

Article

A Bivariate Extension of Type-II Generalized Crack Distribution for Modeling Heavy-Tailed Losses

Taehan Bae ^{1,*}  and Hanson Quarshie ^{2,†}¹ Department of Mathematics and Statistics, University of Regina, Regina, SK S4S 0A2, Canada² University of Regina, Regina, SK S4S 0A2, Canada; hdq233@uregina.ca

* Correspondence: taehan.bae@uregina.ca

† These authors contributed equally to this work.

Abstract: As an extension of the (univariate) Birnbaum–Saunders distribution, the Type-II generalized crack (GCR2) distribution, built on an appropriate base density, provides a sufficient level of flexibility to fit various distributional shapes, including heavy-tailed ones. In this paper, we develop a bivariate extension of the Type-II generalized crack distribution and study its dependency structure. For practical applications, three specific distributions, GCR2-Generalized Gaussian, GCR2-Student's t , and GCR2-Logistic, are considered for marginals. The expectation-maximization algorithm is implemented to estimate the parameters in the bivariate GCR2 models. The model fitting results on a catastrophic loss dataset show that the bivariate GCR2 distribution based on the generalized Gaussian density fits the data significantly better than other alternative models, such as the bivariate lognormal distribution and some Archimedean copula models with lognormal or Pareto marginals.

Keywords: heavy-tailed distribution; type-II generalized crack distribution; Spearman's rho; Kendall's tau; EM algorithm; catastrophic loss

MSC: 60E05; 62H10



Citation: Bae, T.; Quarshie, H. A. Bivariate Extension of Type-II Generalized Crack Distribution for Modeling Heavy-Tailed Losses. *Mathematics* **2024**, *12*, 3718. <https://doi.org/10.3390/math12233718>

Academic Editor: Himchan Jeong

Received: 30 October 2024

Revised: 23 November 2024

Accepted: 24 November 2024

Published: 27 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Researchers have examined diverse classes of distributions to study various facets of problems. Most loss datasets in the context of actuarial loss modeling share some common characteristics, such as being skewed to the right, unimodal (multimodal in certain situations), and having a thin left tail and a moderate to extremely thick right tail. In recent years, various classes of heavy-tailed distributions, including the subexponential distribution class, have been studied for modeling heavy-tailed (or extreme) data [1–3]. On the other hand, a different stream of problems, such as the periodic vibrations in commercial aircraft, have motivated the introduction of the Birnbaum–Saunders (BS) distribution [4]. The BS distribution models the total time elapsed until a critical threshold is exceeded by fatigue accumulated on a subject (material) of interest, causing the failure event (or a crack) of the material to occur. Due to its ability to fit right-skewed data, the BS distribution is highly effective for modeling numerous scenarios, e.g., situations where there is an accumulation of a certain factor that drives a quantifiable characteristic to surpass a critical threshold. See [5] for details on the theoretical properties and applications of the BS distribution.

Various extensions of the BS distribution have been discussed in the literature. In [6], an extended version of Birnbaum–Saunders distribution family is introduced using the density of elliptical distribution in place of the standard normal density that quantifies the amount of the stress per cycle of material use in the BS setting. In [7], the extreme value version of the generalized Birnbaum–Saunders (GBS) distribution, whose tail thickness is determined by that of the auxiliary distribution (i.e., an elliptical distribution), has been discussed. Some applications of the extreme value BS models can be found in [5,8].

The (three-parameter) Gaussian crack (lifetime) distribution introduced in [9] is another important extension of the BS distribution. The Gaussian crack (CR) distribution is a two-component mixture of inverse Gaussian and length-biased inverse Gaussian distributions with a weight parameter p , and it features increased flexibility to fit various datasets due to the additional mixture weight parameter. However, the Gaussian crack distribution relies on the standard normal base density, and thus it lacks heavy-tailedness. The limited applicability of the Gaussian crack distribution for modeling heavy-tailed data such as insurance losses motivates the construction of a large class of generalized crack (GCR) distributions [10]. The GCR distribution class contains the Gaussian crack distribution as a specific member, and each member of the class is built on a specific choice of a base-density function that determines the tail characteristics of the resulting GCR distribution. In [10,11], the GCR distributions with the Student's t and the generalized Gaussian base-density functions are applied to catastrophic losses and heavy-tailed precipitation time series, respectively.

In [12], the GCR distribution class has recently been further extended to the class of Type-II generalized crack (GCR2) distributions in which an additional shape parameter τ is included to increase flexibility over the GCR class. The key distributional properties, such as the tail characteristics of GCR2 distribution, depend on the specification of the base-density function and the shape parameters involved in each model.

In the literature, several important modeling frameworks allow for the creation of new distribution families from given ones, including the Azzalini method [13], Lehmann-type distributions [14], and Topp-Leone families [15]. While each framework possesses its own distributional characteristics, these general frameworks are common in the sense that the specification of the baseline distribution function (or a density function) plays a crucial role in determining key distributional properties of the constructed model. For instance, the Lehmann-type I setting with Stoppa (baseline) distribution function renders a model that can be effectively used for actuarial data with extreme observations [16]. In this sense, the modeling approach used in the construction of the GCR/GCR2 distribution class is in line with these general frameworks. Typical parametric distributions built under the aforementioned general frameworks have simple (closed) forms, and thus, they may not perform well when data features complex (i.e., multimodal) shapes. Due to its mixture structure, the GCR/GCR2 distribution with an appropriate baseline density can be advantageous over the simple models in such cases.

Regarding applications for bivariate data with heavy-tailed marginals, a bivariate GCR distribution in the form of a four-component mixture of independent models has been constructed in [17]. The authors demonstrate that bivariate GCR models can exhibit useful dependence structures and serve as valuable models for diverse real-world situations.

This paper aims to extend the univariate GCR2 models to bivariate cases by employing the mixture model structure used in [17]. Three specific examples of GCR2 distributions, i.e., a newly constructed GCR2 model based on the logistic density in addition to the GCR2- t and the GCR2-GG models introduced in [12], are used as marginals to effectively model heavy-tailed insurance/catastrophic loss data. We investigate some theoretical properties of the proposed bivariate models, such as the conditional distribution and the dependence structure, and discuss the expectation-maximization (EM) algorithm for model estimation. The applicability of the proposed bivariate GCR2 models and the estimation method is illustrated through a model fitting to a real disaster loss dataset.

The rest of this paper is organized as follows. Section 2 gives a brief review of the origin, definition, examples, and key properties of the univariate Type-II generalized crack distributions for the reader's convenience. In Section 3, the bivariate Type-II generalized crack distribution is introduced with some detailed discussions of its theoretical properties. A method of model estimation and its application on a real catastrophic loss dataset are presented in Sections 4 and 5, respectively. Finally, Section 6 provides some concluding remarks.

2. Type-II Generalized Crack Distribution

2.1. Birnbaum–Saunders Distribution

Birnbaum–Saunders distribution (also known as the fatigue life distribution) introduced in [4] is one of the most popular distributions for modeling the time of a crack (or a failure) occurrence when a material specimen is used repeatedly and experiences material fatigue due to the gradual accumulation of stress/damages. The failure event occurs when the accumulated stress on the material specimen hits a critical threshold $w > 0$, and the failure time T is the first hitting time of the accumulated stress to the critical threshold. The derivation of the Birnbaum–Saunders distribution as an approximate distribution of the first hitting time of the Brownian motion can be found in [12]. In fact, the BS distribution is a two-point mixture of the inverse Gaussian (IG) and the length-biased inverse Gaussian (LB-IG) distributions with equal mixture weights. Formally, the cumulative distribution function (cdf) and probability density function (pdf) of BS distribution, respectively, are given as follows:

$$F_{BS}(x; \alpha, \beta) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)\right), \quad x > 0, \tag{1}$$

$$f_{BS}(x; \alpha, \beta) = \frac{1}{2}f_{IG}(x; \alpha, \beta) + \frac{1}{2}f_{LB-IG}(x; \alpha, \beta), \quad x > 0, \tag{2}$$

where

$$f_{IG}(x; \alpha, \beta) = \frac{\sqrt{\beta}}{\alpha} x^{-3/2} \phi\left[\frac{1}{\alpha}\left\{\left(\frac{x}{\beta}\right)^{1/2} - \left(\frac{\beta}{x}\right)^{1/2}\right\}\right] \tag{3}$$

$$f_{LB-IG}(x; \alpha, \beta) = \frac{1}{\alpha\sqrt{\beta}} x^{-1/2} \phi\left[\frac{1}{\alpha}\left\{\left(\frac{x}{\beta}\right)^{1/2} - \left(\frac{\beta}{x}\right)^{1/2}\right\}\right], \tag{4}$$

in which $\beta > 0$ and $\alpha > 0$ are the scale and shape parameters, respectively; $\Phi(\cdot)$ and $\phi(\cdot)$ denote the cdf and pdf of the standard normal distribution, and f_{IG} and f_{LB-IG} denote the cdf and pdf of the two-parameter IG and the LB-IG distributions, respectively.

2.2. Gaussian Crack Distribution

The Gaussian crack distribution introduced in [9] is an extension of BS distribution by introducing the weight parameter $p \in [0, 1]$ in place of the fixed weight of 1/2 in the BS distribution. Naturally, the CR distribution allows for higher flexibility compared to the classical BS distribution. The pdf of the Gaussian crack distribution is given as follows:

$$\begin{aligned} f_{CR}(x; \alpha, \beta, p) &= p f_{IG}(x; \alpha, \beta) + q f_{LB-IG}(x; \alpha, \beta) \\ &= \left(p \frac{\sqrt{\beta}}{\alpha} x^{-3/2} + (1-p) \frac{1}{\alpha\sqrt{\beta}} x^{-1/2}\right) \phi\left[\frac{1}{\alpha}\left\{\left(\frac{x}{\beta}\right)^{1/2} - \left(\frac{\beta}{x}\right)^{1/2}\right\}\right], \end{aligned}$$

for $x > 0$, where $\beta > 0$, $\alpha > 0$ and $0 \leq p \leq 1$ are scale, shape, and mixture weight parameters, respectively, and $q = 1 - p$.

Even though the right tail of the CR distribution thickens as p decreases, the tail maintains the shape of an exponential distribution. It is easy to see that the limit of the hazard rate function of the CR distribution converges to a constant that is greater than zero, i.e.,

$$\lim_{x \rightarrow \infty} \frac{f_{CR}(x)}{1 - F_{CR}(x)} = \frac{1}{2\beta\alpha^2} > 0,$$

which suggests that the Gaussian crack distribution does not belong to heavy-tailed distribution class [3,18].

2.3. Generalized Crack Distribution

In [10], the Gaussian CR distribution family has been extended to a large class of generalized crack (GCR) distribution whose members depend on the specification of the base-density function $g(\cdot)$, where $g(\cdot)$ is a base (or auxiliary) density function which is symmetric about zero. The base-density function replaces the standard normal density function $\phi(\cdot)$ used for the Gaussian crack density.

Specifically, a random variable T has the generalized crack(GCR) distribution with base density g , denoted as $GCR(\alpha, \beta, p; g)$ where the parameters $\alpha > 0$, $\beta > 0$ and the mixture weight $0 \leq p \leq 1$, if its pdf is given as

$$f_{GCR}(x; \alpha, \beta, p; g) = pf_{IS}(x; \alpha, \beta; g) + qf_{LB-IS}(x; \alpha, \beta; g),$$

where

$$f_{IS}(x; \alpha, \beta; g) = \frac{\sqrt{\beta}}{\alpha} x^{-3/2} g \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta} \right)^{1/2} - \left(\frac{\beta}{x} \right)^{1/2} \right\} \right]$$

$$f_{LB-IS}(x; \alpha, \beta; g) = \frac{1}{\alpha\sqrt{\beta}} x^{-1/2} g \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta} \right)^{1/2} - \left(\frac{\beta}{x} \right)^{1/2} \right\} \right].$$

The functions f_{IS} and f_{LB-IS} denotes the pdfs of the inverse symmetric (IS) and the length-biased inverse symmetric (LB-IS) distribution, respectively, on the base-density function $g(\cdot)$. The cdf of the $GCR(\alpha, \beta, p; g)$ distribution is given as:

$$F_{GCR}(x; \alpha, \beta, p; g) = \bar{G}(b(x)) + (2p - 1) \int_{b(x)}^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds$$

where $\bar{G}(x) = 1 - G(x)$ is the survival function of the distribution with the cdf $G(x) = \int_0^x g(s) ds$ corresponding to the base-density function $g(\cdot)$, and $b(x) = \frac{1}{\alpha} [\sqrt{\beta/x} - \sqrt{x/\beta}]$.

A general expression for the n th raw moments of the IS and LB-IS random variable, along with details on the tail behavior of the GCR distribution, is provided in [10]. Their findings demonstrated that, with an appropriate choice of the base-density function, the heavy-tailed generalized crack distribution performs better than many well-known parametric distributions, such as lognormal, Pareto type II, and Weibull distributions, which are frequently used in modeling positively skewed and heavy-tailed extreme data sets.

2.4. Type-II Birnbaum–Saunders (BS2) Distribution

The Type-II Birnbaum–Saunders distribution extends the BS distribution by introducing another shape parameter τ . Applying the inverse transform method to Equation (1), it can easily be verified that the following stochastic relationship holds. Let us define a random variable T as

$$T = \beta \left(\frac{\alpha Z + \sqrt{\alpha^2 Z^2 + 4}}{2} \right)^2$$

where Z is a standard normal random variable. The Type-II Birnbaum–Saunders (BS2) distribution includes an additional shape parameter $\tau > 0$ to the above expression, defined formally as

$$T = \beta \left(\frac{\alpha Z + \sqrt{\alpha^2 Z^2 + 4}}{2} \right)^{1/\tau}.$$

The pdf of the Type-II Birnbaum–Saunders (BS2) distribution is

$$\begin{aligned}
 f_{BS2}(x; \alpha, \beta, \tau) &= \left(\frac{2\tau}{\alpha\beta}\right) \left[\frac{1}{2} \left(\frac{\beta}{x}\right)^{\tau+1} + \frac{1}{2} \left(\frac{x}{\beta}\right)^{\tau-1} \right] \phi \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau \right\} \right] \\
 &= \frac{1}{2} f_{BS2}^{(1)}(x; \alpha, \beta, \tau) + \frac{1}{2} f_{BS2}^{(2)}(x; \alpha, \beta, \tau),
 \end{aligned}$$

where $\alpha > 0, \beta > 0, \tau > 0$ are model parameters, and

$$\begin{aligned}
 f_{BS2}^{(1)}(x; \alpha, \beta, \tau) &= \left(\frac{2\tau}{\alpha\beta}\right) \left(\frac{\beta}{x}\right)^{\tau+1} \phi \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau \right\} \right] \\
 f_{BS2}^{(2)}(x; \alpha, \beta, \tau) &= \left(\frac{2\tau}{\alpha\beta}\right) \left(\frac{x}{\beta}\right)^{\tau-1} \phi \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau \right\} \right].
 \end{aligned}$$

The cdf of BS2 distribution is

$$F_{BS2}(x; \alpha, \beta, \tau) = \Phi \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau \right\} \right]. \tag{5}$$

Like the BS distribution, the BS2 distribution is a two-point mixture of densities $f_{BS2}^{(1)}$ and $f_{BS2}^{(2)}$ with equal weights. However, with the extra shape parameter τ , the BS2 distribution provides more flexibility over the BS distribution. The BS2 distribution reduces to the Birnbaum–Saunders distribution when $\tau = 1/2$.

2.5. Type-II Generalized Crack Distribution

The Type-II generalized crack distribution class introduced in [12] can be seen as a natural extension of the Type-II Birnbaum–Saunders distribution by replacing the standard normal density with a symmetric base density and including the mixture weight parameter p . Specifically, the pdf of GCR2 distribution with base density g is given as follows:

$$f_{GCR2}(x; \alpha, \beta, \tau, p; g) = p f_{GCR2}^{(1)}(x; \alpha, \beta, \tau; g) + q f_{GCR2}^{(2)}(x; \alpha, \beta, \tau; g)$$

where

$$\begin{aligned}
 f_{GCR2}^{(1)}(x; \alpha, \beta, \tau; g) &= \left(\frac{2\tau}{\alpha\beta}\right) \left(\frac{\beta}{x}\right)^{\tau+1} g \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau \right\} \right] \\
 f_{GCR2}^{(2)}(x; \alpha, \beta, \tau; g) &= \left(\frac{2\tau}{\alpha\beta}\right) \left(\frac{x}{\beta}\right)^{\tau-1} g \left[\frac{1}{\alpha} \left\{ \left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau \right\} \right],
 \end{aligned}$$

and the cdf is given as

$$\begin{aligned}
 F_{GCR2}(x; \alpha, \beta, \tau, p; g) &= p F_{GCR2}^{(1)}(x; \alpha, \beta, \tau; g) + q F_{GCR2}^{(2)}(x; \alpha, \beta, \tau; g) \\
 &= \bar{G}(b(x)) + (2p - 1) \int_{b(x)}^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \tag{6}
 \end{aligned}$$

where

$$\begin{aligned}
 F_{GCR2}^{(1)}(x; \alpha, \beta, \tau; g) &= \bar{G}(b(x)) + \int_{b(x)}^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\
 F_{GCR2}^{(2)}(x; \alpha, \beta, \tau; g) &= \bar{G}(b(x)) - \int_{b(x)}^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds,
 \end{aligned}$$

and $b(x) := b(x; \alpha, \beta, \tau) = \alpha^{-1}\{(\beta/x)^\tau - (x/\beta)^\tau\}$. By the symmetry of the base density g and using L'Hopital's rule, it is easy to verify the following asymptotic relation:

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{\bar{F}_{GCR2}(x; \alpha, \beta, \tau, p; g)}{G(b(x))} &= \lim_{x \rightarrow \infty} \frac{g(b(x))b'(x) \left\{ 1 + (1 - 2p) \frac{-b(x)}{\sqrt{(b(x))^2 + 4/\alpha^2}} \right\}}{g(b(x))b'(x)} \\ &= 2(1 - p). \end{aligned}$$

Due to the symmetry of the base density g , $G(x) = \bar{G}(-x)$, and thus, the GCR2 distribution with base density g has the following asymptotic tail:

$$\bar{F}_{GCR2}(x; \alpha, \beta, \tau, p; g) \sim 2(1 - p)G(b(x)) = 2(1 - p)\bar{G}\left\{\alpha^{-1}[(x/\beta)^\tau - (\beta/x)^\tau]\right\},$$

as $x \rightarrow \infty$. Here we write $a(x) \sim b(x)$ as $x \rightarrow \infty$ if $\lim_{x \rightarrow \infty} a(x)/b(x) = 1$. Specifically when $p = 0$, $\bar{F}_{GCR2}(x; \alpha, \beta, \tau, p; g) \sim 2\bar{G}\{\alpha^{-1}[(x/\beta)^\tau - (\beta/x)^\tau]\}$, and, when $p = 1$, the tail of the GCR2 distribution decays to zero faster than $G(b(x))$ as $x \rightarrow \infty$.

2.6. Specific Examples of GCR2 Distributions

The specification of the base-density function on which the GCR2 distribution is built may depend on some key distributional features, such as tail characteristics that are required for each specific application. Two practical examples, the GCR2- t and the GCR2-GG distributions whose base densities are the Student's t and the generalized Gaussian (normal) distributions, respectively, are given in [12]. In this paper, we also consider the GCR2 distribution with the logistic base density, referred to as the GCR2-LG distribution, as another member of the GCR2 distribution class.

Example 1 (GCR2- t distribution). *The Student's t distribution has a regular varying tail, and its density function is given as*

$$g_t(x; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}, \nu > 0.$$

From this the density of the GCR2- t distribution is expressed as

$$\begin{aligned} f_{GCR2}^t(x; \alpha, \beta, \tau, p; \nu) &= \frac{2\tau\Gamma\left(\frac{\nu+1}{2}\right)}{\alpha\beta\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left[p\left(\frac{\beta}{x}\right)^{\tau+1} + q\left(\frac{x}{\beta}\right)^{\tau-1} \right] \\ &\quad \times \left[1 + \frac{1}{\alpha^2\nu} \left\{ \left(\frac{x}{\beta}\right)^{2\tau} + \left(\frac{\beta}{x}\right)^{2\tau} - 2 \right\} \right]^{-\left(\frac{\nu+1}{2}\right)}. \end{aligned}$$

The Student's t distribution belongs to the Maximum Domain of Attraction (MDA) of the Fréchet distribution with index ν , which means that the distribution of a properly normalized maximum of independent and identically distributed (i.i.d.) Student's t random variables converge to a Fréchet distribution asymptotically. In [12], it is shown that the tail of the GCR2- t distribution is regularly varying with index $\tau\nu$, and thus, it also belongs to the MDA of the Fréchet distribution. Compared to the tail of the GCR- t distribution that has the index $\frac{1}{2}\nu$, the tail of GCR2- t becomes heavier than that of the GCR- t when τ becomes smaller than $1/2$. Figures 1–3 illustrate the shapes of the GCR2- t density function for a few sets of prescribed parameter values.

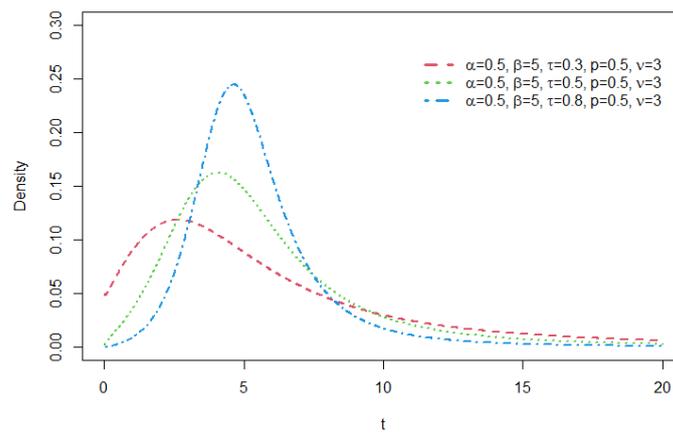


Figure 1. Density functions of GCR2- t distribution with $\alpha = 0.5, \beta = 5, \tau \in \{0.3, 0.5, 0.8\}, p = 0.5,$ and $\nu = 3.$

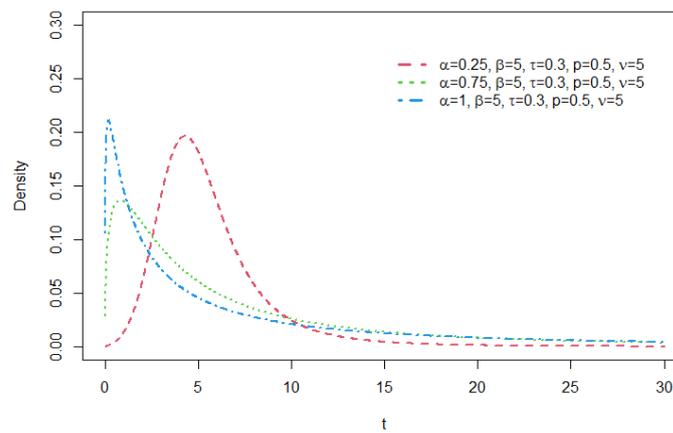


Figure 2. Density functions of GCR2- t distribution with $\alpha \in \{0.25, 0.75, 1\}, \beta = 5, \tau = 0.3, p = 0.5,$ and $\nu = 5.$

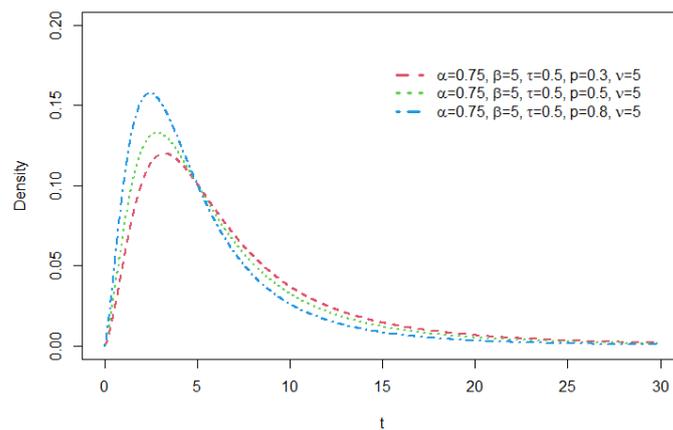


Figure 3. Density functions of GCR2- t distribution with $\alpha = 0.75, \beta = 5, \tau = 0.5, p \in \{0.3, 0.5, 0.8\}$ and $\nu = 5.$

Example 2 (GCR2-GG distribution). *The generalized Gaussian (GG) distribution is a large distribution family that encompasses both thin to moderately heavy-tailed ones, and it can be useful for various practical applications. In particular, when $\theta < 1$, the GG distribution has a subexponential tail (see [3] for details on the subexponential distribution class).*

The pdf of the generalized Gaussian distribution is given as

$$g_{GG}(x; \theta) = \frac{\theta}{2\lambda\Gamma\left(\frac{1}{\theta}\right)} \exp\left(-\left|\frac{x-\mu}{\lambda}\right|^\theta\right).$$

Recall that the base-density function is required to be symmetric for the construction of the GCR2 model. For this, we set $\mu = 0$, and, for the identification of parameters, we further set $\lambda = \sqrt{\Gamma(1/\theta)/\Gamma(3/\theta)}$. The pdf of the resulting GCR2-GG distribution is given as follows.

$$f_{GCR2}^{GG}(x; \alpha, \beta, \tau, p; \theta) = \frac{\tau\theta}{\alpha\beta\lambda\Gamma\left(\frac{1}{\theta}\right)} \left[p\left(\frac{\beta}{x}\right)^{\tau+1} + q\left(\frac{x}{\beta}\right)^{\tau-1} \right] \exp\left(-\left|\frac{1}{\alpha\lambda}\left[\left(\frac{x}{\beta}\right)^\tau - \left(\frac{\beta}{x}\right)^\tau\right]\right|^\theta\right).$$

Figures 4–6 illustrate the shapes of the GCR2-GG density function for a few prescribed sets of parameter values.

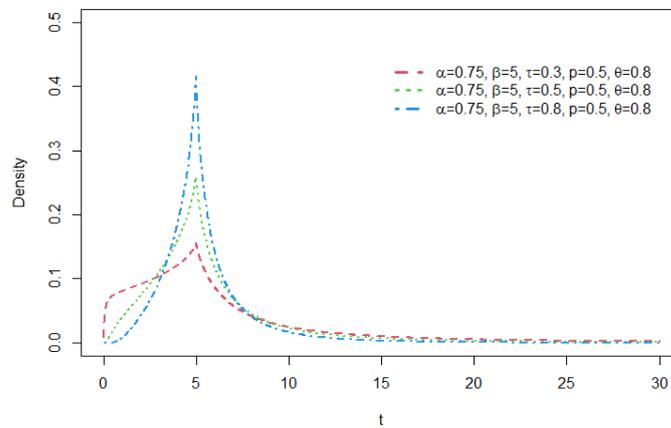


Figure 4. Density functions of GCR2-GG distribution with $\alpha = 0.75$, $\beta = 5$, $\tau \in \{0.3, 0.5, 0.8\}$, $p = 0.5$, and $\theta = 0.8$.

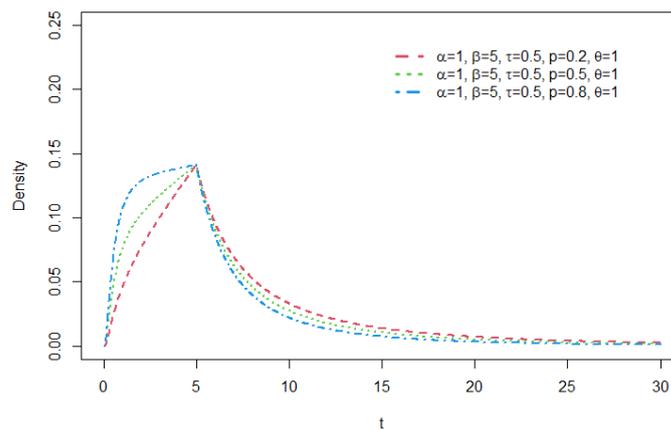


Figure 5. Density functions of GCR2-GG distribution with $\alpha = 1$, $\beta = 5$, $\tau = 0.5$, $p \in \{0.2, 0.5, 0.8\}$, and $\theta = 1$.

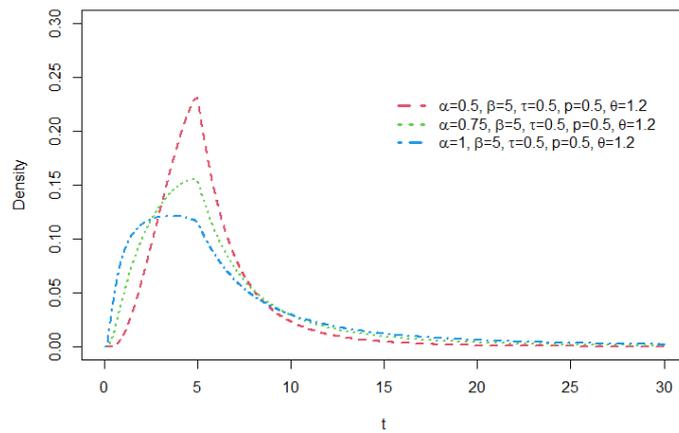


Figure 6. Density functions of GCR2-GG distribution with $\alpha \in \{0.5, 0.75, 1\}$, $\beta = 5$, $\tau = 0.5$, $p = 0.5$, and $\theta = 1.2$.

Example 3 (GCR2-LG distribution). The (symmetric) logistic (LG) distribution has the pdf

$$g_{LG}(x) = \frac{e^{-x/s}}{s(1 + e^{-x/s})^2}.$$

For parameter identification, we set $s = \frac{\sqrt{3}}{\pi}$, and thus, the resulting GCR2-LG distribution has one less parameter than the GCR2-t and GCR2-GG distributions. The pdf of the GCR2-LG distribution is given as

$$f_{GCR2}^{LG}(x; \alpha, \beta, \tau, p) = \left(\frac{2\pi\tau}{\sqrt{3}\alpha\beta} \right) \left[p \left(\frac{\beta}{x} \right)^{\tau+1} + q \left(\frac{x}{\beta} \right)^{\tau-1} \right] \frac{\exp \left\{ -\frac{\pi}{\sqrt{3}\alpha} \left[\left(\frac{x}{\beta} \right)^\tau - \left(\frac{\beta}{x} \right)^\tau \right] \right\}}{\left(1 + \exp \left\{ -\frac{\pi}{\sqrt{3}\alpha} \left[\left(\frac{x}{\beta} \right)^\tau - \left(\frac{\beta}{x} \right)^\tau \right] \right\} \right)^2}.$$

Due to the fact that the tail of the logistic distribution is heavier than that of normal distribution, the corresponding GCR2-LG distribution can be an effective model for datasets with thin to moderately heavy tails.

Figures 7–9 illustrate the shapes of the GCR2-LG density function for a few sets of prescribed parameter values.

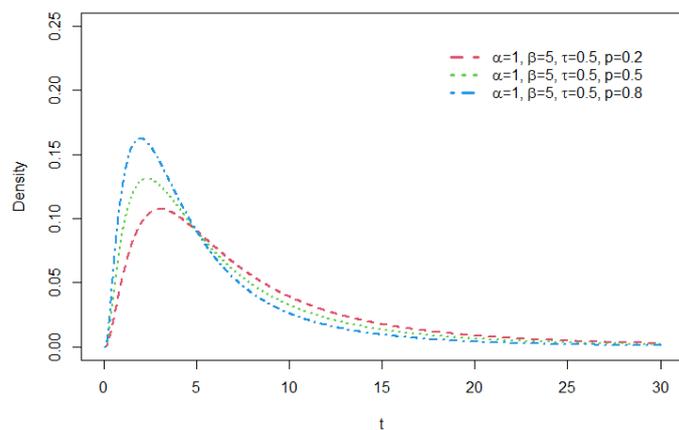


Figure 7. Density functions of GCR2-LG distribution with $\alpha = 1$, $\beta = 5$, $\tau = 0.5$ and $p \in \{0.2, 0.5, 0.8\}$.

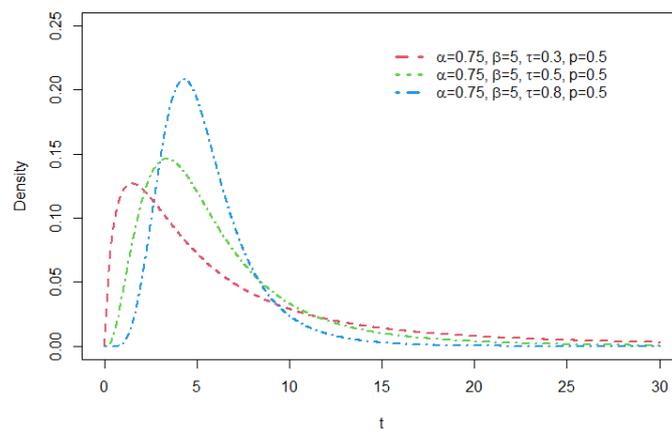


Figure 8. Density functions of GCR2-LG distribution with $\alpha = 0.75$, $\beta = 5$, $\tau \in \{0.5, 0.75, 1\}$ and $p = 0.5$.

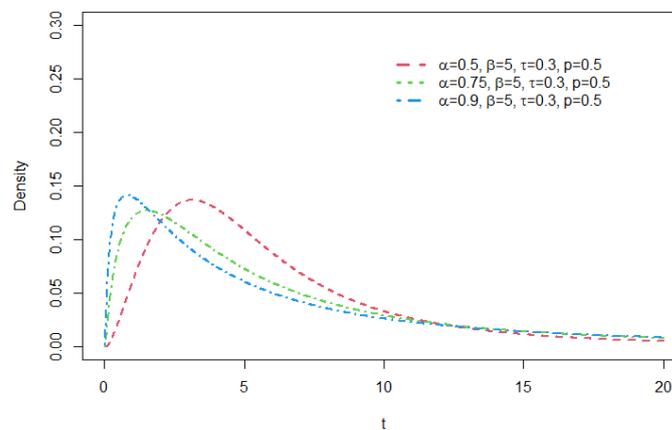


Figure 9. Density functions of GCR2-LG distribution with $\alpha \in \{0.5, 0.75, 0.9\}$, $\beta = 5$, $\tau = 0.3$ and $p = 0.5$.

As can be seen from Figures 1–9, various density shapes can be formed depending on the distributional features of the base-density function and the values of shape parameters. For each specific GCR2 distribution family, the right tail of the distribution becomes heavier as the mixture weight parameter p becomes smaller. The larger the shape parameter α is, the heavier the tail becomes, and the opposite is true for the shape parameter τ . For details of tail properties of the GCR2 distribution, see Theorem 1 and Theorem 2 in [12].

3. Bivariate GCR2 Distribution

In this section, we introduce a bivariate distribution with GCR2 marginals and study its key theoretical properties. Formally, the bivariate Type-II generalized crack (BVGCR2) distribution is defined as follows.

Definition 1. A pair of random variables $\mathbf{T} = (T_1, T_2)$ has a bivariate Type-II generalized crack distribution with base density g and parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, $\boldsymbol{\beta} = (\beta_1, \beta_2)$, $\boldsymbol{\tau} = (\tau_1, \tau_2)$ and $\mathbf{p} = (p_{11}, p_{12}, p_{21}, p_{22})$, denoted as $BVGCR2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\tau}, \mathbf{p}; g)$, if and only if, its joint pdf is given as follows:

$$\begin{aligned}
 f_{BVGCR2}(t_1, t_2; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\tau}, \mathbf{p}; g) &= p_{11}f_{GCR2}^{(1)}(t_1; \alpha_1, \beta_1, \tau_1; g)f_{GCR2}^{(1)}(t_2; \alpha_2, \beta_2, \tau_2; g) \\
 &+ p_{12}f_{GCR2}^{(1)}(t_1; \alpha_1, \beta_1, \tau_1; g)f_{GCR2}^{(2)}(t_2; \alpha_2, \beta_2, \tau_2; g) \\
 &+ p_{21}f_{GCR2}^{(2)}(t_1; \alpha_1, \beta_1, \tau_1; g)f_{GCR2}^{(1)}(t_2; \alpha_2, \beta_2, \tau_2; g) \\
 &+ p_{22}f_{GCR2}^{(2)}(t_1; \alpha_1, \beta_1, \tau_1; g)f_{GCR2}^{(2)}(t_2; \alpha_2, \beta_2, \tau_2; g), \quad t_1, t_2 > 0,
 \end{aligned}$$

where the mixture weight parameters satisfy $0 \leq p_j \leq 1, j \in J := \{11, 12, 21, 22\}$, $\sum_{j \in J} p_j = 1$, and, for each $i \in \{1, 2\}$,

$$f_{\text{GCR2}}^{(1)}(t_i; \alpha_i, \beta_i, \tau_i; g) = \left(\frac{2\tau_i}{\alpha_i\beta_i}\right) \left(\frac{\beta_i}{t_i}\right)^{\tau_i+1} g \left[\frac{1}{\alpha_i} \left\{ \left(\frac{t_i}{\beta_i}\right)^{\tau_i} - \left(\frac{\beta_i}{t_i}\right)^{\tau_i} \right\} \right],$$

and

$$f_{\text{GCR2}}^{(2)}(t_i; \alpha_i, \beta_i, \tau_i; g) = \left(\frac{2\tau_i}{\alpha_i\beta_i}\right) \left(\frac{t_i}{\beta_i}\right)^{\tau_i-1} g \left[\frac{1}{\alpha_i} \left\{ \left(\frac{t_i}{\beta_i}\right)^{\tau_i} - \left(\frac{\beta_i}{t_i}\right)^{\tau_i} \right\} \right].$$

Clearly, $\text{BVGCR2}(\alpha, \beta, \tau, p; g)$ is a mixture of four combinations of independent bivariate distributions. It is easy to see that the marginal distributions of T_1 and T_2 are $\text{GCR2}(\alpha_1, \beta_1, \tau_1, p_1 = p_{11} + p_{12}; g)$ and $\text{GCR2}(\alpha_2, \beta_2, \tau_2, p_2 = p_{11} + p_{21}; g)$, respectively. Please note that for simplicity, we assume that the marginal distributions are built on the same base-density function g , but any model parameters involved in g are distinct for each marginal.

3.1. Conditional Distribution

The conditional distribution is useful in simulating a pair of random variables from the BVGCR2 distribution. From the following relationship between the two mixture components of the GCR2 distribution,

$$\frac{f_{\text{GCR2}}^{(2)}(t; \alpha, \beta, \tau; g)}{f_{\text{GCR2}}^{(1)}(t; \alpha, \beta, \tau; g)} = \left(\frac{t}{\beta}\right)^{2\tau},$$

the conditional density of T_2 given $T_1 = t_1$, denoted as $f_{2|1}(t_2|t_1; g)$, can be expressed as

$$\begin{aligned} f_{2|1}(t_2|t_1; g) &:= \frac{p_{11}f^{(1)}(t_1; g)f^{(1)}(t_2; g) + p_{12}f^{(1)}(t_1; g)f^{(2)}(t_2; g)}{p_1f^{(1)}(t_1; g) + q_1f^{(2)}(t_1; g)} \\ &+ \frac{p_{21}f^{(2)}(t_1; g)f^{(1)}(t_2; g) + p_{22}f^{(2)}(t_1; g)f^{(2)}(t_2; g)}{p_1f^{(1)}(t_1; g) + q_1f^{(2)}(t_1; g)} \\ &= \frac{p_{11} + p_{21}\left(\frac{t_1}{\beta_1}\right)^{2\tau_1} f^{(1)}(t_2; g) + p_{12} + p_{22}\left(\frac{t_1}{\beta_1}\right)^{2\tau_1} f^{(2)}(t_2; g)}{p_1 + q_1\left(\frac{t_1}{\beta_1}\right)^{2\tau_1}} \\ &= \left(\frac{p_{11} + p_{21}\left(\frac{t_1}{\beta_1}\right)^{2\tau_1}}{p_1 + q_1\left(\frac{t_1}{\beta_1}\right)^{2\tau_1}}\right) f^{(1)}(t_2; g) + \left(\frac{p_{12} + p_{22}\left(\frac{t_1}{\beta_1}\right)^{2\tau_1}}{p_1 + q_1\left(\frac{t_1}{\beta_1}\right)^{2\tau_1}}\right) f^{(2)}(t_2; g). \end{aligned}$$

Here the subscripts and parameters in the density functions have been dropped for notational convenience. From the expression, we can see that the conditional distribution of T_2 given $T_1 = t_1$ is also $\text{GCR2}(\alpha_2, \beta_2, \tau_2, p_{2|1}; g)$ where

$$p_{2|1} = \frac{p_{11} + p_{21}(t_1/\beta_1)^{2\tau_1}}{p_1 + q_1(t_1/\beta_1)^{2\tau_1}}.$$

Using the conditional distribution, one can easily simulate a pair of random variates (T_1, T_2) from the BVGCR2 model; first simulate T_1 from the $\text{GCR2}(\alpha_1, \beta_1, \tau_1, p_1; g)$ using the acceptance-rejection method as given in [12], and then simulate, T_2 from the conditional distribution $\text{GCR2}(\alpha_2, \beta_2, \tau_2, p_{2|1}; g)$.

3.2. Dependence Measures

In this section, we derive expressions for Spearman’s rho and Kendall’s tau of BVGCR2 random variables.

3.2.1. Spearman’s Rho

Spearman’s rho is a commonly used measure of dependence between two random variables. Due to the invariance under monotone transformations, Spearman’s rho provides a broad interpretation of the dependence structure for any bivariate distributions. With marginal densities $f_{T_1}(t_1)$ and $f_{T_2}(t_2)$ for the random variables T_1 and T_2 , respectively, and the joint distribution $F(t_1, t_2)$ on $(t_1, t_2) \in \mathbb{R}_+^2$, the (population version) Spearman’s rho is defined as

$$\begin{aligned} \rho_s &= \frac{E(U_1 U_2) - E(U_1)E(U_2)}{\sqrt{\text{Var}(U_1)}\sqrt{\text{Var}(U_2)}} \\ &= 12 \int_0^\infty \int_0^\infty F(t_1, t_2) f_{T_1}(t_1) f_{T_2}(t_2) dt_1 dt_2 - 3, \end{aligned} \tag{7}$$

where $U_1 = F_{T_1}(T_1)$ and $U_2 = F_{T_2}(T_2)$ are uniform random variables, i.e., Spearman’s rho is Pearson’s correlation between transformations of the original random variables into standard uniform marginals. The following provides an expression for Spearman’s rho of a pair of random variables with a BVGCR2 distribution.

Proposition 1. *Suppose $(T_1, T_2) \sim \text{BVGCR2}(\alpha, \beta, \tau, p; g)$. Then, Spearman’s rho between T_1 and T_2 is expressed as*

$$\rho_s = 48(p_{11}p_{22} - p_{12}p_{21})\gamma_1\gamma_2, \tag{8}$$

where

$$\gamma_i = \int_{-\infty}^\infty \frac{t}{\sqrt{t^2 + 4/\alpha_i^2}} G_i(t) g_i(t) dt, \quad i = 1, 2.$$

Proof. See Appendix A. \square

3.2.2. Kendall’s Tau

Kendall’s tau is a measure of association (concordance/discordance) between two random variables. Formally, for the random variables T_1 and T_2 with the joint distribution $F(t_1, t_2)$ on $(t_1, t_2) \in \mathbb{R}_+^2$, the (population version) Kendall’s tau is defined as

$$\begin{aligned} \tau_k &= P((T_1 - T'_1)(T_2 - T'_2) > 0) - P((T_1 - T'_1)(T_2 - T'_2) < 0) \\ &= 4 \int_0^\infty \int_0^\infty F(t_1, t_2) f(t_1, t_2) dt_1 dt_2 - 1, \end{aligned} \tag{9}$$

where the pair (T'_1, T'_2) has the joint distribution F and is independent to (T_1, T_2) .

The following provides an expression for Kendall’s tau of random variables with a BVGCR2 distribution.

Proposition 2. *Suppose $(T_1, T_2) \sim \text{BVGCR2}(\alpha, \beta, \tau, p; g)$. Then, Kendall’s tau between T_1 and T_2 is expressed as*

$$\tau_k = 32(p_{11}p_{22} - p_{12}p_{21})\gamma_1\gamma_2, \tag{10}$$

where

$$\gamma_i = \int_{-\infty}^\infty \frac{t}{\sqrt{t^2 + 4/\alpha_i^2}} G_i(t) g_i(t) dt, \quad i = 1, 2.$$

Proof. See Appendix B. \square

Remark 1. *It is important to note that*

$$0 \leq \gamma \leq \frac{1}{4}. \tag{11}$$

These inequalities can be proven using the symmetry of $g(t)$, $G(-t) = 1 - G(t)$ and integration by parts. Specifically,

$$\begin{aligned} \gamma &= \int_{-\infty}^{\infty} \frac{t}{\sqrt{t^2 + 4/\alpha^2}} G(t)g(t) dt \\ &= \int_{-\infty}^0 \frac{t}{\sqrt{t^2 + 4/\alpha^2}} G(t)g(t) dt + \int_0^{\infty} \frac{t}{\sqrt{t^2 + 4/\alpha^2}} G(t)g(t) dt \\ &= \int_0^{\infty} \frac{-s}{\sqrt{s^2 + 4/\alpha^2}} G(-s)g(-s) ds + \int_0^{\infty} \frac{t}{\sqrt{t^2 + 4/\alpha^2}} G(t)g(t) dt \\ &= -\int_0^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} (1 - G(s))g(s) ds + \int_0^{\infty} \frac{t}{\sqrt{t^2 + 4/\alpha^2}} G(t)g(t) dt \\ &= -\int_0^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds + 2 \int_0^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} G(s)g(s) ds \\ &= \int_0^{\infty} (2G(s) - 1) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds. \end{aligned}$$

Clearly $2G(s) - 1 \geq 0$ and $0 \leq \frac{s}{\sqrt{s^2 + 4/\alpha^2}} \leq 1$, for $0 \leq s < \infty$. Then, by the fact $\int_0^{\infty} G(s)g(s)ds = 3/8$, we have

$$0 \leq \int_0^{\infty} (2G(s) - 1) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \leq \int_0^{\infty} (2G(s) - 1)g(s)ds = 2\left(\frac{3}{8}\right) - \frac{1}{2} = \frac{1}{4}.$$

From this and by Equations (8) and (10), we obtain the following bounds for ρ_s and τ_k :

$$|\rho_s| \leq \frac{3}{4}, \quad |\tau_k| \leq \frac{1}{2},$$

Note that ρ_s and τ_k become the maximum when $p_{11} = p_{22} = 1/2$ and γ attains its maximum of $1/4$.

We also remark that if $p_{11}p_{22} = p_{12}p_{21}$, and thus $\rho_s = \tau_k = 0$, the joint density of the BVGCR2 model can be expressed as a product of two GCR2 marginals, i.e., the two random variables are independent.

3.3. Tail Independence

As remarked in the previous section, the dependency measures of the proposed BVGCR2 model are bounded, and thus, the model may not be suitable for cases where extreme dependency is required, i.e., market turmoil. Here, we further investigate the tail dependence of the BVGCR2 model in terms of the upper-tail dependence, which is defined as

$$\begin{aligned} \lambda_U &= \lim_{q \rightarrow 1} \Pr[T_1 > F_{T_1}^{-1}(q) | T_2 > F_{T_2}^{-1}(q)] = \lim_{q \rightarrow 1} \frac{\Pr[T_1 > F_{T_1}^{-1}(q), T_2 > F_{T_2}^{-1}(q)]}{1 - q} \\ &= \lim_{q \rightarrow 1} \frac{1 - 2q + F_{T_1, T_2}(F_{T_1}^{-1}(q), F_{T_2}^{-1}(q))}{1 - q} \\ &= 2 - \lim_{q \rightarrow 1} \frac{d}{dq} F_{T_1, T_2}(F_{T_1}^{-1}(q), F_{T_2}^{-1}(q)), \quad (12) \end{aligned}$$

where the last equality is held by L'Hopital's rule. In fact,

$$\begin{aligned} & \frac{d}{dq} F_{T_1, T_2}(F_{T_1}^{-1}(q), F_{T_2}^{-1}(q)) \\ &= p_{11} \frac{d}{dq} \left\{ F_{T_1}^{(1)}(F_{T_1}^{-1}(q)) F_{T_2}^{(1)}(F_{T_2}^{-1}(q)) \right\} + p_{12} \frac{d}{dq} \left\{ F_{T_1}^{(1)}(F_{T_1}^{-1}(q)) F_{T_2}^{(2)}(F_{T_2}^{-1}(q)) \right\} \\ & \quad + p_{21} \frac{d}{dq} \left\{ F_{T_1}^{(2)}(F_{T_1}^{-1}(q)) F_{T_2}^{(1)}(F_{T_2}^{-1}(q)) \right\} + p_{22} \frac{d}{dq} \left\{ F_{T_1}^{(2)}(F_{T_1}^{-1}(q)) F_{T_2}^{(2)}(F_{T_2}^{-1}(q)) \right\} \\ &= p_{11} \left\{ F_{T_2}^{(1)}(F_{T_2}^{-1}(q)) \frac{f_{T_1}^{(1)}(F_{T_1}^{-1}(q))}{f_{T_1}(F_{T_1}^{-1}(q))} + F_{T_1}^{(1)}(F_{T_1}^{-1}(q)) \frac{f_{T_2}^{(1)}(F_{T_2}^{-1}(q))}{f_{T_2}(F_{T_2}^{-1}(q))} \right\} \\ & \quad + p_{12} \left\{ F_{T_2}^{(2)}(F_{T_2}^{-1}(q)) \frac{f_{T_1}^{(1)}(F_{T_1}^{-1}(q))}{f_{T_1}(F_{T_1}^{-1}(q))} + F_{T_1}^{(1)}(F_{T_1}^{-1}(q)) \frac{f_{T_2}^{(2)}(F_{T_2}^{-1}(q))}{f_{T_2}(F_{T_2}^{-1}(q))} \right\} \\ & \quad + p_{21} \left\{ F_{T_2}^{(1)}(F_{T_2}^{-1}(q)) \frac{f_{T_1}^{(2)}(F_{T_1}^{-1}(q))}{f_{T_1}(F_{T_1}^{-1}(q))} + F_{T_1}^{(2)}(F_{T_1}^{-1}(q)) \frac{f_{T_2}^{(1)}(F_{T_2}^{-1}(q))}{f_{T_2}(F_{T_2}^{-1}(q))} \right\} \\ & \quad + p_{22} \left\{ F_{T_2}^{(2)}(F_{T_2}^{-1}(q)) \frac{f_{T_1}^{(2)}(F_{T_1}^{-1}(q))}{f_{T_1}(F_{T_1}^{-1}(q))} + F_{T_1}^{(2)}(F_{T_1}^{-1}(q)) \frac{f_{T_2}^{(2)}(F_{T_2}^{-1}(q))}{f_{T_2}(F_{T_2}^{-1}(q))} \right\}. \end{aligned}$$

Since $f_{T_1}(t_1) = (p_{11} + p_{12})f^{(1)}(t_1) + (p_{21} + p_{22})f^{(2)}(t_1)$, $f_{T_2}(t_2) = (p_{11} + p_{21})f^{(1)}(t_2) + (p_{12} + p_{22})f^{(2)}(t_2)$, and $f^{(2)}(t_i) = (t_i/\beta_i)^{2\tau_i}, i = 1, 2$, we have

$$\begin{aligned} \lim_{q \rightarrow 1} F_{T_i}^{(1)}(F_{T_i}^{-1}(q)) &= 1, \quad \lim_{q \rightarrow 1} \frac{f_{T_i}^{(1)}(F_{T_i}^{-1}(q))}{f_{T_i}(F_{T_i}^{-1}(q))} = 0, \quad i = 1, 2, \\ \lim_{q \rightarrow 1} \frac{f_{T_1}^{(2)}(F_{T_1}^{-1}(q))}{f_{T_1}(F_{T_1}^{-1}(q))} &= \frac{1}{p_{21} + p_{22}}, \quad \lim_{q \rightarrow 1} \frac{f_{T_2}^{(2)}(F_{T_2}^{-1}(q))}{f_{T_2}(F_{T_2}^{-1}(q))} = \frac{1}{p_{12} + p_{22}}. \end{aligned}$$

Therefore,

$$\lim_{q \rightarrow 1} \frac{d}{dq} F_{T_1, T_2}(F_{T_1}^{-1}(q), F_{T_2}^{-1}(q)) = \frac{p_{12}}{p_{12} + p_{22}} + \frac{p_{21}}{p_{21} + p_{22}} + \frac{p_{22}}{p_{21} + p_{22}} + \frac{p_{22}}{p_{12} + p_{22}} = 2.$$

Then, by Equation (12), we have $\lambda_U = 2 - \lim_{q \rightarrow 1} \frac{d}{dq} F_{T_1, T_2}(F_{T_1}^{-1}(q), F_{T_2}^{-1}(q)) = 0$, i.e., the BVGCR2 model lacks upper-tail dependence.

4. Model Parameter Estimation

In this section, we discuss parameter estimation for the bivariate Type-II generalized crack distribution using the expectation-maximization algorithm. We briefly review the EM algorithm in a general setting and provide a specific application to the BVGCR2 models.

4.1. Maximum Likelihood Estimation

Suppose (X_1, X_2, \dots, X_n) is an independent and identically distributed random sample drawn from a density $f(x|\theta)$. Likelihood function is given as

$$L(\theta) = f(x|\theta) = \prod_{i=1}^n f(x_i|\theta).$$

The maximum likelihood estimation (MLE) aims to find the parameter estimate that maximizes the likelihood function, or equivalently, the log-likelihood function:

$$\hat{\theta}_{MLE} = \arg \max_{\theta} \{ \log f(x|\theta) \}.$$

In the presence of latent (hidden/unobserved) values, however, the direct maximization method often does not provide reliable estimates; hence, other alternative methods are usually considered, and one such method is the expectation-maximization algorithm.

4.2. Expectation-Maximization Algorithm

For simple mixture models with a small number of mixture components, the direct optimization of the log-likelihood function may be used to obtain the maximum likelihood estimates of the model parameters. However, the direct optimization often fails to converge when the number of mixture components is large relative to the sample size.

The expectation-maximization (EM) algorithm [19] is an efficient iterative procedure used to compute the maximum likelihood estimates in the presence of missing or hidden data. When applied to finite mixture model settings, the expectation step renders the separation of the mixture weight parameters from other model parameters for optimization, and the maximization step gives an explicit solution for updating the mixture weights. Due to this, the algorithm is the most widely used maximum likelihood estimation technique for finite mixture models. For details on recent developments of EM-type algorithms for the Poisson mixture model and multivariate Gaussian mixture models in complex data settings, see [20,21], respectively.

Here, we briefly review the general form of the EM algorithm. Suppose we have observed data $\underline{x} = (x_1, x_2, \dots, x_n)$ with density $p(\underline{x}|\theta)$ and some latent (hidden/unobserved data) $\underline{z} = (z_1, z_2, \dots, z_n)$ with density $p(\underline{z}|\theta)$. The density of the complete data is denoted by $p(\underline{x}, \underline{z}|\theta)$. The goal of the EM algorithm is to find the MLE, i.e., the maximum of the observed data likelihood function,

$$L(\theta) = \sum_{\underline{z}} p(\underline{x}, \underline{z}|\theta).$$

The EM algorithm proceeds by iterating between the following two steps;

- E-Step: This step calculates the expectation of the likelihood with respect to the conditional distribution of \underline{z} given \underline{x} and the initial parameter estimate $\theta^{(m)}$, i.e.,

$$Q(\theta|\theta^{(m)}) = E_{Z|\underline{x}, \theta^{(m)}} [\log p(\underline{x}, Z|\theta)].$$

- M-Step: Choose $\hat{\theta}^{(m+1)} = \arg \max_{\theta} Q(\theta|\theta^{(m)})$.

Lemma 1. *The EM algorithm improves $Q(\theta|\theta^{(m)})$. That is, if $Q(\theta^{(m+1)}|\theta^{(m)}) \geq Q(\theta^{(m)}|\theta^{(m)})$, then $l(\theta^{(m+1)}) \geq l(\theta^{(m)})$.*

Proof. See Appendix C. \square

We now provide the EM algorithm for the estimation of the parameters in the BVGCR2 model. Let $\mathbf{t} = (t_1, t_2, \dots, t_n)$ be a random sample drawn from a BVGCR2($\alpha, \beta, \tau, p; g$) distribution, where $t_i = (t_{i1}, t_{i2})$ is a pair of observations for each $i = 1, \dots, n$, and the base density g may have its own parameter(s) such as ν involved in the Student's t density. We denote the vector of parameters involved in the base densities by $\theta = (\theta_1, \theta_2)$. Letting $\gamma = (\alpha, \beta, \tau, p, \theta)$ be the vector of all model parameters, the likelihood function based on the incomplete data is

$$L(\gamma|\mathbf{t}) = \prod_{i=1}^n \sum_{j \in J} q_j f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau, \theta; g)$$

where $J = \{11, 12, 21, 22\}$ is the set of indexes and

$$f_j(t_i; \gamma; g) = \begin{cases} f_{GCR2}^{(1)}(t_{i1}; \alpha_1, \beta_1, \tau_1; \theta_1; g) f_{GCR2}^{(1)}(t_{i2}; \alpha_2, \beta_2, \tau_2; \theta_2; g), & j = 11 \\ f_{GCR2}^{(1)}(t_{i1}; \alpha_1, \beta_1, \tau_1; \theta_1; g) f_{GCR2}^{(2)}(t_{i2}; \alpha_2, \beta_2, \tau_2; \theta_2; g), & j = 12 \\ f_{GCR2}^{(2)}(t_{i1}; \alpha_1, \beta_1, \tau_1; \theta_1; g) f_{GCR2}^{(1)}(t_{i2}; \alpha_2, \beta_2, \tau_2; \theta_2; g), & j = 21 \\ f_{GCR2}^{(2)}(t_{i1}; \alpha_1, \beta_1, \tau_1; \theta_1; g) f_{GCR2}^{(2)}(t_{i2}; \alpha_2, \beta_2, \tau_2; \theta_2; g), & j = 22. \end{cases}$$

On the other hand, letting $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ be a set of latent variables where $\Pr(Z_i = j) = q_j, j = 1, 2$, for each $i \in \{1, \dots, n\}$, the likelihood function based on the (augmented) complete data is

$$L(\gamma | \mathbf{t}, \mathbf{Z}) = \prod_{i=1}^n \prod_j [q_j f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g)]^{I(Z_i=j)}$$

where $I(\cdot)$ denotes the indicator function.

Let $\gamma^{(m)} = (\alpha^{(m)}, \beta^{(m)}, \tau^{(m)}, \mathbf{p}^{(m)}, \theta^{(m)})$ denote the current estimate of γ after m -th iteration of the EM algorithm. Then, by Bayes' theorem, we have

$$p_{ij}^{(m)} := \Pr(Z_i = j | t_i, \gamma^{(m)}) = \frac{q_j^{(m)} f_j(t_{i1}, t_{i2}; \alpha^{(m)}, \beta^{(m)}, \tau^{(m)}; \theta^{(m)})}{\sum_{j \in J} q_j^{(m)} f_j(t_{i1}, t_{i2}; \alpha^{(m)}, \beta^{(m)}, \tau^{(m)}; \theta^{(m)})}$$

The expectation step (E-step) of the EM algorithm follows.

$$\begin{aligned} E_{Z | t, \gamma^{(m)}} [\log L(\gamma | \mathbf{t}, \mathbf{Z})] &= \sum_{i=1}^n E_{Z | t, \gamma^{(m)}} \left(\log \prod_j [q_j f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g)]^{I(Z_i=j)} \right) \\ &= \sum_{i=1}^n E_{Z | t, \gamma^{(m)}} \left(\sum_{j \in J} I(Z_i = j) \log [q_j f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g)] \right) \\ &= \sum_{i=1}^n \sum_{j \in J} \Pr(Z_i = j | t_i, \gamma^{(m)}) \log [q_j f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g)] \\ &= \sum_{i=1}^n \sum_{j \in J} p_{ij}^{(m)} \log [q_j f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g)] \\ &= \sum_{i=1}^n \sum_{j \in J} p_{ij}^{(m)} \log q_j + \sum_{i=1}^n \sum_{j \in J} p_{ij}^{(m)} \log f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g). \end{aligned} \tag{13}$$

The maximization step (M-step) finds the updated parameter estimates that maximize the objective function (13), which separates q_j and the other parameters $(\alpha, \beta, \tau; \theta)$. The update of q_j can be dealt with separately by applying the method of Lagrange multiplier. The updated estimate is

$$q_j^{(m+1)} = \arg \max_q \left\{ \sum_{i=1}^n \sum_{j \in J} p_{ij}^{(m)} \log q_j \right\} = \left(\frac{1}{n} \sum_{i=1}^n p_{i,11}^{(m)}, \dots, \frac{1}{n} \sum_{i=1}^n p_{i,22}^{(m)} \right)'$$

The updated estimates $\alpha^{(m+1)}, \beta^{(m+1)}, \tau^{(m+1)}$ and $\theta^{(m+1)}$ are the maximizers of the objective function

$$\sum_{i=1}^n \sum_{j \in J} p_{ij}^{(m)} \log f_j(t_{i1}, t_{i2}; \alpha, \beta, \tau; \theta; g) = (Q_1) + (Q_2),$$

where

$$(Q_1) = \sum_{i=1}^n \left[(p_{i,11}^{(m)} + p_{i,12}^{(m)}) \log f_{GCR2}^{(1)}(t_{i1}; \alpha_1, \beta_1, \tau_1; \theta_1; g) + (p_{i,21}^{(m)} + p_{i,22}^{(m)}) \log f_{GCR2}^{(2)}(t_{i1}; \alpha_1, \beta_1, \tau_1; \theta_1; g) \right],$$

and

$$(Q_2) = \sum_{i=1}^n \left[(p_{i,11}^{(m)} + p_{i,21}^{(m)}) \log f_{GCR2}^{(1)}(t_{i2}; \alpha_2, \beta_2, \tau_2; \theta_2; g) + (p_{i,12}^{(m)} + p_{i,22}^{(m)}) \log f_{GCR2}^{(2)}(t_{i2}; \alpha_2, \beta_2, \tau_2; \theta_2; g) \right].$$

That is, the maximization for $(\alpha_1, \beta_1, \tau_1; \theta_1)$ can proceed separately from that for $(\alpha_2, \beta_2, \tau_2; \theta_2)$ and thus, the dimensionality of the optimization problem is reduced significantly.

5. Applications

In [12], the usefulness of the univariate GCR2 models for heavy-tailed data modeling has been demonstrated through an application with a real loss dataset. In this section, we fit several bivariate Type-II generalized crack distributions on a real catastrophic loss dataset compiled from the International Disaster Database (EM-DAT, www.emdat.be, accessed on 22 November 2024). Specifically, each observation in the dataset is composed of two variables: ‘Meteo’ and ‘Hydro’. Marginally, ‘Meteo’ is a quarterly time series of (estimated) losses from meteorological disasters such as storms and extreme temperatures, spanning from 1950 to 2022 in Asia, and ‘Hydro’ is a series of (estimated) losses due to hydrological disasters such as flood and landslide for the same geographical area and the time span. For bivariate model fitting, we remove the pairs with missing observations, resulting in a bivariate dataset with 166 observations. The losses are inflation-adjusted to be equivalent to the US dollar values in 2021.

Table 1 presents summary statistics of losses due to meteorological and hydrological disasters in Asia. The descriptive statistics, along with the histogram and the normal Q-Q plots (Figures 10 and 11), suggest that the marginal distributions of the two variables are both positively skewed and heavy-tailed. Many time series include some deterministic components such as long-term trends and seasonality. To isolate the deterministic components, each quarterly time series is decomposed under the multiplicative model assumption. The time series decompositions given in Figure 12 show the presence of strong seasonality in both time series and weak evidence of long-term trend. Since the proposed GCR2 models do not assume any deterministic seasonality, we deseasonalize each time series by dividing it by its estimated seasonal component.

Table 1. Descriptive summary statistics of the dataset (Unit: 100 million USD).

Data	<i>n</i>	Min	1st Qu.	Median	Mean	3rd Qu.	Max	Skewness	Kurtosis
Meteo	166	0.003	2.005	7.275	30.799	35.282	287.090	2.820	11.304
Hydro	166	0.007	1.692	12.172	45.538	53.330	621.207	4.014	23.188

Figure 13 gives the scatter plots of the two seasonally adjusted variables and their log-transformations. The figure shows some evidence of a dependent relationship between the two variables. The sample Spearman’s rho and Kendall’s tau of the variables are 0.425 and 0.291, respectively.

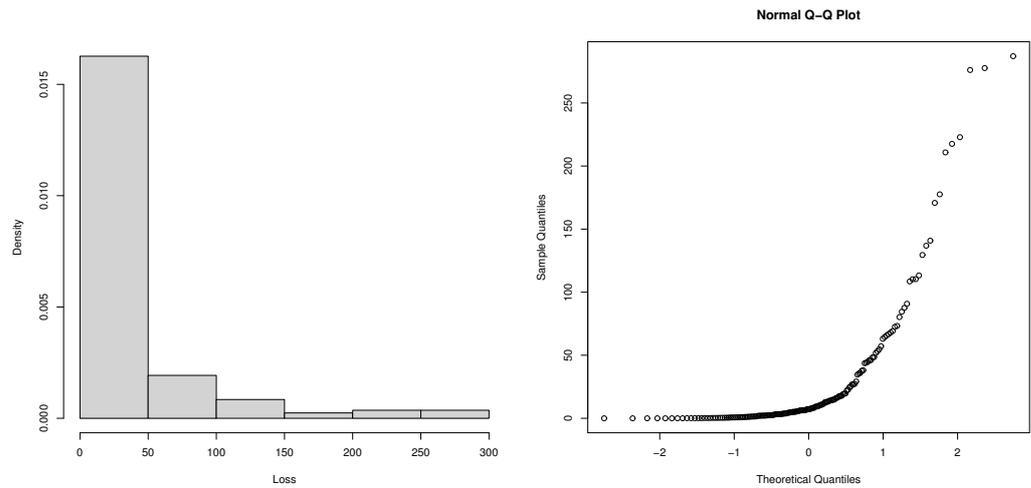


Figure 10. Histogram of Meteo loss dataset and its normal Q-Q plot.

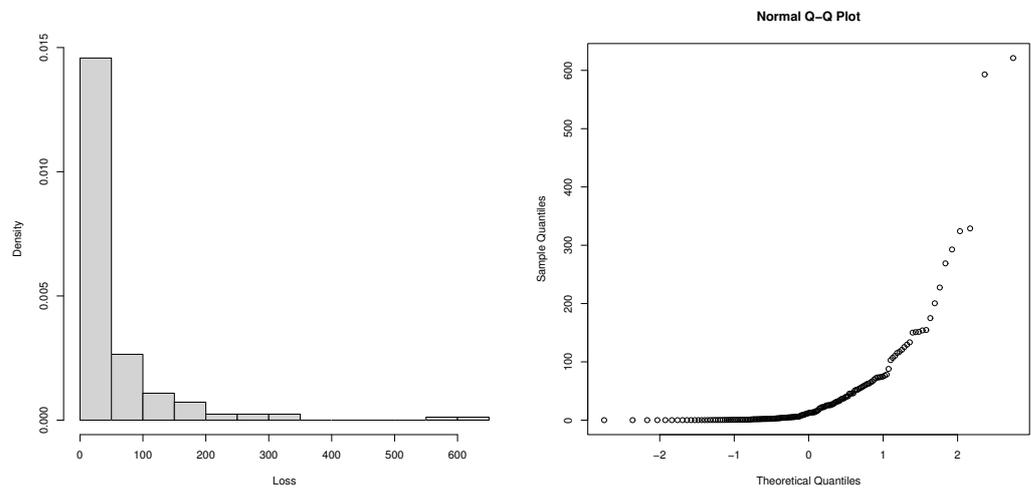


Figure 11. Histogram of Hydro loss dataset and its normal Q-Q plot.

We apply the EM algorithm described in Section 4.2 to fit the (seasonally adjusted) data using six specific bivariate models: GCR-GG, GCR- t , GCR-LG, GCR2-GG, GCR2- t , and GCR2-LG. For each model fitting, the EM algorithm requires initialization of the parameter values. To obtain a reliable result, we first fit the marginal models separately using the estimation method given in [12], and the fitted values of the marginal parameters, i.e., $(\alpha_i, \beta_i, \tau_i, \theta)$ for GCR models and $(\alpha_i, \beta_i, \tau_i, \theta)$, $i = 1, 2$, for the GCR2 models, are used for the initialization of the BVGCR (or BVGCR2) model parameters. For mixture weight parameter initialization, we use a large set of possible parameter values satisfying $p_{11} + p_{12} = p_1$ and $p_{11} + p_{21} = p_2$, where p_1 and p_2 are the mixture weight parameter estimates for the marginal models. The log-likelihood function values of the EM fits under the set of initializations are compared, and the one that gives the largest log-likelihood value is selected for the final fit. To compare the performance of BVGCR2 models with some other benchmark bivariate models commonly used in loss modeling, we further implement the maximum likelihood estimation of the following seven bivariate models: BVLNorm (bivariate lognormal), Clayton-LNorm (bivariate Clayton copula with lognormal marginals), Gumbel-LNorm (bivariate Gumbel copula with lognormal marginals), Frank-LNorm (bivariate Frank copula with lognormal marginals), Clayton-Pareto (bivariate Clayton copula with two-parameter Pareto marginals), Gumbel-Pareto (bivariate Gumbel copula with two-parameter Pareto marginals), and Frank-Pareto (bivariate Frank copula with two-parameter Pareto marginals) models.

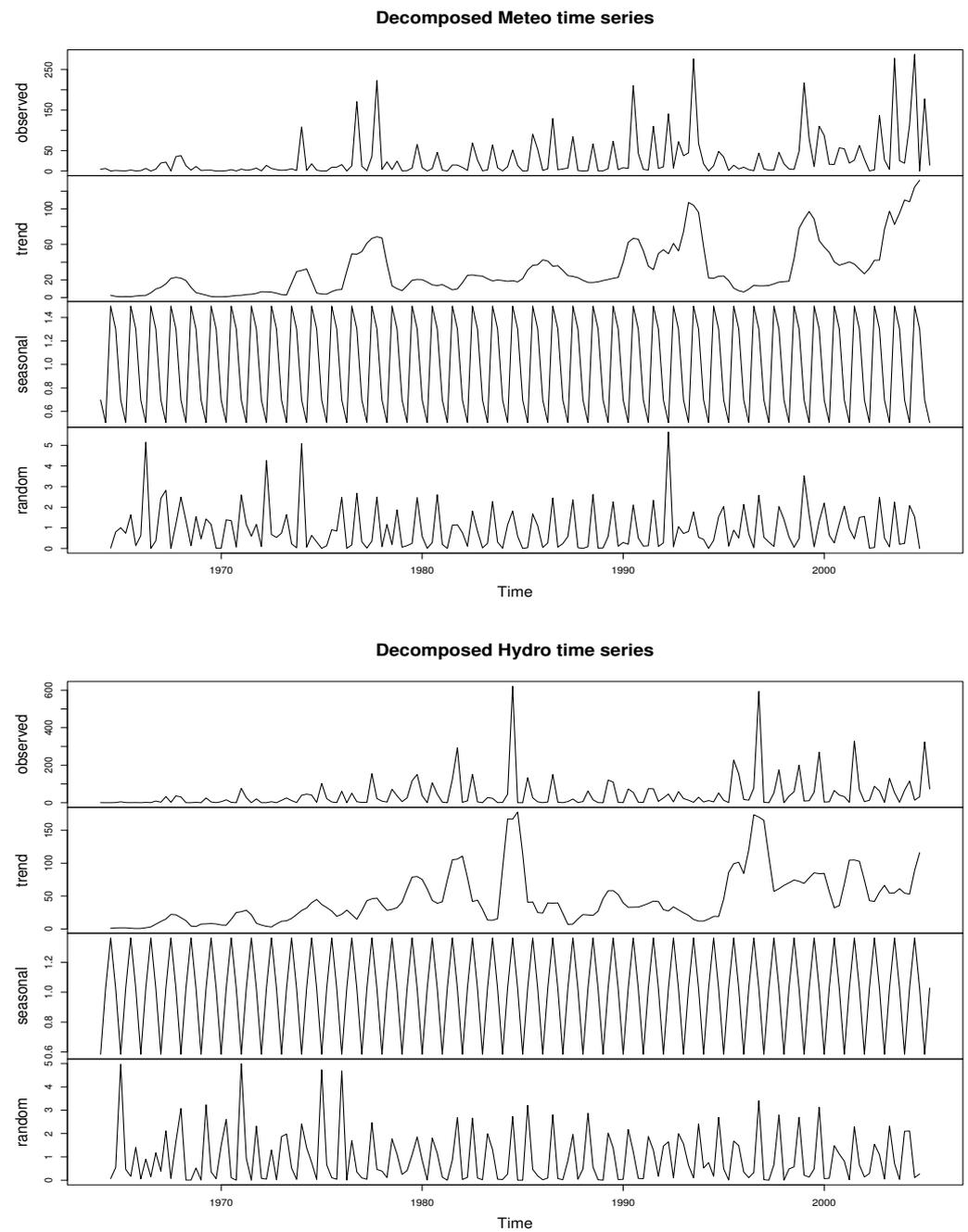


Figure 12. Multiplicative decompositions of Meteo and Hydro time series.

Since the number of estimated model parameters differs by model, we compare the fits of candidate models in terms of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), defined as

$$AIC = -2(\text{Log-Likelihood}) + 2k$$

$$BIC = -2(\text{Log-Likelihood}) + k \log n,$$

respectively, where k is the number of estimated parameters and n is the sample size. Among candidate models, the preferred model is the one with the smallest value of either of these criteria.

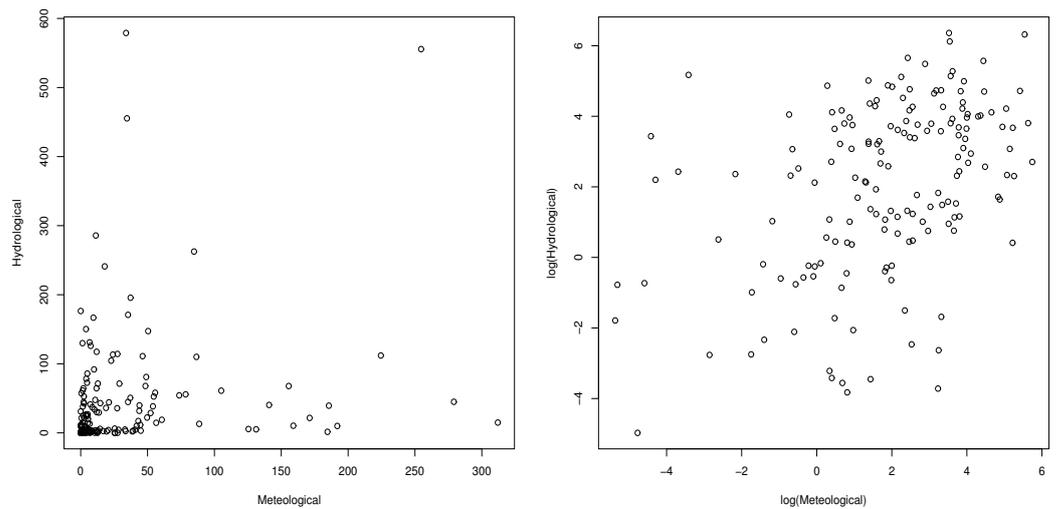


Figure 13. Scatter plots of the deseasonalized data and their log-transformations.

From Table 2, we see that, based on the Akaike information criterion, the fitted BVGCR2-GG model outperforms all the other alternative bivariate models. However, the BVGCR-LG model is preferred in terms of the Bayesian information criterion, which heavily penalizes complex models. Comparing the BVGCR-GG to the BVGCR2-GG, one can see that the BVGCR2-GG significantly improves model fitting due to the additional shape parameter.

Table 2. Model comparison with AIC and BIC.

Model	Log-Likelihood	AIC	BIC
BVGCR-GG	−1370.332	2758.664	2786.672
BVGCR- <i>t</i>	−1372.381	2762.762	2790.770
BVGCR-LG	−1372.066	2758.132	2779.916
BVGCR2-GG	−1367.346	2756.692	2790.924
BVGCR2- <i>t</i>	−1371.782	2765.564	2799.796
BVGCR2-LG	−1371.707	2761.414	2789.422
BVLNorm	−1387.839	2785.678	2796.126
Clayton-LNorm	−1393.159	2796.318	2811.878
Gumbel-LNorm	−1385.676	2781.352	2796.912
Frank-LNorm	−1386.976	2783.952	2799.512
Clayton-Pareto	−1405.479	2820.958	2836.518
Gumbel-Pareto	−1403.067	2816.134	2831.694
Frank-Pareto	−1405.029	2820.058	2835.618

Table 3 gives parameter estimates of the fitted marginal distributions and the bivariate GCR2-GG model.

Table 3. Parameter estimates of marginals and Bivariate GCR2-GG models.

	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\tau}_1$	$\hat{\theta}_1$	\hat{p}_1	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\tau}_2$	$\hat{\theta}_2$	\hat{p}_2	\hat{p}_{11}	\hat{p}_{12}	\hat{p}_{21}	\hat{p}_{22}
Meteo	20.075	0.915	0.737	0.623	0.094									
Hydro						11.152	1.964	0.669	0.809	0.189				
Bivariate	22.644	1.494	0.791	0.477		11.442	1.964	0.675	0.798		0.124	0.040	0.072	0.764

Please note that the values of $(p_{11} + p_{12})$ and $(p_{11} + p_{21})$ based on the fitted BVGCR2-GG model are larger than the corresponding estimates of p_1 and p_2 . Spearman’s rho and Kendall’s tau under the fitted BVGCR2-GG model are 0.173 and 0.115, respectively, which are lower than the empirical counterparts. This may be because the empirical data contains some spurious dependence due to the deterministic trends.

To test the statistical significance of the parameter estimates of the BVGCR2-GG model, we construct the bootstrap confidence intervals by computing the 2.5th and the 97.5th percentiles of the estimates based on the 100 bootstrap samples. Table 4 shows that all the parameter estimates on the original dataset fall in the corresponding 95% bootstrap confidence intervals.

Table 4. The 95% bootstrap confidence intervals for the parameters in Bivariate GCR2-GG model.

	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\tau}_1$	$\hat{\theta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\tau}_2$	$\hat{\theta}_2$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$
Original estimate	22.644	1.494	0.791	0.477	11.442	1.964	0.675	0.798	0.124	0.040	0.072	0.764
Bootstrap mean	27.673	1.524	0.735	0.583	14.189	2.065	0.680	0.988	0.131	0.037	0.067	0.766
2.5th percentile	7.031	0.762	0.542	0.347	2.356	1.500	0.322	0.418	0.059	0.000	0.000	0.687
97.5th percentile	52.149	2.252	0.951	1.013	42.898	2.803	0.913	3.027	0.192	0.080	0.147	0.843

Figure 14 gives the contour plots of the fitted BVGCR2-GG model, and that of the log-transformed random variables, and Figure 15 presents the scatter plots of the simulated random variables from the BVGCR2-GG model and their log-transformations. Comparing these plots with Figure 13, we can see that the fitted model explains the dependence structure of the empirical data well.

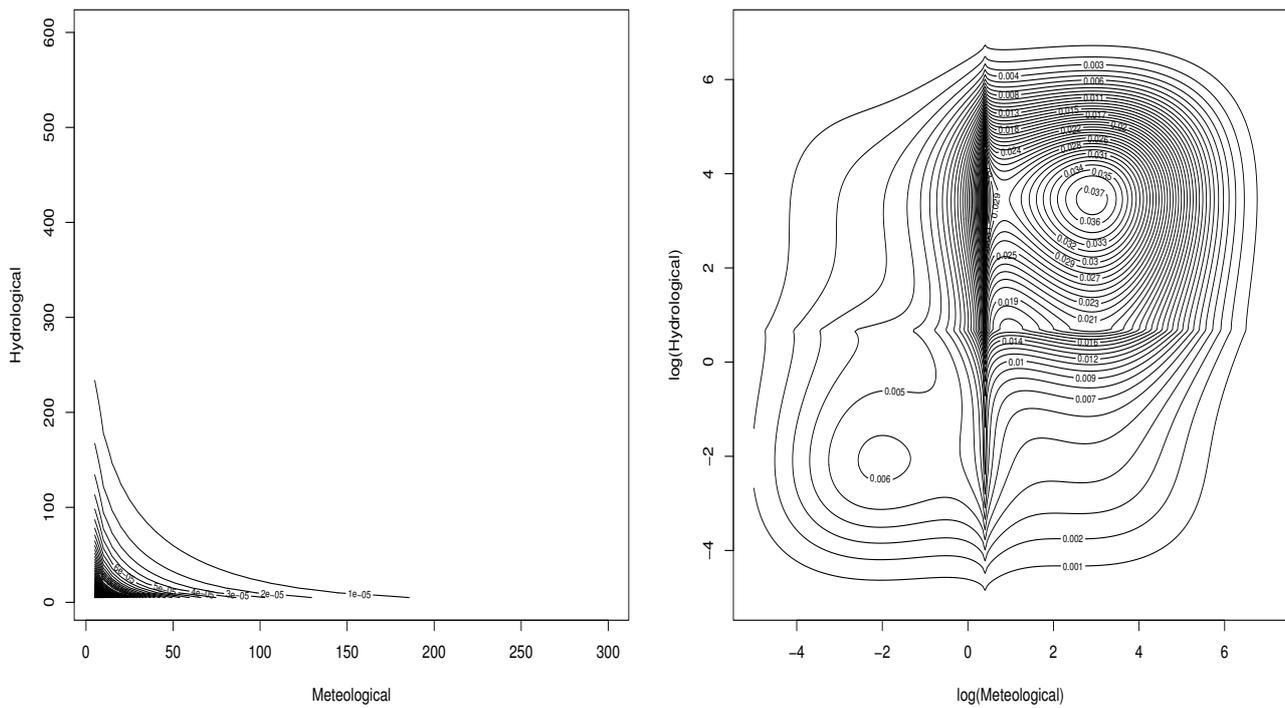


Figure 14. Contour plots of the fitted BVGCR2-GG density and the density of the log-transformed random variables.

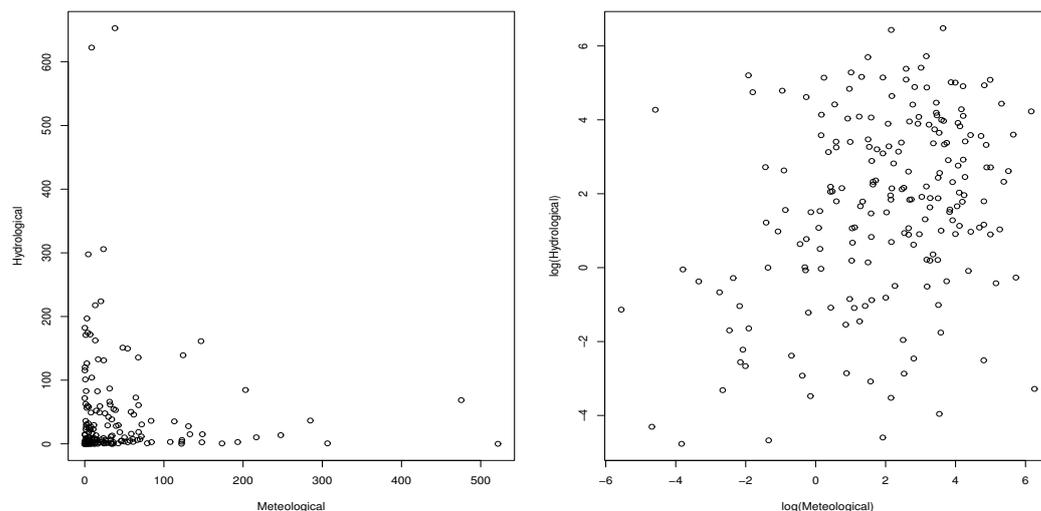


Figure 15. Scatter plots of simulated random variables from the fitted BVGCR2-GG model and their log-transformations.

6. Concluding Remarks

In this paper, we constructed a bivariate extension of the Type-II generalized crack distribution and studied a few specific examples of the bivariate GCR2 distributions based on the generalized Gaussian, Student's t , and logistic densities to demonstrate the applicability of the constructed model. Specifically, our main theoretical finding is that the level of dependence of the constructed BVGCR2 model in terms of Kendall's tau and Spearman's rho is a weak to medium association. The model fitting to catastrophic loss data showed that the fitted BVGCR2-GG model outperformed all the other alternative models based on the Akaike information criterion. Especially when compared to the BVGCR-GG model, the BVGCR2-GG model has shown a significant improvement due to the increased flexibility. With an appropriate choice of base-density function, the proposed BVGCR2 model can be effectively used for various applications that require a weak to moderate level of dependence.

With the lack of the upper-tail dependence, the bivariate GCR2 distributions may not be applicable for the situations where variables are (or assumed to be) asymptotically dependent in the upper tail, e.g., stress testing for market/credit portfolios. This limitation can be alleviated using a common parameter for both marginals and randomizing the parameter. For example, we may set $\tau = \tau_1 = \tau_2$, the shape parameters in the GCR2 marginals, and assume τ follows a Gamma distribution. The use of a common random parameter is expected to widen the range of dependence levels and allow for upper-tail dependence. We will pursue this approach in future research.

Author Contributions: Conceptualization, T.B.; methodology, T.B. and H.Q.; software, H.Q.; validation, T.B.; formal analysis, T.B. and H.Q.; investigation, T.B.; resources, T.B.; data curation, H.Q.; writing—original draft preparation, T.B. and H.Q.; writing—review and editing, T.B.; visualization, H.Q.; supervision, T.B.; project administration, T.B.; funding acquisition, T.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Discovery Grant program of the Natural Science and Engineering Research Council of Canada (NSERC) with a funding reference number RGPIN-2022-03428.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors thank the editors and the anonymous reviewers for their helpful comments that improved the readability and quality of this manuscript.

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

Appendix A. Proof of Proposition 1

We drop parameters in the expressions of the functions for notational convenience and write $\overline{G}(\cdot) = 1 - G(\cdot)$ and $b(t_i) = \alpha_i^{-1}[(\beta_i/t_i)^\tau - (t_i/\beta_i)^\tau]$, $i = 1, 2$. By $f_{BVGCR2}(t_1, t_2; \alpha, \beta, \tau, p)$ the joint distribution of (T_1, T_2) having $BVGCR2(\alpha, \beta, \tau, p; g)$, is

$$F(t_1, t_2) = p_{11}F_{GCR2}^{(1)}(t_1)F_{GCR2}^{(1)}(t_2) + p_{12}F_{GCR2}^{(1)}(t_1)F_{GCR2}^{(2)}(t_2) + p_{21}F_{GCR2}^{(2)}(t_1)F_{GCR2}^{(1)}(t_2) + p_{22}F_{GCR2}^{(2)}(t_1)F_{GCR2}^{(2)}(t_2),$$

and the marginal density of $T_i, i = 1, 2$, is

$$f_{T_i}(t_i) = p_i f_{GCR2}^{(1)}(t_i) + q_i f_{GCR2}^{(2)}(t_i),$$

also, the cdf of $GCR2(t; \alpha, \beta, \tau, p)$ distribution can be written as

$$F_{GCR2}(t; \alpha, \beta, \tau, p) = 1 - G(b(t)) + (2p - 1)H(t; \alpha, \beta, \tau, p),$$

where $p_1 = p_{11} + p_{12}, p_2 = p_{11} + p_{21}, q_i = 1 - p_i, i = 1, 2, G(x) = \int_{-\infty}^x g(s)ds$ is the cdf of g and

$$H(t; \alpha, \beta, \tau, p) = \int_{b(t)}^{\infty} \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s)ds.$$

By expanding the integrand in (7) using these expressions and taking the double integration, we have

$$\begin{aligned} & \int_0^\infty \int_0^\infty F(t_1, t_2) f_{T_1}(t_1) f_{T_2}(t_2) dt_1 dt_2 \\ &= p_{11} \{ p_1 p_2 E[F_{S_1}(S_1)] E[F_{S_2}(S_2)] + p_1 q_2 \zeta_2 E[F_{S_1}(S_1)] + q_1 p_2 \zeta_1 E[F_{S_2}(S_2)] + q_1 q_2 \zeta_1 \zeta_2 \} \\ &+ p_{12} \{ p_1 p_2 \eta_2 E[F_{S_1}(S_1)] + p_1 q_2 E[F_{S_1}(S_1)] E[F_{V_2}(V_2)] + q_1 p_2 \zeta_1 \eta_2 + q_1 q_2 \zeta_1 E[F_{V_2}(V_2)] \} \\ &+ p_{21} \{ p_1 p_2 \eta_1 E[F_{S_2}(S_2)] + p_1 q_2 \eta_1 \zeta_2 + q_1 p_2 E[F_{V_1}(V_1)] E[F_{S_2}(S_2)] + q_1 q_2 \zeta_2 E[F_{V_2}(V_2)] \} \\ &+ p_{22} \{ p_1 p_2 \eta_1 \eta_2 + p_1 q_2 \eta_1 E[F_{V_2}(V_2)] + q_1 p_2 \eta_2 E[F_{V_1}(V_1)] + q_1 q_2 E[F_{V_1}(V_1)] E[F_{V_2}(V_2)] \}. \end{aligned}$$

where for each $i = 1, 2$,

$$\begin{aligned} E[F_{S_i}(S_i)] &= \int_0^\infty F_{GCR2}^{(1)}(t_i) f_{GCR2}^{(1)}(t_i) dt_i = \frac{1}{2} \\ E[F_{V_i}(V_i)] &= \int_0^\infty F_{GCR2}^{(2)}(t_i) f_{GCR2}^{(2)}(t_i) dt_i = \frac{1}{2} \\ \zeta_i &:= \int_0^\infty F_{GCR2}^{(1)}(t_i) f_{GCR2}^{(2)}(t_i) dt_i \\ &= \int_0^\infty [F_{GCR2}^{(2)}(t_i) + 2H(t_i)] f_{GCR2}^{(2)}(t_i) dt_i \\ &= E[F_{V_i}(V_i)] + 2 \int_0^\infty [H(t_i) f_{GCR2}^{(2)}(t_i)] dt_i = \frac{1}{2} + 2 \int_0^\infty [H(t_i) f_{GCR2}^{(2)}(t_i)] dt_i \\ \eta_i &:= \int_0^\infty F_{GCR2}^{(2)}(t_i) f_{GCR2}^{(1)}(t_i) dt_i \\ &= \int_0^\infty [F_{GCR2}^{(1)}(t_i) + 2H(t_i)] f_{GCR2}^{(1)}(t_i) dt_i \\ &= E[F_{S_i}(S_i)] - 2 \int_0^\infty [H(t_i) f_{GCR2}^{(1)}(t_i)] dt_i = \frac{1}{2} - 2 \int_0^\infty [H(t_i) f_{GCR2}^{(1)}(t_i)] dt_i. \end{aligned}$$

That is, the double integral in (7) reduces to the following two integrals: $\int_0^\infty H(t)f_{GCR2}^{(1)}(t)dt$ and $\int_0^\infty H(t)f_{GCR2}^{(2)}(t)dt$. Due to the expression $F_{GCR2}^{(1)}(t) = 1 - G(b(t)) + H(t)$ and

$$\begin{aligned} & \int_{-\infty}^\infty H(b^{-1}(s)) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty \left(\int_s^\infty \frac{z}{\sqrt{z^2 + 4/\alpha^2}} \right) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds = 0, \end{aligned}$$

and changing the order of integrations, we obtain

$$\begin{aligned} \int_0^\infty H(t)f_{GCR2}^{(1)}(t)dt &= \int_0^\infty \int_{b(t)}^\infty \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds f_{GCR2}^{(1)}(t) dt \\ &= \int_{-\infty}^\infty \left(\int_{b^{-1}(s)}^\infty f_{GCR2}^{(1)}(t) dt \right) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty (1 - F_{GCR2}^{(1)}(b^{-1}(s))) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty (G(s) - H(b^{-1}(s))) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty G(s) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds - \int_{-\infty}^\infty H(b^{-1}(s)) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty \frac{s}{\sqrt{s^2 + 4/\alpha^2}} G(s) g(s) ds = \gamma. \end{aligned}$$

Similarly, due to the expression $F_{GCR2}^{(2)}(t) = 1 - G(b(t)) - H(t)$, we obtain

$$\begin{aligned} \int_0^\infty H(t)f_{GCR2}^{(2)}(t)dt &= \int_0^\infty \int_{b(t)}^\infty \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds f_{GCR2}^{(2)}(t) dt \\ &= \int_{-\infty}^\infty \left(\int_{b^{-1}(s)}^\infty f_{GCR2}^{(2)}(t) dt \right) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty (1 - F_{GCR2}^{(2)}(b^{-1}(s))) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty (G(s) + H(b^{-1}(s))) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty G(s) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds + \int_{-\infty}^\infty H(b^{-1}(s)) \frac{s}{\sqrt{s^2 + 4/\alpha^2}} g(s) ds \\ &= \int_{-\infty}^\infty \frac{s}{\sqrt{s^2 + 4/\alpha^2}} G(s) g(s) ds = \gamma. \end{aligned}$$

Combining these results and some simplifications give the expression (8).

Appendix B. Proof of Proposition 2

As in the proof of Proposition 1, expanding the integrand in (9) using the joint cdf and pdf of the BVGCR2 and taking double integrals, give

$$\begin{aligned} & \int_0^\infty \int_0^\infty F(t_1, t_2) f(t_1, t_2) dt_1 dt_2 \\ &= (p_{11})^2 E[F_{S_1}(S_1)] E[F_{S_2}(S_2)] + p_{11} p_{12} \zeta_2 E[F_{S_1}(S_1)] + p_{11} p_{21} \zeta_1 E[F_{S_2}(S_2)] \\ &+ p_{11} p_{22} \zeta_1 \zeta_2 + p_{12} p_{11} \eta_2 E[F_{S_1}(S_1)] + (p_{12})^2 E[F_{S_1}(S_1)] E[F_{V_2}(V_2)] \\ &+ p_{12} p_{21} \zeta_1 \eta_2 + p_{12} p_{22} \zeta_1 E[F_{V_2}(V_2)] + p_{21} p_{11} \eta_1 E[F_{S_2}(S_2)] + p_{21} p_{12} \eta_1 \zeta_2 \\ &+ (p_{21})^2 E[F_{V_1}(V_1)] E[F_{S_2}(S_2)] + p_{21} p_{22} \zeta_2 E[F_{V_2}(V_2)] + p_{22} p_{11} \eta_1 \eta_2 \\ &+ p_{21} p_{11} \eta_1 E[F_{V_2}(V_2)] + p_{22} p_{21} \eta_2 E[F_{V_1}(V_1)] + (p_{22})^2 E[F_{V_1}(V_1)] E[F_{V_2}(V_2)]. \end{aligned}$$

As given in the Proof of Proposition 1, $E[F_{S_1}(S_1)] = E[F_{V_2}(V_2)] = \frac{1}{2}$, $\zeta_i = \frac{1}{2} + 2\gamma_i$ and $\eta_i = \frac{1}{2} - 2\gamma_i$ for each $i = 1, 2$. With these and after some simplification, the above integral can be expressed as

$$\int_0^\infty \int_0^\infty F(t_1, t_2) f(t_1, t_2) dt_1 dt_2 = \frac{1}{4} + 8(p_{11}p_{22} - p_{12}p_{21})\gamma_1\gamma_2.$$

Appendix C. Proof of Lemma 1

Given the observed data \underline{x} , the log-likelihood function $l(\theta) := \log p(\underline{x}|\theta)$. By the law of total probability, we have

$$l(\theta) = \log p(\underline{x}|\theta) = \log \sum_{\underline{z}} p(\underline{x}|\underline{z}, \theta)p(\underline{z}|\theta) = \log \sum_{\underline{z}} p(\underline{x}, \underline{z}|\theta).$$

Then, by using Jensen’s inequality, we can obtain the following inequality on the log-likelihood function:

$$\begin{aligned} l(\theta) &= \log \sum_{\underline{z}} p(\underline{z}|\underline{x}, \theta^{(m)}) \frac{p(\underline{x}, \underline{z}|\theta)}{p(\underline{z}|\underline{x}, \theta^{(m)})}, \\ &= \log E_{Z|\underline{x}, \theta^{(m)}} \left[\frac{p(\underline{x}, Z|\theta)}{p(Z|\underline{x}, \theta^{(m)})} \right] \\ &\geq E_{Z|\underline{x}, \theta^{(m)}} \left[\log \frac{p(\underline{x}, Z|\theta)}{p(Z|\underline{x}, \theta^{(m)})} \right] \\ &= E_{Z|\underline{x}, \theta^{(m)}} [\log p(\underline{x}, Z|\theta)] - E_{Z|\underline{x}, \theta^{(m)}} [\log p(Z|\underline{x}, \theta^{(m)})] \\ &= Q(\theta|\theta^{(m)}) - E_{Z|\underline{x}, \theta^{(m)}} [\log p(Z|\underline{x}, \theta^{(m)})] = Q(\theta|\theta^{(m)}) - h(\theta^{(m)}). \end{aligned}$$

Note that $Q(\theta|\theta^{(m)})$ is the only term that depends on θ in the inequality

$$l(\theta) \geq Q(\theta|\theta^{(m)}) - h(\theta^{(m)}),$$

and the inequality holds for all θ including the situation where $\theta = \theta^{(m)}$. Specifically, when $\theta = \theta^{(m)}$,

$$\begin{aligned} Q(\theta^{(m)}|\theta^{(m)}) - h(\theta^{(m)}) &= E_{Z|\underline{x}, \theta^{(m)}} \left[\log \frac{p(\underline{x}, Z|\theta)}{p(Z|\underline{x}, \theta^{(m)})} \right] \\ &= \sum_{\underline{z}} p(\underline{z}|\underline{x}, \theta^{(m)}) \log \frac{p(\underline{x}, \underline{z}|\theta^{(m)})}{p(\underline{z}|\underline{x}, \theta^{(m)})} \\ &= \sum_{\underline{z}} p(\underline{z}|\underline{x}, \theta^{(m)}) \log p(\underline{x}|\theta^{(m)}) \\ &= \log p(\underline{x}|\theta^{(m)}) \sum_{\underline{z}} p(\underline{z}|\underline{x}, \theta^{(m)}) \\ &= \log p(\underline{x}|\theta^{(m)}) = l(\theta^{(m)}). \end{aligned}$$

From these we can deduce that, if $\theta = \theta^{(m+1)}$ satisfies $Q(\theta|\theta^{(m)}) \geq Q(\theta^{(m)}|\theta^{(m)})$, then

$$l(\theta^{(m+1)}) + h(\theta^{(m)}) \geq Q(\theta^{(m+1)}|\theta^{(m)}) \geq Q(\theta^{(m)}|\theta^{(m)}) = l(\theta^{(m)}) + h(\theta^{(m)}),$$

and thus, $l(\theta^{(m+1)}) \geq l(\theta^{(m)})$ for each $m \in \{1, 2, \dots\}$.

References

1. Bingham, N.H.; Goldie, C.M.; Teugels, J.L. *Regular Variation*; Cambridge University Press: Cambridge, UK, 1987.
2. Embrechts, P.; Klüppelberg, C.; Mikosch, T. *Modelling Extremal Events for Insurance and Finance*; Springer: Berlin, Germany, 1997.
3. Foss, S.; Korshunov, D.; Zachary, S. *An Introduction to Heavy-Tailed and Subexponential Distributions*; Springer Science+Business Media: Berlin, Germany, 2013.
4. Birnbaum, Z.W.; Saunders, S.C. A new family of life distributions. *J. Appl. Probab.* **1969**, *6*, 319–327. [[CrossRef](#)]
5. Leiva, V. *The Birnbaum-Saunders Distribution*; Academic Press: Cambridge, MA, USA, 2015.
6. Diaz-García, J.A.; Leiva-Sánchez, V. A new family of life distributions based on the elliptically contoured distributions. *J. Stat. Plan. Inference* **2005**, *128*, 445–457. [[CrossRef](#)]
7. Gomes, M.I.; Ferreira, M.; Leiva, V. The extreme value Birnbaum-Saunders model, its moments and an application in biometry. *Biom. Lett.* **2012**, *49*, 81–94. [[CrossRef](#)]
8. Ferreira, M.; Gomes, M.I.; Leiva, V. On an extreme value version of the Birnbaum–Saunders distribution. *REVSTAT-Stat. J.* **2012**, *10*, 181–210.
9. Volodin, I.N.; Dzhungurova, O.A. On limit distributions emerging in the generalized Birnbaum-Saunders model. *J. Math. Sci.* **2000**, *99*, 1348–1366. [[CrossRef](#)]
10. Bae, T.; Chen, J. On heavy-tailed crack distribution for loss severity modeling. *Int. J. Stat. Probab.* **2017**, *6*, 92–110. [[CrossRef](#)]
11. Mazjini, M.; Bae, T. Statistical modelling of precipitation data in Canadian Prairies with a dynamic mixture structure. *Theor. Appl. Climatol.* **2023**, *153*, 173–192. [[CrossRef](#)]
12. Bae, T.; Volodin, A. Type-II Generalized Crack Distribution with Application to Heavy-Tailed Data Modeling. *J. Stat. Theory Pract.* **2022**, *16*, 53. [[CrossRef](#)]
13. Azzalini, A. A Class of Distributions Which Includes the Normal Ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
14. Jones, M.C. Families of distributions arising from distributions of order statistics. *Test* **2004**, *13*, 1–43. [[CrossRef](#)]
15. Al-Shomrani, A.A.; Arif, O.H.; Ibrahim, S.; Shahbaz, M.Q. Topp–Leone Family of Distributions: Some Properties and Application. *Pak. J. Stat. Oper. Res.* **2016**, *12*, 443–451. [[CrossRef](#)]
16. Gómez-Déniz, E.; Calderín-Ojeda, E. On the Use of Lehmann’s Alternative to Capture Extreme Losses in Actuarial Science. *Risks* **2023**, *12*, 6. [[CrossRef](#)]
17. Bae, T.; Choi, Y.H. A bivariate extension of three-parameter generalized crack distribution for loss severity modelling. *J. Korean Stat. Soc.* **2022**, *51*, 378–402. [[CrossRef](#)]
18. Klugman, S.A.; Panjer, H.H.; Willmot, G.E. *Loss Models: From Data to Decisions*; John Wiley & Sons: Hoboken, NJ, USA, 2012; Volume 71.
19. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–22. [[CrossRef](#)]
20. Wang, Q.; Guo, G.; Qian, G.; Jiang, X. Distributed online expectation-maximization algorithm for Poisson mixture model. *Appl. Math. Model.* **2023**, *124*, 734–748. [[CrossRef](#)]
21. Guo, G.; Wang, Q.; Allison, J.; Qian, G. Accelerated distributed expectation-maximization algorithms for the parameter estimation in multivariate Gaussian mixture models. *Appl. Math. Model.* **2025**, *137*, 115709. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.