

Partitionnement des variables de séries temporelles multivariées selon la dépendance de leurs extrêmes

Soutenance de thèse

Alexis Boulin^{1,2}

¹Université Côte d'Azur, CNRS, LJAD, Nice, France.

²Inria, Lemon.

30 septembre 2024

Membres du jury

Examinateurs : Stéphan Cléménçon

Johan Segers

Carlo Gaetan

Armelle Guillou

Philippe Naveau

Vincent Vandewalle

Encadrants : Thomas Laloë

Elena Di Bernardino

Gwladys Toulemonde



- Élaborer des outils pour la réduction de dimension en utilisant le partitionnement des variables.

- Élaborer des outils pour la réduction de dimension en utilisant le partitionnement des variables.
- Obtenir des garanties statistiques solides pour les méthodes d'estimation proposées.

- Élaborer des outils pour la réduction de dimension en utilisant le partitionnement des variables.
- Obtenir des garanties statistiques solides pour les méthodes d'estimation proposées.
- Appliquer les méthodes proposées sur des bases de données dans le domaine des sciences climatiques.

Dans l'étude classique de la statistique, deux théorèmes prévalent :

Dans l'étude classique de la statistique, deux théorèmes prévalent :

Loi forte des grands nombres

Considérons X_1, \dots, X_n comme un échantillon indépendant et sidentiquement distribués selon une variable aléatoire X intégrable. Alors,

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}[X].$$

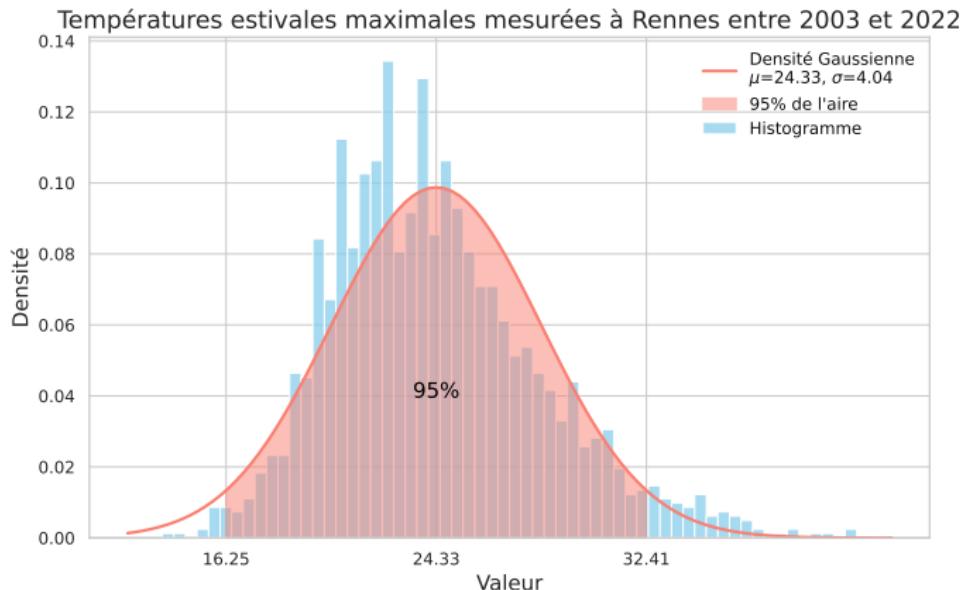
Dans l'étude classique de la statistique, deux théorèmes prévalent :

Théorème de la limite centrale

Toujours avec un échantillon X_1, \dots, X_n indépendant et identiquement distribué selon une variable aléatoire X de carré intégrable, nous obtenons

$$(a_n)^{-1} \left(\sum_{i=1}^n X_i - b_n \right) \xrightarrow[n \rightarrow \infty]{d} Z,$$

où Z est une variable aléatoire gaussienne centrée et réduite, $b_n = n\mathbb{E}[X]$ et $a_n = \sqrt{n/\text{Var}(X_1)}$.



La théorie des valeurs extrêmes,
c'est quoi



- **Théorie :** Sous des hypothèses minimales, la distribution des maxima ou des excès de variables aléatoires converge vers une certaine classe de distributions.

- **Théorie** : Sous des hypothèses minimales, la distribution des maxima ou des excès de variables aléatoires converge vers une certaine classe de distributions.
- **Modélisation** : Utiliser ces distributions limites pour modéliser les maxima ou les excès de variables aléatoires.

- Soit $\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(d)})$ pour $i = 1, \dots, n$, des vecteurs aléatoires i.i.d. en ayant pour fonction de répartition $F_{\mathbf{X}}$.

- Soit $\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(d)})$ pour $i = 1, \dots, n$, des vecteurs aléatoires i.i.d. en ayant pour fonction de répartition $F_{\mathbf{X}}$.
- Les maxima composante par composante sont définis par :

$$\bigvee_{i=1}^n \mathbf{X}_i = \left(\bigvee_{i=1}^n X_i^{(1)}, \dots, \bigvee_{i=1}^n X_i^{(d)} \right).$$

- Soit $\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(d)})$ pour $i = 1, \dots, n$, des vecteurs aléatoires i.i.d. en ayant pour fonction de répartition $F_{\mathbf{X}}$.
- Les maxima composante par composante sont définis par :

$$\bigvee_{i=1}^n \mathbf{X}_i = \left(\bigvee_{i=1}^n X_i^{(1)}, \dots, \bigvee_{i=1}^n X_i^{(d)} \right).$$

- Nous obtenons :

$$\mathbb{P} \left\{ \bigvee_{i=1}^n \frac{\mathbf{X}_i - \mathbf{b}_n}{\mathbf{a}_n} \leq \mathbf{x} \right\} = F_{\mathbf{X}}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n),$$

pour certains vecteurs $\mathbf{a}_n > 0$ et \mathbf{b}_n .

- Soit $\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(d)})$ pour $i = 1, \dots, n$, des vecteurs aléatoires i.i.d. en ayant pour fonction de répartition $F_{\mathbf{X}}$.
- Les maxima composante par composante sont définis par :

$$\bigvee_{i=1}^n \mathbf{X}_i = \left(\bigvee_{i=1}^n X_i^{(1)}, \dots, \bigvee_{i=1}^n X_i^{(d)} \right).$$

- Nous obtenons :

$$\mathbb{P} \left\{ \bigvee_{i=1}^n \frac{\mathbf{X}_i - \mathbf{b}_n}{\mathbf{a}_n} \leq \mathbf{x} \right\} = F_{\mathbf{X}}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n),$$

pour certains vecteurs $\mathbf{a}_n > 0$ et \mathbf{b}_n .

- Les distributions des valeurs extrêmes, notées $H_{\mathbf{X}}$, correspondent aux limites de cette dernière distribution,

$$F_{\mathbf{X}}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) \xrightarrow{n \rightarrow \infty} H_{\mathbf{X}}(\mathbf{x}). \quad (\text{MDA})$$

- Soit $\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(d)})$ pour $i = 1, \dots, n$, des vecteurs aléatoires i.i.d. en ayant pour fonction de répartition $F_{\mathbf{X}}$.
- Les maxima composante par composante sont définis par :

$$\bigvee_{i=1}^n \mathbf{X}_i = \left(\bigvee_{i=1}^n X_i^{(1)}, \dots, \bigvee_{i=1}^n X_i^{(d)} \right).$$

- Nous obtenons :

$$\mathbb{P} \left\{ \bigvee_{i=1}^n \frac{\mathbf{X}_i - \mathbf{b}_n}{\mathbf{a}_n} \leq \mathbf{x} \right\} = F_{\mathbf{X}}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n),$$

pour certains vecteurs $\mathbf{a}_n > 0$ et \mathbf{b}_n .

- Les distributions des valeurs extrêmes, notées $H_{\mathbf{X}}$, correspondent aux limites de cette dernière distribution,

$$F_{\mathbf{X}}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) \xrightarrow{n \rightarrow \infty} H_{\mathbf{X}}(\mathbf{x}). \quad (\text{MDA})$$

- Les seules limites possibles sont les loi max-stables,
 $H_{\mathbf{X}}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) = H_{\mathbf{X}}(\mathbf{x})$ pour tout $n \in \mathbb{N}$.

- Pour simplifier, nous considérons des distributions max-stables avec marges de Fréchet, c'est-à-dire $H_\alpha^{(j)}(x) = e^{-x^{-\alpha}}$ pour $x > 0$, $j = 1, \dots, d$ et $\alpha > 0$.

- Pour simplifier, nous considérons des distributions max-stables avec marges de Fréchet, c'est-à-dire $H_\alpha^{(j)}(x) = e^{-x^{-\alpha}}$ pour $x > 0$, $j = 1, \dots, d$ et $\alpha > 0$.

Représentation des lois max-stables

Soit $H_{\mathbf{X}}$ une fonction de répartition. Les énoncés suivants sont équivalents :

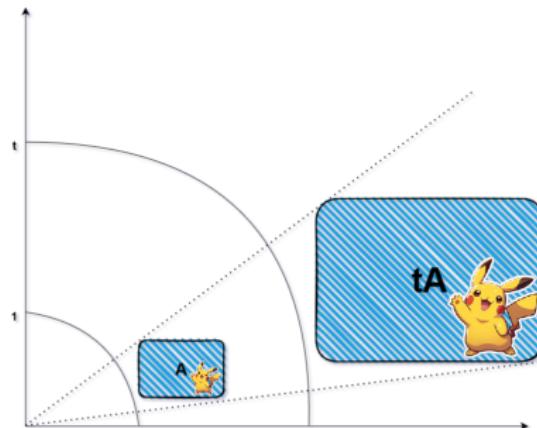
- $H_{\mathbf{X}}$ est une fonction de répartition max-stable avec des marges de Fréchet $H_\alpha^{(j)}$ pour $\alpha > 0$ et $j = 1, \dots, d$.
- Il existe $\Lambda_{\mathbf{X}}$ une mesure borélienne sur $\mathcal{E}_+ = [0, \infty)^d \setminus \{0\}$ satisfaisant $\Lambda_{\mathbf{X}}(\mathcal{E}_+ \setminus [0, \mathbf{x}]) < \infty$.

Cette correspondance est donnée par :

$$H_{\mathbf{X}}(\mathbf{x}) = \begin{cases} e^{-\Lambda_{\mathbf{X}}(\mathcal{E}_+ \setminus [0, \mathbf{x}])), & \mathbf{x} > 0 \\ 0 & \mathbf{x} \leq 0. \end{cases}$$

- La mesure exponentielle Λ_X caractérise la distribution max-stable ainsi que le comportement des valeurs extrêmes de \mathbf{X} .

- La mesure exponentielle $\Lambda_{\mathbf{X}}$ caractérise la distribution max-stable ainsi que le comportement des valeurs extrêmes de \mathbf{X} .
- La mesure exponentielle est homogène $\Lambda_{\mathbf{X}}(tA) = t^{-\alpha} \Lambda_{\mathbf{X}}(A)$ pour un borélien $A \subset \mathbb{R}^d$ qui ne contient pas l'origine.

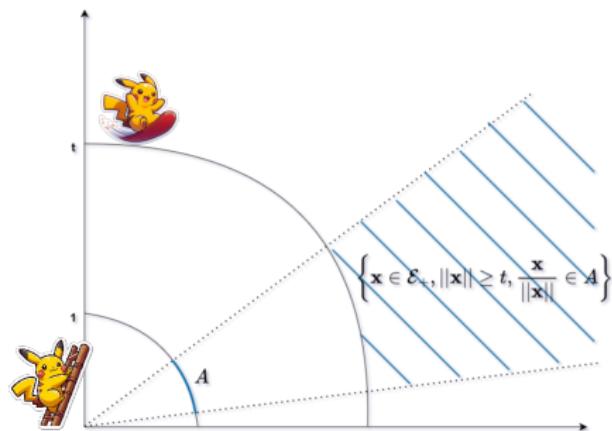


- Soit $\|\cdot\|$ une norme sur \mathbb{R}^d . Pour tout $t \geq 0$ et tout borélien A de $\mathbb{S}_{\mathcal{E}_+} = \{\mathbf{z} \in \mathcal{E}_+ : \|\mathbf{z}\| = 1\} \subset \mathcal{E}_+$.

- Soit $\|\cdot\|$ une norme sur \mathbb{R}^d . Pour tout $t \geq 0$ et tout borélien A de $\mathbb{S}_{\mathcal{E}_+} = \{\mathbf{z} \in \mathcal{E}_+ : \|\mathbf{z}\| = 1\} \subset \mathcal{E}_+$.
- L'homogénéité de la mesure exponentielle donne l'identité suivante :

$$\Lambda_{\mathbf{x}} \left(\left\{ \mathbf{x} \in \mathcal{E}_+ : \|\mathbf{x}\| \geq t, \frac{\mathbf{x}}{\|\mathbf{x}\|} \in A \right\} \right) = t^{-\alpha} \Phi(A),$$

où Φ est la mesure spectrale.



Représentation des lois max-stables

Considérons \mathbf{X} un vecteur aléatoire avec comme distribution $H_{\mathbf{X}}$. Les énoncés suivants sont équivalents :

- ① $H_{\mathbf{X}}$ est une fonction de répartition max-stable avec des marges de Fréchet $H_{\alpha}^{(j)}$ pour $\alpha > 0$ et $j = 1, \dots, d$.
- ② Il existe une mesure finie Φ sur $\mathbb{S}_{\varepsilon_+}$ telle que pour tout $\mathbf{x} \in \mathbb{R}_+^d$

$$H_{\mathbf{X}}(\mathbf{x}) = \exp \left\{ - \int_{\mathbb{S}_{\varepsilon_+}} \left(\bigvee_{j=1}^d \frac{a^{(j)}}{x^{(j)}} \right)^{\alpha} \Phi(d\mathbf{a}) \right\}.$$

Fonction de dépendance caudale

En normalisant toutes les marges par des Fréchet standardisés, c'est à dire $H_1^{(j)}(x) = e^{-x^{-1}}$ pour $x > 0$, nous pouvons définir la fonction de dépendance caudale comme suit :

$$L(z^{(1)}, \dots, z^{(d)}) = \Lambda_{\mathbf{X}}(\mathcal{E}_+ \setminus [0, 1/\mathbf{z}]) = \int_{\mathbb{S}_{\mathcal{E}_+}} \bigvee_{j=1}^d a^{(j)} z^{(j)} \Phi(d\mathbf{a}).$$

Fonction de Pickands

Nous pouvons également définir la fonction de Pickands sur le simplexe unité $\Delta_{d-1} = \{(w^{(1)}, \dots, w^{(d-1)}) \in [0, 1]^{d-1} : \sum_{j=1}^{d-1} w^{(j)} \leq 1\}$ comme suit :

$$\mathcal{A}(w^{(1)}, \dots, w^{(d-1)}) = \int_{\mathbb{S}_{\mathcal{E}_+}} \bigvee_{j=1}^d a^{(j)} w^{(j)} \Phi(d\mathbf{a}),$$

où $w^{(j)} = z^{(j)} / \sum_{j=1}^d z^{(j)}$ avec $\mathbf{z} \in \mathbb{R}_+^d$.

La copule des valeurs extrêmes

La copule des valeurs extrêmes s'écrit :

$$C_{\infty}(\mathbf{u}) = \left(\prod_{j=1}^d u^{(j)} \right)^{\mathcal{A}\left(\frac{\ln u^{(1)}}{\sum_{j=1}^d \ln u^{(j)}}, \dots, \frac{\ln u^{(d-1)}}{\sum_{j=1}^d \ln u^{(j)}} \right)}, \quad \mathbf{u} \in (0, 1]^d.$$

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Ce chapitre propose une estimation non paramétrique de la fonction de Pickands en présence de données manquantes.

Clustering et extrêmes composés

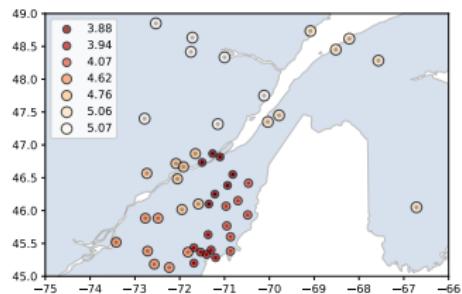
Clustering souple et variable latentes

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Clustering et extrêmes composés

Clustering souple et variable latentes



Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Ce chapitre propose une nouvelle classe de modèles pour le partitionnement des variables ainsi qu'un algorithme adapté à cette problématique.

Clustering et extrêmes composés

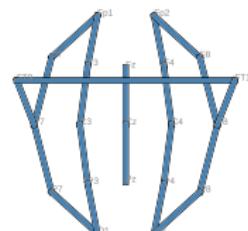
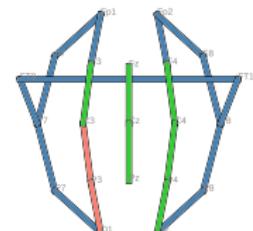
Clustering souple et variable latentes

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Clustering et extrêmes composés

Clustering souple et variable latentes



Madogramme et données manquantes

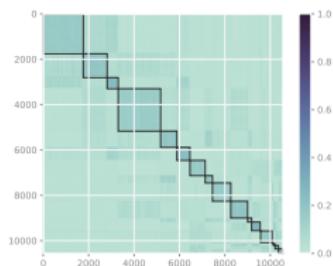
Clustering des variables d'un processus multivarié

Ce chapitre vise à résumer les tendances spatiales d'extrêmes composés de précipitations et de vitesses du vent en Europe à l'aide de méthodes de partitionnement de variables.

Clustering et extrêmes composés

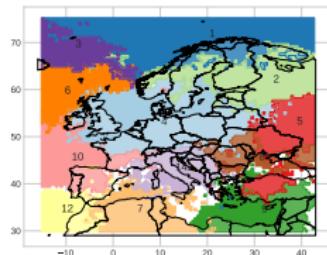
Clustering souple et variable latentes

Madogramme et données manquantes



Clustering des variables d'un processus multivarié

Clustering et extrêmes composés



Clustering souple et variable latentes

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Ce chapitre propose une nouvelle méthode d'estimation pour le modèle à facteurs linéaires, dans lequel la mesure spectrale est discrète et caractérisée par une matrice $A \in \mathbb{R}_+^{d \times K}$.

Clustering et extrêmes composés

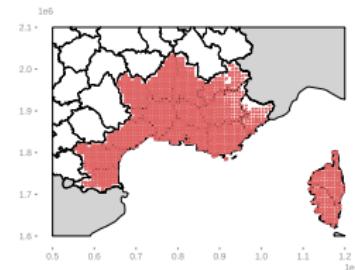
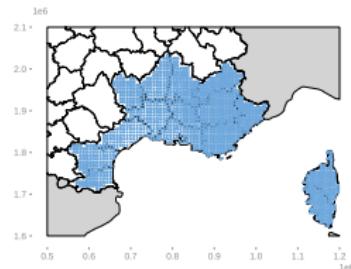
Clustering souple et variable latentes

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Clustering et extrêmes composés

Clustering souple et variable latentes

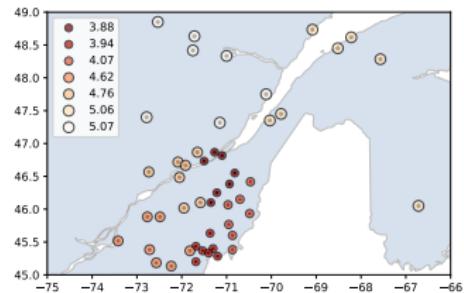


Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Clustering et extrêmes composés

Clustering souple et variable latentes



Définition ([Marcon et al., 2017])

Soit \mathbf{X} un vecteur aléatoire en ayant pour fonction de répartition max-stable $H_{\mathbf{X}}$, le **w**-madogramme est défini par

$$\nu(\mathbf{w}) = \mathbb{E} \left[\bigvee_{j=1}^d \left\{ H^{(j)}(X^{(j)}) \right\}^{1/w^{(j)}} - \frac{1}{d} \sum_{j=1}^d \left\{ H^{(j)}(X^{(j)}) \right\}^{1/w^{(j)}} \right],$$

où $\mathbf{w} \in \Delta_{d-1}$.

Définition ([Marcon et al., 2017])

Soit \mathbf{X} un vecteur aléatoire en ayant pour fonction de répartition max-stable $H_{\mathbf{X}}$, le **w**-madogramme est défini par

$$\nu(\mathbf{w}) = \mathbb{E} \left[\sqrt[d]{\left\{ H^{(j)}(X^{(j)}) \right\}^{1/w^{(j)}}} - \frac{1}{d} \sum_{j=1}^d \left\{ H^{(j)}(X^{(j)}) \right\}^{1/w^{(j)}} \right],$$

où $\mathbf{w} \in \Delta_{d-1}$.

Le **w**-madogramme permet d'identifier la fonction de Pickands \mathcal{A} . En effet, on a

$$\mathcal{A}(\mathbf{w}) = \frac{\nu(\mathbf{w}) + c(\mathbf{w})}{1 - \nu(\mathbf{w}) - c(\mathbf{w})},$$

avec $c(\mathbf{w}) = d^{-1} \sum_{j=1}^d w^{(j)} / (1 + w^{(j)})$.

- Soit $\mathbf{X}_1, \dots, \mathbf{X}_n$ des copies i.i.d. de \mathbf{X} .

- Soit $\mathbf{X}_1, \dots, \mathbf{X}_n$ des copies i.i.d. de \mathbf{X} .
- On introduit le vecteur $\tilde{\mathbf{X}}_i \in \bigotimes_{j=1}^d (\mathbb{R} \cup \text{NA})$ défini par

$$\tilde{X}_i^{(j)} = I_i^{(j)} X_i^{(j)} + (1 - I_i^{(j)}) \text{NA}, \quad j \in \{1, \dots, d\},$$

où $I_i^{(j)} \sim \mathcal{B}(p^{(j)}(\mathbf{X}))$ sont i.i.d. avec

$$p^{(j)}(\mathbf{X}) = \mathbb{P}\{I_i^{(j)} = 1 | \mathbf{X}\},$$

$$p(\mathbf{X}) = \mathbb{P}\{I_i^{(1)} = 1, \dots, I_i^{(d)} = 1 | \mathbf{X}\}.$$

- Soit $\mathbf{X}_1, \dots, \mathbf{X}_n$ des copies i.i.d. de \mathbf{X} .
- On introduit le vecteur $\tilde{\mathbf{X}}_i \in \bigotimes_{j=1}^d (\mathbb{R} \cup \text{NA})$ défini par

$$\tilde{X}_i^{(j)} = I_i^{(j)} X_i^{(j)} + (1 - I_i^{(j)}) \text{NA}, \quad j \in \{1, \dots, d\},$$

où $I_i^{(j)} \sim \mathcal{B}(p^{(j)}(\mathbf{X}))$ sont i.i.d. avec

$$p^{(j)}(\mathbf{X}) = \mathbb{P}\{I_i^{(j)} = 1 | \mathbf{X}\},$$

$$p(\mathbf{X}) = \mathbb{P}\{I_i^{(1)} = 1, \dots, I_i^{(d)} = 1 | \mathbf{X}\}.$$

Hypothèse (A)

- ① Les marges de la fonction de distribution $H_{\mathbf{X}}$ sont continues.
- ② Pour tout $j \in \{1, \dots, d\}$, les dérivées partielles $\dot{C}_{\infty}^{(j)}$ de la copule des valeurs extrêmes C_{∞} existent et sont continues sur l'ensemble $\{\mathbf{u} \in [0, 1]^d : u^{(j)} \in (0, 1)\}$.

- Soit $\mathbf{X}_1, \dots, \mathbf{X}_n$ des copies i.i.d. de \mathbf{X} .
- On introduit le vecteur $\tilde{\mathbf{X}}_i \in \bigotimes_{j=1}^d (\mathbb{R} \cup \text{NA})$ défini par

$$\tilde{X}_i^{(j)} = I_i^{(j)} X_i^{(j)} + (1 - I_i^{(j)}) \text{NA}, \quad j \in \{1, \dots, d\},$$

où $I_i^{(j)} \sim \mathcal{B}(p^{(j)}(\mathbf{X}))$ sont i.i.d. avec

$$p^{(j)}(\mathbf{X}) = \mathbb{P}\{I_i^{(j)} = 1 | \mathbf{X}\},$$

$$p(\mathbf{X}) = \mathbb{P}\{I_i^{(1)} = 1, \dots, I_i^{(d)} = 1 | \mathbf{X}\}.$$

Hypothèse (B)

$\forall i \in \{1, \dots, n\}$, les vecteurs \mathbf{I}_i et \mathbf{X}_i sont indépendants (Missing Completely At Random).

- Sous l'**Hypothèse (B)**, les estimateurs des marges et de la fonction de répartition jointe sont donnés respectivement par :

$$\hat{H}_n^{(j)}(x) = \frac{1}{\sum_{i=1}^n I_i^{(j)}} \sum_{i=1}^n \mathbb{1}_{\{\tilde{X}_i^{(j)} \leq x\}} I_i^{(j)}, \quad x \in \mathbb{R},$$

$$\hat{H}_n(\mathbf{x}) = \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_i^{(j)}} \sum_{i=1}^n \mathbb{1}_{\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\}} \prod_{j=1}^d I_i^{(j)}, \quad \mathbf{x} \in \mathbb{R}^d.$$

- Sous l'**Hypothèse (B)**, les estimateurs des marges et de la fonction de répartition jointe sont donnés respectivement par :

$$\begin{aligned}\hat{H}_n^{(j)}(x) &= \frac{1}{\sum_{i=1}^n I_i^{(j)}} \sum_{i=1}^n \mathbb{1}_{\{\tilde{X}_i^{(j)} \leq x\}} I_i^{(j)}, \quad x \in \mathbb{R}, \\ \hat{H}_n(\mathbf{x}) &= \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_i^{(j)}} \sum_{i=1}^n \mathbb{1}_{\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\}} \prod_{j=1}^d I_i^{(j)}, \quad \mathbf{x} \in \mathbb{R}^d.\end{aligned}$$

- Un estimateur du **w**-madogramme est donné par :

$$\begin{aligned}\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) &= \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_i^{(j)}} \sum_{i=1}^n \left[\bigvee_{j=1}^d \left\{ \hat{H}_n^{(j)}(\tilde{X}_i^{(j)}) \right\}^{1/w^{(j)}} \right. \\ &\quad \left. - \frac{1}{d} \sum_{j=1}^d \left\{ \hat{H}_n^{(j)}(\tilde{X}_i^{(j)}) \right\}^{1/w^{(j)}} \right] \prod_{j=1}^d I_i^{(j)}.\end{aligned}$$

- Sous l'**Hypothèse (B)**, les estimateurs des marges et de la fonction de répartition jointe sont donnés respectivement par :

$$\hat{H}_n^{(j)}(x) = \frac{1}{\sum_{i=1}^n I_i^{(j)}} \sum_{i=1}^n \mathbb{1}_{\{\tilde{X}_i^{(j)} \leq x\}} I_i^{(j)}, \quad x \in \mathbb{R},$$

$$\hat{H}_n(\mathbf{x}) = \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_i^{(j)}} \sum_{i=1}^n \mathbb{1}_{\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\}} \prod_{j=1}^d I_i^{(j)}, \quad \mathbf{x} \in \mathbb{R}^d.$$

- Un estimateur corrigé du **w**-madogramme est donné par :

$$\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) = \hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) -$$

$$\begin{aligned} & \sum_{j=1}^d \frac{\lambda_j(\mathbf{w})(d-1)}{d} \left[\frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_{i,j}} \sum_{i=1}^n \left(\left\{ \hat{H}_n^{(j)}(\tilde{X}_i^{(j)}) \prod_{j=1}^d I_i^{(j)} \right\}^{1/w^{(j)}} \right) \right. \\ & \left. - \frac{w^{(j)}}{1+w^{(j)}} \right] \end{aligned}$$

TCL fonctionnel

Sous les hypothèses **(A)** et **(B)**, nous avons la convergence faible dans $\ell^\infty(\Delta_{d-1})$ de l'estimateur hybride $\hat{\nu}_n^{\mathcal{H}}$ lorsque $n \rightarrow \infty$

$$\begin{aligned} \sqrt{n} \left(\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w}) \right)_{\Delta_{d-1}} &\rightsquigarrow \\ & \left(\frac{1}{d} \sum_{j=1}^d \int_{[0,1]} \alpha \left(\mathbf{1}^{(j)} \left(x^{w^{(j)}} \right) \right) - \beta^{(j)} \left(x^{w^{(j)}} \right) dx \right. \\ & \quad \left. - \int_{[0,1]} S_{C_\infty} \left(x^{w^{(1)}}, \dots, x^{w^{(d)}} \right) \right)_{\mathbf{w} \in \Delta_{d-1}}, \end{aligned}$$

où $S_{C_\infty} = \alpha(\mathbf{u}) - \sum_{j=1}^d \beta^{(j)}(u^{(j)}) \dot{C}_\infty^{(j)}(\mathbf{u})$ et les fonctions de covariance des processus $\alpha, \beta^{(j)}$ pour $j \in \{1, \dots, d\}$ dépendent des probabilités p et $p^{(j)}$.

TCL fonctionnel

Sous les hypothèses **(A)** et **(B)**, nous avons la convergence faible dans $\ell^\infty(\Delta_{d-1})$ de l'estimateur hybride $\hat{\nu}_n^{\mathcal{H}}$ lorsque $n \rightarrow \infty$

$$\begin{aligned} \sqrt{n} \left(\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) - \nu(\mathbf{w}) \right)_{\Delta_{d-1}} &\rightsquigarrow \\ & \left(\frac{1}{d} \sum_{j=1}^d (1 + \lambda^{(j)}(\mathbf{w})(d-1)) \int_{[0,1]} \alpha \left(\mathbf{1}^{(j)} \left(x^{w^{(j)}} \right) \right) - \beta^{(j)} \left(x^{w^{(j)}} \right) dx \right. \\ & \quad \left. - \int_{[0,1]} S_{C_\infty} \left(x^{w^{(1)}}, \dots, x^{w^{(d)}} \right) \right)_{\mathbf{w} \in \Delta_{d-1}}, \end{aligned}$$

où $S_{C_\infty} = \alpha(\mathbf{u}) - \sum_{j=1}^d \beta^{(j)}(u^{(j)}) C_\infty^{(j)}(\mathbf{u})$ et les fonctions de covariance des processus $\alpha, \beta^{(j)}$ pour $j \in \{1, \dots, d\}$ dépendent des probabilités p et $p^{(j)}$.

- Pour $\mathbf{w} \in \Delta_{d-1}$, nous avons

$$\sqrt{n} \left(\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{S}^{\mathcal{H}}(\mathbf{p}, \mathbf{w})),$$

$$\sqrt{n} \left(\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) - \nu(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{S}^{\mathcal{H}*}(\mathbf{p}, \mathbf{w})),$$

où $\mathbf{p} = (p^{(1)}, \dots, p^{(d)}, p, (p^{(ij)})_{1 \leq i, j \leq d})$.

- Pour $\mathbf{w} \in \Delta_{d-1}$, nous avons

$$\sqrt{n} \left(\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, S^{\mathcal{H}}(\mathbf{p}, \mathbf{w})),$$

$$\sqrt{n} \left(\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) - \nu(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, S^{\mathcal{H}^*}(\mathbf{p}, \mathbf{w})),$$

où $\mathbf{p} = (p^{(1)}, \dots, p^{(d)}, p, (p^{(ij)})_{1 \leq i, j \leq d})$.

- Nous pouvons écrire ces deux variances asymptotiques comme

$$S^{\mathcal{H}}(\mathbf{p}, \mathbf{w}) = f(\mathbf{p}, \mathcal{A}, \mathbf{w}), \quad S^{\mathcal{H}^*}(\mathbf{p}, \mathbf{w}) = g(\mathbf{p}, \mathcal{A}, \mathbf{w}),$$

où f et g sont deux fonctionnelles dont nous avons détaillé l'expression.

Nous définissons l' estimateur de la fonction de dépendance de Pickands pour l'échantillon $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)$ comme suit :

$$\hat{\mathcal{A}}_n^{\mathcal{H}^*}(\mathbf{w}) = \frac{\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) + c(\mathbf{w})}{1 - \hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) - c(\mathbf{w})}.$$

Nous définissons l'estimateur de la fonction de dépendance de Pickands pour l'échantillon $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)$ comme suit :

$$\hat{\mathcal{A}}_n^{\mathcal{H}^*}(\mathbf{w}) = \frac{\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) + c(\mathbf{w})}{1 - \hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) - c(\mathbf{w})}.$$

TCL

Sous les hypothèses **(A)** et **(B)**, l'estimateur hybride $\hat{\mathcal{A}}_n^{\mathcal{H}^*}(\mathbf{w})$ converge faiblement lorsque $n \rightarrow \infty$ et $\mathbf{w} \in \Delta_{d-1}$:

$$\sqrt{n} \left(\hat{\mathcal{A}}_n^{\mathcal{H}^*}(\mathbf{w}) - \mathcal{A}(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{V}(\mathbf{p}, \mathbf{w})),$$

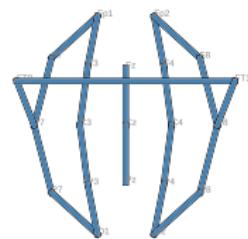
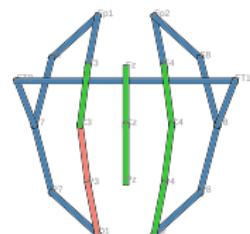
où $\mathcal{V}(\mathbf{p}, \mathbf{w}) = (1 + \mathcal{A}(\mathbf{w}))^4 \mathcal{S}^{\mathcal{H}^*}(\mathbf{p}, \mathbf{w})$.

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Clustering et extrêmes composés

Clustering souple et variable latentes



- On considère $(\mathbf{Z}_t, t \in \mathbb{Z})$ comme un processus multivarié strictement stationnaire avec une loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$.

- On considère $(\mathbf{Z}_t, t \in \mathbb{Z})$ comme un processus multivarié strictement stationnaire avec une loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$.
- Notons $\mathbf{M}_m = (M_m^{(j)})_{j=1}^d$ les maxima composante par composante avec $M_m^{(j)} = \vee_{i=1}^m Z_i^{(j)}$.

- Il existe $\mathbf{a}_m = (a_{m,1}, \dots, a_{m,d})$ et $\Lambda_{\mathbf{Z}}$ tel que

$$\lim_{m \rightarrow \infty} \mathbb{P}\{\mathbf{M}_m \leq \mathbf{a}_m(\mathbf{x})\} = e^{-\Lambda_{\mathbf{Z}}(\mathcal{E}_+ \setminus [0, \mathbf{x}])}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

- Il existe $\mathbf{a}_m = (a_{m,1}, \dots, a_{m,d})$ et $\Lambda_{\mathbf{Z}}$ tel que

$$\lim_{m \rightarrow \infty} \mathbb{P}\{\mathbf{M}_m \leq \mathbf{a}_m(\mathbf{x})\} = e^{-\Lambda_{\mathbf{Z}}(\mathcal{E}_+ \setminus [0, \mathbf{x}])}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

- Il existe une partition $O = \{O_1, \dots, O_G\}$ de $\{1, \dots, d\}$ tel que

$$\Lambda_{\mathbf{Z}} = \sum_{g=1}^G \delta_0 \otimes \cdots \otimes \Lambda^{(O_g)} \otimes \cdots \otimes \delta_0,$$

où $\Lambda^{(O_g)}(A) = \Lambda_{\mathbf{Z}}(A \times [0, \infty)^{d-|O_g|})$ avec A un Borélien de $\mathbb{R}^{(O_g)}$.

- Il existe $\mathbf{a}_m = (a_{m,1}, \dots, a_{m,d})$ et $\Lambda_{\mathbf{Z}}$ tel que

$$\lim_{m \rightarrow \infty} \mathbb{P}\{\mathbf{M}_m \leq \mathbf{a}_m(\mathbf{x})\} = e^{-\Lambda_{\mathbf{Z}}(\mathcal{E}_+ \setminus [0, \mathbf{x}])}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

- Il existe une partition $O = \{O_1, \dots, O_G\}$ de $\{1, \dots, d\}$ tel que

$$\Lambda_{\mathbf{Z}} = \sum_{g=1}^G \delta_0 \otimes \cdots \otimes \Lambda^{(O_g)} \otimes \cdots \otimes \delta_0,$$

où $\Lambda^{(O_g)}(A) = \Lambda_{\mathbf{Z}}(A \times [0, \infty)^{d-|O_g|})$ avec A un Borélien de $\mathbb{R}^{(O_g)}$.

Alors, $\mathcal{L}((\mathbf{Z}, t \in \mathbb{Z}))$ suit un modèle

Asymptotiquement Indépendant (AI) par blocs.

- Cadre probabiliste : On souhaite retrouver la partition la plus fine \bar{O} telle que $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ soit un modèle AI-blocs.

- Cadre probabiliste : On souhaite retrouver la partition la plus fine \bar{O} telle que $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ soit un modèle AI-blocs.
- Cadre statistique : On observe un échantillon du processus $(\mathbf{Z}_t, t \in \mathbb{Z})$.

- **Cadre probabiliste** : On souhaite retrouver la partition la plus fine \bar{O} telle que $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ soit un modèle AI-blocs.
- **Cadre statistique** : On observe un échantillon du processus $(\mathbf{Z}_t, t \in \mathbb{Z})$.
- **Approche par les maxima par blocs** : On divise notre échantillon de taille n en k blocs de taille m . On calcule les maxima du i -ème bloc pour la variable j -ème est notée par

$$M_{m,i}^{(j)} = \max \{ Z_{t,j} : t \in (im - m, im] \cap \mathbb{Z} \} .$$

Hypothèse (A)

Chaque paire d'indice (a, b) dans le cluster \bar{O}_g dans la partition \bar{O} avec $g = 1, \dots, G$ est asymptotiquement dépendant.

Hypothèse (A)

Chaque paire d'indice (a, b) dans le cluster \bar{O}_g dans la partition \bar{O} avec $g = 1, \dots, G$ est asymptotiquement dépendant.

Une condition suffisante est que pour chaque cluster \bar{O}_g dans la partition \bar{O} , la mesure exponentielle $\Lambda^{(\bar{O}_g)}$ possède une densité par rapport à la mesure de Lebesgue sur le quadrant positif $\mathbb{R}_+^{(\bar{O}_g)} \setminus \{\mathbf{0}\}$.

Corrélation extrémale

Soit $(Z^{(a)}, Z^{(b)})$ un couple de variables aléatoires dans le domaine d'attraction d'une distribution max-stable $H^{(a,b)}$, alors la corrélation extrémale est définie par

$$\chi(a, b) = 2 - \theta(a, b) := 2 - L^{(a,b)}(1, 1) = 2 - 2\mathcal{A}^{(a,b)}(1/2),$$

avec $\chi(a, b) \in [0, 1]$. Si $\chi(a, b) = 0$, alors le vecteur $(Z^{(a)}, Z^{(b)})$ est asymptotiquement indépendant.

Dans un modèle AI-blocs, l'énoncé

$$\mathbf{Z}^{(O_g)} \perp\!\!\!\perp_{ext} \mathbf{Z}^{(O_h)}, \quad g \neq h,$$

est équivalent à l'énoncé impliquant la corrélation extrémale χ ,

$$\chi(a, b) = \chi(b, a) = 0, \quad \forall a \in O_g, \forall b \in O_h, \quad g \neq h.$$

Dans un modèle AI-blocs, l'énoncé

$$\mathbf{Z}^{(O_g)} \perp\!\!\!\perp_{ext} \mathbf{Z}^{(O_h)}, \quad g \neq h,$$

est équivalent à l'énoncé impliquant la corrélation extrémale χ ,

$$\chi(a, b) = \chi(b, a) = 0, \quad \forall a \in O_g, \forall b \in O_h, \quad g \neq h.$$

On notera par $\mathcal{X} = [\chi(a, b)]_{a, b \in \{1, \dots, d\}}$ la matrice formée par les corrélations extrémales bivariées de \mathbf{Z} .

Rappel

$$\chi(a, b) = 2 - \theta(a, b) := 2 - 2\mathcal{A}^{(a,b)}(1/2).$$

- Nous estimons la corrélation extrémale en utilisant l'approche suivante :

$$\hat{\chi}_{n,m}(a, b) = 2 - \hat{\theta}_{n,m}(a, b), \quad \hat{\theta}_{n,m}(a, b) = \frac{1/2 + \hat{\nu}_{n,m}(a, b)}{1/2 - \hat{\nu}_{n,m}(a, b)}.$$

Rappel

$$\chi(a, b) = 2 - \theta(a, b) := 2 - 2\mathcal{A}^{(a,b)}(1/2).$$

- Nous estimons la corrélation extrémale en utilisant l'approche suivante :

$$\hat{\chi}_{n,m}(a, b) = 2 - \hat{\theta}_{n,m}(a, b), \quad \hat{\theta}_{n,m}(a, b) = \frac{1/2 + \hat{\nu}_{n,m}(a, b)}{1/2 - \hat{\nu}_{n,m}(a, b)}.$$

- La quantité $\hat{\nu}_{n,m}(a, b)$ est l'estimateur du madogramme, donné par

$$\hat{\nu}_{n,m}(a, b) = \frac{1}{k} \sum_{i=1}^k \left[\bigvee_{j \in \{a,b\}} \hat{U}_{n,m,i}^{(j)} - \frac{1}{2} \sum_{j \in \{a,b\}} \hat{U}_{n,m,i}^{(j)} \right],$$

où $(\hat{U}_{n,m,1}^{(j)}, \dots, \hat{U}_{n,m,k}^{(j)})$ sont les rangs échelonnés des maxima par blocs.

Rappel

$$\chi(a, b) = 2 - \theta(a, b) := 2 - 2\mathcal{A}^{(a,b)}(1/2).$$

- Nous estimons la corrélation extrémale en utilisant l'approche suivante :

$$\hat{\chi}_{n,m}(a, b) = 2 - \hat{\theta}_{n,m}(a, b), \quad \hat{\theta}_{n,m}(a, b) = \frac{1/2 + \hat{\nu}_{n,m}(a, b)}{1/2 - \hat{\nu}_{n,m}(a, b)}.$$

- La quantité $\hat{\nu}_{n,m}(a, b)$ est l'estimateur du madogramme, donné par

$$\hat{\nu}_{n,m}(a, b) = \frac{1}{k} \sum_{i=1}^k \left[\bigvee_{j \in \{a,b\}} \hat{U}_{n,m,i}^{(j)} - \frac{1}{2} \sum_{j \in \{a,b\}} \hat{U}_{n,m,i}^{(j)} \right],$$

où $(\hat{U}_{n,m,1}^{(j)}, \dots, \hat{U}_{n,m,k}^{(j)})$ sont les rangs échelonnés des maxima par blocs.

- On notera par $\hat{\mathcal{X}} = [\hat{\chi}_{n,m}(a, b)]_{a,b \in \{1, \dots, d\}}$ la matrice des corrélations extrémales estimées.

Initialisation :

- $S = \{1, \dots, d\}$,

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

Initialisation :

- $S = \{1, \dots, d\}$,
- $\hat{\mathcal{X}}$,

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

Initialisation :

- $S = \{1, \dots, d\}$,
- $\hat{\mathcal{X}}$,
- $\tau = 0.14 \in [0, 1]$.

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

Première instance :

- $(a_e, b_e) = \arg \max_{a, b \in S} \hat{\chi}_{n,m}(a, b)$.

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

Deuxième instance :

- Si $\hat{\chi}_{n,m}(a_\ell, b_\ell) \geq \tau$,
alors :

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

Deuxième instance :

- Si $\hat{\chi}_{n,m}(a_\ell, b_\ell) \geq \tau$, alors :
 - $\hat{O}_1 = \{s \in S : \hat{\chi}_{n,m}(a_\ell, s) \wedge \hat{\chi}_{n,m}(b_\ell, s) \geq \tau\}$.

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

Troisième instance :

- $S = S \setminus \hat{O}_1$

		4	5
4		1	0.08
5		0.08	1

Première instance :

- $(a_\ell, b_\ell) = \arg \max_{a, b \in S} \hat{\chi}_{n,m}(a, b)$.

	4	5
4	1	0.08
5	0.08	1

- Deuxième instance :
- Si $\hat{\chi}_{n,m}(a_\ell, b_\ell) < \tau$,
alors :

	4	5
4	1	0.08
5	0.08	1

- Deuxième instance :
- Si $\hat{\chi}_{n,m}(a_\ell, b_\ell) < \tau$, alors :
 - $\hat{O}_2 = \{a_\ell\}$.

		4	5
4		1	0.08
5		0.08	1

Troisième instance :

- $S = S \setminus \hat{O}_2$

$$\begin{array}{c|c} & 5 \\ \hline 5 & 1 \end{array}$$

- Première instance :
- Puisque $|S| = 1$, alors :

$$\begin{array}{r|l} & 5 \\ \hline 5 & 1 \end{array}$$

Première instance :

- Puisque $|S| = 1$, alors :
 - $\hat{O}_3 = \{a_\ell\}$.

$$\begin{array}{r} & 5 \\ \hline 5 & | 1 \end{array}$$

Sortie de l'algorithme :

- $\hat{O} = \{\{1, 2, 3\}, \{4\}, \{5\}\}$

	1	2	3	4	5
1	1	0.33	0.34	0.09	0.1
2	0.33	1	0.57	0.05	0.11
3	0.34	0.57	1	0.05	0.15
4	0.09	0.05	0.05	1	0.08
5	0.1	0.11	0.15	0.08	1

- La dépendance serielle de $(\mathbf{Z}_t, t \in \mathbb{Z})$ est contrôlée à l'aide de coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\varphi(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2, \mathbb{P}(A_1) \neq 0} |\mathbb{P}(A_2 | A_1) - \mathbb{P}(A_2)|.$$

- La dépendance serielle de $(\mathbf{Z}_t, t \in \mathbb{Z})$ est contrôlée à l'aide de coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\varphi(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2, \mathbb{P}(A_1) \neq 0} |\mathbb{P}(A_2|A_1) - \mathbb{P}(A_1)|.$$

- Le coefficient de mélange pour $n \geq 1$ est défini par :

$$\varphi(n) = \varphi(\mathcal{F}_t, \mathcal{G}_{t+n})$$

où $\mathcal{F}_k = \sigma(\mathbf{Z}_t, t \leq k)$ et $\mathcal{G}_\ell = \sigma(\mathbf{Z}_t, t \geq \ell)$.

- La dépendance serielle de $(\mathbf{Z}_t, t \in \mathbb{Z})$ est contrôlée à l'aide de coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\varphi(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2, \mathbb{P}(A_1) \neq 0} |\mathbb{P}(A_2|A_1) - \mathbb{P}(A_1)|.$$

- Le coefficient de mélange pour $n \geq 1$ est défini par :

$$\varphi(n) = \varphi(\mathcal{F}_t, \mathcal{G}_{t+n})$$

où $\mathcal{F}_k = \sigma(\mathbf{Z}_t, t \leq k)$ et $\mathcal{G}_\ell = \sigma(\mathbf{Z}_t, t \geq \ell)$.

- Si $\varphi(n) \rightarrow 0$ lorsque $n \rightarrow \infty$, alors $(\mathbf{Z}_t, t \in \mathbb{Z})$ est un processus φ -mélangeant.

Théorème

- Définissons les deux quantités suivantes

$$\text{MECO}(\mathcal{X}) := \min_{\substack{a \sim b \\ a \in \bar{\mathcal{O}}}} \chi(a, b), \quad d_m = \max_{a \neq b} |\chi_m(a, b) - \chi(a, b)|.$$

Théorème

- Définissons les deux quantités suivantes

$$\text{MECO}(\mathcal{X}) := \min_{\substack{a \sim b \\ a \sim b}} \chi(a, b), \quad d_m = \max_{a \neq b} |\chi_m(a, b) - \chi(a, b)|.$$

- Soit (τ, η) deux paramètres vérifiant

$$\tau = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right), \quad \eta = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right).$$

Théorème

- Définissons les deux quantités suivantes

$$\text{MECO}(\mathcal{X}) := \min_{\substack{a \sim b}} \chi(a, b), \quad d_m = \max_{a \neq b} |\chi_m(a, b) - \chi(a, b)|.$$

- Soit (τ, η) deux paramètres vérifiant

$$\tau = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right), \quad \eta = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right).$$

- $(\mathbf{Z}_t, t \in \mathbb{Z})$ suivant un modèle AI-blocs.

Théorème

- Définissons les deux quantités suivantes

$$\text{MECO}(\mathcal{X}) := \min_{\substack{a \sim b}} \chi(a, b), \quad d_m = \max_{a \neq b} |\chi_m(a, b) - \chi(a, b)|.$$

- Soit (τ, η) deux paramètres vérifiant

$$\tau = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right), \quad \eta = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right).$$

- $(\mathbf{Z}_t, t \in \mathbb{Z})$ suivant un modèle AI-blocs.
- Le processus considéré est algébriquement φ -mélangeant, avec $\varphi(n) \leq \lambda n^{-\zeta}$ où $\lambda > 0$ et $\zeta > 1$.

Théorème

- Définissons les deux quantités suivantes

$$\text{MECO}(\mathcal{X}) := \min_{\substack{a \sim b}} \chi(a, b), \quad d_m = \max_{a \neq b} |\chi_m(a, b) - \chi(a, b)|.$$

- Soit (τ, η) deux paramètres vérifiant

$$\tau = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right), \quad \eta = O \left(d_m + \sqrt{\frac{\ln(d)}{k}} \right).$$

- $(\mathbf{Z}_t, t \in \mathbb{Z})$ suivant un modèle AI-blocs.
- Le processus considéré est algébriquement φ -mélangeant, avec $\varphi(n) \leq \lambda n^{-\zeta}$ où $\lambda > 0$ et $\zeta > 1$.
- Si $\text{MECO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant, dans le sens où

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}},$$

avec $\gamma > 0$.

Rappel

Soit (τ, η) deux paramètres vérifiant

$$\tau = O\left(d_m + \sqrt{\frac{\ln(d)}{k}}\right), \quad \eta = O\left(d_m + \sqrt{\frac{\ln(d)}{k}}\right),$$

si $\text{MECO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant,

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}}.$$

Rappel

Soit (τ, η) deux paramètres vérifiant

$$\tau = O\left(d_m + \sqrt{\frac{\ln(d)}{k}}\right), \quad \eta = O\left(d_m + \sqrt{\frac{\ln(d)}{k}}\right),$$

si $\text{MEO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant,

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}}.$$

- Nous n'avons pas supposé que la loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ est max-stable. Un lien est implicitement donné par un terme de biais d_m .

Rappel

Soit (τ, η) deux paramètres vérifiant

$$\tau = O\left(\textcolor{red}{d_m} + \sqrt{\frac{\ln(d)}{k}}\right), \quad \eta = O\left(\textcolor{red}{d_m} + \sqrt{\frac{\ln(d)}{k}}\right),$$

si $\text{MECO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant,

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}}.$$

- Nous n'avons pas supposé que la loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ est max-stable. Un lien est implicitement donné par un terme de biais $\textcolor{red}{d_m}$.

Rappel

Soit (τ, η) deux paramètres vérifiant

$$\tau = O\left(d_m + \sqrt{\frac{\ln(d)}{k}}\right), \quad \eta = O\left(d_m + \sqrt{\frac{\ln(d)}{k}}\right),$$

si $\text{MEO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant,

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}}.$$

- Nous n'avons pas supposé que la loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ est max-stable. Un lien est implicitement donné par un terme de biais d_m .
- Une plus grande dimension d ou un plus grand biais d_m impliquent un seuillage τ plus élevé. Un nombre d'observations plus grand n diminue le seuillage τ .

Rappel

Soit (τ, η) deux paramètres vérifiant

$$\tau = O\left(\textcolor{red}{d_m} + \sqrt{\frac{\ln(d)}{k}}\right), \quad \eta = O\left(\textcolor{red}{d_m} + \sqrt{\frac{\ln(d)}{k}}\right),$$

si $\text{MECO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant,

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}}.$$

- Nous n'avons pas supposé que la loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ est max-stable. Un lien est implicitement donné par un terme de biais d_m .
- Une plus grande dimension d ou un plus grand biais $\textcolor{red}{d_m}$ impliquent un seuillage τ plus élevé. Un nombre d'observations plus grand n diminue le seuillage τ .

Rappel

Soit (τ, η) deux paramètres vérifiant

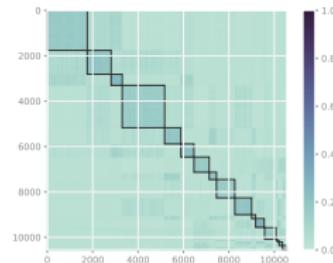
$$\tau = O\left(\textcolor{red}{d_m} + \sqrt{\frac{\ln(d)}{k}}\right), \quad \eta = O\left(\textcolor{red}{d_m} + \sqrt{\frac{\ln(d)}{k}}\right),$$

si $\text{MECO}(\mathcal{X}) > \eta$, alors l'algorithme est consistant,

$$\mathbb{P}\{\hat{O} = \bar{O}\} \geq 1 - \frac{2(1 + \sqrt{e})}{d^{2\gamma}}.$$

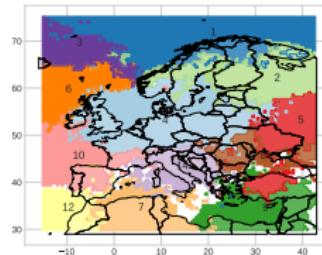
- Nous n'avons pas supposé que la loi stationnaire $\mathcal{L}((\mathbf{Z}_t, t \in \mathbb{Z}))$ est max-stable. Un lien est implicitement donné par un terme de biais d_m .
- Une plus grande dimension d ou un plus grand biais d_m impliquent un seuillage τ plus élevé. Un nombre d'observations plus grand $\textcolor{red}{n}$ diminue le seuillage τ .

Madogramme et données manquantes



Clustering des variables d'un processus multivarié

Clustering et extrêmes composés



Clustering souple et variable latentes

- On utilise la base de données ERA5 pour étudier les extrêmes composées des sommes de précipitations journalières et des maxima de vent journalier durant l'hiver étendu (de novembre à mars, de 1979 à 2022).

- On utilise la base de données ERA5 pour étudier les extrêmes composées des sommes de précipitations journalières et des maxima de vent journalier durant l'hiver étendu (de novembre à mars, de 1979 à 2022).
- Disponible avec une résolution spatiale de $0,25^\circ$ sur une grille régulière, nous nous concentrons sur la zone $[-15^\circ E, 42,5^\circ E] \times [30^\circ N, 75^\circ N]$ qui couvre l'Europe.

- On utilise la base de données ERA5 pour étudier les extrêmes composées des sommes de précipitations journalières et des maxima de vent journalier durant l'hiver étendu (de novembre à mars, de 1979 à 2022).
- Disponible avec une résolution spatiale de $0,25^\circ$ sur une grille régulière, nous nous concentrerons sur la zone $[-15^\circ E, 42,5^\circ E] \times [30^\circ N, 75^\circ N]$ qui couvre l'Europe.

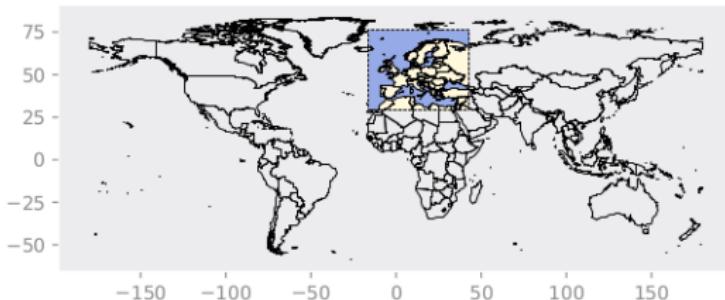


Figure – Région considérée pour l'analyse.

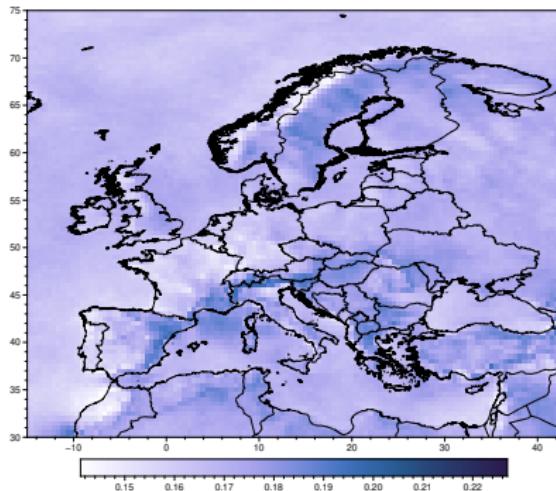


Figure – Proportion du total des sommes de précipitations journalières et des maxima des vitesses du vent qui dépassent leurs 90^e quantiles respectifs.

- Supposons que $\mathbf{Z}_i = (\mathbf{Z}_i^{(1)}, \dots, \mathbf{Z}_i^{(d)})$, $1 \leq i \leq n$ (ici, $n = 6655$) soit un processus stationnaire ayant d (dans notre base de données, $d = 10556$) marges multivariées $\mathbf{Z}_i^{(j)} = (\mathcal{Z}_n^{(j,1)}, \mathcal{Z}_n^{(j,2)})$ avec $j = 1, \dots, d$.

- Supposons que $\mathbf{Z}_i = (\mathbf{Z}_i^{(1)}, \dots, \mathbf{Z}_i^{(d)})$, $1 \leq i \leq n$ (ici, $n = 6655$) soit un processus stationnaire ayant d (dans notre base de données, $d = 10556$) marges multivariées $\mathbf{Z}_i^{(j)} = (\mathcal{Z}_n^{(j,1)}, \mathcal{Z}_n^{(j,2)})$ avec $j = 1, \dots, d$.
- On suppose que le processus $(\mathbf{Z}_i, i \in \mathbb{N})$ appartient au domaine d'attraction d'une distribution max-stable $H_{\mathbf{Z}}$ et qu'il existe une partition $O = \{O_g\}_{g=1}^G$ de $\{1, \dots, d\}$ où $|O_g| = d_g$ tel que $H_{\mathbf{Z}} = \prod_{g=1}^G H^{(O_g)}$.

- Définissons respectivement le coefficient extrémal entre deux pixels et celui sur un pixel par

$$\theta(a, b) = L^{(a, b)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^4}), \quad \theta(j) = L^{(j)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^2}), \quad j \in \{a, b\}.$$

- Définissons respectivement le coefficient extrémal entre deux pixels et celui sur un pixel par

$$\theta(a, b) = L^{(a, b)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^4}), \quad \theta(j) = L^{(j)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^2}), \quad j \in \{a, b\}.$$

- La métrique du SECO est définie par

$$\text{SECO}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)}) = \theta(a) + \theta(b) - \theta(a, b),$$

- Définissons respectivement le coefficient extrémal entre deux pixels et celui sur un pixel par

$$\theta(a, b) = L^{(a, b)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^4}), \quad \theta(j) = L^{(j)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^2}), \quad j \in \{a, b\}.$$

- La métrique du SECO est définie par

$$\text{SECO}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)}) = \theta(a) + \theta(b) - \theta(a, b),$$

- La métrique du SECO est positive et vaut 0 si et seulement si les deux vecteurs sont asymptotiquement indépendants.

- Définissons respectivement le coefficient extrémal entre deux pixels et celui sur un pixel par

$$\theta(a, b) = L^{(a, b)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^4}), \quad \theta(j) = L^{(j)}(\underbrace{\mathbf{1}}_{\in \mathbb{R}^2}), \quad j \in \{a, b\}.$$

- La métrique du SECO est définie par

$$\text{SECO}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)}) = \theta(a) + \theta(b) - \theta(a, b),$$

- La métrique du SECO est positive et vaut 0 si et seulement si les deux vecteurs sont asymptotiquement indépendants.
- De plus, le SECO se réduit à la corrélation extrémale dans le cas où les vecteurs aléatoires sont en réalité des variables aléatoires :

$$\text{SECO}(Z^{(1)}, Z^{(2)}) = 2 - \theta(1, 2) = \chi(1, 2).$$

- L' **estimateur** du SECO est noté par $\widehat{\text{SECO}}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)})$ et est défini par :

$$\widehat{\text{SECO}}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)}) = \hat{\theta}(a) + \hat{\theta}(b) - \hat{\theta}(a, b),$$

où $\hat{\theta}$ est un estimateur du coefficient extrémal θ .

- L' **estimateur** du SECO est noté par $\widehat{\text{SECO}}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)})$ et est défini par :

$$\widehat{\text{SECO}}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)}) = \hat{\theta}(a) + \hat{\theta}(b) - \hat{\theta}(a, b),$$

où $\hat{\theta}$ est un estimateur du coefficient extrémal θ .

- De plus, nous avons

$$\hat{\Theta}(a, b) := \widehat{\text{SECO}}(\mathbf{Z}^{(a)}, \mathbf{Z}^{(b)}) / \min\{\hat{\theta}(a), \hat{\theta}(b)\} \in [0, 1].$$

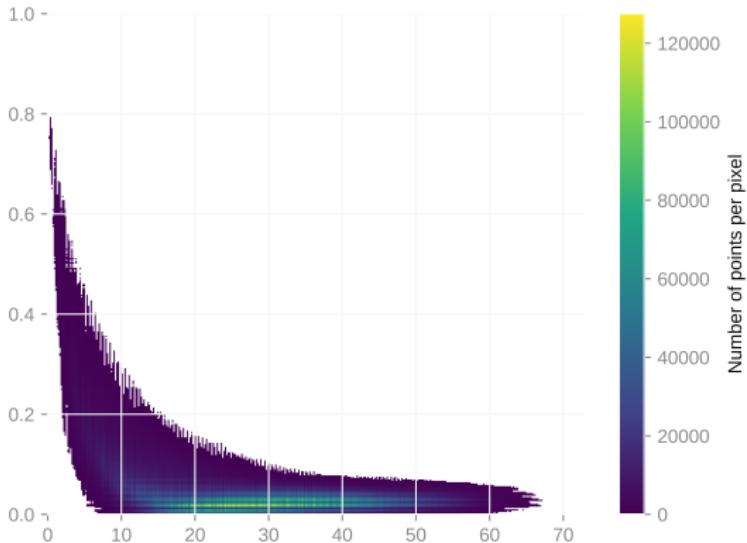


Figure – SECO empirique calculé sur toutes les paires des $d = 10556$ pixels.

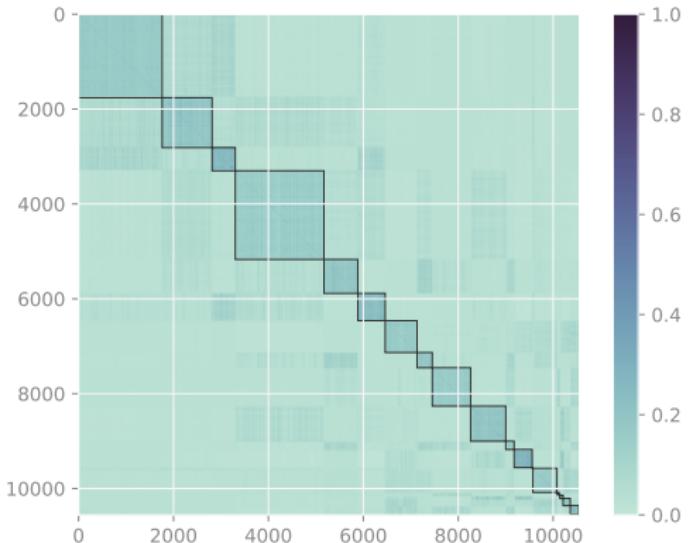


Figure – Partitionnement des variables en utilisant la matrice du SECO avec un seuil $\tau = 0,08$. Les carrés représentent les groupes de variables.

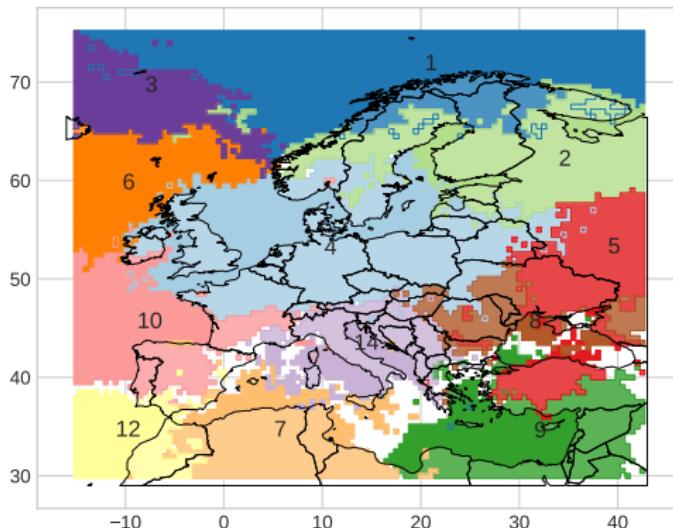


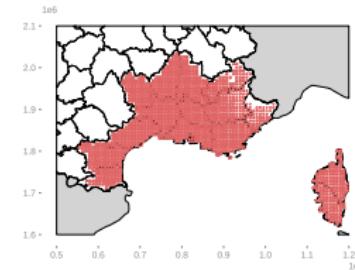
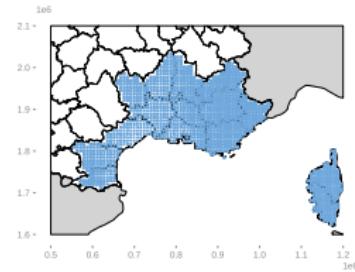
Figure – Représentation des 12 plus grands clusters issus de la matrice de SECO empirique entre les sommes de précipitations journalières et des maxima de vents journaliers en utilisant le seuil $\tau = 0,08$.

Madogramme et données manquantes

Clustering des variables d'un processus multivarié

Clustering et extrêmes composés

Clustering souple et variable latentes



Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Un vecteur aléatoire observable à valeurs dans $\mathbf{X} \in \mathbb{R}^d$;

Modèle à Variables Latentes
 $\mathbf{X} = A\mathbf{Z} + \mathbf{E}$

- Un vecteur aléatoire observable à valeurs dans $\mathbf{X} \in \mathbb{R}^d$;
- Un vecteur aléatoire latent à valeurs dans \mathbb{R}^K qui est dans le domaine d'attraction d'une max-stable avec comme indice caudal $\alpha = 1$ et

$$\Lambda_{\mathbf{Z}} = \sum_{k=1}^K \delta_0 \otimes \cdots \otimes \Lambda_{Z^{(k)}} \otimes \cdots \otimes \delta_0, \quad \Lambda_{Z^{(k)}}(dy) = y^{-2} dy.$$

Modèle à Variables Latentes
 $\mathbf{X} = A\mathbf{Z} + \mathbf{E}$

- Un vecteur aléatoire observable à valeurs dans $\mathbf{X} \in \mathbb{R}^d$;
- Un vecteur aléatoire latent à valeurs dans \mathbb{R}^K qui est dans le domaine d'attraction d'une max-stable avec comme indice caudal $\alpha = 1$ et

$$\Lambda_{\mathbf{Z}} = \sum_{k=1}^K \delta_0 \otimes \cdots \otimes \Lambda_{Z^{(k)}} \otimes \cdots \otimes \delta_0, \quad \Lambda_{Z^{(k)}}(dy) = y^{-2} dy.$$

- Un bruit à queue légère $\mathbf{E} \in \mathbb{R}^d$.

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Un vecteur aléatoire observable à valeurs dans $\mathbf{X} \in \mathbb{R}^d$;
- Un vecteur aléatoire latent à valeurs dans \mathbb{R}^K qui est dans le domaine d'attraction d'une max-stable avec comme indice caudal $\alpha = 1$ et

$$\Lambda_{\mathbf{Z}} = \sum_{k=1}^K \delta_0 \otimes \cdots \otimes \Lambda_{Z^{(k)}} \otimes \cdots \otimes \delta_0, \quad \Lambda_{Z^{(k)}}(dy) = y^{-2} dy.$$

- Un bruit à queue légère $\mathbf{E} \in \mathbb{R}^d$.
- Ce modèle a l'avantage d'être aisément interprétable :

$$X^{(1)} = \begin{array}{c} 0.5 \\ \text{La moitié des extremes} \\ \text{est due à } Z^{(1)} \end{array} Z^{(1)} + \begin{array}{c} 0.5 \\ \text{La moitié des extremes} \\ \text{est due à } Z^{(2)} \end{array} Z^{(2)}$$

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- La mesure spectrale Φ de la loi max-stable $H_{\mathbf{X}}$ a la représentation discrète suivante :

$$\Phi(\cdot) = w^{-1} \sum_{k=1}^K \|A_{\cdot k}\| \delta_{\frac{A_{\cdot k}}{\|A_{\cdot k}\|}}(\cdot), \quad w = \sum_{k=1}^K \|A_{\cdot k}\|,$$

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- La mesure spectrale Φ de la loi max-stable $H_{\mathbf{X}}$ a la représentation discrète suivante :

$$\Phi(\cdot) = w^{-1} \sum_{k=1}^K \|A_{\cdot k}\| \delta_{\frac{A_{\cdot k}}{\|A_{\cdot k}\|}}(\cdot), \quad w = \sum_{k=1}^K \|A_{\cdot k}\|,$$

- Le modèle à variables latentes est une adaptation linéaire du modèle max-linéaire éponyme :

Modèle max-linéaire

$$\mathbf{X} = \left(\sqrt[a=1]{A_{1a}} Z^{(a)}, \dots, \sqrt[a=1]{A_{da}} Z^{(a)} \right)$$

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Nous proposons un model-based clustering via A .

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Nous proposons un **model-based clustering** via A .
- Deux composantes $X^{(i)}$ et $X^{(j)}$ sont considérées comme **communes** si elles présentent une **association positive**.

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Nous proposons un **model-based clustering** via A .
- Deux composantes $X^{(i)}$ et $X^{(j)}$ sont considérées comme **communes** si elles présentent une **association positive**.
- Les variables présentant cette similitude sont **groupées ensemble** dans un cluster nommé G_a :

$$G_a = \{j \in \{1, \dots, d\} : A_{ja} > 0\}, \text{ pour chaque } a \in \{1, \dots, K\}.$$

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Nous proposons un **model-based clustering** via A .
- Deux composantes $X^{(i)}$ et $X^{(j)}$ sont considérées comme **communes** si elles présentent une **association positive**.
- Les variables présentant cette similitude sont **groupées ensemble** dans un cluster nommé G_a :

$$G_a = \{j \in \{1, \dots, d\} : A_{ja} > 0\}, \text{ pour chaque } a \in \{1, \dots, K\}.$$

- Nous imposons les **contraintes supplémentaires** :

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Nous proposons un **model-based clustering** via A .
- Deux composantes $X^{(i)}$ et $X^{(j)}$ sont considérées comme **communes** si elles présentent une **association positive**.
- Les variables présentant cette similitude sont **groupées ensemble** dans un cluster nommé G_a :

$$G_a = \{j \in \{1, \dots, d\} : A_{ja} > 0\}, \text{ pour chaque } a \in \{1, \dots, K\}.$$

- Nous imposons les **contraintes supplémentaires** :

- **Condition (i)** $\sum_{a=1}^K A_{ja} = 1$;

Modèle à Variables Latentes

$$\mathbf{X} = A\mathbf{Z} + \mathbf{E}$$

- Nous proposons un **model-based clustering** via A .
- Deux composantes $X^{(i)}$ et $X^{(j)}$ sont considérées comme **communes** si elles présentent une **association positive**.
- Les variables présentant cette similitude sont **groupées ensemble** dans un cluster nommé G_a :

$$G_a = \{j \in \{1, \dots, d\} : A_{ja} > 0\}, \text{ pour chaque } a \in \{1, \dots, K\}.$$

- Nous imposons les **contraintes supplémentaires** :
 - **Condition (i)** $\sum_{a=1}^K A_{ja} = 1$;
 - **Condition (ii)** Pour chaque $a \in \{1, \dots, K\}$, il existe au moins un indice $j \in \{1, \dots, d\}$ tel que $A_{ja} = 1$ et $A_{jb} = 0$, $\forall b \neq a$.

- Pour une matrice A quelconque, l'ensemble de variables pures s'écrit :

$$I = \bigcup_{a=1}^K I_a, \quad I_a := \{i \in \{1, \dots, d\} : A_{ia} = 1, A_{ib} = 0 \ \forall b \neq a\}.$$

- Pour une matrice A quelconque, l'ensemble de variables pures s'écrit :

$$I = \bigcup_{a=1}^K I_a, \quad I_a := \{i \in \{1, \dots, d\} : A_{ia} = 1, A_{ib} = 0 \forall b \neq a\}.$$

- L'ensemble des variables impures sera noté :

$$J := \{1, \dots, d\} \setminus I.$$

Initialisation :

- $S = \{1, \dots, d\}$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Initialisation :

- $S = \{1, \dots, d\}$,
- $\hat{\mathcal{X}}$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Initialisation :

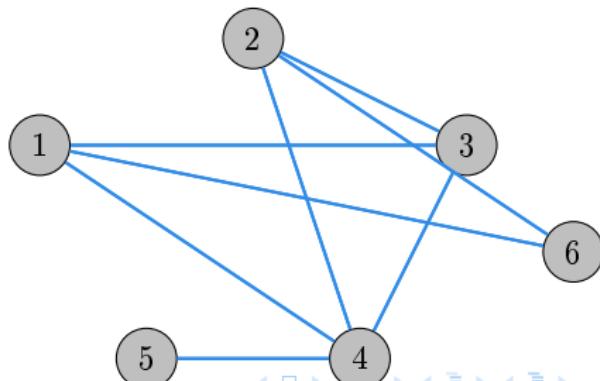
- $S = \{1, \dots, d\}$,
- $\hat{\mathcal{X}}$,
- $\delta = 0.1385 \in [0, 1]$.

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation des variables pures :

- Graphe $G = (V, E)$ où $V = S$ et $(i, j) \in E$ si $\hat{\chi}_{n,m}(i, j) \leq \delta$,

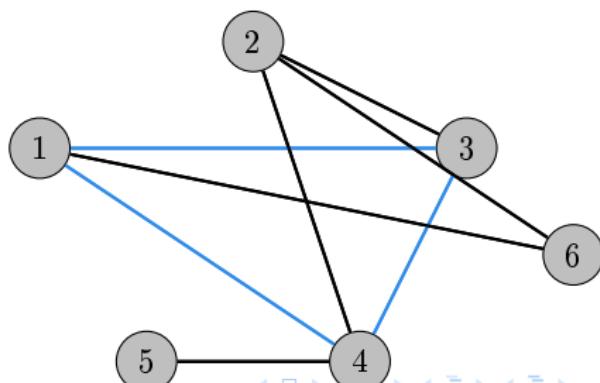
	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1



Estimation des variables pures :

- Graph $G = (V, E)$ où
 $V = S$ et $(i, j) \in E$ si
 $\hat{\chi}_{n,m}(i, j) \leq \delta$,
 - Clique maximale, \bar{G} , de

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1



Estimation des variables pures :

- Graphe $G = (V, E)$ où $V = S$ et $(i, j) \in E$ si $\hat{\chi}_{n,m}(i, j) \leq \delta$,
- Clique maximale, \bar{G} , de G .
- $\hat{I}^{(1)} = \{j \in \{1, \dots, d\} : 1 - \hat{\chi}_{n,m}(1, j) \leq \delta\} = \{1, 2\}$.

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation des variables pures :

- Graphe $G = (V, E)$ où $V = S$ et $(i, j) \in E$ si $\hat{\chi}_{n,m}(i, j) \leq \delta$,
- Clique maximale, \bar{G} , de G .
- $\hat{I}^{(3)} = \{j \in \{1, \dots, d\} : 1 - \hat{\chi}_{n,m}(3, j) \leq \delta\} = \{3\}$.

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation des variables pures :

- Graphe $G = (V, E)$ où $V = S$ et $(i, j) \in E$ si $\hat{\chi}_{n,m}(i, j) \leq \delta$,
- Clique maximale, \bar{G} , de G .
- $\hat{I}^{(4)} = \{j \in \{1, \dots, d\} : 1 - \hat{\chi}_{n,m}(4, j) \leq \delta\} = \{4\}$.

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation des variables pures :

- Graphe $G = (V, E)$ où $V = S$ et $(i, j) \in E$ si $\hat{\chi}_{n,m}(i, j) \leq \delta$,
- Clique maximale, \bar{G} , de G .
- $\mathcal{I} = \{\textcolor{blue}{\{1, 2\}}, \textcolor{red}{\{3\}}, \textcolor{green}{\{4\}}\}$.

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la
sous-matrice A_I :

Estimation de la sous-matrice A_I :

- $\hat{A}_{11} = \hat{A}_{21} = 1$.

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Estimation de la sous-matrice A_I :

- $\hat{A}_{32} = 1$.

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Estimation de la sous-matrice A_I :

- $\hat{A}_{43} = 1$.

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Estimation de la
sous-matrice A_J :

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(5)} = \left(\frac{1}{2}(0.45 + 0.45), 0.62, 0.03\right)$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(5)} = (\frac{1}{2}(0.45 + 0.45), 0.62, 0.03)$,
- $\bar{\beta}^{(5)} = (0.45, 0.62, 0)$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(5)} = (\frac{1}{2}(0.45 + 0.45), 0.62, 0.03)$,
- $\bar{\beta}^{(5)} = (0.45, 0.62, 0)$,
- $\hat{\beta}^{(5)} = (0.415, 0.585, 0)$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(5)} = (\frac{1}{2}(0.45 + 0.45), 0.62, 0.03)$,
- $\bar{\beta}^{(5)} = (0.45, 0.62, 0)$,
- $\hat{\beta}^{(5)} = (0.415, 0.585, 0)$,
- $\hat{A}_5 = \hat{\beta}^{(5)}$.

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.415 & 0.585 & 0 \end{pmatrix}$$

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(6)} = (\frac{1}{2}(0 + 0), 0.37, 0.7)$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(6)} = \left(\frac{1}{2}(0+0), 0.37, 0.7\right)$,
- $\bar{\beta}^{(6)} = (0, 0.37, 0.7)$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(6)} = \left(\frac{1}{2}(0+0), 0.37, 0.7\right)$,
- $\bar{\beta}^{(6)} = (0, 0.37, 0.7)$,
- $\hat{\beta}^{(6)} = (0, 0.335, 665)$,

	1	2	3	4	5	6
1	1	0.97	0.03	0	0.45	0
2	0.97	1	0.03	0	0.45	0
3	0.03	0.03	1	0.06	0.62	0.37
4	0	0	0.06	1	0.03	0.7
5	0.45	0.45	0.62	0.03	1	0.35
6	0	0	0.37	0.7	0.35	1

Estimation de la sous-matrice A_J :

- $\bar{\chi}^{(6)} = (\frac{1}{2}(0 + 0), 0.37, 0.7)$,
- $\bar{\beta}^{(6)} = (0, 0.37, 0.7)$,
- $\hat{\beta}^{(6)} = (0, 0.335, 665)$,
- $\hat{A}_{6\cdot} = \hat{\beta}^{(6)}$.

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.415 & 0.585 & 0 \\ 0 & 0.335 & 0.665 \end{pmatrix}$$

Estimation des clusters
souple :

- $\hat{G}_1 = \{1, 2, 5\}$,

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.415 & 0.585 & 0 \\ 0 & 0.335 & 0.665 \end{pmatrix}$$

Estimation des clusters
souple :

- $\hat{G}_1 = \{1, 2, 5\}$,
- $\hat{G}_2 = \{3, 5, 6\}$,

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.415 & 0.585 & 0 \\ 0 & 0.335 & 0.665 \end{pmatrix}$$

Estimation des clusters
souple :

- $\hat{G}_1 = \{1, 2, 5\}$,
- $\hat{G}_2 = \{3, 5, 6\}$,
- $\hat{G}_3 = \{4\}$.

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.415 & 0.585 & 0 \\ 0 & 0.335 & 0.665 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.41 & 0.59 & 0 \\ 0 & 0.30 & 0.7 \end{pmatrix}$$

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.415 & 0.585 & 0 \\ 0 & 0.335 & 0.665 \end{pmatrix}$$

- La dépendance serielle de $(\mathbf{X}_t, t \in \mathbb{Z})$ est contrôlée via des coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\alpha(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2} |\mathbb{P}(A_2 \cap A_1) - \mathbb{P}(A_1)\mathbb{P}(A_2)|.$$

- La dépendance sérielle de $(\mathbf{X}_t, t \in \mathbb{Z})$ est contrôlée via des coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\alpha(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2} |\mathbb{P}(A_2 \cap A_1) - \mathbb{P}(A_1)\mathbb{P}(A_2)|.$$

- Le coefficient de mélange pour $n \geq 1$ est défini par :

$$\alpha(n) = \alpha(\mathcal{F}_t, \mathcal{G}_{t+n})$$

où $\mathcal{F}_k = \sigma(\mathbf{Z}_t, t \leq k)$ et $\mathcal{G}_\ell = \sigma(\mathbf{Z}_t, t \geq \ell)$.

- La dépendance sérielle de $(\mathbf{X}_t, t \in \mathbb{Z})$ est contrôlée via des coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\alpha(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2} |\mathbb{P}(A_2 \cap A_1) - \mathbb{P}(A_1)\mathbb{P}(A_2)|.$$

- Le coefficient de mélange pour $n \geq 1$ est défini par :

$$\alpha(n) = \alpha(\mathcal{F}_t, \mathcal{G}_{t+n})$$

où $\mathcal{F}_k = \sigma(\mathbf{Z}_t, t \leq k)$ et $\mathcal{G}_\ell = \sigma(\mathbf{Z}_t, t \geq \ell)$.

- Si $\alpha(n) \rightarrow 0$ lorsque $n \rightarrow \infty$, alors $(\mathbf{Z}_t, t \in \mathbb{Z})$ est un processus α -mélangeant.

- La dépendance sérielle de $(\mathbf{X}_t, t \in \mathbb{Z})$ est contrôlée via des coefficients de mélange. Pour deux tribus \mathcal{A}_1 et \mathcal{A}_2 d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit

$$\alpha(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A_1, A_2 \in \mathcal{A}_1 \times \mathcal{A}_2} |\mathbb{P}(A_2 \cap A_1) - \mathbb{P}(A_1)\mathbb{P}(A_2)|.$$

- Le coefficient de mélange pour $n \geq 1$ est défini par :

$$\alpha(n) = \alpha(\mathcal{F}_t, \mathcal{G}_{t+n})$$

où $\mathcal{F}_k = \sigma(\mathbf{Z}_t, t \leq k)$ et $\mathcal{G}_\ell = \sigma(\mathbf{Z}_t, t \geq \ell)$.

- Si $\alpha(n) \rightarrow 0$ lorsque $n \rightarrow \infty$, alors $(\mathbf{Z}_t, t \in \mathbb{Z})$ est un processus α -mélangeant.
- Si $(\mathbf{X}_t, t \in \mathbb{Z})$ est φ -mélangeant alors $(\mathbf{X}_t, t \in \mathbb{Z})$ est α -mélangeant.

Théorème

Posons $s = \max_{j \in \{1, \dots, d\}} \|A_{j\cdot}\|_0$. Soit $(\mathbf{X}_t, t \in \mathbb{Z})$ un processus avec $\alpha(n) \leq \exp(-2cn)$, $c > 0$, et dont les maxima sont dans le domaine d'attraction d'une loi max-stable $H_{\mathbf{X}}$ dont la mesure spectrale est discrète. Sous des conditions de signaux forts sur la matrice A et si

$$\delta = d_m + c_1 \left(\sqrt{\frac{\ln(kd)}{k}} + \frac{\ln(k) \ln \ln(k) \ln(kd)}{k} \right),$$

alors l'estimateur \hat{A} vérifie :

- ① Récupérations des facteurs latents :

$$\hat{K} = K,$$

avec probabilité plus grande que $1 - d^{-c_0}$ pour une constante positive c_0 .

Théorème

Posons $s = \max_{j \in \{1, \dots, d\}} \|A_{j\cdot}\|_0$. Soit $(\mathbf{X}_t, t \in \mathbb{Z})$ un processus avec $\alpha(n) \leq \exp(-2cn)$, $c > 0$, et dont les maxima sont dans le domaine d'attraction d'une loi max-stable $H_{\mathbf{X}}$ dont la mesure spectrale est discrète. Sous des conditions de signaux forts sur la matrice A et si

$$\delta = d_m + c_1 \left(\sqrt{\frac{\ln(kd)}{k}} + \frac{\ln(k) \ln \ln(k) \ln(kd)}{k} \right),$$

alors l'estimateur \hat{A} vérifie :

- ② Une borne supérieure :

$$L_2(\hat{A}, A) \leq 4\sqrt{s}\delta,$$

où $L_2(A, A') := \min_{P \in S_K} \|AP - A'\|_{\infty, 2}$, et $\|A\|_{\infty, 2} := \max_{1 \leq j \leq d} \|A_{j\cdot}\|_2$,

avec probabilité plus grande que $1 - d^{-c_0}$ pour une constante positive c_0 .

Rappel

Si

$$\delta = d_m + c_1 \left(\sqrt{\frac{\ln(kd)}{k}} + \frac{\ln(k) \ln \ln(k) \ln(kd)}{k} \right),$$

Alors l'estimateur \hat{A} de la matrice vérifie :

$$L_2(\hat{A}, A) \leq 4\sqrt{s}\delta,$$

avec probabilité plus grande que $1 - d^{-c_0}$.

Rappel

Si

$$\delta = d_m + c_1 \left(\sqrt{\frac{\ln(kd)}{k}} + \frac{\ln(k) \ln \ln(k) \ln(kd)}{k} \right),$$

Alors l'estimateur \hat{A} de la matrice vérifie :

$$L_2(\hat{A}, A) \leq 4\sqrt{s}\delta,$$

avec probabilité plus grande que $1 - d^{-c_0}$.

- Une plus grande dimension d ou un plus grand biais d_m augmente la distance entre \hat{A} et A . Un nombre d'observation plus grand n diminue la distance entre \hat{A} et A .

Rappel

Si

$$\delta = \textcolor{red}{d_m} + c_1 \left(\sqrt{\frac{\ln(k\textcolor{red}{d})}{k}} + \frac{\ln(k) \ln \ln(k) \ln(k\textcolor{red}{d})}{k} \right),$$

Alors l'estimateur \hat{A} de la matrice vérifie :

$$L_2(\hat{A}, A) \leq 4\sqrt{s}\delta,$$

avec probabilité plus grande que $1 - d^{-c_0}$.

- Une plus grande dimension $\textcolor{red}{d}$ ou un plus grand biais $\textcolor{red}{d_m}$ augmente la distance entre \hat{A} et A . Un nombre d'observation plus grand n diminue la distance entre \hat{A} et A .

Rappel

Si

$$\delta = \textcolor{red}{d_m} + c_1 \left(\sqrt{\frac{\ln(\textcolor{red}{k}d)}{k}} + \frac{\ln(\textcolor{red}{k}) \ln \ln(\textcolor{red}{k}) \ln(\textcolor{red}{k}d)}{k} \right),$$

Alors l'estimateur \hat{A} de la matrice vérifie :

$$L_2(\hat{A}, A) \leq 4\sqrt{s}\delta,$$

avec probabilité plus grande que $1 - d^{-c_0}$.

- Une plus grande dimension d ou un plus grand biais d_m augmente la distance entre \hat{A} et A . Un nombre d'observation plus grand $\textcolor{red}{n}$ diminue la distance entre \hat{A} et A .

Rappel

Si

$$\delta = d_m + c_1 \left(\sqrt{\frac{\ln(kd)}{k}} + \frac{\ln(k) \ln \ln(k) \ln(kd)}{k} \right),$$

Alors l'estimateur \hat{A} de la matrice vérifie :

$$L_2(\hat{A}, A) \leq 4\sqrt{s}\delta,$$

avec probabilité plus grande que $1 - d^{-c_0}$.

- Une plus grande dimension d ou un plus grand biais d_m augmente la distance entre \hat{A} et A . Un nombre d'observation plus grand n diminue la distance entre \hat{A} et A .
- Plus la matrice A est parcimonieuse dont le degré est indiqué par s et plus la distance entre \hat{A} et A est faible.

Rappel

Si

$$\delta = d_m + c_1 \left(\sqrt{\frac{\ln(kd)}{k}} + \frac{\ln(k) \ln \ln(k) \ln(kd)}{k} \right),$$

Alors l'estimateur \hat{A} de la matrice vérifie :

$$L_2(\hat{A}, A) \leq 4\sqrt{\textcolor{red}{s}}\delta,$$

avec probabilité plus grande que $1 - d^{-c_0}$.

- Une plus grande dimension d ou un plus grand biais d_m augmente la distance entre \hat{A} et A . Un nombre d'observation plus grand n diminue la distance entre \hat{A} et A .
- Plus la matrice A est parcimonieuse dont le degré est indiqué par $\textcolor{red}{s}$ et plus la distance entre \hat{A} et A est faible.

- Modèle de régression linéaire pour les extrêmes :

$$Y = \theta^\top \mathbf{X} + \xi,$$

avec $\theta \in \mathbb{R}_+^d$ et le vecteur \mathbf{X} est distribuée selon des Pareto indépendantes.

- Modèle de régression linéaire pour les extrêmes :

$$Y = \theta^\top \mathbf{X} + \xi,$$

avec $\theta \in \mathbb{R}_+^d$ et le vecteur \mathbf{X} est distribuée selon des Pareto indépendantes.

- Sous certaines conditions sur le bruit, alors

$$\inf_{\hat{\theta}} \sup_{\theta \in \mathbb{B}_0(s)} \mathbb{P}_{\mathbf{X}}^n \left\{ \|\hat{\theta} - \theta\|_2 \geq C \sqrt{\frac{s}{n^2} \ln(d/s)} \right\} \geq \beta,$$

$$\inf_{\hat{\theta}} \sup_{\theta \in \mathbb{B}_0(s)} \mathbb{E}_{\mathbf{X}}^n \left[\|\hat{\theta} - \theta\|_2 \right] \geq \tilde{C} \sqrt{\frac{s}{n^2} \ln(d/s)},$$

où $0 < \beta < 1$, $C > 0$ et $\tilde{C} > 0$ sont des constantes absolues.

- Modèle de régression linéaire pour les extrêmes :

$$Y = \theta^\top \mathbf{X} + \xi,$$

avec $\theta \in \mathbb{R}_+^d$ et le vecteur \mathbf{X} est distribuée selon des Pareto indépendantes.

- Sous certaines conditions sur le bruit, alors

$$\inf_{\hat{\theta}} \sup_{\theta \in \mathbb{B}_0(s)} \mathbb{P}_{\mathbf{X}}^n \left\{ \|\hat{\theta} - \theta\|_2 \geq C \sqrt{\frac{s}{n^2} \ln(d/s)} \right\} \geq \beta,$$

$$\inf_{\hat{\theta}} \sup_{\theta \in \mathbb{B}_0(s)} \mathbb{E}_{\mathbf{X}}^n \left[\|\hat{\theta} - \theta\|_2 \right] \geq \tilde{C} \sqrt{\frac{s}{n^2} \ln(d/s)},$$

où $0 < \beta < 1$, $C > 0$ et $\tilde{C} > 0$ sont des constantes absolues.

- Compléxifier le modèle ?

$$\mathbf{X} = A\mathbf{Z} + E, \quad Y = \theta^\top \mathbf{X} + \xi,$$

où $A \in \mathbb{R}_+^{d \times K}$. On peut alors montrer que θ est identifiable.

- Détection de point de rupture dans un modèle à variable latente pour la dépendance extrémale :

$$\mathbf{X}_t = A^{(1)} \mathbf{Z}_t \mathbb{1}_{\{t \leq n_1\}} + A^{(2)} \mathbf{Z}_t \mathbb{1}_{\{t > n_1\}} + \mathbf{E}_t, \quad t = 1, \dots, n.$$

- Détection de point de rupture dans un modèle à variable latente pour la dépendance extrémale :

$$\mathbf{X}_t = A^{(1)} \mathbf{Z}_t \mathbb{1}_{\{t \leq n_1\}} + A^{(2)} \mathbf{Z}_t \mathbb{1}_{\{t > n_1\}} + \mathbf{E}_t, \quad t = 1, \dots, n.$$

- Notons par $\|A\|_{(k)} = \sum_{\ell=1}^k \sigma_\ell(A)$ la norme de Ky-Fan(k) d'une matrice $A \in \mathbb{R}^{d \times K}$ où $k \leq K$ et $\sigma_\ell(A)$ correspondent à la ℓ ième plus grande valeur singulière de A .

- Détection de point de rupture dans un modèle à variable latente pour la dépendance extrémale :

$$\mathbf{X}_t = A^{(1)} \mathbf{Z}_t \mathbb{1}_{\{t \leq n_1\}} + A^{(2)} \mathbf{Z}_t \mathbb{1}_{\{t > n_1\}} + \mathbf{E}_t, \quad t = 1, \dots, n.$$

- Notons par $\|A\|_{(k)} = \sum_{\ell=1}^k \sigma_\ell(A)$ la norme de Ky-Fan(k) d'une matrice $A \in \mathbb{R}^{d \times K}$ où $k \leq K$ et $\sigma_\ell(A)$ correspondent à la ℓ ième plus grande valeur singulière de A .
- Pour des observations $\mathbf{X}_1, \dots, \mathbf{X}_n \in \mathbb{R}^d$, nous notons la matrice estimée seulement avec l'échantillon $1, \dots, n_1$ et $n_1 + 1, \dots, n$ par $\hat{A}_{n_1}^{(1)}$ et $\hat{A}_{n_1}^{(2)}$ respectivement.

- Détection de point de rupture dans un modèle à variable latente pour la dépendance extrémale :

$$\mathbf{X}_t = A^{(1)} \mathbf{Z}_t \mathbb{1}_{\{t \leq n_1\}} + A^{(2)} \mathbf{Z}_t \mathbb{1}_{\{t > n_1\}} + \mathbf{E}_t, \quad t = 1, \dots, n.$$

- Notons par $\|A\|_{(k)} = \sum_{\ell=1}^k \sigma_\ell(A)$ la norme de Ky-Fan(k) d'une matrice $A \in \mathbb{R}^{d \times K}$ où $k \leq K$ et $\sigma_\ell(A)$ correspondent à la ℓ ième plus grande valeur singulière de A .
- Pour des observations $\mathbf{X}_1, \dots, \mathbf{X}_n \in \mathbb{R}^d$, nous notons la matrice estimée seulement avec l'échantillon $1, \dots, n_1$ et $n_1 + 1, \dots, n$ par $\hat{A}_{n_1}^{(1)}$ et $\hat{A}_{n_1}^{(2)}$ respectivement.
- Une statistique de test de H_0 contre H_1 pourrait prendre la forme suivante :

$$T = \max_{n_1=2, \dots, n-2} \max_{k=1, \dots, K} \sqrt{n} \|\hat{A}_{n_1}^{(1)} - \hat{A}_{n_1}^{(2)}\|_{(k)}.$$

Merci de votre attention !



Degemer mat !



Issa Nissa !