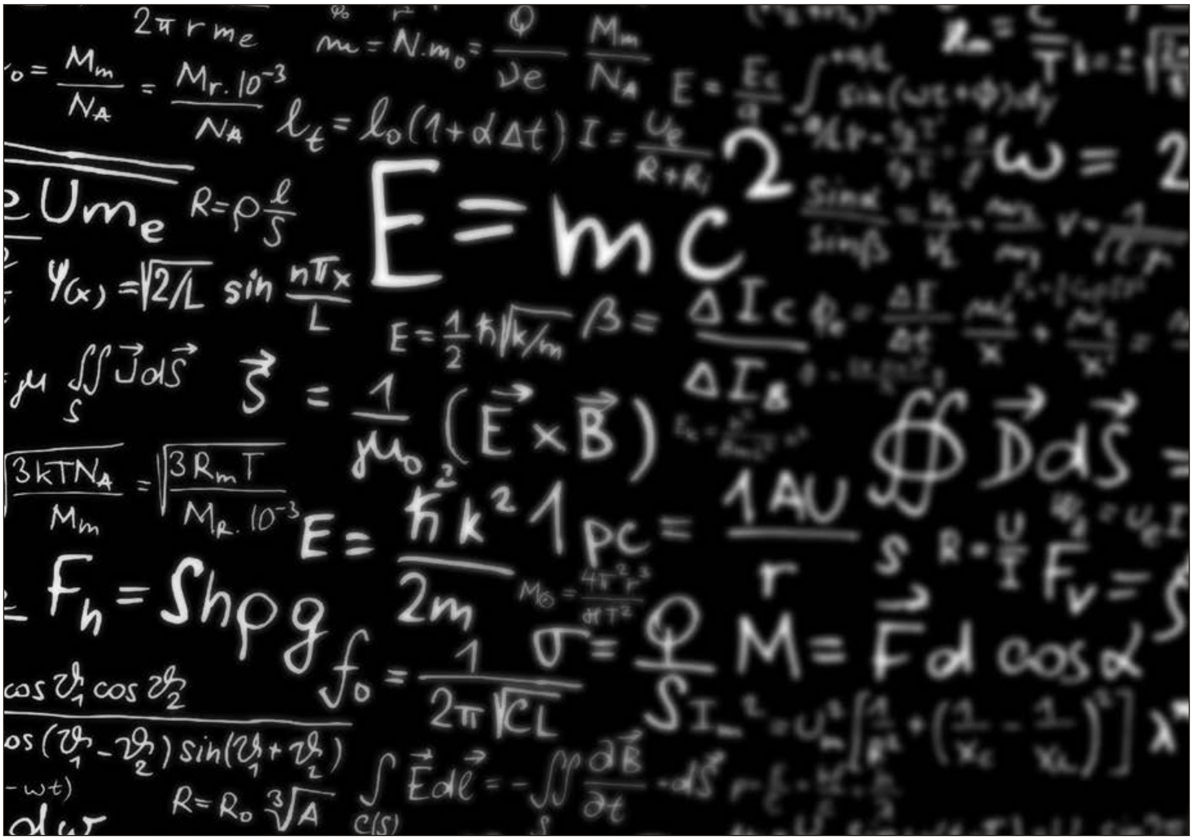

Asian Journal of Mathematics & Statistics

ISSN 1994-5418

science alert

ANSInet
an open access publisher
http://ansinet.com

Asian Journal of Mathematics and Statistics 2 (2): 20-32, 2009
ISSN 1994-5418

Conditional Dependence of Trivariate Generalized Pareto Distributions

Diakarya Barro
Department of Academic, Laboratoire LANIBIO,
UFR-SEA, Université de Ouagadougou, Burkina Faso

Abstract: In this study we consider the dependence of the family of multivariate generalized Pareto distributions under given conditions on lower dimensional margins. A new function which describes this conditional dependence is built via Pickands dependence function. This function provides a new characterization of the basic subfamilies of trivariate generalized Pareto distributions.

Key words: Pickands dependence function, discordance degree, angular distribution function, discordance function, trivariate Pareto distributions

INTRODUCTION

Extreme Values Theory (EVT) is based on modelling and measuring events which occur with very small probability. Mainly two methods have been developed in this theory: the block maxima method (Beirlant *et al*., 2005) and the Peacks-Over-Threshold (POT) approach (Coles, 2001; Resnick, 1987). The block maxima method is interested to asymptotic behavior of the laws of the component-wise maxima appropriately normalized under the condition that the univariate margins are independent and identically distributed (iid). This method shows that these asymptotic laws are the Multivariate Extreme Value Distributions (MEVDs). Suggested originally by hydrologists, the POT approach is rather based on modelling of exceedances of a random sample over a large threshold within a time period.

Earlier studies have developed statistical structures to describe the dependence of the multivariate distributions arising from these two approaches. In many latest books and reviews on the topic (Beirlant *et al*., 2005; Gaume, 2005), it has been shown that no single parametric family can summary the MEVDs like does do the Generalized Extreme Value (GEV) family in the univariate EVT. Nevertheless, if the univariate margins are given the dependence of these distributions can be characterized by equivalent measures like Pickands dependence function, exponent measure or stable tail dependence function (Degen, 2006). Furthermore, Tajvidi has shown (Tajvidi, 1996) that for a sample of random vectors $\{X_n; n \geq 1\} = \{(X_{n,1},...,X_{n,m}); m \geq 1\}$ the law H of the exceedances over a large threshold is the multivariate Generalized Pareto Distribution (GPD). Moreover, this excess distribution H is linked to the asymptotic component-wise maxima model G of the same sample by Eq. 1:

$$H(x) = \frac{-1}{\log G(0)} \log\left[\frac{G(x)}{G(\min(x,0))}\right];\ x = (x_1,...,x_n) \in IR^n \tag{1}$$

The aim and particularity of this study is to build a new structure which describes the dependence of multivariate GPDs but under given conditions made on the marginal distributions. This new conditional dependence function enable us to characterize the basic parametric subfamilies of three-dimensional GPDs.

MATERIALS AND METHODS

In this study, we consider the following problem: Let consider a situation where $X = \{(X_1,...,X_n);n\geq 1\}$ is a random vector with a multivariate GPD function H and we are interested to model a structure which describes both the dependence of H under the condition $X_I>x_{0,I}$ and the dependence of the survival function $\overline{H}$ of H under the condition $X_J\leq x_{0,J}$; $x_{0,I}$ and $x_{0,J}$ being given realizations of the complementary lower dimensional margins X_I and X_J of X. Therefore, it is desirable to model the structure which gives at any realization $x = (x_1,...,x_n)$ of X the probability of the discordances $\{X_I \leq x_{0,I} / X_J > x_{0,J}\}$ and $\{X_I > x_{0,I} / X_J \leq x_{0,J}\}$. For this purpose Pickands dependence function of a MEVD would be useful (Coles, 2001; Beirlant *et al*., 2005). Let:

$$M_n = (M_{n,1},...,M_{n,m}) = \left(\max_{1\leq i\leq n}(X_{i,1}),...,\max_{1\leq i\leq n}(X_{i,m})\right)$$

be the component-wise maxima of a random vector $\{X_n, n\geq 1\} = \{(X_{n,1},...,X_{n,m}); m\geq 1\}$ with univariate iid variables and with distribution function F. A m-dimensional continuous and non-degenerate function G is a MEVD if there exists vectors of normalizing sequences $\sigma_n = (\sigma_{n,1},...,\sigma_{n,m}) \in IR^m$ with $\{\sigma_{n,j} > 0, 1\leq j\leq m\}$ and $\mu_n = (\mu_{n,1},...,\mu_{n,m}) \in IR^m$ such that (in component-wise algebraic notations), for all $x = (x_1,...,x_m) \in IR^m$:

$$\lim_{n\to+\infty} P\left(\frac{M_n - \mu_n}{\sigma_n} \leq x\right) = \lim_{n\to+\infty} F^n(\sigma_n x + \mu_n) = G(x) \tag{2}$$

If Eq. 2 holds F is said to belong to the max-domain of attraction of G. Therefore, from the link established between G and H by Eq. 1, we obtain, for all $x = (x_1,...,x_m)\in IR^m$, the following characterization:

$$H(x) = 1 + \left\{\left\{-\sum_{i=1}^{m} y_i(x_i)\right\} A\left(\frac{y_1(x_1)}{\sum_{i=1}^{m} y_i(x_i)},\ldots,\frac{y_{m-1}(x_{m-1})}{\sum_{i=1}^{m} y_i(x_i)}\right)\right\} = 1 + \log G(x) \tag{3}$$

where, A is the Pickands dependence function of H defined on the unit simplex

$$S_{m-1} = \left\{t = (t_1,...,t_{m-1}) \in [0,1]^{m-1}; \sum_{i=1}^{m-1} t_i \leq 1\right\}$$

in IR^{m-1} such that:

$$A(t) = \int_{S_{m-1}} \max\left(-q_1 t_1,...,-q_{m-1}t_{m-1}, -q_m\left(1-\sum_{i=1}^{m-1} t_i\right)\right)\mu d(q)$$

verifying for $t\in S_{m-1}$, the condition:

$$\max\left(t_1,\ldots,t_{m-1}, 1-\sum_{i=1}^{m-1} t_i\right) \leq A(t) \leq 1$$

μ being the angular measure of H on S_{m-1} (Beirlant *et al*., 2005). In addition, the y_i are defined by the transformations:

$$y_i(x_i) = \left[1+\xi_i\left(\frac{x_i-\mu_i}{\sigma_i}\right)\right]_+^{-\frac{1}{\xi_i}},\ i=1,..,m;\ \text{with } x_+ = \max(x,0) \qquad (4)$$

where, $\mu_i \in IR$. ξ_i. $\in IR$ and $\sigma_i>0$ are respectively the location, shape and scale parameters of the univariate margins G_i of G. The different values of the parameter ξ_i allow the GEV distribution defined by:

$$GEV(x_i) = G_i(x_i) = \begin{cases} \exp\left\{-\left[1+\xi_i\left(\frac{x_i-\mu_i}{\sigma_i}\right)\right]_+^{-\frac{1}{\xi_i}}\right\} & \text{if } \xi_i \neq 0 \\ \exp\left\{-\exp\left[-\left(\frac{x_i-\mu_i}{\sigma_i}\right)_+\right]\right\} & \text{if } \xi_i = 0 \end{cases} \qquad (5)$$

to describe the three types of asymptotic extreme behavior such as:

$$GEV_i(x_i) = \begin{cases} \exp\left\{-\exp\left[-\left(\frac{x_i-\mu_i}{\sigma_i}\right)\right]\right\} = \Lambda(x_i); & \text{if } \xi_i = 0,\ x_i \in IR \\ \exp\left\{-\left[1+\xi_i\left(\frac{x_i-\mu_i}{\sigma_i}\right)\right]_+^{-\frac{1}{\xi_i}}\right\} = \Phi_{\frac{1}{\xi_i}}(x_i); & \text{if } \xi_i > 0,\ x_i > \mu_i \\ \exp\left\{-\left[-\left(\frac{x_i-\mu_i}{\sigma_i}\right)\right]_+^{\frac{1}{\xi_i}}\right\} = \Psi_{\frac{1}{\xi_i}}(x_i); & \text{if } \xi_i < 0,\ x_i < \mu_i \end{cases} \qquad (6)$$

The laws **Λ**, **Φ** and **Ψ** are from Gumbel, Fréchet and Weibull, respectively.

RESULTS

Here, the main three theorems of this study will be presented and proved.

Angular Distribution of Multivariate GPDs

The following theorem gives the angular distribution of a multivariate GPD.

Theorem 1

Let H be a multivariate GPD. Then, there exists a function L(.) defined on S_{m-1} in IR^{m-1} such as, for all $x = (x_1,...,x_m) \in IR^m$:

$$H(x_1,\dots,x_m) = 1+\left\{-L\left(\frac{x_1}{\sum_{i=1}^m x_i},\dots,\frac{x_{m-1}}{\sum_{i=1}^m x_i}\right)\right\} \qquad (8)$$

Moreover, if H is continuously differentiable of order m, the density function l of L fulfills, for all $t = (t_1,...,t_{m-1}) \in S_{m-1}$, the equality:

$$l(t_1,\dots,t_{m-1}) = -\left(\sum_{i=1}^{m} t_i\right)^{-(m+1)} \times \frac{\partial^m}{\partial t_1 \dots \partial t_m} H\left(t_1^{-1},\dots,t_m^{-1}\right) \tag{9}$$

The function L(.) is the angular distribution of the multivariate GPD H.

Proof

Let V be the exponent measure function of the distribution H with unit Fréchet margins (Michel, 2006; Resnick, 1987). Therefore, for all $x_i>0$ we have:

$$H(x_1,\dots,x_m) = 1 - V\left(\frac{-1}{x_1},\dots,\frac{-1}{x_m}\right) \tag{10}$$

It is known that:

$$\frac{\partial^m}{\partial x_1 \dots \partial x_m} \tilde{V}(x_1,\dots,x_m) = -\left(\sum_{i=1}^{m} x_i\right)^{-(m+1)} \times l\left(\frac{x_1}{\sum_{i=1}^{m} x_i},\dots,\frac{x_{m-1}}{\sum_{i=1}^{m} x_i}\right) \tag{11}$$

Where:

$$\tilde{V}(x_1,\dots,x_m) = V\left(\frac{-1}{x_1},\dots,\frac{-1}{x_m}\right)$$

$l(.)$ being the angular density of H. Furthermore, we have:

$$\frac{\partial^m}{\partial x_1 \dots \partial x_m} \tilde{V}(x_1,\dots,x_m) = \frac{1}{x_i^2 \dots x_m^2}\left(\frac{\partial^m}{\partial x_1 \dots \partial x_m} V\right)\left(\frac{-1}{x_1},\dots,\frac{-1}{x_m}\right)$$

This result inserted in Eq. 11 gives :

$$\left(\frac{\partial^m}{\partial x_1 \dots \partial x_m} V\right)\left(\frac{-1}{x_1},\dots,\frac{-1}{x_m}\right) = \frac{-1}{x_i^2 \dots x_m^2}\left(\sum_{i=1}^{m} x_i\right)^{-(m+1)} \times l\left(\frac{x_1}{\sum_{i=1}^{m} x_i},\dots,\frac{x_{m-1}}{\sum_{i=1}^{m} x_i}\right)$$

By taking:

$$t_i = \frac{x_i}{\sum_{i=1}^{m} x_i}$$

We have:

$$\left(\frac{\partial^m}{\partial t_1 \dots \partial t_m} V\right)(t_1,\dots,t_m) = \left(\sum_{i=1}^{m} t_i\right)^{-(m+1)} \times l(t_1,\dots,t_{m-1})$$

Replacing $l(t_1,\dots,t_{m-1})$ in Eq. 10 we see that Eq. 9 assertion holds.

For proving the following theorems we define a conditional dependence measure for the family of multivariate distributions.

A Conditional Measure of a Multivariate Distribution

Let n, k be natural numbers such that $\{n\geq 2; 1\leq k<n\}$ and let N_k be a given subset of k elements of $N = \{1,...,n\}$, the set of the first n natural numbers.

Definition 1

We define N_k-partition of a random vector $X = \{(X_1,...,X_n), n\geq 2\}$ (or the partition of X in the direction of N_k) by the pairwise vector $\tilde{X} = \left(\tilde{X}_{N_k}, \tilde{X}_{\bar{N}_k}\right)$ as:

- $\tilde{X}_{N_k} = \left(X_{N_k,1},...,X_{N_k,k}\right)$

  is the k-dimensional marginal vector of X whose component indexes are ordered in the subset N_k
- $\tilde{X}_{\bar{N}_k} = \left(X_{\bar{N}_k,1},...,X_{\bar{N}_k,n-k}\right)$

  is the (n-k)-dimensional marginal vector of X whose component indexes are ordered in $\bar{N}_k = C_N^{N_k}$, the complementary of N_k in N

Similarly, every realization $x = (x_1,..,x_n)$ of X can be decomposed into two parts

$$x = (\tilde{x}_{N_k}, \tilde{x}_{\bar{N}_k}) \text{ where } \tilde{x}_{N_k} = \left(x_{N_k,1},...,x_{N_k,k}\right) \text{ and } \tilde{x}_{\bar{N}_k} = \left(x_{\bar{N}_k,1},...,x_{\bar{N}_k,n-k}\right)$$

are, respectively realizations of vectors $\tilde{X}_{N_k}$ et $\tilde{X}_{\bar{N}_k}$. If H, H_{N_k} and $H_{\bar{N}_k}$ denote the distribution functions of the random vectors X, $\tilde{X}_{N_k}$ and $\tilde{X}_{\bar{N}_k}$ then for all realization $x = (x_1,...,x_n)$ of X we have

$$H_{N_k}(\tilde{x}_{N_k}) = \lim_{\tilde{x}_{\bar{N}_k}\to\tilde{x}^*_{\bar{N}_k}} H(x) \text{ and } H_{\bar{N}_k}\left(\tilde{x}_{\bar{N}_k}\right) = \lim_{\tilde{x}_{N_k}\to\tilde{x}^*_{N_k}} H(x) \text{ where } \tilde{x}^*_{N_k} = \left(x^*_{N_k,1},...,x^*_{N_k,k}\right) \text{ and } \tilde{x}^*_{\bar{N}_k} = \left(x^*_{\bar{N}_k,1},...,x^*_{\bar{N}_k,n-k}\right)$$

are the upper endpoints of the functions H_{N_k} and $\tilde{X}_{\bar{N}_k}$.

Definition 2

Given a N_k-partition $\tilde{X} = \{(\tilde{X}_{N_k}, \tilde{X}_{\bar{N}_k}), 1\leq k<n\}$ of $X = (X_1,...,X_n)$ we define the upper N_k-discordance degree of X as the conditional probability given for all $x = (x_1,..., x_n) \in IR^n$ by $\delta^+_{N_k}(x) = P(\tilde{X}_{N_k} > \tilde{x}_{N_k} / \tilde{X}_{\bar{N}_k} \leq \tilde{x}_{\bar{N}_k})$. Similarly, the lower N_k-discordance degree of X is defined, for all $x = (x_1,...,x_n)$ in IR^n by $\delta^-_{N_k}(x) = P(\tilde{X}_{N_k} \leq \tilde{x}_{N_k} / \tilde{X}_{\bar{N}_k} > \tilde{x}_{\bar{N}_k})$.

The following definition characterizes the probability that one of the margins $\tilde{X}_{N_k}$ and $\tilde{X}_{\bar{N}_k}$ exceeds 1/2, while the values taken by the other are less than 1/2.

Definition 3

Given the distribution H of a multivariate random vector $X = \{(X_1,...,X_n), n\geq 2\}$ with univariate margins H_i, $1\leq i\leq n$ we define the upper N_k-median discordance degree of H by the real number denoted by $\delta^+_{N_k,H}$ such as: $\delta^+_{N_k,H} = \delta^+_{N_k}\left[\left(H_1^{-1}(\tfrac{1}{2}),..., H_n^{-1}(\tfrac{1}{2})\right)\right]$ where H_i^{-1} is quantile function of H_i. Similarly, the lower N_k-median discordance degree of H is defined by $\delta^-_{N_k,H} = \delta^-_{N_k}\left[\left(H_1^{-1}(\tfrac{1}{2}),..., H_n^{-1}(\tfrac{1}{2})\right)\right]$.

Example 1

Let $X = (X_1, X_2, X_3)$ be a trivariate random vector. Let's consider $N_2 = \{1, 3\}$. The lower N_2-discordance degree of X is given, for $x = (x_1, x_2, x_3) \in IR^3$ by $\delta^-_{N_2}(x_1,x_2,x_3) = P(X_1 \leq x_1, X_3 \leq x_3 / X_2 > x_2)$. Particularly, if H is a continuous distribution function of X we verify easily that $\delta^-_{N_2}(x_1,x_2,x_3) = \left[H_{1,3}(x_1,x_3) - H(x_1,x_2,x_3)\right]\bar{H}_2^{-1}(x_2)$, while for $N_1 = \{1\}$, the upper N_1-median discordance degree is given by $\delta^+_{N_1,H} = 1 - H\left[H_1^{-1}(\tfrac{1}{2}), H_2^{-1}(\tfrac{1}{2}), H_3^{-1}(\tfrac{1}{2})\right]\bar{H}_{2,3}^{-1}\left[H_1^{-1}(\tfrac{1}{2}), H_2^{-1}(\tfrac{1}{2})\right]$where

$H_2>0$ and $H_{1,3}>0$ are, respectively the distribution functions of the margins X_2 and (X_1, X_3); $\bar{H}_i^{-1}$ being the inverse of the survival function of H_i. Let's suppose, in addition that:

$$H_\theta(x_1, x_2, x_3) = \exp\left\{-\left[(-\ln x_1)^\theta + (-\ln x_2)^\theta + (-\ln x_3)^\theta\right]^{\frac{1}{\theta}}\right\}, \theta > 0$$

for all $x_i \in]0,1]$ (distribution whose univariate margins are uniform on 0,1]), we check easily that, for $\theta = 1/2$ we get :

$$\delta^+_{N_1,H} = \delta^+_{N_1}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = 0.96875$$

The following result shows that the N_k-marginal distribution of a MEVD is also a MEVD.

Theorem 2

Suppose there exists a MEVD G describing the asymptotic behavior of the component-wise maxima of X suitably normalized. Then, there exists a k-dimensional MEVD G_k and a (n-k)-dimensional MEVD G_{n-k} associated, respectively to the component-wise maxima of the marginal vectors $\tilde{X}_{N_k}$ and $\tilde{X}_{\bar{N}_k}$. Moreover G_k and G_{n-k} are the marginal distributions of G.

Proof

Let $\sigma_n = (\sigma_1,...,\sigma_n) \in IR^n$; $\sigma_i>0$ and $\mu_n = (\mu_1;...;\mu_n) \in IR^n$ be the vectors of normalizing sequences of the component-wise maxima $M_n = (M_1,...,M_n)$ associated to the MEVD G by previous Eq. 2. Then, if

$$x^*_{\bar{N}_k} = \left(x^*_{1,\bar{N}_k};...;x^*_{n-k,\bar{N}_k}\right)$$

is the upper endpoint of marginal vector $\tilde{X}_{\bar{N}_k}$, we have:

$$\lim_{x_{\bar{N}_k} \to x^*_{\bar{N}_k}} G(x_1,...,x_m) = \lim_{x_{\bar{N}_k} \to x^*_{\bar{N}_k}} \left[\lim_{n\to+\infty} P\left(\frac{M_1-\mu_1}{\sigma_1} \le x_1;...;\frac{M_n-\mu_n}{\sigma_n} \le x_n\right)\right]$$

$$= \lim_{n\to+\infty}\left[\lim_{x_{\bar{N}_k} \to x^*_{\bar{N}_k}} P\left(\frac{M_1-\mu_1}{\sigma_1} \le x_1;...;\frac{M_n-\mu_n}{\sigma_n} \le x_n\right)\right]$$

$$= \lim_{n\to+\infty}\left[P\left(\frac{M_{1,N_k}-\mu_{1,N_k}}{\sigma_{1,N_k}} \le x_{1,N_k};...;\frac{M_{k,N_k}-\mu_{k,N_k}}{\sigma_{k,N_k}} \le x_{k,N_k}\right)\right]$$

$$= G_k\left(x_{1,N_k},...,x_{k,N_k}\right)$$

Therefore, there exists two vectors of normalizing sequences $\sigma_{N_k} = (\sigma_{1,N_k},...,\sigma_{k,N_k}) \in IR^k$, $\sigma_{i,N_k}>0$ and $\mu_{N_k} = (\mu_{1,N_k};...;\mu_{k,N_k}) \in IR^k$ such as the marginal component-wise maxima $M_{N_k} = (M_{1,N_k},...,M_{k,N_k})$ of M_n converges to G_{N_k} according to Eq. 2. Thereby G_k is a MEVD with k variables.

Similarly, we establish that G_{n-k} also arises as the limiting distribution of the marginal component-wise maxima $M_{\bar{N}_k} = (M_{1,\bar{N}_k},...,M_{n-k,\bar{N}_k})$ linearly normalized with vectors of sequences $\sigma_{\bar{N}_k} = (\sigma_{1,\bar{N}_k},...,\sigma_{n-k,\bar{N}_k}) \in IR^{n-k}$ with $\sigma_{j,\bar{N}_k} > 0$, $1 \le j \le n-k$ and $\mu_{\bar{N}_k} = (\mu_{1,\bar{N}_k},...,\mu_{n-k,\bar{N}_k}) \in IR^{n-k}$.

A Conditional Dependence Function of a Multivariate GPD

Note that, in the above, each of the discordance degrees $\delta^{+}_{N_k}$ and $\delta^{-}_{N_k}$ of a random vector X can be obtained by functional transformations of the other. Therefore, the following characterizations will be restricted to the upper $\delta^{+}_{N_k}$ which will be denoted by δ in the simplest case $\delta^{+}_{N_1}$ i.e., k = 1.

Theorem 3

Let G be a MEVD with discordance degree δ. Then, there exists a convex function D defined on the unit simplex S_{m-1} by:

$$\delta(x_1,...,x_m) = 1 - \exp\left\{\left\{-\sum_{i=1}^{m} y_i(x_i)\right\} D\left(\frac{y_1(x_1)}{\sum_{i=1}^{m} y_i(x_i)},...,\frac{y_{m-1}(x_{m-1})}{\sum_{i=1}^{m} y_i(x_i)}\right)\right\} \tag{7}$$

for all $(x_1,...,x_m)$ in IR^n ; where the $y_i(x_i)$ satisfy Eq. 4, for i = 1,…, m.
D is called the discordance function of G or of its corresponding GPD H.

Taking a MEVD with unit Gumbel margin, $G_i(x_i) = \exp\{-\exp(-x_i)\}$; $x_i>0$, the following corollary characterizes the simplest upper median degree, $\delta^{+}_{N_1,G}$.

Corollary

Let G be a MEVD with unit Gumbel. Then, the upper median discordande degree of G, denoted by $\tilde{\delta}_G$ is given by

$$\tilde{\delta}_G = \delta^{+}_{N_1,G} = 1 - (\ln 2)^{nD\left(\frac{1}{n},...,\frac{1}{n}\right)}$$

Example 2

Let G_θ, θ>1 be the logistic model of MEVD given for $(x_1, x_2, x_3)\in IR^3$ by

$$G_\theta(x_1,x_2,x_3) = \exp\left\{-\left[y_1^\theta(x_1) + y_2^\theta(x_2) + y_3^\theta(x_3)\right]^{\frac{1}{\theta}}\right\}$$

$y_i(x_i)$ satisfying Eq. 4. Its discordance function is given, for all $(t_1, t_2)\in S_2$ by

$$D_\theta(t_1,t_2) = \left[t_1^\theta + t_2^\theta + (1-t_1-t_2)^\theta\right]^{\frac{1}{\theta}} - \left[t_2^\theta + (1-t_2)^\theta\right]^{\frac{1}{\theta}}$$

and the median discordance degree

$$\bar{\delta}_{G_\theta} = 1 - (\ln 2)^{3D_\theta\left(\frac{1}{3},\frac{1}{3}\right)}$$

Particularly, for θ = 2, we get $\bar{\delta}_{G_\theta} = 0.864$.

Proof of Theorem 3

Let's suppose that, for i = 1,…,m, the univariate margins G_i of the MEVD G have the generalized form: $G_i(x_i) = \exp\{-y_i(x_i)\}$ where the $y_i(x_i)$ satisfy Eq. 4. Therefore, for all $(x_1,...,x_n)\in IR^n$;

$$\delta(x_1,...,x_n) = P(X_1 > x_1/X_i \le x_i; 2 \le i \le n) = 1 - \frac{P(X_j \le x_j; 1 \le j \le n)}{P(X_i \le x_i; 2 \le i \le n)} \tag{12}$$

Thus, $G_{\bar{N}_1}$ being the joint distribution of the margin vector $(X_2,\ldots,X_n)$ of X, we have:

$$\delta(x_1,\ldots,x_n) = 1 - \frac{G(x_1,\ldots,x_n)}{G_{\bar{N}_1}(x_2,\ldots,x_n)}$$

Furthermore, due to theorem 1, the function $G_{\bar{N}_1}$ is a k-dimensional extreme value distribution. Therefore, if A and $A_{\bar{N}_1}$ are the Pickands dependence functions of G and $G_{\bar{N}_1}$, respectively, then, in Eq. 12 we have:

$$\frac{G(x_1,\ldots,x_n)}{G_{\bar{N}_1}(x_2,\ldots,x_n)} = \exp\left\{-\sum_{i=1}^{m} y_i A\left(\frac{y_1}{\sum_{i=1}^{m} y_i},\ldots,\frac{y_{m-1}}{\sum_{i=1}^{m} y_i}\right) + \sum_{i=2}^{m} y_i A_{\bar{N}_1}\left(\frac{y_2}{\sum_{i=2}^{m} y_i},\ldots,\frac{y_{m-1}}{\sum_{i=2}^{m} y_i}\right)\right\}$$

Furthermore, we have:

$$A_{\bar{N}_1}\left(\frac{y_2}{\sum_{i=2}^{m} y_i},\ldots,\frac{y_{m-1}}{\sum_{i=2}^{m} y_i}\right) = A_{\bar{N}_1}\left(\frac{t_2}{1-t_1},\ldots,\frac{t_{m-1}}{1-t_1}\right)$$

where $t_j = \frac{y_j}{\sum_{i=1}^{m} y_i} \in]0,1[$, for all $j = 2,..,m-1$. Therefore, it follows that :

$$\frac{G(x_1,\ldots,x_n)}{G_{\bar{N}_1}(x_2,\ldots,x_n)} = \exp\left\{-\sum_{i=1}^{m} y_i \left[A(t_1,\ldots,t_{m-1}) + (1-t_1) A_{\bar{N}_1}\left(\frac{t_2}{1-t_1},\ldots,\frac{t_{m-1}}{1-t_1}\right)\right]\right\}$$

$$= \exp\left\{-\sum_{i=1}^{m} y_i D\left(\frac{y_1}{\sum_{i=1}^{m} y_i},\ldots,\frac{y_{m-1}}{\sum_{i=1}^{m} y_i}\right)\right\}$$

where, D is the convex function such that $0 \le 1\text{-}t_1 \le 1$ and defined on the unit simplex

$$S_{m-1} = \left\{t = (t_1,\ldots,t_{m-1}) \in [0,1]^{m-1}; \sum_{i=1}^{m} t_i \le 1\right\} \subset IR^{m-1}$$

by

$$D(t) = A(t_1,\ldots,t_{m-1}) + (1-t_1) A_{\bar{N}_1}\left(\frac{t_2}{1-t_1},\ldots,\frac{t_{m-1}}{1-t_1}\right)$$

for all $t \in S_{m-1}$. Particularly for the trivariate case, we have:

$$D(t_1, t_2) = A(t_1,t_2) + (1\text{-}t_1)\ A_{\bar{N}_1}\left[\frac{t_2}{1-t_1}\right]$$

defined on S_2.

APPLICATION TO THE TRIVARIATE MODELS OF GPDs

The logistic model is the most important family of multivariate GPDs.

The Family of Trivariate GPD of Logistic Type

Let $X = (X_1, X_2, X_3)$ be a trivariate random vector with a parametric distribution H_θ, $\theta>1$. The above Eq. 3 enable us to characterize H_θ by its discordance function D_θ via its Pickands dependence function A_θ (Michel, 2006).

Definition 4

The trivariate parametric function H_θ is a MGPD of Logistic Type if H_θ has, for all $(x_1, x_2, x_3)\in IR^3$ the representation:

$$H_\theta(x_1,x_2,x_3) = 1 + \left\{\left\{-\sum_{i=1}^{3} y_i(x_i)\right\} D_\theta\left(\frac{y_1(x_1)}{\sum_{i=1}^{3} y_i(x_i)}, \frac{y_2(x_2)}{\sum_{i=1}^{3} y_i(x_i)}\right)\right\}$$

With the $y_i(x_i)$ satisfying Eq. 4 and where the discordance function D_θ of H_θ is given for all:

$$(t_1,t_2)\in S_2, \text{ by } D_\theta(t_1,t_2) = \left[t_1^\theta + t_2^\theta + (1-t_1-t_2)^\theta\right]^{\frac{1}{\theta}} - \left[t_2^\theta + (1-t_2)^\theta\right]^{\frac{1}{\theta}}$$

We give here three basic trivariate GPDs of Logistic Type (Joe, 1997; Hüsler and Reiss, 1989) and we build their discordance functions:

- **The trivariate family of GPD of Logistic Type of Gumbel**

$$H_\theta(x_1,x_2,x_3) = 1 + \left\{-\left[y_1^\theta(x_1) + y_2^\theta(x_2) + y_3^\theta(x_3)\right]^{\frac{1}{\theta}}\right\}, \text{ for } (x_1,x_2,x_3)\in IR^3 \text{ and } \theta>1, \text{ we have}$$

$$D_\theta(t_1,t_2) = \left[t_1^\theta + t_2^\theta + (1-t_1-t_2)^\theta\right]^{\frac{1}{\theta}} - \left[t_2^\theta + (1-t_2)^\theta\right]^{\frac{1}{\theta}}, \text{ for all } (t_1, t_2)\in S_2$$

Particularly if $\theta\to1^-$ we obtain the trivariate Pareto independent model $H(x_1, x_2, x_3)=1+\{-y_1(x_1) +y_2(x_2)+y_3(x_3)]\}$ with $D(t_1, t_2) = 2-t_1$ for $(t_1, t_2)\in S_2$

- **The trivariate family of GPD of Logistic Type of Galambos**

$$H_\theta(x_1,x_2,x_3) = 1 + \left\{-\left(\left[y_1^{\theta_1} + y_2^{\theta_1} + y_3^{\theta_1}\right] - \left(y_1^{-\theta_1\theta_2} + y_2^{-\theta_1\theta_2} + y_3^{-\theta_1\theta_2}\right)^{\frac{-1}{\theta_2}}\right)^{\frac{1}{\theta_1}}\right\}$$

$(x_1,x_2,x_3)\in IR^3$ and $\theta_1\geq1$, $\theta_2\geq1$. We have, for all $(t_1,t_2)\in S_2$

$$D_\theta(t_1,t_2) = \left(t_2^{\theta_1} + (1-t_2)^{\theta_1} - \left(t_2^{-\theta_1\theta_2} + (1-t_2)^{-\theta_1\theta_2}\right)^{\frac{-1}{\theta_2}}\right)^{\frac{1}{\theta_1}} + \left(\left[t_1^{\theta_1} + t_2^{\theta_1} + (1-t_1-t_2)^{\theta_1}\right] - \left(t_1^{-\theta_1\theta_2} + t_2^{-\theta_1\theta_2} + (1-t_1-t_2)^{-\theta_1\theta_2}\right)^{\frac{-1}{\theta_2}}\right)^{\frac{1}{\theta_1}}$$

- **The trivariate family of GPD of Logistic Type of Hüsler-Réiss**

  The GPD which describes the behavior of the exceedances of the trivariate normal distribution over a large threshold is given for $(x_1, x_2, x_3)\in IR^3$; $\theta = (\theta_1, \theta_2, \theta_3)$ by:

$$
\begin{aligned}
H_{\theta_1,\theta_2,\theta_3}(x_1,x_2,x_3) &= 1+\Big\{-\left[y_1+y_2+y_3\right] \\
&+y_1\left[\Phi\left(\frac{1}{\theta_1}+\frac{\theta_1}{2}\log\left(\frac{y_1}{y_2}\right)\right)+\Phi\left(\frac{1}{\theta_3}+\frac{\theta_3}{2}\log\left(\frac{y_1}{y_3}\right)\right)\right] \\
&+y_2\left[\Phi\left(\frac{1}{\theta_1}+\frac{\theta_1}{2}\log\left(\frac{y_2}{y_1}\right)\right)+\Phi\left(\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{y_2}{y_3}\right)\right)\right] \\
&+y_3\left[\Phi\left(\frac{1}{\theta_3}+\frac{\theta_3}{2}\log\left(\frac{y_3}{y_1}\right)\right)+\Phi\left(\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{y_3}{y_2}\right)\right)\right] \\
&+\int_0^{y_3}\bar{\Phi}_2\left(\frac{1}{\theta_3}+\frac{\theta_3}{2}\log\left(\frac{t}{y_1}\right),\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{t}{y_2}\right);\rho\right)dt\Big\}
\end{aligned}
$$

where, Φ notes the distribution function of the standard normal law and $\bar{\Phi}_2(.,.,\rho)$, the survival function of the bivariate normal distribution function with covariance matrix:

$$
\rho=\rho(\theta_1,\theta_2,\theta_3)=\begin{pmatrix} 0 & 2\left(\frac{1}{\theta_1^2}+\frac{1}{\theta_2^2}-\frac{1}{\theta_3^2}\right) \\ 2\left(\frac{1}{\theta_1^2}+\frac{1}{\theta_2^2}-\frac{1}{\theta_3^2}\right) & 0 \end{pmatrix}
$$

Thus, for all (t_1, t_2) the corresponding discordance function is:

$$
\begin{aligned}
D_\theta(t_1,t_2) = 2 &- t_1\left[\Phi\left(\frac{1}{\theta_1}+\frac{\theta_1}{2}\log\left(\frac{t_1}{t_2}\right)\right)+\Phi\left(\frac{1}{\theta_3}+\frac{\theta_3}{2}\log\left(\frac{t_1}{1-t_1-t_2}\right)\right)\right] \\
&- t_2\left[\Phi\left(\frac{1}{\theta_1}+\frac{\theta_1}{2}\log\left(\frac{t_2}{t_1}\right)\right)+\Phi\left(\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{t_2}{1-t_1-t_2}\right)\right)\right] \\
&- (1-t_1-t_2)\left[\Phi\left(\frac{1}{\theta_3}+\frac{\theta_3}{2}\log\left(\frac{1-t_1-t_2}{t_1}\right)\right)+\Phi\left(\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{1-t_1-t_2}{t_2}\right)\right)\right] \\
&- t_2\Phi\left(\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{t_2}{1-t_2}\right)\right)-(1-t_2)\Phi\left(\frac{1}{\theta_2}+\frac{\theta_2}{2}\log\left(\frac{1-t_2}{t_2}\right)\right)+R(t_1,t_2,\theta)
\end{aligned}
$$

where, $R(t_1, t_2, \theta)$ is an integral rest defined for all $(t_1, t_2)\in S_2$.

The Family of Trivariate GPD of Nested Type

The Nested Logistic Type is an asymmetric subfamily of logistic model. It generalizes this model to allow different degrees of dependence between the components of the underlying random vector. For $(x_1, x_2, x_3)\in IR^3$ and $\theta_1, \theta_2>1$, define now, recursively the following norm

$$
\|x\|_{\theta_1,\theta_2}=\|(x_1,x_2,x_3)\|_{\theta_1,\theta_2}=\left\|\|(x_1,x_2)\|_{\theta_1},x_3\right\|_{\theta_2}
$$

where, $\|.\|_\theta$ is the usual θ-norm with the convention that the absolute value is taken if the norm does not have an index (Joe, 1997; Michel, 2006).

Definition 5

The distribution function given, for all $(x_1, x_2, x_3)\in IR^3$, by

$$H_{\theta_1,\theta_2}(x_1,x_2,x_3) = 1 + \left\{-\left\|\left(y_1(x_1), y_2(x_2), y_3(x_3)\right)\right\|_{\theta_1,\theta_2}\right\}$$

is called the generalized Pareto distribution of Nested logistic type.

The basic trivariate GPD of Nested Logistic Type is given, for $\theta_1, \theta_2 \geq 1$ by:

$$H_{\theta_1,\theta_2}(x_1,x_2,x_3) = 1 + \left\{-\left(\left[y_1^{\theta_1}(x_1) + y_2^{\theta_1}(x_2)\right]^{\frac{\theta_2}{\theta_1}} + y_3^{\theta_2}(x_3)\right)^{\frac{1}{\theta_2}}\right\};\ (x_1,x_2,x_3) \in IR^3$$

$$\text{We have: } D_{\theta_1,\theta_2}(t_1,t_2) = \left(\left(t_1^{\theta_1} + t_2^{\theta_1}\right)^{\frac{\theta_2}{\theta_1}} + (1-t_1-t_2)^{\theta_2}\right)^{\frac{1}{\theta_2}} + t_2 - 1 \text{ with } (t_1,t_2) \in S_2.$$

The Family of Trivariate GPD of Asymmetric Logistic Type

The asymmetric distributions arise as the models which describe the asymptotic behavior of the maxima of storms recorded at different locations along a coastline (Gaume, 2005). They generalize the logistic model but does not include the nested logistic model from the previous section (Michel, 2006).

Definition 6

Let B be a non-empty subset of {1,2,3} and let $\lambda_C \geq 1$ be arbitrary numbers for every $C\subset B$ with $|C|\geq 2$ and $\lambda_C = 1$ if $|C| = 1$. Furthermore, let $0 \leq p_{i,C} \leq 1$ where, $p_{i,C} = 0$ if $I\notin C$ and the side condition

$$\sum_{C\subset B} p_{i,C} = 1$$

is fulfilled for i=1,2,3. Then the distribution function

$$H(x_1,x_2,x_3) = 1 + \left\{\left\{-\sum_{C\in B} p_{i,C}\right\}\left\{\sum_{i\in C}\left[p_{i,C}\, y_i(x_i)\right]^{\frac{1}{\lambda_C}}\right\}\right\}$$

is a generalized Pareto distribution of Asymmetric Logistic Type.

The basic trivariate GPDs of Asymmetric Logistic Type with their discordance functions follow (Joe, 1997):

- **The trivariate Asymmetric GPD of Logistic Type of Gumbel**

$$H_{\theta_1,\theta_2}(x_1,x_2,x_3) = 1 + \left\{-\left(\left[y_1^{\theta_1\theta_2} + 2^{-\theta_2} y_2^{\theta_1\theta_2}\right]^{\frac{1}{\theta_2}} + \left[2^{-\theta_2} y_2^{\theta_1\theta_2} + y_3^{\theta_1\theta_2}\right]^{\frac{1}{\theta_2}}\right)^{\frac{1}{\theta_1}}\right\}$$

For all $(x_1, x_2, x_3)\in IR^3$ and $\theta_1, \theta_2>0$ we have for all $(t_1, t_2)\in S_2$:

$$D_{\theta_1,\theta_2}(t_1,t_2) = \left(\left[t_1^{\theta_1\theta_2} + 2^{-\theta_2} t_2^{\theta_1\theta_2}\right]^{\frac{1}{\theta_2}} + \left[2^{-\theta_2} t_2^{\theta_1\theta_2} + (1-t_1-t_2)_3^{\theta_1\theta_2}\right]^{\frac{1}{\theta_2}}\right)^{\frac{1}{\theta_1}}$$

$$-\left(\left[2^{-\theta_2} t_2^{\theta_1\theta_2} + (1-t_2)^{\theta_1\theta_2}\right]^{\frac{1}{\theta_2}}\right)^{\frac{1}{\theta_1}}$$

- **The trivariate Asymmetric GPD of Logistic Type of Galambos**

$$H_{\theta_1,\theta_2}(x) = 1 + \left\{ -\left(\left[y_1^{\theta_1} + y_2^{\theta_1} + y_3^{\theta_1} \right] - \left[y_1^{-\theta_1\theta_2} + 2^{\theta_2} y_2^{-\theta_1\theta_2} \right]^{\frac{-1}{\theta_2}} - \left[y_3^{-\theta_1\theta_2} + 2^{\theta_2} y_2^{-\theta_1\theta_2} \right]^{\frac{-1}{\theta_2}} \right)^{\frac{1}{\theta_1}} \right\}$$

for $(x_1, x_2, x_3) \in IR^3$ and $\theta_1, \theta_2 > 0$. We have, for $(t_1, t_2), \in S_2$

$$D_{\theta_1,\theta_2}(t_1,t_2) = \left(t_1^{\theta_1} + t_2^{\theta_1} + (1-t_1-t_2)^{\theta_1} - \left[t_1^{-\theta_1\theta_2} + 2^{\theta_2} t_2^{-\theta_1\theta_2} \right]^{\frac{-1}{\theta_2}} \right.$$

$$\left. - \left[(1-t_1-t_2)^{-\theta_1\theta_2} + 2^{\theta_2} t_2^{-\theta_1\theta_2} \right]^{\frac{-1}{\theta_2}} \right)^{\frac{1}{\theta_1}} - \left(\left[t_2^{\theta_1} + (1-t_2)^{\theta_1} \right] - \frac{1}{2} t_2^{\theta_1} - \left((1-t_2)^{-\theta_1\theta_2} + 2^{\theta_2} t_2^{-\theta_1\theta_2} \right) \right)^{\frac{1}{\theta_1}}$$

DISCUSSION

The results of the study show that the dependence of all trivariate GDP, under given conditions on the lower dimensional margins, is totally described by its discordance function. These results are similar to the characterizations of the multivariate GPDs developed by Tajvidi (1996) or to the equivalent dependence measures for MEVDs (Resnick, 1987; Beirlant *et al*., 2005). But the particularity of this study is the that the new measure and function describe the joint dependence under any condition made on the support of a lower dimensional margin. Moreover, the applications of the study determine clearly the three main subfamilies of the models of trivariate GPDs by characterizing them by their discordance function.

We found that the results conform to the solution of the problem considered earlier. This is seen at all realisations $x = (x_1,\ldots,x_n)$ of the random vector X. We also note that the theorem 3 establishes a link between the new dependence structure and the previous dependence measures via Pickands dependence one.

CONCLUSION

In this research we have investigated about characterization of a conditional dependence of multivariate families of generalized Pareto distributions. We have built a new measure and function which describe this conditional dependence. Basic trivariate subfamilies of multivariate GPDs have been characterized by this function. Moreover, we have computed the expressions of this function for the usual trivariate subfamilies of GPDs.

REFERENCES

Beirlant, J., Y. Goegebeur, J. Segers and J. Teugels, 2005. Statistics of Extremes Theory and Applications. John Wiley and Sons, Chichester.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. 1st Edn., Springer, New York, ISBN-13: 978-1852334598.

Degen, M., 2006. On multivariate generalised pareto distributions and high risk scenarios. Diploma Thesis, Department of Mathematics, ETH Zürich.

Gaume, E., 2005. On the asymptotic behavior of flood peak distributions-theoretical Hydrol. Earth Syst. Sci. Discuss., 2: 1835-1864.

Hüsler, J. and R.D. Reiss, 1989. Extreme value theory. Proceedings of a Conference Held in Oberwolfach Dec. 6-12, 1987. Springer, Berlin.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Monographs on Statistics and Applied Probabilty 73. Chapman and Hall, London, ISBN: 9780412073311.

Michel, R., 2006. Simulation and Estimation in Multivariate Generalized Pareto Dissertation, Fakultät für Mathematik und Informatikn, Universität Würzburg, Würzburg.

Resnick, S., 1987. Extreme Values, Regular Variation and Point Processes.Extreme Values, Regular Variation and Point Processes.Extreme Values, Regular Variation and Point Processes. Springer Series of the Applied Probability Trust, Springer, New York.

Tajvidi, N., 1996. Characterisation and Some Statistical Aspects of Univariate and Multivariate Generalised Pareto Distributions. Dissertation, Department of Mathematics, Chalmers Tekniska Högskola Göteborg, Swedish.