Res_tw = goodness.fit(data = g_p, pdf = pdf_tw, cdf = cdf_tw, starts = c(a,b),
  method = "N", mle = NULL, domain = c(0,Inf))

Res_tiiw = goodness.fit(data = g_p, pdf = pdf_tiiw, cdf = cdf_tiiw, starts = c(a,b),
  method = "N", mle = NULL, domain = c(0,Inf))

# Descriptive statistics.

ds = data.frame(mean=mean(g_p),
  median=median(g_p),
  SD=sd(g_p),
  Variance=var(g_p),
  Skewness=skewness(g_p),
  Kurtosis=kurtosis(g_p),
  Min=min(g_p),
  Max=max(g_p))

ds

# MLEs and some statistics.

df = data.frame(

```

```

dist = c("G", "SG", "CG", "TG", "TIITG"),
a = c(Res_w$mle[2]**(-Res_w$mle[1]), Res_sw$mle[2]**(-Res_sw$mle[1]), Res_cw$mle[2]
**(-Res_cw$mle[1]), Res_tw$mle[2]**(-Res_tw$mle[1]), Res_titw$mle[2]**(-Res_titw$mle[1])),
b = c(Res_w$mle[1], Res_sw$mle[1], Res_cw$mle[1], Res_tw$mle[1], Res_titw$mle[1]),
lambda = c(Res_w$Value, Res_sw$Value, Res_cw$Value, Res_tw$Value, Res_titw$Value),
A = c(Res_w$A, Res_sw$A, Res_cw$A, Res_tw$A, Res_titw$A),
W = c(Res_w$W, Res_sw$W, Res_cw$W, Res_tw$W, Res_titw$W),
KS = c(Res_w$KS$statistic, Res_sw$KS$statistic, Res_cw$KS$statistic, Res_tw$KS$statistic,
Res_titw$KS$statistic),
pv = c(Res_w$KS$p.value, Res_sw$KS$p.value, Res_cw$KS$p.value, Res_tw$KS$p.value,
Res_titw$KS$p.value),
AIC = c(Res_w$AIC, Res_sw$AIC, Res_cw$AIC, Res_tw$AIC, Res_titw$AIC),
CAIC = c(Res_w$CAIC, Res_sw$CAIC, Res_cw$CAIC, Res_tw$CAIC, Res_titw$CAIC),
BIC = c(Res_w$BIC, Res_sw$BIC, Res_cw$BIC, Res_tw$BIC, Res_titw$BIC),
HQIC = c(Res_w$HQIC, Res_sw$HQIC, Res_cw$HQIC, Res_tw$HQIC, Res_titw$HQIC))

df

# Graphics: pdf and cdf

par(mfrow=c(2,1))
par(bg="gray99")
hist(g_p, prob=T, ylim =c(0,0.08), col="gray88", xlab = "x",
ylab = "pdf", main = "");box();grid()
curve(pdf_w(Res_w$mle,x), add = T, col=2, lwd=2)
curve(pdf_sw(Res_sw$mle,x), add = T, col=3, lwd=2, lty=2)
curve(pdf_cw(Res_cw$mle,x), add = T, col=1, lwd=2, lty=3)
curve(pdf_tw(Res_tw$mle,x), add = T, col=4, lwd=2, lty=4)
curve(pdf_titw(Res_titw$mle,x), add = T, col=6, lwd=2, lty=5)

legend("topright", c(expression("W"), expression("SW"), expression("CW"), expression("TW"),
expression("TIITW")), lwd=c(2,2,2,2,2), lty=c(1,2,3,4,5), col=c(2,3,1,4,6), cex=1, bty = "o",
bg = "azure2")

plot(ecdf(g_p),xlab = "x",ylab = "cdf",main = "",do.points=F,verticals=T,lwd=2)
box();grid()
curve(cdf_w(Res_w$mle,x),add = T, col=2,lwd=2)
curve(cdf_sw(Res_sw$mle,x), add = T, col=3, lwd=2, lty=2)
curve(cdf_cw(Res_cw$mle,x), add = T, col=1, lwd=2, lty=3)
curve(cdf_tw(Res_tw$mle,x), add = T, col=4, lwd=2, lty=4)
curve(cdf_titw(Res_titw$mle,x), add = T, col=6, lwd=2, lty=5)

legend("topleft", c(expression("W"), expression("SW"), expression("CW"), expression("TW"),
expression("TIITW")), lwd=c(2,2,2,2,2), lty=c(1,2,3,4,5), col=c(2,3,1,4,6), cex=1, bty = "o",
bg = "azure2")

# For the results of the other data sets, it is enough to copy-paste the data in the vector g_p.

```