

RESEARCH ARTICLE

# Beyond the Sin-G family: The transformed Sin-G family

Farrukh Jamal<sup>1</sup>, Christophe Chesneau<sup>2</sup>, Dalal Lala Bouali<sup>2,3</sup>, Mahmood UI Hassan<sup>4</sup>\*

**1** Department of Statistics, The Islamia University, Bahawalpur, Pakistan, **2** Department of Mathematics, Université de Caen, Caen, France, **3** Laboratory of Applied Mathematics, Mohamed Khider University of Biskra, Biskra, Algeria, **4** Stockholm University, Stockholm, Sweden

☞ These authors contributed equally to this work.  
\* scenic555@gmail.com



## Abstract

In recent years, the trigonometric families of continuous distributions have found a place of choice in the theory and practice of statistics, with the Sin-G family as leader. In this paper, we provide some contributions to the subject by introducing a flexible extension of the Sin-G family, called the transformed Sin-G family. It is constructed from a new polynomial-trigonometric function presenting a desirable “versatile concave/convex” property, among others. The modelling possibilities of the former Sin-G family are thus multiplied. This potential is also highlighted by a complete theoretical work, showing stochastic ordering results, studying the analytical properties of the main functions, deriving several kinds of moments, and discussing the reliability parameter as well. Then, the applied side of the proposed family is investigated, with numerical results and applications on the related models. In particular, the estimation of the unknown model parameters is performed through the use of the maximum likelihood method. Then, two real life data sets are analyzed by a new extended Weibull model derived to the considered trigonometric mechanism. We show that it performs the best among seven comparable models, illustrating the importance of the findings.

## OPEN ACCESS

**Citation:** Jamal F, Chesneau C, Bouali DL, UI Hassan M (2021) Beyond the Sin-G family: The transformed Sin-G family. PLoS ONE 16(5): e0250790. <https://doi.org/10.1371/journal.pone.0250790>

**Editor:** Alan D Hutson, Roswell Park Cancer Institute, UNITED STATES

**Received:** November 23, 2020

**Accepted:** April 13, 2021

**Published:** May 11, 2021

**Copyright:** © 2021 Jamal et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## 1 Introduction

Recent advances in probability distribution theory and applications have seen the rise of various general families of distributions, successfully applied for different statistical problems. In this regard, a nice survey can be found in [1]. Here, we put the light on the trigonometric families of continuous distributions, i.e., those defined by a cumulative distribution function (cdf) involving trigonometric functions (sine, cosine, tangent, cotangent, and various combinations of these). The pioneer work is about the Sin-G family developed by [2–5]. As indicated by its name, it is defined around the sine function; the corresponding cdf is given by

$$F(x; \zeta) = \sin \left[ \frac{\pi}{2} G(x; \zeta) \right], \quad x \in \mathbb{R}, \tag{1}$$

where  $G(x; \zeta)$  is a baseline cdf of a continuous distribution with parameter(s) vector denoted

by  $\zeta$ . It is now demonstrated that the Sin-G family has the ability to provide flexible statistical models to fit data of various nature. Also, it is a simple alternative to the model derived to the baseline distribution, without the addition of parameter. For instance, in [2], the exponential distribution is used as a baseline to construct the SinE model, which reveals to suitably fit the famous bladder cancer patients data of [6]. Also, he has the better fit as compared to some classical models such as the former exponential one, having better Akaike information criteria (AIC), Bayesian information criteria (BIC) and Kolmogorov-Smirnov (KS) test values. On the other side, based on the inverse Weibull distribution (see [7]), the SinIW model was introduced by [4], with application to the so-called Guinea pigs data by [8], providing better BIC in comparison to some other solid models. A “free for all” R package on the SinIW model is provided in [9]. As a matter of fact, the qualities of the models derived to the Sin-G family have inspired other general families of continuous distributions also centered around trigonometric functions, such as the Cos-G family by [5], CS-G family by [10], NSin-G family by [11], TransSC-G family by [12], SinTL-G family by [13], SinKum-G family by [14], and SinEOF-G family by [15]. The majority of these families are based on the Sin-G structure, with no additional tuning parameters or transformations.

In this paper, we go further the Sin-G family by proposing a new extended version of it, called the transformed Sin-G (TS-G) family. The corresponding cdf is derived to (1), with the use of a simple one-parameter polynomial-trigonometric transformation. This transformation has the following features: (i) it is analytically simple and includes the non-transformed case, (ii) it has the properties of a continuous cdf, that is, has its values into the unit interval, is continuous, almost everywhere differentiable and increasing, and (iii) it can be convex or concave, or none of them, for well-identified values of the parameter. Thanks to its versatility, this transformation significantly enhances the flexible properties of (1), and the baseline cdf as well. Thus, the TS-G family distinguishes itself from other modified Sin-G families by its overall simplicity, original polynomial-trigonometric functions, and the advantage of flexible kurtosis, skewness, versatile distribution tails, and various hazard rate shapes, as a result of the considered transformation. Thus, the TS-G family can provide interesting models for diverse fitting purposes. This practical aspect, along with important theoretical results, are developed in this study.

The rest of the paper is organized as follows. The basics on the TS-G family are presented in Section 2. Also, an emphasis is put on a special distribution of the family based on the Weibull distribution, motivated by its desirable shapes characteristics in the modelling sense. In Section 3, interesting properties of the TS-G family are studied, including stochastic ordering results, equivalence properties, critical points analysis, series expansion involving known exponentiated functions, moments, and reliability parameter. In Section 4, by adopting a statistical approach, the TS-G model parameters are estimated with the maximum likelihood method, supported by a simulation study. Then, applications of this special model are addressed in Section 5, showing how the new family can be of interest to fit various data sets, outperforming seven other solid extended or modified Weibull models of the literature. Section 6 formulates concluding remarks.

## 2 Basics on the TS-G family

In this section, the TS-G family is defined, with motivations and discussions.

### 2.1 On a special polynomial-trigonometric function

The following result presents some interesting features of a simple polynomial-trigonometric function, which will be at the basis of the TS-G family.

**Proposition 1** Let  $\lambda \in [0, 1]$  and  $T_\lambda(x)$  be the following parametric function:

$$T_\lambda(x) = \sin\left(\frac{\pi}{2}x\right) - \lambda \frac{\pi}{2}x \cos\left(\frac{\pi}{2}x\right), \quad x \in [0, 1], \tag{2}$$

with  $T_\lambda(x) = 0$  if  $x < 0$  and  $T_\lambda(x) = 1$  if  $x > 1$ . Then, the following properties hold:

- $T_\lambda(x)$  has the properties of a continuous cdf,
- $T_\lambda(x)$  can be convex or concave according to the values of  $\lambda$ . In particular, for  $\lambda \in [0, 1/3]$ ,  $T_\lambda(x)$  is concave and, for  $\lambda \in [1/2, 1]$ ,  $T_\lambda(x)$  is convex.
- For  $\lambda \in (1/3, 1/2)$ ,  $T_\lambda(x)$  can be neither convex nor concave.

**Proof.** First of all, the following inequality holds: for  $y \in [0, \pi/2]$ , we have

$$\sin(y) \geq y \cos(y), \tag{3}$$

(see [16]). Let us now prove the first point of the proposition. Since  $\lambda \in [0, 1]$ , it follows from (3) that  $0 \leq T_1(x) \leq T_\lambda(x) \leq \sin[(\pi/2)x] \leq 1$ . Also,  $T_\lambda(x)$  satisfies  $T_\lambda(0) = 0$  and  $T_\lambda(1) = 1$ , it is continuous, differentiable and, by differentiating on  $x$ , we have

$$\frac{d}{dx} T_\lambda(x) = \frac{\pi}{2} \left[ (1 - \lambda) \cos\left(\frac{\pi}{2}x\right) + \lambda \frac{\pi}{2}x \sin\left(\frac{\pi}{2}x\right) \right].$$

As a sum of positive functions, we have  $dT_\lambda(x)/dx \geq 0$ , so  $T_\lambda(x)$  is increasing. We conclude that  $T_\lambda(x)$  has the properties of a continuous cdf. For the second point of the proof, let us notice that, by differentiating on  $x$ , we have

$$\frac{d^2}{dx^2} T_\lambda(x) = \frac{\pi^2}{4} \left[ (2\lambda - 1) \sin\left(\frac{\pi}{2}x\right) + \lambda \frac{\pi}{2}x \cos\left(\frac{\pi}{2}x\right) \right].$$

Therefore, if  $\lambda \in [0, 1/3]$ , it follows from  $2\lambda - 1 \leq -1/3$  and (3) that

$$\frac{d^2}{dx^2} T_\lambda(x) \leq -\frac{\pi^2}{12} \left[ \sin\left(\frac{\pi}{2}x\right) - \frac{\pi}{2}x \cos\left(\frac{\pi}{2}x\right) \right] \leq 0.$$

That is,  $T_\lambda(x)$  is concave. On the other hand, if  $\lambda \in [1/2, 1]$ , we have  $d^2 T_\lambda(x)/dx^2 \geq 0$  as a sum of positive functions, implying that  $T_\lambda(x)$  is convex.

Now, for  $\lambda = 2/5 \in (1/3, 1/2)$ , we have

$$\frac{d^2}{dx^2} T_\lambda(0.1) = 0.07592538 > 0, \quad \frac{d^2}{dx^2} T_\lambda(0.8) = -0.08606892 < 0,$$

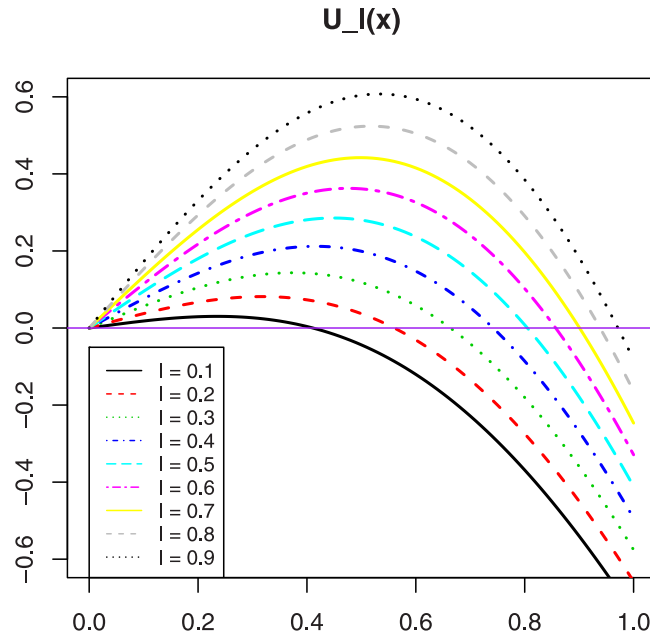
implying that  $T_\lambda(x)$  can be neither convex nor concave. As a visual approach, if we set  $U_\ell(x) = d^2 T_{\lambda_\ell}(x)/dx^2$ , with  $\lambda_\ell = \ell/2 + (1 - \ell)/3$  and  $\ell \in \{0.1, 0.2, \dots, 0.9\}$ , so that  $\lambda_\ell \in (1/3, 1/2)$ , Fig 1 shows that  $U_\ell(x)$  can be positive and negative, implying that  $T_\lambda(x)$  is neither convex nor concave for the considered values of  $\lambda$ . This concludes the proof of Proposition 1.

One can remark that the function  $T_\lambda(x)$  defined by (2) can be written as  $T_\lambda(x) = T_\lambda^*\{\sin[(\pi/2)x]\}$ , where

$$T_\lambda^*(x) = x - \lambda \arcsin(x) \sqrt{1 - x^2}, \quad x \in [0, 1].$$

One can establish that the function  $T_\lambda^*(x)$  has the properties of a cdf, which is not mentioned in the existing literature.

In view of Proposition 1, the transformation function  $T_\lambda^*(x)$  allows to “convexify (or not)” the convex cdf  $s(x) = \sin[(\pi/2)x]$ ,  $x \in [0, 1]$ , while keeping its cdf properties. This ability is not



**Fig 1. Plots of the function  $U_\ell(x)$  for  $\ell \in \{0.1, 0.2, \dots, 0.9\}$ .**

<https://doi.org/10.1371/journal.pone.0250790.g001>

satisfied by some other simple transformation functions, as the power transformation, i.e.,  $T_\gamma^{**}(x) = x^\gamma$  with  $\gamma > 0$ , for instance. This aspect is the driving force behind the TS-G family, which aims to expand the Sin-G family in a straightforward manner to open new statistical perspectives. We show the convex/concave properties of the function  $T_\lambda(x)$  given by (2) in Fig 2, by considering several values for  $\lambda$ .

### 2.2 Definition

By taking the benefits of the flexibility of  $T_\lambda(x)$  given by (2) as described in Proposition 1, the proposed TS-G family of continuous distributions is defined by the following cdf:

$$F(x; \lambda, \zeta) = \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] - \lambda \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right], \quad x \in \mathbb{R}, \quad (4)$$

where  $\lambda \in [0, 1]$  and, as usual,  $G(x; \zeta)$  is a baseline cdf of a continuous distribution with parameter(s) vector denoted by  $\zeta$ .

That is, by considering the transformations  $T_\lambda(x)$  and  $T_\lambda^*(x)$  discussed above, we have  $F(x; \lambda, \zeta) = T_\lambda[G(x; \zeta)]$  or, equivalently,  $F(x; \lambda, \zeta) = T_\lambda^* \{ \sin [(\pi/2)G(x; \zeta)] \}$ , motivating the name of “transformed Sin-G family”. One can notice that the cdf of the former Sin-G family is derived by taking  $\lambda = 0$ . Also, based on Proposition 1 and the convex/concave properties of  $T_\lambda^*(x)$ , we argue that the overall flexibility of the cdf of the former Sin-G family provided by (1) is enhanced. This is concretized by the addition of the modulating polynomial-cosine term  $\lambda(\pi/2)G(x; \zeta)\cos[(\pi/2)G(x; \zeta)]$ , which opens up a whole new world of possibilities.

Also, one can write  $F(x; \lambda, \zeta)$  as a simple mixture of two cdfs of the TS-G family itself:  $F(x; 0, \zeta)$  and  $F(x; 1, \zeta)$ , with the weights  $1 - \lambda$  and  $\lambda$ , respectively, i.e.,

$$F(x; \lambda, \zeta) = (1 - \lambda)F(x; 0, \zeta) + \lambda F(x; 1, \zeta).$$

Hence, the role of  $\lambda$  is to balance  $F(x; 0, \zeta)$  and  $F(x; 1, \zeta)$ , each reaching different targets in terms of statistical modelling.

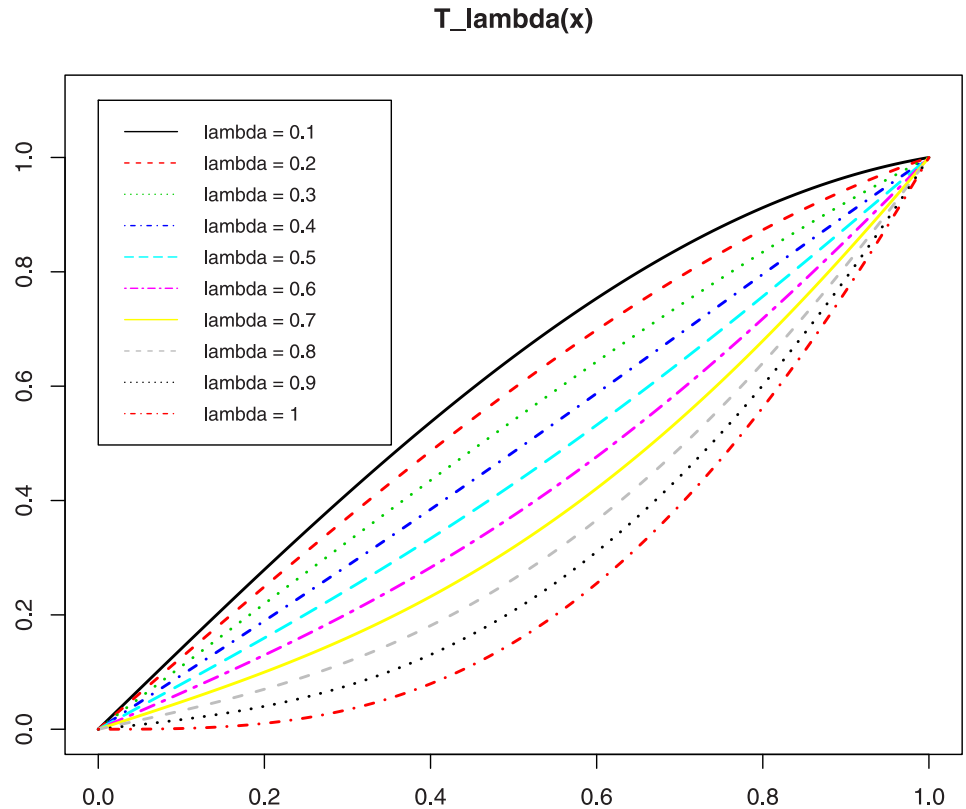


Fig 2. Plots of the function  $T_\lambda(x)$  for  $\lambda \in \{0.1, 0.2, \dots, 1\}$ .

<https://doi.org/10.1371/journal.pone.0250790.g002>

Among the other functions of interest, the survival function (sf) of the TS-G family is given by

$$S(x; \lambda, \zeta) = 1 - \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + \lambda \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right], \quad x \in \mathbb{R}.$$

Upon an almost everywhere differentiation of  $F(x; \lambda, \zeta)$  with respect to  $x$ , the corresponding probability density function (pdf) is given by

$$f(x; \lambda, \zeta) = \frac{\pi}{2} g(x; \zeta) \left\{ \lambda \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\}, \quad (5)$$

where  $g(x; \zeta)$  is the pdf of the baseline distribution, i.e., obtained by an almost everywhere differentiation of  $G(x; \zeta)$ .

Another important function of the TS-G family, specially when the support of the baseline distribution is  $(0, +\infty)$ , is the hazard rate function (hrf) defined by

$$h(x; \lambda, \zeta) = \frac{\frac{\pi}{2} g(x; \zeta) \left\{ \lambda \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\}}{1 - \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + \lambda \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]}, \quad x \in \mathbb{R}. \quad (6)$$

For the importance of the sf and hrf, in reliability analysis mainly, we may refer the reader to [17], and the references therein.

### 2.3 A special distribution: The TSW distribution

Naturally, each choice for  $G(x; \zeta)$  gives a new TS-G distribution. Here, we focus our attention on the Weibull distribution as baseline, i.e., defined by the following cdf:

$$G(x; \alpha, \beta) = 1 - e^{-\alpha x^\beta}, \quad x > 0, \tag{7}$$

and  $G(x; \alpha, \beta) = 0$  if  $x \leq 0$ , where  $\alpha > 0$  and  $\beta > 0$  are scale and shape parameters, respectively. As a main interest, the Weibull distribution is known to be an alternative to the exponential distribution, offering more flexible hazard rate shapes; decreasing and increasing shapes can be observed. It has been involved with success in a plethora of applications requiring the analysis of lifetime and reliability data. In this regard, we may refer the reader to [18–20].

We thus aim to extend the Weibull distribution, along with their properties, via the use of the TS-G family. That is, by inserting (7) into (4), we introduce the TSW distribution defined by the following cdf:

$$\begin{aligned} F(x; \lambda, \alpha, \beta) &= \sin \left[ \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \right] - \lambda \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \cos \left[ \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \right] \\ &= \cos \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] - \lambda \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \sin \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right], \quad x > 0, \end{aligned}$$

and  $F(x; \lambda, \alpha, \beta) = 0$  if  $x \leq 0$ , where the second expression is obtained after some trigonometric manipulations.

Also, the corresponding sf, pdf and hrf are, respectively, given by

$$S(x; \lambda, \alpha, \beta) = 1 - \cos \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] + \lambda \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \sin \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right], \quad x > 0,$$

and  $S(x; \lambda, \alpha, \beta) = 1$  if  $x \leq 0$ ,

$$f(x; \lambda, \alpha, \beta) = \frac{\pi}{2} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} \left\{ \lambda \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \cos \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] + (1 - \lambda) \sin \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] \right\}, \quad x > 0,$$

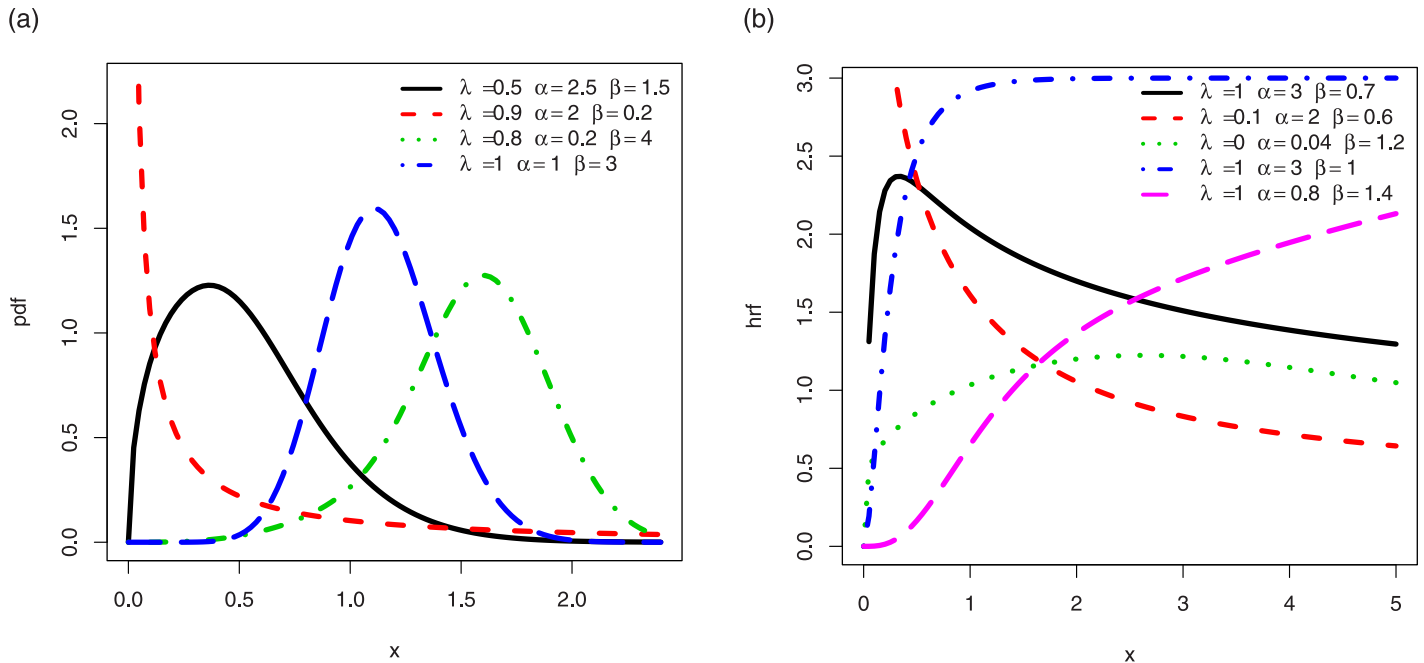
and  $f(x; \lambda, \alpha, \beta) = 0$  if  $x \leq 0$ , and

$$\begin{aligned} h(x; \lambda, \alpha, \beta) &= \frac{\frac{\pi}{2} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} \left\{ \lambda \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \cos \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] + (1 - \lambda) \sin \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] \right\}}{1 - \cos \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right] + \lambda \frac{\pi}{2} (1 - e^{-\alpha x^\beta}) \sin \left[ \frac{\pi}{2} e^{-\alpha x^\beta} \right]}, \\ &x > 0, \end{aligned}$$

and  $h(x; \lambda, \alpha, \beta) = 0$  if  $x \leq 0$ .

After some graphical investigations, the curvature properties of the functions of the TSW distribution reveal to be desirably versatile. Evidence can be seen in Fig 3, which displays some plots of the corresponding pdf and hrf for various values of the parameters.

In particular, Fig 3(a) indicates that the pdf of the TSW distribution has various skewness shapes (near symmetrical, left, right, bathtub, reversed-J shapes, mainly), along with different kurtosis properties. Fig 3(b) reveals that the corresponding hrf possesses versatile shapes, such as decreasing, increasing, bathtub (classic and upside-down) and reversed-J shapes. These observations imply that the TSW distribution is adequate to fit heterogeneous data sets. In our study, this aspect will be developed in Section 5, where the TSW distribution is used to fit two real life data sets. Also, it will be compared with other extended or modified Weibull models, and the results will be quite favorable to the TSW model.



**Fig 3. Selection of plots for (a) the pdf and (b) the hrf of the TSW distribution.**

<https://doi.org/10.1371/journal.pone.0250790.g003>

### 3 Notable mathematical properties

Here, we explore some mathematical properties of interest satisfied by the TS-G family.

#### 3.1 Stochastic ordering results

Stochastic ordering results are crucial to understand a certain hierarchy existing between the distributions, with consequence on their comparison from the modelling point of view. In the framework of the TS-G family, the following result presents some relations involving the cdf of the TS-G family (beyond the following immediate stochastic ordering property:  $F(x; \lambda, \zeta) \leq F(x; 0, \zeta)$ ).

**Proposition 2** *The following inequalities hold:*

- If  $\lambda_2 \geq \lambda_1 \geq 0$ , we have  $F(x; \lambda_2, \zeta) \leq F(x; \lambda_1, \zeta)$ .
- For  $\lambda \in [0, 2/\pi]$ , we have  $F(x; \lambda, \zeta) \geq F_*(x; \zeta)$ , where

$$F_*(x; \zeta) = G(x; \zeta) \left\{ 1 - \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\}$$

is a valid cdf.

**Proof.** Based on (4), since  $\lambda_2 \geq \lambda_1$  and the involved functions are positive, we have

$$\sin \left[ \frac{\pi}{2} G(x; \zeta) \right] - \lambda_2 \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \leq \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] - \lambda_1 \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right],$$

implying the desired inequality.

For the second point, the following inequality holds: for  $y \in [0, \pi/2]$ , we have  $\sin(y) \geq y(2/\pi)$  (see [16]). Hence, based on (4), since  $\lambda \in [0, 2/\pi]$ , we have

$$\begin{aligned} F(x; \lambda, \zeta) &\geq G(x; \zeta) - \lambda \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \\ &\geq G(x; \zeta) - G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] = F_*(x; \zeta). \end{aligned}$$

Then, one can remark that  $F_*(x; \zeta)$  is the cdf of the rv  $Z = \max(X, Y)$ , where  $X$  is a rv having the (baseline) cdf  $G(x; \zeta)$  and  $Y$  is a rv having the cdf of the Cos-G family (see [5]), with  $X$  and  $Y$  independent.

The following result is about a likelihood stochastic ordering of the TS-G family. We refer the reader to [21] for the details on the concept of likelihood stochastic order.

**Proposition 3** *Let  $X_1$  be a rv having the cdf  $F(x; \lambda_1, \zeta)$  and  $X_2$  be a rv having the cdf  $F(x; \lambda_2, \zeta)$ . Then, if  $\lambda_2 \geq \lambda_1$ , we have  $X_1 \leq X_2$  in the likelihood stochastic ordering sense.*

**Proof.** Following [21], we have  $X_1 \leq X_2$  in the likelihood stochastic ordering sense if and only if the following ratio function is decreasing with respect to  $x$ :

$$r(x; \lambda_1, \lambda_2, \zeta) = \frac{f(x; \lambda_1, \zeta)}{f(x; \lambda_2, \zeta)}, \quad x \in \mathbb{R},$$

where  $f(x; \lambda_1, \zeta)$  and  $f(x; \lambda_2, \zeta)$  are the corresponding pdfs of  $F(x; \lambda_1, \zeta)$  and  $F(x; \lambda_2, \zeta)$ , respectively. That is, by using (5), we have

$$r(x; \lambda_1, \lambda_2, \zeta) = \frac{\lambda_1 \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda_1) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]}{\lambda_2 \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda_2) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]}, \quad x \in \mathbb{R}.$$

Upon an almost everywhere differentiation with respect to  $x$ , after some developments, we get

$$\frac{d}{dx} r(x; \lambda_1, \lambda_2, \zeta) = (\lambda_1 - \lambda_2) \frac{\pi g(x; \zeta) [\pi G(x; \zeta) + \sin [\pi G(x; \zeta)]]}{4 \left\{ \lambda_2 \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda_2) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\}^2},$$

which is negative if and only if  $\lambda_2 \geq \lambda_1$ , implying the desired result.

### 3.2 Equivalence properties

Here, some equivalence properties of crucial functions of the TS-G family are discussed, which can be helpful to find their limits and also, understand the tails properties of the distribution.

As  $G(x; \zeta) \rightarrow 0$ , we establish that

$$F(x; \lambda, \zeta) \sim \frac{\pi}{2} (1 - \lambda) G(x; \zeta), \quad f(x; \lambda, \zeta) \sim \frac{\pi}{2} (1 - \lambda) g(x; \zeta), \quad h(x; \lambda, \zeta) \sim \frac{\pi}{2} (1 - \lambda) g(x; \zeta).$$

Also, as  $G(x; \zeta) \rightarrow 1$ , we have

$$F(x; \lambda, \zeta) \sim 1 - \frac{\pi^2}{4} \lambda [1 - G(x; \zeta)], \quad f(x; \lambda, \zeta) \sim \frac{\pi^2}{4} \lambda g(x; \zeta), \quad h(x; \lambda, \zeta) \sim \frac{g(x; \zeta)}{1 - G(x; \zeta)}.$$

In each case, we see how the new parameter  $\lambda$  modulates the limits; it has a strong effect in this regard, except for the hrf when  $G(x; \zeta) \rightarrow 1$ .



In the case of the TSW distribution as described in Subsection 2.3, the following equivalence hold. As  $x \rightarrow 0$ , we have

$$F(x; \lambda, \alpha, \beta) \sim \frac{\pi}{2}(1 - \lambda)\alpha x^\beta, \quad f(x; \lambda, \alpha, \beta) \sim \frac{\pi}{2}(1 - \lambda)\alpha\beta x^{\beta-1}$$

and

$$h(x; \lambda, \alpha, \beta) \sim \frac{\pi}{2}(1 - \lambda)\alpha\beta x^{\beta-1}.$$

Therefore, we obtain  $\lim_{x \rightarrow 0} f(x; \lambda, \alpha, \beta) = \lim_{x \rightarrow 0} f(x; \lambda, \alpha, \beta) = \ell$  with  $\ell = +\infty$  if  $\beta < 1$ ,  $\ell = (\pi/2)(1 - \lambda)\alpha$  if  $\beta = 1$  and  $\ell = 0$  if  $\beta > 1$ .

Also, as  $x \rightarrow +\infty$ , we have

$$F(x; \lambda, \alpha, \beta) \sim 1 - \frac{\pi^2}{4}\lambda e^{-\alpha x^\beta}, \quad f(x; \lambda, \alpha, \beta) \sim \frac{\pi^2}{4}\lambda\alpha\beta x^{\beta-1} e^{-\alpha x^\beta}$$

and

$$h(x; \lambda, \alpha, \beta) \sim \alpha\beta x^{\beta-1}.$$

Hence, we have  $\lim_{x \rightarrow +\infty} f(x; \lambda, \alpha, \beta) = 0$  in all the situations, and  $\lim_{x \rightarrow +\infty} h(x; \lambda, \alpha, \beta) = \ell$  with  $\ell = 0$  if  $\beta < 1$ ,  $\ell = \alpha$  if  $\beta = 1$  and  $\ell = +\infty$  if  $\beta > 1$ .

### 3.3 Critical points

Some analytical facts about the critical points of functions of the TS-G family are now presented. First of all, the study of critical point(s), i.e., mode(s), of  $f(x; \lambda, \zeta)$  informs us on the possible singularities of the related model. A critical point of  $f(x; \lambda, \zeta)$ , say  $x^*$ , is solution of the following non-linear equation:  $df(x; \lambda, \zeta)/dx = 0$ , which is equivalent to be solution of the following more tractable non-linear equation:  $d\{\log[f(x; \lambda, \zeta)]\}/dx = 0$ , i.e.,

$$\frac{dg(x; \zeta)/dx}{g(x; \zeta)} + \frac{\pi}{2}g(x; \zeta) \frac{(2\lambda - 1) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + \frac{\pi}{2} \lambda G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]}{\lambda \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]} = 0.$$

Then, the nature of  $x^*$  depends on the values of  $\eta = d^2\{\log[f(x; \lambda, \zeta)]\}/dx^2|_{x=x^*}$ . More specifically,  $x^*$  is designated as a local maximum point if  $\eta < 0$ , an inflection point if  $\eta = 0$ , and a local minimum point if  $\eta > 0$ .

The same methodology can be applied to study the critical points for  $h(x; \lambda, \zeta)$ , which can be useful to identify specific hazard rate shapes (monotonic, bathtub, S. . .) for a modelling aim. Let us just mention that a critical point for  $h(x; \lambda, \zeta)$  is solution of the following non-linear equation:  $d\{\log[h(x; \lambda, \zeta)]\}/dx = 0$ , i.e.,

$$\begin{aligned} & \frac{dg(x; \zeta)/dx}{g(x; \zeta)} + \frac{\pi}{2}g(x; \zeta) \frac{(2\lambda - 1) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + \frac{\pi}{2} \lambda G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]}{\lambda \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]} \\ & + \frac{\frac{\pi}{2}g(x; \zeta) \left\{ \lambda \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\}}{1 - \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + \lambda \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right]} = 0. \end{aligned}$$

Clearly, in most of the cases, the critical points of  $f(x;\lambda, \zeta)$  and  $h(x;\lambda, \zeta)$  have not closed-forms. They can however be determined as numerical values by using mathematical softwares, as Mathematica, Python, R, Matlab. . .

Concerning the TSW distribution, numerical investigations, supported by Fig 3 as well, show that it is unimodal, with a corresponding hrf that can have one critical point.

### 3.4 A series expansion

The following result establishes a new representation of the pdf of the TS-G family involving exponentiated baseline pdfs. Such results are common for the pdfs of modern general families of continuous distributions (see, e.g., [4, 11, 22]).

**Proposition 4** For any  $x$  such that  $G(x;\zeta) < 1$ , the following series expansion holds:

$$f(x; \lambda, \zeta) = \sum_{k=0}^{+\infty} a_k v_{2k+1}(x; \zeta),$$

where  $a_k = (\pi/2)^{2k+1} (-1)^k [1 - \lambda(2k + 1)] / (2k + 1)!$  and  $v_\gamma = \gamma g(x;\zeta) G(x;\zeta)^{\gamma-1}$ , with  $\gamma = 2k + 1$ .

**Proof.** Owing to the series expansions of the sine and cosine functions, after some developments, we get

$$\begin{aligned} F(x; \lambda, \zeta) &= \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] - \lambda \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right], \\ &= \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k + 1)!} \left[ \frac{\pi}{2} G(x; \zeta) \right]^{2k+1} - \lambda \frac{\pi}{2} G(x; \zeta) \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!} \left[ \frac{\pi}{2} G(x; \zeta) \right]^{2k} \\ &= \sum_{k=0}^{+\infty} a_k [G(x; \zeta)]^{2k+1}. \end{aligned}$$

We end the proof of Proposition 4 by differentiating the above function with respect to  $x$ .

Proposition 4 is of interest because the properties of most of the exponentiated standard distributions are well known, and thus, can be used to determine those of the TS-G family. Also, from the practical point of view, it allows us to define some integral terms by the means of (infinite) sums, which sometimes give less error than compute the integral directly. In this regard, we refer to the discussion in [22].

In the setting of the TSW distribution, we have

$$f(x; \lambda, \alpha, \beta) = \sum_{k=0}^{+\infty} a_k v_{2k+1}(x; \alpha, \beta),$$

where  $v_{2k+1}(x;\zeta)$  denotes the pdf of the exponentiated Weibull distribution, defined with power parameter  $2k + 1$  (see [23]), i.e.,

$$v_{2k+1}(x; \alpha, \beta) = (2k + 1) \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} (1 - e^{-\alpha x^\beta})^{2k}, \quad x > 0,$$

and  $v_{2k+1}(x;\alpha, \beta) = 0$  if  $x \leq 0$ . Further details about the exponentiated Weibull distribution can also be found in [24, 25].

### 3.5 Generalities on the moments

Let  $X$  be a rv having the cdf  $F(x;\alpha, \beta, \zeta)$  given by (4) (and the pdf  $f(x;\alpha, \beta, \zeta)$  given by (5)) and  $\phi(x)$  be a function. Then, assuming that it makes mathematical sense, the expectation of  $\phi(X)$  is

obtained as

$$\begin{aligned} \Theta_\phi(X) &= E[\phi(X)] = \int_{-\infty}^{+\infty} \phi(x)f(x; \lambda, \zeta)dx \\ &= \int_{-\infty}^{+\infty} \phi(x) \frac{\pi}{2}g(x; \zeta) \left\{ \lambda \frac{\pi}{2}G(x; \zeta) \sin \left[ \frac{\pi}{2}G(x; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2}G(x; \zeta) \right] \right\} dx \\ &= I_\phi^{(1)} + I_\phi^{(2)}, \end{aligned}$$

where, by denoting  $Q_G(u; \zeta)$  the inverse function of  $G(x; \zeta)$ ,

$$\begin{aligned} I_\phi^{(1)} &= \lambda \frac{\pi^2}{4} \int_{-\infty}^{+\infty} \phi(x)g(x; \zeta)G(x; \zeta) \sin \left[ \frac{\pi}{2}G(x; \zeta) \right] dx \\ &= \lambda \frac{\pi^2}{4} \int_0^1 \phi[Q_G(u; \zeta)]u \sin \left( \frac{\pi}{2}u \right) du \end{aligned}$$

and

$$\begin{aligned} I_\phi^{(2)} &= (1 - \lambda) \frac{\pi}{2} \int_{-\infty}^{+\infty} \phi(x)g(x; \zeta) \cos \left[ \frac{\pi}{2}G(x; \zeta) \right] dx \\ &= (1 - \lambda) \frac{\pi}{2} \int_0^1 \phi[Q_G(u; \zeta)] \cos \left( \frac{\pi}{2}u \right) du. \end{aligned}$$

These two integrals can be determined analytically, depending on the complexity of the function  $\phi[Q_G(u; \zeta)]$ . In all the situations, for given baseline cdf and  $\lambda$ ,  $\Theta_\phi(X)$  can be calculated by the means of numerical techniques, implemented in any mathematical software.

Also, for an alternative analytical treatment, Proposition 4 implies that

$$\Theta_\phi(X) = \sum_{k=0}^{+\infty} a_k \int_{-\infty}^{+\infty} \phi(x)v_{2k+1}(x; \zeta)dx. \tag{8}$$

For practical purposes, the sum can be truncated to a large enough integer  $K$ , providing a suitable approximation of  $\Theta_\phi(X)$ . Some derivations of  $\Theta_\phi(X)$  are presented in Table 1, which follow from several specific choices of  $\phi(x)$ . As an example of application, the  $m^{th}$  raw moments of a rv  $X$  following the TSW distribution can be derived from (8) and the  $m^{th}$  raw moments of the exponentiated Weibull distribution with power parameter  $2k + 1$  as established in [26].

**Table 1. Specific measures and functions derived to  $\Theta_\phi(X)$  according to the choice of  $\phi(x)$ .**

$\Theta_\phi(X)$	$\phi(x)$
mean ( $\mu_*$ )	$x$
variance	$(x - \mu_*)^2$
$m^{th}$ raw moment	$x^m$
$m^{th}$ central moment	$(x - \mu_*)^m$
$m^{th}$ inverse moment	$x^{-m}$
$m^{th}$ logarithmic moment	$[\log(x)]^m$
$m^{th}$ descending factorial moment	$x(x - 1)(x - 2) \dots (x - m + 1)$
$m^{th}$ incomplete moment with respect to $t$	$x^m$ if $x \leq t$ , and 0 elsewhere
$(m, q)^{th}$ probability weighted moment	$x^m F(x; \lambda, \zeta)^q$
moment generating function with respect to $t$	$e^{tx}$
characteristic function with respect to $t$	$e^{itx}$

<https://doi.org/10.1371/journal.pone.0250790.t001>

### 3.6 Reliability parameter

The general definition of the reliability parameter can be formulated as follows. Let  $X_1$  and  $X_2$  be two continuous rvs that can be compared based on a scenario that makes sense in a random system. Then, the corresponding reliability parameter can be defined as

$$R = P(X_2 < X_1) = \int \int_{\{y < x\}} f(x, y; \xi) dx dy,$$

where  $f(x, y; \xi)$  denotes the joint pdf of  $(X_1, X_2)$ , with  $\xi$  as parameter(s) vector. Details and applications of  $R$  in a concrete setting can be found in [27, 28], and the references therein.

The following result concerns the expression of  $R$  for the TS-G family in a specific setting.

**Proposition 5** *Let  $X_1$  and  $X_2$  be two independent rvs having the cdfs  $F(x; \lambda_1, \zeta)$  and  $F(x; \lambda_2, \zeta)$ , respectively. Then, we have*

$$R = \frac{1}{2} + \frac{1}{16} (\lambda_1 - \lambda_2) (\pi^2 - 4).$$

**Proof.** Owing to the independence of  $X_1$  and  $X_2$ , and (4) and (5), and after some integral calculus, we arrive at

$$\begin{aligned} R &= P(X_2 < X_1) = \int_{-\infty}^{+\infty} F(x; \lambda_2, \zeta) f(x; \lambda_1, \zeta) dx \\ &= \int_{-\infty}^{+\infty} \left\{ \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] - \lambda_2 \frac{\pi}{2} G(x; \zeta) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\} \times \\ &\quad \frac{\pi}{2} g(x; \zeta) \left\{ \lambda_1 \frac{\pi}{2} G(x; \zeta) \sin \left[ \frac{\pi}{2} G(x; \zeta) \right] + (1 - \lambda_1) \cos \left[ \frac{\pi}{2} G(x; \zeta) \right] \right\} dx \\ &= \frac{\pi}{2} \int_0^1 \left\{ \sin \left( \frac{\pi}{2} u \right) - \lambda_2 \frac{\pi}{2} u \cos \left( \frac{\pi}{2} u \right) \right\} \left\{ \lambda_1 \frac{\pi}{2} u \sin \left( \frac{\pi}{2} u \right) + (1 - \lambda_1) \cos \left( \frac{\pi}{2} u \right) \right\} du \\ &= \frac{1}{2} + \frac{1}{16} (\lambda_1 - \lambda_2) (\pi^2 - 4). \end{aligned}$$

This ends the proof of Proposition 5.

In Proposition 5, when  $X_1$  and  $X_2$  are identically distributed, i.e.,  $\lambda_1 = \lambda_2$ , we get  $R = 1/2$ . Also, Proposition 5 is useful to have a simple estimate of  $R$  based on estimates of  $\lambda_1$  and  $\lambda_2$ . Indeed, if  $\hat{\lambda}_1$  and  $\hat{\lambda}_2$  are estimates of  $\lambda_1$  and  $\lambda_2$ , respectively, then the plugging approach suggests the following estimate for  $R$ :

$$\hat{R} = \frac{1}{2} + \frac{1}{16} (\hat{\lambda}_1 - \hat{\lambda}_2) (\pi^2 - 4).$$

However, more research into the application of this formula to real-world data is needed.

## 4 Maximum likelihood estimation

Here, an inferential study of the TS-G family is proposed, estimating the parameters of the TS-G model by the maximum likelihood method.

### 4.1 The basics

The maximum likelihood method is commonly employed in parametric estimation because of its overall simplicity and the theoretical guarantees ensuring strong convergence properties on the obtained estimates. In this regard, the reader will find everything in [29]. We may also refer to [30–32] for modern applications of this method. In the context of the TS-G family, the

theoretical basics of the maximum likelihood method are described below. Let  $x_1, \dots, x_n$  be  $n$  observed values of a rv having the cdf given by (4). Then, the log-likelihood function for the parameters  $\lambda$  and  $\zeta$ , supposed to be unknown, is defined as

$$\ell(\lambda, \zeta) = \sum_{i=1}^n \log [f(x_i; \lambda, \zeta)] = n \log \left( \frac{\pi}{2} \right) + \sum_{i=1}^n \log [g(x_i; \zeta)] + \sum_{i=1}^n \log \left\{ \lambda \frac{\pi}{2} G(x_i; \zeta) \sin \left[ \frac{\pi}{2} G(x_i; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x_i; \zeta) \right] \right\}.$$

Then, the maximum likelihood method suggests the estimates given by  $\hat{\lambda}$  and  $\hat{\zeta}$ , which is possibly a vector of estimates, making  $\ell(\hat{\lambda}, \hat{\zeta})$  maximal, among all the possible values for  $\lambda$  and  $\zeta$ . They are called maximum likelihood estimates (MLEs). From the ideal mathematical point of view, they are the solutions the following system of equations:  $\partial \ell(\lambda, \zeta) / \partial \lambda = 0$  and  $\partial \ell(\lambda, \zeta) / \partial \zeta = 0$ , with

$$\frac{\partial \ell(\lambda, \zeta)}{\partial \lambda} = \sum_{i=1}^n \frac{\frac{\pi}{2} G(x_i; \zeta) \sin \left[ \frac{\pi}{2} G(x_i; \zeta) \right] - \cos \left[ \frac{\pi}{2} G(x_i; \zeta) \right]}{\lambda \frac{\pi}{2} G(x_i; \zeta) \sin \left[ \frac{\pi}{2} G(x_i; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x_i; \zeta) \right]}$$

and

$$\frac{\partial \ell(\lambda, \zeta)}{\partial \zeta} = \sum_{i=1}^n \frac{\partial g(x_i; \zeta)}{\partial \zeta} \frac{1}{g(x_i; \zeta)} + \frac{\pi}{2} \sum_{i=1}^n \frac{\partial g(x_i; \zeta)}{\partial \zeta} \frac{(2\lambda - 1) \sin \left[ \frac{\pi}{2} G(x_i; \zeta) \right] + \frac{\pi}{2} \lambda G(x_i; \zeta) \cos \left[ \frac{\pi}{2} G(x_i; \zeta) \right]}{\lambda \frac{\pi}{2} G(x_i; \zeta) \sin \left[ \frac{\pi}{2} G(x_i; \zeta) \right] + (1 - \lambda) \cos \left[ \frac{\pi}{2} G(x_i; \zeta) \right]}.$$

In most of the cases, the analytical expressions for  $\hat{\lambda}$  and  $\hat{\zeta}$  seem not possible. However, for given baseline cdf and  $\lambda$ , they can be approximated numerically by iterative techniques. Common routines are the `optim` function of the R software or `PROC NLMIXED` of the SAS (Statistical Analysis System) software. Also, one can determine the standard errors (SEs) of the MLEs which follow from the inverse of the observed information matrix. By assuming that  $\zeta$  contains several parameters, say  $m$ , this observed information matrix is defined by

$J = \{-\partial^2 \ell(\hat{\lambda}, \hat{\zeta}) / \partial \psi_i \partial \psi_j\}_{i,j=1, \dots, m+1}$ , where  $\psi_1 = \lambda$ , and  $\psi_{1+r}$  denotes the  $r^{th}$  component of  $\zeta$ . From the SEs, one can construct asymptotic confidence intervals of the parameters, among others.

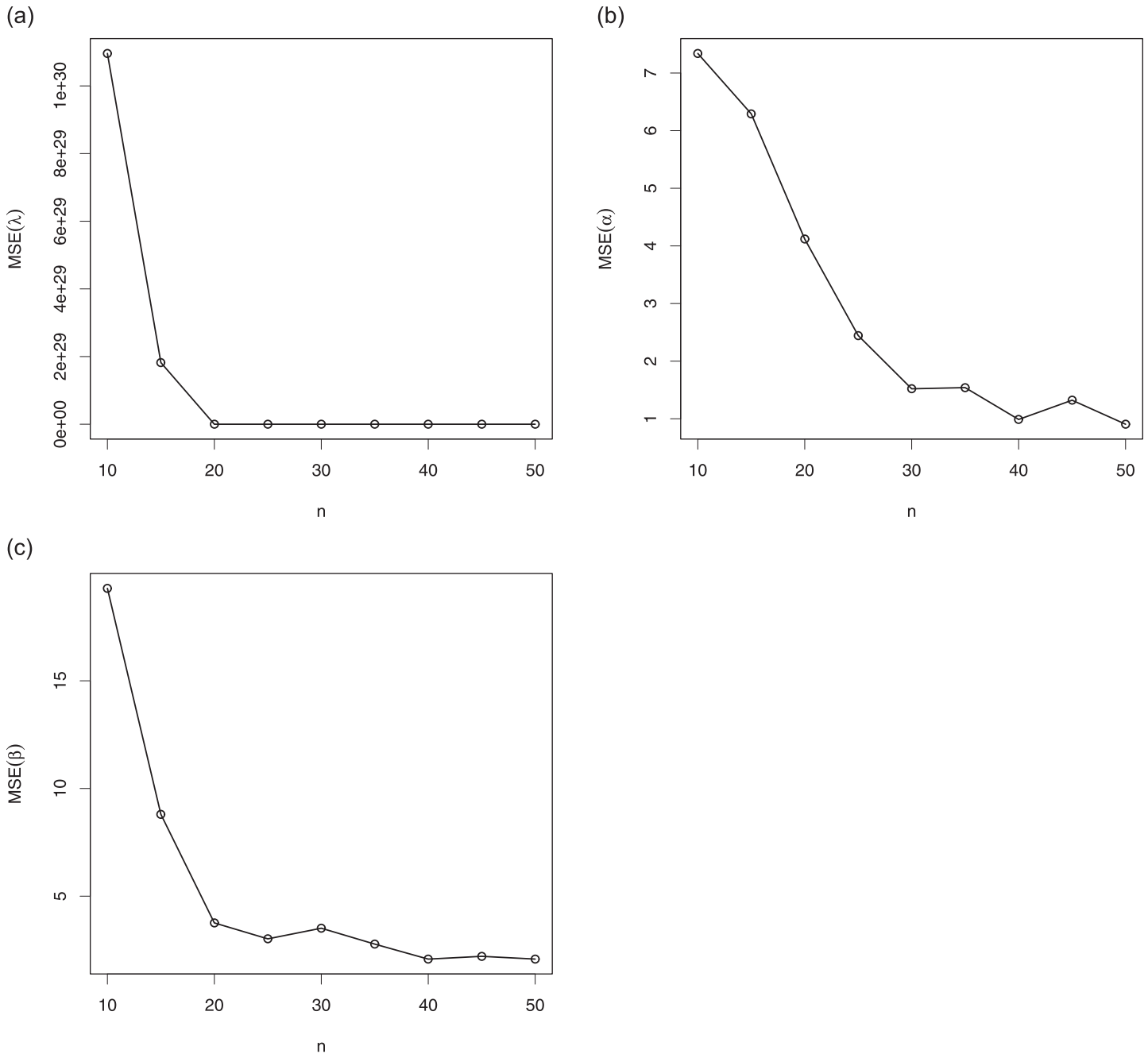
A statistical aspect of the TS-G family that is not investigated in this study is the identifiability. Numerical experiments show no particular problem on this property, but the rigorous theory remains to be developed with precise mathematical tools.

The rest of the study is devoted to the empirical and real life applications of the TS-G model with the consideration of the MLEs of the parameters.

### 4.2 Simulation

Here, we illustrate the practical aspect of the MLEs in the setting of the TSW model, i.e., based on the TSW distribution presented in Subsection 2.3. More precisely, we propose a graphical simulation approach illustrating the numerical behavior of the MLEs  $\hat{\lambda}$ ,  $\hat{\alpha}$  and  $\hat{\beta}$ , of  $\lambda$ ,  $\alpha$  and  $\beta$ , respectively. The R software is used in this regard.

We proceed as follows. We generate  $N = 3000$  samples  $(x_1, \dots, x_n)$  of size  $n = 10$  to 50 from a rv following the TSW distribution with the two following sets of parameters:  $\mathcal{S}_1 : (\lambda =$



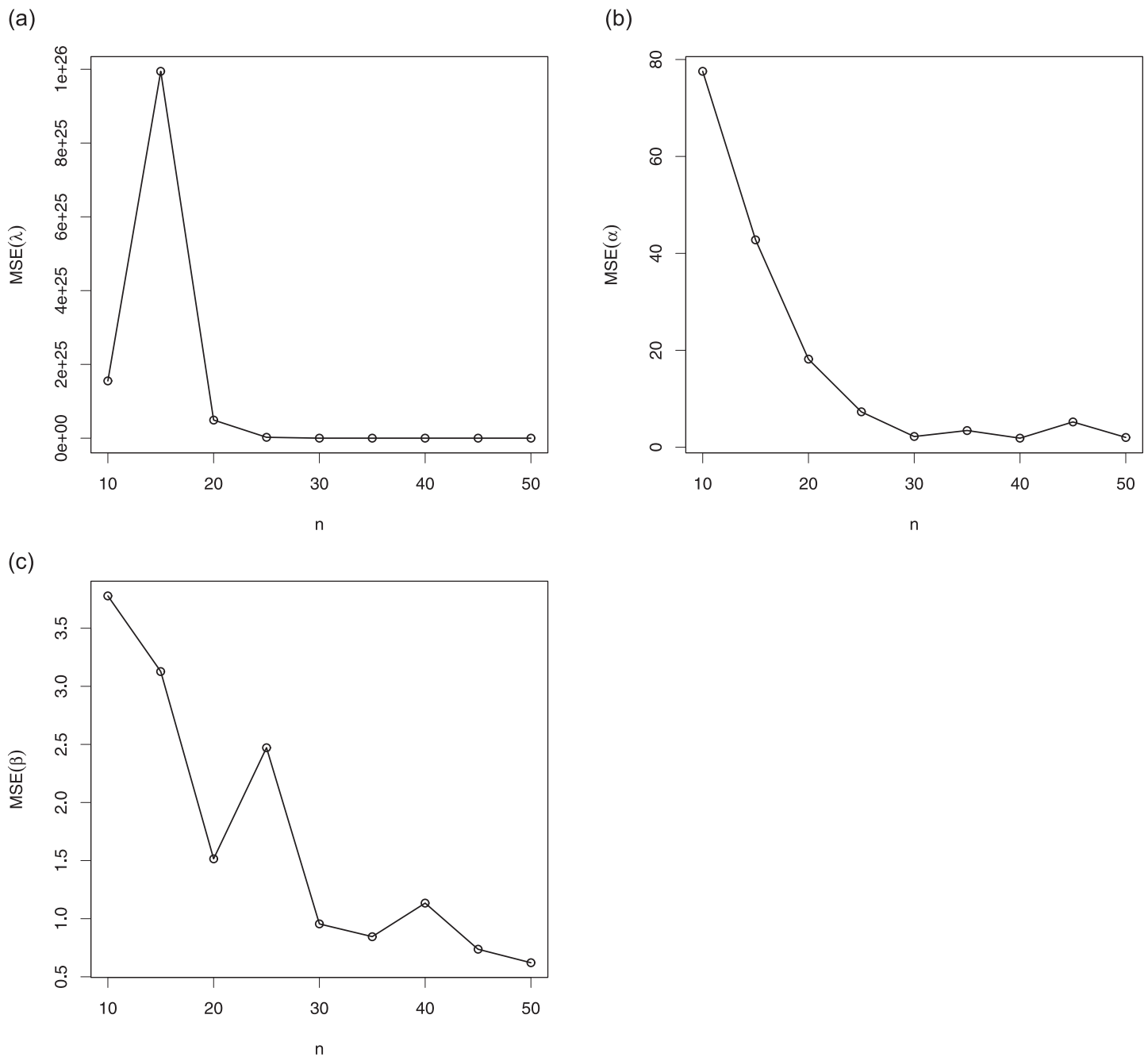
**Fig 4. Plots of the empirical MSEs of the TSW model parameters for  $\mathcal{S}_1$  : ( $\lambda = 0.3, \alpha = 3, \beta = 5$ ) for (a)  $\lambda$ , (b)  $\alpha$  and (c)  $\beta$ .**

<https://doi.org/10.1371/journal.pone.0250790.g004>

$0.3, \alpha = 3, \beta = 5$ ) and  $\mathcal{S}_2$  : ( $\lambda = 0.1, \alpha = 3.5, \beta = 4.5$ ). We also calculate the empirical mean squared errors (MSEs) of the MLEs defined as, for  $h = \lambda, \alpha, \beta$ ,

$$\widehat{MSE}_h = \frac{1}{N} \sum_{i=1}^N (\hat{h}_i - h)^2,$$

where the index  $i$  refers to the  $i^{th}$  generated samples. The results of this simulation study are presented in Figs 4 and 5 for  $\mathcal{S}_1$  and  $\mathcal{S}_2$ , respectively.



**Fig 5. Plots of the empirical MSEs of the TSW model parameters for  $\mathcal{S}_2$  : ( $\lambda = 0.1, \alpha = 3.5, \beta = 4.5$ ) for (a)  $\lambda$ , (b)  $\alpha$  and (c)  $\beta$ .**

<https://doi.org/10.1371/journal.pone.0250790.g005>

As a prime observation, we see that, in all the situations, when the sample size increases, the empirical MSEs approach the axis  $y = 0$ . This illustrates the “numerical convergence” of the MLEs to the true values of the parameters.

### 5 Applications

Thanks to its desirable flexible properties, the TSW model aims to be applied in concrete scenarios, such as the fit of real life data. We share this finding by considered the two following well-referenced real life data sets.

“The first data set”. The first data set finds its source in [28]. It contains the tensile strength (with unit in GPa) for single carbon fibers. This data set is given by: {0.312, 0.314, 0.479, 0.552, 0.700, 0.803, 0.861, 0.865, 0.944, 0.958, 0.966, 0.997, 1.006, 1.021, 1.027, 1.055, 1.063, 1.098, 1.140, 1.179, 1.224, 1.240, 1.253, 1.270, 1.272, 1.274, 1.301, 1.301, 1.359, 1.382, 1.382, 1.426, 1.434, 1.435, 1.478, 1.490, 1.511, 1.514, 1.535, 1.554, 1.566, 1.570, 1.586, 1.629, 1.633, 1.642, 1.648, 1.684, 1.697, 1.726, 1.770, 1.773, 1.800, 1.809, 1.818, 1.821, 1.848, 1.880, 1.954, 2.012, 2.067, 2.084, 2.090, 2.096, 2.128, 2.233, 2.433, 2.585, 2.585}.

“The second data set”. The second data set, often called breaking stress of carbon fibers data set, was used by [33]. This data set is given by: {3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 3.56, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 1.57, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.89, 2.88, 2.82, 2.05, 3.65, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.35, 2.55, 2.59, 2.03, 1.61, 2.12, 3.15, 1.08, 2.56, 1.80, 2.53}.

In addition, seven successful models are considered for comparison, also defined as extended or modified versions of the Weibull model and having two, or three or four tuning parameters. Namely, we consider the four-parameter generalized modified Weibull (GMW) model by [34], four-parameter Kumaraswamy Weibull (KW) model by [35], four-parameter beta Weibull (BW) model by [6], four-parameter odd log-logistic modified Weibull (OLLMW) model by [36], three-parameter transmuted Weibull (TW) model by [37], three-parameter modified Weibull (MW) model by [38] and the former two-parameter sine Weibull (SW) model by [3].

As criteria of goodness-of-fits to compare these models, we chose the Cramér-Von Mises (CVM), Anderson-Darling (AD) and KS statistics, with the corresponding KS p-values. Also, the AIC is calculated. For the use of the AIC in applied frameworks, one may refer to [39–41]. The global rule is the following ones. The smaller the values of the CVM, AD, KS statistics and AIC, and the larger the values of the KS p-values, the better the fit of the corresponding model to the considered data. The R software is used.

**Table 2. CVM, AD, KS with p-value, MLEs and SEs for the first data set.**

Model	CVM	AD	KS	p-value			MLEs (SEs)	
TSW	0.0152	0.1322	0.0385	1.0000	0.7896	0.4807	2.4826	-
$(\lambda, \alpha, \beta)$					(0.1687)	(0.1702)	(0.4224)	-
GMW	0.0184	0.1649	0.0421	0.9960	4.5031	0.4927	0.3401	0.8561
$(a, \alpha, \gamma, \lambda)$					(9.1201)	(0.7698)	(1.0690)	(0.2833)
KW	0.0226	0.1984	0.0475	0.9977	0.7268	0.1621	1.0308	3.5369
$(\alpha, \beta, \gamma, \theta)$					(0.0052)	(0.0186)	(0.0218)	(0.0086)
BW	0.0256	0.2217	0.0480	0.9973	0.3585	3.7827	0.7813	5.7953
$(\alpha, \beta, \gamma, \theta)$					(1.9772)	(1.2906)	(0.4103)	(18.9342)
OLLW	0.0228	0.1664	0.04675	0.9982	0.0729	0.5845	0.0146	22.3637
$(\alpha, \beta, \gamma, \theta)$					(0.1025)	(0.1487)	(0.0384)	(29.5884)
TW	0.0428	0.3266	0.3145	0.0000	2.7732	1.4508	-0.5636	-
$(\alpha, \beta, \lambda)$					(0.4919)	(0.1420)	(0.4269)	-
MW	0.0195	0.1733	0.0431	0.9995	0.0180	0.1892	3.3740	-
$(\alpha, \beta, \theta)$					(0.0609)	(0.0780)	(0.5227)	-
SW	0.0236	0.2076	0.0442	0.9902	0.1291	3.0852	-	-
$(\alpha, \beta)$					(0.0275)	(0.2951)	-	-

<https://doi.org/10.1371/journal.pone.0250790.t002>



**Table 3. CVM, AD, KS with p-value, MLEs and SEs for the second data set.**

Model	CVM	AD	KS	p-value			MLEs (SEs)	
TSW	0.0554	0.3538	0.0694	0.9079	0.7440	0.0679	2.7694	-
( $\lambda, \alpha, \beta$ )					(0.1915)	(0.0504)	(0.5096)	-
GMW	0.0653	0.3939	0.0760	0.8394	5.4737	0.4343	0.1493	0.5167
( $a, \alpha, \gamma, \lambda$ )					(7.9525)	(0.6457)	(0.5395)	(0.1722)
KW	0.0703	0.4501	0.0825	0.7591	0.6536	0.1738	0.0664	3.8782
( $\alpha, \beta, \gamma, \theta$ )					(0.0230)	(0.0416)	(0.0142)	(0.0171)
BW	0.0846	0.5041	0.0812	0.7761	0.1864	4.0715	0.7592	6.9449
( $\alpha, \beta, \gamma, \theta$ )					(0.4201)	(1.2708)	(0.3673)	(2.7517)
OLLW	0.1032	0.54558	0.0780	0.8160	0.0729	0.5845	0.0146	22.3637
( $\alpha, \beta, \gamma, \theta$ )					(0.0301)	(0.0828)	(0.0256)	(23.3981)
TW	0.1260	0.6700	0.3669	0.0000	2.9256	2.7531	-0.5906	-
( $\alpha, \beta, \lambda$ )					(0.4828)	(0.2294)	(0.3744)	-
MW	0.0640	0.40134	0.0795	0.7974	0.0165	0.0144	3.7146	-
( $\alpha, \beta, \theta$ )					(0.0206)	(0.0079)	(0.4069)	-
SW	0.0863	0.4937	0.09003	0.6584	0.0165	3.1733	-	-
( $\alpha, \beta$ )					(0.0064)	(0.3016)	-	-

<https://doi.org/10.1371/journal.pone.0250790.t003>

Tables 2 and 3 list the values of the CVM, AD, KS with p-value, and the MLEs and their corresponding SEs of the models parameters for the first and second data sets, respectively.

Tables 2 and 3 indicate that the smallest CVM, AD and KS and the largest KS p-value are for the TSW model; it is the best model with the considered criteria. In particular, it

**Table 4. The  $-\hat{\ell}$  and AIC for the first data set.**

Distribution	$-\hat{\ell}$	AIC
TSW	48.4389	102.8779
GMW	48.7195	106.0641
KW	48.7684	105.5368
BW	48.8954	105.7908
OLLW	49.3799	106.7600
TW	48.7059	103.4118
MW	48.9583	103.9166
SW	49.5012	103.0024

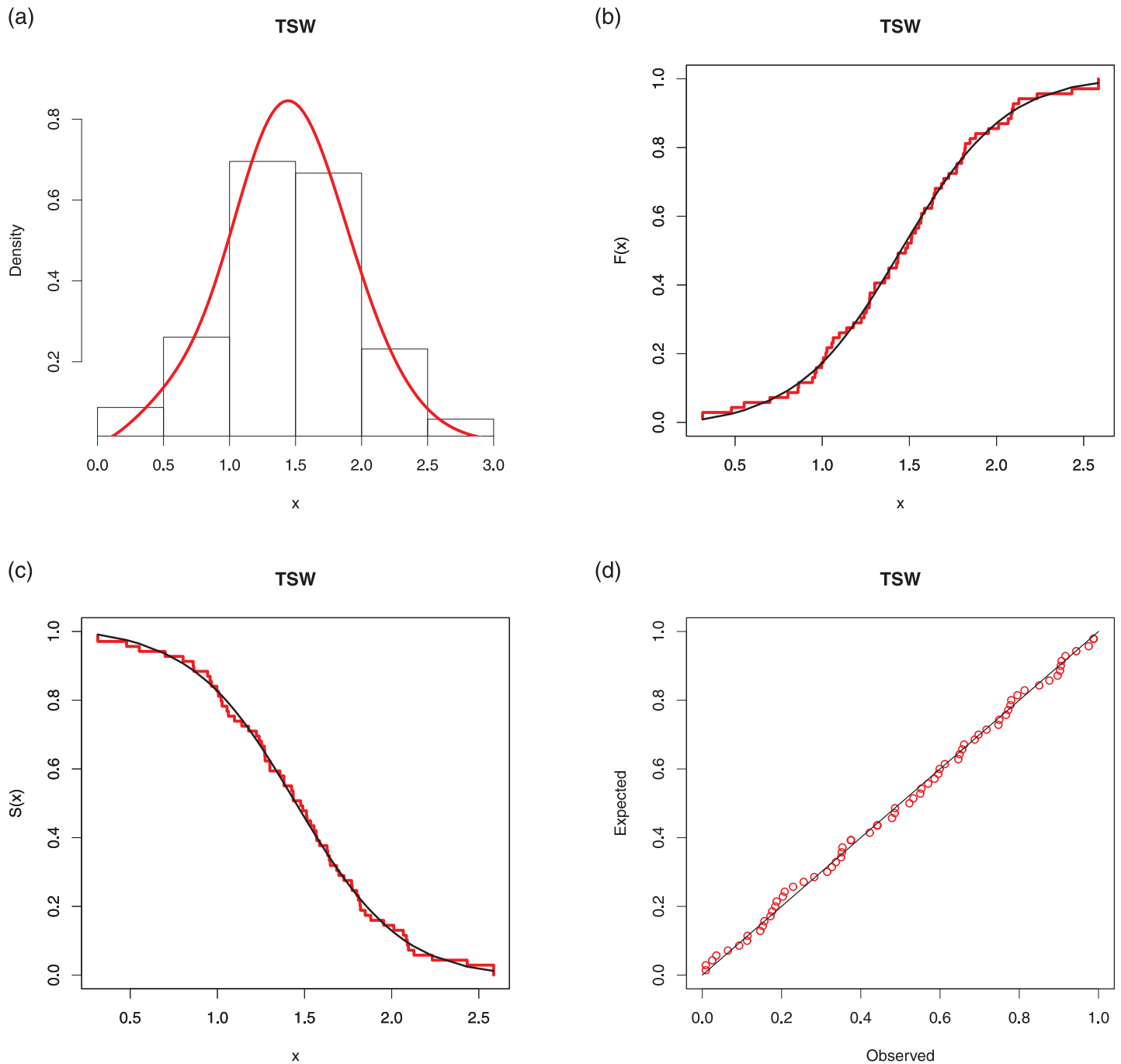
<https://doi.org/10.1371/journal.pone.0250790.t004>

**Table 5. The  $-\hat{\ell}$  and AIC for the second data set.**

Distribution	$-\hat{\ell}$	AIC
TSW	85.1358	176.2717
GMW	85.3731	178.7462
KW	85.60939	179.2188
BW	85.9184	179.8368
OLLW	85.5593	179.1187
TW	85.5453	177.0907
MW	85.52304	177.0461
SW	86.6910	177.3820

<https://doi.org/10.1371/journal.pone.0250790.t005>

outperforms the former SW model corresponding to  $\lambda = 0$ . That is, we see that the parameter  $\lambda$  of the TSW model is estimated “far from zero”, i.e., the corresponding MLEs are  $\hat{\lambda} = 0.7896$  and  $\hat{\lambda} = 0.7440$ , for the first and second data set, respectively. This points out the importance of the transformed sine technique to obtain suitable fits of these data, in comparison to the former SW model.



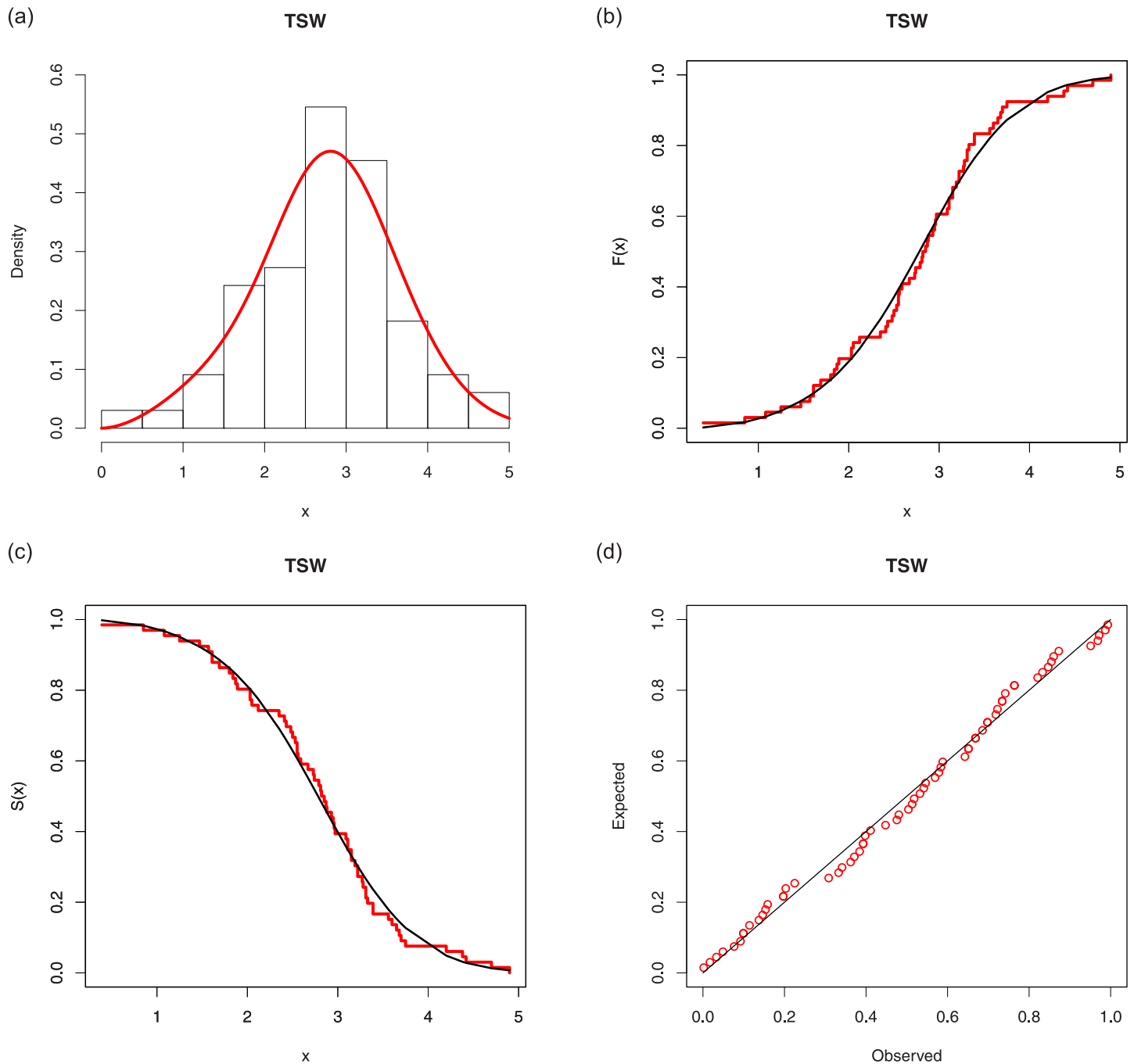
**Fig 6. Several fits of the TSW model for the first data set: (a) estimated pdf over the histogram, (b) estimated cdf over the empirical cdf, (c) estimated sf over the empirical sf and (d) P-P plot.**

<https://doi.org/10.1371/journal.pone.0250790.g006>

Tables 4 and 5 present the minus estimated log-likelihood, i.e.,  $-\hat{\ell} = -\ell(\hat{\lambda}, \hat{\alpha}, \hat{\beta})$  for the TSW model, and AIC values of the model parameters for the first and second data set, respectively.

According to Tables 4 and 5, since it has the lowest AIC for the two data sets, the TSW model can be considered as the best one.

We now provide a graphical visualization of the nice fitting results of the TSW model. That is, Figs 6 and 7 display several fits of the TSW model. In particular, the histograms of



**Fig 7. Several fits for the TSW model for the second data set: (a) estimated pdf over the histogram, (b) estimated cdf over the empirical cdf, (c) estimated sf over the empirical sf and (d) P-P plot.**

<https://doi.org/10.1371/journal.pone.0250790.g007>

the both data sets are plotted, along with the curves of the corresponding estimated pdfs, i.e.,  $f(x; \hat{\lambda}, \hat{\alpha}, \hat{\beta})$ , the curves of the estimated cdfs, i.e.,  $F(x; \hat{\lambda}, \hat{\alpha}, \hat{\beta})$ , are plotted over the ones of the corresponding empirical cdfs of the data, the curves of the estimated sfs, i.e.,  $S(x; \hat{\lambda}, \hat{\alpha}, \hat{\beta})$ , are plotted over the curves of corresponding empirical sfs of the data, and Probability-Probability (P-P) plots are provided.

In all the graphics, we see that the red curves fit well the corresponding black curves, attesting the efficiency of the TSW model in this data fitting exercise.

## 6 Concluding remarks

Based on a new one-parameter transformation function, we provide an original extension of the Sin-G family of continuous distributions, introducing the transformed Sin-G (TS-G) family. We discuss how an additional parameter  $\lambda$  can enhance the flexibility of the cdf of the former Sin-G family, with nice consequences for modelling purposes. An emphasis is put on the transformed Sin Weibull (TSW) distribution, showing a high potential in the analysis and modelling of lifetime data. Some general mathematical features of the TS-G family are established. Then, a statistical approach is adopted; the maximum likelihood estimates (MLEs) for the TS-G model parameters are discussed. The TSW model is highlighted, demonstrating that it is more capable of fitting data than seven rival models, some of which have more parameters. The TS-G family can find a broader use in all areas dealing with modern data as a result of its qualities. For example, it can be used to construct models in multivariate analysis, regression, classification, and other statistical fields of importance. In addition, the transformation  $T_\lambda(x)$  or  $T_\lambda^*(x)$  can be used to efficiently extend other existing families of distributions. These viewpoints necessitate additional developments, which we plan to incorporate in future works.

## Acknowledgments

The authors thank the two reviewers for their detailed and constructive comments.

## Author Contributions

**Conceptualization:** Christophe Chesneau, Mahmood Ul Hassan.

**Formal analysis:** Farrukh Jamal, Dalal Lala Bouali.

**Investigation:** Farrukh Jamal, Christophe Chesneau, Dalal Lala Bouali.

**Methodology:** Mahmood Ul Hassan.

**Writing – original draft:** Farrukh Jamal, Christophe Chesneau, Mahmood Ul Hassan.

**Writing – review & editing:** Dalal Lala Bouali.

## References

1. Brito CR, Rêgo LC, Oliveira WR and Gomes-Silva F. Method for generating distributions and classes of probability distributions: The univariate case, Hacettepe Journal of Mathematics and Statistics. 2019, 48, 897–930.
2. Kumar D, Singh U and Singh SK. A new distribution using sine function: its application to bladder cancer patients data, Journal of Statistics Applications and Probability. 2015, 4, 417–427.
3. Souza L. New trigonometric classes of probabilistic distributions, Thesis, Universidade Federal Rural de Pernambuco, 2015.
4. Souza L, Junior WRO, de Brito CCR, Chesneau C, Ferreira TAE and Soares L. On the Sin-G class of distributions: theory, model and application, Journal of Mathematical Modeling. 2019, 7, 3, 357–379.
5. Souza L, Junior WRO, de Brito CCR, Chesneau C, Ferreira TAE and Soares L. General properties for the Cos-G class of distributions with applications, Eurasian Bulletin of Mathematics. 2019, 2, 2, 63–79.

6. Lee C, Famoye F and Olumolade O. Beta-Weibull distribution: Some properties and applications to censored data, *Journal of modern applied statistical methods*. 2007, 6, 1, 173–186. <https://doi.org/10.22237/jmasm/1177992960>
7. Nelson W. *Applied life data analysis*, John Wiley and Sons, New York, 1982.
8. Bjerkedal T. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli, *American Journal of Hygiene*. 1960, 72, 130–148. PMID: [13801076](https://pubmed.ncbi.nlm.nih.gov/13801076/)
9. Souza L, Gallindo L and Serafim-de-Souza L. SinIW: The SinIW distribution. R package version 0.2. 2016, Available at <https://CRAN.R-project.org/package=SinIW>.
10. Chesneau C, Bakouch HS and Hussain T. A new class of probability distributions via cosine and sine functions with applications, *Communications in Statistics—Simulation and Computation*. 2019, 48, 8, 2287–2300. <https://doi.org/10.1080/03610918.2018.1440303>
11. Mahmood Z, Chesneau C and Tahir MH. A new sine-G family of distributions: properties and applications, *Bulletin of Computational Applied Mathematics*. 2019, 7, 1, 53–81.
12. Jamal F and Chesneau C. A new family of polyno-expo-trigonometric distributions with applications, *Infinite Dimensional Analysis, Quantum Probability and Related Topics*. 2019, 22, 04, 1950027, 1–15. <https://doi.org/10.1142/S0219025719500279>
13. Al-Babtain AA, Elbatal I, Chesneau C and Elgarhy M. Sine Topp-Leone-G family of distributions: Theory and applications, *Open Physics*. 2020, 18, 1, 574–593. <https://doi.org/10.1515/phys-2020-0180>
14. Jamal F and Chesneau C. The sine Kumaraswamy-G family of distributions, 2020, preprint: <https://hal.archives-ouvertes.fr/hal-02120197>.
15. Jamal F, Chesneau C and Aidi K. The sine extended odd Fréchet-G family of distribution with applications to complete and censored data, *Mathematica Slovaca*. (2021), (to appear).
16. Mitrinović DS. *Analytic Inequalities*, Springer-Verlag, Berlin, 1970.
17. Klein JP and Moeschberger ML. *Survival analysis: Techniques for censored and truncated data*, 2nd ed.; Springer: Berlin, Germany, 2003.
18. Wang Y, Chan Y, Gui Z, Webb D and Li L. Application of Weibull distribution analysis to the dielectric failure of multilayer ceramic capacitors, *Materials Science and Engineering: B*. 1997, 47, 3, 197–203. [https://doi.org/10.1016/S0921-5107\(97\)00041-X](https://doi.org/10.1016/S0921-5107(97)00041-X)
19. Corzo O and Bracho N. Application of Weibull distribution model to describe the vacuum pulse osmotic dehydration of sardine sheets, *LWT-Food Science and Technology*. 2008, 41, 6, 1108–1115. <https://doi.org/10.1016/j.lwt.2007.06.018>
20. Aslam M, Arif OH and Jun CH. An attribute control chart for a Weibull distribution under accelerated hybrid censoring, *PLoS one*. 2017, 12, 3, e0173406. <https://doi.org/10.1371/journal.pone.0173406> PMID: [28257479](https://pubmed.ncbi.nlm.nih.gov/28257479/)
21. Shaked M and Shanthikumar JG. *Stochastic orders*, Wiley, New York, 2007.
22. Cordeiro GM and de Castro M. A new family of generalized distributions, *Journal of Statistical Computation and Simulation*. 2011, 81, 883–893. <https://doi.org/10.1080/00949650903530745>
23. Mudholkar GS and Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*. 1993, 42, 2, 299–302. <https://doi.org/10.1109/24.229504>
24. Mudholkar GS, Srivastava DK and Freimer M. The exponentiated Weibull family; a reanalysis of the bus motor failure data, *Technometrics*. 1995, 37, 4, 436–445. <https://doi.org/10.1080/00401706.1995.10484376>
25. Mudholkar GS and Hutson AD. The exponentiated Weibull family: some properties and a flood data application, *Communications in Statistics—Theory and Methods*. 1996, 25, 3059–3083. <https://doi.org/10.1080/03610929608831886>
26. Choudhury A. A simple derivation of moments of the exponentiated Weibull distribution, *Metrika*. 2005, 62, 1, 17–22. <https://doi.org/10.1007/s001840400351>
27. Kotz S, Lumelskii Y and Penskey M. *The stress-strength model and its generalizations and applications*, World Scientific, Singapore, 2003.
28. Raqab M, Madi T and Debasis K. Estimation of P (Y < X) for the 3-parameter generalized exponential distribution, *Communications in Statistics-Theory and Methods*. 2008, 37, 2854–2864. <https://doi.org/10.1080/03610920802162664>
29. Casella G and Berger RL. *Statistical inference*, Duxbury Advanced Series Thomson Learning: Pacific Grove, 2002.
30. Dong B, Ma X, Chen F and Chen S. Investigating the differences of single- and multi-vehicle accident probability using mixed logit model. *Journal of Advanced Transportation*. 2018, Article ID 2702360, 9 pages.

31. Chen F and Chen S. Injury severities of truck drivers in single- and multi-vehicle accidents on rural highway. *Accident Analysis and Prevention*. 2011, 43, 1677–1688. <https://doi.org/10.1016/j.aap.2011.03.026> PMID: 21658494
32. Chen F, Chen S and Ma X. Analysis of hourly crash likelihood using unbalanced panel data mixed logit model and real-time driving environmental big data. *Journal of Safety Research*. 2018, 65, 153–159. <https://doi.org/10.1016/j.jsr.2018.02.010> PMID: 29776524
33. Cordeiro GM and Lemonte A. The  $\beta$ -Birnbbaum-Saunders distribution: An improved distribution for fatigue life modeling, *Computational Statistics and Data Analysis*. 2011, 55, 1445–1461. <https://doi.org/10.1016/j.csda.2010.10.007>
34. Carrasco JMF, Edwin Ortega MM and Cordeiro GM A generalized modified Weibull distribution for life-time modeling, *Computational Statistics and Data Analysis*. 2008, 53, 450–462. <https://doi.org/10.1016/j.csda.2008.08.023>
35. Cordeiro GM, Ortega EMM and Nadarajah S. The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*. 2010, 347, 8, 1399–1429. <https://doi.org/10.1016/j.jfranklin.2010.06.010>
36. Saboor A, Alizadeh M, Khan MN, Gosh I and Cordeiro GM. Odd Log-Logistic modified Weibull distribution, *Mediterranean Journal of Mathematics*. 2017, 14, 96. <https://doi.org/10.1007/s00009-017-0880-3>
37. Aryal GR and Tsokos CP. Transmuted Weibull distribution: A generalization of the Weibull probability distribution, *European Journal of Pure and Applied Mathematics*. 2011, 4, 2, 89–102.
38. Sarhan AM and Zain-din M. Modified Weibull distribution, *Applied Sciences*. 2009, 11, 123–136.
39. Huang H, Song B, Xu P, Zeng Q, Lee J and Abdel-Aty M. Macro and micro models for zonal crash prediction with application in hot zones identification. *Journal of Transport Geography*. 2016, 54, 248–256. <https://doi.org/10.1016/j.jtrangeo.2016.06.012>
40. Wen H, Zhang X, Zeng Q and Sze NN. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. *Accident Analysis and Prevention*. 2019, 132, 105249. <https://doi.org/10.1016/j.aap.2019.07.025> PMID: 31415995
41. Zeng Q, Wang X, Wen H and Yuan Q. An empirical investigation of the factors contributing to local-vehicle and non-local-vehicle crashes on freeway. *Journal of Transportation Safety & Security*. 2020, 1–15. <https://doi.org/10.1080/19439962.2020.1779422> PMID: 27648455