symmetry

MDPI

Article

Estimation of $P(X < Y)$ Stress—Strength Reliability Measures for a Class of Asymmetric Distributions: The Case of Three-Parameter p-Max Stable Laws

Felipe Sousa Quintino [1], Pushpa Narayan Rathie [1], Luan Carlos de Sena Monteiro Ozelim [2,*] and Tiago Alves da Fonseca [3]

[1] Department of Statistics, University of Brasilia, Brasilia 70910-900, Brazil; felipe.quintino@unb.br (F.S.Q.); pushpanrathie@yahoo.com (P.N.R.)
[2] Department of Civil and Environmental Engineering, University of Brasilia, Brasilia 70910-900, Brazil
[3] Gama Engineering College, University of Brasilia, Brasilia 70910-900, Brazil; fonsecafga@unb.br
* Correspondence: luanoz@gmail.com or ozelim@unb.br

Abstract: Asymmetric distributions are frequently seen in real-world datasets due to a number of factors, such as sample biases and nonlinear interactions between the variables observed. Thus, in order to better characterize real-world phenomena, studying asymmetric distribution is of great interest. In this work, we derive stress–strength reliability formulas of the type $P(X < Y)$ when both X and Y follow p-max stable laws with three parameters, which are inherently asymmetric. The new relations are given in terms of extreme-value H-functions and have been obtained under fewer parameter restrictions when compared to similar results in the literature. We estimate the parameters of the p-max stable laws by a stochastic optimization method and the stress–strength probability by a maximum likelihood procedure. The performance of the analytical models is evaluated through simulations and real-life dataset modeling.

Keywords: stress–strength reliability; $\mathcal{H}$-function; p-max stable laws

Citation: Quintino, F.S.; Rathie, P.N.; Ozelim, L.C.d.M.; da Fonseca, T.A. Estimation of $P(X < Y)$ Stress—Strength Reliability Measures for a Class of Asymmetric Distributions: The Case of Three-Parameter p-Max Stable Laws. *Symmetry* **2024**, *16*, 837. https://doi.org/10.3390/sym16070837

Academic Editor: Chong Wang

Received: 30 April 2024
Revised: 25 May 2024
Accepted: 30 May 2024
Published: 3 July 2024

1. Introduction

Reliability measures of the type $P(X < Y)$, often represented as R and referred to as stress–strength reliability, are important for evaluating the performance of different systems and processes. This measure indicates the probability that a random variable X, which can represent a general performance metric or quality indicator, is less than another random variable Y, which could signify a threshold or standard to be met. In this context, X and Y are not limited to engineering concepts like stress and strength but are applicable in any scenario where two quantities are compared. A higher R value signifies a more reliable system, indicating a greater likelihood that the performance metric X will be below the threshold Y. Calculating $P(X < Y)$ requires an understanding of the joint distribution of X and Y, which can be determined through various methods, including simulation, analytical solutions, or the use of copulas to model dependencies between variables. We refer the reader to [1] for further details on this subject.

Let Y and X be independent continuous random variables from probability density function (PDF) f_Y and cumulative distribution function (CDF) F_X, respectively. We can write the stress–strength reliability measure as:

$$R = P(X < Y) = \int_{-\infty}^{\infty} F_X(x) f_Y(x) dx. \quad (1)$$

Thus, R is a measure of component reliability, and it may be interpreted as the probability of a system failure when the applied stress Y is greater than its strength X. It is often assumed that X and Y are independent random variables and that they belong to the same family of probability distributions. Rathie et al. [2] present a recent survey on the subject.

Studying reliability measures such as $P(X < Y)$ for asymmetric marginal distributions is crucial for understanding a variety of real-world scenarios that require data-driven solutions. In the case of finance applications, where risk assessment is critical, asymmetric distributions play a central role, especially heavy tailed ones [3]. For example, in stock market analysis, knowing the likelihood that a particular stock will do better than another is crucial to making wise investment choices. Investors can efficiently manage risk and optimize their portfolios with the aid of reliability metrics, which assist in quantifying these probabilities [4].

Furthermore, asymmetry in distributions is common in a wide range of social and ecological phenomena, including the spread of illnesses and the distribution of money [5]. Asymmetric distributions can be used to describe the different scenarios of disease transmission within populations in epidemiology [6]. Researchers can better understand the likelihood of particular outcomes and aid in the creation of tailored intervention methods by examining reliability measurements in such circumstances. In short, studying reliability measures of the type $P(X < Y)$ for asymmetric distributions makes modeling and prediction more precise, which in turn helps one to make more informed decisions across a variety of domains.

In particular, reliability measures of the stress–strength type for classic extreme value distributions were studied by [7], who derived expressions for R in terms of special functions for l-max stable laws (Fréchet, Weibull, and Gumbel). Several authors have worked on the estimation and application of stress–strength for the l-max stable distributions (e.g., [8–11]). Some generalizations of l-max stable distributions have been proposed to either allow better data fitting or provide more convenient mathematical properties. In the work of Aryal and Tsokos [12], for example, the generalized extreme value distribution (GEV) was extended to a model named transmuted GEV (TGEV). Bivariate data were also considered like bimodal Weibull [13], bimodal Gumbel [14], and bimodal GEV (BGEV) [15] distributions.

The $l-$max stable distributions are derived as a limiting distribution of linearly normalized partial maxima. Another approach to generalize such distributions is by non-linearly normalizing partial maxima of independent identically distributed random variables (iid RVs). This way, for a given CDF $F(\cdot)$, suppose there exists sequences of real numbers $\{\gamma_n\}$ and $\{\beta_n\}$ with $\gamma_n, \beta_n > 0$ such that

$$\lim_{n\to\infty} F^n\left(\gamma_n|x|^{\beta_n} sign(x)\right) = H(x), \tag{2}$$

weakly, where $H(\cdot)$ is a non-degenerate CDF. The three-parameter p-max stable laws can be obtained from CDF (2) by the definition of the same p-type. That means that we assume there exist positive constants $\alpha, \beta, and\ \gamma$ such that

$$H_i(x;\alpha,\beta,\gamma) = H_i(\gamma|x|^{\beta}\text{sign}(x);\alpha), \ \ i = 1,\cdots,6.$$

where $\alpha = 1$ for $i = 5,6$. It was shown in [16] that H is of the same p-type as one of the following distributions: log-Fréchet, log-Weibull, inverse log-Fréchet, inverse log-Weibull, standard Fréchet, and standard Weibull. Such limiting distributions are heavy tailed and asymmetric. Therefore, the convergence in (2) is usually studied by assessing the approximation on the tails, as discussed, for example, by Feng and Chen [17] and references therein.

To the best of our knowledge, the literature lacks previous in-depth studies on reliability inference for p-max stable distributions, and this work stands as a contribution by providing estimation methods for R based on stochastic optimization for this class of distributions. Thus, in this paper, we consider the problem of estimating the stress–strength parameter R when X and Y are independent three-parameter p-max stable random variables with the same CDF but different parameters. In order to validate our results, a robust framework was proposed and applied to model real and synthetic datasets, rigorously indicating the capacities of the p-max models and the usability of the analytical formulas hereby derived to calculate R.

Our main contributions are as follows:

1. to analytically derive R in terms of special functions, for each three-parameter p-max stable law with fewer parameter restrictions compared to previous results in the literature;
2. to propose an estimator for R;
3. to apply the results to the modeling of real datasets. In particular, two real scenarios are investigated, showing the versatility of stress–strength reliability (SSR) modeling approaches using p-max models. First, soccer pass completion proportions of two different championships (UEFA Champions League and 2022 FIFA World Cup) were compared, allowing scouting professionals to use the SSR results as a proxy for technical level comparison of teams that competed at those tournaments. Then, a second application involved the modeling and comparison of the strength of carbon fibers of different lengths when subjected to tension efforts. In both modeling scenarios, the best fitting p-max stable distribution (both qualitatively (by graphical methods) and quantitatively (by information criteria)) was taken as a starting point.

This paper is organized as follows: Section 2 introduces preliminaries, especially the definition of the $\mathcal{H}$-function, the $\mathbb{H}$-function, and the three-parameter p-max stable laws. Section 3, on the other hand, deals with the derivation of R when X and Y are independent p-max stable random variables. The maximum likelihood estimation for R is presented in Section 4. In Section 5, we deal with Monte Carlo simulations as well as with the modeling of two real situations involving football datasets and different-length carbon fibers. The Section 6 deals with conclusions.

2. Preliminaries

In this section, we give some definitions and results, which will be used subsequently.

2.1. Special Functions

The $\mathcal{H}$-function is defined by

$$\mathcal{H}^{m,n}_{p,q}\left[z\middle| \begin{array}{c}(a_1,A_1),\cdots,(a_p,A_p)\\(b_1,B_1),\cdots,(b_q,B_q)\end{array}\right] = \frac{1}{2\pi i}\int_L \frac{\prod_{k=1}^m \Gamma(b_j+B_js)\prod_{j=1}^n\Gamma(1-a_j-A_js)}{\prod_{k=m+1}^q\Gamma(1-b_j-B_js)\prod_{j=n+1}^p\Gamma(a_j+A_js)} z^{-s}ds,$$

where $i=\sqrt{-1}$, $0\le m\le q$, $0\le n\le p$ (not both m and n simultaneously zero), $A_j>0$ $(j=1,\cdots,p)$, $B_k>0$ $(k=1,\cdots,q)$, and a_j and b_k are complex numbers such that no poles of $\Gamma(b_k+B_ks)$ $(k=1,\cdots,m)$ coincide with poles of $\Gamma(1-a_j-A_js)$ $(j=1,\cdots,n)$. L is a suitable contour $w-i\infty$ to $w+i\infty$, $w\in\mathbb{R}$, separating the poles of the two types mentioned above. For more details, see [18]. As a special case, for $a>0$, $b>0$, and $c>0$, we have

$$\int_0^\infty \exp\{-ay-by^c\}dy = \frac{1}{b^{1/c}c}\mathcal{H}^{1,1}_{1,1}\left[\frac{a}{b^{1/c}}\middle| \begin{array}{c}(\frac{c-1}{c},\frac{1}{c})\\(0,1)\end{array}\right]. \tag{3}$$

Next, let us consider the extreme-value H-function, recently defined in [19]. Thus, this function, hereby denoted as $\mathbb{H}$, can be defined as

$$\mathbb{H}(a_1,a_2,a_3,a_4,a_5,a_6) := \int_0^\infty y^{a_6}\exp\{-a_1y-(a_2y^{a_3}+a_4)^{a_5}\}dy, \tag{4}$$

where $\Re(a_1),\Re(a_2),\Re(a_4)\in\mathbb{R}_+, a_3,a_5\in\mathbb{C}$, not both $\Re(a_1)$ and $\Re(a_2)$ can be equal to zero at the same time, $\Re(a_6)>-1$ when $a_1\neq 0$ or $a_1=0$ and $\text{sign}(a_3)=\text{sign}(a_5)$, and

$\Re(a_6) < -1$ when $a_1 = 0$ and $\text{sign}(a_3) \neq \text{sign}(a_5)$. In this paper, $\mathbb{R}$, $\mathbb{C}$, and $\Re$ denote the real numbers, complex numbers, and the real part of a complex number, respectively.

An important special case of this function is obtained by taking $a_4 = 0$, which represents an upper (or lower) bound for its value depending on the sign of a_5. This case is, therefore, an extreme value of the function and can be written in terms of the $\mathcal{H}$-function as [19]:

$$\begin{aligned}
\mathbb{H}(a_1, a_2, a_3, 0, a_5, a_6) &= \int_0^\infty y^{a_6} \exp\{-a_1 y - a_2^{a_5} y^{a_3 a_5}\} dy \\
&= \frac{1}{a_2^{(1+a_6)/a_3} a_3 a_5} \mathcal{H}_{1,1}^{1,1}\left[a_1 a_2^{-1/a_3} \middle| \begin{matrix} (1 - \frac{(1+a_6)}{a_3 a_5}, \frac{1}{a_3 a_5}) \\ (0,1) \end{matrix} \right] \\
&= \frac{1}{a_1^{a_6+1}} \mathcal{H}_{1,1}^{1,1}\left[\left(\frac{a_2}{a_1^{a_3}} \right)^{a_5} \middle| \begin{matrix} (-a_6, a_3 a_5) \\ (0,1) \end{matrix} \right],
\end{aligned} \tag{5}$$

when $\text{sign}(a_3) = \text{sign}(a_5)$ and

$$\begin{aligned}
\mathbb{H}(a_1, a_2, a_3, 0, a_5, a_6) &= \frac{1}{a_2^{(1+a_6)/a_3} |a_3 a_5|} \mathcal{H}_{0,2}^{2,0}\left[a_1 a_2^{-1/a_3} \middle| \begin{matrix} - \\ (0,1), (\frac{(1+a_6)}{a_3 a_5}, \frac{1}{|a_3 a_5|}) \end{matrix} \right] \\
&= \frac{1}{a_1^{a_6+1}} \mathcal{H}_{0,2}^{2,0}\left[\left(\frac{a_2}{a_1^{a_3}} \right)^{a_5} \middle| \begin{matrix} - \\ (0,1), (1 + a_6, |a_3 a_5|) \end{matrix} \right],
\end{aligned} \tag{6}$$

otherwise.

In the next sections, we prove that all stress–strength probabilities involving three-parameter p-max stable laws can be written as $\mathbb{H}-$functions and, in addition, some parameter restrictions allow Equation (3) to be readily used.

2.2. Three-Parameter p-Max Stable Laws

As indicated, the three-parameter p-max stable laws are obtained by taking CDFs of the types H_i, $i = 1, \cdots, 6$, where all H_i satisfy (2) such that:

$$H_1(x;\alpha) = \begin{cases} 0, & \text{if} \quad x < 1, \\ \exp\{-(\log x)^{-\alpha}\}, & \text{if} \quad x \geq 1, \end{cases} \tag{7}$$

$$H_2(x;\alpha) = \begin{cases} 0, & \text{if} \quad x < 0, \\ \exp\{-(-\log x)^{\alpha}\}, & \text{if} \quad 0 \leq x < 1, \\ 1, & \text{if} \quad x \geq 1, \end{cases} \tag{8}$$

$$H_3(x;\alpha) = \begin{cases} 0, & \text{if} \quad x < -1, \\ \exp\{-(-\log(-x))^{-\alpha}\}, & \text{if} \quad -1 \leq x < 0, \\ 1, & \text{if} \quad x \geq 0, \end{cases} \tag{9}$$

$$H_4(x;\alpha) = \begin{cases} \exp\{-(\log(-x))^{\alpha}\}, & \text{if} \quad x < -1, \\ 1, & \text{if} \quad x \geq -1, \end{cases} \tag{10}$$

$$H_5(x) = \begin{cases} 0, & \text{if} \quad x < 0, \\ \exp\{-x^{-1}\} & \text{if} \quad x \geq 0, \end{cases} \tag{11}$$

$$H_6(x) = \begin{cases} \exp\{x\}, & \text{if} \quad x < 0, \\ 1 & \text{if} \quad x \geq 0, \end{cases} \tag{12}$$

where $\alpha > 0$. These distributions are called Extreme Value Distributions under power normalization or p-max stable laws, and they are known, respectively, as log-Fréchet, log-Weibull, inverse log-Fréchet, inverse log-Weibull, standard Fréchet, and standard Weibull. For a complete characterization of these distributions see [20].

The corresponding three-parameter PDFs are given by

$$h_1(x;\alpha,\beta,\gamma) = \exp\left\{-[\log\gamma x^\beta]^{-\alpha}\right\}\frac{\alpha\beta}{x}[\log(\gamma x^\beta)]^{-\alpha-1}\mathbb{1}_{(\gamma^{-1/\beta},\infty)}(x),$$

$$h_2(x;\alpha,\beta,\gamma) = \exp\left\{-\left[-\log(\gamma x^\beta)\right]^{\alpha}\right\}\left[-\log(\gamma x^\beta)\right]^{\alpha-1}\frac{\beta\alpha}{x}\mathbb{1}_{(0,\gamma^{-1/\beta})}(x),$$

$$h_3(x;\alpha,\beta,\gamma) = \exp\left\{-\left[-\log(\gamma(-x)^\beta)\right]^{-\alpha}\right\}\frac{\alpha\beta}{-x}\left[-\log(\gamma(-x)^\beta)\right]^{-\alpha-1}\mathbb{1}_{(-\gamma^{-1/\beta},0)}(x),$$

$$h_4(x;\alpha,\beta,\gamma) = \exp\{-\left[\log(\gamma(-x)^\beta)\right]^{\alpha}\}\left[\log(\gamma(-x)^\beta)\right]^{\alpha-1}\frac{\alpha\beta}{-x}\mathbb{1}_{(-\infty,-\gamma^{-1/\beta})}(x),$$

$$h_5(x;\beta,\gamma) = \exp\left\{-(\gamma x^\beta)^{-1}\right\}\frac{\beta}{\gamma}x^{-\beta-1}\mathbb{1}_{[0,\infty)}(x),$$

$$h_6(x;\beta,\gamma) = \exp\{-\gamma(-x)^\beta\}\gamma\beta(-x)^{\beta-1}\mathbb{1}_{(-\infty,0)}(x),$$

where $\alpha,\beta,\gamma \in \mathbb{R}_+$, and $\mathbb{1}_A(x)$ denote the indicator function of the set A.

In Section 5, we apply the PDFs h_1, h_2, and h_5 to the modeling of data with positive support. Furthermore, the supports of h_1 and h_2 depend on the parameters. Thus, the maximum likelihood estimation is not as straightforward as in the usual cases. Figures 1–3 show the behavior of these densities for some choices of parameters, revealing the asymmetry and heavy tails of the PDFs.

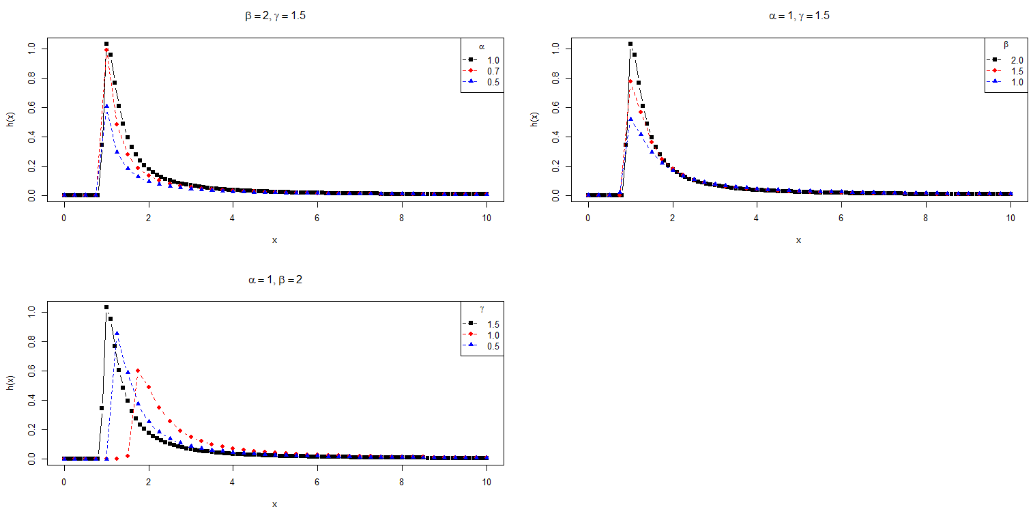

Figure 1. Plots for the PDF $h_1(x;\alpha,\beta,\gamma)$.

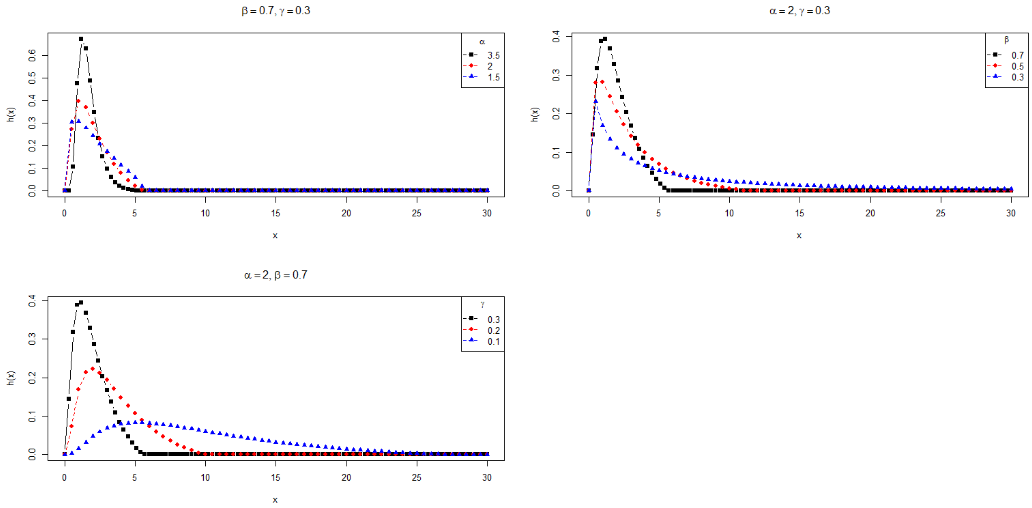

Figure 2. Plots for the PDF $h_2(x;\alpha,\beta,\gamma)$.

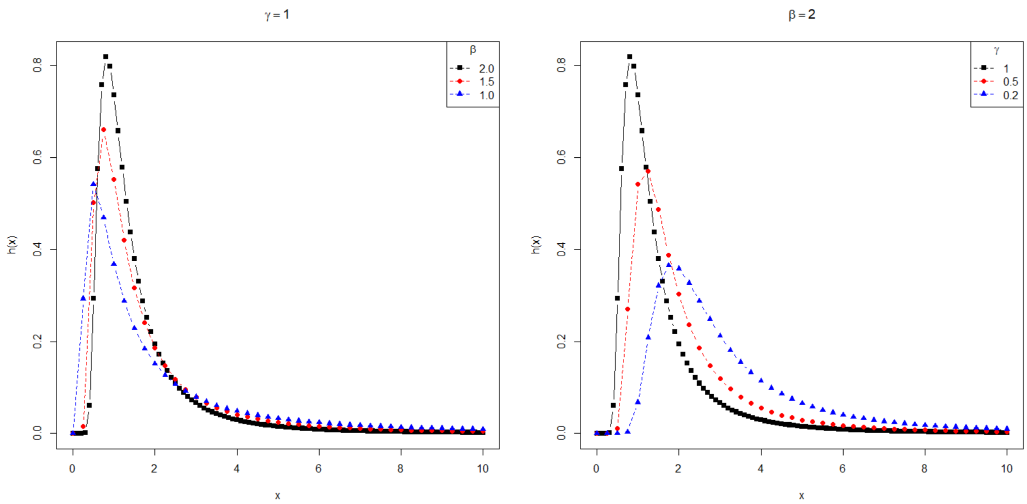

Figure 3. Plots for the PDF $h_5(x;\beta,\gamma)$.

3. Reliability $P(X<Y)$ for Three-Parameter p-Max Stable Laws

In this section, the reliability of two independent three-parameter p-max stable random variables is derived in terms of the $\mathbb{H}-$function. In addition, with suitable parameter restrictions, the $\mathcal{H}$-function and a simpler form in terms of standard functions are obtained. Firstly, we consider the case of two independents $H_1(\cdot;\alpha,\beta,\gamma)$.

Theorem 1. *Let Y and X be independent random variables, respectively, with CDF $H_1(\cdot;\alpha_1,\beta_1,\gamma_1)$ and $H_1(\cdot;\alpha_2,\beta_2,\gamma_2)$, $\alpha_j,\beta_j,\gamma_j\in\mathbb{R}_+$, and $j=1,2$. Then,*

$$R=P(X<Y)=\mathbb{H}\left(1,\frac{\beta_2}{\beta_1},-\frac{1}{\alpha_1},\log\gamma_2-\frac{\beta_2}{\beta_1}\log\gamma_1,-\alpha_2,0\right), \tag{13}$$

provided that $\gamma_1^{-1/\beta_1}\geq\gamma_2^{-1/\beta_2}$. In particular, if $\gamma_1^{1/\beta_1}=\gamma_2^{1/\beta_2}$, then

$$R=\left(\frac{\beta_2}{\beta_1}\right)^{\alpha_1}\frac{\alpha_1}{\alpha_2}\mathcal{H}^{1,1}_{1,1}\left[\left(\frac{\beta_2}{\beta_1}\right)^{\alpha_1}\middle|\begin{array}{c}(\frac{\alpha_2-\alpha_1}{\alpha_2},\frac{\alpha_1}{\alpha_2})\\(0,1)\end{array}\right]. \tag{14}$$

When $\alpha = \alpha_1 = \alpha_2$, R can be written explicitly as

$$R = \frac{\beta_2^\alpha}{\beta_1^\alpha + \beta_2^\alpha}. \tag{15}$$

Proof. Set $\alpha_j, \beta_j, and \gamma_j \in \mathbb{R}_+$ $(j = 1, 2)$. Then,

$$\begin{aligned} R &= P(X < Y) = \int_{-\infty}^{\infty} H_1(x; \alpha_2, \beta_2, \gamma_2) h_1(x; \alpha_1, \beta_1, \gamma_1) dx \\ &= \int_M^\infty \exp\left\{ -\left[\log(\gamma_2 x^{\beta_2})\right]^{-\alpha_2} - \left[\log(\gamma_1 x^{\beta_1})\right]^{-\alpha_1} \right\} \frac{\alpha_1 \beta_1}{x} [\log(\gamma_1 x^{\beta_1})]^{-\alpha_1 - 1} dx, \end{aligned} \tag{16}$$

where $M = \max\{\gamma_1^{-1/\beta_1}, \gamma_2^{-1/\beta_2}\}$. Substituting $y = \left[\log(\gamma_1 x^{\beta_1})\right]^{-\alpha_1}$ and taking $\gamma_1^{-1/\beta_1} \geq \gamma_2^{-1/\beta_2}$, we can rewrite (16) as

$$R = \int_0^\infty \exp\left\{ -y - \left[\frac{\beta_2}{\beta_1} y^{-1/\alpha_1} + \log \gamma_2 - \frac{\beta_2}{\beta_1} \log \gamma_1 \right]^{-\alpha_2} \right\} dy. \tag{17}$$

Hence, (13) follows from (17) and (4). In addition, applying (3) with $\gamma_1^{-1/\beta_1} = \gamma_2^{-1/\beta_2}$, we obtain (14). In the case where $\alpha = \alpha_1 = \alpha_2$, we have the explicit form (15). □

Secondly, we consider the case of two independents $H_2(\cdot; \alpha, \beta, \gamma)$.

Theorem 2. *Let Y and X be independent random variables, respectively, with CDF $H_2(\cdot; \alpha_1, \beta_1, \gamma_1)$ and $H_2(\cdot; \alpha_2, \beta_2, \gamma_2)$, $\alpha_j, \beta_j, \gamma_j \in \mathbb{R}_+$, and $j = 1, 2$. Then,*

$$R = P(X < Y) = \mathbb{H}\left(1, \frac{\beta_2}{\beta_1}, \frac{1}{\alpha_1}, -\log \gamma_2 + \frac{\beta_2}{\beta_1} \log \gamma_1, \alpha_2, 0\right), \tag{18}$$

provided that $\gamma_1^{-1/\beta_1} \leq \gamma_2^{-1/\beta_2}$. In particular, if $\gamma_1^{1/\beta_1} = \gamma_2^{1/\beta_2}$, then

$$R = \left(\frac{\beta_1}{\beta_2}\right)^{\alpha_1} \frac{\alpha_1}{\alpha_2} \mathcal{H}_{1,1}^{1,1}\left[\left(\frac{\beta_1}{\beta_2}\right)^{\alpha_1} \middle| \begin{array}{c} (\frac{\alpha_2 - \alpha_1}{\alpha_2}, \frac{\alpha_1}{\alpha_2}) \\ (0, 1) \end{array} \right]. \tag{19}$$

When $\alpha = \alpha_1 = \alpha_2$, R can be written explicitly as

$$R = \frac{\beta_1^\alpha}{\beta_1^\alpha + \beta_2^\alpha}. \tag{20}$$

Proof. Set $\alpha_j, \beta_j, and \gamma_j \in \mathbb{R}_+$ $(j = 1, 2)$. Then,

$$\begin{aligned} R &= P(X < Y) = \int_{-\infty}^{\infty} H_2(x; \alpha_2, \beta_2, \gamma_2) h_2(x; \alpha_1, \beta_1, \gamma_1) dx \\ &= \int_0^m \exp\left\{ -\left[-\log(\gamma_2 x^{\beta_2})\right]^{\alpha_2} - \left[-\log(\gamma_1 x^{\beta_1})\right]^{\alpha_1} \right\} \frac{\alpha_1 \beta_1}{x} [-\log(\gamma_1 x^{\beta_1})]^{\alpha_1 - 1} dx, \end{aligned} \tag{21}$$

where $m = \min\{\gamma_1^{-1/\beta_1}, \gamma_2^{-1/\beta_2}\}$. Substituting $y = \left[-\log(\gamma_1 x^{\beta_1})\right]^{\alpha_1}$ and taking $\gamma_1^{-1/\beta_1} \leq \gamma_2^{-1/\beta_2}$, we can rewrite (21) as

$$R = \int_0^\infty \exp\left\{ -y - \left[\frac{\beta_2}{\beta_1} y^{1/\alpha_1} - \log \gamma_2 + \frac{\beta_2}{\beta_1} \log \gamma_1 \right]^{\alpha_2} \right\} dy. \tag{22}$$

Hence, (18) follows from (22) and (4). In addition, applying (3) with $\gamma_1^{-1/\beta_1} = \gamma_2^{-1/\beta_2}$, we obtain (19). In the case where $\alpha = \alpha_1 = \alpha_2$, we have the explicit form (20). □

Thirdly, we consider the case of two independents $H_3(\cdot;\alpha,\beta,\gamma)$. The proofs of Theorems 3 and 4 are similar to those of Theorems 1 and 2, respectively. The details are omitted.

Theorem 3. *Let Y and X be independent random variables, respectively, with CDF $H_3(\cdot;\alpha_1,\beta_1,\gamma_1)$ and $H_3(\cdot;\alpha_2,\beta_2,\gamma_2)$, $\alpha_j,\beta_j,\gamma_j \in \mathbb{R}_+$, $j=1,2$. Then,*

$$R = P(X<Y) = \mathbb{H}\left(1, \frac{\beta_2}{\beta_1}, -\frac{1}{\alpha_1}, -\log\gamma_2 + \frac{\beta_2}{\beta_1}\log\gamma_1, -\alpha_2, 0\right), \tag{23}$$

provided that $-\gamma_1^{-1/\beta_1} \geq -\gamma_2^{-1/\beta_2}$. In particular, if $\gamma_1^{1/\beta_1} = \gamma_2^{1/\beta_2}$, then

$$R = \left(\frac{\beta_2}{\beta_1}\right)^{\alpha_1} \frac{\alpha_1}{\alpha_2} \mathcal{H}_{1,1}^{1,1}\left[\left(\frac{\beta_2}{\beta_1}\right)^{\alpha_1} \middle| \begin{array}{c} (\frac{\alpha_2-\alpha_1}{\alpha_2}, \frac{\alpha_1}{\alpha_2}) \\ (0,1) \end{array}\right].$$

When $\alpha=\alpha_1=\alpha_2$, R can be written explicitly as

$$R = \frac{\beta_2^\alpha}{\beta_1^\alpha + \beta_2^\alpha}.$$

Now, we consider the case of two independents $H_4(\cdot;\alpha,\beta,\gamma)$.

Theorem 4. *Let Y and X be independent random variables, respectively, with CDF $H_4(\cdot;\alpha_1,\beta_1,\gamma_1)$ and $H_4(\cdot;\alpha_2,\beta_2,\gamma_2)$, $\alpha_j,\beta_j,\gamma_j \in \mathbb{R}_+$, and $j=1,2$. Then,*

$$R = P(X<Y) = \mathbb{H}\left(1, \frac{\beta_2}{\beta_1}, \frac{1}{\alpha_1}, \log\gamma_2 - \frac{\beta_2}{\beta_1}\log\gamma_1, \alpha_2, 0\right), \tag{24}$$

provided that $-\gamma_1^{-1/\beta_1} \leq -\gamma_2^{-1/\beta_2}$. In particular, if $\gamma_1^{1/\beta_1} = \gamma_2^{1/\beta_2}$, then

$$R = \left(\frac{\beta_1}{\beta_2}\right)^{\alpha_1} \frac{\alpha_1}{\alpha_2} \mathcal{H}_{1,1}^{1,1}\left[\left(\frac{\beta_1}{\beta_2}\right)^{\alpha_1} \middle| \begin{array}{c} (\frac{\alpha_2-\alpha_1}{\alpha_2}, \frac{\alpha_1}{\alpha_2}) \\ (0,1) \end{array}\right].$$

When $\alpha=\alpha_1=\alpha_2$, R can be written explicitly as

$$R = \frac{\beta_1^\alpha}{\beta_1^\alpha + \beta_2^\alpha}.$$

Lastly, we consider the cases of two independents $H_j(\cdot;\beta,\gamma)$ $(j=5,6)$.

Theorem 5. *Let Y and X be independent random variables, respectively, with CDF $H_i(\cdot;\beta_j,\gamma_j)$, $j=1,2$, and $i=5,6$. Then,*

(a) for $i=5$,

$$R = P(X<Y) = \frac{1}{\gamma_1}\mathbb{H}\left(\frac{1}{\gamma_1}, \frac{1}{\gamma_2}, \frac{\beta_2}{\beta_1}, 0, 1, 0\right) \tag{25}$$

$$= \frac{\gamma_2^{\beta_1/\beta_2}\beta_1}{\gamma_1\beta_2} \mathcal{H}_{1,1}^{1,1}\left[\frac{\gamma_2^{\beta_1/\beta_2}}{\gamma_1} \middle| \begin{array}{c} (\frac{\beta_2-\beta_1}{\beta_2}, \frac{\beta_1}{\beta_2}) \\ (0,1) \end{array}\right]. \tag{26}$$

In particular, if $\beta=\beta_1=\beta_2$, we have

$$R = \frac{\gamma_2}{\gamma_1+\gamma_2}; \tag{27}$$

(b) *for* $i = 6$,

$$R = P(X < Y) = \gamma_1 \mathbb{H}\left(\gamma_1, \gamma_2, \frac{\beta_2}{\beta_1}, 0, 1, 0\right) = \frac{\gamma_1 \beta_1}{\gamma_2^{\beta_1/\beta_2} \beta_2} \mathcal{H}_{1,1}^{1,1}\left[\frac{\gamma_1}{\gamma_2^{\beta_1/\beta_2}} \middle| \begin{array}{c} (\frac{\beta_2-\beta_1}{\beta_2}, \frac{\beta_1}{\beta_2}) \\ (0,1) \end{array} \right].$$

In particular, if $\beta = \beta_1 = \beta_2$, *we have*

$$R = \frac{\gamma_1}{\gamma_1 + \gamma_2};$$

Proof. We prove case $i = 5$, and case $i = 6$ follows analogously. We have

$$\begin{aligned} R &= P(X < Y) = \int_{-\infty}^{\infty} H_5(x; \alpha_2, \beta_2, \gamma_2) h_5(x; \alpha_1, \beta_1, \gamma_1) dx \\ &= \int_0^{\infty} \exp\left\{-(\gamma_2 x^{\beta_2})^{-1} - (\gamma_1 x^{\beta_1})^{-1}\right\} \frac{\beta_1}{\gamma_1} x^{-\beta_1 - 1} dx. \end{aligned} \tag{28}$$

Substituting $y = x^{-\beta_1}$ in (28), we obtain

$$R = \frac{1}{\gamma_1} \int_0^{\infty} \exp\left\{-\gamma_1^{-1} y - \gamma_2^{-1} y^{\beta_2/\beta_1}\right\} dy. \tag{29}$$

Therefore, (25) follows from (29) and (4) (alternatively, (26) follows from (29) and (3)). In particular, taking $\beta = \beta_1 = \beta_2$, (27) follows from (29). □

By combining all the Theorems from 1 to 4, it is possible to state the following Corollary:

Corollary 1. *Let* Y *and* X *be independent random variables, respectively, with CDF* $H_w(\cdot; \alpha_1, \beta_1, \gamma_1)$ *and* $H_w(\cdot; \alpha_2, \beta_2, \gamma_2)$, $\alpha_j, \beta_j, \gamma_j \in \mathbb{R}_+$, $j = 1, 2$, *and* $w = 1, \ldots, 4$. *Then,*

$$\begin{aligned} R_w &= P(X < Y) \\ &= \mathbb{H}\left(1, \frac{\beta_2}{\beta_1}, \frac{(-1)^w}{\alpha_1}, \frac{\cos\left(\frac{\pi}{4} + \frac{(w-1)\pi}{2}\right)}{\cos\left(\frac{\pi}{4}\right)}\left(\log \gamma_2 - \frac{\beta_2}{\beta_1} \log \gamma_1\right), (-1)^w \alpha_2, 0\right), \end{aligned} \tag{30}$$

provided that $\gamma_1^{-1/\beta_1} \geq \gamma_2^{-1/\beta_2}$. *In particular, if* $\gamma_1^{1/\beta_1} = \gamma_2^{1/\beta_2}$, *then*

$$R_w = \left(\frac{\beta_2}{\beta_1}\right)^{(-1)^{w+1}\alpha_1} \frac{\alpha_1}{\alpha_2} \mathcal{H}_{1,1}^{1,1}\left[\left(\frac{\beta_2}{\beta_1}\right)^{(-1)^{w+1}\alpha_1} \middle| \begin{array}{c} (\frac{\alpha_2-\alpha_1}{\alpha_2}, \frac{\alpha_1}{\alpha_2}) \\ (0,1) \end{array} \right]. \tag{31}$$

When $\alpha = \alpha_1 = \alpha_2$, R_w *can be written explicitly as*

$$R_w = \frac{\sin^2\left(\frac{w\pi}{2}\right)\beta_2^{\alpha} + \cos^2\left(\frac{w\pi}{2}\right)\beta_1^{\alpha}}{\beta_1^{\alpha} + \beta_2^{\alpha}}. \tag{32}$$

We finish this section by noting that Theorems 1–5 can be generalized to random samples of a given F distribution that is in the domain of attraction of one of the p-max stable laws (see [20] for a complete characterization of the domains of attraction of the p-max stable laws). We describe below these generalizations.

Let $X_1, \cdots, X_n$ be a sample from the CDF F and assume that there exist sequences of real numbers $\{a_n\}$ and $\{b_n\}$ with $a_n, b_n > 0$ such that (2) holds for some H_i $(i = 1, \cdots, 6)$. Set $M_n = \max\{X_1, \cdots, X_n\}$ and $\tilde{M}_n = \left|\frac{M_n}{a_n}\right|^{1/b_n} sign(M_n)$. Equations (1) and (2) imply

$$\begin{aligned}
\lim_{n\to\infty} P(\tilde{M}_n < Y) &= \lim_{n\to\infty} \int_{-\infty}^{\infty} F_{\tilde{M}_n}(x) f_Y(x) dx \\
&= \int_{-\infty}^{\infty} \lim_{n\to\infty} F^n\left(a_n |x|^{b_n} sign(x)\right) h_i(x; \alpha_1, \beta_1, \gamma_1) dx \\
&= \int_{-\infty}^{\infty} H_i(x; \alpha_2, \beta_2, \gamma_2) h_i(x; \alpha_1, \beta_1, \gamma_1) dx \qquad (33)
\end{aligned}$$

where $Y \sim H_i(\cdot; \alpha_1, \beta_1, \gamma_1)$. Using the corresponding Theorems (1–5), (33) can be obtained in terms of the1 $\mathbb{H}$-function.

4. Estimation

This section deals with parameter estimation for the p-max stable laws via a random optimization method and bootstrap confidence intervals.

Several authors (e.g., [8,9,11]) have estimated R by maximum likelihood. However, they relied on strong parameter restrictions to obtain an explicit form for R. Thus, the estimation of the parameters must be done jointly in the two samples. In our case, such restrictions were not necessary since we worked with expressions of R in terms of functions $\mathbb{H}$ and $\mathcal{H}_{1,1}^{1,1}$, releasing any requirements about similar parameters between different samples.

To the best of our knowledge, there are few studies concerning parameter estimation, although the literature suggests several theoretical studies of p-max stable distributions (e.g., [21]). Here, we present a different approach for parameter estimation for the p-max stable laws.

We initially consider the PDF $h_2(\cdot; \alpha, \beta, \gamma)$. For the other p-max stable laws, similar expressions are obtained using the PDFs presented in Section 2.2. Take $\mathbf{x} = (x_1, \cdots, x_n)$ as a sample of n observations. The likelihood function for the PDF $h_2(\cdot; \alpha, \beta, \gamma)$ is given by the following:

$$\begin{aligned}
L_2(\alpha, \beta, \gamma; \mathbf{x}) &= \alpha^n \beta^n \exp\left\{ - \sum_{i=1}^{n} \left[- \log(\gamma x_i^{\beta}) \right]^{\alpha} \right\} \times \\
&\quad \times \prod_{i=1}^{n} \frac{\left[- \log(\gamma x_i^{\beta}) \right]^{\alpha - 1}}{x_i} \mathbb{1}_{(0, \gamma^{-1/\beta})}(x_i). \qquad (34)
\end{aligned}$$

Note that $\prod_{i=1}^{n} \mathbb{1}_{(0,\gamma^{-1/\beta})}(x_i) > 0$ if and only if $x_i \in (0, \gamma^{-1/\beta})$ for all $i = 1, \cdots, n$. Then, we are not able to obtain the MLE explicitly, so an additional condition is required in the likelihood maximization process.

Remark 1. *The MLE of R is obtained using the invariance property of MLE. This is due to the Theorems 1–5 that describe R in terms of the function $\mathbb{H}$ (which is an integral, hence a continuous and measurable function).*

4.1. A Random Optimization Method for Approximating the MLE

Now, we describe the optimization methodology to be implemented for parameter estimation. Let $L : \Theta \subset \mathbb{R}^d \to \mathbb{R}$ be a likelihood function for which the maximum $y_0 = \max\{L(\theta); \theta \in \Theta\}$ is assumed to be finite.

Algorithm 1 can find the point of maximum θ_0 for which $L(\theta_0) = y_0$. Particularly, unlike conventional algorithms, random points in space are generated according to a generic distribution G (not necessarily uniform) on the parameter space Θ. This allows us to introduce weights in some regions of the parameter space, as a kind of prior information.

Algorithm 1. *Let $\xi_1, \xi_2, \cdots$ be independent and identically distributed random vectors with common distribution G on Θ. Let $(\theta_1, Y_1), (\theta_2, Y_2), \cdots$ be defined by*

Step 1 . $\theta_1 = \xi_1$ and $Y_1 = L(\theta_1)$.
Step $k+1$. Having defined (θ_k, Y_k), let (θ_{k+1}, Y_{k+1}) be defined as

$$\begin{cases} \theta_{k+1} = \xi_{k+1} \text{ and } Y_{k+1} = L(\xi_{k+1}), & \text{if } L(\xi_{k+1}) \geq Y_k; \\ \theta_{k+1} = \theta_k \text{ and } Y_{k+1} = Y_k, & \text{otherwise.} \end{cases}$$

It was proved by [22] that for given $\varepsilon > 0$ and $0 < \delta_\varepsilon < 1$

$$P(|L(\theta_n) - y_0| \leq \varepsilon) \geq 1 - \delta_\varepsilon,$$

that is, the $\varepsilon-$region of attraction of y_0 has been attained with probability $1 - \delta_\varepsilon$, provided that the stop rule consists of terminating the algorithm for k such that

$$k \geq -\frac{m(\Theta)\log\delta_\varepsilon}{\varepsilon}, \tag{35}$$

where m denotes the Lebesgue measure on $\mathbb{R}^N$. This means that with high probability, the algorithm reaches the desired maximum.

4.2. Bootstrap

The bootstrap method used in the next section to obtain bootstrap confidence intervals of R is described below.

Algorithm 2 describes the approach used in the next section to obtain bootstrap confidence intervals of R.

1. Generate independent bootstrap samples **X** and **Y** of sizes n_x and n_y, respectively.
2. Compute the parameter estimation based on **X** and **Y**.
3. Obtain $\hat{R}$.
4. Repeat steps $1-3$ $M = 1000$ times.
5. The approximate $100(1-\alpha)\%$ confidence interval of R is given by $[\hat{R}_M(\alpha/2), \hat{R}_M(1-\alpha/2)]$, where $\hat{R}_M(\alpha) \approx \hat{G}^{-1}(\alpha)$ and $\hat{G}$ are the cumulative distribution function of $\hat{R}$.

Algorithm 2. *Let* **X** *and* **Y** *be samples of sizes n_x and n_y, respectively, and a positive integer M.*

Step 1 *Generate independent bootstrap samples* **X** *and* **Y**.
Step 2 *Compute the parameter estimation based on* **X** *and* **Y**.
Step 3 *Obtain $\hat{R}$.*
Step 4 *Repeat steps* **1** *to* **3** *M times.*
Step 5 *The approximate $100(1-\nu)\%$ confidence interval of R is given by $[\hat{R}_M(\nu/2), \hat{R}_M(1-\nu/2)]$, where $\hat{R}_M(\nu) \approx \hat{G}^{-1}(\nu)$ and $\hat{G}$ is the cumulative distribution function of $\hat{R}$.*

5. Applications

In this section, we present Monte Carlo simulations as well as the modeling of two real situations involving football datasets and different-length carbon fibers. In order to enable readers to apply the methodology hereby proposed, the codes are available at a public repository [23] (link to be shared after acceptance).

5.1. Simulation Results

To illustrate the behavior of the random optimization method for approximating the MLE described in Algorithm 1 and to evaluate the performance of the estimate $\hat{R}$, we simulate random samples from CDFs H_1, H_2, and H_5 (the other distributions could be used as well). The random samples are simulated using the generalized inverse of the CDFs applied to uniform random variables.

Values of $\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \varepsilon, \delta_\varepsilon$, and n are pre-specified, where n is the sample size and ε and δ_ε are the parameters of the random optimization Algorithm 1.

Monte Carlo simulations were implemented in the language and environment for statistical computing R-4.4.0 [24], and each simulation outcome is based on $M = 1000$ samples of the parameter settings. In Tables 1 and 2, we study the performance of the estimator $\hat{R}$ when the PDF is h_2. The PDFs h_1 and h_5 are treated in Tables 3 and 4.

Table 1. Mean, bias, and RMSE of $\hat{R}$ for PDF h_2 ($\varepsilon = 0.1$, $\delta_\varepsilon = 0.1$, and $n = 30$).

N	α_2	β_2	γ_2	α_1	β_1	γ_1	R	$\hat{R}_{MC}$	Bias	RMSE
2	1	0.50	0.30	0.70	0.50	0.30	0.5240	0.5203	−0.0036	0.1550
2	1	0.50	0.30	0.70	0.75	0.30	0.4057	0.3716	−0.0342	0.1487
2	1	0.50	0.30	1.00	0.75	0.30	0.4017	0.3524	−0.0493	0.1480
2	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2873	−0.0888	0.1537
2	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2915	−0.0846	0.1575
2	1	0.30	0.55	1.00	0.55	0.90	0.3769	0.2860	−0.0910	0.1558
2	1	1.00	1.00	0.90	0.50	0.50	0.6641	0.6647	0.0007	0.1308
2	1	1.00	1.00	0.95	0.50	0.50	0.6654	0.6643	−0.0011	0.1287
2	1	1.00	1.00	0.97	0.50	0.50	0.6659	0.6650	−0.0009	0.1319
2	1	1.00	1.00	1.00	0.50	0.50	0.6667	0.6668	0.0002	0.1287
5	1	0.50	0.30	0.70	0.50	0.30	0.5240	0.5392	0.0152	0.1501
5	1	0.50	0.30	0.70	0.75	0.30	0.4057	0.4051	−0.0007	0.1379
5	1	0.50	0.30	1.00	0.75	0.30	0.4017	0.3606	−0.0411	0.1435
5	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.3107	−0.0654	0.1380
5	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.3110	−0.0651	0.1412
5	1	0.30	0.55	1.00	0.55	0.90	0.3769	0.3056	−0.0713	0.1438
5	1	1.00	1.00	0.90	0.50	0.50	0.6641	0.6621	−0.0020	0.1255
5	1	1.00	1.00	0.95	0.50	0.50	0.6654	0.6537	−0.0117	0.1309
5	1	1.00	1.00	0.97	0.50	0.50	0.6659	0.6575	−0.0085	0.1301
5	1	1.00	1.00	1.00	0.50	0.50	0.6667	0.6552	−0.0115	0.1292
10	1	0.50	0.30	0.70	0.50	0.30	0.5240	0.5320	0.0080	0.1492
10	1	0.50	0.30	0.70	0.75	0.30	0.4057	0.4078	0.0021	0.1412
10	1	0.50	0.30	1.00	0.75	0.30	0.4017	0.3765	−0.0251	0.1419
10	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.3204	−0.0558	0.1319
10	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.3248	−0.0514	0.1325
10	1	0.30	0.55	1.00	0.55	0.90	0.3769	0.3110	−0.0660	0.1390
10	1	1.00	1.00	0.90	0.50	0.50	0.6641	0.6396	−0.0244	0.1376
10	1	1.00	1.00	0.95	0.50	0.50	0.6654	0.6398	−0.0256	0.1360
10	1	1.00	1.00	0.97	0.50	0.50	0.6659	0.6400	−0.0259	0.1342
10	1	1.00	1.00	1.00	0.50	0.50	0.6667	0.6401	−0.0266	0.1286

Remark 2. *(a) We start our study of simulations by PDF h_2 inspired by the applications presented in Section 5.2. As it will be seen, h_2 presented a good modeling performance in those cases.*

(b) Recall that Algorithm 1 depends on parameters $(\varepsilon, \delta_\varepsilon)$ and on a distribution G on the parameter space Θ. To estimate (α, β, γ), Θ must be a subspace of $\mathbb{R}^3_+$. We start by fixing the values $\varepsilon = 0.1, \delta_\varepsilon = 0.1$, the search region $\Theta_0 = [0, N]^3$ for a fixed N value, and the uniform distribution on Θ_0. Table 1 gives the results for different values of N. Changes in the constants $\varepsilon = 0.1$ and $\delta_\varepsilon = 0.1$ or on the upper bound of N could result in better or worse estimation. Table 2 shows results for $\varepsilon = 0.01$ and $\delta_\varepsilon = 0.05$.

(c) If we want to search the entire parameter space, we should use another distribution with support on R^3_+ (e.g., gamma distribution).

For the simulation, we fix a search region Θ_0 of the parameter space Θ, and for each line in the table

1. 1000 random samples of $X \sim H_2(\alpha_2, \beta_2, \gamma_2)$ and $Y \sim H_2(\alpha_1, \beta_1, \gamma_1)$ are simulated;
2. for each simulation, the parameter $R = P(X < Y)$ is estimated, according to the likelihood function (34) and Algorithm 1;

3. the mean of the 1000 corresponding $\hat{R}$ (denote $\hat{R}_{MC}$) is obtained;
4. the Bias and the Root Mean Squared Error (RMSE) are computed.

Table 1 shows that

- in general, the estimation of R had good results, indicated by the small value of the bias;
- the bias values were within the fixed range $\varepsilon = 0.1$;
- RMSE did not increase as we increased the search space Θ_0.

Table 2 shows that by reducing the value of ε and δ_ε while increasing n, we are able to reduce the RMSE values of Table 1, although we did not obtain significant reductions in the Bias values.

Table 2. Mean, bias, and RMSE of $\hat{R}$ for PDF h_2 ($\varepsilon = 0.01$, $\delta_\varepsilon = 0.05$, and $n = 200$).

N	α_2	β_2	γ_2	α_1	β_1	γ_1	R	$\hat{R}_{MC}$	Bias	RMSE
2	1	0.50	0.30	0.70	0.50	0.30	0.5240	0.5331	0.0090	0.0940
2	1	0.50	0.30	0.70	0.75	0.30	0.4057	0.3671	−0.0390	0.0880
2	1	0.50	0.30	1.00	0.75	0.30	0.4017	0.3418	−0.0600	0.0940
2	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2811	−0.0950	0.1170
2	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2801	−0.0960	0.1170
2	1	0.30	0.55	1.00	0.55	0.90	0.3769	0.2785	−0.0980	0.1180
2	1	1.00	1.00	0.90	0.50	0.50	0.6641	0.6628	−0.0010	0.0720
2	1	1.00	1.00	0.95	0.50	0.50	0.6654	0.6621	−0.0030	0.0690
2	1	1.00	1.00	0.97	0.50	0.50	0.6659	0.6602	−0.0060	0.0680
2	1	1.00	1.00	1.00	0.50	0.50	0.6667	0.6649	−0.0020	0.0660
5	1	0.50	0.30	0.70	0.50	0.30	0.5240	0.5299	0.0060	0.0920
5	1	0.50	0.30	0.70	0.75	0.30	0.4057	0.3632	−0.0430	0.0900
5	1	0.50	0.30	1.00	0.75	0.30	0.4017	0.3395	−0.0620	0.0960
5	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2854	−0.0910	0.1140
5	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2820	−0.0940	0.1160
5	1	0.30	0.55	1.00	0.55	0.90	0.3769	0.2767	−0.1000	0.1200
5	1	1.00	1.00	0.90	0.50	0.50	0.6641	0.6606	−0.0030	0.0720
5	1	1.00	1.00	0.95	0.50	0.50	0.6654	0.6601	−0.0050	0.0710
5	1	1.00	1.00	0.97	0.50	0.50	0.6659	0.6626	−0.0030	0.0680
5	1	1.00	1.00	1.00	0.50	0.50	0.6667	0.6663	0.0000	0.0670
10	1	0.50	0.30	0.70	0.50	0.30	0.5240	0.5311	0.0072	0.0898
10	1	0.50	0.30	0.70	0.75	0.30	0.4057	0.3664	−0.0394	0.0907
10	1	0.50	0.30	1.00	0.75	0.30	0.4017	0.3365	−0.0651	0.0984
10	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2839	−0.0923	0.1155
10	1	0.30	0.55	0.70	0.55	0.90	0.3762	0.2827	−0.0935	0.1167
10	1	0.30	0.55	1.00	0.55	0.90	0.3769	0.2770	−0.1000	0.1211
10	1	1.00	1.00	0.90	0.50	0.50	0.6641	0.6634	−0.0007	0.0703
10	1	1.00	1.00	0.95	0.50	0.50	0.6654	0.6635	−0.0019	0.0692
10	1	1.00	1.00	0.97	0.50	0.50	0.6659	0.6620	−0.0040	0.0688
10	1	1.00	1.00	1.00	0.50	0.50	0.6667	0.6619	−0.0047	0.0694

As in the h_2 case, for the PDFs h_1 and h_5, the search space is $\Theta_0 = [0, N]^3$. Thus, based on the results from Tables 1 and 2 and on the computational difficulties with the increasing of N, we restrict our analysis for $N = 2$ when the PDF is h_1 (for h_5, which has only two parameters to be estimated, we keep $N = 10$).

Tables 3 and 4 present the mean, bias, and RMSE of $\hat{R}_{MC}$ for $M = 1000$ Monte Carlo simulations of X and Y from H_1 and H_5, respectively, with $n = 30$, $\varepsilon = 0.01$, and $\delta_\varepsilon = 0.05$. For the simulations, we followed the same procedure used in the generation of Table 1.

Table 3. Mean, bias, and RMSE of $\hat{R}$ for PDF h_1 ($\varepsilon = 0.01$, $\delta_\varepsilon = 0.05$, $N = 2$, and $n = 30$).

α_2	β_2	γ_2	α_1	β_1	γ_1	R	$\hat{R}_{MC}$	Bias	RMSE
1	0.50	0.30	0.70	0.50	0.30	0.5240	0.4161	−0.1078	0.2443
1	0.50	0.30	0.70	0.75	0.30	0.3721	0.3728	0.0007	0.2103
1	0.50	0.30	1.00	0.75	0.30	0.3064	0.4284	0.1219	0.2032
1	0.30	0.55	0.70	0.55	0.90	0.3182	0.3056	−0.0125	0.1478
1	0.30	0.55	0.70	0.55	0.90	0.3182	0.3097	−0.0085	0.1562
1	0.30	0.55	1.00	0.55	0.90	0.2414	0.2493	0.0079	0.1214
1	1.00	1.00	0.90	0.50	0.50	0.7961	0.7569	−0.0391	0.1347
1	1.00	1.00	0.95	0.50	0.50	0.7949	0.7661	−0.0288	0.1215
1	1.00	1.00	0.97	0.50	0.50	0.7944	0.7668	−0.0276	0.1151
1	1.00	1.00	1.00	0.50	0.50	0.7937	0.7657	−0.0279	0.1236

Table 4. Mean, bias, and RMSE of $\hat{R}$ for PDF h_5 ($\varepsilon = 0.01$, $\delta_\varepsilon = 0.05$, $N = 10$, and $n = 30$).

β_2	γ_2	β_1	γ_1	R	$\hat{R}_{MC}$	Bias	RMSE
0.30	0.50	0.20	0.30	0.7362	0.7811	0.0450	0.0783
0.30	0.50	0.60	0.50	0.3443	0.3309	−0.0134	0.0743
0.30	0.50	0.90	1.00	0.2050	0.1337	−0.0713	0.0848
0.50	0.70	0.20	0.30	0.8706	0.9080	0.0374	0.0545
0.50	0.70	0.60	0.50	0.5473	0.5249	−0.0225	0.0796
0.50	0.70	0.90	1.00	0.3507	0.2960	−0.0546	0.0853
1.00	1.00	0.20	0.30	0.9393	0.9666	0.0273	0.0344
1.00	1.00	0.60	0.50	0.7370	0.7050	−0.0320	0.0755
1.00	1.00	0.90	1.00	0.5071	0.5247	0.0176	0.0778
1.00	1.00	1.00	1.00	0.5000	0.4997	−0.0003	0.0755

Table 4 shows that the estimator $\hat{R}_{MC}$ has a better performance with a more precise estimation when the PDF is h_5, also with less RMSEs. This was expected since there are only two parameters to be estimated from a random sample of X and two others from Y.

5.2. Real Dataset Applications

We now give two applications using real data analyzed earlier in the literature.

5.2.1. Medium Pass Completion Proportion

We use the data for the UEFA Champions League and 2022 FIFA World Cup datasets (Available at https://www.kaggle.com/ accessed on 13 February 2024) to illustrate the model developed in the preceding sections. We compare the medium pass completion proportion, that is, relative frequency of successful passes between 14 and 18 m - thus a number from 0 (none of the passes) to 1 (all the passes). These datasets were modeled before in [25]. For the convenience of the reader, datasets from UEFA (X) and FIFA (Y) are presented below:

$$\begin{aligned} X = (&0.289, 0.700, 0.211, 0.733, 0.444, 0.544, 0.089, 0.767, 0.433, 0.911,\\ &0.800, 0.733, 0.278, 0.456, 0.178, 0.200, 0.244, 0.467, 0.022, 0.400, 0.378,\\ &0.589, 0.600, 0.567, 0.844, 0.711, 0.289, 0.178, 0.489, 0.278, 0.611, 0.544,\\ &0.267, 0.489, 0.467, 0.300, 0.311) \end{aligned}$$

and

$$\begin{aligned} Y = (&0.888, 0.815, 0.907, 0.891, 0.827, 0.898, 0.856, 0.861, 0.890, 0.860,\\ &0.920, 0.894, 0.913, 0.849, 0.781, 0.828, 0.864, 0.820, 0.846, 0.879,\\ &0.860, 0.885, 0.862, 0.769, 0.845, 0.846, 0.931, 0.863, 0.856, 0.879,\\ &0.812, 0.841). \end{aligned}$$

Descriptive statistics for X and Y are presented in Table 5. The boxplot shown in Figure 4 shows that X is more dispersed than Y and that Y values tend to be larger than X values (1st quartile of Y is greater than 3rd quartile of Y). Computing the value of the statistic R is important to probabilistically measure such a difference observed in the datasets.

Table 5. Descriptive statistics for X and Y.

Dataset	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	n
X	0.02	0.28	0.46	0.45	0.60	0.91	37
Y	0.77	0.84	0.86	0.86	0.89	0.93	32

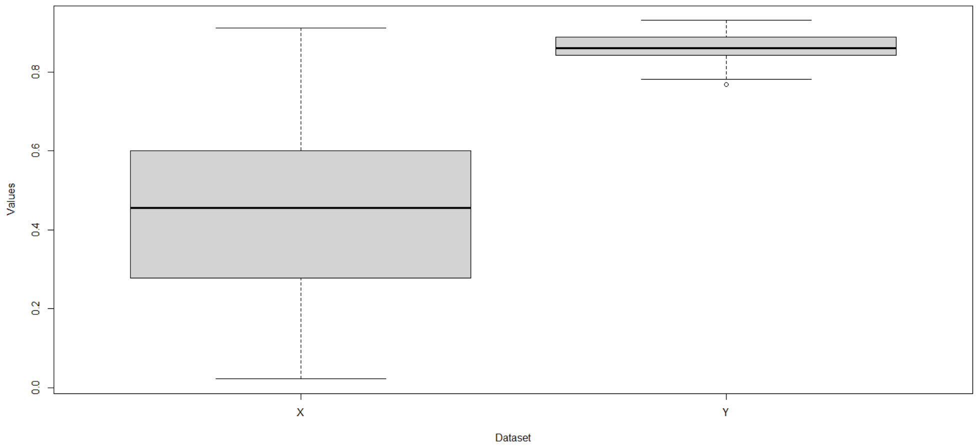

Figure 4. Boxplots of X (**left**) and Y (**right**).

As the datasets have positive support, the PDF candidates to model such datasets are h_1, h_2, and h_5. After estimating the parameters for the three distributions, the information criteria AIC, BIC, and EDC were applied, justifying the choice of PDF h_2 (see Table 6). This choice was also supported by the Kolmogorov–Smirnov test, whose p-values were 0.8954 and 0.726, for X and Y, respectively, indicating that we could not reject the null hypothesis that the CDF is H_2. Figure 5 shows the fit of distributions to datasets.

Table 6. Estimated parameters and information criteria for model selection.

Dataset	PDF	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	AIC	BIC	EDC
X	h_1	0.7698	0.4113	5.4562	103.95	88.29	102.25
	$\mathbf{h_2}$	**1.8231**	**0.7822**	**0.8249**	**−8.26**	**−23.93**	**−9.97**
	h_5	–	0.9778	3.5948	29.66	19.22	28.53
Y	h_1	17.9328	1.1439	3.3069	−113.48	−128.27	−114.94
	$\mathbf{h_2}$	**2.1254**	**9.7882**	**1.8085**	**−127.37**	**−142.17**	**−128.83**
	h_5	–	17.8660	19.9746	−112.29	−122.15	−113.26

Using Theorem 2, we obtain $\hat{R} = 0.9235$ and the 95% Bootstrap confident interval is $(0.8921, 1.0000)$. That indicates a high probability that the proportion of successful passes between 14 and 18 m in UEFA matches was lower than in FIFA.

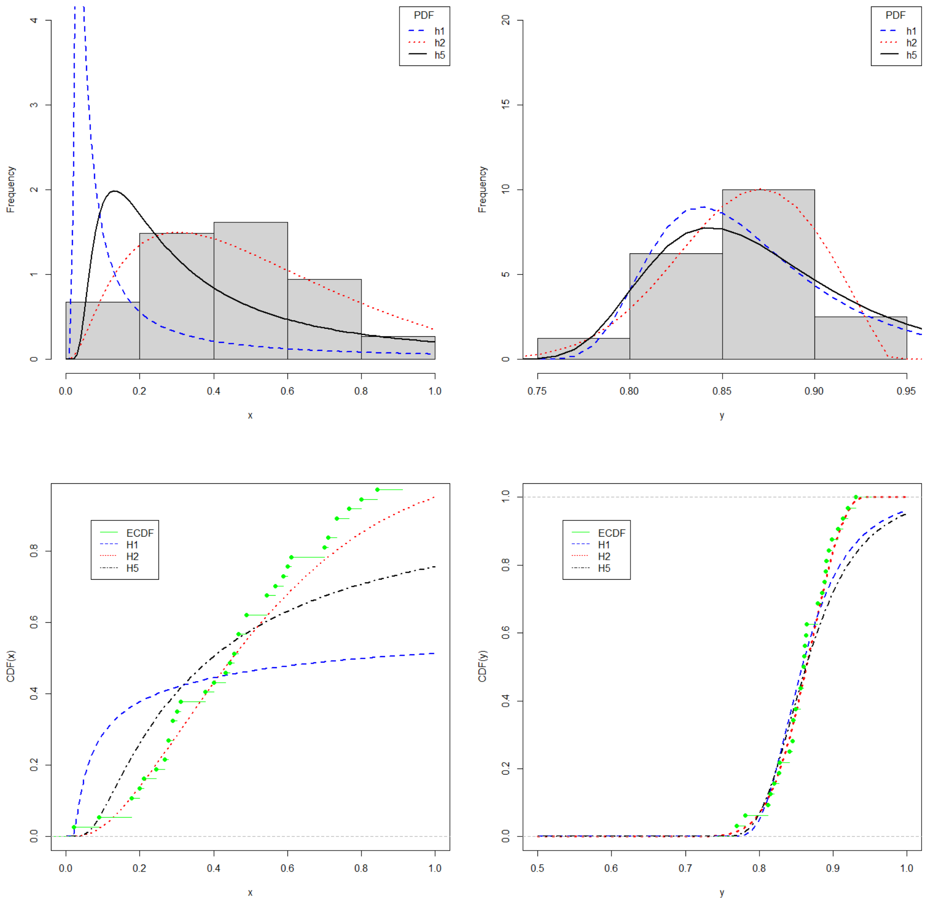

Figure 5. Plots for X (**left**) and Y (**right**). On (**top**), histogram and fitted PDF; on (**bottom**), empirical CDF and fitted CDF.

5.2.2. Carbon Fibers

We now present an application of stress–strength probability calculation in the modeling and comparison of carbon fibers of length 20 mm (X) and 10 mm (Y). X and Y represent the strength data measured in GPa (Gigapascal) for single carbon fibers tested under tension and are also presented below in addition to being frequently used in the literature (e.g., [26]).

$$
\begin{aligned}
X \quad = \quad & (1.312, 1.314, 1.479, 1.552, 1.700, 1.803, 1.861, 1.865, 1.944, 1.958, 1.966, \\
& 1.977, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.140, 2.179, 2.224, 2.240, \\
& 2.253, 2.270, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382, 2.382, 2.426, 2.434, \\
& 2.435, 2.478, 2.490, 2.511, 2.514, 2.535, 2.554, 2.566, 2.570, 2.586, 2.629, \\
& 2.633, 2.642, 2.648, 2.684, 2.697, 2.726, 2.770, 2.773, 2.800, 2.809, 2.818, \\
& 2.821, 2.848, 2.880, 2.954, 3.012, 3.067, 3.084, 3.090, 3.096, 3.128, 3.233, \\
& 3.433, 3.585, 3.585)
\end{aligned}
$$

$$Y = (1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020)$$

Table 7 and Figure 6 show that the descriptive profile of X and Y in which it is possible to observe that Y (carbon fibers of length 10 mm) tends to have greater values than X (carbon fibers of length 20 mm). This indicates that we expect a probability $R = P\,(X < Y)$ greater than $1/2$.

Table 7. Descriptive statistics for X and Y.

Dataset	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	n
X	1.31	2.10	2.48	2.45	2.77	3.58	69
Y	1.90	2.55	3.00	3.06	3.42	5.02	63

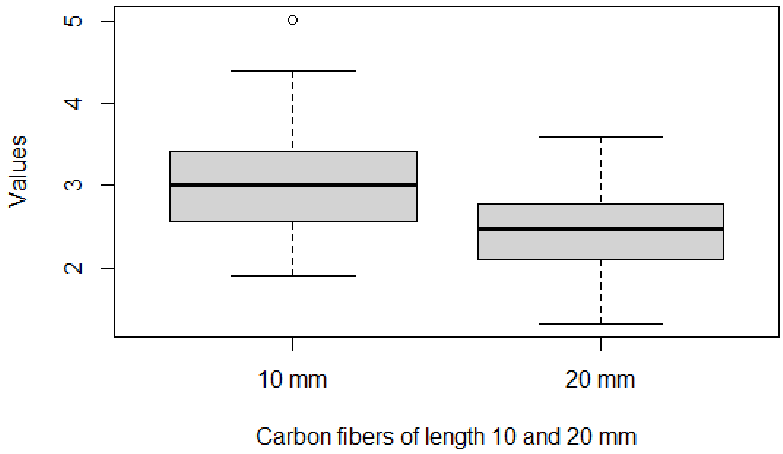

Figure 6. Boxplots of X (**left**) and Y (**right**).

Considering that both X and Y are positive datasets, for the estimation of R, we must choose a p-max stable distribution with positive support. In this case, the candidates are h_1, h_2, and h_5, whose estimated parameters are shown in Table 8. In this same table, we also compared the fitted p-max distributions to the fittings obtained for the Weibull (WB) and the Exponentiated Weibull (EWB) distributions, with the latter being introduced in [27]. As a selection criterion for the best distribution, the AIC, BIC, and EDC criteria are evaluated, and we choose PDF h_2. For the Y random variable, h_2 is found to be the best distribution according to all the metrics used. On the other hand, for the X random variable, by a small margin, the Weibull distribution is the best according to AIC e BIC; on the other hand, h_2 shows up as the best model according to EDC. Considering that EDC is a generalization of both AIC and BIC and that it encompasses an optimal penalization term, h_2 was chosen as the best model for the X random variable as well. The fit of the p-max distributions considered, in particular H_2, to the data can be seen in Figure 7.

Table 8. Estimated parameters and information criteria for model selection.

Dataset	PDF	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	AIC	BIC	EDC
X	h_1	26.1763	0.1583	2.4240	138.20	157.60	140.86
	$\mathbf{h_2}$	**2.1882**	**2.0072**	**0.0706**	104.57	123.98	**107.23**
	h_5	–	3.3015	0.1254	153.71	166.65	155.48
	WB	3.8428 (shape)	–	11.3142 (scale)	**101.74**	**114.68**	114.68
	EWB	0.4839	1.8690	1.84582	150.83	163.77	163.77
Y	h_1	32.7124	0.1640	2.3043	113.27	94.41	110.74
	$\mathbf{h_2}$	**4.2267**	**1.1982**	**0.1073**	**106.26**	**87.40**	**103.73**
	h_5	–	2.4785	0.0639	162.05	149.48	160.37
	WB	3.9090 (shape)	–	38.5449 (scale)	124.30	136.88	125.99
	EWB	0.3860	1.7889	1.4415	183.50	196.07	185.18

Legend: Weibull (WB), Exponentiated Weibull (EWB).

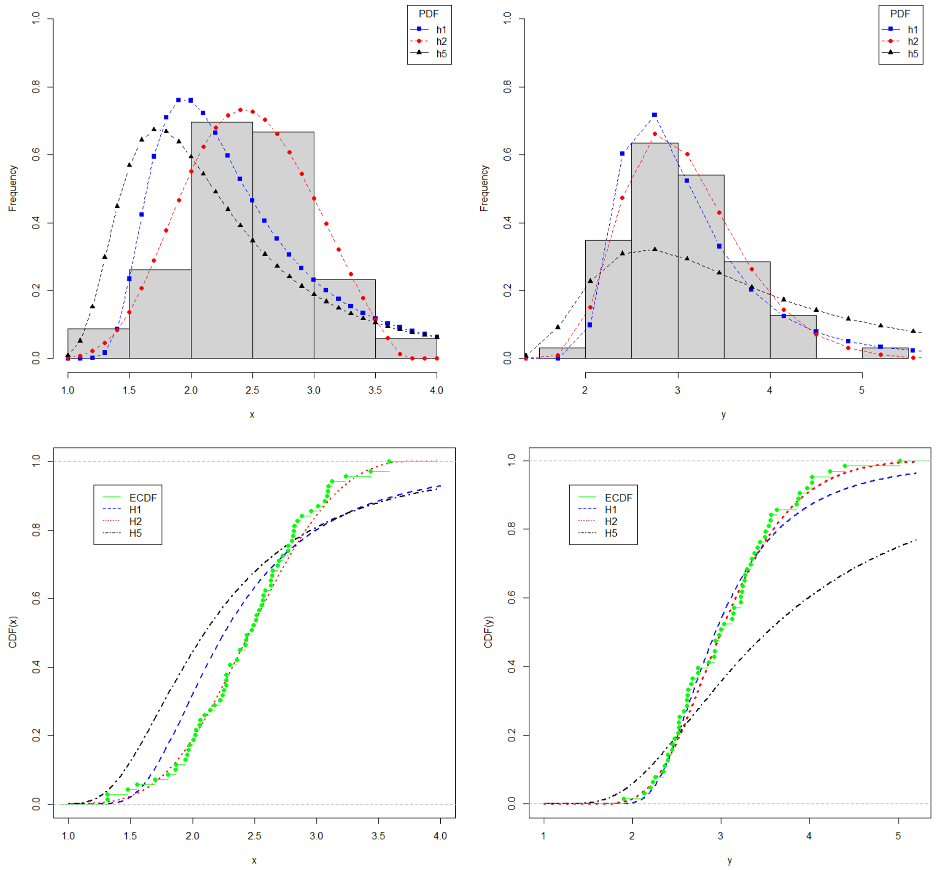

Figure 7. Plots for X (**left**) and Y (**right**). On top, histogram and fitted PDF; on bottom, empirical CDF and fitted CDF.

We present the following conclusions:

- According to the AIC, BIC, and EDC criteria, PDF h_2 is the one that best fits data X and Y. This was expected since the same data were already modeled via Weibull distribution (see [11,28]) and having positive right endpoint, Theorem 3.1 in [20] establishes that H_2 would be the corresponding p-max stable distribution;
- The p-values of the Kolmogorov–Smirnov test are 0.9404 and 0.8390, respectively, which indicate that we cannot reject the null hypotheses that the X and Y CDFs are H_2.
- The same conclusion for Table 8 can be obtained from the analysis of Figure 7, which presents the adjustment of the PDFs h_1, h_2, and h_3 to the datasets.

- Based on the choice of H_2 to model the data X and Y and on the estimated parameters (Table 8), the estimated value of R calculated from (18) is $\hat{R} = 0.7701$, and the 95% Bootstrap confident interval for R is $(0.7017, 0.8425)$.

6. Conclusions

Our study aimed to investigate the estimation of the $R = P(X < Y)$ for independent marginals X and Y following p-max stable distributions. In order to do so, we obtained exact expressions for R. By using the new formulas proposed, direct and exact reliability applications are made possible for an important class of asymmetric distributions.

We discuss the application of a novel class of special functions , the so-called extreme value $\mathbb{H}-$function, which allows us to write the expressions of R explicitly and with minimal restrictions. In particular, by imposing additional parameter restrictions, R can be calculated in terms of $\mathcal{H}-$functions as well as even more compact expressions.

To the best of our knowledge, there are no previous works in the literature aiming to provide expressions and frameworks to perform reliability statistical inference for p-max stable distributions, and this work stands as a contribution by providing estimation methods based on stochastic optimization.

A restraint of our estimation method is the fact that it relies on compact search spaces $[0, N]^3$ for fixed N. However, we tested the performance of the proposed estimator by a Monte Carlo simulation study. Even though the search range N exponentially governs the computational effort required, the reported results reveal the correctness of the methodological approach hereby proposed.

Two applications to real datasets were carried out to show the performance of the p-max stable laws in reliability scenarios. Future work may explore other extreme value distributions and their reliability calculations, such as bimodal Weibull, bimodal Gumbel, bimodal GEV, and the extreme-value Birnbaum-Saunders distribution.

Overall, it is possible to summarize the strengths of the present paper as follows:

1. General expressions were analytically derived for $R = P(X < Y)$ when X and Y follow three-parameter p-max stable laws with fewer parameter restrictions compared to previous results in the literature;
2. A stochastic optimization procedure was proposed to build an estimator for R based on the novel expressions derived;
3. The validity of the expressions and of the general methodological framework developed were demonstrated by Monte Carlo simulations;
4. The suitability of the p-max distributions to model real datasets was attested by study cases.

On the other hand, the main weaknesses of the present paper are as follows:

1. The stochastic optimization procedure relies on compact search spaces $[0, N]^3$ for fixed N, whose impact is exponential on the computational effort required;
2. The amount of data used to illustrate the methodology and equations developed in the paper is limited; thus, the superiority of the p-max distributions over other possible models needs to be assessed in a case-by-case fashion.

Author Contributions: Conceptualization, F.S.Q.; methodology, F.S.Q., T.A.d.F., L.C.d.S.M.O. and P.N.R.; software, F.S.Q. and T.A.d.F.; validation, F.S.Q., T.A.d.F., L.C.d.S.M.O. and P.N.R.; formal analysis, F.S.Q., T.A.d.F., L.C.d.S.M.O. and P.N.R.; investigation, F.S.Q.; writing—original draft preparation, F.S.Q. and T.A.d.F.; writing—review and editing, L.C.d.S.M.O. and P.N.R.; supervision, P.N.R. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was partially waived by the Editorial Office of Symmetry, and the remaining value was fully funded by University of Brasilia (Edital DPI/DPG/BCE Nº 01/2024).

Data Availability Statement: The data presented in the study are contained with the article and openly available in Zenodo at https://zenodo.org/doi/10.5281/zenodo.11416971.

Acknowledgments: The authors acknowledge the support provided by the University of Brasilia (UnB).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kotz, S.; Lumelskii, Y.; Pensky, M. *The Stress-Strength Model and Its Generalizations: Theory and Applications*; World Scientific: Singapore, 2003.
2. Rathie, P.; Freitas, S.; Nojosa, R.; Mendes, A.; Silva, T. Stress-Strength Reliability Models Involving H-Function Distributions. *J. Ramanujan Soc. Math. Math. Sci.* **2022**, *9*, 217–234.
3. Rathie, P.N.; Ozelim, L.C.S.M. Exact and approximate expressions for the reliability of stable Lévy random variables with applications to stock market modelling. *J. Comput. Appl. Math.* **2017**, *321*, 314–322. [CrossRef]
4. Quintino, F.S.; Oliveira, M.; Rathie, P.N.; Ozelim, L.C.S.M.; da Fonseca, T.A. Asset selection based on estimating stress-strength probabilities: The case of returns following three-parameter generalized extreme value distributions. *AIMS Math.* **2023**, *9*, 2345–2368. [CrossRef]
5. Domma, F.; Giordano, S. A copula-based approach to account for dependence in stress-strength models. *Stat. Pap.* **2012**, *54*, 807–826. [CrossRef]
6. Huang, C.Y.; Chin, W.C.B.; Wen, T.H.; Fu, Y.H.; Tsai, Y.S. EpiRank: Modeling Bidirectional Disease Spread in Asymmetric Commuting Networks. *Sci. Rep.* **2019**, *9*, 5415.
7. Nadarajah, S. Reliability for extreme value distributions. *Math. Comput. Model.* **2003**, *37*, 915–922. [CrossRef]
8. Abbas, K.; Tang, Y. Objective Bayesian analysis of the Frechet stress–strength model. *Stat. Probab. Lett.* **2014**, *84*, 169–175. [CrossRef]
9. Jia, X.; Nadarajah, S.; Guo, B. Bayes estimation of $P(Y < X)$ for the Weibull distribution with arbitrary parameters. *Appl. Math. Model.* **2017**, *47*, 249–259.
10. Krishnamoorthy, K.; Lin, Y. Confidence limits for stress–strength reliability involving Weibull models. *J. Stat. Plan. Inference* **2010**, *140*, 1754–1764. [CrossRef]
11. Kundu, D.; Raqab, M. Estimation of $R = P(Y < X)$ for three-parameter Weibull distribution. *Stat. Probab. Lett.* **2009**, *79*, 1839–1846.
12. Aryal, G.R.; Tsokos, C.P. On the transmuted extreme value distribution with application. *Nonlinear Anal. Theory Methods Appl.* **2009**, *71*, e1401–e1407. [CrossRef]
13. Vila, R.; Niyazi Çankaya, M. A bimodal Weibull distribution: Properties and inference. *J. Appl. Stat.* **2022**, *49*, 3044–3062. [CrossRef]
14. Otiniano, C.; Vila, R.; Brom, P.; Bourguignon, M. On the Bimodal Gumbel Model with Application to Environmental Data. *Austrian J. Stat.* **2023**, *52*, 45–65. [CrossRef]
15. Otiniano, C.E.; Paiva, B.S.; Vila, R.; Bourguignon, M. A bimodal model for extremes data. *Environ. Ecol. Stat.* **2023**, *30*, 261–288. [CrossRef]
16. Pantcheva, E. Limit theorems for extreme order statistics under nonlinear normalization. In *Stability Problems for Stochastic Models*; Springer: Berlin/Heidelberg, Germany, 1985; pp. 284–309.
17. Feng, B.; Chen, S. On large deviations of extremes under power normalization. *Stat. Probab. Lett.* **2015**, *99*, 27–35. [CrossRef]
18. Mathai, A.; Saxena, R.; Haubold, H. *The H-Function: Theory and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.
19. Rathie, P.N.; Ozelim, L.C.d.S.M.; Quintino, F.; Fonseca, T.A.d. On the Extreme Value H-Function. *Stats* **2023**, *6*, 802–811. [CrossRef]
20. Mohan, N.; Ravi, S. Max domains of attraction of univariate and multivariate p-max stable laws. *Theory Probab. Its Appl.* **1993**, *37*, 632–643. [CrossRef]
21. Barakat, H.; Nigm, E.; Khaled, O. Extreme value modeling under power normalization. *Appl. Math. Model.* **2013**, *37*, 10162–10169. [CrossRef]
22. Dorea, C.C. Stopping Rules for a Random Optimization. *SIAM J. Control Optim.* **1990**, *28*, 841–850. [CrossRef]
23. Quintino, F.; Rathie, P.; Ozelim, L.; Fonseca, T. Codes for Estimation of $P(X < Y)$ stress–strength reliability measures for a class of asymmetric distributions: the case of three-parameter p-max stable laws. *Zenodo* 2024. Available online: https://zenodo.org/records/11416972 (accessed on 13 February 2024).
24. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022.
25. Vila, R.; Balakrishnan, N.; Saulo, H.; Zörnig, P. Family of bivariate distributions on the unit square: Theoretical properties and applications. *J. Appl. Stat.* **2023**, *51*, 1729–1755. [CrossRef]
26. Valiollahi, R.; Asgharzadeh, A.; Raqab, M.Z. Estimation of P (Y < X) for Weibull distribution under progressive Type-II censoring. *Commun. Stat.-Theory Methods* **2013**, *42*, 4476–4498.
27. Mudholkar, G.; Srivastava, D. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Trans. Reliab.* **1993**, *42*, 299–302. [CrossRef]
28. Kundu, D.; Gupta, R. Estimation of $P[Y < X]$ for Weibull distributions. *IEEE Trans. Reliab.* **2006**, *55*, 270–280. [CrossRef]