Hoeffding–Sobol and Möbius decompositions for (tail-)dependence analysis

Cécile Mercadier

Université Claude Bernard Lyon 1

Institut Camille Jordan / UMR CNRS 5208

mercadier@math.univ-lyon1.fr

Recent coauthors

Olivier Roustant
Professor in Applied Mathematics at INSA Toulouse
Member of the Institut de Mathématiques de Toulouse

Paul Ressel
Professor emer. at Mathematisch-Geographische Fakultät
Kath. Universität Eichstätt-Ingolstadt, Eichstätt, Germany

Christian Genest
Professor in the Department of Mathematics and Statistics
McGill University, Montréal (Québec) Canada

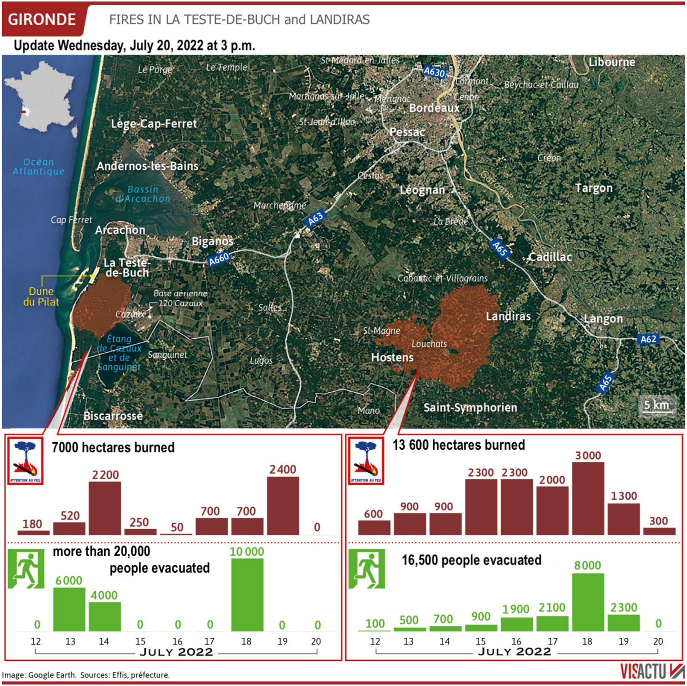
A recent example of simultaneous occurrence of extremes

“Very complicated” weather conditions

For the first time since Tuesday, the situation is "extremely unfavorable" this Sunday evening, say the firefighters. They are worried about

- the 44°C expected on Monday
- combined with very low humidity in the air (10% hygrometry)
- winds of up to 60 km/h which could prevent bomber planes from flying
- Especially since this wind is turning

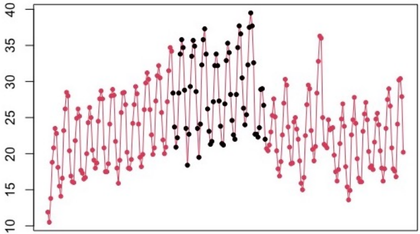
source : https ://www.francebleu.fr/infos/


GIRONDE
FIRES IN LA TESTE-DE-BUCH and LANDIRAS
Update Wednesday, July 20, 2022 at 3 p.m.
Libourne
Bordeaux
Pessac
Lège-Cap-Ferret
Andernos-les-Bains
Océan Atlantique
Bassin d'Arcachon
Cap Ferret
Arcachon
Biganos
La Teste-de-Buch
Dune du Pilat
Base aérienne 120 Cazaux
Cazaux
Étang de Cazaux et de Sanguinet
Sanguinet
Biscarrosse
Salles
Lugos
Marcheprime
Cestas
Léognan
La Brède
Targon
Cadillac
Cabanac-et-Villagrains
Landiras
Langon
St-Magne
Louchats
Hostens
Mano
Saint-Symphorien
Créon
Beychac-et-Caillau
Lormont
Cenon
Mérignac
Martignas-sur-Jalle
St-Jean-d'Illac
St-Médard-en-Jalles
Le Porge
Le Temple
Garonne
A630
A63
A660
A65
A62
5 km
ATTENTION AU FEU
7000 hectares burned
180
520
2 200
250
50
700
700
2 400
0
more than 20,000 people evacuated
0
6 000
4000
0
0
0
10 000
0
0
12
13
14
15
16
17
18
19
20
JULY 2022
ATTENTION AU FEU
13 600 hectares burned
600
900
900
2300
2300
2000
3 000
1300
300
16,500 people evacuated
100
500
700
900
1 900
2100
8 000
2300
0
12
13
14
15
16
17
18
19
20
JULY 2022
Image: Google Earth. Sources: Effis, préfecture.
VISACTU

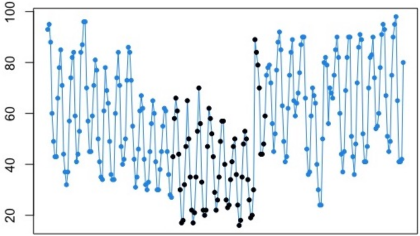
Bordeaux (July 2022)

source : donneespubliques.meteofrance.fr

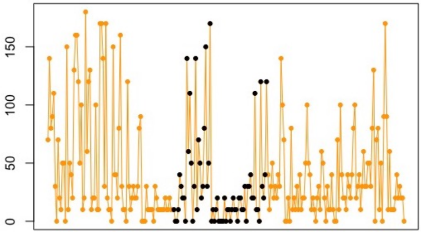
Temperature



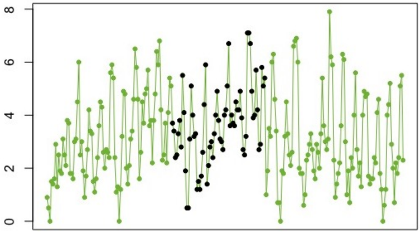
Humidity



Changes wind direction



Wind Speed



1. Motivation with recent forest fires
2. Short introduction to multivariate extreme value theory
3. The tail dependograph
4. More... and bounds for the tail dependograph
5. Links between the Hoeffding-Sobol and the Möbius decompositions

Extreme value theory

Whatever the domain of application is, Extreme value theory is concerned with what happens in the tail of a probability distribution.

It is motivated by this analysis of univariate, **multivariate**, temporal, spatial, and spatio-temporal datasets.

Understanding extreme features implies margins + **dependence modelling**

The questions are

- How much extremes are *dependent*?
- How to *measure and represent* the tail dependence?

Domain of attraction

Assumption

For $j = 1, \ldots, m$, let $\mathbf{X}^{(j)} = (X_1^{(j)}, \ldots X_d^{(j)})$ be i.i.d. copies of $\mathbf{X} = (X_1, \ldots, X_d)$ a rv $\sim F$. One assumes that there exist $\mathbb{R}^d$-sequences $\mathbf{a}^{(m)}$ and $\mathbf{b}^{(m)}$, where $a_i^{(m)} > 0$ for all $i = 1, \ldots, d$ and a d.f. G with nondegenerate margins such that, as m tends to infinity,

$$\mathbb{P}\left(\left(\max_{j=1,\ldots,m} \mathbf{X}^{(j)} - \mathbf{b}^{(m)}\right) / \mathbf{a}^{(m)} \leq \mathbf{x}\right) = F^m\left(\mathbf{a}^{(m)}\mathbf{x} + \mathbf{b}^{(m)}\right) \xrightarrow[m\to\infty]{} G(\mathbf{x})\,.$$

Then G is called a multivariate extreme value (MEV) distribution.

Its univariate margins $G_1, \ldots, G_d$ are generalized extreme value distributions

$$GEV(x) = \exp\left\{-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+\right\}$$

and G has several possible representations.

We say that F (or $\mathbf{X}$) is in the domain of attraction of G, $F \in DA(G)$.

Representations of multivariate extreme value (MEV) distributions

In terms of the exponent measure $\mu_\star$ homogeneous of order -1

$$G(\mathbf{x}) = \exp\left(-\mu_\star\left([\mathbf{0}, \infty] \setminus [\mathbf{0}, (-1/\ln G_1(x_1)), \ldots, -1/\ln G_d(x_d))]\right)\right)$$

In terms of the spectral measure H on $\mathbb{S}_1^+$ s.t. $\int_{\mathbb{S}_1^+} w_i H(d\mathbf{w}) = 1$ $(i = 1, \ldots, d)$

$$G(\mathbf{x}) = \exp\left(\int_{\mathbb{S}_1^+} \min_{i=1,\ldots,d}\{w_i \ln G_i(x_i)\} H(d\mathbf{w})\right)$$

- **Balkema and Resnick (1977)** *Max-Infinite divisibility*
- **de Haan and Resnick (1977)** *Limit theory for multivariate sample extremes*
- **Resnick (1987)** Extreme Values, Regular Variation and Point Processes
- **Beirlant, Goegebeur, Teugels and Segers (2004)** Statistics of Extremes
- **Fougères (2004)** *Multivariate extremes (chapter 7)* in Extreme Values in Finance, Telecommunications and the Environment.
- **de Haan and Ferreira (2006)** Extreme Value Theory

Representations of multivariate extreme value (MEV) distributions (continuation)

In terms of the stable tail dependence function **(stdf)** ℓ

$$G(\mathbf{x}) = \exp\left(-\ell\left(-\ln G_1(x_1), \ldots, -\ln G_d(x_d)\right)\right)$$

$$\ell(\mathbf{x}) = \int_{\mathbb{S}_{\mathbf{1}}^{+}} \max_{i=1,\ldots,d} \{w_i x_i\} H(d\mathbf{w}) \qquad \mathbf{x} \in [\mathbf{0}, \infty]$$

homogeneous of order 1 ; equals 1 at the unit vectors ; and fully d-max decreasing.

- **Ressel (2013)** *Homogeneous distributions – And a spectral representation of classical mean values and stable tail dependence functions*
- **Ressel (2022)** *Stable tail dependence functions – some basic properties*

The link between the stable tail dependence function ℓ and the original distribution functions $F, F_1, \ldots, F_d$ is given under the assumption $F \in DA(G)$:

For $\mathbf{x}$ such that $1 - F_i(x_i)$ is small

$$F(\mathbf{x}) \approx \exp\left(-\ell\left(-\ln F_1(x_1), \ldots, -\ln F_d(x_d)\right)\right)$$

Well-known measures for tail dependence

$\mathbb{P}(M_i \leq x) = \exp(-1/x)$

Assume that $\mathbf{M} = (M_1, \dots, M_d)$ has a MEV distribution G with Fréchet margins and stdf ℓ. Then, for any subset I from $\{1, ..., d\}$, define

$$\mathbb{P}(\max_{i \in I} M_i < x) = \mathbb{P}(M_i < x)^{\theta_I} .$$

The parameter θ_I is the **extremal coefficient associated with** I. It satisfies $1 \leq \theta_I \leq |I|$ where **stronger dependence** corresponds to **smaller extremal coefficients**. Equivalently,

$$\theta_I = \ell(\mathbf{1}_I, \mathbf{0}_{I^c}) = \int_{\mathbb{S}_{\mathbf{1}}^+} \max_{i \in I}\{w_i\} H(d\mathbf{w}) .$$

Furthermore, when $\mathbf{X} \sim F \in DA(G)$, and p such that $1 - p$ is small

$$\mathbb{P}(X_i \leq F_i^{-1}(p), \forall i \in I) \simeq p^{\theta_I} .$$

- **Tiago de Oliviera (1962/63)** *Structure theory of bivariate extremes, extensions*
- **Smith (1990)** *Max-stable processes and spatial extremes*
- **Schlather and Tawn (2002)** *Inequalities for the Extremal Coefficients of MEVD*
- **Ferreira and Ferreira (2018)** *Multidimensional extremal dependence coefficients*

Well-known measures for tail dependence (continuation)

Assume n observations of $\mathbf{X} = (X_1, \ldots, X_d) \in DA(G)$ denoted $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ and let $k = k(n)$ be an **intermediate sequence**. The empirical estimate of ℓ is

$$\hat{\ell}_{k,n}(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^{n} \mathbf{1}_{\left\{X_{\mathbf{1}}^{(i)} \geq X_{\mathbf{1},n-[kx_{\mathbf{1}}]+\mathbf{1},n}^{(i)} \text{ or } \ldots \text{ or } X_d^{(i)} \geq X_{d,n-[kx_d]+\mathbf{1},n}^{(i)}\right\}}$$

Under suitable conditions, it can be shown the following asymptotic expansion

$$\hat{\ell}_{k,n}(\mathbf{x}) = \ell(\mathbf{x}) + \frac{\alpha(n/k)M(\mathbf{x})}{\sqrt{k}} + \frac{Z_\ell(\mathbf{x})}{\sqrt{k}}$$

where Z_ℓ is a continuous centered Gaussian process (its covariance expression is known via ℓ), α is a function that tends to 0 at infinity, and M is a continuous function.

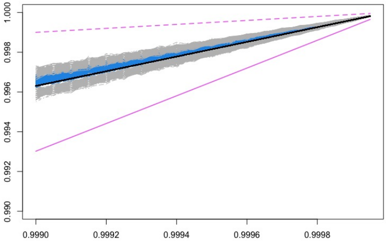
Consequently, one can define $\hat{\theta}_I = \hat{\ell}_{k,n}(\mathbf{1}_I, \mathbf{0}_{I^c})$ and derive confidence interval for q^{θ_I} using previous asymptotic normality and delta method.

- **Huang (1992) Statistics of bivariate extreme values**
- **Fougères, de Haan, Mercadier (2015)** *Bias correction in multivariate extremes*

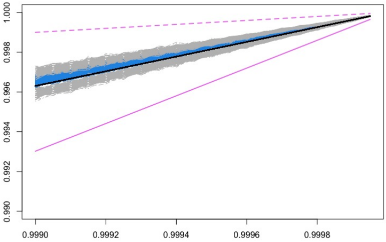
Well-known measures for tail dependence (continuation)... at work !

Estimation of $p^{\theta_{\{1,\dots,7\}}}$, for 2 disctincts models, p given on $x-$axis.

Values : $n = 1000$, $b = 0.2 \times n$, $k = n \times 5/100$, $k_b = b \times 5/100$, $|I| = d = 7$
In black, the true value.
In blue, the estimations $p^{\hat{\ell}_{k,n}(\mathbf{1})}$.
In gray, the confidence intervals obtained by **subsampling**.
In purple, complete tail independence/complete tail dependence (false assumptions).



$F = MEVD$+margins transformation

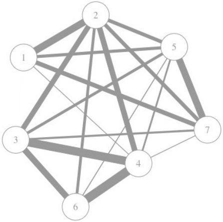


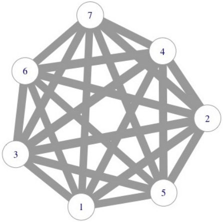
$F \in DA(MEVD)$

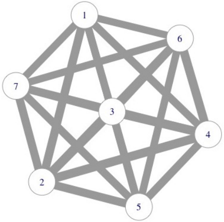
1. Motivation with recent forest fires
2. Short introduction to multivariate extreme value theory
3. The tail dependograph
4. More... and bounds for the tail dependograph
5. Links between the Hoeffding-Sobol and the Möbius decompositions

Construction of a graph of tail dependence

Consider again the MEVD G already used in the simulation.
Below, representation of theoretical values as a graph.







$$i \leftrightarrow j \qquad i \leftrightarrow j \qquad i \leftrightarrow j$$

$$2 - \theta_{\{ij\}} \qquad \sum_{I, I \supseteq \{ij\}} |I| - \theta_I \qquad \sum_{I, I \supseteq \{ij\}} (|I| - \theta_I)/|I|$$

Extremal coefficients are not what we need!

The Hoeffding–Sobol decomposition

The Hoeffding–Sobol decomposition of a **square integrable** function $f : [0,1]^d \to \mathbb{R}$ is an expansion with terms of increasing complexity

$$f(u_1, \ldots, u_d) = f_\emptyset + \sum_{i=1}^{d} f_i(u_i) + \sum_{i<j} f_{ij}(u_i, u_j) + \cdots + f_{1\ldots d}(u_1, \ldots, u_d)$$

where, for each $I \subseteq \{1, \ldots, d\}$, the sub-function

- f_I is defined on $[0,1]^{|I|}$
- f_I is centered $\int f_I \otimes_{i \in I} d\mu_i = 0$
- f_I and $f_{I'}$ are orthogonal as soon as $I \neq I'$

Main references

- **Hoeffding (1948)** *A class of statistics with asymptotically normal distribution. Ann. Math. Statist., 19, 293–325.*
- **Efron and Stein (1981)** *The Jackknife Estimate of Variance. Ann. Statist., 9 :3, 586–596.*
- **M. Sobol (1993)** *Sensitivity estimates for nonlinear mathematical models. Math. Modeling Comput. Exp., 1, 407–414.*
- **A. W. van der Vaart (1998)** *Asymptotic Statistics. Cambridge Univ. Press.*
- **G. Chastaing, F. Gamboa & C. Prieur (2012)** *Generalized Hoeffding–Sobol decomposition for dependent variables : Application to sensitivity analysis. Electron. J. Statist., 6, 2420–2448.*

Stochastic view of the Hoeffding–Sobol decomposition

Let $\mathbf{U} = (U_1, \dots, U_d)$ be a random vector with

- support $[0, 1]^d$
- **independent margins**

Its distribution is denoted by $\mu = \mu_1 \otimes \cdots \otimes \mu_d$.

Existence by recursive construction

- $f_\emptyset = \mathbb{E}[f(\mathbf{U})]$ the mean of the response

  the expectation refers to $d\mu$

- $f_i(u_i) = \mathbb{E}[f(\mathbf{U})|U_i = u_i] - f_\emptyset$ the main effect of component $\{i\}$

  the expectation is with respect to $\otimes_{k \neq i} d\mu_k$

Stochastic view of the Hoeffding–Sobol decomposition (continuation)

- $f_{ij}(u_i, u_j) = \mathbb{E}[f(\mathbf{U})|U_i = u_i, U_j = u_j] - f_i(u_i) - f_j(u_j) - f_\emptyset$ the second-order interaction from the pair of components $\{i, j\}$

  the expectation is associated to $\otimes_{k \neq i,j} d\mu_k$ *;*

- and so on for the third-order, and higher-order interactions...

So that, the terms in the Hoeffding–Sobol decomposition of f are then

$$f_I(\mathbf{u}_I) = \mathbb{E}[f(\mathbf{U})|\mathbf{U}_I = \mathbf{u}_I] - \sum_{I' \subsetneq I} f_{I'}(\mathbf{u}_{I'})$$

where $\mathbf{u}_I$ concatenates the components of $\mathbf{u}$ whose indices are in I.

Stochastic view of the Hoeffding–Sobol decomposition (continuation)

Efron and Stein (1981)

The decomposition of the variable $f(U_1, \ldots, U_d)$ is such that

$f(U_1, \ldots, U_d) = f_\emptyset + \sum_i f_i(U_i) + \sum_{i<j} f_i(U_i, U_j) + \ldots + f_{1,\ldots,d}(U_1, \ldots, U_d)$

where all the $2^d - 1$ variables on the right side above

- *have mean zero and*
- *are mutually uncorrelated.*

The decomposition is unique in the sense that once given the following property "For non-empty subset of indices $I \subseteq \{1, \ldots, d\}$, one has in addition

$$\mathbb{E}[f_I(\mathbf{U}_I)|\mathbf{U}_{I'}] = 0 \text{ when } I' \subsetneq I. \quad "$$

then the terms constant, of first effects, and so on, must be given by the recursive construction.

Functional decomposition of the variance

The Hoeffding–Sobol decomposition is orthogonal and corresponds to a variance decomposition, viz.

$$\mathrm{var}[f(\mathbf{U})] = \mathbb{E}[\{f(\mathbf{U}) - \ell_\emptyset\}^2] = \sum_{I\subseteq\{1,\dots,d\}} \mathbb{E}[f_I^2(\mathbf{U}_I)] = \sum_{I\subseteq\{1,\dots,d\}} \mathrm{var}[f_I(\mathbf{U}_I)]$$

This can be written in abbreviated form as

$$D(f) = \sum_{I\subseteq\{1,\dots,d\}} D_I(f)$$

where

- $D(f)$ denotes the global variance
- $D_I(f)$ is the variance associated to the term f_I

Superset combination of variances

One combination of such variances is of prime interest.

Superset importance coefficients

The superset importance coefficients associated with subset I is the sum of all variances of terms that contain I

$$\Upsilon_I^2(f) = \sum_{J \supseteq I} D_J(f)$$

The focus will be on pairwise quantities. For the pair $\{i, j\}$ with $i \neq j$, set

$$\Upsilon_{\{i,j\}}^2(f) = \sum_{J \supseteq \{i,j\}} D_J(f)$$

The latter has the integral representation

$$\frac{1}{4} \int_{[0,1]^{d+2}} \left\{ f(\mathbf{x}) - T_i[f](\mathbf{x}, u) - T_j[f](\mathbf{x}, v) + T_{i,j}[f](\mathbf{x}, u, v) \right\}^2 d\mu(\mathbf{x}) d\mu_i(u) d\mu_j(v),$$

where $T_i[f](\mathbf{x}, u) = f(\mathbf{x} + (u - x_i)\mathbf{e}_i)$ and $T_{i,j}[f](\mathbf{x}, u, v) = f(\mathbf{x} + (u - x_i)\mathbf{e}_i + (v - x_j)\mathbf{e}_j)$

One can refer to

- **R. Liu and A. B. Owen (2006)** *Estimating mean dimensionality of analysis of variance decompositions. J. Amer. Statist. Assoc., 101, 712–721.*

- **J. Fruth, O. Roustant, and S. Kuhnt (2014)** *Total interaction index : A variance-based sensitivity index for interaction screening. J. Statist. Plann. Inf., 147, 212–223.*

The FANOVA graph

It is a valued graph derived from the superset importance coefficients, where

- *the vertices are the input variables ;*
- *an edge exists iff the superset importance coefficient is strictly positive ;*
- *the edge's weight is the value of the superset importance coefficient.*

- **J. Fruth (2015).** *Sensitivity analysis and graph-based methods for black-box functions with an application to sheet metal forming. (PhD Thesis).*

Definition of the tail dependograph (Mercadier and Roustant, 2019)

The existence of a domain of attraction G is equivalent to the existence of a non degenerate limit $\ell = \ell_{(X_1,\dots,X_d)}$ satisfying

$$\ell(\mathbf{x}) = \lim_{\lambda\to\infty} \lambda \mathbb{P}\left\{1 - F_1(X_1) \le x_1/\lambda \text{ or } \dots \text{ or } 1 - F_d(X_d) \le x_d/\lambda\right\}$$

equivalently written as

$$\ell(\mathbf{x}) = \lim_{\lambda\to\infty} \lambda \left\{1 - F[F_1^{-1}(1 - x_1/\lambda), \dots, F_d^{-1}(1 - x_d/\lambda)]\right\}$$

The tail dependograph

The tail dependograph is a FANOVA graph applied to the stable tail dependence function

$$f = \ell_{(X_1,\dots,X_d)}$$

Fact 1 The role of the tail dependograph is to represent the structure of the tail dependence between variables with **non-oriented graphs of asymptotic dependence.**

Fact 2 Tail independence from the tail dependograph is **concordant** with the intuitive meaning : Let A, B, C a partition of $\{1, \ldots, d\}$, then $\mathbf{X}_A$ and $\mathbf{X}_B$ are asymptotically independent if

$$\ell(\mathbf{u}) = \ell(\mathbf{u}_A, \mathbf{0}, \mathbf{u}_C) + \ell(\mathbf{0}, \mathbf{u}_B, \mathbf{u}_C) \quad \forall \mathbf{u} = (\mathbf{u}_A, \mathbf{u}_B, \mathbf{u}_C) \; .$$

No edge of the tail dependograph goes from any vertex in A to any vertex in B.

Fact 3 Non parametric inference leads to **ranked-based estimators** of the tail superset indices.

Fact 4 Asymptotic normality under standard second order condition.

Assume first that $\ell(\mathbf{u}) = \ell(\mathbf{u}_A, \mathbf{0}_B) + \ell(\mathbf{0}_A, \mathbf{u}_B)$.
Set $f(\mathbf{u}_A) = \ell(\mathbf{u}_A, \mathbf{0}_B)$ and $g(\mathbf{u}_B) = \ell(\mathbf{0}_A, \mathbf{u}_B)$.
The functions f and g admit the Hoeffing–Sobol decomposition $f(\mathbf{u}_A) = \sum_{I \subseteq A} f_I(\mathbf{u}_I)$ and $g(\mathbf{u}_B) = \sum_{I \subseteq B} g_I(\mathbf{u}_I)$ so that, from the uniqueness of the decomposition,

$$\ell(\mathbf{u}) = \sum_{I \subseteq \{1,\dots,d\}} \ell_I(\mathbf{u}_I)$$

with $\ell_I = f_I$ if $I \subseteq A$, $\ell_I = g_I$ if $I \subseteq B$ and $\ell_I \equiv 0$ if $I \cap A \neq \emptyset$ and $I \cap B \neq \emptyset$. It yields

$$\Upsilon^2_{\{i,j\}}(\ell) = \begin{cases} \Upsilon^2_{\{i,j\}}(f) & \text{if } \{i,j\} \subseteq A \\ \Upsilon^2_{\{i,j\}}(g) & \text{if } \{i,j\} \subseteq B \\ 0 & \text{if } (i,j) \in A \times B \text{ or } (i,j) \in B \times A\,. \end{cases}$$

No edge goes from any vertex in A to any vertex in B.
Now, assume that $\Upsilon^2_{\{i,j\}}(\ell) = 0$ for all $i \in A$ and $j \in B$. Then, since it is a sum of positive terms, all terms vanish : $D_K = 0$ and $\ell_K \equiv 0$ $\forall K$ that contains $\{i,j\}$. Thus $\ell(\mathbf{u}) = f(\mathbf{u}_A) + g(\mathbf{u}_B)$. In particular $\ell(\mathbf{u}_A, \mathbf{0}_B) = f(\mathbf{u}_A) + g(\mathbf{0}_B)$ and $\ell(\mathbf{0}_A, \mathbf{u}_B) = f(\mathbf{0}_A) + g(\mathbf{u}_B)$. By adding these terms, we obtain

$$\ell(\mathbf{u}_A, \mathbf{0}_B) + \ell(\mathbf{0}_A, \mathbf{u}_B) = f(\mathbf{u}_A) + g(\mathbf{0}_B) + f(\mathbf{0}_A) + g(\mathbf{u}_B) = \ell(\mathbf{u}) + \ell(\mathbf{0}) = \ell(\mathbf{u})\,.$$

Properties of the tail dependograph (check Fact 3) (Mercadier and Roustant, 2019)

Tensor-product function Lemma

If $f(\mathbf{u}) = \prod_{t=1}^{d} f_t(u_t)$ *then* $\Upsilon_I^2(f) = \prod_{t\in I} var(f_t(U_t)) \prod_{t\notin I} \mathbb{E}[f_t^2(U_t)]$.

Expanding the product $f(\mathbf{u}) = \prod_{t=1}^{d}\{(f_t(u_t) - m_t) + m_t\} = \sum_{I\subseteq\{1,\ldots,d\}} f_I(\mathbf{u}_I)$ with

$$f_I(\mathbf{u}_I) = \prod_{t\in I}\{f_t(u_t) - m_t\} \prod_{t\notin I} m_t$$

conclusion follows $\mathbb{E}[f_I(\mathbf{u}_I)|\mathbf{u}_J] = 0 \ \forall J \subsetneq I$ + unicity of the decomposition.

$$\begin{aligned} f_I^{\mathrm{super}}(\mathbf{u}) &= \prod_{t\in I}(f_t(u_t) - m_t)\left\{\sum_{J\supseteq I}\prod_{t\in J\setminus I}(f_t(u_t) - m_t)\prod_{t\notin J} m_t\right\} \\ &= \prod_{t\in I}(f_t(u_t) - m_t)\left\{\sum_{K\subseteq\{1,\ldots,d\}\setminus I}\prod_{t\in K}(f_t(u_t) - m_t)\prod_{t\notin K} m_t\right\} \\ &= \prod_{t\in I}(f_t(u_t) - m_t)\prod_{t\notin I} f_t(u_t). \end{aligned}$$

$$\Upsilon_I^2(f) = \mathbb{E}[f_I^{\mathrm{super}}(\mathbf{U})^2]$$

Properties of the tail dependograph (check Fact 3) (Mercadier and Roustant, 2019)

Let $k = k(n)$ be an **intermediate sequence**. Recall that

$$\hat{\ell}_{k,n}(\mathbf{x}) = \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}_{\left\{X_{\mathbf{1}}^{(i)} \geq X_{\mathbf{1},n-[kx_{\mathbf{1}}]+\mathbf{1},n}^{(i)} \text{ or } \dots \text{ or } X_d^{(i)} \geq X_{d,n-[kx_d]+\mathbf{1},n}^{(i)}\right\}}$$

$$= \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}_{\left\{x_{\mathbf{1}} \geq \tilde{R}_{\mathbf{1}}^{(i)} \text{ or } \dots \text{ or } x_d \geq \tilde{R}_d^{(i)}\right\}}$$

in terms of $\tilde{R}_t^{(i)} := (n - R_t^{(i)} + 1)/k$ where $R_t^{(i)}$ is the rank of $X_t^{(i)}$ among $X_t^{(1)}, \dots, X_t^{(n)}$. One can also write

$$\hat{\ell}_{k,n}(\mathbf{x}) = \frac{1}{k}\sum_{i=1}^{n}\left(1 - \prod_{t=1}^{d} \mathbf{1}_{\{x_t < \tilde{R}_t^{(i)}\}}\right)$$

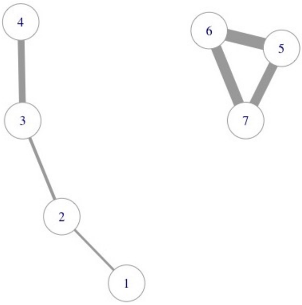
$+$ Tensor-product lemma $\Rightarrow$ The pairwise tsic have **rank-based expressions**

$$\Upsilon_{i,j}^2(\hat{\ell}_{k,n}) = \frac{1}{k^2}\sum_{s=1}^{n}\sum_{s'=1}^{n}\left\{\prod_{t\in\{i,j\}}\left(\bar{R}_t^{(s)} \wedge \bar{R}_t^{(s')} - \bar{R}_t^{(s)}\bar{R}_t^{(s')}\right)\prod_{t\notin\{i,j\}} \bar{R}_t^{(s)} \wedge \bar{R}_t^{(s')}\right\}$$

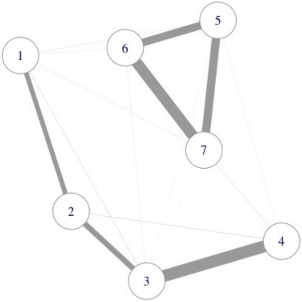
Properties of the tail dependograph (check Fact 1) (Mercadier and Roustant, 2019)

Rank-based estimate of the tail dependograph at work!
$\mathbf{X} = (X_1, \ldots, X_7) \sim F$
such that its distribution F is in the **domain of attraction of** G with

$$\ell(x_1, \ldots, x_7) = \ell_{12}(x_1, x_2) + \ell_{23}(x_2, x_3) + \ell_{34}(x_3, x_4) + \ell_{567}(x_5, x_6, x_7)$$



Theoretical



Empirical with $n = 1000$

1. Motivation with recent forest fires

2. Short introduction to multivariate extreme value theory

3. The tail dependograph

4. More... and bounds for the tail dependograph

5. Links between the Hoeffding-Sobol and the Möbius decompositions

Choquet representation of the HS decomposition of stdf

The Sobol effects and associated variances of stable tail dependence function are expressed on $\mathbb{S}_\vee^+ \times [0,1]$ *as* **integrals of rank-one functions**.

Set $C = \{\mathbf{w} \in [0,1]^d, \max \mathbf{w} = 1\}$. Change to the $L^\infty-$norm allows to write

$$\ell(\mathbf{x}) = \ell(\mathbf{1}) \int_C \max(\mathbf{x} \cdot \mathbf{w}) d\nu(\mathbf{w}) = \ell(\mathbf{1}) \int_C \int_0^1 ds \mathbf{1}_{s<\max(\mathbf{x}\cdot\mathbf{w})} d\nu(\mathbf{w})$$

$$= \ell(\mathbf{1}) \int_C d\nu(\mathbf{w}) \int_0^1 ds(1 - \mathbf{1}_{s\geq \mathbf{x}\cdot\mathbf{w}}) = \ell(\mathbf{1}) - \ell(\mathbf{1}) \int_C d\nu(\mathbf{w}) \int_0^1 ds \mathbf{1}_{s\geq \mathbf{x}\cdot\mathbf{w}}$$

The equality

$$\sum_{v \subseteq u} \ell_v(\mathbf{x}) = \int_{[0,1]^{-u}} \ell(\mathbf{x}) d\lambda_{-u}(\mathbf{x})$$

combined with the Fubini-Tonelli theorem yields and after simplification

$$\Upsilon_u^2 = \ell(\mathbf{1})^2 \int_C d\nu(\mathbf{w}) \int_C d\nu(\mathbf{v}) \int_0^1 ds \int_0^1 dt \prod_{i \notin u} K_i(s,t) \prod_{i \in u} (K_i(s,t) - k_i(s)k_i(t)) \, .$$

Recall that

$$\Upsilon_u^2 = 2^{-|u|} \int_{[0,1]^{d+|u|}} (D_{\mathbf{z}^u}^{\mathbf{x}^u} \ell(\cdot, \mathbf{z}^{-u}))^2 d\mathbf{x}^u d\mathbf{z}$$

Set

$$AME := \frac{1}{n} \sum_{i=1}^{n} |\hat{\Upsilon_u^2}(\ell) - \Upsilon_u^2(\ell)|$$

The AME for the estimation of $\Upsilon_u^2(\ell)$ when $\ell(\mathbf{x}) = \max(\mathbf{x})$:

	$d=5$		$d=10$	
	$u=\{1,2\}$	$u=\{1,\dots,d\}$	$u=\{1,2\}$	$u=\{1,\dots,d\}$
$N=1000$	12.15×10^{-5}	79.41×10^{-5}	31.18×10^{-6}	159×10^{-10}
$N=1000$	7.15×10^{-5}	0.40×10^{-5}	14.71×10^{-6}	24.39×10^{-10}
$N=10,000$	5.43×10^{-5}	0.26×10^{-5}	12.07×10^{-6}	–
$N=10,000$	1.99×10^{-5}	0.11×10^{-5}	3.68×10^{-6}	7.53×10^{-10}

Upper bound for the superset importance coefficients of stdf

Let ℓ be a d-variate stable tail dependence function.
Then,

$$\Upsilon_I^2(\ell) \leq \Upsilon_I^2(\ell_{\vee,I}) \leftarrow \textit{known}$$

for any non-empty $I \subseteq \{1,\dots,d\}$ where $\ell_{\vee,I}(\mathbf{x}_I) = \max_{i\in I} x_i$.

If ℓ is a d-variate stdf with equality $\Upsilon_u^2(\ell) = \Upsilon_u^2(\ell_{\vee,u})$ for a given $\emptyset \neq u \subseteq \{1,\dots,d\}$, then its projection on the variables $\mathbf{x}^u$ is equal to

$$\ell(\mathbf{x}^u, \mathbf{0}^{-u}) = \ell^{\vee,u}(\mathbf{x}^u) = \max_{i\in u} x_i$$

Proof based on

- a generalization of Hölder's inequality
- tricky justification from multivariate monotonicity

Example with $\ell_\vee(\mathbf{x}) = \max_{i=1,\dots,d} x_i$

$$\text{Sobol effect} \quad \ell_{\vee,\emptyset} = \frac{d}{d+1}$$

$$\text{Sobol effects} \quad \ell_{\vee,I}(\mathbf{x}) = -\int_0^1 \prod_{i\in I} (\mathbf{1}_{s\geq x_i} - s) s^{d-|I|} ds$$

$$\text{Global variance of } \ell_\vee \quad \sigma^2 = \frac{d}{(d+1)^2(d+2)}$$

$$\text{Individual variance of } \ell_\vee \quad \sigma_I^2 = \frac{2(2d-|I|+1)!|I|!}{(d+1)(2d+2)!}$$

$$\text{Superset coefficients of } \ell_\vee \quad \Upsilon_I^2 = \frac{2d!|I|!}{(d+|I|+2)!}$$

$$\text{Superset pairwise coefficients of } \ell_\vee \quad \Upsilon_{\{i,j\}}^2 = \frac{4}{(d+1)(d+2)(d+3)(d+4)}$$

1. Motivation with recent forest fires
2. Short introduction to multivariate extreme value theory
3. The tail dependograph
4. More... and bounds for the tail dependograph
5. Links between the Hoeffding-Sobol and the Möbius decompositions

Projections

Let T be a random variable and $\{Z, Z \in \mathcal{F}\}$ a set of random variables, all defined on the same probability space, with a second order moment.

Definition

The projection $P_{\mathcal{F}}T$ of T on $\mathcal{F}$ is the argmin on $\mathcal{F}$ of $\mathbb{E}[(T-Z)^2]$.

Theorem

Let $\mathcal{F}$ a linear space of squared integrable variables.
$P_{\mathcal{F}}T$ is the projection of T on $\mathcal{F}$ if and only if

- $P_{\mathcal{F}}T \in \mathcal{F}$
- *and its error $T - P_{\mathcal{F}}T$ is orthogonal to $\mathcal{F}$: $\mathbb{E}[(T - P_{\mathcal{F}}T)Z] = 0$ for all $Z \in \mathcal{F}$.*

Furthermore, $P_{\mathcal{F}}T$ is almost surely unique.

Projections: examples

- $\mathcal{F} = [1]$
  leads to the mathematical expectation $\mathbb{E}[T]$

- $\mathcal{F} = [g_1(X_1), g_1 \text{ measurable and } \mathbb{E}[g_1(X_1)^2] < \infty]$
  ... the conditional expectation $\mathbb{E}[T|X_1]$

- $\mathcal{F} = [g_1(X_1), \ldots, g_d(X_d), g_i \text{ measurable and } \mathbb{E}[g_i(X_i)^2] < \infty]$
  with independent X_i
  leads to $\sum_{i=1}^d \mathbb{E}[T|X_i] - (d-1)\mathbb{E}[T]$ also called *Hájek Projection*

- $\mathcal{F} = [g_A(X_i, i \in A) \forall A \subseteq \{1, \ldots, d\}, \text{ s. t. } g_A \text{ measurables and } \mathbb{E}[g_A(X_A)^2] < \infty]$
  with independent X_i
  one obtains the *Hoeffding decomposition*; in particuliar the part associated to the subset A is given by

$$\sum_{B \subseteq A} (-1)^{|A|-|B|} \mathbb{E}[T|(X_i, i \in B)]$$

➡ **Van der Vaart (1996)** *Comments at the end of Chapter 11 Projections*

The Möbius decomposition

- **Deheuvels (1981)** *An Asymptotic Decomposition for Multivariate Distribution-Free Tests of Independence.*

focuses on independence test

$$(H_0) C(u_1, \ldots, u_d) = u_1 \ldots u_d \qquad \text{against} \qquad (H_1) C(u_1, \ldots, u_d) \neq u_1 \ldots u_d$$

From Theorem 2 (Deheuvels, 1981)

Let C_n stand for the empirical copula estimate of C. Set

$$\Delta_{n,ij} = C_n(u_{\{ij\}}, \mathbf{1}_{-\{ij\}}) - u_i u_j$$

and

$$\Delta_{n,A} = C_n(u_{\{A\}}, \mathbf{1}_{-A}) - \sum_{i \in A} u_i \Delta_{n,A \setminus \{i\}}$$

Then the empirical processes $\{\sqrt{n}\Delta_{n,A}\}_{A,|A|\geq 2}$ converge jointly under (H_0) to independent centered Gaussian processes whose covariances are characterized.

The Möbius decomposition (continuation)

Let $v \subseteq \{1, \ldots, d\}$ and $\mathbf{x} = (x_1, \ldots, x_d) \in [0,1]^d$.
Set $\mathbf{x}_{v,1} = (t_1, \ldots, t_d)$ where $t_i = x_i$ if $i \in v$ and $t_i = 1$ otherwise.

The **Möbius decomposition** of a copula C can be written as

$$C(\mathbf{x}) = \Pi(\mathbf{x}) + \sum_{u \subseteq \{1,\ldots,d\}, |u| \geq 2} \mathcal{M}_u(C)(\mathbf{x}) \times \prod_{k \notin u} x_k,$$

with

$$\mathcal{M}_u(C)(\mathbf{x}) = \sum_{v \subseteq u} (-1)^{|u \setminus v|} C(\mathbf{x}_{v,1}) \times \prod_{k \in u \setminus v} x_k.$$

Note that $\mathcal{M}_u(C) = 1$ whenever $|u| = 0$ and $\mathcal{M}_u(C) = 0$ if $|u| = 1$.

The first formula can also be written as

$$C(\mathbf{x}) = \sum_{u \subseteq \{1,\ldots,d\}} C_u(\mathbf{x}) \ .$$

The Möbius decomposition (continuation)

- **Ghoudi, Kulperger et Rémillard (2001)**
  A Nonparametric Test of Serial Independence for Time Series and Residuals.
  Proposition 2.1 characterizes $C = \Pi$ with $\mathcal{M}_u(C) = 0$ for any $u \neq \emptyset$.

- **Genest and Rémillard (2004)**
  Test of independence and randomness based on the empirical copula process.
  Rewriting and generalization of Deheuvels (1981) to serial dependence.

- **Kojadinovic and Holmes (2009)**
  Tests of independence among continuous random vectors based on Cramér–von Mises functionals of the empirical copula process.
  Generalization to independence by blocks.

...

- **Kuo et al. (2010)** *On Decompositions of multivariate functions.*
  Focus is done on decompositions of the form

$$f = \sum_{u \subseteq \{1,\dots,d\}} f_u$$

where f_u only depends on the subset of variables $\{x_j : j \in u\}$.

Let $(P_j)_{j=1,\dots,d}$ be a commuting family of linear idempotent maps acting on $\mathcal{F}$ the linear space of functions f defined on $[0,1]^d$.

Assumptions of Kuo et al. (2010) :
For any $f \in \mathcal{F}$ and $j \in \{1,\dots,d\}$, one assumes that

- $P_j(f) \in \mathcal{F}_{-j}$ subspace of $\mathcal{F}$ with functions not depending on x_j
- and if $f \in \mathcal{F}_{-j} \Rightarrow P_j(f) = f$

...

Theorem 2.1 (Kuo et al., 2010)

For any function $f \in \mathcal{F}$ and any subset $u \subseteq \{1, \dots, d\}$, let define

$$f_u = \left(\prod_{j \in u} (I - P_j) \right) P_{\{1,\dots,d\} \setminus u}(f) \, .$$

Then

$$f = \sum_{u \subseteq \{1,\dots,d\}} f_u$$

and f_u only depends on variables which component indices are in u.

Moreover, this decomposition has three equivalent expressions.
As above, as well as following :

- $g_0 = P_{\{1,\dots,d\}}(f)$ *and* $g_u = P_{\{1,\dots,d\} \setminus u}(f) - \sum_{v \subsetneq u} g_v$ *then* $f_u = g_u$
- $h_u = \sum_{v \subseteq u} (-1)^{|u|-|v|} P_{\{1,\dots,d\} \setminus v}(f)$ *then* $f_u = h_u$

. . .

Which maps for the decomposition under study

Hoeffding-Sobol is a particular case of Kuo et al. (2010) for

$$P_j(f)(\mathbf{x}) = \int_0^1 f(x_1, \ldots, x_{j-1}, z, x_{j+1}, \ldots, x_d)dz$$

whereas in the context of **Möbius**, one has

$$P_j(f)(\mathbf{x}) = f(\mathbf{x}_{\{j\}}, \mathbf{1}_{-\{j\}}) \times f(\mathbf{x}_{-\{j\}}, \mathbf{1}_{\{j\}})$$

that does not verify the assumptions of Kuo et al. (2010).

Good news! (Mercadier, Roustant and Genest, 2022)

Both are particular cases of a common context

Proposition 1

Allowing any term to be dependent of any variable, the generalization is obtained

- *without the linear assumption on P_j*

and

Proposition 2

Allowing any term to be dependent of any variable, the generalization is obtained

- *without the linear assumption on P_j*
- *without the idempotent assumption on P_j*

However, the maps P_j still need to **commute**.

- Hoeffding-Sobol and Möbius both decompose a function of several variables as a sum of terms with increasing complexity.
- Both are obtained as **particular cases** of a generalization of Kuo et al. (2010).
- In the Hoeffding-Sobol decomposition, the map P_u **neutralizes the effect** of the variables in u, average with respect to any of them.
- In the Möbius decomposition, the map P_u **erases the stochastic dependence** between variables that are in u from those out from u.
- An advantage of Hoeffding-Sobol is to lead to **orthogonal terms**, allowing the decomposition of the variance.
- Möbius takes another advantage. Its terms are very **easy to compute**, as combination of several evaluations.

1) General Conclusion

2) Thank you for your attention!