

RESEARCH ARTICLE

Return Period of Nonconcurrent Climate Compound Events: A Nonparametric Bivariate Generalized Pareto Approach

Grégoire Jacquemin^{1,2}  | Denis Allard³  | Xavier Freulon¹ | Mathieu Vrac² 

¹Centre for Geosciences and Geoengineering, Mines Paris, PSL University, Fontainebleau, France | ²Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL), CNRS/CEA/UVSQ, UMR8212, Université Paris-Saclay, Gif-sur-Yvette, France | ³Biostatistiques et Processus Spatiaux (BioSP), INRAE, Avignon, France

Correspondence: Grégoire Jacquemin (gregoire.jacquemin@minesparis.psl.eu)

Received: 20 May 2025 | **Revised:** 24 September 2025 | **Accepted:** 1 December 2025

Keywords: antecedent precipitation index | asymptotic dependence | bivariate extremes | extended generalized Pareto distribution | temporal dependence

ABSTRACT

Compound events (CEs), commonly defined as the “combination of multiple drivers and/or hazards that contributes to societal or environmental risk”, often result in amplified impacts compared to individual hazards. In order to estimate the return period of bivariate CEs, a novel nonparametric approach employing bivariate Generalized Pareto distributions (bi-GPD) is proposed and compared to a copula-based approach. Special attention is given to account for temporal dependencies and nonconcurrent compound events. The latter are defined as excess of variables over a threshold at a relatively close time. The return period of such bivariate events is carefully defined and closed-form expressions are obtained for both approaches. Simulations reveal the bi-GPD approach is effective in case of positive asymptotic dependence and should be avoided in case of asymptotic independence. The novel approach is then applied to ERA5 reanalysis data to analyze two types of compound events: a spatial CE with simultaneous floods due to accumulated precipitation across two large watersheds in France and a preconditioned CE consisting of a devastating flood triggered by extreme precipitation over a saturated soil.

1 | Introduction

Recent advances in the study of extreme events have highlighted the necessity of examining multi-hazard events, and the interrelation between these hazards (Gill and Malamud 2014; Terzi et al. 2019). These multi-hazard events are often referred to as compound events, which are defined in Zscheischler et al. (2020) as “combination of multiple drivers and/or hazards that contributes to societal or environmental risk”. This combination often results in amplified impacts compared to individual hazards. Compound events represent a distinct category of extreme

events, characterized by the interaction between multiple components for a comprehensive description. “Simpler” events, typically described using a single variable, have been more extensively studied than CEs, even though the research in compound weather and climate events grew largely in the last decade (Brett et al. 2025). CE have also been studied by means of indices aggregating several variables into a single one, for example the case of wildfires in Villalobos-Herrera et al. (2021). More thorough and complex modeling of the statistical dependencies between the variables enables the consideration of multiple variables in the study of compound events (Bevacqua et al. 2017; Yang et al. 2023;

Xu et al. 2023). In this study, we will restrict ourselves to the study of bivariate compound events, that is, compound events involving two variables.

Investigating compound events involves analyzing extreme values and their joint behavior at high quantiles, therefore requiring the application of multivariate Extreme Value Theory (EVT, Beirlant et al. 2006). Tilloy et al. (2020) provided a comprehensive review and comparison of various multivariate extremal approaches for compound events. Multivariate extremes may be represented using joint tail models, but they require a complex parametric modeling of a slowly varying function in the tail, and can only provide a model for the joint excess region (Ledford and Tawn 1997; Ramos and Ledford 2008). The conditional extreme approach estimates the dependence structure between two variables, conditional on one being extreme (Wadsworth and Tawn 2022; Winter and Tawn 2015), but has some limitations (Drees and Janßen 2017). First, this approach assumes an event in which both variables are extreme, as opposed to an event where at least one of the variables is extreme. Secondly it does not directly provide the excess probability, which must be estimated through simulations (Heffernan and Tawn 2004). Kernel density estimation (Cooley et al. 2019), a nonparametric approach, can estimate joint excess probabilities and isolines of return periods based on underlying regularly varying functions. However, it depends critically on the selection of the bandwidth matrix (Duong 2007), a key hyper-parameter that significantly influences the performance of the estimator. Optimal selection of the bandwidth is complex and requires substantial expertise.

In the context of compound events, copulas are commonly used to model data in more than one dimension. The use of copulas allows for a decoupling between the marginals and the dependence structure. There exists three primary frameworks involving copulas: a parametric approach, a semi-parametric approach with joint densities and a nonparametric approach. The parametric copula framework relies on a parametric modeling of the margins, and a parametric modeling of the copula (Cammalleri et al. 2024; Bennet et al. 2025), through maximum likelihood estimation or inference function for margins (Joe 2005). The semi-parametric copula joint density framework, proposed in Genest et al. (1995), is based on empirical margins and parametric copulas, and is compared to the parametric framework in Kim et al. (2007). It is used for instance in Tavakol et al. (2020) and Páscoa et al. (2024). The nonparametric copula approach uses nonparametric methodologies, such as Bernstein copulas or kernel density estimations, some of them being detailed in Provost and Zang (2024). All these frameworks usually lead to a misrepresentation of the univariate extremes, because the margins are often modeled with their empirical distribution or with a parametric distribution not suited for extremes. This issue can be addressed by using an extreme value mixture model (MacDonald et al. 2011) to model the marginals in the parametric copula framework, as in Bevacqua et al. (2017). However, this could still lead to biasing the representation of extremes by the bulk of the bivariate distribution (Bevacqua et al. 2019).

To more accurately capture the extremes and their associated probabilities, we follow the approach presented in Bevacqua et al. (2019, 2020). In this approach, only the tails of the margins are modeled with a Generalized Pareto Distribution (GPD) and

a parametric copula is estimated only on bivariate exceedances, that is, when both variables exceed a threshold. This approach will be referred to as the GPD-copula approach in this article. However, the GPD-copula approach requires the selection of a copula family and the estimation of its parameters (Nelsen 2006). The approach is significantly influenced by various choices, including the threshold, copula family, and estimator. Moreover, the accurate estimation of copula parameters is sensitive to the number of data points in the upper tail of the distribution. This method is also highly sensitive to the potential temporal dependence inherent in the process.

The primary objective of this work is to propose a novel bivariate stochastic approach for compound events which avoids the shortcomings of the above mentioned approaches. This new approach is semi-parametric, flexible, allows the computation of the probability of excess and does not rely on the fine tuning of hyper-parameters. The proposed bivariate approach is based on the Multivariate Generalized Pareto Distribution (MGPD, Rootzén and Tajvidi 2006; Rootzén et al. 2018) and incorporates a re-parametrization introduced by Legrand et al. (2023). This approach is semi-parametric: the marginal distributions are modeled using the extended GPD (Naveau et al. 2016), while the joint exceedance probability is expressed nonparametrically. As in Tilloy et al. (2020), we will evaluate our bivariate Generalized Pareto Distribution (bi-GPD) approach against the GPD-copula approach in the tail using several classical criteria commonly employed by practitioners, such as joint exceedance probability in the tail, return periods, and coefficients of extremal dependence (Coles et al. 1999).

This work has been motivated by two recent floods that can be interpreted as compound events (CE). The first examined CE is the simultaneous flooding of the Seine and Loire rivers in France during late spring 2016, and the second one is the devastating flooding of the Ahr catchment, Germany, in July 2021. In both CEs, the accumulation of precipitation over several days played a critical role in the occurrence of the floods. Accounting for precipitation accumulation introduces temporal dependence among the relevant variables, a feature addressed through specific methodological approaches that will be discussed later in this work. Moreover, the 2016 flood event in France was triggered by an unusually high amount of precipitation over several days at the end of May, which resulted in the two rivers reaching their peak levels *a few days apart*.

From a statistical point of view, a compound event is defined as *the simultaneous exceedance of both variables over some threshold*. This definition, which considers the probability of joint excess, does not take into account a possible time lag between the two variables. However, in the Seine/Loire motivating example, the flooding of two major rivers within a few days of each other may not be considered as two distinct events but rather as a single compound event by many stakeholders, state services, insurance and reinsurance companies, media, and so forth. To our knowledge, the analysis of compound events has not yet explicitly considered the existence of a time lag between the excesses of the different variables. This observation motivated the definition adopted in this work: *an exceedance of both variables over a threshold at a relatively close time*. We refer to these events as “nonconcurrent CEs”. The second objective of this study is thus to account

for nonconcurrency in the notion of return period. New expressions are therefore obtained, both for the GPD-copula and the bi-GPD approach.

This article is organized as follows: Section 2 recalls some necessary background material on univariate extremal models (GPD and EGPD), along with the GPD-copula approach used to benchmark our new approach. In Section 3, our new bivariate Generalized Pareto Distribution (bi-GPD) approach is presented. Section 4 discusses different approaches to address the temporal dependence of the variables, and introduces the concept of nonconcurrency. Section 5 explores the implications of the nonconcurrency for the return periods, proposing a first definition for the bivariate return period of nonconcurrent compound events, which is declined for both approaches presented earlier. A comparison of the two extremal approaches (the GPD-copula approach and the bi-GPD approach) is provided through statistical simulations in Section 6. Section 7 applies the bi-GPD approach and the GPD-copula approach to the two compound floodings motivating this work, and discusses the results and the nonconcurrency. Finally, Section 8 provides some conclusions and perspectives of future work.

2 | Reminders on Extreme Value Theory

In the following, all the bold notations correspond to vectors. The bivariate cumulative distribution function of $X = (X_1, X_2)$ is $F(x_1, x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2)$ and the marginal cumulative distribution function of X_i is denoted F_i , $i = 1, 2$. We further denote \bar{F} the joint survival function: $\bar{F}(x_1, x_2) = \mathbb{P}(X_1 > x_1, X_2 > x_2)$ and $\bar{F}_i = 1 - F_i$ for $i = 1, 2$. All operations involving vectors, like maximum, inequalities, and so forth, are meant component-wise. The set of consecutive integers $\{n, n + 1, \dots, m\}$ will be denoted $\llbracket n, m \rrbracket$, with $n \leq m$. Finally, for a a scalar, we note $a_+ = \max(0, a)$ and $a_- = \min(0, a)$.

2.1 | Univariate Modeling

Let X_t be a random stationary process indexed along time, characterized by its cumulative distribution function (cdf) F_i , $i = 1, 2$. In this work, motivated by the modeling of climate extreme events, the process is indexed along discrete time, for example it represents daily values. In this Section and in Section 3, we further assume that the $X_{i,t}$ s are independent, with $X_{i,t}$ the value of the process at time t . This assumption will be relaxed in Section 4, where time dependence will be accounted for.

As $n \rightarrow \infty$, the extremal behavior of $(X_{i,t})_{t=1,n}$ can be expressed either through the distribution of its maximum $M_n = \max_{t \leq n}(X_{i,t})$, or through the distribution of the excess over a high threshold u . The first approach, called block-maxima, states that, with appropriate normalizing sequences, the limit distribution of M_n is a Generalized Extreme Value (GEV) distribution (Leadbetter 1974). According to the second approach, called peaks-over-threshold, the conditional distribution of the excess, $\mathbb{P}(X_i \leq x | X_i > u)$, can be approximated by a Generalized Pareto Distribution (GPD, Pickands 1975) for any positive scalar $x > u$ and with u close to the upper-end point of F_i . The cdf of the GPD is of the form:

$$G(x - u; \xi, \sigma) = \begin{cases} 1 - \left(1 + \xi \frac{x-u}{\sigma}\right)^{-\frac{1}{\xi}}, & \text{for } \xi \neq 0 \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right), & \text{for } \xi = 0, \end{cases} \quad (1)$$

where the support of G is $[0, \infty)$ for $\xi \geq 0$ and $[0, -\sigma/\xi]$ for $\xi < 0$, with $\xi \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ being the shape and scale parameters, respectively.

The new bivariate nonparametric approach that will be presented in Section 3 requires a complete modeling of the distribution, that is, both for the bulk and for the tail of the distribution. The Extended Generalized Pareto Distribution (EGPD) proposed in Naveau et al. (2016) allows such a modeling. An interesting feature of this family is that no threshold is required to separate the bulk from the tail, the latter behaving by construction as a GPD at very high values. Denoting G and G_E the cumulative distribution functions of the GPD and of the EGPD respectively, we will consider the mixture of powers given by: $G_E(x) = pG(x; \xi, \sigma)^{\kappa_1} + (1 - p)G(x; \xi, \sigma)^{\kappa_2}$, with $\kappa_1, \kappa_2 > 0$ and $0 \leq p \leq 1$. Note that when $p = 0$ or $p = 1$, this model reduces to a simple power. See Naveau et al. (2016) for more complex models and general conditions on G_E .

2.2 | Bivariate Modeling

2.2.1 | Coefficients of Extremal Dependence

The upper tail dependence coefficient χ is the probability that one variable reaches high values, conditional on the other variable being high (Joe 1993; Coles et al. 1999).

Definition 1 (Upper tail dependence). Let X_1 and X_2 be two random variables characterized by their cdf F_1 and F_2 , respectively. Let us denote $V_1 = F_1(X_1)$ and $V_2 = F_2(X_2)$ their uniform transforms.

The upper tail dependence function, $\chi(\cdot)$, is:

$$\chi(v) = \mathbb{P}(V_2 > v | V_1 > v), \quad 0 \leq v < 1, \quad (2)$$

and the upper tail dependence coefficient is $\chi = \lim_{v \rightarrow 1} \chi(v)$.

The upper extremal dependence function, $\bar{\chi}(\cdot)$, is:

$$\bar{\chi}(v) = \frac{2 \log \mathbb{P}(V_1 > v)}{\log \mathbb{P}(V_1 > v, V_2 > v)} - 1, \quad 0 \leq v < 1, \quad (3)$$

and the upper extremal dependence coefficient is $\bar{\chi} = \lim_{v \rightarrow 1} \bar{\chi}(v)$.

The couple $(\chi, \bar{\chi})$ characterizes the extremal dependence. Clearly, $\chi \in [0, 1]$. Positive values correspond to asymptotic dependence, and the value of χ indicates the strength of the dependence up to $\chi = 1$ which corresponds to perfect dependence. When $\chi = 0$, X_1 and X_2 are asymptotically independent, the strength of the (sub-asymptotic) dependence being given by $\bar{\chi} \in (-1, 1)$. $\bar{\chi} = 0$ corresponds to total independence. Positive (resp. negative) values correspond to asymptotic independence with positive (resp. negative) association. If $\bar{\chi} = 1$, there is asymptotic dependence, and $\chi \in (0, 1]$ is the relevant coefficient to

TABLE 1 | Summary of the different combinations of $(\chi, \bar{\chi})$.

	$\chi = 0$	$0 < \chi < 1$	$\chi = 1$
	Asympt. indep.	Asympt. dep.	Perfect dep.
$-1 < \bar{\chi} < 0$	Asympt. indep (-)	NP	NP
$\bar{\chi} = 0$	Independence	NP	NP
$0 < \bar{\chi} < 1$	Asympt. indep (+)	NP	NP
$\bar{\chi} = 1$	NP	Asympt. dep.	Perfect dep.

Abbreviation: NP, not possible.

describe the strength of dependence. The different configurations are summarized in Table 1. See also Coles et al. (1999) for more details on these coefficients.

For simplicity, we will use the terms *asymptotically dependent event* when the estimated $\bar{\chi}$ of a compound event is approximately equal to 1, and *asymptotically independent event* when the estimated χ is approximately equal to 0.

2.2.2 | The GPD-Copula Approach

A classical approach to represent multivariate extreme values and their dependence is the GPD-copula approach. According to Sklar's theorem (Sklar 1959), there exists a function $C : [0, 1]^2 \rightarrow [0, 1]$ called the copula of $\mathbf{X} = (X_1, X_2)$, defined for all $\mathbf{x} = (x_1, x_2) > 0$ such that

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (4)$$

Moreover, if the marginals of \mathbf{X} are continuous, then the copula is unique. Later, we will make use of the survival copula \bar{C} defined as: $\bar{F}(x_1, x_2) = \bar{C}(F_1(x_1), F_2(x_2))$. It is easy to verify that for $0 \leq v_1, v_2 \leq 1$, $\bar{C}(v_1, v_2) = 1 - v_1 - v_2 + C(v_1, v_2)$.

Let us assume that F is in the domain of attraction of a multivariate extreme value distribution. Then, for each X_i , $i = 1, 2$, its excess over high thresholds, $F_i^{u_i}(y) := \mathbb{P}(X_i - u_i \leq y | X_i > u_i)$, can be approximated by a GPD as $u_i \rightarrow \infty$. Define $\mathbf{u} = (u_1, u_2)$. Applying Sklar's theorem to the bivariate excess $F^{\mathbf{u}}(y_1, y_2) = \mathbb{P}(X_1 - u_1 \leq y_1, X_2 - u_2 \leq y_2 | X_1 > u_1, X_2 > u_2)$, one defines the copula of joint excesses $C_{\mathbf{u}}$ with $F^{\mathbf{u}}(y_1, y_2) = C_{\mathbf{u}}(F_1^{u_1}(y_1), F_2^{u_2}(y_2))$. In Section 5, this will allow us to express the joint probabilities of the excesses as well as bivariate return periods.

The coefficients of extremal dependence can be expressed in terms of copula. The margins (X_1, X_2) are transformed into (V_1, V_2) to obtain uniform margins. For $0 \leq v < 1$, the upper tail dependence coefficient becomes:

$$\chi = \lim_{v \rightarrow 1} \frac{\mathbb{P}(V_1 > v, V_2 > v)}{\mathbb{P}(V_1 > v)} = \lim_{v \rightarrow 1} \frac{\bar{C}(v, v)}{1 - v}. \quad (5)$$

Regarding the upper extremal dependence coefficient, the definition in (3) leads immediately to

$$\bar{\chi}(v) = \frac{2 \log(1 - v)}{\log \bar{C}(v, v)} - 1. \quad (6)$$

The GPD-copula approach necessitates either using an empirical copula or selecting a copula model and estimating its parameters. As will be illustrated later, the parametric modeling is unfortunately hampered by instabilities due to small or moderate sample sizes and/or the difficulty of selecting one copula family among the many possible families. To overcome these difficulties, we propose a new nonparametric approach for bivariate extreme exceedances.

3 | A Nonparametric Approach Using Bivariate Generalized Pareto Distribution

The proposed nonparametric approach is based on the characterization of multivariate GPDs by Rootzén et al. (2018) and further worked out in Kiriliouk et al. (2019). This characterization of multivariate GPDs was later used by Legrand et al. (2023) for the simultaneous stochastic simulation of extreme coastal and offshore significant wave heights. It is used here for the nonparametric estimation of bivariate probabilities. For completeness, we briefly recall the approach used in Legrand et al. (2023) before presenting our contributions.

3.1 | Joint Exceedance Probability With Nonparametric Bivariate GPD

Let us consider a high probability $0 \ll p < 1$ and $\mathbf{u} = (F_1^{-1}(p), F_2^{-1}(p))$. Then, in a peak-over-threshold approach, $[\mathbf{X} - \mathbf{u} | \mathbf{X} \not\leq \mathbf{u}]$ follows a bivariate GPD with marginal parameters ξ and σ (Beirlant et al. 2006; Rootzén and Tajvidi 2006). $\mathbf{X} \not\leq \mathbf{u}$ means that $X_1 > u_1$ or $X_2 > u_2$, that is, at least one of the components is extreme. We further denote the exponential transform by $\mathbf{X}^E = (X_1^E, X_2^E) = (-\log(1 - F_1(X_1)), -\log(1 - F_2(X_2)))$. Then, the vector $\mathbf{Z} = (Z_1, Z_2)$ defined by $\mathbf{Z} := [\mathbf{X}^E - \mathbf{u}^E | \mathbf{X}^E \not\leq \mathbf{u}^E]$ follows a bivariate GPD with parameters $\xi = \mathbf{0}$ and $\sigma = \mathbf{1}$, and $u_i^E = -\log(1 - F_i(u_i)) = -\log(1 - p)$.

Rootzén et al. (2018) established a stochastic representation of \mathbf{Z} . They showed that there exists a random vector $\mathbf{T} = (T_1, T_2)$ and a unit exponential random variable E independent of \mathbf{T} , such that

$$\mathbf{Z} \stackrel{\text{law}}{=} E + \mathbf{T} - \max(\mathbf{T}) \quad (7)$$

where $\stackrel{\text{law}}{=}$ denotes the equality in distribution.

To avoid the parametric modeling of \mathbf{T} as in Kiriliouk et al. (2019), in a bivariate context, Legrand et al. (2023) defined the random variable

$$\Delta = Z_1 - Z_2 = T_1 - T_2 \quad (8)$$

with cumulative distribution F_{Δ} . The decomposition (7) thus becomes:

$$Z_1 = E + \Delta_{-}, \quad Z_2 = E - \Delta_{+}. \quad (9)$$

This decomposition, illustrated on Figure 1, is used to compute the joint exceedance probability function without any parametric assumption on F_{Δ} , as stated in Theorem 1.

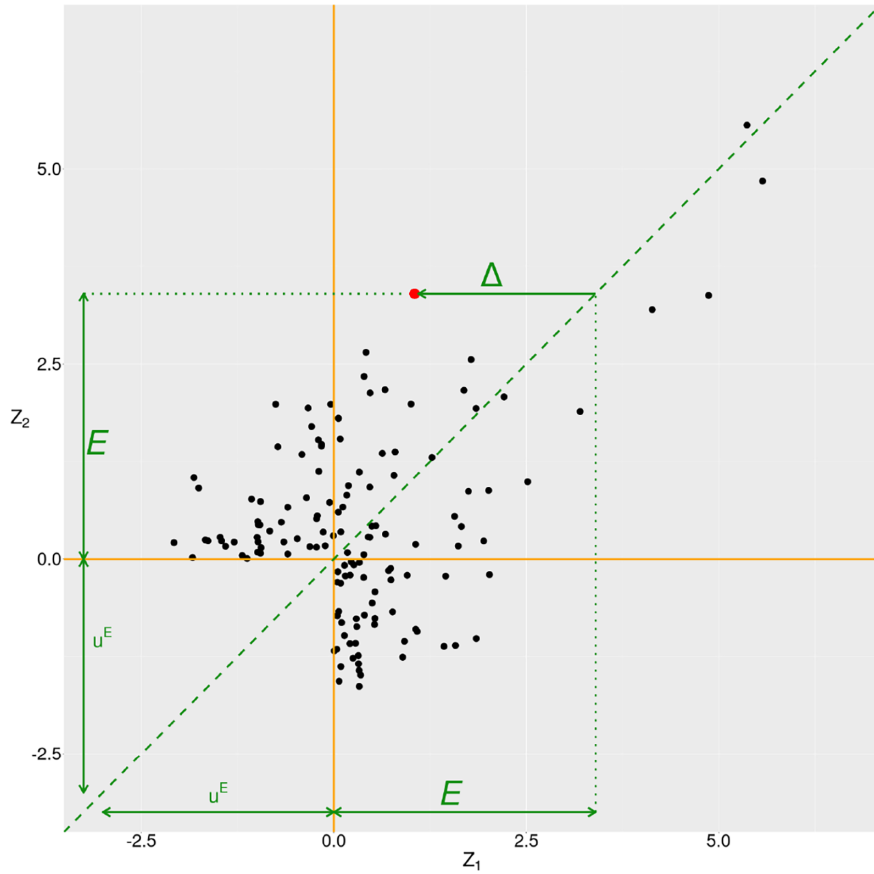


FIGURE 1 | Illustration of the bivariate GPD representation: data are from the Seine/Loire event, studied in Section 7 with a representation of E and Δ , as in Equation (9).

Theorem 1. Let u_1 and u_2 be two thresholds corresponding to high quantiles of X_1 and X_2 , respectively, with $p = F_1(u_1) = F_2(u_2) < 1$ and let $u^E = -\log(1 - p)$. Let us assume that, in the decomposition leading to (9), F_Δ is a continuous cdf. Then, for $y_1, y_2 \geq 0$:

$$\begin{aligned} & \bar{F}(u_1 + y_1, u_2 + y_2) \\ &= \frac{\bar{F}(u_1, u_2) \left[(1 - F_2(u_2 + y_2)) \int_{\delta_+}^{+\infty} e^{-t} F_\Delta(t) dt - (1 - F_1(u_1 + y_1)) \int_{-\infty}^{\delta_-} e^t F_\Delta(t) dt \right]}{(1 - p) \int_0^{+\infty} e^{-t} (F_\Delta(t) - F_\Delta(-t)) dt} \end{aligned} \quad (10)$$

where $\delta = z_1 - z_2 = \log \left[(1 - F_2(u_2 + y_2)) / (1 - F_1(u_1 + y_1)) \right]$.

The proof is given in Appendix A. The joint survival function can thus be easily computed as an integral of the cumulative distribution F_Δ weighted by the exponential. In practice, the empirical cdf is used in equation (10). As will be shown in Section 6 on simulations, thanks to this weighting, the integrals computed in samples of moderate size are robust estimates of the true value.

3.2 | Extremal Dependence With Bivariate GPD

As shown in Equation (5), the upper tail dependence function admits a simple expression involving the survival copula,

with $\chi(v) = \bar{C}(v, v) / (1 - v)$. More generally, the coefficients of extremal dependence can also be expressed as a function of the survival cumulative density function, \bar{F} .

Proposition 1. Let $\mathbf{u} = (u_1, u_2)$ be two thresholds corresponding to high quantiles of X_1 and X_2 respectively, with $p = F_1(u_1) = F_2(u_2)$, such that $[X - \mathbf{u} | X \not\leq \mathbf{u}]$ follows a bivariate GPD. Let further assume that X_1 and X_2 are not perfectly negatively rank-correlated, that is, that there is no decreasing one-to-one mapping ψ such that $F_2 = F_1 \circ \psi$. Then, $\forall v \geq p$:

$$\chi(v) = \frac{\bar{F}(u_1, u_2)}{1 - p} = \chi > 0 \quad \text{and} \quad \bar{\chi} = 1. \quad (11)$$

The proof is given in Appendix B. Notice that $\chi(v) = \chi$ is a constant completely defined by p and \bar{F} . Moreover, χ as given in Equation (11) is equal to $\chi(v)$ as defined in Equation (2) for $v = 1 - p = \bar{F}_1(u_1) = \bar{F}_2(u_2)$. Proposition 1 thus shows that our new approach based on the bivariate GPD representation is not able to represent independence or asymptotic independence. This is a known result for bivariate extremal distribution (Michel 2008; Coles et al. 2001). In the light of this result, Equation (10) can be decomposed in two terms: the first one corresponds to χ and the second one, the ratio of integrals, corresponds to a distance from the bisector.

4 | Compound Events Defined on Temporal Dependent Times Series

As it will be detailed in Section 7.1, the Seine/Loire event is defined by variables showing temporal dependence, in particular in the extremes. This section presents a possibility to incorporate dependent times series in the theory presented in the previous section, before introducing the notion of nonconcurrent compound events.

4.1 | Accounting for Temporal Dependence: The Extremal Index

From now on, the X_i s will be identically distributed but not independent, except when explicitly stated otherwise. Temporally dependent variables tend to reach high values in groups, often called clusters (see e.g., Bousquet and Bernardara 2021), that usually correspond to a single physical event. Elements of extreme value theory presented in Section 2 rely on the hypothesis of independent series. The study of these clusters is necessary to modify the theory to deal with temporal dependence (Leadbetter et al. 1989).

We require that some independence property holds for random variables that are far apart. We will thus assume that the so-called $D(u_n)$ condition (Leadbetter 1974; Beirlant et al. 2006) is verified. This condition states that the events $\{\max_{k \in I_1}(X_k) \leq u_n\}$ and $\{\max_{k \in I_2}(X_k) \leq u_n\}$ can become approximately independent as $n \rightarrow \infty$, when the disjoint sets of indices $I_i \subset \llbracket 1, n \rrbracket, i = 1, 2$ are separated by a distance $d(n)$ such that $d(n) = o(n)$. In essence, the assumption is made that the long range dependence of the process is limited, and, in practice, that the maxima on disjoint sets of indices can be considered as approximately independent.

To account for temporal dependence, Coles et al. (2001) and Beirlant et al. (2006) proposed an approach called declustering: inference is made only on maxima of clusters that exceed the chosen threshold. Declustering has been criticized for introducing bias in the estimates in many practical cases (Fawcett and Walshaw 2007). Eastoe and Tawn (2012) proposed to correct the distribution estimated on cluster maxima, but this approach relies on complex modeling of the extremal index, the latter being defined in the next paragraph. Moreover, their correction, with the estimation of the extremal index described in Section 6.1, did not lead to significant changes in our results. Buriticá and Naveau (2023) introduced a novel approach based on large deviation to model extremes of temporally dependent variables. Their approach is however quite complex. It needs large temporal blocks and stands on a strong hypothesis for the multivariate aspect (margins are supposed asymptotically equivalent up to a constant) that is not matched in our examples. For these reasons we opted for the widely used declustering approach.

Declustering relies on the estimation of the extremal index, defined in the following way. Consider an independent and identically distributed sample $(\tilde{X}_i)_{i=1, n}$ following the same marginal distribution as $(X_i)_{i=1, n}$. We denote \tilde{M}_n the maximum of $(\tilde{X}_i)_{i=1, n}$, which is the counterpart of M_n (see Section 2.1). F_{\max} and \tilde{F}_{\max} are the CDFs of M_n and \tilde{M}_n , respectively. Since the X_i s are not independent but the \tilde{X}_i s are, these two functions are not equal

in general. The extremal index is the scalar θ , with $0 < \theta \leq 1$, for which (Leadbetter et al. 1983):

$$F_{\max} = \tilde{F}_{\max}^\theta \tag{12}$$

This index quantifies the temporal dependence of a process in its extreme values, allowing to proceed as if only a fraction of the data was available, but that independence can be assumed on this fraction of data. The estimation of the extremal index is addressed in Section 6.1.

M. Leadbetter (1991) characterizes the distribution of the excesses of cluster maxima over a threshold u_n . For X a stationary process with temporal dependence, the excesses of cluster maxima over u_n converges in distribution to a generalized Pareto distribution $G(\cdot; \xi, \sigma)$ as $n \rightarrow \infty$. For this result, u_n should be a sequence of thresholds for which the condition $D(u_n)$ is satisfied and such that: $n(1 - F(u_n)) \xrightarrow{n \rightarrow \infty} \tau < +\infty$. In this case, the distribution of the cluster maxima over a threshold is actually equal to the distribution of the tail of the process X (Beirlant et al. 2006).

The principal advantage in estimating the parameters of the GPD on cluster maxima rather than directly on process exceedances is that the cluster maxima are, by construction, independent. Moreover, $1/\theta$ can be seen as the mean cluster size of the process (Moloney et al. 2019). Let q denote the smallest integer greater than $1/\theta$; blocks of size q represent the smallest block sizes for which the maxima can be considered independent.

In the general multivariate case, the analogue of the extremal index is a function (Smith and Weissman 1996; Nandagopalan 1994; Martins and Ferreira 2005; Robert 2008). Here, we restrict ourselves to the bivariate case. Let $\{X_n\}$ be a stationary sequence of random vectors in \mathbb{R}_+^2 with cdf F . Assume that there are vectors $a_n > 0$ and b_n , and bivariate extreme value distributions K and \tilde{K} (the natural extensions of F_{\max} and \tilde{F}_{\max}) such that, as $n \rightarrow \infty$:

$$\begin{aligned} \mathbb{P}[a_n^{-1}(M_n - b_n) \leq x] &\rightarrow K(x) \\ \mathbb{P}[a_n^{-1}(\tilde{M}_n - b_n) \leq x] &= F^n(a_n x + b_n) \rightarrow \tilde{K}(x). \end{aligned}$$

Definition 2 (Bivariate extremal index). For $\tau \in [0, \infty) \setminus \{0\}$, let $x(\tau)$ be the one-to-one mapping such that $\tau_j = -\log \tilde{K}_j(x_j)$ for $j = 1, 2$ when $\tau_j > 0$. When $\tau_j = 0$, we set $x_j = \sup\{x \in \mathbb{R} : \tilde{K}_j(x) < 1\}$. The extremal index function $\theta(\cdot)$ of the sequence $\{X_n\}$ is defined by:

$$\theta(\tau) = \frac{\log K(x(\tau))}{\log \tilde{K}(x(\tau))} \Leftrightarrow K(x(\tau)) = \tilde{K}^{\theta(\tau)}(x(\tau)), \quad \tau \in [0, \infty) \setminus \{0\}. \tag{13}$$

Clearly, the definition in (13) generalizes the extremal index defined in (12), since $\theta(\tau_1, 0) = \theta_1$ and $\theta(0, \tau_2) = \theta_2$ for any $\tau_i > 0$. We refer to Beirlant et al. (2006) or Nandagopalan (1994) for a complete construction and a detailed exposition of the properties of the multivariate extremal index. Since $x(\tau)$ is one-to-one mapping, one can also write $\tau(x) = (-\log \tilde{K}_1(x_1), -\log \tilde{K}_2(x_2))$. Thus, the bivariate extremal index $\theta(\tau)$ can be more simply rewritten:

$$\theta(\tau) = \theta(\tau(x_1, x_2)) := \theta(x_1, x_2).$$

4.2 | Nonconcurrent Compound Events

For the spatial compound events motivating this work, the exact time at which each marginal extreme occurs can differ by some interval. As an example, the spatial compound event consisting of the almost simultaneous extreme precipitation of the Seine and the Loire watersheds in France in 2016 was separated by 4 days. Even though it did not happen the exact same day, they were considered to belong to the same compound event (van Oldenborgh et al. 2016). We call these “nonconcurrent” compound events. In Section 5, it will be shown that taking into account the time interval between these events has an impact on how one defines and computes the return period.

Extending the definition of return periods to a bivariate context depends on the chosen event (Brunner et al. 2016). Here we impose that both variables exceed their respective thresholds within some time window. Let us denote h the size of the time window, that is, the maximum number of time steps (here, days) between the exceedance of the two variables. We call h the nonconcurrency parameter. The probability that both variables exceed $(x_1, x_2) > 0$ respectively *within the same block of size h* is:

$$\mathbb{P}\left[\max_{i \in B} (X_{1,i}, X_{2,i}) > (x_1, x_2)\right], \quad (14)$$

where B is a block of size h . Notice that h has to be such that the maxima within different blocks can be considered independent, that is, $h \geq 1/\theta(x_1, x_2)$ for any $(x_1, x_2) > 0$. When $\theta(x_1, x_2) = 1$, the case $h = 1$ corresponds to co-occurrence, and the derived formulas that will be presented in Sections 5.3 and 5.4 align with the conventional return period formulas. In the rest of this work, we will use the probability defined in (14) to compute the return period of a nonconcurrent compound event.

5 | Return Periods for Nonconcurrent Compound Events

A natural way to quantify the occurrence of an extreme event is through its return period. For a single variable, an extreme event is defined as the exceedance of a certain value x_T called the “return level”. The return period T is then defined as the expectation of the waiting time between two such events. For temporally independent data, this leads to the classical formula: $T = 1/N\mathbb{P}(X > x_T)$, where N is the number of days per year to get T in years. This section will present specific developments for the return period when accounting for temporally dependent variables in a bivariate context for nonconcurrent compound events. Theoretical expressions will be derived, both for the GPD-copula and for the bi-GPD approaches.

5.1 | Return Periods as the Expectation of a Waiting Time

Let us consider n_b blocks $(B_j)_{j=1, \dots, n_b}$, each of size h , and let us define the bivariate maxima $\mathbf{M}(B_j) = \max_{i \in B_j} (X_{1,i}, X_{2,i})$, $j = 1, \dots, n_b$. Following Resnick (2008), we define the waiting time $\eta(x_1, x_2)$ as the index of the first block of size h with a (bivariate) maxima exceeding $x_1 > 0$ and $x_2 > 0$:

$$\eta(x_1, x_2) = \inf_{j=1, \dots, n_b} (\mathbf{M}(B_j) > (x_1, x_2)). \quad (15)$$

By convention, $\eta(x_1, x_2) = n_b + 1$ if no block meets the requirement. The definition of the return period is the expectation of the waiting time. Since $\mathbb{E}(\eta(x_1, x_2))$ is a number of blocks, it needs to be renormalized, that is, multiplied by h/N , to get T in years. We thus get the general definition:

Definition 3. The nonconcurrent return period associated to the return level vector $(x_{1,T}, x_{2,T})$ is the time T such that:

$$T = \frac{h\mathbb{E}(\eta(x_{1,T}, x_{2,T}))}{N}$$

All bivariate random variables $\mathbf{M}(B_j)$ are identically distributed as $\mathbf{M}(B) = (M_1(B), M_2(B))$ and we note their common cdf H , and \bar{H} the corresponding survival function. We further denote H_i the marginal cdf of M_i , $i = 1, 2$. Under the conditions $h\theta_1 \geq 1$, $h\theta_2 \geq 1$ and $h\theta(x_1, x_2) \geq 1 \forall (x_1, x_2) > 0$, the variables $(\mathbf{M}(B_j))_{j=1, \dots, n_b}$ can be considered independent.

Proposition 2 (Waiting time). *In the context above, for an integer $1 \leq k \leq n_b$ and scalars $x_1, x_2 > 0$, we have:*

$$\mathbb{P}[\eta(x_1, x_2) \leq k] = 1 - \left[1 - \bar{H}(x_1, x_2)\right]^k.$$

Furthermore,

$$\lim_{n_b \rightarrow \infty} \mathbb{E}(\eta(x_1, x_2)) = \frac{1}{\bar{H}(x_1, x_2)}.$$

The proof is given in Appendix C. Proposition 2, together with the Definition 3 of the return period, provides a way of computing the return period of nonconcurrent bivariate compound events. Specifically, as the number n_b of blocks becomes very large, we get

$$T = \frac{h}{N\bar{H}(x_{1,T}, x_{2,T})}. \quad (16)$$

Even though nonconcurrent compound events are bivariate in nature, it is interesting to see how Equation (16) would behave in a univariate context. In this case, when $x_i \rightarrow \infty$, $H_i(x_i) \simeq F_i^{h\theta_i}(x_i) \simeq 1 - h\theta_i \bar{F}_i(x_i)$ for $i = 1, 2$. Hence, the univariate return period reduces to $T \simeq 1/N\theta_i \bar{F}_i(x_{i,T})$ for very large return values. As expected, it does not depend on h . Compared to the usual expression $T_i = 1/N\mathbb{P}(X_i > x_{i,T})$, there is a factor θ_i in the denominator relating to the temporal dependence in the time series $(X_i)_{i=1, \dots, n}$, as discussed in Section 4.1.

5.2 | Properties of the Nonconcurrency

In a bivariate context, the nonconcurrent return period can be influenced by the value of h and Proposition 3 will show that it actually depends of the nature of the asymptotic dependence. As seen in Section 2.2.1, the process (X_1, X_2) can be asymptotically dependent or independent. In this specific part, for ease of exposition, we suppose that $\theta(x_1, x_2) = 1$ for any $(x_1, x_2) > 0$, that is, the variables are temporally independent. Similar expressions can be obtained for general extremal indices smaller than 1 at the price of heavier expressions, but without modifying the general behavior with respect to h , χ and $\bar{\chi}$. For simplicity, they are not presented here.

Proposition 3. *For scalars $x_{1,T}, x_{2,T} > 0$ such that $F_1(x_{1,T}) = F_2(x_{2,T}) = v$ is close to 1, the bivariate nonconcurrent return period can be expressed:*

- In case of asymptotic dependence (i.e., when $\bar{\chi} = 1$): $T \simeq \frac{1}{N\chi(v)(1-v)}$.
- In case of asymptotic independence and positive association (i.e., when $\chi = 0$ and $0 \leq \bar{\chi} < 1$): $T \simeq \frac{1}{N((1-v)^{\alpha(v)} + (h-1)(1-v)^2)}$, with $\alpha(v) = 2/(\bar{\chi}(v) + 1)$.
- Perfect independence corresponds to $\chi = 0$ and $\bar{\chi} = 0$. The previous case thus reduces to: $T \simeq \frac{1}{Nh(1-v)^2}$.
- In case of asymptotic independence and negative association (i.e., when $\chi = 0$ and $-1 < \bar{\chi} < 0$): $T \simeq \frac{1}{N(h-1)(1-v)^2}$.

The proof is given in Appendix D. For an asymptotically dependent compound event, the return period is thus approximately independent of the nonconcurrency parameter h , and it is proportional to the univariate return period with a factor $1/\chi$. For perfect dependence (i.e., $\chi = 1$), the bivariate return period is quite logically equal to the univariate one. For an asymptotically independent compound event, the nonconcurrent return period is approximately proportional to $1/h$, the value of $\bar{\chi}$ greatly influencing the exact nature of the dependence of $T(h)$. It is important to emphasize that the expressions in Proposition 3 should not be used directly because they rely on asymptotic arguments. Their purpose is to illustrate how the return period varies with the nonconcurrency block size h .

$$T \simeq \frac{h}{N \left[1 - F_1(x_{1,T})^{h\theta_1} - F_2(x_{2,T})^{h\theta_2} + (F_1(x_{1,T}) + F_2(x_{2,T}) - 1 + \bar{F}(x_{1,T}, x_{2,T}))^{h\theta(x_{1,T}, x_{2,T})} \right]} \quad (19)$$

In practice, and this will be illustrated on the case studies, the notion of nonconcurrency in case of asymptotic dependence is of little use because the bivariate return period is approximately constant with respect to h . In contrast, the nonconcurrency of a compound event is relevant in case of asymptotic independence.

5.3 | Application to the GPD-Copula Approach

When using the GPD-copula approach, the tail distribution of each marginal is modeled by a GPD, denoted G_i , with $i = 1, 2$ and the dependence is represented by a copula, C . We denote $H_i^{u_i}(x_i) = \mathbb{P}(M_i \leq x_i | M_i > u_i)$ the cdf of the tail of M_i for $i = 1, 2$ and $x_i \geq u_i$, hence the approximation $H^u(x_i) \simeq G_i(x_i)$. We further denote $H^{u_1, u_2}(x_1, x_2) = \mathbb{P}(M_1 \leq x_1, M_2 \leq x_2 | M_1 > u_1, M_2 > u_2)$.

Proposition 4. *In the GPD-copula approach, for each variable $i = 1, 2$, the univariate return period is*

$$T_i = \frac{h}{N\bar{H}_i(u_i)(1 - H^u(x_{i,T}))}, \quad (17)$$

with $\bar{H}_i \simeq 1 - F_i^{h\theta_i}$. The bivariate return period is:

$$T = \frac{h}{N\bar{H}(u_1, u_2)[1 + C(H_1^{u_1}(x_{1,T}), H_2^{u_2}(x_{2,T})) - C(H_1^{u_1}(x_{1,T}), 1) - C(1, H_2^{u_2}(x_{2,T}))]} \quad (18)$$

with $\bar{H}(u_1, u_2) \simeq 1 - F_1^{h\theta_1}(u_1) - F_2^{h\theta_2}(u_2) + F^{h\theta(u_1, u_2)}(u_1, u_2)$.

The proof is given in Appendix E. Notice that in the light of Section 4.1, the approximation $H^u(x_i) \simeq G_i(x_i)$ holds for any h . Hence, the quantity $N[1 + C(H_1^{u_1}(x_{1,T}), H_2^{u_2}(x_{2,T})) - C(H_1^{u_1}(x_{1,T}), 1) - C(1, H_2^{u_2}(x_{2,T}))]$ in the denominator of Equation (18) does not depend on h , and can be denoted N^* . The return period thus becomes $T = h/N^*\bar{H}(u_1, u_2)$. Proposition 3 can then be applied with N^* instead of N and $v^* = F_1(u_1) = F_2(u_2)$.

5.4 | Application to the bi-GPD Approach

As seen in Section 5.1, the univariate return period is $T_i \simeq 1/N\theta_i\bar{F}_i(x_{i,T})$. Based on the EGPD modeling, F_i is directly accessible and the computation is straightforward.

Proposition 5. *In the bivariate generalized Pareto approach, the bivariate return period becomes:*

The proof is a straightforward application of the general expression of $\bar{H}(x_1, x_2)$. It is detailed in Appendix F. Results of Proposition 3 can directly be applied with the bi-GPD approach. It should however be kept in mind that the bi-GPD approach is unable to represent asymptotic independence.

5.5 | From Return Period Curves to Return Period Surfaces

The definition of the waiting time used in Proposition 2 paves the way for a generalization of the relationship between the return period and the return level. For $k \in \mathbb{N}^*$, we denote $t = kh/N$ and we define the probability $p(x_1, x_2, t)$ by:

$$p(x_1, x_2, t) = \mathbb{P} \left[\eta(x_1, x_2) \leq \frac{Nt}{h} \right] = 1 - \left[1 - \bar{H}(x_1, x_2) \right]^{Nt/h}. \quad (20)$$

It is also possible to see t as a function of (x_1, x_2, p) :

$$t(x_1, x_2, p) = \frac{h \log(1-p)}{N \log \left[1 - \bar{H}(x_1, x_2) \right]} \simeq \frac{-h \log(1-p)}{N\bar{H}(x_1, x_2)}. \quad (21)$$

Setting $p = 1 - e^{-1}$ in Equation (21) leads to the usual return period defined in 3. Indeed, when considering the probability that the waiting time is less than its expectation, we get:

$$\begin{aligned}
 p(x_1, x_2, h\mathbb{E}(\eta(x_1, x_2)/N)) &= \mathbb{P}[\eta(x_1, x_2) \leq \mathbb{E}(\eta(x_1, x_2))] \\
 &= 1 - [1 - \overline{H}(x_1, x_2)]^{\mathbb{E}(\eta(x_1, x_2))} \\
 &= 1 - \exp\left(\frac{1}{\overline{H}(x_1, x_2)} \log(1 - \overline{H}(x_1, x_2))\right) \simeq 1 - e^{-1}.
 \end{aligned}$$

Therefore, the return period T defined by the expectation of the waiting time corresponds to $t(x_1, x_2, 1 - e^{-1})$.

Instead of using $p = 1 - e^{-1}$ for defining the return period, any probability $p \in (0, 1)$ could be considered for the relationships between (x_1, x_2) , t and p in Equations (20) and (21). This is somehow similar to the well-known Intensity-Duration-Frequency (IDF) curves in hydrology, (Schlef et al. 2023) with an important difference: the ‘‘Intensity’’ and the ‘‘Frequency’’ are the x_i s and t respectively in (21), but the ‘‘Duration’’ is replaced by the probability p . Despite these differences, the usage is similar. It will be illustrated in Section 7.

6 | Simulations

Simulations are performed to compare the bi-GPD and the GPD-copula approaches on controlled conditions and to assess the suitability of each of them for the analysis of the CEs studied in Section 7. Figure 2 shows a workflow diagram detailing the main steps (with their associated Sections) for both approaches. To compare the approaches, three quantities of interest are computed for each approach: the return period T , the joint excess probability $\overline{H}(x_1, x_2)$, and the estimators of the coefficients of extremal dependence $(\chi, \overline{\chi})$ presented in Section 2.2. The estimators are denoted $(\hat{\chi}, \hat{\overline{\chi}})$. These quantities are computed based on the generated random samples and compared to the theoretical values deduced from the probability models. First, the estimation

of the parameters in practice is explained. Then, the simulation protocol is presented and the results are discussed.

6.1 | Estimation in Practice

6.1.1 | Estimation for the Extremal Index

In the univariate context, the extremal index is estimated using the Dgaps estimator proposed in Holešovský and Fusek (2020), available in the R package `exdex` (Northrop and Christodoulides 2023). The bivariate extremal index is estimated using an idea from Smith and Weissman (1996) and continued in Robert (2008): the bivariate extremal index on (x_1, x_2) is equal to the univariate extremal index of the sequence $\max(-\tau_1(x_1)/\log F_1(X_1), -\tau_2(x_2)/\log F_2(X_2))$. It is then also estimated with the Dgaps estimator. The estimators are denoted $\hat{\theta}$. In order to ensure that the bivariate return period is positive and greater than the univariate ones, a constraint is placed on the estimation of the bivariate extremal index. We impose:

$$\begin{aligned}
 \max[\hat{\theta}_1 \overline{F}_1(x_1)/(1 - F(x_1, x_2)), \hat{\theta}_2 \overline{F}_2(x_2)/(1 - F(x_1, x_2))] \\
 \leq \hat{\theta}(x_1, x_2) \leq \hat{\theta}_1 \overline{F}_1(x_1)/(1 - F(x_1, x_2)) \\
 + \hat{\theta}_2 \overline{F}_2(x_2)/(1 - F(x_1, x_2))
 \end{aligned}$$

6.1.2 | Estimation for the GPDs

The Maximum Likelihood Estimation (MLE) of the parameters of the GPD is done with the R package `tea`. Desirable properties of this estimator, such as consistency, asymptotic normality and reliable confidence intervals, rely on an independence assumption. The data under scrutiny in this article are inherently temporally dependent, as will be apparent in the next section. Considering the maxima of clusters, (see Section 4.1), an operation known as

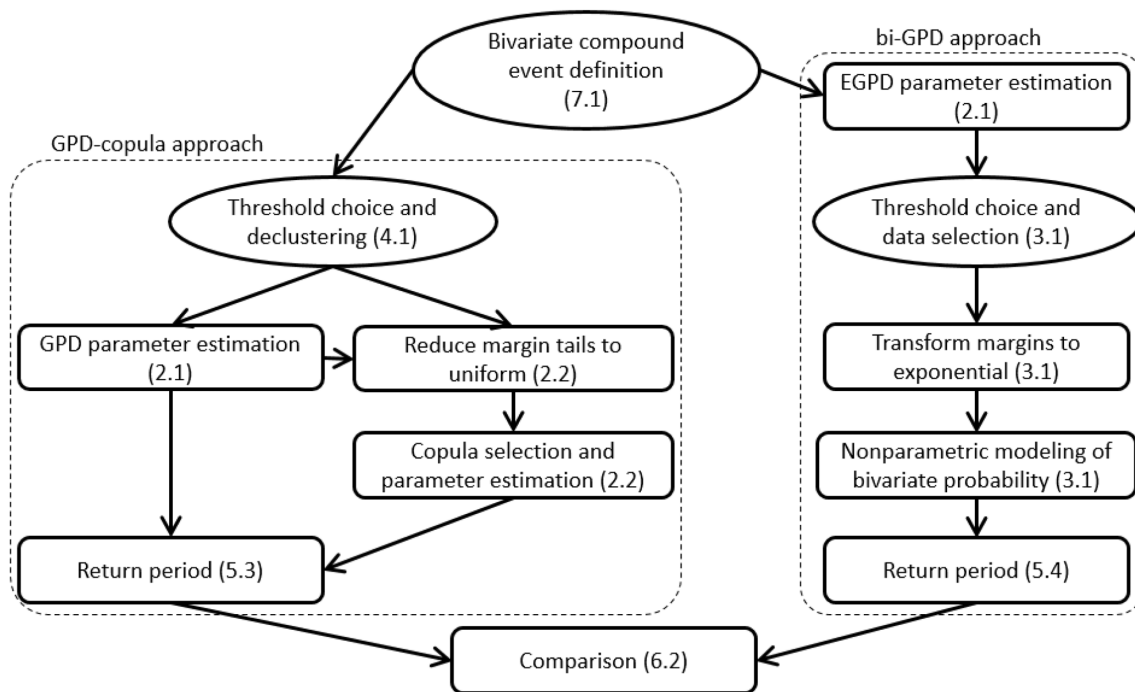


FIGURE 2 | Workflow diagram of the main steps (with their associated Sections) for the two proposed approaches. Elliptical boxes represent modeling decisions and rectangular boxes represent statistical treatment.

declustering allows us to assume independence. The size of these clusters is taken as the first integer larger than $1/\hat{\theta}$, with $\hat{\theta}$ being estimated as presented above.

Finally, a threshold equal to the 95th quantile is employed for all variables. This value is regarded as sufficiently high to apply the theorems while also ensuring sufficient data points for reliable estimation. Other high quantiles have been considered, without significant impact on the estimates.

6.1.3 | Estimation for the EGDs

The estimation method proposed in Naveau et al. (2016) available in the R package `mev` (Belzile 2023) is restricted to $\xi \geq 0$. Legrand et al. (2023) argued that the EGD can handle $\xi < 0$, relaxing the constraint in Naveau et al. (2016). In this work, a modified version of the package `mev` allowing for $\xi \geq -0.5$ is used to find the MLE of the parameters of the EGD.

6.1.4 | Estimation for the GPD-Copula Approach

The parameters of the copula are estimated using the R package `VineCopula` with MLE. As in the univariate case, declustering is necessary and is performed in a similar manner: the selected points are the maxima of each variable within a block of size h , provided each maximum exceeds its respective threshold. Notice that these points may not correspond to actual data values, because the maxima of the variables may not occur at the same time step. This is standard procedure in multivariate extreme value analysis, as outlined in Beirlant et al. (2006) and in Haan and Ferreira (2006).

The copula is selected from a set of classical parametric copula families, see the reference of the `VineCopula` package. The optimal copula family is chosen based on the Bayesian Information Criterion (BIC), averaged over quantiles ranging from the 0.90 to the 0.98 probability levels, with a lag of 0.01. The BIC of the copula families requiring an association (positive or negative) different from the association estimated on the data are set to 0 at this quantile. Only copula families that have the same association than the data at the 0.95 quantile level are considered for selection. This process aims to ensure a more robust copula, as the arbitrary threshold may not adequately represent the dependence structure at high levels.

According to Kass and Raftery (1995), a BIC difference less than 2 units between two models indicates that these models are non significantly different with respect to the data. When this happens with the Gumbel being second, the Gumbel copula is nonetheless preferred because it is the only extreme value copula within the considered families.

Note that the probability $F(u_1, u_2)$ appearing in the bivariate return period expression in Equation (18) (Proposition 4) is estimated by its empirical frequency.

6.1.5 | Estimation for the bi-GPD Approach

The bivariate joint excess probability $\bar{F}(x_{1,T}, x_{2,T})$ in Proposition 5 is computed using Theorem 1. The cdf of Δ is its

empirical cdf and the integrals are numerically computed. The univariate cdfs appearing in Theorem 1 and in Proposition 5 are computed according to the EGDs estimated previously. The bivariate exceedance probability is estimated by its empirical frequency.

6.2 | Results Without Temporal Dependence

We first simulate data with no temporal dependence. Therefore, all extremal indices are equal to 1. Two random uniform variables are first generated according to a chosen copula: Independent, Gaussian, Gumbel and Joe. Except with the Independent, each copula is parametrized such that its Kendall's τ is equal to 0.2, 0.5 and 0.8. Then, EGD margins are obtained by marginally transforming the random variables obtained in the first step. Simple power EGDs are used (see Section 2.1), with parameters: $(\kappa, \sigma, \xi) = (2, 4, 0.02)$ and $(\kappa, \sigma, \xi) = (1.2, 3, 0.15)$ for the first and second margins, respectively.

Since the simulated margins are EGDs, the tail of the marginal distributions show by construction GPD behaviors. The GPD-copula approach is thus clearly applicable. The bi-GPD approach is non parametric, hence potentially suitable for any model of random variables. The only restriction is that it is based on a representation valid at high quantiles in presence of asymptotic dependence only. It is therefore expected that the bi-GPD works best for the higher values of Kendall's τ . The estimation procedures described in the previous section are then run for the GPD-copula and for the bi-GPD approaches at thresholds corresponding to a probability 0.95. The return levels are set at 36 and 45 on the two margins, which corresponds to quantiles of probability 0.9995 on both margins. The nonconcurrent parameter h is taken equal to 1, which represents co-occurrence. The impact of h will be illustrated and discussed in Section 7.3. 500 draws are performed, and each draw is made of 61 days \times 30 years = 1830 points to mimic the conditions of the first compound event studied in Section 7. The coefficients of extremal dependence χ and $\bar{\chi}$ are limits of the extremal dependence functions $\chi(v)$ and $\bar{\chi}(v)$ as $v \rightarrow 1$. Hence, for numerical reasons, the coefficients are approximated by their associated function computed for $v = 0.9999$. For the bi-GPD approach, Equation (11) is used to compute $\hat{\chi}$ in order to assess the difference to the theoretical value $\bar{\chi} = 1$ stated in Proposition 1.

For the simulations, the two proposed approaches are also compared to another parametric copula approach. In this approach, the margins are modeled by the EGD to model the entire range of the data and still have an accurate representation of the extremes. The copula is hence learned on all the data. This approach is named in this article the EGD-copula approach. Here, the copula is chosen to be an extreme value copula to better model the dependence in the extremes. The Gumbel, Galambos and Husler-Reiss copulas are selected for evaluation. The EGD-copula, the GPD-copula and the bi-GPD approaches are then compared in terms of joint exceedance probability and coefficients of extremal dependence.

The results of the bivariate nonconcurrent excess probability are presented in Figure 3 and the continuous rank probability score

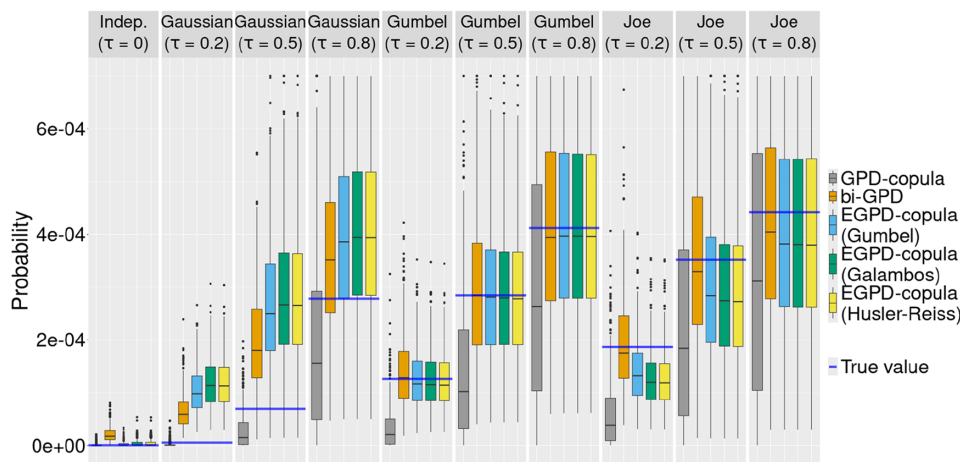


FIGURE 3 | Boxplots of the bivariate exceedance probabilities estimated on 60×31 bivariate random vectors with EGD marginals distributed according to an independent, Gaussian, Gumbel or Joe copula (500 repetitions). The value of the Kendall's tau (τ) is indicated in parenthesis.

TABLE 2 | Continuous rank probability score (CRPS) of the estimated bivariate exceedance probabilities derived from the simulations of Figure 3.

$\times 10^{-5}$	Indep.	Gaussian				Gumbel			Joe		
τ	0	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8	
GPD-copula	0.01 (0.001)	0.35 (0.01)	3.20 (0.14)	7.56 (0.38)	7.57 (0.22)	10.40 (0.49)	9.45 (0.46)	9.96 (0.34)	9.59 (0.54)	9.81 (0.50)	
bi-GPD	1.26 (0.05)	4.29 (0.11)	7.72 (0.33)	4.93 (0.34)	1.59 (0.08)	3.32 (0.15)	4.97 (0.22)	2.07 (0.09)	4.30 (0.18)	5.28 (0.23)	
Gumbel	0.06 (0.01)	7.51 (0.17)	13.26 (0.46)	6.85 (0.44)	1.41 (0.06)	3.11 (0.14)	4.93 (0.22)	3.23 (0.18)	4.68 (0.24)	5.43 (0.24)	
Galambos	0.10 (0.01)	8.69 (0.18)	14.50 (0.49)	7.23 (0.51)	1.37 (0.06)	3.08 (0.14)	4.91 (0.22)	4.11 (0.20)	4.94 (0.26)	5.43 (0.25)	
H.-R.	0.10 (0.01)	8.67 (0.19)	14.47 (0.52)	7.22 (0.45)	1.37 (0.06)	3.06 (0.14)	4.91 (0.22)	4.17 (0.19)	4.97 (0.27)	5.43 (0.25)	

Note: Values in parentheses represent standard errors. In each column, the lowest value and values not significantly different from the lowest value are in bold. A value is considered to be not significantly different from another value if it is in the standard deviation interval of this other value.

(CRPS, Matheson and Winkler 1976), calculated between the distribution of the estimated probabilities and the true probability for each simulating copula, are summarized in Table 2. On Figure 3, the blue lines represent the true values, computed from the theoretical distribution used to generate the data and the boxplots correspond to the 500 repetitions of 1830 values. Notice that the bivariate excess probability increases with the strength of the dependence parametrized by τ . The EGD-copula approach leads to very similar results with either the Gumbel, the Galambos or the Husler-Reiss copula. From now on, “EGD-copula approach” will refer to the approach with any of these three copulas. In 7 out of 10 cases, the bi-GPD approach leads to a median closer to the true value than the GPD-copula approach or the EGD-copula approach. The true value is included 7 times out of 10 in the 50% box for the bi-GPD approach, whereas it is only the case 4 times out of 10 for the GPD-copula approach and 6 times for the EGD-copula approach.

In these 7 cases (Gaussian with $\tau = 0.8$, Gumbel and Joe), the bi-GPD approach also presents a lower CRPS than the GPD-copula approach (see Table 2). The bi-GPD approach has

the lowest CRPS, or a CRPS not significantly different from the lowest one, when simulating either from the Gaussian copula with $\tau = 0.8$ or from the Joe copula. When simulating from the Gumbel copula with $\tau = 0.2$ or $\tau = 0.5$, the CRPS of the EGD-copula and bi-GPD approaches are almost identical. It can be noted that the CRPS of the EGD-copula approach is slightly lower with a Galambos or Husler-Reiss copula than with a Gumbel copula when simulating from a Gumbel copula, but their difference is not significant when considering the bootstrap standard deviations. Further analyses (not shown) indicated that it can be explained by biases in the estimation of the parameters of the EGD, which slightly modify the dependence structure, and lead to the Gumbel copula not always having the lowest CRPS. Indeed, the same simulations using the true parameters for the EGD (not shown) lead to the CRPS of the EGD-copula approach with the Gumbel copula to be the lowest.

The EGD-copula approach never returns a probability equal to 0, which would lead to infinite return periods. The bi-GPD approach very rarely returns a probability equal to 0 (14 occurrences over 5000 simulations), whereas the GPD-copula

approach is more likely to do so (126 zeros and 33 NAs). The NAs returned by the GPD-copula approach are due to insufficient points above the thresholds to select a copula family and to estimate its parameters. These NAs can be considered as zeros. It was shown in Section 2.2 that the bi-GPD approach is not adapted for independent or asymptotically independent variables: the GPD-copula approach easily outperforms the bi-GPD approach and the EGD-copula approach when simulating from the Independent copula and from the Gaussian copula with a low Kendall's τ . Conversely, the bi-GPD approach performs better for asymptotically dependent data (Gumbel and Joe copulas) as well as for asymptotically independent data with a high Kendall's τ .

The results on the bivariate return period, shown in Figure S1, lead to similar interpretation, because the bivariate return period is proportional to the inverse of the bivariate exceedance probability (see Equation (16)). The results on the coefficients of extremal dependence, shown on Figure 4, highlight the inability of the bi-GPD and EGD-copula approaches to represent independence or asymptotic independence. When simulating from the Independent copula or from the Gaussian copula, the estimators of $\hat{\chi}$ and $\hat{\bar{\chi}}$ with the bi-GPD and EGD-copula approaches are highly biased, while the estimators obtained with the GPD-copula approach are unbiased or less biased. In contrast, for asymptotic dependent copulas, such as the Gumbel and the Joe copulas, the bi-GPD estimators show less bias and smaller variance than the copula ones. The estimators of the EGD-copula approach estimate the values for the Gumbel copula with high accuracy, but show biased results for the Joe copula.

In light of these results, the bi-GPD approach is the best approach in case of asymptotic dependence or asymptotic independence when $\tau \geq 0.8$. The GPD-copula approach is favored in case of independence or asymptotic independence when $\tau < 0.8$. The EGD-copula approach showed some rigidity: as the copula is imposed to be an extreme value copula, this approach is not suited for all data. The EGD-copula approach is not considered in the rest of this work as it is outperformed by the two other approaches.

In order to better understand the variability in each approach, we decomposed the contribution of the error entailed by the univariate component from that entailed by the bivariate component. Firstly, on both margins, the bias and the RMSE between the estimated univariate exceedance probabilities and the true values are computed. For the GPD-copula approach, which involves marginal GPDs, the averaged bias on the first margin (resp. second margin) is -1.4×10^{-5} (resp. 1.4×10^{-6}), which is less than 3% of the true value and the RMSEs are between 3.1×10^{-4} and 4.2×10^{-4} . For the bi-GPD approach, which involves marginal EGDs, the averaged bias on the first margin (resp. second margin) is 2.7×10^{-5} (resp. 2.1×10^{-5}), which is less than 6% of the true value, and the RMSEs are between 2.1×10^{-4} and 2.7×10^{-4} . In terms of univariate exceedance probabilities at such extreme quantiles, the EGD modeling thus appears to be slightly more biased (but still with a low bias) and to have a lower estimation variance, hence a lower RMSE than the GPD modeling. Secondly, the bivariate exceedance probabilities are computed directly using the true univariate parameters, hence the margins are not estimated. The results (see Figures S2, S3, and S4)

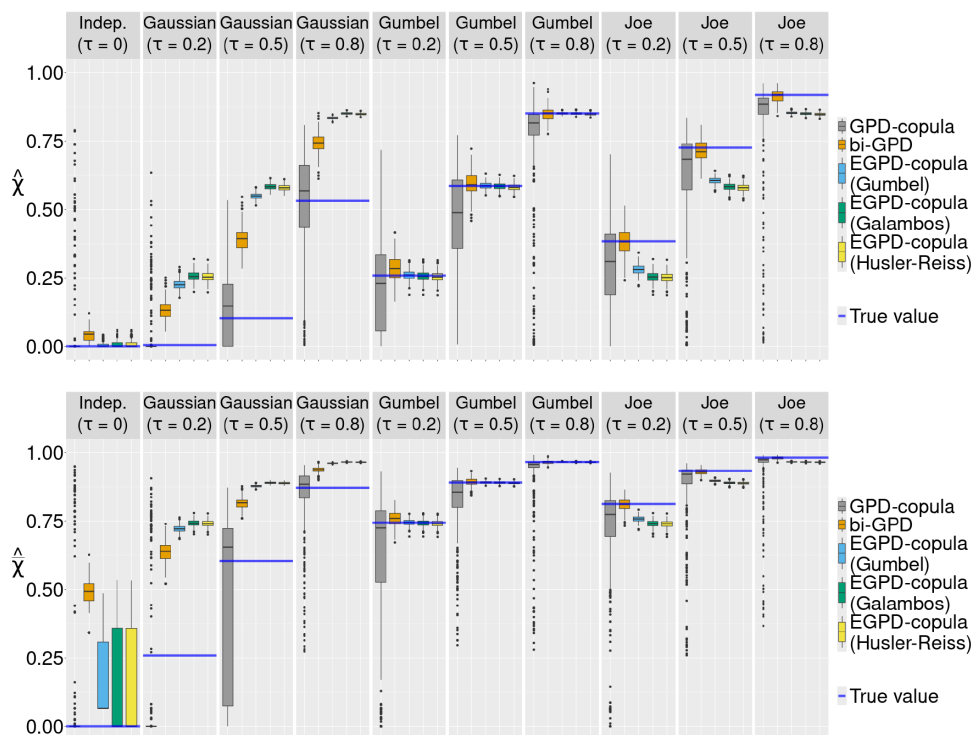


FIGURE 4 | Boxplots of $\hat{\chi}$ (top) and $\hat{\bar{\chi}}$ (bottom) estimated on 60×31 bivariate random vectors with EGD marginals distributed according to an independent, Gaussian, Gumbel or Joe copula (500 repetitions). In each panel, estimation with the GPD-copula approach are in gray (on the left) and with the bi-GPD approach are in orange (on the right). The theoretical values deduced from the probability models are represented by blue lines. The Kendall's tau (τ) are indicated in parenthesis.

are similar to those obtained when the margins were estimated, but with a lower variance. The bi-GPD approach remains unbiased and it has a lower variance than the GPD-copula approach for the asymptotically dependent copulas, and the GPD-copula approach remains the least biased in case of independence and asymptotic independence. A great part of the estimation error can therefore be attributed to the estimation of the marginals.

6.3 | Results With Temporal Dependence

Finally, for the Gaussian copula ($\tau = 0.5$ and $\tau = 0.8$) a comparison between variables with and without temporal dependence is performed. The temporal dependence along the 61 consecutive days is modeled by a stationary process with a temporal correlation function equal to $1/(1 + |i - j|/r)$, with i and j being days, and r being a temporal range. Here, we set $r = 10$. The bivariate extremal index is between 0.09 and 0.20, hence $h = 11$ verifies the hypothesis $h \geq 1/\theta(x_1, x_2)$. The results of the bivariate nonconcurrent excess probability are presented in Figure 5. The first observation is that the true value of joint excess probability is different with and without temporal dependence. If we could set $h = 1/\theta(x_1, x_2)$, the true values would be equal with or without temporal dependence because h would exactly balance out $\theta(x_1, x_2)$. Because the extremal index is not known but estimated, and due to discretization effect, this is rarely the case, and the true exceedance probability differs. The estimation is less accurate in presence of temporal dependence because, for both approaches, there is less effective data to estimate the parameters. With fewer effective data we also obtain more probabilities exactly equal to 0. For $\tau = 0.5$ ($\tau = 0.8$), in presence of temporal dependence we get 97 (90) zeroes over 200 draws for the GPD-copula approach and only 22 (0) zeroes for the bi-GPD approach. In contrast, without temporal dependence, we get 6 (5) zeroes for the GPD-copula approach and 0 (0) zeroes for the bi-GPD approach. The temporal dependence significantly decreasing the accuracy of the approaches is also illustrated in the Supporting Information, in Figure S5 for the bivariate return period and in Figure S6 for the coefficients of extremal dependence.

Overall, the bi-GPD approach seems more accurate and robust than the GPD-copula approach as long as the data present some

sort of asymptotic dependence, or asymptotic independence with strong positive association.

7 | Application to Compound Extreme Precipitations

7.1 | Events and Data Presentation

In May 2016, repeated precipitations over central France caused the flooding of the Seine and Loire rivers (van Oldenborgh et al. 2016): the accumulation of precipitation was maximal in both catchment areas over a 20-day period. Peak river levels were reached within a few days interval, on June 3 for the Seine in Paris and on June 7 for the Loire in Nantes. This time lag motivates the definition of nonconcurrent compound events, as seen in Section 4. This event, referred to as the Seine/Loire event, will be analyzed as a bivariate spatially compound event according to the classification proposed in Zscheischler et al. (2020), involving one variable for each watershed. To describe a catchment's response to precipitation accumulation and natural runoff, the Antecedent Precipitation Index (API) is used. It is defined on day j as:

$$API_j = \sum_{i=1}^m k^{i-1} PR_{j-i}, \quad (22)$$

where PR_d is the total precipitation at day d . This definition depends on two parameters, k and m . The integer m represents the number of days over which the API is calculated, and k is a proxy for the watershed's runoff. The API and its parameters are discussed by Kohler and Linsley (1951), with the optimal value for k being analyzed in detail by Li et al. (2021). After preliminary analyses based on model stability, $m = 17$ and $k = 0.87$ were set for both watersheds. For this study, ERA5 data (Hersbach et al. 2020) from 1992 to 2021 on a $1^\circ \times 1^\circ$ grid are used. The geographic extent of both watershed is discretized using the ERA5 grid (see Figure 6). Climate is supposed stationary over this 30-year period. Daily precipitation is spatially averaged across each watershed and two time series of daily APIs are calculated between May 1st and June 30th for the 30 years. Only May and June are considered to conserve a similar rainfall regime in the

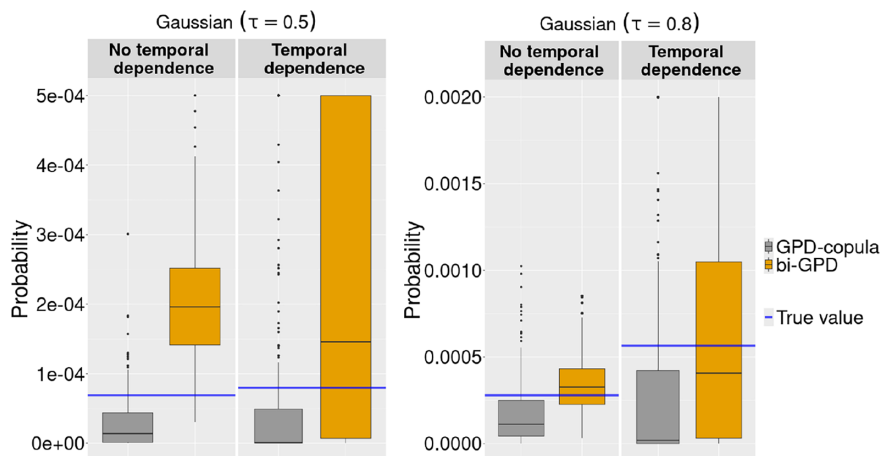


FIGURE 5 | Boxplots of the bivariate exceedance probabilities estimated on 1830 bivariate random vectors (200 repetitions) with a Gaussian copula, with and without temporal dependence.

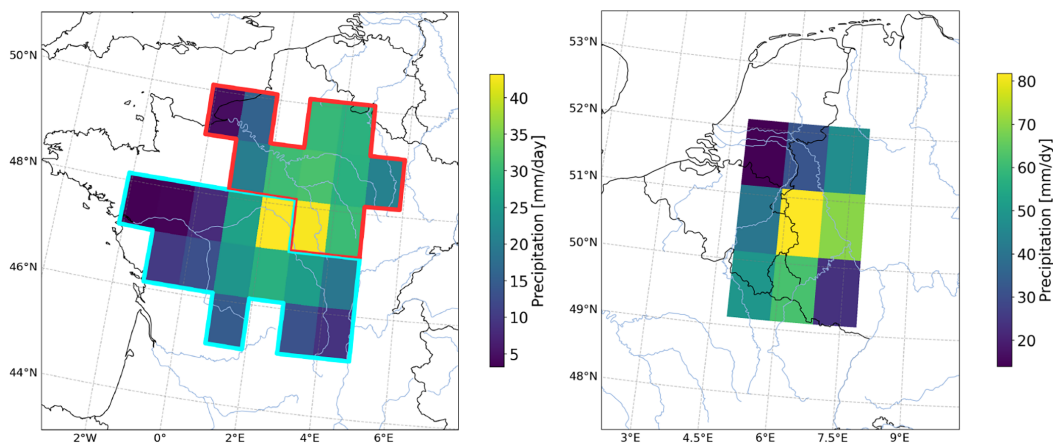


FIGURE 6 | Left: Seine (red area) and Loire watersheds (blue area) on a $1^\circ \times 1^\circ$ grid with the precipitation on the 30/05/2016 (ERA5 reanalysis data). Right: selected gridded area for the Germany/Belgium event of 14/07/2021, and the associated precipitation values.

TABLE 3 | Estimation of return periods and extremal coefficients for spring flooding events across different methods and basins.

Approach	Sample size			Return periods			Extremal coefficients	
	Seine	Loire	Both	Seine	Loire	Both	$\hat{\chi}$	$\hat{\bar{\chi}}$
Empirical							0.49	0.61
GPD-copula	43	41	26	123	111	979	0.42	0.83
bi-GPD	1830	1830	139	651	1160	2156	0.49	0.86
	PR	API	Both	PR	API	Both	$\hat{\chi}$	$\hat{\bar{\chi}}$
Empirical							0.08	0.08
GPD-copula	131	43	19	317	6	22458	0.00	0.00
bi-GPD	2760	2760	265	1395	7.5	1639	0.08	0.57

data. The maximum API values reached during the May/June 2016 period are considered as the event's return levels: 39.6 mm for the Seine and 41.1 mm for the Loire. We consider that there is a compound event when both APIs exceed their respective 2016 maxima within a sufficiently close time frame. The nonconcurrent parameter for this event is set to $h = 4$ days, which is also the first integer larger than $1/\min(\theta)$.

On July 14, 2021, extreme precipitations over the Ahr basin, a river between Belgium and Germany, led to an important flood that claimed many lives. This event, referred to as the Germany/Belgium event, has been analyzed as a preconditioned compound event in Mohr et al. (2023): the soil moisture level a few days before rainfall was a decisive factor in the occurrence of flooding. This event is modeled using the averaged daily precipitation (PR) over a selected area (see Figure 6) as the main variable and the API in the same area as the preconditioning variable. Here, the API is used as a proxy for soil moisture, with parameters $m = 30$ and $k = 0.9$ (Blanchard et al. 1981; Teng et al. 1993; Schröter et al. 2015; De Moraes et al. 2024). To account for seasonality, only daily values over the summer months (June, July and August) of each year are considered, yielding two daily time series. We consider that there is a (preconditioned) compound event when both values exceed the values recorded on the 14th of July 2021. The nonconcurrent parameter is set to $h = 7$ days, the first integer larger than $1/\min(\theta)$.

For both examples, the variables are assumed to satisfy the conditions allowing the use of extreme value theory (see Section 4.1). Since variables are time dependent, a declustering algorithm is first applied, as presented in Section 4, and then the model parameters are estimated as described in Section 6.1.

7.2 | Results

The return periods for both events are shown in Table 3 and in Figure 7 with red diamonds. To assess the estimation uncertainty, data are bootstrapped 200 times (i.e., 200 resamples with replacement of the same number of data). Notice that the bootstrapping is limited to the data involved in the estimation of the parameters and in the computation of the return periods. For the GPD-copula approach, this corresponds to the couples exceeding a threshold (equal to the 95th quantile) on at least one margin, used both for the estimation of the GPD parameters and the copula parameters. For the bi-GPD approach, all data are bootstrapped since they are all used for the estimation of the parameters of the EGPD. Return periods are then computed for each resample, and represented by a boxplot. Empirical estimates of χ and $\bar{\chi}$ are computed at the 95th quantile.

7.2.1 | Seine/Loire Event

In the case of the Seine/Loire event (upper panel of Figure 7), the GPD-copula approach yields lower return periods than

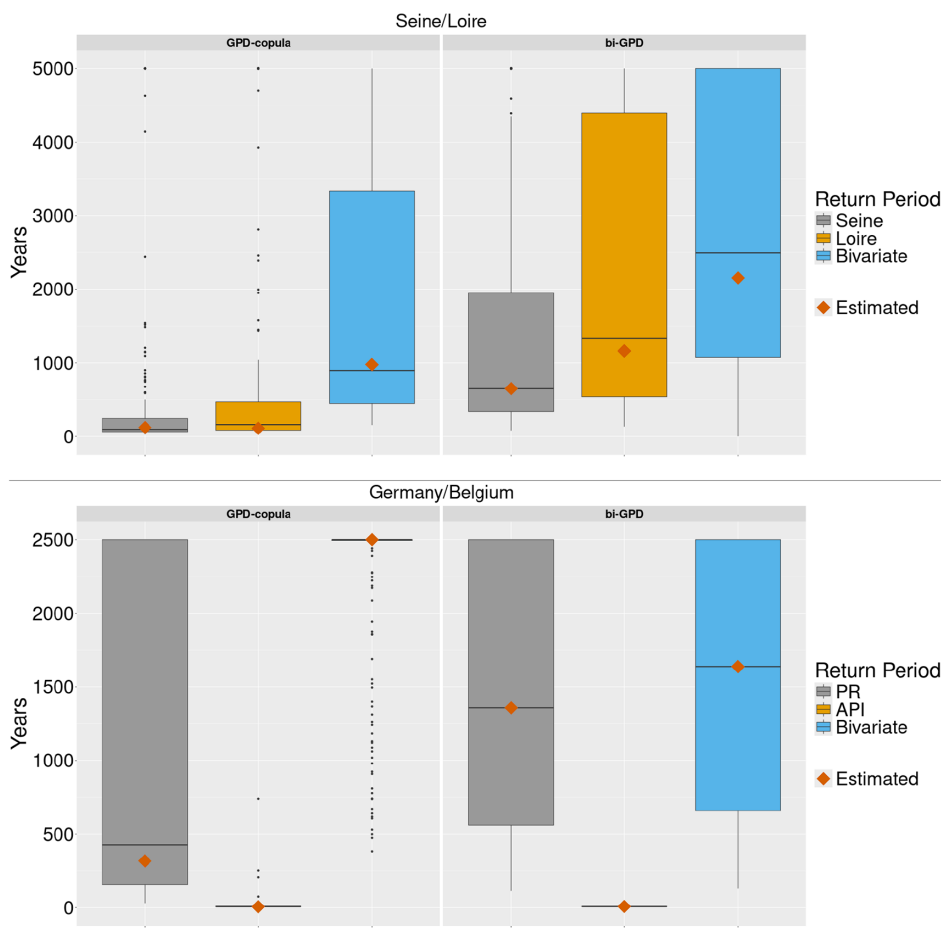


FIGURE 7 | Estimated univariate and bivariate return periods (red diamonds) and estimation uncertainty (200 bootstraps). Top: Seine/Loire event; bottom: Germany/Belgium event. For clarity, any value that exceeds the displayed limit (5000 years) on the Y-axis is censored to this value. This adjustment does not affect the statistical representation but may distort the visual interpretation of the variance of the boxplot.

the bi-GPD approach for both univariate and bivariate events. Specifically, for the GPD-copula approach (resp. for the bi-GPD approach), we get the following return periods: 123 (651) years for the Seine event, 111 (1160) years for the Loire event, and 979 (2156) years for the bivariate CE.

The empirical estimates of the coefficients of extremal dependence are $(\hat{\chi}, \hat{\bar{\chi}}) = (0.49, 0.61)$. For the GPD-copula and the bi-GPD approaches, the values computed according to Equations (5), (6) and (11) are equal to $(\hat{\chi}, \hat{\bar{\chi}}) = (0.42, 0.83)$ and $(\hat{\chi}, \hat{\bar{\chi}}) = (0.49, 0.86)$ respectively. According to Table 1, either $\chi = 0$ or $\bar{\chi} = 1$. Practical estimation may however differ from the theoretical result and values as high as 0.83 and 0.86 can be interpreted as being equal to 1, and therefore correspond to positive asymptotic dependence, considering that $\hat{\chi} \approx 0.5$. Plots of the upper tail dependence function defined in Equation (2) further confirm this result (not shown). As discussed in Section 6, the bi-GPD approach is better suited for events characterized by strong dependence between variables, particularly when there is temporal dependence. Consequently, the return periods estimated using the bi-GPD approach are likely to be less biased than those estimated with the GPD-copula approach for this event.

The boxplots shown in Figure 7 are wider for the bi-GPD approach than for the GPD-copula approach, thus indicating that the estimation variance is higher with the bi-GPD approach. However, it should be remembered that on the simulations reported in Section 6, the RMSE was smaller for the bi-GPD approach than for the GPD-copula approach. Consequently, it is expected that the bi-GPD return periods on the bivariate event is less biased than that computed with the GPD-copula approach.

Finally, although not clearly visible on Figure 7, the GPD-copula approach leads more frequently to infinite values (specifically: 7, 16, and 22 infinite values for the Seine event, Loire event, and bivariate CE, respectively, out of 200 resamples), whereas such occurrences are rarer for the bi-GPD approach (1, 3, and 9 infinite values for the Seine event, Loire event, and bivariate CE, respectively).

7.2.2 | Germany/Belgium Event

For the Germany/Belgium event (lower panel of Figure 7), the API return periods are similar across both approaches (6 years for the GPD modeling versus 7.5 years for the EGPD modeling). The return periods for the precipitation (PR) differ more significantly

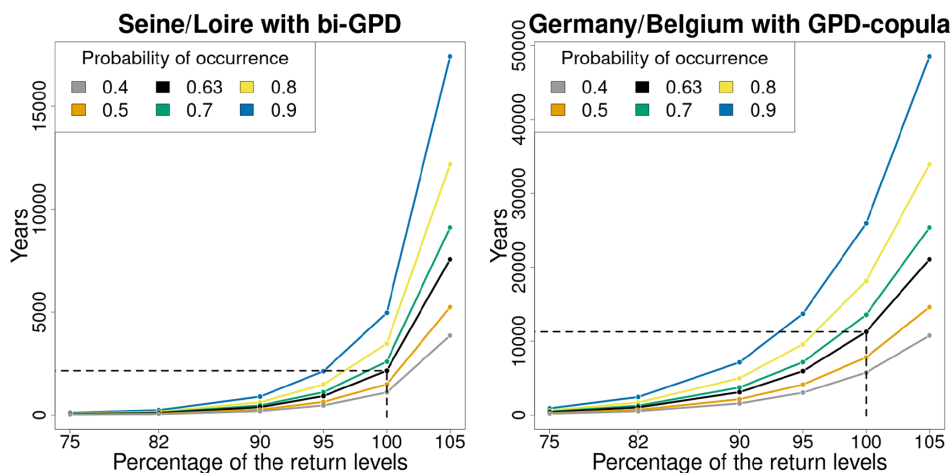


FIGURE 8 | Return period surfaces for different probabilities and percentage of the maximal observed value for the Seine/Loire event with the bi-GPD approach (Left) and for the Germany/Belgium event with the GPD-copula approach (Right).

(317 years for the GPD modeling versus 1395 years for the EGPD modeling). The bootstrapped boxplots for PR and API are essentially comparable between the EGPD and the GPD modeling. The most striking difference lies in the estimation of the nature of the asymptotic dependence.

Empirical estimates of the coefficients of extremal dependence give $(\hat{\chi}, \hat{\chi}) = (0.08, 0.08)$, which, for both coefficients, can be interpreted as equal to 0. Both approaches yield different coefficients of extremal dependence: $(\hat{\chi}, \hat{\chi}) = (0, 0)$ for the GPD-copula approach versus $(\hat{\chi}, \hat{\chi}) = (0.08, 0.57)$ for the bi-GPD approach. The averages computed from the bootstrapped data differ less but still significantly, with $(\hat{\chi}, \hat{\chi}) = (0.06, 0.13)$ for the GPD-copula approach and $(\hat{\chi}, \hat{\chi}) = (0.08, 0.57)$ for the bi-GPD approach. Hence, the Germany/Belgium CE is very likely an independent CE. As already discussed, the GPD-copula approach is more adapted to represent it since the bi-GPD is unable to model asymptotic independence.

This impacts the estimation of the bivariate return period values (22,458 years for the copula versus 1639 years for the bi-GPD). It also impacts the variability in the bootstrapped samples: we observe 28 infinite values and only 28 values below 2500 years for the GPD-copula approach as compared to only 3 infinite value for the bi-GPD approach. In case of asymptotic independence, the bivariate return period is proportional to the product of the univariate ones, thus leading to high variability and extremely high return period values if at least one of the univariate return period presents a high variability and high values. In contrast, the bi-GPD approach estimates the data as asymptotically dependent, since it inherently cannot estimate otherwise.

This event highlights two key conclusions: the bi-GPD approach is unsuitable for independent or asymptotically independent events, and the GPD-copula approach is highly sensitive to data. Therefore, due to its higher precision and stability, the bi-GPD approach should be preferred whenever the data show asymptotic dependence, while the GPD-copula approach should be favored in case of asymptotic independence, albeit with caution regarding to its sensitivity to data.

7.2.3 | Return Period Surfaces

As introduced in Section 5.5, return period surfaces can be computed and used for decision making. Examples of such return period surfaces are presented on Figure 8. For visual representations, values of the APIs and PR are on the x-axis, years are on the y-axis, and the surface is represented by a few curves corresponding to probability of occurrence in the time interval $[0, T_{\text{Years}}]$. Recall that the usual return period corresponds to $p = 0.63$. The API values reached in May/June 2016 for the Seine and the Loire watersheds are 39.6 and 41.1 respectively. The PR value reached on July 14 is 46.6 and the API value is 64.4 for the Germany/Belgium event. The other values correspond to fractions of these two values: 0.75, 0.82, 0.90, 0.95, 1, and 1.05 in that order. Only a few probabilities of occurrence are represented, in particular for $p = 1 - e^{-1} \approx 0.63$, where the values and time correspond to the return levels and usual return period. The dashed line shows how to read the bivariate return period value for the studied example.

Notice that other probabilities of occurrence or other return levels can be considered. Because of the convexity of the represented curves, a small change in the values or in the probability greatly affects the resulting “return period”. For example, on the Seine/Loire event analyzed with the bi-GPD approach, for return levels equal to 39.6 and 41.1 for the API of the Seine and the Loire respectively, the usual return period is 2 156 years (thus corresponding to $p = 0.63$). If the lower probability $p = 0.5$ was considered, the associated “return period” would decrease to 1 495 years (Table 4).

7.3 | Discussion Over the Nonconcurrency

Figure 9 presents the evolution of the return period as a function of the nonconcurrency parameter h for the two studied events and with the two approaches.

In case of asymptotic dependence, according to Proposition 3, the bivariate return period should be constant with respect to h . This is the case for the Seine/Loire CE, for which χ was estimated

TABLE 4 | Values and probability of empirical quantiles of percentage of return levels.

Fraction		75	82	90	95	100	105
Seine	Value (mm)	29.7	32.5	35.6	37.6	39.6	41.6
	Percentile	0.997	0.997	0.998	0.999	1.000	—
Loire	Value (mm)	30.8	33.7	37	39.1	41.1	43.2
	Percentile	0.996	0.998	0.999	0.999	1.000	—
PR	Value (mm/day)	35	38.2	42	44.3	46.6	49
	Percentile	0.999	0.999	0.999	0.999	1.000	—
API	Value (mm)	48.3	52.8	57.9	61.2	64.4	67.6
	Percentile	0.963	0.979	0.991	0.993	1.000	—

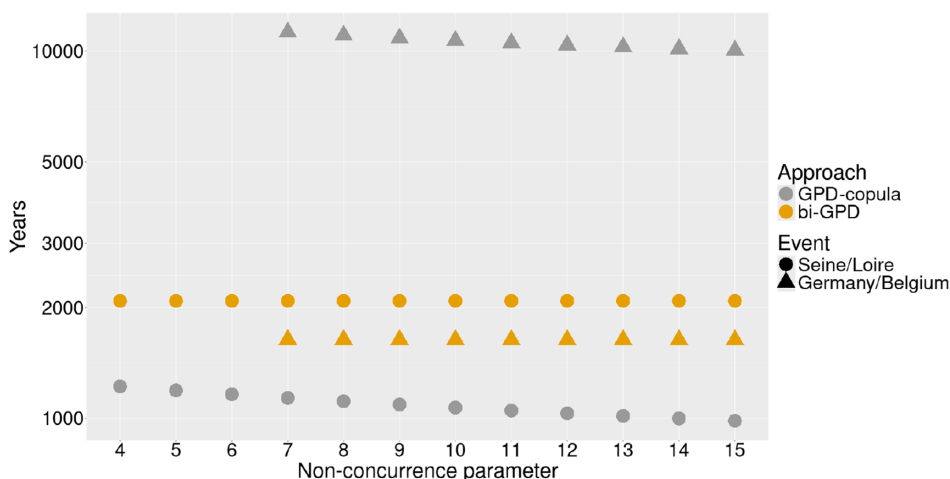


FIGURE 9 | Evolution of the bivariate return period as a function of the nonconcurrency parameter for both events (ERA5 data, logarithmic scale for the years).

around 0.50 in the previous section. Figure 9 shows that it is indeed the case when using the bi-GPD approach (orange circles). However, when using the GPD-copula approach, the bivariate return period for the Seine/Loire CE is found to decrease with h (Figure 9, gray circles). To explain this behavior we need to go back to the proof of Proposition 3. The assumption $\bar{F}_i(x_i) \simeq 0$ allowed us to find that the first order approximation of $\bar{H}(x_1, x_2)$ is proportional to h . In the GPD-copula approach, the declustering approach chosen to account for temporal dependence leads to assume that the behavior above the threshold u_i is independent of the choice of h . As a consequence, the first order expansion of $\bar{H}(x_1, x_2)$ is not a good approximation and the denominator in Equation (18) is a second order polynomial instead of being proportional to h .

In case of asymptotic independence, the bivariate return period should be approximately proportional to $1/h$, see Proposition 3. The bivariate return period with the bi-GPD approach for the Germany/Belgium event is found to be constant (Figure 9, orange triangles). It is important to recall that the bi-GPD approach presented in the first part of this work is not able to represent asymptotic independence. Any event represented with the bi-GPD will therefore be considered asymptotically dependent and, according to Proposition 3, will present a constant bivariate return period with respect to h . Hence, if the event is suspected to be asymptotically independent or if the nonconcurrency of the

compound event is of importance, the GPD-copula approach should be preferred.

8 | Conclusion and Perspectives

8.1 | Conclusion

In this paper, a representation of bivariate compound events based on bivariate generalized Pareto distribution was proposed and compared to the GPD-copula approach and the EGPD-copula approach. Based on statistical simulations, the bi-GPD approach showed to be less biased and more accurate than the GPD-copula approach for asymptotically dependent variables. A focus was made on temporally dependent variables in a peaks-over-threshold context, handled with the extremal index. The Seine/Loire CE motivated the introduction of “nonconcurrent” compound events, defined as variables exceeding a threshold at a relatively close time, for which formulas and discussions were proposed in this work. The return period of nonconcurrent CE is independent of the nonconcurrency parameter h in case of asymptotic dependence, whereas it is inversely proportional to h in case of asymptotic independence, which makes this new concept of particular interest for events represented by asymptotically independent data. Finally, the two approaches were applied

to two examples of nonconcurrent compound events, whose return periods have been detailed.

8.2 | Perspectives

However, some improvements remain possible. In the Germany/Belgium event, which is a preconditioned CE, the API had to reach a high threshold *before* the other variable (PR) reaches a high threshold in order to trigger the devastating event witnessed on the 14th of July, 2021. In both approaches presented above, we did not make this distinction and we considered that the two variables must belong to the same time window, regardless of their order of appearance. Future work should be devoted to account for this asymmetry. This could be done during the declustering step for the GPD-copula approach (not done here). It should also be accounted for in the formula of the nonconcurrent return period. This will probably necessitate a change in the definition of the waiting time (15) in future work.

Another shortcoming of the presented approaches is that they can only be used to analyze bivariate CE, whereas some CE involve more than two variables, like wildfires or convective storms (Villalobos-Herrera et al. 2021; Prein and Holland 2018). The GPD-copula approach can be extended to more variables through vine copulas (Kurowicka and Joe 2010). The bi-GPD approach could also be extended by considering pairs $\Delta_{ij} = Z_i - Z_j$ as in Madhar et al. (2024), which is the natural extension of Equation (8) in more than two dimensions. Alternatively, another approach, the Pareto processes, can be used to model bivariate CE, and can be extended in higher dimensions (Ferreira and De Haan 2014).

Moreover, recent developments in the study of the evolution of extreme events have highlighted the strong link between global warming and the occurrences of these events (Seneviratne et al. 2016; Li et al. 2019). Projecting the evolution of the frequency and magnitude of these events in a future climate has become a central issue (Kharin et al. 2018), and this is usually done by using the outputs of climate models. It is known that climate model outputs are statistically biased (Christensen et al. 2008), therefore it is common practice to correct them against observations, for example through quantile mapping (Haddad and Rosenfeld 1997), which is a popular univariate bias correction method. Compound events and multivariate representations naturally call for multivariate bias correction methods (François et al. 2020). However, questions still remain about the capacity of these univariate and multivariate bias correction methods to accurately correct the multivariate extremes. Can bias correction help better represent CE in the future? Are multivariate bias correction methods necessary when addressing CE?

Lastly, our proposed modeling framework could be applied in many different contexts. For instance, it could be useful in an attribution context to more effectively represent compound events. Extreme event attribution tries to measure how ongoing Human-induced climate change affects extreme events. It aims to determine whether such recent events can be explained by, or linked to, a warming atmosphere and are not simply due to natural variations. Precise modeling of complex events is thus needed, both in factual scenario (current climate) and counterfactual scenario (climate simulations without human

CO₂ emissions), and the approach proposed here would then be relevant for that.

Acknowledgments

We thank the Copernicus Climate Change Services for making the ERA5 reanalyses available. This work has been supported by the chair Geolearning, funded by ANDRA, BNP Paribas, CCR and the SCOR Foundation for Science. We also acknowledge the support of the COESION project funded by the French National program LEFE (Les Enveloppes Fluides et l'Environnement). This work also benefited from state aid managed by the National Research Agency under France 2030 bearing the references ANR-22-EXTR-0005 (Agence National de la Recherche (ANR)).

Funding

This work was supported by Agence National de la Recherche (ANR) (ANR-22-EXTR-0005).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in the Climate Data Store of Copernicus at <https://doi.org/10.24381/cds.adbb2d47>.

References

- Beirlant, J., Y. Goegebeur, J. Segers, and J. L. Teugels. 2006. *Statistics of Extremes: Theory and Applications*. John Wiley & Sons.
- Belzile, L. 2023. MEV: Modelling of Extreme Values.
- Bennet, M. J., D. G. Kingston, and N. J. Cullen. 2025. "Detection of Compound and Seesaw Hydrometeorological Extremes in New Zealand: A Copula-Based Approach." *EGU Sphere* 2025: 1–30.
- Bevacqua, E., D. Maraun, I. Hobæk Haff, M. Widmann, and M. Vrac. 2017. "Multivariate Statistical Modelling of Compound Events via Pair-Copula Constructions: Analysis of Floods in Ravenna (Italy)." *Hydrology and Earth System Sciences* 21, no. 6: 2701–2723.
- Bevacqua, E., D. Maraun, M. I. Vousdoukas, et al. 2019. "Higher Probability of Compound Flooding From Precipitation and Storm Surge in Europe Under Anthropogenic Climate Change." *Science Advances* 5, no. 9: eaaw5531.
- Bevacqua, E., M. I. Vousdoukas, T. G. Shepherd, and M. Vrac. 2020. "Brief Communication: The Role of Using Precipitation or River Discharge Data When Assessing Global Coastal Compound Flooding." *Natural Hazards and Earth System Sciences* 20, no. 6: 1765–1782.
- Blanchard, B. J., M. J. McFarland, T. J. Schmutge, and E. Rhoades. 1981. "Estimation of Soil Moisture With API Algorithms and Microwave Emission." *JAWRA Journal of the American Water Resources Association* 17, no. 5: 767–774.
- Bousquet, N., and P. Bernardara. 2021. *Extreme Value Theory With Applications to Natural Hazards*. Springer.
- Brett, L., C. J. White, D. I. V. Domeisen, B. van den Hurk, P. Ward, and J. Zscheischler. 2025. "Review Article: The Growth in Compound Weather and Climate Event Research in the Decade Since Srex." *Natural Hazards and Earth System Sciences* 25, no. 8: 2591–2611.
- Brunner, M. I., J. Seibert, and A.-C. Favre. 2016. "Bivariate Return Periods and Their Importance for Flood Peak and Volume Estimation." *WIREs Water* 3, no. 6: 819–833.
- Buritica, G., and P. Naveau. 2023. "Stable Sums to Infer High Return Levels of Multivariate Rainfall Time Series." *Environmetrics* 34, no. 4: e2782.

- Cammalleri, C., C. De Michele, and A. Toreti. 2024. "Exploring the Joint Probability of Precipitation and Soil Moisture Over Europe Using Copulas." *Hydrology and Earth System Sciences* 28, no. 1: 103–115.
- Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher. 2008. "On the Need for bias Correction of Regional Climate Change Projections of Temperature and Precipitation." *Geophysical Research Letters* 35, no. 20: L20709.
- Coles, S., J. Bawa, L. Trenner, and P. Dorazio. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Vol. 208. Springer.
- Coles, S., J. Heffernan, and J. Tawn. 1999. "Dependence Measures for Extreme Value Analyses." *Extremes* 2: 339–365.
- Cooley, D., E. Thibaud, F. Castillo, and M. F. Wehner. 2019. "A Nonparametric Method for Producing Isolines of Bivariate Exceedance Probabilities." *Extremes* 22: 373–390.
- De Moraes, M. A. E., W. M. Mendes Filho, C. A. Bortolozzo, et al. 2024. "Antecedent Precipitation Index to Estimate Soil Moisture and Correlate as a Triggering Process in the Occurrence of Landslides." *International Journal of Geosciences* 15, no. 1: 70–86.
- Drees, H., and A. Janßen. 2017. "Conditional Extreme Value Models: Fallacies and Pitfalls." *Extremes* 20: 777–805.
- Duong, T. 2007. "Ks: Kernel Density Estimation and Kernel Discriminant Analysis for Multivariate Data in r ." *Journal of Statistical Software* 21, no. 7: 1–16.
- Eastoe, E. F., and J. A. Tawn. 2012. "Modelling the Distribution of the Cluster Maxima of Exceedances of Subasymptotic Thresholds." *Biometrika* 99, no. 1: 43–55.
- Fawcett, L., and D. Walshaw. 2007. "Improved Estimation for Temporally Clustered Extremes." *Environmetrics* 18, no. 2: 173–188.
- Ferreira, A., and L. De Haan. 2014. "The Generalized Pareto Process; With a View Towards Application and Simulation." *Bernoulli* 20, no. 4: 1717–1737.
- François, B., M. Vrac, A. J. Cannon, Y. Robin, and D. Allard. 2020. "Multivariate Bias Corrections of Climate Simulations: Which Benefits for Which Losses?" *Earth System Dynamics* 11, no. 2: 537–562.
- Genest, C., K. Ghoudi, and L.-P. Rivest. 1995. "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions." *Biometrika* 82, no. 3: 543–552.
- Gill, J. C., and B. D. Malamud. 2014. "Reviewing and Visualizing the Interactions of Natural Hazards." *Reviews of Geophysics* 52, no. 4: 680–722.
- Haan, L., and A. Ferreira. 2006. *Extreme Value Theory: An Introduction*. Vol. 3. Springer.
- Haddad, Z. S., and D. Rosenfeld. 1997. "Optimality of Empirical z - r Relations." *Quarterly Journal of the Royal Meteorological Society* 123, no. 541: 1283–1293.
- Heffernan, J. E., and J. A. Tawn. 2004. "A Conditional Approach for Multivariate Extreme Values (With Discussion)." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 66, no. 3: 497–546.
- Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The era5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146, no. 730: 1999–2049.
- Holešovský, J., and M. Fusek. 2020. "Estimation of the Extremal Index Using Censored Distributions." *Extremes* 23, no. 2: 197–213.
- Joe, H. 1993. "Parametric Families of Multivariate Distributions With Given Margins." *Journal of Multivariate Analysis* 46, no. 2: 262–282.
- Joe, H. 2005. "Asymptotic Efficiency of the Two-Stage Estimation Method for Copula-Based Models." *Journal of Multivariate Analysis* 94, no. 2: 401–419.
- Kass, R. E., and A. E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90, no. 430: 773–795.
- Kharin, V. V., G. M. Flato, X. Zhang, N. P. Gillett, F. Zwiers, and K. J. Anderson. 2018. "Risks From Climate Extremes Change Differently From 1.5°C to 2.0°C Depending on Rarity." *Earth's Future* 6, no. 5: 704–715.
- Kim, G., M. J. Silvapulle, and P. Silvapulle. 2007. "Comparison of Semiparametric and Parametric Methods for Estimating Copulas." *Computational Statistics & Data Analysis* 51, no. 6: 2836–2850.
- Kirilouk, A., H. Rootzén, J. Segers, and J. L. Wadsworth. 2019. "Peaks Over Thresholds Modeling With Multivariate Generalized Pareto Distributions." *Technometrics* 61, no. 1: 123–135.
- Kohler, M. A., and R. K. Linsley. 1951. *Predicting the Runoff From Storm Rainfall*. Vol. 30. US Department of Commerce.
- Kurowicka, D., and H. Joe. 2010. *Dependence Modeling*. World Scientific.
- Leadbetter, M. 1991. "On a Basis for 'Peaks Over Threshold' Modeling." *Statistics & Probability Letters* 12, no. 4: 357–362.
- Leadbetter, M., I. Weissman, L. De Haan, and H. Rootzén. 1989. "On Clustering of High Values in Statistically Stationary Series." *Meet. Statistical Climatology* 16: 217–222.
- Leadbetter, M. R. 1974. "On Extreme Values in Stationary Sequences." *Zeitschrift für Wahrscheinlichkeitstheorie Und Verwandte Gebiete* 28, no. 4: 289–303.
- Leadbetter, M. R., G. Lindgren, and H. Rootzén. 1983. *Extremes and Related Properties of Random Sequences and Processes*. Springer Science & Business Media.
- Ledford, A. W., and J. A. Tawn. 1997. "Modelling Dependence Within Joint Tail Regions." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 59, no. 2: 475–499.
- Legrand, J., P. Ailliot, P. Naveau, and N. Raillard. 2023. "Joint Stochastic Simulation of Extreme Coastal and Offshore Significant Wave Heights." *Annals of Applied Statistics* 17, no. 4: 3363–3383.
- Li, C., F. Zwiers, X. Zhang, and G. Li. 2019. "How Much Information is Required to Well Constrain Local Estimates of Future Precipitation Extremes?" *Earth's Future* 7, no. 1: 11–24.
- Li, X., Y. Wei, and F. Li. 2021. "Optimality of Antecedent Precipitation Index and Its Application." *Journal of Hydrology* 595: 126027.
- MacDonald, A., C. Scarrott, D. Lee, B. Darlow, M. Reale, and G. Russell. 2011. "A Flexible Extreme Value Mixture Model." *Computational Statistics & Data Analysis* 55, no. 6: 2137–2157.
- Madhar, N., J. Legrand, and M. Thomas. 2024. Assessing Extreme Risk Using Stochastic Simulation of Extremes. <https://arxiv.org/abs/2406.08019>.
- Martins, A. P., and H. Ferreira. 2005. "The Multivariate Extremal Index and the Dependence Structure of a Multivariate Extreme Value Distribution." *TEST* 14: 433–448.
- Matheson, J. E., and R. L. Winkler. 1976. "Scoring Rules for Continuous Probability Distributions." *Management Science* 22, no. 10: 1087–1096.
- Michel, R. 2008. "Some Notes on Multivariate Generalized Pareto Distributions." *Journal of Multivariate Analysis* 99, no. 6: 1288–1301.
- Mohr, S., U. Ehret, M. Kunz, et al. 2023. "A Multi-Disciplinary Analysis of the Exceptional Flood Event of July 2021 in Central Europe – Part 1: Event Description and Analysis." *Natural Hazards and Earth System Sciences* 23, no. 2: 525–551.
- Moloney, N. R., D. Faranda, and Y. Sato. 2019. "An Overview of the Extremal Index." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29, no. 2: 022101.
- Nandagopalan, S. 1994. "On the Multivariate Extremal Index." *Journal of Research of the National Institute of Standards and Technology* 99, no. 4: 543.

- Naveau, P., R. Huser, P. Ribereau, and A. Hannart. 2016. "Modeling Jointly Low, Moderate, and Heavy Rainfall Intensities Without a Threshold Selection." *Water Resources Research* 52, no. 4: 2753–2769.
- Nelsen, R. B. 2006. *An Introduction to Copulas*. Springer.
- Northrop, P. J., and C. Christodoulides. 2023. "Exdex: Estimation of the Extremal Index." *R Package Version 1*, no. 2: 3.
- Páscoa, P., C. M. Gouveia, A. F. S. Ribeiro, and A. Russo. 2024. "Compound Drought and Hot Events Assessment in Australia Using Copula Functions." *Environmental Research Communications* 6, no. 3: 031002.
- Pickands, J. 1975. "Statistical Inference Using Extreme Order Statistics." *Annals of Statistics* 3, no. 1: 119–131.
- Prein, A. F., and G. J. Holland. 2018. "Global Estimates of Damaging Hail Hazard." *Weather and Climate Extremes* 22: 10–23.
- Provost, S. B., and Y. Zang. 2024. "Nonparametric Copula Density Estimation Methodologies." *Mathematics* 12, no. 3: 398.
- Ramos, A., and A. Ledford. 2008. "A New Class of Models for Bivariate Joint Tails." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 71, no. 1: 219–241.
- Resnick, S. I. 2008. *Extreme Values, Regular Variation, and Point Processes*. Vol. 4. Springer Science & Business Media.
- Robert, C. Y. 2008. "Estimating the Multivariate Extremal Index Function." *Bernoulli* 14, no. 4: 1027–1064.
- Rootzén, H., J. Segers, and J. L. Wadsworth. 2018. "Multivariate Generalized Pareto Distributions: Parametrizations, Representations, and Properties." *Journal of Multivariate Analysis* 165: 117–131.
- Rootzén, H., and N. Tajvidi. 2006. "Multivariate Generalized Pareto Distributions." *Bernoulli* 12, no. 5: 917–930.
- Schlef, K. E., K. E. Kunkel, C. Brown, et al. 2023. "Incorporating Non-Stationarity From Climate Change Into Rainfall Frequency and Intensity-Duration-Frequency (IDF) Curves." *Journal of Hydrology* 616: 128757.
- Schröter, K., M. Kunz, F. Elmer, B. Mühr, and B. Merz. 2015. "What Made the June 2013 Flood in Germany an Exceptional Event? A Hydro-Meteorological Evaluation." *Hydrology and Earth System Sciences* 19, no. 1: 309–327.
- Seneviratne, S., M. Donat, A. Pitman, R. Knutti, and R. Wilby. 2016. "Allowable Co2 Emissions Based on Regional and Impact-Related Climate Targets." *Nature* 529, no. 7587: 477–483.
- Sklar, M. 1959. "Fonctions de Répartition à n Dimensions et Leurs Marges." *Publications de l'Institut Statistique de l'Université de Paris* 8, no. 3: 229–231.
- Smith, R. L., and I. Weissman. 1996. *Characterization and Estimation of the Multivariate Extremal Index*. Manuscript.
- Tavakol, A., V. Rahmani, and J. Harrington Jr. 2020. "Probability of Compound Climate Extremes in a Changing Climate: A Copula-Based Study of Hot, Dry, and Windy Events in the Central United States." *Environmental Research Letters* 15, no. 10: 104058.
- Teng, W., J. Wang, and P. Doraiswamy. 1993. "Relationship Between Satellite Microwave Radiometric Data, Antecedent Precipitation Index, and Regional Soil Moisture." *International Journal of Remote Sensing* 14, no. 13: 2483–2500.
- Terzi, S., S. Torresan, S. Schneiderbauer, A. Critto, M. Zebisch, and A. Marcomini. 2019. "Multi-Risk Assessment in Mountain Regions: A Review of Modelling Approaches for Climate Change Adaptation." *Journal of Environmental Management* 232: 759–771.
- Tilloy, A., B. D. Malamud, H. Winter, and A. Joly-Laugel. 2020. "Evaluating the Efficacy of Bivariate Extreme Modelling Approaches for Multi-Hazard Scenarios." *Natural Hazards and Earth System Sciences* 20, no. 8: 2091–2117.
- van Oldenborgh, G. J., S. Philip, E. Aalbers, et al. 2016. "Rapid Attribution of the May/June 2016 Flood-Inducing Precipitation in France and Germany to Climate Change." *Hydrology and Earth System Sciences Discussions* 2016: 1–23.
- Villalobos-Herrera, R., E. Bevacqua, A. F. S. Ribeiro, et al. 2021. "Towards a Compound-Event-Oriented Climate Model Evaluation: A Decomposition of the Underlying Biases in Multivariate Fire and Heat Stress Hazards." *Natural Hazards and Earth System Sciences* 21, no. 6: 1867–1885.
- Wadsworth, J., and J. Tawn. 2022. "Higher-Dimensional Spatial Extremes via Single-Site Conditioning." *Spatial Statistics* 51: 100677.
- Winter, H., and J. Tawn. 2015. "Modelling Heatwaves in Central France: A Case-Study in Extremal Dependence." *Journal of the Royal Statistical Society: Series C: Applied Statistics* 65, no. 3: 345–365.
- Xu, K., C. Wang, and L. Bin. 2023. "Compound Flood Models in Coastal Areas: A Review of Methods and Uncertainty Analysis." *Natural Hazards* 116, no. 1: 469–496.
- Yang, Y., D. Maraun, A. Ossó, and J. Tang. 2023. "Increased Spatial Extent and Likelihood of Compound Long-Duration Dry and Hot Events in China, 1961–2014." *Natural Hazards and Earth System Sciences* 23, no. 2: 693–709.
- Zscheischler, J., O. Martius, S. Westra, et al. 2020. "A Typology of Compound Weather and Climate Events." *Nature Reviews Earth & Environment* 1, no. 7: 333–347.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1.** Supporting Information.

Appendix A

Proof of Theorem 1

Lemma 1. Let us define $\mathbf{Z} = (Z_1, Z_2) := \mathbf{X}^E - \mathbf{u}^E | \mathbf{X}^E \not\leq \mathbf{u}^E$, and let us denote \bar{F}_Z and \bar{F}_E the survival cdf of \mathbf{Z} and \mathbf{X}^E respectively. We have for any scalars $z_1, z_2 \geq 0$:

$$\bar{F}_E(z_1 + u_1^E, z_2 + u_2^E) = \bar{F}_Z(z_1, z_2)(1 - F_E(u_1^E, u_2^E)) \quad (\text{A1})$$

Proof. By application of the conditioning on \mathbf{Z} , one gets

$$\begin{aligned} \bar{F}_Z(z_1, z_2) &:= \mathbb{P}(Z_1 > z_1, Z_2 > z_2) \\ &= \mathbb{P}[X_1^E - u_1^E > z_1, X_2^E - u_2^E > z_2 | \mathbf{X}^E \not\leq \mathbf{u}^E] \\ &= \frac{\mathbb{P}[X_1^E - u_1^E > z_1, X_2^E - u_2^E > z_2]}{\mathbb{P}[\mathbf{X}^E \not\leq \mathbf{u}^E]} \\ &= \frac{\mathbb{P}[X_1^E > z_1 + u_1^E, X_2^E > z_2 + u_2^E]}{\mathbb{P}[(X_1^E > u_1^E) \text{ or } (X_2^E > u_2^E)]} \\ &= \frac{\bar{F}_E(z_1 + u_1^E, z_2 + u_2^E)}{1 - F_E(u_1^E, u_2^E)}. \end{aligned}$$

□

Proof of Theorem 1. For $y_1, y_2 \geq 0$, we have

$$\begin{aligned} \bar{F}(u_1 + y_1, u_2 + y_2) &= \mathbb{P}(X_1 > u_1 + y_1, X_2 > u_2 + y_2) \\ &= \mathbb{P}(X_1^E > -\log(1 - F_1(u_1 + y_1)), \\ &\quad X_2^E > -\log(1 - F_2(u_2 + y_2))) \quad (\text{A2}) \\ &= \bar{F}_E(z_1 + u_1^E, z_2 + u_2^E) \\ &= \mathbb{P}(Z_1 > z_1, Z_2 > z_2)(1 - F_E(u_1^E, u_2^E)), \end{aligned}$$

by application of Lemma 1 with $z_i = -\log(1 - F_i(u_i + y_i)) - u_i^E = -\log\left(\frac{1 - F_i(u_i + y_i)}{1 - p}\right)$. Using the representation in Equation (9), the first factor of the product is

$$\mathbb{P}(Z_1 > z_1, Z_2 > z_2) = \mathbb{P}(E + \Delta_- > z_1, E - \Delta_+ > z_2).$$

By conditioning on E , which is an exponential random variable independent of Δ , one gets

$$\begin{aligned} \mathbb{P}(Z_1 > z_1, Z_2 > z_2) &= \int_0^{+\infty} \mathbb{P}(E + \Delta_- > z_1, E - \Delta_+ > z_2 | E = t) e^{-t} dt \\ &= \int_0^{+\infty} \mathbb{P}(\Delta_- > z_1 - t, \Delta_+ < t - z_2) e^{-t} dt. \end{aligned}$$

The above probability is trivially non equal to zero when either $\{\Delta < 0\} \cap \{\Delta > z_1 - t\} \cap \{0 < t - z_2\}$ or $\{\Delta \geq 0\} \cap \{0 > z_1 - t\} \cap \{\Delta < t - z_2\}$. These two events can be re-written $\{t > \max(z_1, z_2)\} \cap \{z_1 - t < \Delta < t - z_2\}$. Hence, the probability now becomes

$$\begin{aligned} \mathbb{P}(Z_1 > z_1, Z_2 > z_2) &= \int_{\max(z_1, z_2)}^{+\infty} \mathbb{P}(z_1 - t < \Delta < t - z_2) e^{-t} dt \\ &= \int_{\max(z_1, z_2)}^{+\infty} [F_\Delta(t - z_2) - F_\Delta(z_1 - t)] e^{-t} dt \\ &= \int_{\max(z_1, z_2)}^{+\infty} F_\Delta(t - z_2) e^{-t} dt - \int_{\max(z_1, z_2)}^{+\infty} F_\Delta(z_1 - t) e^{-t} dt \\ &= e^{-z_2} \int_{(z_1 - z_2)_+}^{+\infty} e^{-t} F_\Delta(t) dt - e^{-z_1} \int_{-\infty}^{(z_1 - z_2)_-} e^t F_\Delta(t) dt \\ &= \frac{(1 - F_2(u_2 + y_2)) \int_{(z_1 - z_2)_+}^{+\infty} e^{-t} F_\Delta(t) dt - (1 - F_1(u_1 + y_1)) \int_{-\infty}^{(z_1 - z_2)_-} e^t F_\Delta(t) dt}{1 - p}. \end{aligned}$$

The second equality holds because of the continuity of F_Δ . In the special case $y_1 = y_2 = 0$, one gets:

$$\begin{aligned} \bar{F}(u_1, u_2) &= \mathbb{P}(Z_1 > 0, Z_2 > 0) (1 - F_E(u_1^E, u_2^E)) \\ &= \left[\int_0^{+\infty} e^{-t} F_\Delta(t) dt - \int_{-\infty}^0 e^t F_\Delta(t) dt \right] (1 - F_E(u_1^E, u_2^E)) \\ &= (1 - F_E(u_1^E, u_2^E)) \int_0^{+\infty} e^{-t} (F_\Delta(t) - F_\Delta(-t)) dt \end{aligned}$$

from which we find the second factor of the product in Equation (A2):

$$1 - F_E(u_1^E, u_2^E) = \frac{\bar{F}(u_1, u_2)}{\int_0^{+\infty} e^{-t} (F_\Delta(t) - F_\Delta(-t)) dt}.$$

Plugging these in Equation (A2), we find

$$\bar{F}(u_1 + y_1, u_2 + y_2) = \frac{\bar{F}(u_1, u_2) \left[(1 - F_2(u_2 + y_2)) \int_{(z_1 - z_2)_+}^{+\infty} e^{-t} F_\Delta(t) dt - (1 - F_1(u_1 + y_1)) \int_{-\infty}^{(z_1 - z_2)_-} e^t F_\Delta(t) dt \right]}{(1 - p) \int_0^{+\infty} e^{-t} (F_\Delta(t) - F_\Delta(-t)) dt}.$$

Then, denoting $\delta = z_1 - z_2$, the final result is obtained. \square

Appendix B

Proof of Proposition 1

Let $\mathbf{u} = (u_1, u_2)$ be two thresholds corresponding to high quantiles of X_1 and X_2 respectively, with $p = F_1(u_1) = F_2(u_2)$, such that $[\mathbf{X} - \mathbf{u} | \mathbf{X} \not\leq \mathbf{u}]$

follows a bivariate GPD. \mathbf{X} is also supposed not to have a perfect negative dependence. Then, for $v \geq p$:

$$\begin{aligned} \chi(v) &= \mathbb{P}(V_2 > v | V_1 > v) = \frac{\mathbb{P}(V_2 > v, V_1 > v)}{\mathbb{P}(V_1 > v)} \\ &= \frac{\mathbb{P}(F_2(X_2) > F_1(x), F_1(X_1) > F_1(x))}{1 - v} \\ &= \frac{\mathbb{P}(X_2 > F_2^{-1}(F_1(x)), X_1 > x)}{1 - v}. \end{aligned}$$

Using Theorem 1 with $u_1 + y_1 = x$ and $u_2 + y_2 = F_2^{-1}(F_1(x))$, we have:

$$\begin{aligned} \chi(v) &= \frac{\bar{F}(u_1, u_2) \left[(1 - F_2(F_2^{-1}(F_1(x)))) \int_0^{+\infty} e^{-t} F_\Delta(t) dt - (1 - F_1(x)) \int_{-\infty}^0 e^t F_\Delta(t) dt \right]}{(1 - v)(1 - p) \int_0^{+\infty} e^{-t} (F_\Delta(t) - F_\Delta(-t)) dt} \\ &= \frac{\bar{F}(u_1, u_2) \left[(1 - v) \int_0^{+\infty} e^{-t} F_\Delta(t) dt - (1 - v) \int_{-\infty}^0 e^t F_\Delta(t) dt \right]}{(1 - v)(1 - p) \int_0^{+\infty} e^{-t} (F_\Delta(t) - F_\Delta(-t)) dt} \\ &= \frac{\bar{F}(u_1, u_2)}{1 - p}. \end{aligned}$$

According to Sklar's theorem, there exists a copula C such that: $\bar{F}(u_1, u_2) = 1 - 2p + C(p, p)$. The Fréchet-Hoeffding theorem for copula bounds states that $1 - 2p + \max(2p - 1, 0) \leq \bar{F}(u_1, u_2) \leq 1 - 2p + \min(p, p)$. When p is a high probability ($p > 0.5$), the Fréchet-Hoeffding bounds thus become $0 \leq \bar{F}(u_1, u_2) \leq 1 - p$ and hence $0 \leq \chi \leq 1$. The lower bound is reached when $\bar{F}(u_1, u_2) = 0$, which is only possible for perfectly negatively dependent variables. As a consequence, χ is always positive when the variables are not perfectly negatively dependent. Conversely, the upper bound $\chi = 1$ is reached when $\bar{F}(u_1, u_2) = 1 - p$, that is, when the variables are perfectly positively dependent. Notice that when X_1 and X_2 are independent, then $C(p, p) = p^2$ and

$$\chi = \frac{1 - 2p + p^2}{1 - p} = 1 - p > 0.$$

If \mathbf{X} are not perfectly negatively dependent, that is, if $\chi > 0$, we have

$$\begin{aligned} \bar{\chi}(v) &= \frac{2 \log(\mathbb{P}(V_1 > v))}{\log(\mathbb{P}(V_1 > v, V_2 > v))} - 1 = \frac{2 \log(1 - v)}{\log(\chi \mathbb{P}(V_1 > v))} - 1 \\ &= \frac{\log(1 - v) - \log(\chi(v))}{\log(1 - v) + \log(\chi(v))}. \end{aligned}$$

Hence, $\bar{\chi} = \lim_{v \rightarrow 1} \bar{\chi}(v) = 1$.

Appendix C

Proof of Proposition 2

In the following, we will often work with disjoint events, that is, events with an empty intersection. In order to distinguish the union of such events to the general case, we will denote \sqcup the union of two disjoint events, for which we have $\mathbb{P}(A \sqcup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Recall that there are p blocks of size h denoted B_1, \dots, B_p , and that $M_1(B_i)$ and $M_2(B_i)$ are the maximum of the two variables in block B_i , with

$\mathbf{M}(B_i) = (M_1(B_i), M_2(B_i))$. For $j = 1, 2$, the $(M_j(B_i))_{i=1, \dots, p}$ are identically distributed to a random variable M_j , with cdf H_j , H being the joint cdf of $\mathbf{M} = (M_1, M_2)$. Let us consider an integer $0 < k \leq p + 1$ and two scalars $x_1, x_2 > 0$ and define the event

$$\begin{aligned} A_i &= \{ \mathbf{M}(B_i) \leq (x_1, x_2) \} \sqcup \{ (M_1(B_i) \leq x_1) \cap (M_2(B_i) > x_2) \} \\ &\quad \sqcup \{ (M_1(B_i) > x_1) \cap (M_2(B_i) \leq x_2) \}. \end{aligned}$$

Then,

$$\begin{aligned} \mathbb{P}[\eta(x_1, x_2) \leq k] &= \mathbb{P}[\exists k_0 \in \llbracket 1, k \rrbracket \text{ such that } \mathbf{M}(B_{k_0}) > (x_1, x_2)] \\ &= 1 - \mathbb{P}[\forall i \in \llbracket 1, k \rrbracket, A_i \text{ is true}], \end{aligned}$$

Since independence between blocks is assumed and since the maxima $\mathbf{M}(B_i)$ are identically distributed, we get:

$$\begin{aligned} \mathbb{P}[\eta(x_1, x_2) \leq k] &= 1 - [\mathbb{P}(\mathbf{M} \leq (x_1, x_2)) \\ &\quad + \mathbb{P}(M_1 \leq x_1 \cap M_2 > x_2) + \mathbb{P}(M_2 \leq x_2 \cap M_1 > x_1)]^k. \end{aligned}$$

By the decomposition $\mathbb{P}(M_1 \leq x_1) = \mathbb{P}[(M_1 \leq x_1) \cap (M_2 > x_2) \cup (M_2 \leq x_2)]$, it is straightforward to establish that $\mathbb{P}(M_1 \leq x_1 \cap M_2 > x_2) = \mathbb{P}(M_1 \leq x_1) - \mathbb{P}[M_1 \leq x_1 \cap M_2 \leq x_2]$, and similarly for $\mathbb{P}(M_2 \leq x_2 \cap M_1 > x_1)$. Hence,

$$\begin{aligned} \mathbb{P}[\eta(x_1, x_2) \leq k] &= 1 - [\mathbb{P}(M_1 \leq x_1) + \mathbb{P}(M_2 \leq x_2) - \mathbb{P}(\mathbf{M} \leq (x_1, x_2))]^k \\ &= 1 - [H_1(x_1) + H_2(x_2) - H(x_1, x_2)]^k \\ &= 1 - [1 - \overline{H}(x_1, x_2)]^k. \end{aligned}$$

Then,

$$\begin{aligned} \mathbb{P}[\eta(x_1, x_2) = k] &= \mathbb{P}[\eta(x_1, x_2) \leq k] - \mathbb{P}[\eta(x_1, x_2) \leq k - 1] \\ &= \overline{H}(x_1, x_2) [1 - \overline{H}(x_1, x_2)]^{k-1}. \end{aligned}$$

For $k \leq p$, this is the expression of a geometric law with parameter $(1 - \overline{H}(x_1, x_2))$. Hence, as $p \rightarrow \infty$, the expectation is

$$\mathbb{E}(\eta(x_1, x_2)) = \frac{1}{\overline{H}(x_1, x_2)}.$$

Appendix D

Proof of Proposition 3

To simplify, we will consider the case where the variables are temporally independent ($\theta = 1$). The results also hold for temporally dependent variables, but lead to complex expressions that are not shown in this work. We consider $h \in \mathbb{N}^*$ and scalars $x_1, x_2 > 0$ such that $F_1(x_1) = F_2(x_2) = v$ and $w = 1 - v$ close to 0, so that w^2 is negligible compared to w .

- We first consider the case of asymptotic dependence, that is, $\overline{\chi} = 1$. From the definition $\chi(v) = \mathbb{P}(V_2 > v | V_1 > v) = \mathbb{P}(X_2 > x_2 | X_1 > x_1)$, we get $\overline{F}(x_1, x_2) = \chi(v)\overline{F}_1(x_1) = \chi(v)w$. For the return period $T(h) = h/(N\overline{H}(x_1, x_2))$, we need to compute

$$\begin{aligned} \overline{H}(x_1, x_2) &= 1 - H_1(x_1) - H_2(x_2) + H(x_1, x_2) \\ &= 1 - F_1^h(x_1) - F_2^h(x_2) + F^h(x_1, x_2) \\ &= 1 - 2v^h + (1 - \overline{F}_1(x_1) - \overline{F}_2(x_2) + \overline{F}(x_1, x_2))^h \\ &= 1 - 2(1 - w)^h + (1 - 2w + \chi(v)w)^h. \end{aligned}$$

Since w is close to 0, we can approximate the above expression by its second order polynomial in h :

$$\overline{H}(x_1, x_2) = h\chi(v)w + \frac{h(h-1)}{2} \times ((\chi(v) - 2)^2 - 2)w^2 + o(w^2).$$

The second term is proportional to w^2 . It is thus negligible with respect to the first leading term which is proportional to w . Hence we get $T(h) \simeq 1/(N\chi(v)w)$. At the first order, $T(h)$ does not depend on h .

- We now consider the case of asymptotic independence, that is, $\chi = 0$ and $-1 < \overline{\chi} < 1$. Definition (3) rewrites $\overline{F}(x_1, x_2)\overline{\chi}^{(v)+1} = \overline{F}_1(x_1)^2$. Therefore, $\overline{H}(x_1, x_2)$ becomes:

$$\overline{H}(x_1, x_2) = 1 - 2(1 - w)^h + (1 - 2w + w\frac{2}{\overline{\chi}^{(v)+1}})^h$$

Let us denote $\alpha(v) = \frac{2}{\overline{\chi}^{(v)+1}}$. The second order polynomial leads to

$$\begin{aligned} \overline{H}(x_1, x_2) &\simeq h \left(w^{\alpha(v)} - \frac{1}{2} (w^{2\alpha(v)} - 4w^{\alpha(v)+1} + 2w^2) \right) \\ &\quad + \frac{h^2}{2} (w^{2\alpha(v)} - 4w^{\alpha(v)+1} + 2w^2). \end{aligned}$$

In case of positive association, that is, when $0 \leq \overline{\chi} < 1$, we observe that $1 < \alpha(v) \leq 2$, with $\alpha(v) = 2$ when $\overline{\chi}(v) = 0$ and $\alpha(v) = 1$ when $\overline{\chi}(v) = 1$. It follows immediately that for positive association, $2 < \alpha(v) + 1 \leq 3$ and $2 < 2\alpha(v) \leq 4$. Hence, the term h^2w^2 is not always negligible as compared to $hw^{\alpha(v)}$, in particular for the higher values of $\overline{\chi}(v)$. Then,

$$\overline{H}(x_1, x_2) = hw^{\alpha(v)} + h(h-1)w^2 + o(w^2)$$

and

$$T(h) \simeq \frac{1}{N(w^{\alpha(v)} + (h-1)w^2)}.$$

In case of complete independence, $\overline{\chi}(v) = 0$, and $\alpha(v) = 2$. Hence, $T(h) \simeq 1/Nhw^2$.

In case of negative association, that is, $-1 < \overline{\chi} < 0$, we have $\alpha(v) > 2$. Hence $w^{\alpha(v)}$ is negligible as compared to w^2 , and the probability becomes: $\overline{H}(x_1, x_2) = h(h-1)w^2 + o(w^2)$. The return period is thus:

$$T(h) \simeq \frac{1}{N(h-1)w^2}.$$

Appendix E

Proof of Proposition 4

From the definition of the extremal index, we have $H_i(x) \simeq F_i^{h\theta_i}(x)$ for $i = 1, 2$ and $x > 0$. For bivariate maxima, there is a similar expression which follows from the definition of the bivariate extremal index in Equation (13), with $H(x_1, x_2) = F^{h\theta(x_1, x_2)}(x_1, x_2)$, for $x_1, x_2 > 0$.

Univariate Return Periods

For each variable $i = 1, 2$, the upper tail of the cdf of the block maxima is approximately a GPD, that is, for $x_i > u_i$, u_i being a high threshold, $G_i(x_i) \simeq H_i^{u_i}(x_i) = \mathbb{P}[M_i \leq x_i | M_i > u_i]$, see Section 4.1. It follows that $\overline{H}_i^{u_i}(x_i) = \mathbb{P}[M_i > x_i | M_i > u_i] = \mathbb{P}[M_i > x_i] / \mathbb{P}[M_i > u_i]$. Hence,

$$\overline{H}_i(x_i) = \overline{H}_i(u_i)\overline{H}_i^{u_i}(x_i). \tag{E1}$$

With the univariate return period formula from Section 5.1, we thus get

$$T_i \simeq \frac{1}{N\theta_i F_i(x_i)} \simeq \frac{h}{N\overline{H}_i(u_i)\overline{H}_i^{u_i}(x_i)}. \tag{E2}$$

Bivariate Return Periods

The combination of Definition 3 and Proposition 2 gives the following expression for the bivariate return period for any $(x_1, x_2) \geq (u_1, u_2)$:

$$T = \frac{h}{N\overline{H}(x_1, x_2)}.$$

The conditioning in the definition of \overline{H} provides:

$$\overline{H}(x_1, x_2) = \mathbb{P}[(M_1, M_2) > (x_1, x_2) | (M_1, M_2) > (u_1, u_2)]\overline{H}(u_1, u_2) \tag{E3}$$

We need to have an expression for both factors.

Regarding the first factor in Equation (E3), according to Sklar's theorem, there exists a copula C such that:

$$\begin{aligned} \mathbb{P}[(M_1, M_2) \leq (x_1, x_2) | (M_1, M_2) > (u_1, u_2)] \\ = H^{u_1, u_2}(x_1, x_2) = C(H_1^{u_1}(x_1), H_2^{u_2}(x_2)). \end{aligned}$$

On the other hand,

$$\begin{aligned} H^{u_1, u_2}(x_1, x_2) &= \mathbb{P}[(M_1, M_2) \leq (x_1, x_2) | (M_1, M_2) > (u_1, u_2)] \\ &= \mathbb{P}[M_1 \leq x_1 | M_1 > u_1, M_2 > u_2] \\ &\quad + \mathbb{P}[M_2 \leq x_2 | M_1 > u_1, M_2 > u_2] - 1 \\ &\quad + \mathbb{P}[(M_1, M_2) > (x_1, x_2) | (M_1, M_2) > (u_1, u_2)] \end{aligned} \tag{E4}$$

The first two terms of (E4) can be expressed with the copula C :

$$\begin{aligned} \mathbb{P}[M_1 \leq x_1 | M_1 > u_1, M_2 > u_2] \\ = \mathbb{P}[M_1 \leq x_1, M_2 < +\infty | M_1 > u_1, M_2 > u_2] \\ = H^{u_1, u_2}(x_1, +\infty) = C(H_1^{u_1}(x_1), 1), \end{aligned}$$

and similarly for the second component: $\mathbb{P}[M_2 \leq x_2 | M_1 > u_1, M_2 > u_2] = C(1, H_2^{u_2}(x_2))$. Plugging these expressions in Equation (E4) and after some rearrangements, one thus gets:

$$\begin{aligned} \mathbb{P}[(M_1, M_2) > (x_1, x_2) | (M_1, M_2) > (u_1, u_2)] \\ = 1 + C(H_1^{u_1}(x_1), H_2^{u_2}(x_2)) - C(H_1^{u_1}(x_1), 1) - C(1, H_2^{u_2}(x_2)). \end{aligned}$$

Regarding the second factor in Equation (E3), $\overline{H}(u_1, u_2)$ can be expressed with the bivariate extremal index $\theta(u_1, u_2)$ with

$$\begin{aligned} \overline{H}(u_1, u_2) &= 1 - H_1(u_1) - H_2(u_2) + H(u_1, u_2) \\ &\simeq 1 - F_1^{h\theta_1}(u_1) - F_2^{h\theta_2}(u_2) + F^{h\theta(u_1, u_2)}(u_1, u_2). \end{aligned} \tag{E5}$$

The bivariate return period T with a GPD-copula approach thus becomes:

$$T = \frac{h}{N \overline{H}(u_1, u_2) [1 + C(H_1^{u_1}(x_1), H_2^{u_2}(x_2)) - C(H_1^{u_1}(x_1), 1) - C(1, H_2^{u_2}(x_2))]}$$

where $\overline{H}(u_1, u_2)$ is given in Equation (E5).

Appendix F

Proof of Proposition 5

Direct application of the general expression of $\overline{H}(x_1, x_2)$ for $(x_1, x_2) > (u_1, u_2)$ yields:

$$\begin{aligned} \overline{H}(x_1, x_2) &= 1 - H_1(x_1) - H_2(x_2) + H(x_1, x_2) \\ &\simeq 1 - F_1(x_1)^{h\theta_1} - F_2(x_2)^{h\theta_2} + F(x_1, x_2)^{h\theta(x_1, x_2)} \\ &\simeq 1 - F_1(x_1)^{h\theta_1} - F_2(x_2)^{h\theta_2} \\ &\quad + \left(F_1(x_1) + F_2(x_2) - 1 + \overline{F}(x_1, x_2) \right)^{h\theta(x_1, x_2)} \end{aligned}$$

The result follows immediately.