Supplement of

Evaluating the efficacy of bivariate extreme modelling approaches for multi-hazard scenarios

Aloïs Tilloy et al.

Correspondence to: Aloïs Tilloy (alois.tilloy@kcl.ac.uk)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.
TABLE OF CONTENT:

S1 THEORETICAL BACKGROUND ON MULTIVARIATE EXTREME VALUES

S1.1 UNIVARIATE EXTREME VALUE THEORY AND REGULAR VARIATION
S1.1.1 MAXIMUM DOMAIN OF ATTRACTION AND GEV  2
S1.1.2 MAX-STABILITY  3
S1.1.3 REGULAR VARIATION  3

S1.2 MULTIVARIATE EXTREME VALUE STATISTICS
S1.2.1 MAXIMUM DOMAIN OF ATTRACTION AND MAX STABILITY  4
S1.2.2 THE EXPONENT MEASURE  4
S1.2.3 MULTIVARIATE AND HIDDEN REGULAR VARIATION  5

S1.3 BIVARIATE CASE
S1.3.1 GUMBEL COPULA  7
S1.3.2 NORMAL COPULA  7

S2 LEVEL CURVE DENSITY  8

REFERENCES  10
S1 Theoretical background on multivariate extreme values. Associated Section: Section 2.1: Bivariate extreme dependence and Section 2.2: Bivariate models

S1.1 Univariate extreme value theory and regular variation

Extreme value analysis is a statistical approach for analysing extreme data values for a variable of interest. One of the earliest recorded mentions is by Fisher and Tippett (1928). Extreme value analysis was formalized into a statistical method by Gumbel (1958). It has been used extensively in the environmental sciences to overcome the limitations of empirical approaches (e.g., Tiago de Oliveira, 1986; Bingham, 2007). Here we present three main concepts linked to univariate extreme value theory that can be extended to the bivariate case.

S1.1.1 Maximum domain of attraction and GEV

The first principle from which arises extreme value distributions is the maximum domain of attraction: let the random variables $x_1, \ldots, x_n$ be i.i.d. values, with distribution function $F$. Define $M_n = \max(x_1, \ldots, x_n)$ and suppose there exist sequences of normalizing constants $a_n > 0, b_n$ such that (as $n \to \infty$) (Davison and Huser, 2015):

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) = F^n(a_n z + b_n) \xrightarrow{d} G(z)$$  \hspace{1cm} (S1)

where $\xrightarrow{d}$ denotes convergence in the distribution and $G$ is a non-degenerate distribution function. Then $G$ is an extreme value distribution and it is said that $F$ belongs to the maximum domain of attraction of $G$. The constants $a_n$ and $b_n$ are called stabilizing constants. The possible $G$ distributions are then summarized by the Generalized Extreme Value (GEV) distribution (Gümbel, 1958; Coles, 2001; Davison and Huser, 2015):

$$G(x) = P(X \leq x) = \exp\left(-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)$$  \hspace{1cm} (S2)

for $1 + \xi \frac{x - \mu}{\sigma} > 0$, with

- $\mu \in (-\infty, \infty)$ the location parameter
- $\sigma \in [0, \infty)$ the scale parameter
- $\xi \in (-\infty, \infty)$ the shape parameter

The shape parameter $\xi$ controls the heaviness of the tail. It means that the value of this parameter directly affects the estimation of the extremes. The Extreme Type Theorem gives three different families of limiting distributions depending on the sign of the shape parameter (Coles, 2001):

- $\xi = 0$, a Gumbel distribution with and exponential upper tail;
- $\xi > 0$, a Fréchet distribution with a heavy upper tail;
- $\xi < 0$, a reverse Weibull distribution with a light upper tail.

A threshold above which one value is considered as extreme can be set instead of selecting an extreme. In that case, the distribution $G$ of the exceedances above a high threshold $u$ is a Generalized Pareto Distribution (GPD) (Davison and Smith, 1990) of the form:
\[ G(x) = P(X \leq x | X > u) = 1 - \left(1 + \frac{x - u}{\sigma_u} \right)^{-\frac{1}{\xi}} \]  

(S3)

for \( x > u \), with

- \( \sigma_u \in [0, \infty) \) the scale parameter
- \( \xi \in (-\infty, \infty) \) the shape parameter

The shape parameter \( \xi \) of the GPD is equivalent to the shape parameter of the corresponding GEV distribution. This shape parameter changes with the threshold level, which makes the choice of the threshold important (Bernardara et al., 2014). The scale parameter for the GPD is also threshold-dependent.

**S1.1.2 Max-stability**

In the early years of extreme value statistics, Fréchet (1927) identified a functional equation, which he called the stability postulate that provides a mathematical basis for extrapolation and thus lies at the heart of the classical theory of extremes (Davison and Huser, 2015). His stability postulate is now referred to as max-stability (see Eq. A4). Max-stability is a property that is only satisfied by the three families of GEV: the Gumbel, Fréchet and Reverse Weibull families (Coles, 2001). A distribution \( G \) is said to be max-stable if, for every \( n > 0 \), there exist constants \( a_n > 0 \) and \( b_n \) such that:

\[ G^n(a_nz + b_n) = G(z) \]  

(S4)

where \( G^n(z) \) is the distribution function of \( M_n = \max(x_1, \ldots, x_n) \), with the \( x_i \) independent variables for a distribution \( G \). This means that max-stability is satisfied by distributions for which the fact of taking sample maxima leads to the same distribution apart from changes of parameters (Coles, 2001). The maximum domain of attraction and the max-stability property allows one to model any sample maxima distribution with a GEV distribution.

**S1.1.3 Regular variation**

Another important concept linked to extreme value analysis is the theory of regularly varying functions. The link between this concept and extreme values has been mainly discussed by Resnick (1987). A regularly varying function is a function which behave asymptotically like a power function. A function \( F \) is regularly varying at \( \infty \) with index \( \rho \), if for \( x > 0 \) (Resnick, 1987):

\[ \lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^\rho \]  

(S5)

If \( \rho = 0 \), we call \( U \) a slowly varying function. Slowly varying functions are usually denoted by \( L(x) \) The theory of regularly varying functions has links to many mathematical disciplines (Bingham et al., 1987). Moreover, it has been used to understand and investigate maximum domains of attraction in extreme value theory (Bingham et al., 1987; Resnick, 1987; De Haan and Resnick, 1996; Bingham, 2007).
S1.2 Multivariate extreme value statistics

Multivariate extreme value theory is an extension of univariate extreme value theory (Tiago de Oliveira, 1986; Resnick, 1987; Coles, 2001) and various properties of extreme value distributions are analogous in the multivariate framework. Here, the statistics of extremes in a multivariate context are formally presented building on the concepts introduced above (Section A1).

S1.2.1 Maximum domain of attraction and max stability

The maximum domain of attraction can be extended in the multivariate framework. Let the random variables \((X_{j,1}, \ldots, X_{j,d})\), where \(j=1, \ldots, n\), be a collection \(d\)-dimensional vectors of i.i.d. values with a joint distribution \(F\). Define \(M_n = \max(X_{1,k}, \ldots, X_{n,k})\) for \(k=1, \ldots, d\) and suppose there exist sequences of normalizing constants \(a_{n,k} > 0, b_{n,k}\) for \(k=1, \ldots, d\) such that as \(n \to \infty\) (Dutfoy et al., 2014):

\[
P\left(\frac{M_{n,1} - b_{n,1}}{a_{n,1}} \leq z_1, \ldots, \frac{M_{n,d} - b_{n,d}}{a_{n,d}} \leq z_n\right) = F_{n,d}\left(\frac{a_{n,d} z_d + b_{n,d}}{n}\right) \to G(z_1, \ldots, z_d)
\]  

(S6)

where \(\to\) denotes convergence in the distribution and \(G\) is a distribution function with all non-degenerate marginals. Then the limiting distribution \(G\) is a Multivariate Extreme value distribution of dimension \(d\), and \(F\) is said to be in the maximum domain of attraction of \(G\). Each marginal

\[
Z_k = \lim_{n \to \infty} \left(\frac{M_{n,k} - b_{n,k}}{a_{n,k}}\right), \quad k = 1, \ldots, d
\]

(S7)

follows a GEV distribution (Section A1.1) with parameters \((\mu_k, \sigma_k, \xi_k)\). In can also be shown that \(G\) must satisfy the max stability relation (Resnick, 1987; Tawn, 1988, 1990; Coles, 2001).

In practice, two steps are generally required to conduce a multivariate study:

(i) marginal distributions are usually estimated using the univariate extreme value methodology (Section A1);

(ii) the marginal distributions are then transformed to a common distribution, in order to handle the dependence structure using multivariate extreme value theory.

For reason of mathematical elegance and simplicity, but without loss of generality, marginal distributions are usually transformed to standard Fréchet distributions in multivariate extreme value analysis where \(a_{n,k} = k^{-1}\) and \(b_{n,k} = 0\) in (A7). This allows one to focus on the dependence structure between variables (Winter, 2016). From now on, we consider random variables \(Z = (Z_1, \ldots, Z_n)\) with common standard Fréchet margins.

S1.2.2 The exponent measure

The characterization of the dependence structure in the extremes is too complex to be summarized by a parametric family (Davison and Huser, 2015). However, the limiting distribution of \(Z\) with common Fréchet margins is a multivariate extreme value distribution \(G\) with \(z \in \mathbb{R}^d\) and can be written as (Huser, 2013):

\[
G(z) = \exp\{-V(z)\}, \quad z > 0,
\]

(S8)
where \( V(z) \) is a Radon measure called the exponent measure, which contains all the information about dependence among the variables \( Z = (Z_1, \ldots, Z_n) \). The exponent measure can be interpreted as the approximate probability that at least one of the maxima \( Z_{n,k} \) exceeds its threshold (Davison and Huser, 2015):

\[
V(z) = D \int_{SD} \max \left( \frac{w}{z} \right) dH(w)
\]

(S9)

with \( H \) a measure on the \((D-1)\)-dimensional simplex \( SD = w \in \mathbb{R}^d \). The measure \( dH \) is often called the spectral measure.

From the max-stability property with Fréchet margins, the exponent measure is regularly varying and homogeneous of order \(-1\), meaning that:

\[
V(tz) = t^{-1}V(z)
\]

(S10)

Properties of the exponent measure, including its regular variation play a central role when it comes to extrapolation in the upper tail of multivariate variables (Davidson and Huser, 2015). If a bivariate distribution is asymptotically independent, then the exponent measure \( V(t) = 0 \). The theory of regular variation also provides a framework for extrapolation in the upper tail and has been related to multivariate extreme value theory (Resnick, 1987, Cooley et al. 2019).

**S1.2.3 Multivariate and hidden regular variation**

Results presented in Eqs. A7, A8 and A10 can be related to the concept of multivariate regular variation developed and presented by Resnick (Resnick, 1987, 2002). Multivariate variation on the cone \( C = [0, \infty]^d - \{0\} \) can be defined as the following: suppose that \( Z \) is a \( d \)-dimensional random vector in \([0, \infty]^d\), then the distribution of \( Z \) is regularly varying (with unequal components) if there exist functions \( b(t) \rightarrow \infty \), as \( t \rightarrow \infty \) that, for a Radon measure \( \nu \) (i.e., finite on sets bounded away from zero) on \( C \), we have the vague convergence which can be expressed as (Cooley et al., 2019):

\[
\lim_{t \rightarrow \infty} \left[ tP \left( \frac{Z}{b(t)} \in A \right) \right] \rightarrow \nu (A)
\]

(S11)

for any set \( A \subset C \) and where \( b(t) \) is a regularly varying function of some index \( \alpha > 0 \) and \( \nu \) is a Radon measure on the cone \( C = [0, \infty]^d - \{0\} \) which satisfies the homogeneous property

\[
\nu(tA) = t^{-\alpha} \nu(A)
\]

(S12)

for any scaler \( t \) and \( A \subset C \). The limit measure \( \nu(A) \) has a homogeneity property of order \(-\alpha\). The coefficient \( \alpha \) is the index of regular variation and \( \alpha = 1/\xi \) with \( \xi \) the shape parameter of the marginal distributions (see **Section A1**). With a standard Fréchet margins we have \( \xi = 1 \) and therefore \( \alpha = 1 \).

Multivariate extreme value and regular variation theory previously presented provide a rich theory for extremal dependence in the case of asymptotic dependence (Pickands, 1981; Das, 2009) but it is not able to distinguish between asymptotic
independent and actual independence. Ledford and Tawn (1996; 1997) developed a dependence measure that can detect tail dependence in the asymptotic independence setting. The coefficient of tail dependence $\eta$ measures the speed of decay toward independence at high level (Davison and Huser, 2015). The coefficient $\eta$ provides a better understanding of asymptotically independent behaviours and helped develop the concept of hidden regular variation.

Hidden regular variation is a property of the subfamily of distributions having both multivariate regular variation and asymptotic independence. (Resnick, 2002; Maulik and Resnick, 2005) Resnick (2002) Asymptotic independence is a degenerative case for multivariate extreme value theory (Cooley, 2019) The renormalizing sequence $b(t)$ in Eq. A11 grows too rapidly. The latter is replaced by a lighter tailed normalizing sequence $b^0$. Hidden regular variation can therefore be expressed on the cone $C = (0, \infty)^d$ as:

$$\lim_{t \to \infty} \left[ tP \left( \frac{Z}{b^0(t)} \in A \right) \right] \to v_0(A)$$

for any set $A$ bounded away from the axes, $A \subset C$, where $b^0$ is a regularly varying function and $v$ is a Radon measure (i.e., finite on sets bounded away from zero) on the cone $C (0, \infty)^d$ which satisfies

$$v_0(tA) = t^{-1/\eta} v_0(A)$$

for any scaler $t$ and $A \subset C$. Here, $\eta$ is the coefficient of tail dependence $\eta \in (0,1]$. A decreasing value of $\eta$ correspond to weaker dependence.

S1.3 Bivariate case

In the bivariate case, when $d = 2$, the exponent measure (Section A2.2) is expressed as:

$$V(z_1, z_2) = \int_0^1 \max \left( \frac{w}{z_1}, \frac{1-w}{z_2} \right) 2dH(w)$$

with $H$ an arbitrary distribution function on $[0,1]$ satisfying the moment constraint

$$\int_0^1 wdH(w) = 1/2$$

An alternative representation of equation incorporates the Pickands dependence function (Pickands, 1975), denoted by $A(w)$

$$V(z_1, z_2) = (z_1^{-1} + z_2^{-1}) A \left( \frac{z_1}{z_1 + z_2} \right)$$

where $A(w)$ satisfies

$$A(w) = 2 \int_0^1 \max\left( (1-w)q, w(1-q) \right) dH(q)$$

The Pickands dependence function $A(w)$ is a defined on the interval $[0,1]$ and has the following properties: (i) $A(0) = A(1) = 1$, (ii) $A(w)$ is convex and (iii) $A(w)$ is contained in a triangular region $A(w)$ is usually used as a measure of the strength of
dependence between two variables \( z_1 \) and \( z_2 \). The Pickands dependence function can be estimated parametrically through copula functions or with nonparametric estimators (Pickands, 1981; Capéraà et al., 1997).

S1.3.1 Gumbel copula

The Gumbel copula (which is also an Archimedean copula) is one of the oldest extreme value copulas (Eschenburg, 2013). It is also referred to as the bivariate logistic model (with Gumbel margins) in the literature and was first introduced by Gumbel (1961):

\[
C(u, v) = \exp\left\{-\left[\frac{-\ln(u)}{\theta} + \frac{-\ln(v)}{\theta}\right]^{1/\theta}\right\}
\]

(S19)

with \( \theta \in [1, \infty] \) the dependence parameter, \( u \) and \( v \) uniform marginal distributions, and \( \ln \) the natural log and \( \exp \) the exponential.

The Gumbel copula has been used widely in hydrology (Zhang and Singh, 2007; Salvadori and De Michele, 2010; Zheng et al., 2013; Dung et al., 2015) and coastal engineering (Yang and Zhang, 2013; Masina et al., 2015; Mazas and Hamm, 2017). We will also use this copula in our simulation study as a reference for the asymptotic dependence case. Other important extreme value copulas include the Galambos copula which will be used alongside the Gumbel copula as asymptotically dependent models in our simulations (Sect. 3).

S1.3.2 Normal copula

The normal copula has been used in several hazard interrelation studies because of its flexibility (Rueda et al., 2016; Serinaldi, 2016; Sadegh et al., 2017). The normal copula is a single parameter copula with its parameter directly linked to the tail dependence coefficient \( \eta \) presented in Sect. 2.2. As showed by Ledford and Tawn (1997), the normal copula is suitable for the whole range of behaviour within the class of asymptotic independence (i.e. from sub-asymptotic positive to negative association). We use the normal copula as a reference for the asymptotic independence case; the normal copula is expressed as (Sadegh et al., 2017):

\[
C(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)}\right) dxdy
\]

(S20)

with \( \Phi(. \) the standard Gaussian distribution function and \( \rho \in [-1, 1] \) the dependence parameter. The FGM copula exhibits near independent joint tail dependence behaviour, meaning that the coefficient of tail dependence is \( \eta = 0.5 \) (Ledford and Tawn, 1997).
S2 Level curve density. Associated Section: Section 2.3: Return Period in the bivariate framework and Section 3.2: Diagnostic tools

As mentioned in main text, Section 2.3 level curves are composed of an infinite set of bivariate values all corresponding to the same probability of exceedance. In the context of multi-hazards, events with very different properties (e.g., a storm with heavy rain and moderate wind vs. another storm with moderate rain and heavy wind) can have the same return period (Chebana and Ouarda, 2011; Volpi and Fiori, 2012; Sadegh et al., 2018). One approach that has been implemented when using copula models is to use the density of the associated copula to weight \((X_1, X_2)\) pairs on the curves (Volpi and Fiori, 2012). The joint density function of a copula is defined as (Volpi and Fiori, 2012):

\[
 f_{X_1,X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1,X_2}(x_1, x_2)}{\partial x_1 \partial x_2}. \tag{S22}
\]

It is then possible to identify a most-likely scenario (Gräler et al., 2013; Sadegh et al., 2018) which is the coordinate of the level curve with the highest joint density (Fig. S1). Chebana and Ouarda (2011) proposed the decomposition of level curves into a naïve part (tail) and a proper part (central). Volpi and Fiori (2012) defined a level of probability to determine lower and upper limits of the proper part of the level curve. The most likely scenario and proper part of the level curve are shown in Fig. S1, respectively by the purple dot and the curve domain between blue diamonds along the level curve. The joint density probability function of copulas has also been used to estimate joint confidence intervals for level curves (Dung et al., 2015; Zhang et al., 2015; Serinaldi, 2016).

Figure S1: Level curve density from low (green) to high (red) for a probability of joint \((X_1, X_2)\) exceedance \(p = 0.001\) with its density (5000 realisations on a normal copula with log-normal distributions). The purple dot represents the most likely scenario while the two blue diamonds represent upper and lower bound of the proper part of the curve with a 95% confidence level.

The level curve density can be estimated from parametric models (i.e., copula). However, it is also possible to estimate density with a kernel density estimator when enough data are available (main text, Section 2.3.3). For extreme low probability level
curves as the ones we are interested in this study, there are few or no data. The simulation of extreme bivariate data with the conditional extremes model (main text, Section 2.3.2) overcomes this limitation, it is then possible to estimate the level curve density via a kernel density estimator.
References


