



**HAL**  
open science

# Habilitation à diriger des recherches : Contributions à l'étude des extrêmes

Pierre Ribereau

► **To cite this version:**

Pierre Ribereau. Habilitation à diriger des recherches : Contributions à l'étude des extrêmes. Statistiques [math.ST]. Université Claude Bernard Lyon 1, 2021. tel-04888106

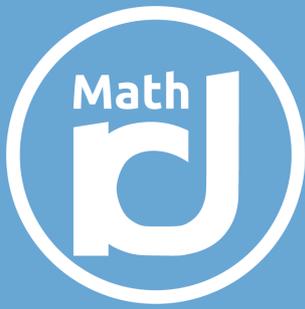
**HAL Id: tel-04888106**

**<https://hal.science/tel-04888106v1>**

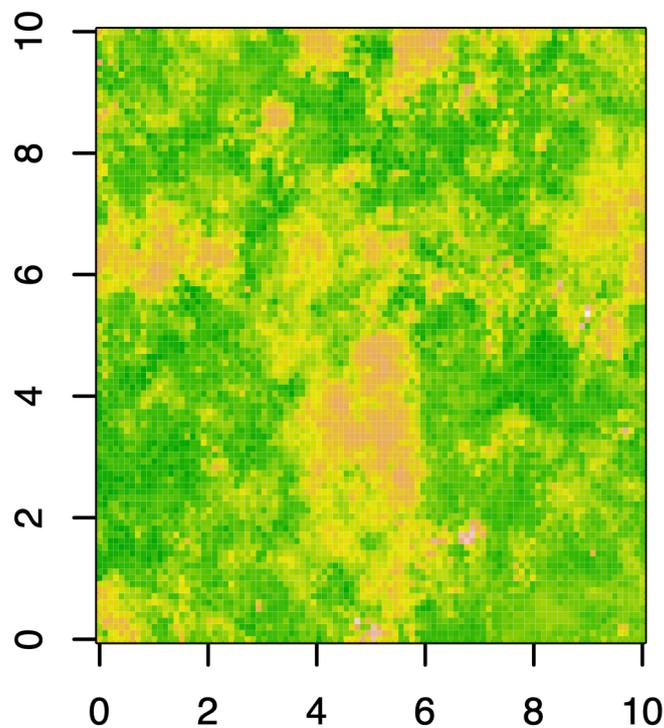
Submitted on 15 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## Contributions à l'étude des extrêmes



**Pierre Ribereau**

Habilitation à diriger des recherches





Université Claude Bernard Lyon 1  
École doctorale **InfoMath**, ED 512  
Spécialité : **Mathématiques**  
N. d'ordre xx-xxxx

# Contributions à l'étude des extrêmes

## Habilitation à diriger des recherches

Soutenue publiquement le 14 janvier 2021 par

**Pierre Ribereau**

devant le Jury composé de :

M. Jean-Noël Bacro	Université de Montpellier	Examineur
Mme. Liliane Bel	AgroParisTech	Rapporteur
M. Clment Dombry	Université de Franche-Comté	Rapporteur
Mme. Anne-Catherine Favre-Pugin	Université de Grenoble	Invitée
M. Carlo Gaetan	Università Ca' Foscari - Venise	Rapporteur
Mme. Armelle Guillou	Université de Strasbourg	Examinatrice
M. Christian Lantuéjoul	MinesParisTech	Invité
Mme Véronique Maume-Deschamps	Université de Lyon	Présidente du jury
M. Philippe Naveau	CNRS - IPSL	Examineur
Mme Gwladys Toulemonde	Université de Montpellier	Examinatrice



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Théorie des Valeurs Extrêmes univariées et multivariées . . . . .	10
1.2	Présentation des processus max-stables . . . . .	15
1.3	Mesures de la dépendance extrême . . . . .	22
1.4	Modèles pour l'indépendance asymptotique . . . . .	26
1.5	Modèles max-mélanges . . . . .	28
1.6	Modèles max-stables spatio-temporels . . . . .	32
<b>2</b>	<b>Extrêmes Univariés</b>	<b>40</b>
2.1	Estimation des paramètres de la GEV par la méthode des moments pondérés généralisés . . . . .	40
2.2	Distribution Skew-GEV . . . . .	42
2.3	Modélisation globale de la distribution des précipitations . . . . .	45
2.4	Mesure de risque basée sur le maximum d'un processus continu . . . . .	48
<b>3</b>	<b>Etude de processus max-stables et max-mélange</b>	<b>56</b>
3.1	Estimation des paramètres d'un processus max-mélange . . . . .	56
3.2	Estimation "non-paramétrique" du paramètre de mélange . . . . .	60
3.3	Estimation semi-paramétrique pour des processus max-stables spatio-temporels . . . . .	67
<b>4</b>	<b>Conclusion et Perspectives</b>	<b>75</b>
	<b>Bibliographie</b>	<b>78</b>

# (Bref) Curriculum Vitae

## **Pierre Ribereau**

Maître de conférences, classe normale  
Université Claude Bernard - Lyon 1  
Section CNU : 26

Né le 19 Juin 1975 à Carcassonne (11)  
3 enfants (2013, 2016, 2017)  
Nationalité Française

Adresse Professionnelle :  
Université Claude Bernard - Lyon 1  
Institut Camille Jordan  
43 boulevard du 11 novembre 1918  
69622 Villeurbanne Cedex

E-mail : pierre.ribereau@univ-lyon1.fr

---

## Carrière

---

- 2011 - :  
**Maître de conférences** à l'Université Claude Bernard Lyon 1.  
Enseignement : Institut de Science Financière et d'Assurance.  
Recherche : Institut Camille Jordan (UMR 5208).
- 2006 - 2011 :  
**Maître de conférences** à l'Université Montpellier 2.
- 2005 - 2006 :  
**Post-doctorat** au Laboratoire des Sciences du Climat et de l'Environnement (avec Pascal You et Philippe Naveau) à Saclay.
- 2000 - 2005 :  
**Thèse de Doctorat** à l'Université Pierre et Marie Curie - Paris 6 sous la direction d'Armelle Guillou et de Paul Deheuvels.
- 1999 - 2000 :  
**DEA de Statistiques** à l'Université Pierre et Marie Curie - Paris 6.

### Articles publiés

- [A-1] Driouchi, D., Ribereau, P. et Treuillet. S. (2004) Ajustements des paramètres d'une chaîne de traitements d'images par un plan d'expériences factoriel fractionnaire  $2^{k-p}$ , *Traitement du Signal*, **Vol. 21**, n°2, pp 141–156.
- [A-2] Blondin, D., Massiani, A. et Ribereau, P. (2005) Vitesses de convergence uniforme presque sûre d'estimateurs non-paramétriques de la régression, *Comptes Rendus de l'Académie des Sciences de Paris, Ser. I*, **340**, pp. 525–528.
- [A-3] Diebolt, J., Guillou, A. et Ribereau, P. (2005) Asymptotic normality of the extreme quantile estimator based on the POT method, *Comptes Rendus de l'Académie des Sciences de Paris, Ser. I*, **341**, pp. 307–312.
- [A-4] Ribereau, P. (2006) Un modèle linéaire à effets mixtes et applications, *Revue de Statistique Appliquée*, **LIV(3)**, pp 65–82.
- [A-5] Yiou, P., Ribereau, P., Naveau, P., Nogaj, M. et Bràdzil, R. (2006) Statistical analysis of floods in Bohemia (Czech Republic) since 1827, *Hydrological Science Journal*, **51(5)**, pp 930–945.
- [A-6] Diebolt, J., Guillou, A. et Ribereau, P. (2007) Asymptotic normality of extreme quantile estimators based on the peak-over-threshold method, *Communications in Statistics : Theory and methods*, **36**, pp 869–886.
- [A-7] Diebolt, J., Guillou, A., Naveau, P. et Ribereau, P. (2008) Improving Probability Weighted Moment Methods for the Generalized Extreme Value Distribution, *REVSTAT, Special issue on "Statistics and related fields"*, 6(1) pp 33–50.

- [A-8] Ribereau, P., Guillou, A. et Naveau, P. (2008) Return Level for Non Stationary GEV distribution based on GPWM, *Non Linear Processes in Geophysics*, **15**(6), pp 1033–1039.
- [A-9] Guillou, A., Naveau, P., Diebolt, J. et Ribereau, P. (2009) Return level bounds for small and moderate sample size, *TEST*, **18**, pp 584–604.
- [A-10] Ribereau, P., Naveau, P. et Guillou, A. (2011) A note of caution when interpreting parameters of excesses distribution, *Water resources research*, vol **34**(10), pp 1215–1221.
- [A-11] Kortschak, D., Loisel, S. et Ribereau, P. (2015) Ruin problems with worsening risks or with infinite mean claims, *Stochastic Models*, **31** :1, pp 119–152.
- [A-12] Opitz, T., Bacro, J.N. et Ribereau, P. (2015) The Spectrogram : A threshold-based inferential tool for extremes of stochastic processes. *Electronic Journal of Statistic*, **9**, pp 842–868.
- [A-13] Ribereau, P., Masiello, E. et Naveau, P. (2016) Skew Generalized Extreme Value Distribution : Probability Weighted Moments Estimation and Application to Block Maxima Procedure, *Communications in Statistics : Theory and Method*, **45** :17, pp 5037–5052.
- [A-14] Naveau, P., Huser, R., Ribereau, P. et Hannart, A. (2016) Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection, *Water Resources Research*, **52** :4, pp 2753-2769.
- [A-15] Blanchet-Scalliet, C., Dorobantu, D., Gay, L., Maume-Deschamps, V. et Ribereau, P. (2018) Risk assessment using suprema data, *Stochastic Environmental Research and Risk Assessment*, **32** :10, pp 2839–2848.
- [A-16] Khadraoui, K. et Ribereau, P. (2019) Bayesian inference with M-splines on spectral measure of bivariate extremes, *Methodology and Computing in Applied Probability*, **21**(3), pp 765–788.
- [A-17] Ahmed, M., Maume-Deschamps, V., Ribereau, P. et Vial, C. (2019) Spatial Risk Measure for max-stable and Max-Mixture Processes, *Stochastics*, **92** :7, pp 1005-1020.

- [A-18] Abu-Awwad, A-F., Maume-Deschamps, V. Ribereau, P. (2020) Censored pairwise likelihood-based tests for mixing coefficient of spatial max-mixture models, *Revista de Investigacion Operacional*, **41**(1), pp 1–26.
- [A-19] Abu-Awwad, A-F., Maume-Deschamps, V. Ribereau, P. (2020) Fitting spatial max-mixture processes with unknown extremal dependence class : an exploratory analysis tool, *TEST*, **29**, pp 479-522.
- [A-20] Abu-Awwad, A-F., Maume-Deschamps, V. et Ribereau, P. (2020) Semi-parametric estimation for space-time max-stable processes :  $F$ -madogram-based estimation approaches, accepté à *Statistical Inference for Stochastic Processes*.
- [A-21] Ahmed, M., Maume-Deschamps, V., Ribereau, P. et Vial, C. (2020) Spatial Risk Measure for Gaussian processes, *Annales de l'ISUP*, **63**(1), pp 3–21.

## Articles soumis

- [S-22] Ahmed, M., Maume-Deschamps, V., Ribereau, P. et Vial, C. A semi-parametric estimation for max-mixture spatial processes, soumis.
- [S-23] Kortschak, D., Masiello, E. et Ribereau, P. Ruin problem in a changing environment with application to the cost of climate change for an insurance company, soumis.
- [S-24] Ahmed, M., Maume-Deschamps, V. et Ribereau, P. Recognizing a Spatial Extreme Model : A Deep Learning approach, soumis.

## Chapitre de livre

- [B-1] Toulemonde, G., Ribereau, P. et Naveau, P. (2016) Applications of extreme value theory to environmental data analysis, *AGU Monographs*, **214**, pp 9–22.

## Articles de vulgarisation

- [V-1] Théron, P. et Ribereau, P. (2012) Profession actuariaire : La Théorie des Valeurs Extrêmes, *L'Actuariel*, **5**, pp 52–54.
- [V-2] Koch, E., Ribereau, P. et Robert, C. (2013) Une tempête dans l'économie, *Mathématiques pour la terre 2013, un jour, une brève*.

---

## Encadrement de thèse

---

**Thomas Opitz** (avec Jean-Noël Bacro) (2010-2013), thèse soutenue le 30 octobre 2013

*Extrêmes multivariés et spatiaux : approches spectrales et modèles elliptiques.*

**Erwan Koch** (avec Christian Y. Robert) (2010-2014), thèse soutenue le 2 Juillet 2014

*Outils et modèles pour l'étude de quelques risques spatiaux et en réseaux : applications aux extrêmes climatiques et à la contagion en finance.*

**Manaf Ahmed** (avec Véronique Maume-Deschamps) (2013-2017), thèse soutenue le 29 Juin 2017

*Sur l'évaluation statistique des risques pour les processus spatiaux.*

**Abdul-Fattah Abu-Awwad** (avec Véronique Maume-Deschamps), thèse soutenue le 20 Juin 2019

*Sur l'inférence statistique pour les processus spatiaux et spatio-temporels.*

---

## Responsabilités/Contrats/Financements

---

1. Contrat de recherche avec la société Ginkyo sur la valorisation statistiques des joueurs de foot. 2019-2020.
2. Organisateur de l'école de recherche CIMPA "Spatial statistics : Extreme Value and Epidemiology" à La Havane (Cuba), Mars 2017.
3. Contrat de recherche avec la société Ginkyo sur la valorisation statistiques des joueurs de foot, 2016-2017.
4. Co-organisateur de la summer school : Extreme value Modeling and water resources, juin 2016.
5. Contrat de recherche avec la société Gerflor sur l'étude sur l'estimation des retours clients pour le système CLIC, 2015.
6. Coordinateur Scientifique du projet MIRACCLE (Mesures et Indicateurs de Risques Adaptés au Changement CLimatiquE), financé par le ministère du Développement Durable via le groupe GICC (Gestion et Impact du Changement Climatique).
7. Co-organisateur des "Entretien Jacques Cartier : Mathématiques Appliquées à la gestion des Risques" (Novembre 2012).
8. Co-organisateur du workshop "Changement climatique et gestion des risques" (Novembre 2011).
9. Correspondant à la mobilité internationale de l'ISFA depuis Novembre 2011.
10. Co-organisateur du Young researchers school on analyzing extreme events distributions (Sept. 2010).

# Chapitre 1

## Introduction

Ce document présente une synthèse de mes travaux de recherche sur la Théorie des Valeurs Extrêmes. Ces travaux se répartissent selon deux grands axes, suivant que l'on considère le cas univarié/bivarié ou le cas spatial.

Dans un premier temps, nous donnerons quelques résultats obtenus en Théorie des Valeurs Extrêmes univariée. Nous présenterons les estimateurs de la distribution des extrêmes généralisée par méthode des moments pondérés généralisés (cf [A-7]). Nous considérerons le cas où les observations sont des maxima pris sur un processus continu d'Ornstein-Uhlenbeck (cf [A-15]) et nous présenterons deux distributions, inspirées de la Théorie des Valeurs Extrêmes, qui permettent une plus grande flexibilité que les distributions "classiques" (cf [A-13]) et [A-14]).

Dans un deuxième temps, nous étudierons le cas spatial. En particulier, nous présenterons les estimateurs des paramètres des processus max-mélanges en se basant sur une méthode des moindres carrés, soit en spécifiant entièrement le modèle (cf [S-22]), soit en estimant uniquement le paramètre de mélange sans poser de modèles complémentaires (cf [A-19]). Nous finirons par la présentation d'estimateur de processus max-stables spatio-temporels (cf [A-20]).

Enfin, nous concluerons par un chapitre présentant quelques perspectives.

Avant de détailler ces résultats, nous rappelons quelques définitions et résultats élémentaires sur la Théorie des Valeurs Extrêmes dans la suite de ce chapitre.

## 1.1 Théorie des Valeurs Extrêmes univariées et multivariées

Les événements extrêmes sont par définition rares et imprévisibles, par conséquent, il est difficile de les gérer. On peut citer par exemple l'étude des maxima annuels ou saisonniers de températures, des périodes de sécheresse, la vitesse maximale de rafales de vent... ou plusieurs de ces variables conjointement. Les sciences du climat sont un des domaines d'applications de la Théorie des Valeurs Extrêmes mais on peut aussi citer l'Hydrologie (cf Katz et al. (2002)) ainsi que l'assurance ou la finance lorsque l'on considère des pertes financières (cf Embrechts et al. (1997)) entre autres.

### Théorie des Valeurs Extrêmes univariées

D'un point de vue probabiliste, considérons une suite  $(X_i)_{i \geq 1}$  de variables aléatoires indépendantes et identiquement distribuées de fonction de répartition commune  $F$ . On sait que si l'on s'intéresse au comportement moyen de la variable, le Théorème Centrale Limite apporte des réponses sur ce problème. D'une manière analogue, nous avons des résultats asymptotiques similaires pour la distribution du maximum de  $n$  variables aléatoires noté  $X_{n,n} = \max_{i=1,\dots,n} X_i$  lorsque la taille de l'échantillon  $n$  augmente. Plus précisément, correctement renormalisé par des suites  $\{a_n\}_n$  et  $\{b_n\}_n$ , le maximum converge vers une distribution des valeurs extrêmes généralisées (GEV), i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{X_{n,n} - a_n}{b_n} \leq x \right) = G_{\mu, \sigma, \xi}(x)$$

où  $G_{\mu, \sigma, \xi}$  satisfait :

$$G_{\mu, \sigma, \xi}(x) = \exp \left( - \left[ 1 + \xi \frac{x - \mu}{\sigma} \right]_+^{-1/\xi} \right). \quad (1.1)$$

Le paramètre  $\mu$  est un paramètre de localisation,  $\sigma$  est un paramètre d'échelle et  $\xi$  est un paramètre de forme. Le cas  $\xi = 0$  s'interprète comme la limite lorsque  $\xi \rightarrow 0$  et correspond à la loi de Gumbel :

$$G_{\mu, \sigma, 0}(x) = \exp \left( - \exp \left[ -\sigma^{-1}(x - \mu) \right] \right).$$

Une propriété fondamentale de cette distribution est la max-stabilité. La distribution  $G$  d'une variable aléatoire  $X$  est max-stable s'il existe des suites  $\{c_n > 0\}$

et  $\{d_n\}$  telles que, pour tout  $n \geq 2$ , le maximum  $X_{n,n}$  sur  $n$  copies indépendantes de  $X$  est égal en distribution à  $c_n X + d_n$ , i.e.,

$$G^n(x) = G\left(\frac{x - d_n}{c_n}\right).$$

Le paramètre de forme  $\xi$  dans (1.1) a un rôle très important en Théorie des Valeurs Extrêmes car son signe et sa valeur donnent une indication sur la décroissance de la fonction queue :  $\bar{F}(x) = 1 - F(x)$  lorsque  $x \rightarrow \infty$ .

Dans la pratique, pour prédire les événements extrêmes ou calculer les probabilités d'apparition d'extrêmes, on a recours au maximum par bloc (*Block Maxima*, cf Chapitre 5, Beirlant et al. (2004)), c'est-à-dire que l'on va considérer le maximum sur des données sur un certain intervalle de temps ou un certain nombre d'observations  $k$ . Ce maximum est supposé être distribué suivant une distribution GEV si  $k$  est suffisamment grand. Si on obtient un nombre suffisant de maxima et qu'on peut supposer qu'ils forment un échantillon indépendant et identiquement distribué, alors on peut estimer les paramètres par maximum de vraisemblance (Smith (1985)) ou en utilisant les moments pondérés (cf Hosking et al. (1985)). Par exemple, dans la Section 2.1 (et [A-6]), nous présenterons l'estimation des paramètres de la distribution GEV par la méthode des moments pondérés généralisés (cf Diebolt et al. (2008)).

Cette approche basée sur le maximum par bloc a parfois été critiquée puisqu'elle ne considère que le maximum sur un certain nombre d'observations et donc peut ignorer d'autres grandes valeurs de l'échantillon, parfois de valeur supérieure à certains maxima retenus. Pour contourner cela, une alternative consiste à considérer l'approche "Excès au delà du seuil" ou PoT (*Peaks-over-Threshold*). Considérons  $X_1, \dots, X_n$  un échantillon de  $n$  variables aléatoires indépendantes et identiquement distribuées suivant une distribution  $F$ . On considère les  $N_u$  variables qui sont supérieures à un seuil  $u$ . L'excès  $Y_i$  correspondant à la variable  $X_i$  est défini par  $Y_i = X_i - u$  si  $X_i > u$ .

La fonction de répartition  $F_u$  de l'excès  $Y$  au delà d'un seuil  $u$  est donné pour tout  $y > 0$  par

$$F_u(y) = \mathbb{P}(Y \leq y | X > u) = \frac{\mathbb{P}(X - u \leq y, X > u)}{\mathbb{P}(X > u)} = \frac{F(u + y) - F(u)}{1 - F(u)}.$$

Si le seuil est suffisamment élevé, on peut approcher cette quantité par la distribution de Pareto Généralisée (GPD), notée  $H_{\sigma_u, \xi}(y)$  dont la fonction de survie est donnée par :

$$\bar{H}_{\sigma_u, \xi}(y) = \begin{cases} \left(1 + \xi \frac{y}{\sigma_u}\right)^{-1/\xi} & \text{si } \xi \neq 0 \\ \exp\left(-\frac{y}{\sigma_u}\right) & \text{sinon .} \end{cases}$$

Cette fonction est définie sur  $\mathbb{R}^+$  si  $\xi \geq 0$  et sur  $[0; -\sigma_u/\xi]$  si  $\xi < 0$ .  $\sigma_u > 0$  est un paramètre d'échelle et  $\xi$  est le paramètre de forme. Le Théorème de Pickands (Pickands (1975)) établit l'équivalence entre les deux approches maxima par blocs et par excès au delà du seuil. Ainsi, nous avons :

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) = G_{\mu, \sigma, \xi}(x)$$

si et seulement si

$$\lim_{u \rightarrow x_F} \sup_{y \in [0; x_F - u]} |\bar{F}_u(y) - \bar{H}_{\xi, \sigma_u}(y)| = 0$$

où  $x_F$  est le point terminal de la distribution, i.e.,  $x_F = \sup\{x : F(x) < 1\}$ . Il est important de remarquer que le paramètre de forme  $\xi$  est le même dans les deux approches alors que le paramètre d'échelle est différent. De fait on a la relation

$$\sigma_u = \sigma + \xi(u - \mu) \tag{1.2}$$

qui permet la construction d'une méthode de choix de taille de bloc pour la méthode *Block Maxima*. On peut préciser qu'il existe une série d'outils diagnostiques et graphiques pour guider le choix du seuil  $u$  (cf Coles (2001)).

Comme dans le cas de la GEV, si on choisit un seuil suffisamment élevé pour avoir l'approximation de la distribution des excès par une GPD et assez bas pour avoir une quantité d'observations  $N_u$  suffisamment grande pour pouvoir faire de l'inférence, on peut estimer les paramètres de la GPD par maximum de vraisemblance ou par la méthode des moments pondérés (cf Hosking & Wallis (1987)).

Dans la Section 2.3 et dans [A-14], nous proposons une extension de la loi GPD : l'*Extended GPD* (EGPD) qui permet de modéliser l'ensemble du spectre des précipitations en étant en accord avec la Théorie des Valeurs Extrêmes (les queues inférieure et supérieure sont approximativement GPD). Par ailleurs, dans [A-3] et [A-5], nous avons montré la normalité asymptotique d'estimateurs de quantiles extrêmes basés sur l'approche PoT et sur la méthode des moments pondérés généralisés. Dans [A-10], nous avons proposé une reparamétrisation de la GPD de façon à rendre les estimateurs des paramètres indépendants. Dans [A-5], nous

avons utilisé l'approche PoT dans un cadre non-stationnaire pour obtenir des estimations de la période et du niveau de retour.

## Théorie des Valeurs Extrêmes multivariés

Nous présentons ici le cas bivarié puisque c'est celui que nous utiliserons dans ce document. Nous appliquons une approche similaire à celle du *Block Maxima*. On suppose que  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  est une suite de vecteurs aléatoires indépendants et identiquement distribués de fonction de répartition commune  $F$ . On note

$$M_{x,n} = \max_{i=1,\dots,n} \{X_i\} \text{ et } M_{y,n} = \max_{i=1,\dots,n} \{Y_i\}$$

$$M_n = (M_{x,n}, M_{y,n})$$

$M_n$  est le vecteur de maxima par composantes. Si on note  $\mathbf{z} = (z_1, z_2)$ , on a

$$\mathbb{P}(M_n \leq \mathbf{z}) = \mathbb{P}(M_{x,n} \leq z_1, M_{y,n} \leq z_2) = F(\mathbf{z})^n.$$

S'il existe deux suites normalisantes de vecteurs  $\{a_n\}_n = \{(a_{n,1}, a_{n,2})\}_n$  et  $\{b_n\}_n = \{(b_{n,1}, b_{n,2})\}_n$  telles que  $a_{n,j} \in \mathbb{R}$  et  $b_{n,j} > 0$  pour  $j = 1, 2$  et une distribution  $G$  de marginales non dégénérées alors

$$F^n(b_n \mathbf{z} + a_n) \rightarrow G(\mathbf{z})$$

où  $b_n \mathbf{z} + a_n = (b_{n,1}z_1 + a_{n,1}, b_{n,2}z_2 + a_{n,2})$ , alors  $G$  est une distribution bivariée généralisée des extrêmes. Avec les mêmes notations, on a

$$G^k(b_k \mathbf{z} + a_k) = G(\mathbf{z})$$

c'est-à-dire  $G$  est max-stable. Ceci implique que les distributions marginales de  $G$  suivent forcément une distribution GEV. On peut donc obtenir une représentation plus simple en supposant que les variables  $X_i$  et  $Y_i$  suivent des lois Fréchet unitaires (i.e., de fonction de répartition  $\exp(-x^{-1})$  pour  $x > 0$ ). On considère alors le vecteur renormalisé

$$M_n^* = (M_{x,n}^*, M_{y,n}^*) = (M_{x,n}/n, M_{y,n}/n)$$

de façon à ce que  $M_n^*$  ait des marginales Fréchet.

**Théorème 1 (Coles (2001))** *Soit  $(X_i, Y_i)$  des vecteurs aléatoires indépendants et identiquement distribués de marginale Fréchet unitaire,  $(M_{x,n}^*, M_{y,n}^*)$  défini comme précédemment. Si :*

$$\mathbb{P}(M_{x,n}^* \leq x, M_{y,n}^* \leq y) \rightarrow G(x, y)$$

où  $G$  est une distribution non dégénérée, alors  $G$  est de la forme

$$G(x, y) = \exp(-V(x, y)), \quad x > 0, y > 0 \quad (1.3)$$

où

$$V(x, y) = 2 \int_0^1 \max \left\{ \frac{\omega}{x}, \frac{1-\omega}{y} \right\} dH(\omega) \quad (1.4)$$

et  $H$  est une distribution sur  $[0, 1]$  qui vérifie la contrainte suivante :

$$\int_0^1 \omega dH(\omega) = 1/2. \quad (1.5)$$

$H$  est appelée la mesure spectrale.

Par exemple, si  $H$  est de la forme

$$H(\omega) = \begin{cases} 1/2 & \text{si } \omega = 0 \text{ ou } 1 \\ 0 & \text{sinon} \end{cases}$$

la distribution bivariable des extrêmes est

$$G(x, y) = \exp(-(x^{-1} + y^{-1}))$$

pour  $x > 0$  et  $y > 0$ . Ce cas correspond à l'indépendance parfaite entre les deux composantes. A l'inverse, le choix pour  $H$  d'une distribution qui met une masse 1 en 1/2 conduit à la distribution des extrêmes suivante :

$$G(x, y) = \exp(-\max\{x^{-1}, y^{-1}\})$$

pour  $x > 0$  et  $y > 0$ , ce qui conduit à la dépendance parfaite, i.e.,  $X = Y$  presque sûrement.

Tout choix de distribution  $H$  qui vérifie (1.5) conduit à une distribution bivariable généralisée. On peut retrouver des familles paramétriques dans Beirlant et al. (2004) ou Coles (2001).

On peut avoir une autre représentation par une approche par processus ponctuel. Soit  $E$  l'ensemble  $E = [0, \infty]^2 \setminus \{0\}$ ,  $\|\cdot\|$  une norme sur  $\mathbb{R}^2$  et  $B \subset E$  la sphère unitaire associée.

**Théorème 2** *Les assertions suivantes sont équivalentes :*

- $G$  est une distribution bivariable des extrêmes de marginale Fréchet unitaire comme dans le Théorème 1.

- Il existe un processus de Poisson non-homogène sur  $[0, \infty) \times E$  d'intensité  $\Lambda$  défini pour tout  $t > 0$  par  $\Lambda([0, t] \times B) = t\mu^*(B)$  où, pour tout  $A \subset B$  et  $r > 0$ ,

$$\mu^* \left( \mathbf{x} \in E : \|\mathbf{x}\| > r ; \frac{\mathbf{x}}{\|\mathbf{x}\|} \in A \right) = 2 \frac{H(A)}{r}, \quad (1.6)$$

où  $\mathbf{x} = (x_1, x_2)$  et  $H$  est une mesure finie telle que (1.5) est vérifiée et

$$G(x) = \exp(-\mu^* \{(x_1, \infty) \times (x_2, \infty)\}).$$

Cette dernière représentation donne une interprétation intéressante de la distribution  $H$ . Soit  $\|\mathbf{x}\| = x_1 + x_2$ , la transformation utilisée dans (1.6) :  $\mathbf{x} = (x_1, x_2) \rightarrow (x_1 + x_2, x_1/(x_1 + x_2)) = (r, \omega)$  est la transformation des coordonnées cartésiennes vers des coordonnées pseudo-polaires, où  $r$  est la mesure de la distance à l'origine et  $\omega$  mesure l'angle sur une échelle sur  $[0, 1]$ . L'équation (1.6) implique que la mesure  $\mu^*$  est la mesure produit entre une fonction de la composante radiale et de la mesure  $H$  de la composante angulaire. En d'autres termes, l'angle est contrôlé par la mesure  $H$  et est indépendant de la distance radiale.

Si  $H$  est dérivable de densité  $h$ , l'interprétation est plus simple. En effet,  $h(\omega)$  mesure alors la fréquence relative des extrêmes dans la direction  $\omega$ . Ainsi, si  $h$  admet de fortes valeurs pour  $\omega$  proche de 0 ou 1 et faibles ailleurs, on est proche de l'indépendance alors que si  $h$  a de fortes valeurs autour de 1/2 et de faibles ailleurs, on est proche de la dépendance parfaite.

Dans [A-16], nous avons proposé un estimateur Bayésien de la mesure spectrale par M-splines et dans [A-12], nous avons "étendu" la mesure spectrale au cas spatial en proposant le spectrogramme comme outil d'exploration et d'inférence spatiales.

Par la suite, nous présentons quelques définitions et notions utiles pour les extrêmes spatiaux.

## 1.2 Présentation des processus max-stables

Même si les processus Gaussiens sont adaptés pour modéliser certains phénomènes spatiaux (cf Gaetan & Guyon (2010)), ils ne sont pas réalistes pour beaucoup de données environnementales, par exemple pour les précipitations. De plus, ces processus ont souvent été critiqués car ils ne sont pas capables de capturer la dépendance asymptotique (cf Davison et al. (2013)). En effet, la distribution Gaussienne est à queue légère, sa fonction de survie décroît exponentiellement et donc

peut sous-estimer les phénomènes extrêmes. De plus, les processus Gaussiens sont asymptotiquement indépendants, ce qui signifie que la dépendance entre deux événements observés en 2 sites différents disparaît au fur et à mesure que les événements deviennent extrêmes. On peut aussi rajouter que lorsque l'on modélise l'occurrence jointe d'extrêmes sur une région, différentes formes de dépendance extrême peuvent apparaître comme l'indépendance asymptotique ou la dépendance asymptotique. La Théorie des Valeurs Extrêmes propose un outil statistique intéressant avec les processus max-stables.

Commençons par définir la fonction coefficient de queue  $\chi(\cdot)$  qui permet de déterminer le type de dépendance extrême entre deux sites.

**Définition 1** Soit  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus spatial stationnaire de distribution marginale commune  $F$ . La fonction coefficient de queue  $\chi$  est définie par

$$\chi(\mathbf{h}) = \lim_{x \rightarrow \infty} \mathbb{P}(X(\mathbf{s}) > x | X(\mathbf{s} + \mathbf{h}) > x), \quad \mathbf{s}, \mathbf{s} + \mathbf{h} \in \mathcal{S}.$$

- Si  $\chi(\mathbf{h}) = 0$ , la paire  $(X(\mathbf{s}), X(\mathbf{s} + \mathbf{h}))$  est dite asymptotiquement indépendante.
- Si  $\chi(\mathbf{h}) \neq 0$ , la paire  $(X(\mathbf{s}), X(\mathbf{s} + \mathbf{h}))$  est dite asymptotiquement dépendante.
- Le processus est dit asymptotiquement indépendant si  $\chi(\mathbf{h}) = 0$  quelque soit  $\mathbf{h} \in \mathcal{S}$ .

Par exemple, un processus max-stable (cf Définition 2) défini sur un espace  $\mathcal{S}$  est soit asymptotiquement dépendant dans le sens où  $\chi(\mathbf{h}) > 0$ , soit indépendant strict. Dans le cas asymptotiquement dépendant, un événement d'intensité forte en  $\mathbf{s} + \mathbf{h}$  implique une probabilité strictement positive d'avoir un événement d'intensité élevée en  $\mathbf{s}$ . Toutefois, dans certaines applications, il semble que  $\chi(\mathbf{h}) = 0$ , comme c'est le cas, par exemple, des épisodes de type Cévenoles qui entraînent des précipitations "extrêmes" très localisées. Dans ce cas, il semble alors préférable d'utiliser un modèle asymptotiquement indépendant, au moins à partir d'une certaine distance entre sites. Comme on l'a dit, les processus Gaussiens sont asymptotiquement indépendants mais ils sont trop restrictifs pour modéliser les événements extrêmes.

Par la suite, nous présentons différents types de processus spatiaux : max-stable, inverse max-stable et max-mélange. Nous présentons aussi des mesures de dépendances adaptées aux processus asymptotiquement dépendants ou indépendants ou qui permettent de discriminer entre eux.

**Définition 2 (Processus max-stable)** Soit  $\{X_i(\mathbf{s}) : \mathbf{s} \in \mathcal{S}\}_{i \geq 0}$  des répliques indépendantes et identiquement distribuées d'un processus aléatoire  $X(\mathbf{s})$ . Le processus est dit max-stable s'il existe deux suites de fonctions continues  $\{a_n(\mathbf{s}) > 0\}$  et  $\{b_n(\mathbf{s})\}$  telles que

$$\left\{ \bigvee_{i=1}^n \frac{X_i(\mathbf{s}) - b_n(\mathbf{s})}{a_n(\mathbf{s})} \right\}_{\mathbf{s} \in \mathcal{S}} \stackrel{\mathcal{D}}{=} \{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}, \quad n = 1, 2, \dots, \quad (1.7)$$

où  $\bigvee$  est l'opérateur maximum.

Cette définition des processus max-stables offre un cadre naturel pour modéliser les extrêmes spatiaux. La théorie des extrêmes univariés implique que la distribution marginale de  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  est une distribution GEV. Sans perte de généralité, on peut transformer les marginales de façon à avoir des marginales Fréchet unitaire. Un processus à marginale Fréchet unitaire est appelé processus max-stable standard. Les processus max-stables standards peuvent être construits avec la représentation spectrale suivante (cf De Haan (1984) ou Schlather (2002)) :

**Théorème 3 (Représentation spectrale pour les processus spatiaux max-stables standards)** Soit  $X$  un processus max-stable sur  $\mathcal{S}$ . Alors il existe  $\{r_i\}_{i \geq 1}$  les réalisations d'un processus de Poisson ponctuel sur  $\mathbb{R}^+$  d'intensité  $dr$  et une suite  $\{U_i(\mathbf{s}), \mathbf{s} \in \mathcal{S}\}_{i \geq 1}$  de copies indépendantes et identiquement distribuées d'un processus aléatoire  $U = \{U(\mathbf{s}), \mathbf{s} \in \mathcal{S}\}$ , indépendant des  $\{r_i\}_{i \geq 1}$  qui satisfait  $\mathbb{E}[U^+(\mathbf{s})] = \mu$  où  $U^+(\mathbf{s}) = \max(U(\mathbf{s}), 0)$ , tels que

$$X(\mathbf{s}) \stackrel{\mathcal{D}}{=} \left\{ \mu^{-1} \bigvee_{i=1}^{\infty} U_i^+(\mathbf{s})/r_i \right\}_{\mathbf{s} \in \mathcal{S}}. \quad (1.8)$$

**Théorème 4 (Maxima Multivariés)** Pour  $D \in \mathbb{N} \setminus \{0\}$ ,  $\mathbf{s}_1, \dots, \mathbf{s}_D \in \mathcal{S}$ , et  $x_1, \dots, x_D > 0$ , la loi  $D$ -dimensionnelle du processus max-stable  $X$  défini en (1.8) est donnée par

$$\begin{aligned} \mathbb{P}(X(\mathbf{s}_1) \leq x_1, \dots, X(\mathbf{s}_D) \leq x_D) &= \exp \left\{ -\mathbb{E} \left[ \bigvee_{i=1}^D \left\{ \frac{U(\mathbf{s}_i)}{x_i} \right\} \right] \right\} \\ &:= \exp \{ -V_{\mathbf{s}_1, \dots, \mathbf{s}_D}(x_1, \dots, x_D) \}, \end{aligned} \quad (1.9)$$

où  $V_{\mathbf{s}_1, \dots, \mathbf{s}_D}(\cdot)$  est la fonction exposant  $D$ -dimensionnelle.

**Définition 3 (Fonction coefficient extrémal, cf Schlather & Tawn (2003))** Soit  $X$  un processus max-stable standard sur  $\mathcal{S}$ . La fonction coefficient extrémal  $D$ -dimensionnelle est définie pour  $D \in \mathbb{N} \setminus \{0\}$ ,  $\mathbf{s}_1, \dots, \mathbf{s}_D \in \mathcal{S}$ , et  $x > 0$ , par

$$\theta_D(\mathbf{s}_1, \dots, \mathbf{s}_D) = -x \log \mathbb{P}(X(\mathbf{s}_1) \leq x, \dots, X(\mathbf{s}_D) \leq x) \in [1, D]. \quad (1.10)$$

La fonction  $\theta_D$  varie de  $\theta_D = 1$  quand les observations  $X_i$  sont parfaitement dépendantes à  $\theta_D = D$  quand les observations sont complètement indépendantes. Elle résume le degré de dépendance sur la diagonale (i.e.,  $x_1 = x_2 = \dots = x_D = x$ ). C'est pourquoi elle ne donne pas une description exhaustive de la dépendance extrême, spécialement lorsque  $D$  est grand.

Nous nous intéresserons dans le Chapitre 3 au cas bivarié. Dans ce cas là, la fonction bivariée de dépendance extrême  $\theta_2(\mathbf{s}_1, \mathbf{s}_2) : \mathbb{R}^+ \mapsto [1, 2]$  vérifie

$$\theta_2(\mathbf{s}_1, \mathbf{s}_2) = V_{\mathbf{s}_1, \mathbf{s}_2}(1, 1) = -x \log \mathbb{P}[\max\{X(\mathbf{s}_1), X(\mathbf{s}_2)\} \leq x], \quad x > 0. \quad (1.11)$$

$\theta_2(\cdot)$  a été étudiée dans Schlather & Tawn (2003). Dans la suite de cette section, nous considérerons des processus stationnaires tels que  $\theta_2(\mathbf{s}_1, \mathbf{s}_2)$  dépend uniquement de  $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ . Dans un souci de simplicité des notations, on notera  $\theta(\mathbf{h})$  pour  $\theta_2(\mathbf{s}_1, \mathbf{s}_2)$  et  $V_{\mathbf{h}}$  pour  $V_{\mathbf{s}_1, \mathbf{s}_2}$ .

Le choix du processus  $U_i$  dans l'Equation (1.8) de la représentation spectrale donnée dans le Théorème 3 conduit à des familles connues de processus max-stables plus ou moins flexibles.

On peut citer par exemple comme choix de  $U_i$  la densité Gaussienne multivariée. On considère alors un processus de Poisson  $\{\xi_i, \mathbf{s}_i\}_{i \geq 1}$  sur  $(0, \infty) \times \mathbb{R}^d$  d'intensité  $\xi^{-2} d\xi ds$ . On considère  $\phi_d(\cdot, \Sigma)$  la densité Gaussienne en dimension  $d$  de moyenne 0 et de matrice de variance-covariance  $\Sigma$ . Le modèle de Smith ou *Gaussian extreme value process* est défini par

$$X(\mathbf{s}) = \bigvee_{i=1}^{\infty} \{\xi_i \phi_d(\mathbf{s} - \mathbf{s}_i; \Sigma)\}.$$

Le choix de  $U(\mathbf{s}) = \sqrt{2\pi} \max\{0, \varepsilon(\mathbf{s})\}$  où  $\varepsilon$  est un processus Gaussien de moyenne nulle, de variance égale à 1 et de fonction de corrélation  $\rho(\cdot)$ , conduit au processus *Extremal Gaussien* ou processus de Schlather (cf Schlather (2002)). Si on choisit  $U(\mathbf{s}) = \exp\{\varepsilon(\mathbf{s}) - \gamma(\mathbf{s})\}$  où  $\varepsilon$  est un processus Gaussien intrinséquement stationnaire de semi-variogramme  $\gamma(\cdot)$ , on trouve le processus de Brown-Resnick (cf Kabluchko et al. (2009)). Le modèle *Extremal-t* (cf Opitz (2013)) est une extension du modèle de Schlather où  $U(\mathbf{s}) = c_\nu \max\{0, \varepsilon(\mathbf{s})\}^\nu$  et  $c_\nu = \sqrt{\pi} 2^{-(\nu-2)/2} \Gamma\left(\frac{\nu+1}{2}\right)^{-1}$ ,  $\nu \geq 1$ ,  $\varepsilon$  étant défini de la même façon que dans le modèle de Schlather.

Le Tableau 1.1 donne un résumé de ces processus max-stables en donnant en particulier la fonction bivariée exponent  $V(\cdot)$  et la fonction coefficient extrême  $\theta(\cdot)$ .

Modèle	$V_{\mathbf{h}}(x_1, x_2)$	$\theta(\mathbf{h})$
Smith	$\frac{1}{x_1} \Phi \left( \frac{\beta(\mathbf{h})}{2} + \frac{\log(\frac{x_2}{x_1})}{\beta(\mathbf{h})} \right) + \frac{1}{x_2} \Phi \left( \frac{\beta(\mathbf{h})}{2} + \frac{\log(\frac{x_1}{x_2})}{\beta(\mathbf{h})} \right)$ $\beta(\mathbf{h}) = \sqrt{\mathbf{h}^T \Sigma^{-1} \mathbf{h}}$ et $\Phi(\cdot)$ est la densité Gaussienne standard.	$2\Phi \left( \frac{\beta(\mathbf{h})}{2} \right)$
Schlather	$\frac{1}{2} \left( \frac{1}{x_1} + \frac{1}{x_2} \right) \left[ 1 + \sqrt{1 - \frac{2(\rho(\mathbf{h})+1)x_1x_2}{(x_1+x_2)^2}} \right]$ $\rho(\mathbf{h})$ est la fct de corrélation.	$1 + \left( \frac{1-\rho(\mathbf{h})}{2} \right)^{1/2}$
TEG	$\left( \frac{1}{x_1} + \frac{1}{x_2} \right) \left[ 1 - \frac{\alpha(\mathbf{h})}{2} \left( 1 - \sqrt{1 - \frac{2(\rho(\mathbf{h})+1)x_1x_2}{(x_1+x_2)^2}} \right) \right]$ $\alpha(\mathbf{h}) = \mathbb{E}[\mathcal{A} \cap \{\mathbf{h} + \mathcal{A}\}] / \mathbb{E}(\mathcal{A})$ , $ \cdot $ : volume de l'ensemble $\mathcal{A}$ . Si $\mathcal{A}$ est le disque de rayon $r$ , $\alpha(\mathbf{h}) \simeq (1 - \mathcal{H}/2r) \mathbb{1}_{[0,2r]}$ .	$2 - \alpha(\mathbf{h}) \left[ 1 - \left( \frac{1-\rho(\mathbf{h})}{2} \right)^{1/2} \right]$
BR	$\frac{1}{x_1} \Phi \left( \sqrt{\frac{\gamma(\mathbf{h})}{2}} + \frac{\log(\frac{x_2}{x_1})}{\sqrt{2\gamma(\mathbf{h})}} \right) + \frac{1}{x_2} \Phi \left( \sqrt{\frac{\gamma(\mathbf{h})}{2}} + \frac{\log(\frac{x_1}{x_2})}{\sqrt{2\gamma(\mathbf{h})}} \right)$ $\gamma(\cdot)$ est le semi-variogramme.	$2\Phi \left[ \left( \frac{\gamma(\mathbf{h})}{2} \right)^{1/2} \right]$
Extremal- $t$	$\frac{1}{x_1} T_{\nu+1} \left( -b\rho(\mathbf{h}) + b \left( \frac{x_2}{x_1} \right)^{1/\nu} \right) + \frac{1}{x_2} T_{\nu+1} \left( -b\rho(\mathbf{h}) + b \left( \frac{x_1}{x_2} \right)^{1/\nu} \right)$ $T_\nu$ est la f.d.r. d'une v.a. Student avec $\text{dl } \nu \geq 1$ . $b^2 = \{\nu + 1\} / \{1 - \rho^2(\mathbf{h})\}$	$2T_{\nu+1} \left( \sqrt{(\nu + 1) \frac{1-\rho(\mathbf{h})}{1+\rho(\mathbf{h})}} \right)$

TABLE 1.1 – Distribution marginale bivariée pour certains processus max-stables.

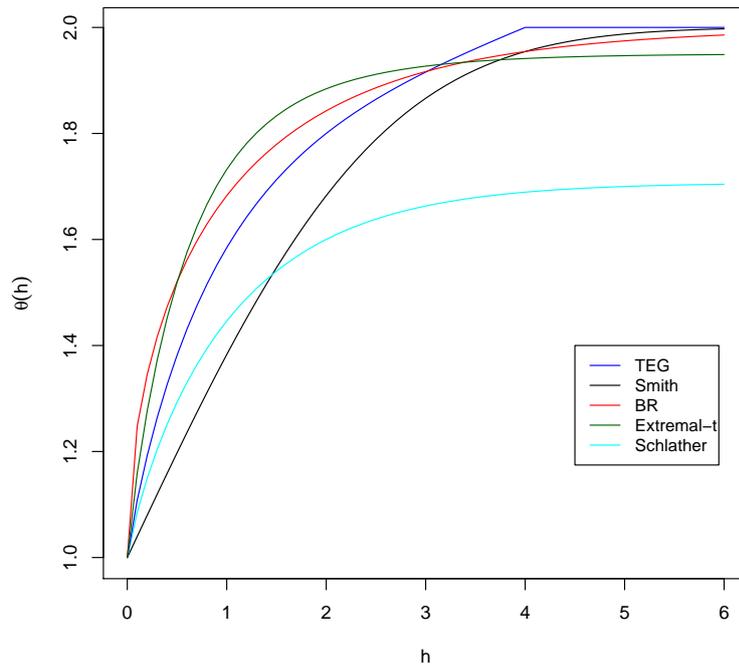


FIGURE 1.1 – Fonction Coefficient Extrémal  $\theta(\cdot)$  pour les processus max-stables isotropiques donnés dans le Tableau 1.1. La matrice de covariance vaut  $\Sigma = \text{Id}_2$  ; la fonction de corrélation est un Whittle-Matérn avec  $c_1 = c_2 = 1$  ; L'ensemble compact  $\mathcal{A}$  est un disque de rayon fixe égale à 2 ;  $\alpha(\|\mathbf{h}\|) \simeq (1 - \|\mathbf{h}\|/4)\mathbb{1}_{[0,4]}$ . Le semi-variogramme est  $\gamma(\|\mathbf{h}\|) = 2\|\mathbf{h}\|$  et les degrés de liberté sont  $\nu = 5$ . On note  $h = \|\mathbf{h}\| \geq 0$ .

Nous allons porter une attention particulière au processus Gaussien Extrémal Tronqué (*Truncated extremal Gaussian* ou TEG), introduit par Schlather (2002) car il sera un élément fondamental dans la construction du processus max-mélange proposé par Bacro et al. (2016) et présenté dans la Définition 11. Comme dans le cas extrémal Gaussien, le TEG est basé sur un processus Gaussien mais il est censuré au delà d'une certaine distance de façon à obtenir l'indépendance. Plus concrètement, on considère les processus  $U_i(\mathbf{s}) = c \max\{0, \varepsilon_i(\mathbf{s})\} \mathbb{1}_{\mathcal{A}_i}(\mathbf{s} - q_i)$  avec  $\varepsilon_i$  des copies indépendantes d'un processus Gaussien stationnaire  $\varepsilon$  (comme dans le cas du modèle extrémal Gaussien) de fonction de corrélation  $\rho(\cdot)$ ,  $\mathbb{1}_{\mathcal{A}}$  est la fonction indicatrice d'un ensemble aléatoire compact  $\mathcal{A} \subset \mathcal{S}$ ,  $\mathcal{A}_i$  sont des copies indépendantes et identiquement distribuées de  $\mathcal{A}$ , les  $q_i$  sont les points d'un processus de Poisson homogènes d'intensité 1 sur  $\mathcal{S}$ , indépendants des  $\varepsilon_i$ . Dans ce cas là, la distribution bivariée du processus TEG est donnée par

$$\begin{aligned} & \mathbb{P}(X(\mathbf{s}) \leq t_1, X(\mathbf{s} + \mathbf{h}) \leq t_2) \\ &= \exp \left\{ - \left( \frac{1}{t_1} + \frac{1}{t_2} \right) \left[ 1 - \frac{\alpha(\mathbf{h})}{2} \left( 1 - \left( 1 - 2 \frac{(\rho(\mathbf{h}) + 1)t_1 t_2}{(t_1 + t_2)^2} \right)^{1/2} \right) \right] \right\} \end{aligned} \quad (1.12)$$

où  $\alpha(\mathbf{h}) = \mathbb{E}[|\mathcal{A} \cap \{\mathbf{h} + \mathcal{A}\}|] / \mathbb{E}(|\mathcal{A}|)$ .

En particulier, quand  $\mathcal{A}$  est un disque de rayon  $r$ , on a l'approximation suivante

$$\alpha(\mathbf{h}) \approx (1 - \|\mathbf{h}\|/(2r)) \mathbb{1}_{[0, 2r]}(h).$$

Avec cette construction, la dépendance à faible distance est largement contrôlée par la fonction de corrélation  $\rho(\cdot)$  alors que la dépendance à longue distance est régulée par la géométrie de  $\mathcal{A}$ . Dans le cas bivarié, on peut choisir  $\mathcal{A}$  tel que  $\alpha(\mathbf{h}) \rightarrow 0$  quand  $\|\mathbf{h}\| \rightarrow \infty$  et donc  $\theta(\mathbf{h}) \rightarrow 2$ . Le défaut de ce modèle est que ses réalisations ne sont pas continues à cause de  $\mathcal{A}$ . Toutefois il peut être adapté pour des phénomènes avec des effets très locaux ou des phénomènes avec une indépendance complète à une certaine distance, ce que l'on peut voir habituellement dans les phénomènes spatio-temporels environnementaux. Le TEG a d'ailleurs été utilisé pour modéliser les maxima annuels de températures quotidiennes dans Davison et al. (2012) ou pour les précipitations quotidiennes dans l'Est de l'Australie dans Bacro et al. (2016). De plus, une version spatio-temporel a été ajustée sur des précipitations horaires dans Huser & Davison (2014).

La Figure 1.1 trace la fonction coefficient extrémal pour des processus max-stables isotropiques présentés et résumés dans le Tableau 1.1. On remarque que le processus de Smith couvre toute la gamme de dépendance puisque la fonction  $\theta$  varie de 1 à 2 quand  $\mathbf{h}$  tend vers  $\infty$  alors que le modèle de Schlather a une borne supérieure de 1.707 si la fonction de covariance n'admet que des valeurs positives.

Autrement dit, dans le cas du processus de Schlather, la dépendance asymptotique persiste même à des distances qui tendent vers  $\infty$ . Pour le modèle Brown-Resnick et l'Extremal- $t$ , deux sites assez fortement distants sont proche de l'indépendance. Pour le TEG, on observe l'indépendance (stricte) si la distance  $\|\mathbf{h}\| \geq 2r$ .

### 1.3 Mesures de la dépendance extrême

Quand on s'intéresse à des données non-extrêmes, un outil classique de mesure de la dépendance spatiale est le semi-variogramme, mais lorsque l'on s'intéresse aux extrêmes, il n'est plus utilisable puisqu'il n'est pas défini. Il est donc nécessaire d'introduire d'autres mesures de dépendances spatiales pour les extrêmes. Il en existe plusieurs dans la littérature, à commencer par la fonction coefficient extrême  $\theta$  pour les processus max-stable, présentée dans la Définition 3. Elle apporte beaucoup d'informations sur la dépendance extrême même si elle ne caractérise pas entièrement la distribution.

Dans Coles et al. (1999), deux mesures indépendantes du modèle ont été proposées pour permettre de distinguer les différents types de dépendance de queue :  $\chi(\mathbf{h})$  et  $\bar{\chi}(\mathbf{h})$ .

**Définition 4 (Fonction Coefficient de queue  $\chi$ ; Coles et al. (1999))**  
*Pour un processus spatial stationnaire  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  de fonction de répartition marginale univarié  $F$ , la fonction coefficient de queue  $\chi$  est définie par*

$$\chi(\mathbf{h}) = \lim_{u \rightarrow 1^-} \mathbb{P}(F(X(\mathbf{s})) > u | F(X(\mathbf{s} + \mathbf{h})) > u), \quad \mathbf{s}, \mathbf{s} + \mathbf{h} \in \mathcal{S}. \quad (1.13)$$

On peut aussi définir, pour tout  $u \in [0, 1]$

$$\chi_u(\mathbf{h}) = \frac{\mathbb{P}(F(X(\mathbf{s})) > u, F(X(\mathbf{s} + \mathbf{h})) > u)}{\mathbb{P}(F(X(\mathbf{s} + \mathbf{h})) > u)}$$

et donc

$$\chi(\mathbf{h}) = \lim_{u \rightarrow 1^-} \chi_u(\mathbf{h})$$

La fonction  $\chi_u(\mathbf{h})$  peut être vue comme la mesure de dépendance entre deux sites distants de  $\mathbf{h}$  au niveau du quantile de niveau  $u$ . On peut remarquer que pour un processus max-stable cette fonction est constante par rapport à  $u$  pour un  $\mathbf{h}$  fixé. On a alors

$$\chi_u(\mathbf{h}) = 2 - \theta(\mathbf{h})$$

comme le montre la Figure 1.2. On a aussi  $\chi(\mathbf{h}) = 2 - \theta(\mathbf{h})$  dans le cas d'un processus max-stable. Les deux fonctions offrent donc une mesure de la dépendance extrême lorsque l'on est dans le cas de processus asymptotiquement dépendants. Dans le cas de l'indépendance asymptotique, elles n'apportent pas d'information et sont donc d'un intérêt limité.

Pour mesurer la force de la dépendance de processus asymptotiquement indépendant, Coles et al. (1999) ont proposé une mesure complémentaire de la dépendance : la fonction  $\bar{\chi}(\cdot)$ .

**Définition 5 (Lower tail dependance function ; Coles et al. (1999))** Soit  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus stationnaire spatial de fonction de répartition marginale  $F$ . La fonction  $\bar{\chi}_u(\cdot)$  est définie comme

$$\bar{\chi}_u(\mathbf{h}) = \frac{2 \log \mathbb{P}(F(X(\mathbf{s})) > u)}{\log \mathbb{P}(F(X(\mathbf{s})) > u, F(X(\mathbf{s} + \mathbf{h})) > u)} - 1 \quad (1.14)$$

$$\bar{\chi}(\mathbf{h}) = \lim_{u \rightarrow 1^-} \bar{\chi}_u(\mathbf{h}).$$

$\bar{\chi}(\mathbf{h})$  varie entre -1 et 1, on est dans le cas de la dépendance asymptotique si  $\bar{\chi}(\mathbf{h}) = 1$  et dans le cas de l'indépendance asymptotique lorsque  $\bar{\chi}(\mathbf{h}) < 1$  et la valeur de  $\bar{\chi}(\mathbf{h})$  donne une mesure de la dépendance. De fait,  $|\bar{\chi}(\mathbf{h})|$  augmente avec la dépendance.

En pratique, les 2 indicateurs  $(\chi, \bar{\chi})$  doivent être considérés ensemble. Pour un  $\mathbf{h}$  donné, le couple  $(\chi_u(\mathbf{h}), \bar{\chi}_u(\mathbf{h}))$  permet de détecter l'indépendance asymptotique lorsque  $u \rightarrow 1$  d'autant que ces quantités peuvent être facilement estimées par leur version empirique.

Par ailleurs, Ledford & Tawn (1996) ont proposé un modèle qui relie la dépendance et l'indépendance asymptotiques basé sur une mesure appelée **coefficient de dépendance de queue** et notée  $\eta(\mathbf{h})$ . Considérons un processus stationnaire de marge Fréchet unitaire. La fonction  $\eta(\mathbf{h})$  détermine la vitesse de décroissance de la probabilité  $\mathbb{P}(X(\mathbf{s}) > z, X(\mathbf{s} + \mathbf{h}) > z)$  pour  $z$  grand. L'approche proposée par Ledford & Tawn (1996) repose sur l'hypothèse que la fonction de queue bivariable et la loi conditionnelle de queue se comportent respectivement comme :

$$\mathbb{P}(X(\mathbf{s}) > x, X(\mathbf{s} + \mathbf{h}) > x) \sim \mathcal{L}_{\mathbf{h}}(x)x^{-1/\eta(\mathbf{h})}, \quad x \rightarrow \infty, \quad (1.15)$$

$$\mathbb{P}(X(\mathbf{s}) > x | X(\mathbf{s} + \mathbf{h}) > x) \sim \mathcal{L}_{\mathbf{h}}(x)x^{1-1/\eta(\mathbf{h})}, \quad x \rightarrow \infty, \quad (1.16)$$

où  $\mathcal{L}_{\mathbf{h}}(\cdot)$  est une fonction univariée à variation lente. De fait, plusieurs types de dépendance peuvent apparaître avec ce modèle (cf Heffernan (2000), Ledford &

Tawn (1996) et Ledford & Tawn (1997)) :

- $\eta(h) = 1$  et  $\mathcal{L}_{\mathbf{h}}(x) > 0$  correspond à la dépendance asymptotique. Par exemple, les processus max-stables ont un  $\eta(\mathbf{h}) = 1$  pour tout  $\mathbf{h}$ .
- $\eta(h) = 1/2$  et  $\mathcal{L}_{\mathbf{h}}(x) \geq 1$  correspond à l'indépendance.
- $\eta(h) < 1$  et  $\mathcal{L}_{\mathbf{h}}(x) > 0$  correspond à l'indépendance asymptotique. Dans ce cas,  $\eta(\mathbf{h})$  mesure le degré de dépendance résiduelle. Si  $\eta(\mathbf{h}) > 1/2$  (resp.  $\eta(\mathbf{h}) < 1/2$ ) indique une association positive (resp. négative).

Sous la condition (1.15), on a  $\bar{\chi}(\mathbf{h}) = 2\eta(\mathbf{h}) - 1$ .

L'extrémogramme est la version "extrême" de la fonction de corrélation d'un processus stationnaire spatial. Il a été introduit par Davis & Mikosch (2009) dans le cas de séries temporelles et étendu au cas spatial par Cho et al. (2016).

**Définition 6 (*Extrémogramme spatial*)** *L'extrémogramme pour un processus stationnaire  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  est défini par*

$$\rho_{\mathcal{A}_1, \mathcal{A}_2}(\mathbf{h}) = \lim_{x \rightarrow \infty} \frac{\mathbb{P}\{x^{-1}X(\mathbf{s}) \in \mathcal{A}_1, x^{-1}X(\mathbf{s} + \mathbf{h}) \in \mathcal{A}_2\}}{\mathbb{P}\{x^{-1}X(\mathbf{s}) \in \mathcal{A}_1\}}, \quad \mathbf{h} \in \mathbb{R}^d, \quad (1.17)$$

où  $\mathcal{A}_1$  et  $\mathcal{A}_2$  sont des Boreliens bornés séparés de l'origine.

Si on choisit  $\mathcal{A}_1 = \mathcal{A}_2 = (1, \infty)$ , on retombe sur la fonction coefficient de dépendance de queue et  $\chi(\mathbf{h}) = \rho_{(1, \infty), (1, \infty)}(\mathbf{h})$ . Donc si  $\rho_{(1, \infty), (1, \infty)}(\mathbf{h}) = 0$  alors  $X(\mathbf{s})$  et  $X(\mathbf{s} + \mathbf{h})$  sont asymptotiquement indépendants et dans le cas  $0 < \rho_{(1, \infty), (1, \infty)}(\mathbf{h}) \leq 1$  ils sont asymptotiquement dépendants.

Finalement, nous présentons le madogramme que nous utiliserons pour faire de l'inférence sur les processus max-stables et max-mélanges dans le Chapitre 3. Il existe différentes versions de celui-ci :  $F$ -madogramme,  $F^\lambda$ -madogramme et  $\lambda$ -madogramme.

Cooley et al. (2006) ont proposé le  $F$ -madogramme comme résumé de la structure de dépendance spatiale extrême.

**Définition 7 (*F-madogramme spatial ; Cooley et al. (2006)*)** *Soit  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus spatial stationnaire de fonction de répartition marginale  $F$ . Le  $F$ -madogramme est défini par :*

$$\nu_F(\mathbf{h}) = \frac{1}{2} \mathbb{E}[|F(X(\mathbf{s})) - F(X(\mathbf{s} + \mathbf{h}))|], \quad 0 \leq \nu_F(\mathbf{h}) \leq 1/6. \quad (1.18)$$

La dépendance parfaite est atteinte lorsque  $\nu_F(\mathbf{h}) = 0$  et à l'inverse, on a l'indépendance complète lorsque  $\nu_F(\mathbf{h}) = 1/6$ . Pour les processus max-stables, il existe une relation entre  $F$ -madogramme et fonction coefficient extrémal  $\theta(\cdot)$  :

$$\nu_F(\mathbf{h}) = \frac{1}{2} - \frac{1}{\theta(\mathbf{h}) + 1}. \quad (1.19)$$

Le  $F$ -madogramme peut facilement être estimé non-paramétriquement par :

$$\hat{\nu}_F(\mathbf{h}) = \frac{1}{2T} \sum_{i=1}^T [|\hat{F}(X_i(\mathbf{s})) - \hat{F}(X_i(\mathbf{s} + \mathbf{h}))|], \quad (1.20)$$

où  $\hat{F}$  est la fonction de répartition empirique. Lorsque le processus est isotrope, il est préférable d'utiliser la version "binned" :

$$\hat{\nu}_F(\mathbf{h}) = \frac{1}{2T|\mathcal{B}_{\mathbf{h}}|} \sum_{(s, s+\mathbf{h}) \in \mathcal{B}_{\mathbf{h}}} \sum_{i=1}^T [|\hat{F}(X_i(s)) - \hat{F}(X_i(s + \mathbf{h}))|], \quad (1.21)$$

où  $\mathcal{B}_{\mathbf{h}}$  est l'ensemble des paires  $s, s + \mathbf{h}$ ,  $\mathbf{h} \in \mathbb{R}^d$  avec une certaine tolérance  $\Delta$ .

Le  $F^\lambda$ -madogramme a été introduit par Bel et al. (2008) et fournit une généralisation du  $F$ -madogramme.

**Définition 8 (Spatial  $F^\lambda$ -madogramme ; Bel et al. (2008))** Pour un processus spatial stationnaire  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  de fonction de répartition marginale  $F$  et pour tout  $\lambda > 0$ , le  $F^\lambda$ -madogramme est défini par :

$$\nu_{F^\lambda}(\mathbf{h}) = \frac{1}{2} \mathbb{E} [ |F^\lambda\{X(\mathbf{s})\} - F^\lambda\{X(\mathbf{s} + \mathbf{h})\}| ]. \quad (1.22)$$

Comme la fonction  $\theta(\cdot)$ , les madogrammes  $\nu_F(\cdot)$  et  $\nu_{F^\lambda}(\cdot)$  ne caractérisent pas complètement la dépendance spatiale du processus puisqu'ils ne se concentrent que sur  $\mathbb{P}(X(\mathbf{s}) \leq x_1, X(\mathbf{s} + \mathbf{h}) \leq x_2)$  avec  $x_1 = x_2 = x$ . Naveau et al. (2009) ont introduit le  $\lambda$ -madogramme pour résoudre ce problème.

**Définition 9 (Spatial  $\lambda$ -madogramme ; Naveau et al. (2009))** Soit  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus spatial stationnaire de fonction de répartition marginale  $F$ . Le  $\lambda$ -madogramme est défini par

$$\nu_\lambda(\mathbf{h}) = \frac{1}{2} \mathbb{E} [ |F^\lambda(X(\mathbf{s})) - F^{1-\lambda}(X(\mathbf{s} + \mathbf{h}))| ], \quad \lambda \in [0, 1]. \quad (1.23)$$

L'idée est que, en faisant varier  $\lambda$ , on se concentre sur  $\mathbb{P}(X(\mathbf{s}) \leq x_1, X(\mathbf{s} + \mathbf{h}) \leq x_2)$  avec  $x_1 = \lambda x$  et  $x_2 = (1 - \lambda)x$ . On obtient ainsi de l'information sur  $V_h(x_1, x_2)$  pour  $x_1 \neq x_2$ . Dans Guillou et al. (2014), les auteurs ont proposé un mado-gramme adapté pour le cas des modèles asymptotiquement indépendants que l'on va développer dans la prochaine section.

## 1.4 Modèles pour l'indépendance asymptotique

Les processus Gaussiens sont les premiers modèles d'indépendance asymptotique qui apparaissent naturellement. Evidemment, tout processus provenant d'une transformation croissante des marginales d'un processus Gaussien est un processus asymptotiquement indépendant. Wadsworth & Tawn (2012) ont proposé une nouvelle classe de modèles asymptotiquement indépendants, plus flexibles que les processus Gaussiens : les processus inverse max-stables. Ils sont obtenus par une "inversion" d'un processus max-stable.

### Définition 10 (*Inverse max-stable processus, Wadsworth & Tawn (2012)*)

Soit  $\{X'(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus max-stable standard de fonction coefficient extrémal  $\theta'(\mathbf{h})$  et de fonction de répartition bivariée  $F_{X'}(\cdot)$ . On définit la transformation

$$\begin{aligned} t(x) : (0, \infty) &\mapsto (0, \infty) \\ x &= -1/\log[1 - \exp\{-1/x\}] \end{aligned}$$

Alors le processus  $Y(s) = t(X'(\cdot))$  est un processus inverse max-stable. Ses marginales sont Fréchet unitaires et vérifient (1.15) avec  $\eta(\mathbf{h}) = 1/\theta'(\mathbf{h})$ , le processus est donc asymptotiquement indépendant.

Chaque modèle max-stable peut donc être transformé en un processus asymptotiquement indépendant via cette transformation. De plus, la distribution bivariée du processus inverse max-stable s'exprime en fonction de la structure de dépendance du processus max-stable original :

$$\begin{aligned} F_Y(x_1, x_2) &= -1 + \exp(-x_1^{-1}) + \exp(-x_2^{-1}) \\ &\quad + \exp\{-V_{X'}\{t(x_1), t(x_2)\}\}, \end{aligned} \tag{1.24}$$

où  $V_{X'}$  est la mesure exponent de la fonction de distribution bivariée  $X'$  entre  $\mathbf{x}$  et  $\mathbf{x} + \mathbf{h}$ .

Dans le Chapitre 3 et dans la suite, on notera  $\theta_Y$  et  $V_Y$  les fonctions  $\theta'$  et  $V'$ . Plus

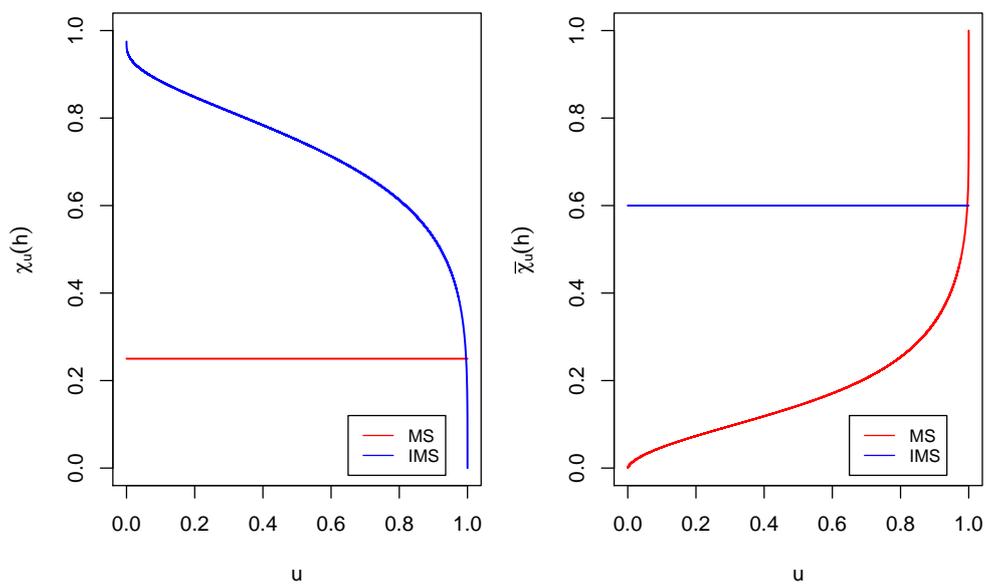


FIGURE 1.2 –  $\chi_u(h)$  et  $\bar{\chi}_u(h)$  pour un processus max-stable avec  $\theta_X(h) = 1.75$  (rouge) et pour un processus inverse max-stable avec  $\theta_Y(h) = 1.25$  (bleu) avec  $h = \|\mathbf{h}\|$ .

généralement, on notera  $\theta_X$  la fonction coefficient extrémal du processus  $X$ .

La Figure 1.2 représente les fonctions  $\chi_u(\mathbf{h})$  et  $\bar{\chi}_u(\mathbf{h})$  pour des processus max-stables et inverse max-stables. Le processus max-stable a une fonction  $\chi_u$  constante égale à  $2 - \theta_X(\mathbf{h})$  et une fonction  $\bar{\chi}_u(\mathbf{h})$  croissante vers 1. A l'inverse, le processus inverse max-stable a une fonction  $\bar{\chi}$  constante égale à  $2/\theta_Y(\mathbf{h}) - 1$  et une fonction  $\chi_u$  de la forme

$$\chi_u(\mathbf{h}) = 2 - \frac{\log(-1 + 2u + (1 - u)^{\theta_Y(\mathbf{h})})}{\log(u)}, \quad u \in [0, 1] \quad (1.25)$$

avec  $\theta_Y(\mathbf{h}) \in [1, 2]$ , décroissante vers 0.

## 1.5 Modèles max-mélanges

Les processus max-mélanges, proposés par Wadsworth & Tawn (2012), présentent la possibilité d'avoir plusieurs types de dépendances dans une seule modélisation. Ils sont une alternative intéressante aux processus max-stable et asymptotiquement indépendant, spécialement lorsque les régions considérées sont étendues. Ils sont obtenus en “mélangeant” un processus max-stable et un processus asymptotiquement indépendant (mais nous considérerons uniquement le cas où ce dernier est un processus inverse max-stable).

### **Définition 11 (Processus spatial max-mélange, Wadsworth & Tawn (2012))**

Soit  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus max-stable standard de coefficient extrémal  $\theta_X(\mathbf{h})$ , de fonction de répartition marginale bivariée  $F_X$ , et soit  $\{Y(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus inverse max-stable de coefficient  $\eta(\mathbf{h}) = 1/\theta_Y(\mathbf{h})$ , de fonction de répartition marginale bivariée  $F_Y$ . On suppose que  $X$  et  $Y$  sont indépendants. Alors le processus

$$Z(\mathbf{s}) = \max\{aX(\mathbf{s}), (1 - a)Y(\mathbf{s})\}, \quad 0 \leq a \leq 1, \quad (1.26)$$

est un processus max-mélange de marginale Fréchet unitaire.

La Figure 1.3 représente 4 réalisations d'un processus max-mélange isotropique, fruit d'un mélange entre un processus de Smith et un processus inverse extremal- $t$  pour différentes valeurs du paramètre de mélange. On voit que le cas  $a = 1$  correspond uniquement au processus max-stable alors que  $a = 0$  correspond uniquement à la partie asymptotiquement indépendante i.e., l'inverse extremal- $t$  dans ce cas.

On peut obtenir directement la distribution de  $(Z(\mathbf{s}), Z(\mathbf{s} + \mathbf{h}))$  pour  $0 < a < 1$  en utilisant l'indépendance entre  $X(\mathbf{s})$  et  $Y(\mathbf{s})$  (voir Wadsworth & Tawn (2012); Bacro et al. (2016)) :

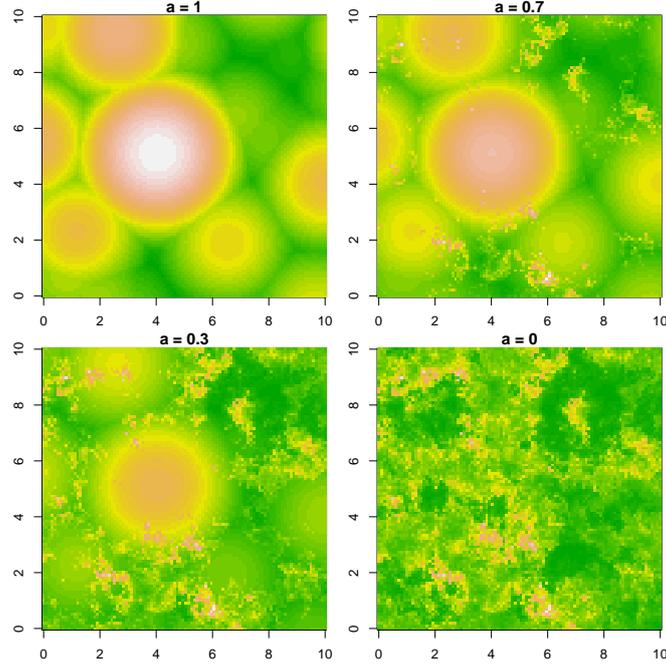


FIGURE 1.3 – Simulations d’un processus max-mélange (Définition 11) sur une échelle logarithmique pour différentes valeurs du paramètre de mélange  $a \in \{1, 0.7, 0.3, 0\}$  avec  $X$  un processus isotropique Smith de paramètre  $\Sigma = \text{Id}_2$  et  $Y$  un processus isotropique inverse max-stable *extremal- $t$*  de degrés de liberté  $v = 2$  et de fonction de corrélation  $\rho(\mathbf{h}) = \exp(-\mathbf{h}/\phi)$  de portée  $\phi = 1$ .

$$F_Z(z_1, z_2) = F_X\left(\frac{z_1}{a}, \frac{z_2}{a}\right) F_Y\left(\frac{z_1}{1-a}, \frac{z_2}{1-a}\right), \quad (1.27)$$

$$\mathbb{P}(Z(\mathbf{s}) > z | Z(\mathbf{s} + \mathbf{h}) > z) \sim a \{2 - \theta_X(\mathbf{h})\} + (1-a)^{\theta_Y(\mathbf{h})} \frac{\mathcal{L}_{\mathbf{h}}\{z/(1-a)\}}{z^{\theta_Y(\mathbf{h})-1}}, \quad z \rightarrow \infty. \quad (1.28)$$

On peut aussi obtenir les fonctions  $\chi$ ,  $\chi_u$  et  $\bar{\chi}_u$  :

$$\chi(\mathbf{h}) = \lim_{z \rightarrow \infty} \mathbb{P}(Z(s) > z | Z(s + \mathbf{h}) > z) = a \{2 - \theta_X(\mathbf{h})\}, \quad (1.29)$$

$$\chi_u(\mathbf{h}) = 2 - a\theta_X(\mathbf{h}) - \frac{\log\left(-1 + 2u^{(1-a)} + (1 - u^{(1-a)})^{\theta_Y(\mathbf{h})}\right)}{\log(u)}, \quad (1.30)$$

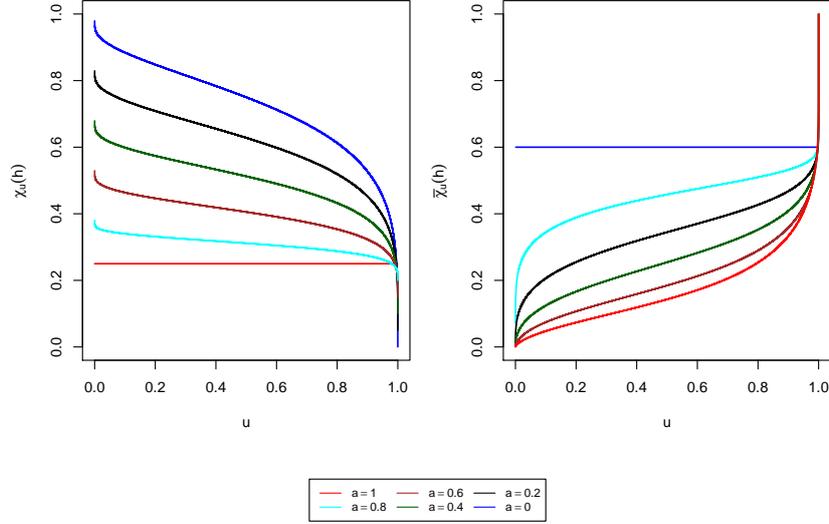


FIGURE 1.4 –  $\chi_u(h)$  et  $\bar{\chi}_u(h)$  pour un processus max-mélange de fonctions coefficient extrémal  $(\theta_X(h), \theta_Y(h)) = (1.75, 1.25)$  et différentes valeurs de  $a \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$  et  $h = \|\mathbf{h}\|$ .

$$\bar{\chi}_u(\mathbf{h}) = \frac{2 \log(1-u)}{\log \left\{ 1 - 2u + \left[ (u^{a\theta_X(\mathbf{h})}) \left( -1 + 2u^{(1-a)} + (1-u^{(1-a)})^{\theta_Y(\mathbf{h})} \right) \right] \right\}} - 1. \quad (1.31)$$

La Figure 1.4 représente  $\chi_u(\mathbf{h})$  et  $\bar{\chi}_u(\mathbf{h})$  pour un processus max-mélange tel que  $(\theta_X(\mathbf{h}), \theta_Y(\mathbf{h})) = (1.75, 1.25)$  pour différentes valeurs de  $a$ . Evidemment pour  $a \in (0, 1)$ ,  $\chi_u(\mathbf{h})$  décroît vers  $a(2 - \theta_X(\mathbf{h}))$  et  $\bar{\chi}_u(\mathbf{h})$  croît vers 1.

Si on part de l'Equation (1.28), on voit que s'il existe un  $h' = \inf\{h : \theta_X(h) = 2\}$  fini alors le processus  $Z$  est asymptotiquement dépendant jusqu'à la distance  $h'$  puis devient asymptotiquement indépendant pour les distances plus longues, contrairement aux processus max-stables où l'on ne pourrait avoir que de l'indépendance à distances plus longues. Parmi les modèles max-stables connus, seul le TEG (voir Table 1.1) a un  $h'$  fini. Si on suppose que  $\mathcal{A}$  est un disque de rayon  $r$  et on utilise un TEG pour la partie max-stable d'un processus max-mélange, la fonction  $\chi$  vaut alors (cf Bacro et al. (2016)) :

$$\chi(\mathbf{h}) = \lim_{z \rightarrow \infty} \mathbb{P}(Z(\mathbf{s}) > z | Z(\mathbf{s} + \mathbf{h}) > z) = a \{2 - \theta_X(\mathbf{h})\} \quad (1.32)$$

Un exemple de modèle qui combine les 3 types de dépendance est le modèle  $M_1$  décrit ci-dessous :

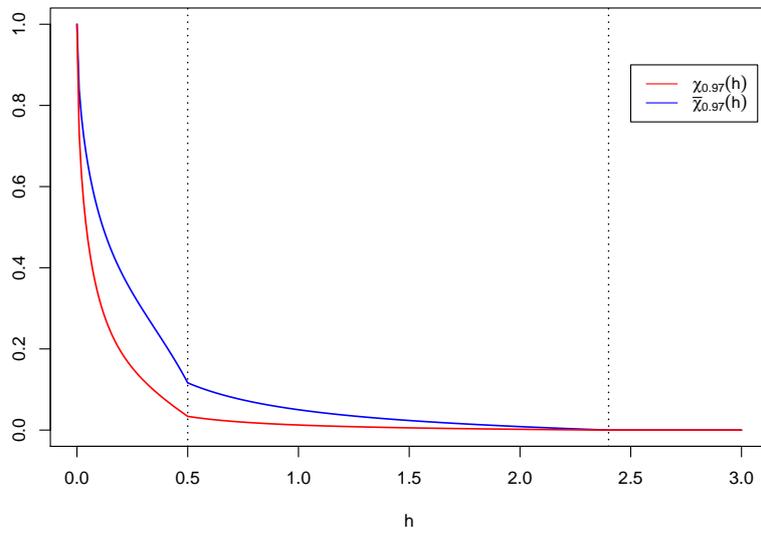


FIGURE 1.5 –  $\chi_{0.97}(h)$  et  $\bar{\chi}_{0.97}(h)$  pour le modèle  $M_1$  (mélange entre un TEG et inverse TEG) avec  $\phi_X = 0.1$ ,  $r_X = 0.25$ ,  $\phi_Y = 0.75$ ,  $r_Y = 1.2$ ,  $\kappa_X = \kappa_Y = 1$  et  $a = 0.5$ . Les lignes verticales en pointillés correspondent aux diamètres des disques dans les processus TEG et on a  $h = \|\mathbf{h}\|$ .

Le modèle  $M_1$  est un modèle max-mélange dans lequel  $X$  est un TEG où  $\mathcal{A}$  est un disque de rayon  $r_X = 0,25$  et la fonction de corrélation  $\rho_X(\mathbf{h}) = \exp\{- (\|\mathbf{h}\|/\phi_X)^{\kappa_X}\}$  de paramètres  $\phi_X = 0.1$  et  $\kappa_X = 1$ . Le processus asymptotiquement indépendant  $Y$  est un inverse TEG avec  $\mathcal{A}_Y$  un disque de rayon  $r_Y = 0.75$  de fonction de corrélation isotropique  $\rho_Y(\mathbf{h}) = \exp\{- (\|\mathbf{h}\|/\phi_Y)^{\kappa_Y}\}$  avec  $\phi_Y = 1.2$  et  $\kappa_Y = 1$ .

La Figure 1.5 représente les deux fonctions  $\chi_{0.97}$  et  $\bar{\chi}_{0.97}$  pour le modèle  $M_1$ . Dans ce cas-là, on considère  $h = \|\mathbf{h}\|$  on constate que le modèle est asymptotiquement dépendant pour  $h < 2r_X = 0.5$  (puisque  $\chi_{0.97} > 0$ ), asymptotiquement indépendant pour  $h$  compris entre  $2r_X = 0.5$  et  $2r_Y = 1.5$  (puisque  $1 > \bar{\chi}_{0.97} > 0$ ) et asymptotiquement indépendant pour  $h > 2r_Y = 1.5$  (puisque  $\bar{\chi}_{0.97} = 0$ ).

Dans [A-18], nous présentons un test basé sur la vraisemblance composite pour la valeur du paramètre de mélange  $a$  d'un processus max-mélange. Nous avons proposé dans [S-22] une méthode d'estimation des paramètres d'un processus max-mélange en utilisant une méthode des moindres carrés non linéaires basée sur le madogramme (cf Section 3.1). Dans [A-19], nous proposons une méthode d'estimation du paramètre de mélange  $a$  sans poser de modèle paramétrique sur les processus  $X$  et  $Y$  intervenant dans la construction du processus max-mélange. Cette procédure est basée sur le  $F^\lambda$ -Madogramme (cf Section 3.2). Dans [A-12], nous avons proposé un nouvel outil pour quantifier la dépendance mais aussi pour estimer les paramètres de processus max-stable. Dans [A-17], nous avons étudié des mesures de risques spatiales pour les processus max-stables et max-mélanges.

## 1.6 Modèles max-stables spatio-temporels

Récemment, les processus max-stables ont été étendus au cas spatio-temporel. Par la suite  $\{X(\mathbf{s}, t) : (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}\}$ ,  $\mathcal{S} \times \mathcal{T} \subset \mathbb{R}^d \times \mathbb{R}^+$  est un processus spatio-temporel max-stable (généralement  $d = 2$ ). Les indices d'espace  $\mathbf{s}$  et de temps  $t$  appartiennent à des ensembles  $\mathcal{S}$  et  $\mathcal{T}$ . On notera  $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2 \in \mathbb{R}^2$  la distance spatiale et  $l = t_1 - t_2 \in \mathbb{R}$  la distance temporelle. Nous allons distinguer les processus max-stables spatio-temporels suivant qu'ils ont une séparabilité spectrale ou non.

## Modèles spatio-temporels max-stables sans séparabilité spectrale

### **Théorème 5 (Représentation spectrale pour les processus max-stables spatio-temporels, De Haan (1984))**

Soit  $X(\mathbf{s}, t)$  un processus max-stable spatio-temporel standard sur  $\mathcal{S} \times \mathcal{T}$ . Alors il existe  $\{\xi_i\}_{i \geq 1}$  des réalisations d'un processus de Poisson sur  $(0, \infty)$  d'intensité  $\xi^{-2} d\xi$  et une suite  $\{U_i(\mathbf{s}, t)\}_{i \geq 1}$  de copies indépendantes et identiquement distribuées d'un processus spatio-temporel  $\{U(\mathbf{s}, t), (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}\}$  avec  $\mathbb{E}\{U(\mathbf{s}, t)\} < \infty$  pour tout  $(\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}$ , et  $U(\mathbf{s}, t) \geq 0$ , qui sont aussi indépendants de  $\xi_i$ , tels que

$$X(\mathbf{s}, t) \stackrel{D}{=} \bigvee_{i=1}^{\infty} \xi_i U_i(\mathbf{s}, t), \quad (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}. \quad (1.33)$$

Ici, standard signifie que les marginales sont Fréchet unitaires, c'est-à-dire,  $F(x) := \mathbb{P}(X(\mathbf{s}, t) \leq x) = \exp\{-x^{-1}\}$ ,  $x > 0$ .

Comme dans le cas des processus max-stables spatiaux, cette représentation est particulièrement utile pour la construction et la simulation de processus max-stables spatio-temporels. Le théorème suivant donne leur distribution  $D$ -dimensionnelle.

**Théorème 6 (Distribution  $D$ -dimensionnelle)** Pour  $D \in \mathbb{N} \setminus \{0\}$ ,  $\mathbf{s}_1, \dots, \mathbf{s}_D \in \mathcal{S}$ ,  $t_1, \dots, t_D \in \mathcal{T}$  et  $x_1, \dots, x_D > 0$ , la distribution  $D$ -dimensionnelle du processus max-stable spatio-temporel  $X$  est donnée par

$$\begin{aligned} \mathbb{P}(X(\mathbf{s}_1, t_1) \leq x_1, \dots, X(\mathbf{s}_D, t_D) \leq x_D) &= \mathbb{P} \left\{ \xi_i \bigvee_{j=1}^D \frac{U_i(\mathbf{s}_j, t_i)}{x_j} \leq 1, \forall i = 1, 2, \dots \right\} \\ &= \exp \left\{ -\mathbb{E} \left( \bigvee_{j=1}^D \frac{U(\mathbf{s}_j, t_j)}{x_j} \right) \right\}. \end{aligned} \quad (1.34)$$

Toutes les distributions multivariées sont des distributions des extrêmes multivariés de marginale Fréchet. En particulier, la loi bivariée d'un processus max-stable spatio-temporel défini en (1.33) peut être exprimée en fonction de la fonction exponent spatio-temporelle bivariée sous jacente :

$$\begin{aligned} -\log F(x_1, x_2) &:= -\log \mathbb{P}[X(\mathbf{s}_1, t_1) \leq x_1, X(\mathbf{s}_2, t_2) \leq x_2] \\ &= V_{\mathbf{s}_1, t_1; \mathbf{s}_2, t_2}(x_1, x_2). \end{aligned} \quad (1.35)$$

Par la suite, nous considérerons des processus stationnaires tels que  $V_{\mathbf{s}_1, t_1; \mathbf{s}_2, t_2}$  dépend uniquement de  $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$  et  $l = t_1 - t_2$ . Donc  $V_{\mathbf{s}_1, t_1; \mathbf{s}_2, t_2}$  sera remplacé par

$V_{\mathbf{h},l}$ .

Nous pouvons étendre les mesures de la dépendance dans le cadre spatial au cadre spatio-temporel.

**Définition 12 (Mesures de la dépendance spatio-temporelle)** *Pour un processus stationnaire max-stable spatio-temporel  $\{X(\mathbf{s}, t) : (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}\}$  de fonction de répartition marginale univariée Fréchet standard  $F$ , on définit :*

1. **Fonction Coefficient extrémal spatio-temporel**

$$\theta(\mathbf{h}, l) = -\log \mathbb{P}(X(\mathbf{s}, t) \leq x, X(\mathbf{s} + \mathbf{h}, t + l) \leq x) \in [1, 2], \quad x > 0. \quad (1.36)$$

2.  **$\chi$  spatio-temporel**

$$\chi_u(\mathbf{h}, l) = 2 - \frac{2 \log \mathbb{P}\{F(X(\mathbf{s}, t)) < u, F(X(\mathbf{s} + \mathbf{h}, t + l)) < u\}}{\log \mathbb{P}\{F(X(\mathbf{s} + \mathbf{h}, t + l)) < u\}} \quad (1.37)$$

et  $\chi(\mathbf{h}, l) = \lim_{u \rightarrow 1^-} \chi_u(\mathbf{h}, l)$ ,  $u \in [0, 1]$ . Comme dans le cas spatial, on a la relation :  $\chi(\mathbf{h}, l) = 2 - \theta(\mathbf{h}, l)$ .

3.  **$\bar{\chi}$  spatio-temporel**

$$\bar{\chi}_u(\mathbf{h}, l) = \frac{2 \log \mathbb{P}(F(X(\mathbf{s}, t)) > u)}{\log \mathbb{P}(F(X(\mathbf{s}, t)) > u, F(X(\mathbf{s} + \mathbf{h}, t + l)) > u)} - 1, \quad (1.38)$$

et  $\bar{\chi}(\mathbf{h}, l) = \lim_{u \rightarrow 1^-} \bar{\chi}_u(\mathbf{h}, l)$ .

4.  **$F$ -madogramme spatio-temporel**

$$\nu_F(\mathbf{h}, l) = \frac{1}{2} \mathbb{E}[|F(X(\mathbf{s}, t)) - F(X(\mathbf{s} + \mathbf{h}, t + l))|] \in [0, 1/6]. \quad (1.39)$$

Le  $F^\lambda$ -madogramme et le  $\lambda$ -madogramme sont définis de manière similaire (cf Définitions 7 et 9).

## Exemple de processus max stables paramétriques spatio-temporels sans séparabilité spectrale

Nous présentons ici un exemple de processus max-stable spatio-temporel que nous utiliserons dans la Section 3.3 et qui vérifie la représentation spectrale de l'Equation (1.33).

## Brown Resnick spatio-temporel

Davis et al. (2013a) ont étendu le modèle de Brown-Resnick au cas spatio-temporel. Un processus stationnaire spatio-temporel Brown-Resnick  $X$  admet la représentation spectrale suivante :

$$X(\mathbf{s}, t) = \bigvee_{i=1}^{\infty} \xi_i \exp(\varepsilon_i(\mathbf{s}, t) - \gamma(\|\mathbf{s}\|, t)), \quad (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}, \quad (1.40)$$

où  $\|\cdot\|$  est la norme Euclidienne sur  $\mathbb{R}^2$ ,  $\{\xi_i\}_{i \geq 1}$  sont des réalisations d'un processus de Poisson  $(0, \infty)$  d'intensité  $\xi^{-2} d\xi$ , les processus  $\{\varepsilon_i(\mathbf{s}, t) : (\mathbf{s}, t) \in (\mathcal{S} \times \mathcal{T})\}$  sont des copies indépendantes d'un processus Gaussien  $\{\varepsilon(\mathbf{s}, t)\}$  d'incrément stationnaires,  $\varepsilon(\mathbf{0}, 0) = 0$ ,  $\mathbb{E}[\varepsilon(\mathbf{s}, t)] = 0$  et de fonction de covariance

$$\text{Cov}(\varepsilon(\mathbf{s}_1, t_1), \varepsilon(\mathbf{s}_2, t_2)) = \gamma(\|\mathbf{s}_1\|, t_1) + \gamma(\|\mathbf{s}_2\|, t_2) - \gamma(\|\mathbf{s}_1 - \mathbf{s}_2\|, |t_1 - t_2|),$$

pour tout  $(\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2) \in \mathcal{S} \times \mathcal{T}$ . La fonction de dépendance  $\gamma$  appelée semi-variogramme spatio-temporel du processus  $\{\varepsilon(\mathbf{s}, t)\}$ , est non négative et conditionnellement définie négative, i.e., pour tout  $k \in \mathbb{N}$ ,  $(\mathbf{s}_1, t_1), \dots, (\mathbf{s}_k, t_k) \in \mathcal{S} \times \mathcal{T}$  et  $a_1, \dots, a_k \in \mathbb{R}$ ,

$$\sum_{i=1}^k \sum_{j=1}^k a_i a_j \gamma(\|\mathbf{s}_i - \mathbf{s}_j\|, |t_i - t_j|) \leq 0, \quad \sum_{i=1}^k a_i = 0.$$

Le processus  $X(\mathbf{s}, t)$  défini dans (1.40) est complètement caractérisé par sa fonction de dépendance  $\gamma$ . En géostatistique,  $\gamma$  est donné par :

$$\gamma(\|\mathbf{h}\|, |l|) = \frac{1}{2} \text{Var}(\varepsilon(\mathbf{s}_1, t_1) - \varepsilon(\mathbf{s}_2, t_2)).$$

Comme on suppose que  $\gamma$  dépend uniquement de la norme de  $\mathbf{s}_1 - \mathbf{s}_2$ , le processus associé est donc isotropique. Pour  $x_1, x_2 > 0$ , la fonction de répartition bivariable  $F$  de  $(X(\mathbf{s}_1, t_1), X(\mathbf{s}_2, t_2))$  est donnée par

$$\begin{aligned} -\log F(x_1, x_2) &= \frac{1}{x_1} \Phi \left( \sqrt{\frac{\gamma(\|\mathbf{h}\|, |l|)}{2}} + \frac{\log\left(\frac{x_2}{x_1}\right)}{\sqrt{2\gamma(\|\mathbf{h}\|, |l|)}} \right) \\ &\quad + \frac{1}{x_2} \Phi \left( \sqrt{\frac{\gamma(\|\mathbf{h}\|, |l|)}{2}} + \frac{\log\left(\frac{x_1}{x_2}\right)}{\sqrt{2\gamma(\|\mathbf{h}\|, |l|)}} \right), \end{aligned} \quad (1.41)$$

où  $\Phi$  est la distribution Gaussienne standard. La fonction coefficient extrémal du modèle est  $\theta(\|\mathbf{h}\|, |l|) = 2\Phi\left\{\sqrt{\gamma(\|\mathbf{h}\|, |l|)/2}\right\}$ .

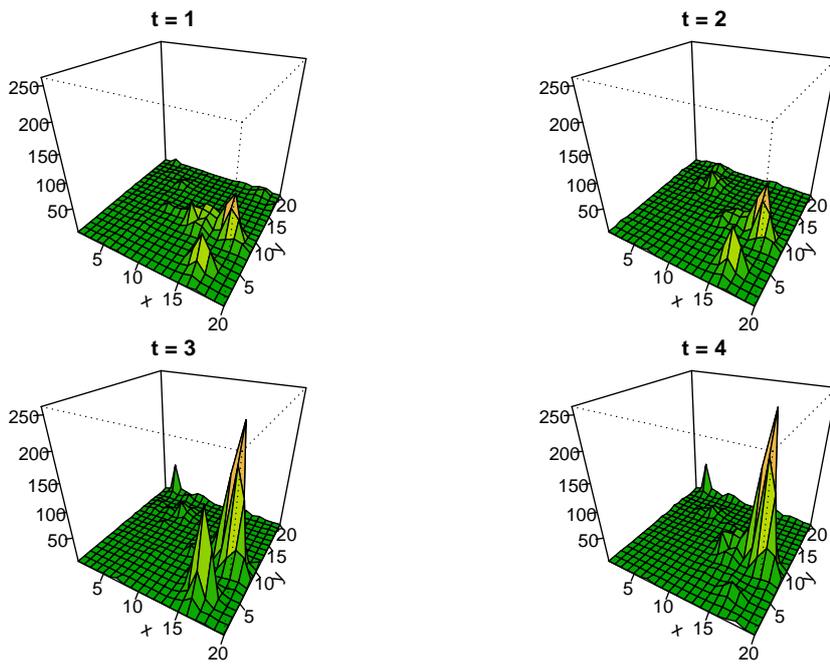


FIGURE 1.6 – Simulation d'un processus spatio-temporel max-stable Brown-Resnick avec un semi-variogramme FBM (*Fractal Brownian Motion*), i.e.,  $\gamma(h, l') = 0.8h^{1.5} + 0.4l'$  à 4 points temporels consécutifs.

La Figure 1.6 représente une simulation du processus spatio-temporel max-stable Brown-Resnick avec un semi-variogramme FBM (*Fractal Brownian Motion*), i.e.,  $\gamma(\|h\|, l') = 0.8h^{1.5} + 0.4l'$ .

Ce modèle a été utilisé par Buhl et al. (2019) pour modéliser le comportement extrême sur des données de précipitations par radar sur la région de la Floride. Ils ont proposé une nouvelle méthode basée sur l'extrémogramme pour estimer les paramètres du modèle.

## Modèles spatio-temporels max-stables avec séparabilité spectrale

L'avantage des modèles proposés précédemment est qu'ils sont construits à partir de processus max-stables spatiaux largement étudiés dans la littérature. Il existe aussi une large littérature sur des fonctions de corrélation spatio-temporelle pour des processus Gaussien, ce qui ouvre la porte à une importante diversité de comportement et de modélisations spatio-temporelles. Toutefois, un problème important est qu'il n'y a pas d'interaction entre les composantes spatiales et temporelles dans la fonction de dépendance sous-jacente. On peut donc voir le temps comme une dimension spatiale supplémentaire puisque le temps et l'espace appartiennent à la même famille de modèles. Une alternative a été proposée par Embrechts et al. (2016) avec des modèles max-stables spatio-temporels avec séparabilité spectrale.

**Définition 13** (*Modèles max-stables spatio-temporels avec séparabilité spectrale, Embrechts et al. (2016)*) La classe de modèles max-stables spatio-temporels avec séparabilité spectrale s'obtient en insérant la décomposition spectrale suivante dans (1.33) :

$$X(\mathbf{s}, t) = \bigvee_{i=1}^{\infty} \xi_i U_t(Q_i) U_{\mathcal{R}(t, Q_i)\mathbf{s}}(W_i), \quad (1.42)$$

où

- $\{\xi_i, Q_i, W_i\}_{i \geq 1}$  sont des points d'un processus de Poisson sur  $(0, \infty) \times E_1 \times E_2$ , et d'intensité  $\xi^{-2} d\xi \times \mu_1(dq) \times \mu_2(dw)$  pour des espaces Polonais  $(E_1, \mathcal{E}_1, \mu_1)$  et  $(E_2, \mathcal{E}_2, \mu_2)$ ,
- la fonction spectrale  $U_t : E_1 \rightarrow (0, \infty)$  est mesurable telle que  $\int_{E_1} U_t(q) \mu_1(dq) = 1$  pour tout  $t \in \mathcal{T}$  et contribue à la dynamique temporelle du processus, alors que la fonction spectrale  $U_s : E_2 \rightarrow (0, \infty)$  est mesurable telle que

$\int_{E_2} U_{\mathbf{s}}(w) \mu_2(dw) = 1$  pour tout  $\mathbf{s} \in \mathcal{S}$  et guide la forme des motifs spatiaux,

- Les opérateurs  $\mathcal{R}(t, q)$  sont bijectifs de  $\mathcal{S}$  dans  $\mathcal{S}$  pour tout  $(t, q) \in \mathcal{T} \times E_1$  et décrivent comment les motifs spatiaux bougent dans l'espace.

La construction (1.42) permet de traiter séparément les aspects spatiaux et temporels. Ainsi la procédure d'estimation peut être simplifiée en estimant séparément les paramètres d'espace et de temps.

Différents exemples de modèles issus de l'Equation (1.42) ont été proposés par Embrechts et al. (2016). Toutefois, les auteurs de cet article se sont concentrés sur un cas particulier Markovien avec une représentation max-autoregressive et un choix naturel de processus max-stable spatio-temporel qui satisfait l'équation suivante :

$$X(\mathbf{s}, t) = \max\{\delta X(\mathbf{s} - \boldsymbol{\tau}, t - 1), (1 - \delta)H(\mathbf{s}, t)\}, \quad (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}, \quad (1.43)$$

où le paramètre  $\delta \in (0, 1)$  mesure l'influence du passé, le paramètre  $\boldsymbol{\tau} \in \mathbb{R}^2$  représente en quelque sorte une direction spécifique de propagation/contagion dans l'espace et  $H(\mathbf{s}, t)$  est une suite de processus max-stables spatiaux indépendants et identiquement distribués. On peut voir ce modèle comme une extension du processus MARMA(1,0) (pour plus de détails, voir Davis & Resnick (1989)).

Par la suite et dans la Section 3.3, nous allons nous concentrer sur les processus qui vérifient l'Equation (1.43). On note  $V_{\mathbf{0}, \mathbf{h} - l\boldsymbol{\tau}}$  la mesure exponentielle qui caractérise la distribution spatiale du processus  $H(\mathbf{s}, t)$ , alors la fonction de distribution bivariable  $F$  de  $(X(0, \mathbf{0}), X(l, \mathbf{h}))$  est donnée, pour  $x_1, x_2 > 0$  par

$$-\log F(x_1, x_2) = V_{\mathbf{0}, \mathbf{h} - l\boldsymbol{\tau}} \left( x_1, \frac{x_2}{\delta^l} \right) + \frac{1 - \delta^l}{x_2}. \quad (1.44)$$

On peut ainsi déduire la fonction coefficient extrémal spatio-temporel en posant  $x_1 = x_2 = x$  :

$$\theta(\mathbf{h}, l) = V_{\mathbf{0}, \mathbf{h} - l\boldsymbol{\tau}} (1, \delta^{-l}) + 1 - \delta^l. \quad (1.45)$$

On voit donc que le temps et l'espace ne sont pas complètement séparés dans la fonction coefficient extrémal même lorsque  $\boldsymbol{\tau} = \mathbf{0}$ . On obtient l'indépendance asymptotique temporel lorsque  $\lim_{l \rightarrow \infty} \theta(\mathbf{h}, l) \rightarrow 2$ .

Nous présentons maintenant le processus spatio-temporel de Smith avec séparabilité spectrale. Soit  $H(\mathbf{s}, t)$  une suite de processus spatiaux de Smith indépendants et identiquement distribués, de matrice de variance-covariance  $\Sigma$  (cf Table 1.1). Alors

la fonction de répartition bivariée  $F$  du modèle spatio-temporel correspondant issu de l'équation (1.43) a la forme

$$\begin{aligned}
-\log F(x_1, x_2) &= \frac{1}{x_1} \Phi \left( \frac{\beta}{2} + \frac{1}{\beta} \log \left( \frac{x_2}{\delta^l x_1} \right) \right) \\
&+ \frac{\delta^l}{x_2} \Phi \left( \frac{\beta}{2} + \frac{1}{\beta} \log \left( \frac{\delta^l x_1}{x_2} \right) \right) + \frac{1 - \delta^l}{x_2},
\end{aligned} \tag{1.46}$$

où  $\beta = \sqrt{(\mathbf{h} - l\boldsymbol{\tau})^t \Sigma^{-1} (\mathbf{h} - l\boldsymbol{\tau})}$ . La fonction coefficient extrémal spatio-temporel associée du modèle est

$$\theta(\mathbf{h}, l) = \Phi \left( \frac{\beta}{2} + \frac{1}{\beta} \log (\delta^{-l}) \right) + \delta^l \Phi \left( \frac{\beta}{2} + \frac{1}{\beta} \log (\delta^l) \right) + 1 - \delta^l. \tag{1.47}$$

Dans la Section 3.3, nous présentons deux procédures d'estimation des paramètres de processus max-stables spatio-temporels basées sur des extensions des madogrammes au cas spatio-temporel (plus de détails dans Abu-Awwad et al. (2021)).

# Chapitre 2

## Modélisation et estimation d'événements extrêmes univariés

### 2.1 Estimation des paramètres de la GEV par la méthode des moments pondérés généralisés

La méthode des moments pondérés a été proposée par Landwehr et al. (1979) et Greenwood et al. (1979). L'idée de la méthode est de comparer, pour une variable aléatoire  $X$  de fonction de répartition  $F$ , les moments

$$\mathbb{E}(X^p(F(X))^r(1-F(X))^s), \quad \text{pour des réels } p, r \text{ et } s,$$

avec leurs versions empiriques, de la même façon que pour la méthode des moments "classiques". Dans le cas d'une variable aléatoire  $X$  suivant une distribution GEV de fonction de répartition  $G$ , Hosking et al. (1985) ont montré que  $\mathbb{E}(X(G(X))^r)$  s'écrit sous la forme

$$\beta_r = \frac{1}{r+1} \left\{ \mu - \frac{\sigma}{\xi} [1 - (1+r)^\xi \Gamma(1-\xi)] \right\}, \quad \xi < 1 \text{ et } \xi \neq 0.$$

Par conséquent, les estimateurs par moments pondérés (PWM)  $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$  des paramètres  $(\mu, \sigma, \xi)$  d'une distribution GEV sont simplement la solution du système d'équations :

$$\begin{cases} \beta_0 = \mu - \frac{\sigma}{\xi} (1 - \Gamma(1-\xi)) \\ 2\beta_1 - \beta_0 = \frac{\sigma}{\xi} \Gamma(1-\xi)(2^\xi - 1) \\ \frac{3\beta_2 - \beta_0}{2\beta_1 - \beta_0} = \frac{3^\xi - 1}{2^\xi - 1} \end{cases}$$

dans lequel  $\beta_r$  doit être remplacé par l'estimateur non biaisé proposé par Landwehr et al. (1979) :

$$\hat{\beta}_r = \frac{1}{n} \sum_{j=1}^n \left( \prod_{l=1}^r \frac{j-l}{n-l} \right) X_{j,n}$$

où  $(X_{1,n}, \dots, X_{n,n})$  représente l'échantillon GEV ordonné. Hosking et al. (1985) ont montré la normalité asymptotique de ces estimateurs pour  $\xi < 0.5$ . Ils concluent aussi que les estimateurs PWM sont meilleurs en terme de RMSE que les estimateurs classiques du maximum de vraisemblance pour les échantillons de petites tailles. La facilité d'implémentation et ses bonnes propriétés expliquent le succès de ces estimateurs en Hydrologie (cf Katz et al. (2002)). Mais le fait que ces estimateurs ne soient valables que pour  $\xi < 1$  et asymptotiquement normaux uniquement pour  $\xi < 1/2$  les rendent inutilisables pour certaines applications.

Diebolt et al. (2007) ont proposé les moments pondérés généralisés (GPWM) définis par

$$\nu_\omega = \mathbb{E}(\omega(G(X))) = \int_{-\infty}^{\infty} x\omega(G(x))dG(x) = \int_0^1 G^{-1}(u)\omega(u)du$$

où  $\omega$  est une fonction continue adaptée et  $G$  est la fonction de répartition de  $X$ . Nous proposons d'estimer  $\nu_\omega$  par

$$\hat{\nu}_\omega = \int_0^1 \mathbb{G}_n^{-1}(u)\omega(u)du$$

où  $\mathbb{G}_n$  est la fonction de répartition empirique basée sur un échantillon  $(X_1, \dots, X_n)$ . Si on s'intéresse aux propriétés asymptotiques de  $\hat{\nu}_\omega$  pour la distribution GEV, on obtient des conditions sur la fonction  $\omega$  au voisinage de 0 et de 1 :

$$\omega(t) = \mathcal{O}((1-t)^b) \text{ pour } t \text{ proche de } 1, \quad b \geq 0$$

et

$$\omega(t) = \mathcal{O}(t^{a'}) \text{ pour } t \text{ proche de } 0, \quad a' \geq 1.$$

Un exemple de fonction est  $\omega(t) = t^a(-\log t)^b, a > a'$ . Dans ce cas, le GPWM de la distribution GEV s'écrit :

$$\nu_\omega = \frac{\sigma}{\xi} \frac{1}{(a+1)^{b-\xi+1}} \Gamma(b-\xi+1) - \left( \frac{\sigma}{\xi} - \mu \right) \frac{1}{(a+1)^{b+1}} \Gamma(b+1).$$

Comme pour la méthode des PWM, un système de trois équations avec 3 différentes valeurs pour  $a$  et/ou  $b$  permet d'obtenir des estimateurs pour les trois paramètres

de la GPWM. En choisissant les trois couples  $(a, b) = (1, 1), (1, 2)(2, 1)$ , on trouve un estimateur de  $\xi$  en résolvant l'équation suivante

$$\frac{\hat{\xi}}{1 - \left(\frac{3}{2}\right)^{\hat{\xi}}} = \frac{2(\hat{\omega}_{11} - \hat{\omega}_{12})}{\hat{\omega}_{11} - \frac{9}{4}\hat{\omega}_{21}},$$

où

$$\hat{\omega}_{ab} = \int_0^1 \mathbb{F}_n^{-1}(u) u^a (-\log u)^b du.$$

Les estimateurs de  $\sigma$  et  $\mu$  s'obtiennent alors

$$\hat{\sigma} = 2^{3-\hat{\xi}} \frac{\hat{\omega}_{11} - \hat{\omega}_{12}}{\Gamma(1 - \hat{\xi})} \quad \text{et} \quad \hat{\mu} = \frac{\hat{\sigma}}{\hat{\xi}} - \frac{\hat{\sigma}}{\hat{\xi}} 2^{\hat{\xi}} \Gamma(2 - \hat{\xi}) + 4\hat{\omega}_{11}.$$

On peut montrer, de manière générale, que les GPWM existent si  $\xi < b + 1$  et sont asymptotiquement normaux si  $\xi < b + \frac{1}{2}$ . Avec le choix effectué ci-dessus pour la fonction  $\omega$  et les valeurs des couples  $(a, b)$  on obtient donc ici que les GPWM existent si  $\xi < 2$  et sont asymptotiquement normaux si  $\xi < \frac{3}{2}$ . Cette méthode étend donc le domaine de validité des estimateurs PWM de la GEV.

Une étude de simulation a montré que les estimations par GPWM sont meilleurs en terme de RMSE que les estimateurs par PWM, en particulier sur les échantillons de petite taille ( $< 50$ ) et pour des valeurs du paramètre  $\xi$  élevées et se comportent comme les estimateurs par maximum de vraisemblance pour les échantillons de taille intermédiaire (entre 50 et 100).

## 2.2 Distribution Skew-GEV

Dans Ribereau et al. (2016), nous proposons une nouvelle distribution : la Skew GEV. Cette distribution est inspirée des travaux de Azzalini (1985) sur la distribution skew-normal afin d'obtenir plus de flexibilité sur la distribution GEV.

**Définition 14** *Une variable aléatoire suit une loi Skew GEV de paramètres  $(\mu, \sigma, \xi, \lambda)$  si sa densité de probabilité est de la forme*

$$f(x) = (\lambda + 1)g(x)G^\lambda(x),$$

où  $\lambda > -1$ ,  $g(\cdot)$  et  $G(\cdot)$  sont respectivement la densité et la fonction de répartition d'une distribution GEV de paramètres  $(\mu, \sigma, \xi)$ . On notera cette distribution SGEV  $(\mu, \sigma, \xi, \lambda)$ .

Il y a plusieurs façons d'interpréter cette distribution.

La première c'est de la considérer simplement comme une distribution un peu plus flexible que la GEV et donc qui s'adapte mieux aux échantillons de maximum surtout si ceux-ci sont pris sur un faible nombre d'observations. On pourrait ainsi construire une procédure pour le choix de la taille de bloc dans la méthode du *Bloc maxima* en choisissant la taille de bloc telle que l'on peut considérer le paramètre  $\lambda$  comme nul.

La deuxième est de voir la SGEV comme la distribution du maximum quand ceux-ci sont pris sur des données dépendantes. Dans ce cas,  $\lambda$  varie entre  $-1$  et  $0$ . L'indice extrémal  $\theta$  (cf Beirlant et al. (2004)) est alors égal à  $\lambda + 1$ . La SGEV est donc plus générale puisque le paramètre  $\lambda$  doit être supérieur à  $-1$ . La SGEV englobe donc la GEV mais aussi la GEV lorsqu'elle est basée sur des données dépendantes. Cela permet entre autre de créer un estimateur *plug-in* du paramètre  $\theta$  par  $\hat{\theta} = \hat{\lambda} + 1$  où  $\hat{\lambda}$  est l'estimateur du paramètre  $\lambda$  de la SGEV obtenu par la méthode des moments pondérés (ou du maximum de vraisemblance).

Une dernière façon de voir la SGEV est la suivante : considérons des variables aléatoires indépendantes et identiquement distribuées  $X_1, X_2, \dots$  de fonction de répartition  $F$  et soit  $N$  une variable aléatoire suivant une loi de Poisson de paramètre  $\lambda^*$ , indépendante de  $\{X_i\}_{i \geq 1}$ . On définit le maximum sur une taille de bloc aléatoire :

$$M = \begin{cases} \max(X_1, \dots, X_N) & \text{si } N \geq 1 \\ 0 & \text{si } N = 0 \end{cases}$$

Après calcul on a

$$\mathbb{P}(M \leq x) = \left\{ \exp(-\bar{F}(x)) \right\}^{\lambda^*}.$$

Si on note  $G(x) = \exp(-\bar{F}(x))$  alors

$$\mathbb{P}(M \leq x) = G^{\lambda^*}(x). \tag{2.1}$$

Si  $\bar{F}$  a une queue GPD de paramètres  $(\mu, \sigma, \xi)$ , alors la fonction de répartition de  $M$  sur l'intervalle  $[\mu, \infty[$  est celle d'une SGEV. Pour  $x < \mu$ , on a  $\mathbb{P}(M \leq x) = \exp(\lambda^*)$ . L'Equation (2.1) donne donc une nouvelle interprétation de cette distribution : la fonction de répartition  $G^{\lambda^*}$ ,  $\lambda^* = \lambda + 1$  sur l'intervalle  $[\mu, \infty[$  peut être vue comme la fonction de répartition du maximum pris sur des excès et sur une taille de bloc aléatoire suivant une loi de Poisson.

On peut estimer les paramètres de cette distribution par maximum de vraisemblance ou par la méthode des moments pondérés. Pour cette dernière méthode, on

peut remarquer que si  $X$  est une variable aléatoire suivant une  $SGEV(\mu, \sigma, \xi, \lambda)$  de fonction de répartition  $F$  et  $Y$  une variable aléatoire suivant une  $GEV(\mu, \sigma, \xi)$  de fonction de répartition  $G$  et si on note  $\nu_a$  les moments pondérés de la  $SGEV$  alors on a

$$\nu_a = (\lambda + 1)\beta_{\lambda(a+1)+a}$$

où  $\beta_r$  est le moment pondéré d'ordre  $r$  d'une  $GEV$ . On a donc

$$\nu_a = \frac{1}{a+1} \left\{ \mu - \frac{\sigma}{\xi} \left[ 1 - \frac{\Gamma(1-\xi)}{((\xi+1)(a+1))^{-\xi}} \right] \right\}.$$

Comme pour la méthode des moments, on choisit 4 valeurs pour  $a$  (typiquement 0, 1, 2 et 3) et on compare les moments théoriques avec leurs versions empiriques :

$$\hat{\nu}_a = \frac{1}{n} \sum_{i=1}^n X_{i,n} \left( \frac{i}{n} \right)^a.$$

où  $X_{1,n}, \dots, X_{n,n}$  représente l'échantillon ordonné. Le système peut être simplifié et après calcul, on trouve l'estimateur de  $\xi$  en résolvant

$$\frac{3\hat{\nu}_2 - \hat{\nu}_0}{2\hat{\nu}_1 - \hat{\nu}_0} = \frac{3^{\hat{\xi}} - 1}{2^{\hat{\xi}} - 1}.$$

On trouve alors les estimateurs de  $\lambda$  et  $\sigma$  en résolvant le système suivant :

$$\begin{cases} 2\hat{\nu}_1 - \hat{\nu}_0 &= \frac{\hat{\sigma}}{\hat{\xi}} \left\{ \Gamma(1-\hat{\xi})(\hat{\lambda}+1)^{\hat{\xi}}(2^{\hat{\xi}}-1) \right\} \\ 4\hat{\nu}_3 - 3\hat{\nu}_2 &= \frac{\hat{\sigma}}{\hat{\xi}} \left\{ \Gamma(1-\hat{\xi})(\hat{\lambda}+1)^{\hat{\xi}}(4^{\hat{\xi}}-3^{\hat{\xi}}) \right\}. \end{cases}$$

Et finalement, on obtient l'estimateur de  $\mu$  par l'équation :

$$\hat{\mu} = \hat{\nu}_0 + \frac{\hat{\sigma}}{\hat{\xi}} \left[ 1 - (\hat{\lambda}+1)^{\hat{\xi}} \Gamma(1-\hat{\xi}) \right].$$

Nous avons montré la normalité asymptotique des estimateurs ainsi obtenus. Sur une étude de simulations, nous avons constaté que les estimateurs PWM étaient meilleurs à ceux obtenus par maximum de vraisemblance surtout pour le paramètre de *skewness*  $\lambda$  quand  $\xi$  est faible. Quand  $\xi$  augmente (i.e.,  $\xi \geq 1$ ), nous remarquons que ni les PWM, ni le maximum de vraisemblance ne donnent de bons résultats.

## 2.3 Modélisation globale de la distribution des précipitations

Il existe une large variété de familles de distribution pour modéliser statistiquement l'intensité des précipitations. Par exemple Katz, Vrac et al. (2007) ou Wilks (2006) suggèrent que la majeure partie de la variabilité des précipitations peut être approchée par des distributions Gamma. Mais il est aussi bien connu (cf Katz et al. (2002)) que la queue de la distribution Gamma est trop légère pour reproduire les précipitations de fortes intensités. Ceci conduit à sous-estimer fortement les niveaux de retour ainsi que toutes les autres quantités liées à des quantiles élevés de précipitations. Pour contourner ce problème, une approche populaire en Hydrologie (cf Katz et al. (2002)) consiste à éliminer les valeurs de précipitations faibles et modérées et de ne se concentrer que sur les plus fortes précipitations. L'avantage de cette approche est de pouvoir utiliser les résultats de la Théorie des Valeurs Extrêmes, i.e., les précipitations au delà d'un seuil  $u$  peuvent être approchées par une loi de Pareto Généralisée si le seuil  $u$  est suffisamment élevé et le nombre d'observations au delà du seuil est suffisant.

La Théorie des Valeurs Extrêmes indique que la probabilité que les fortes précipitations dépassant un seuil  $u$  soient supérieures à  $x$  peut être approchée par la queue d'une loi de Pareto Généralisée  $\bar{H}_\xi(\frac{x-\mu}{\sigma})$ . Le paramètre  $\xi$  va donc déterminer entièrement la forme de la queue de la distribution. Dans le cas des données de précipitations, on suppose que la distribution a une queue légère ( $\xi = 0$ ) ou lourde ( $\xi > 0$ ). Ces conditions semblent être vérifiées pour la plupart des fortes précipitations (cf Katz et al. (2002)).

Il existe quelques extensions de la GPD dans la littérature. Par exemple, dans le livre de Falk et al. (2010), les auteurs étudient les propriétés théoriques des densités de la forme

$$\text{cst} \times \frac{1}{\sigma} h_\xi \left\{ (x - \mu) / \sigma \right\} \left[ 1 + \mathcal{O} \left\{ \bar{H}_\xi^\delta(x/\sigma) \right\} \right] \quad (2.2)$$

où  $\delta > 0$  et  $h_\xi$  est la densité de la GPD. Ce type de distribution a largement été étudiée dans le domaine de recherche des *Skewed distributions* (cf Genton (2004)).

Avant d'expliquer la modélisation pour l'ensemble des précipitations, nous allons regarder un problème souvent peu étudié : celui de l'étude des faibles précipitations. Là encore, on peut utiliser la Théorie des Valeurs Extrêmes en inversant le signe des précipitations, i.e., en posant  $Y = -X$ . De nouveau, la loi des plus grandes observations de  $Y$  peut être ajustée par une GPD avec un paramètre

de forme négatif noté  $-1/\kappa$  et un paramètre d'échelle  $\nu$ , i.e.,

$$\mathbb{P}(Y > -x|Y > -u) \approx \overline{H}_{-1/\kappa} \left( \frac{-x+u}{\nu} \right)$$

quand  $u$  tend vers 0 pour tout  $x$  tel que  $0 < x < u$ . Comme le choix de  $u$  doit vérifier  $\overline{H}_{-1/\kappa}(0) = 0$ , ceci implique que  $u = \kappa\nu$  et donc que

$$\mathbb{P}(Y > -x|Y > -u) \approx \text{cst} \times x^\kappa.$$

En d'autres termes, cela conduit à dire que les précipitations de faible intensité peuvent être ajustées par une loi puissance, i.e.,  $\mathbb{P}(X \leq x) \approx \text{cst} \times x^\kappa$  pour tout "petit"  $x \geq 0$ . On peut remarquer que la loi Gamma respecte cette condition. Le fait d'avoir des paramètres de forme différents pour les précipitations d'intensités faibles et fortes peut se justifier par le fait que ce ne sont pas les mêmes mécanismes physiques qui conduisent à ces types de précipitations différentes.

Nous cherchons donc à modéliser la distribution complète des précipitations mais de façon à ce que les deux extrémités soient en accord avec la Théorie des Valeurs Extrêmes, c'est-à-dire que la queue de la distribution à droite puisse être approchée par une GPD et que la distribution à gauche puisse être approchée par une fonction puissance, i.e.,

$$\mathbb{P}(X \leq x) \approx \begin{cases} 1 - \text{cst} \times \overline{H}_\xi \left( \frac{x}{\sigma} \right), & \text{pour tout } x \text{ "grand"} \\ \text{cst} \times x^\kappa, & \text{pour tout } x \text{ "petit" proche de 0.} \end{cases}$$

Notre point de départ est l'équation utilisée pour simuler des loi GPD :

$$\sigma H_\xi^{-1}(U),$$

où  $U$  est une variable aléatoire suivant une loi uniforme sur  $[0, 1]$  et  $H_\xi^{-1}$  est la fonction quantile d'une GPD de paramètre  $\xi$ . Une façon simple d'ajouter de la flexibilité à ce schéma de simulation est de remplacer la variable aléatoire uniforme  $U$  par  $\zeta^{-1}(U)$  où  $\zeta$  est une fonction de répartition continue sur  $[0, 1]$ . Une famille plus riche de variables aléatoires peut alors être obtenue par

$$X = \sigma H_\xi^{-1}(\zeta^{-1}(U)).$$

On peut exprimer cette dernière équation en fonction de la fonction de répartition de  $X$  et on obtient

$$F(x) = \zeta \left( H_\xi \left( \frac{x}{\sigma} \right) \right). \quad (2.3)$$

Il faut maintenant trouver les fonctions  $\zeta$  telles que le comportement de la queue avec un paramètre de forme  $\xi$  est préservé et que la fonction de répartition de  $X$

pour les valeurs proche de 0 se comporte comme  $x^\kappa$ . Si on note  $\bar{\zeta} = 1 - \zeta$ , on a besoin des contraintes suivantes pour obtenir les propriétés voulues :

$$\lim_{v \downarrow 0} \frac{\bar{\zeta}(1-v)}{v} = a, \text{ pour } 0 < a < \infty, \quad (2.4)$$

$$\lim_{v \downarrow 0} \frac{\zeta(vw(v))}{\zeta(v)} = c, \text{ pour } 0 < c < \infty, \quad (2.5)$$

où  $w(v)$  est une fonction positive telle que  $w(v) = 1 + o(v)$  quand  $v \rightarrow 0$ ,

$$\lim_{v \downarrow 0} \frac{\zeta(v)}{v^\kappa} = d, \text{ pour } 0 < d < \infty, \quad (2.6)$$

Il existe plusieurs types de fonction de répartition  $\zeta$  qui vérifient les contraintes (2.4), (2.5) et (2.6). Nous nous intéresserons ici à la fonction de répartition

$$\zeta(v) = 1 - Q_\delta((1-v)^\delta), \quad \delta > 0, \quad (\text{Modèle (iii)}),$$

où  $Q_\delta$  est la fonction de répartition d'une loi Beta de paramètres  $1/\delta$  et 2, i.e.,

$$Q_\delta(v) = \frac{1+\delta}{\delta} v^{1/\delta} \left(1 - \frac{v}{1+\delta}\right).$$

Ce choix de fonction  $\zeta$  complexe correspond en fait à un cas très simple où  $\mathcal{O}(v) = -v$  et  $u = 0$  dans (2.2), i.e.,

$$f(x; \xi, \sigma, \delta) = \frac{1+\delta}{\delta} \frac{1}{\sigma} h_\xi(x/\sigma) \left\{1 - \bar{H}_\xi^\delta(x/\sigma)\right\}. \quad (2.7)$$

On appellera cette distribution l'*Extended GPD* de paramètre  $(\xi, \sigma, \delta)$ .

La Figure 2.1 permet de comparer la densité de l'Extended GPD munie de cette fonction  $G$  particulière avec la densité de la distribution Gamma et de la GPD.

D'autres choix pour la fonction  $\zeta$  sont possibles comme

$$\zeta(v) = v^\kappa, \quad (\text{Modele (i)}).$$

Ici le paramètre  $\kappa$  contrôle la forme de la queue de la distribution à gauche. Ce choix de fonction  $\zeta$  correspond à la famille de type III introduite par Papastathopoulos & Tawn (2013). La Figure 2.2 permet de comparer la densité obtenue pour différentes valeurs du paramètre  $\kappa$  ainsi que la loi Gamma.

Pour l'estimation des paramètres, nous avons calculé les estimateurs des moments pondérés en suivant la méthode développée pour la GPD par Hosking & Wallis (1987) et nous les avons comparés avec le maximum de vraisemblance. De manière générale, les deux méthodes donnent des résultats similaires en terme de RMSE. Les PWM sont plus performants pour l'estimation de  $\delta$  (mais il est connu que la méthode du maximum de vraisemblance n'est pas optimale pour l'estimation du paramètre de *skewness*) alors que le maximum de vraisemblance est meilleur pour l'estimation du niveau de retour.

Nous avons appliqué ces méthodes à des données horaires de précipitations à Lyon. Les résultats graphiques sont représentés dans la Figure 2.3. Le modèle (i) est celui qui s'ajuste le mieux selon le critère AIC.

Cette distribution a depuis été utilisée dans Tian & Reich (2020) et Turkman et al. (2020) entre autres.

## 2.4 Mesure de risque basée sur le maximum d'un processus continu

Dans cette partie, on s'intéresse aux maxima journaliers de température et on souhaite reconstruire le processus sous-jacent des températures. De cette façon, on peut établir des mesures de risques associées par exemple à la durée de période de canicule.

Comme la température ne dévie pas trop de sa moyenne, un processus avec retour à la moyenne comme le processus d'Ornstein-Uhlenbeck (OU) peut être utilisé pour modéliser le processus des températures (cf Dischel (1998a) et Dischel (1998b)). Ici, nous utiliserons un processus OU stationnaire. Le processus des températures peut alors être écrit par :

$$dX_t = l_0\beta_0(\mu_0 - X_t)dt + \sqrt{\beta_0}dB_t, \quad X_0 \sim \mathcal{N}\left(\mu_0, \frac{1}{l_0}\right)$$

où  $\mu_0 \in \mathbb{R}$ ,  $l_0, \beta_0 \in \mathbb{R}_+^*$  et  $(B_t)_{t \geq 0}$  est un mouvement Brownien standard. On suppose que  $X_0$  et  $(B_t)_{t \geq 0}$  sont indépendants. La mesure  $\mathcal{N}(\mu_0, \frac{1}{2l_0})$  est la mesure stationnaire. Cette modélisation semble raisonnable puisque dans les applications on considère les observations sur une sous-période de l'année (par exemple l'été ou l'hiver). On note  $\theta_0 = (\beta_0, \mu_0, l_0)$ . Le paramètre  $\mu_0$  est la moyenne du processus stationnaire. Le paramètre  $\sqrt{\beta_0}$  est la volatilité du processus, il indique le degré de variabilité. Dans le cas de la température, il indique la tendance à changer rapidement et de manière imprévisible. Si  $\beta_0 = 0$ , le processus est déterministe.

Finalement, le paramètre  $l_0$  montre la “vitesse” à laquelle le processus revient à la moyenne. Le paramètre  $l_0\beta_0$  est parfois appelé paramètre de relaxation. Si  $l_0 = 0$ , le processus est juste un Brownien, standard si de plus  $\beta_0 = 1$ . La Figure 2.4 représente l’influence de ces différents paramètres sur les trajectoires possibles de processus OU.

Notons pour  $s, r \in \mathbb{R}^+$ ,  $S_{[s,r[} = \sup_{s \leq t < r} X_t$  et  $I_{[s,r[} = \inf_{s \leq t < r} X_t$ . On suppose que l’on observe les maxima sur la période  $[0, T]$  avec une partition  $(t_i)_{i \geq 0}$  de pas constant  $h > 0$ . On a ainsi  $n$  suprema  $S_{[t_{i-1}, t_i[}$  pour tout  $i \in \llbracket 1; n \rrbracket$  sur des intervalles disjoints. Remarquons que dans le problème où l’on considère les données de maximum quotidien de températures, le pas est  $h = 1$ .

Les méthodes classiques ne sont pas adaptées ici pour l’estimation des paramètres pour des observations de suprema. En effet, la méthode du maximum de vraisemblance nécessite la densité des suprema. Cette densité ne peut être obtenue que par approximation numérique et cette approximation est très coûteuse en temps de calcul.

C’est pourquoi nous proposons d’utiliser la fonction de répartition du supremum, notée  $F^*$ , dont l’expression est donnée dans la Proposition 1.

**Proposition 1** *Pour  $t \in \mathbb{R}^+$  et  $a \in \mathbb{R}$ , la fonction de répartition  $F^*$  du supremum d’un processus OU stationnaire  $X$  de vecteur de paramètres  $\theta = (\beta, \mu, l) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^*$  est donnée par*

$$\begin{aligned} F^*(a, \theta, t) &= \mathbb{P}(S_{[0,t[} \leq a) \\ &= \Phi\left((a - \mu) \sqrt{2l}\right) \\ &\quad - \int_{-\infty}^a \int_0^{t\beta} e^{-\frac{l}{2}[(a-\mu)^2 + (x-\mu)^2 - u]} \frac{a-x}{\pi \sqrt{\frac{2u^3}{l}}} e^{-\frac{(a-x)^2}{2u}} \mathbb{E}\left[e^{-\frac{l^2}{2} \int_0^u (r_s - a + \mu)^2 ds}\right] du dx \end{aligned}$$

où  $r$  est un pont de Bessel de dimension 3 sur l’intervalle  $[0, u]$  entre 0 et  $a - x$  et  $\Phi$  est la fonction de répartition d’une loi Gaussienne centrée réduite.

Soit  $N_q \in \mathbb{N}$  et  $s_j, j = 1, \dots, N_q$  des réels. On note par  $F_n^*$  la fonction de répartition empirique basée sur l’échantillon de suprema  $S_{[t_{i-1}, t_i[}, i = 1, \dots, n$ .

Une façon d’estimer  $\theta_0$  est d’utiliser une méthode des moindres carrés en minimisant la somme des carrés de la différence entre la fonction de répartition

théorique et empirique des suprema, c'est-à-dire que l'on veut minimiser la fonction  $Q_n$  définie par

$$Q_n(\theta) = \sum_{j=1}^{N_q} [F^*(s_j, \theta, h) - F_n^*(s_j)]^2$$

où  $\theta$  est le vecteur de paramètres du processus OU,  $F^*(S_j, \theta, h) = \mathbb{P}(S_{[0,h]} \leq a)$  et  $s_1, \dots, s_{N_q}$  sont des nombres réels à choisir.

On obtient l'estimateur par

$$\hat{\theta}_n = \left( \hat{\beta}_n, \hat{\mu}_n, \hat{l}_n \right) = \arg \min_{\theta \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^*} Q_n(\theta).$$

Nous avons présenté les résultats pour les suprema mais on peut faire le même raisonnement pour les infima (ou conjointement les suprema et infima).

Sous les conditions classiques, on peut montrer la consistance des estimateurs.

Dans les simulations et dans les applications sur les données réelles, nous avons choisi pour  $s_1, s_2, s_3, s_4$  les quantiles empiriques sur l'échantillon d'ordre 0.2, 0.4, 0.6 et 0.8.

L'idée de départ est de pouvoir évaluer des risques liés aux canicules. Dans Klein Tank et al. (2002), on dispose d'un jeu de données de température à Paris grâce au projet ECA&D (les données sont disponibles à l'adresse <http://www.ecad.eu>). Ce jeu de données est l'un des plus longs pour les mesures de températures puisqu'il commence en 1900 mais il n'enregistre que les maxima, minima et moyenne journaliers. Dans notre cas, nous nous intéressons aux températures en été. Nous sélectionnons ainsi les températures entre le 15 juin et le 14 août (soit 61 observations) chaque année entre 1950 et 1984 (soit 35 années d'observations), pour un total de 2135 observations. Nous avons choisi ces observations pour éviter l'impact du changement climatique et nous pouvons ainsi considérer que les observations sont des maxima d'un processus stationnaire.

En appliquant la procédure d'estimation, on obtient  $\hat{\theta}_0 = (34.35, 19.04, 0.02633)$ . Nous avons utilisé deux critères de validation, l'un basé sur la comparaison de quantiles et l'autre sur la prédiction. Les deux semblent montrer une bonne adéquation du modèle.

On définit une canicule (cf Laaidi et al. (2004)) comme une suite de jours consécutifs ( $\Delta$  journées) pour lesquels la température maximale est supérieure à un seuil élevé ( $a_{max}$ ) et la température minimale est supérieure à un seuil ( $a_{min}$ ).

Ces seuils pour les températures minimales et maximales dépendent de la zone géographique que l'on considère. Pour Paris, ces valeurs sont  $\Delta = 3$ ,  $a_{min} = 21^\circ$  et  $a_{max} = 31^\circ$ . On peut choisir plusieurs type de mesure de risque (qui sont détaillés dans l'article Blanchet-Scaillet et al. (2018)) mais on va s'intéresser ici uniquement à la probabilité d'un événement caniculaire et à sa durée. Comme on s'intéresse aux mesures effectuées en été, on prend  $n = 61$ .

Pour la probabilité d'avoir un épisode caniculaire, on effectue une simulation de Monte-Carlo sur  $10^8$  années de 61 jours. On obtient que la probabilité d'avoir une canicule est de  $2.57 \times 10^{-2}$  pour un été. Il y a eu 2 canicules entre 1985 et 2011 soit une proportion de  $7.41 \times 10^{-2}$ . Ceci souligne l'impact du changement climatique sur les dernières décennies.

Avec  $10^6$  simulations de Monte-Carlo, nous avons obtenu une durée moyenne des canicules de 3.2 jours alors que les 2 dernières ont duré respectivement 3 et 10 jours.

Il y a quelques extensions possibles pour ces travaux. D'abord, on peut s'intéresser à des processus OU non stationnaires, i.e., le paramètre  $\mu$  pourrait être constant par morceaux, typiquement constant chaque été et croissant d'un été à l'autre pour traduire le changement climatique ou le paramètre  $\mu$  pourrait être une fonction du temps de type sinusoïdale pour traduire les saisons. Une autre possibilité serait d'estimer les paramètres par maximum de vraisemblance simulé. La vraisemblance étant difficile à calculer mais la simulation des observations étant relativement facile, cela consiste à effectuer une simulation avec un jeu de paramètres, implémenter la vraisemblance empirique basée sur les observations, calculée sur les vraies observations. Ensuite, on cherche le jeu de paramètres qui maximise cette vraisemblance simulée.

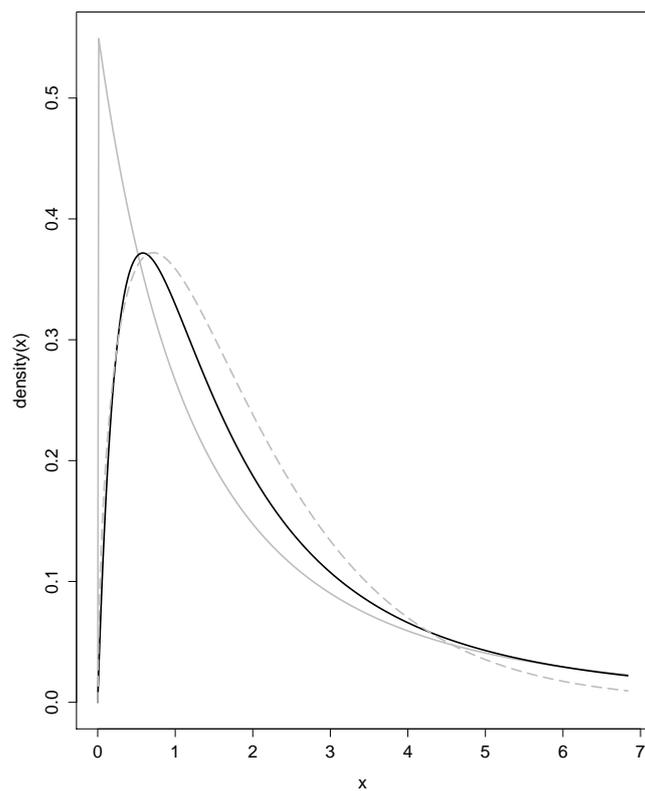


FIGURE 2.1 – Comparaison entre 3 densités différentes. Les courbes en gris, pointillé et noire correspondent respectivement à la GPD de paramètres  $(u, \sigma, \xi) = (0, 1.8, 0.5)$ , à la distribution Gamma de paramètres  $(\gamma_1, \gamma_2) = (1.2, 1.6)$  et à la *Extended GPD* avec  $\zeta(v) = 1 - Q_\delta((1-v)^\delta)$ ,  $\delta > 0$  de paramètres  $(\delta, \sigma, \xi) = (1, 1, 0.5)$ .

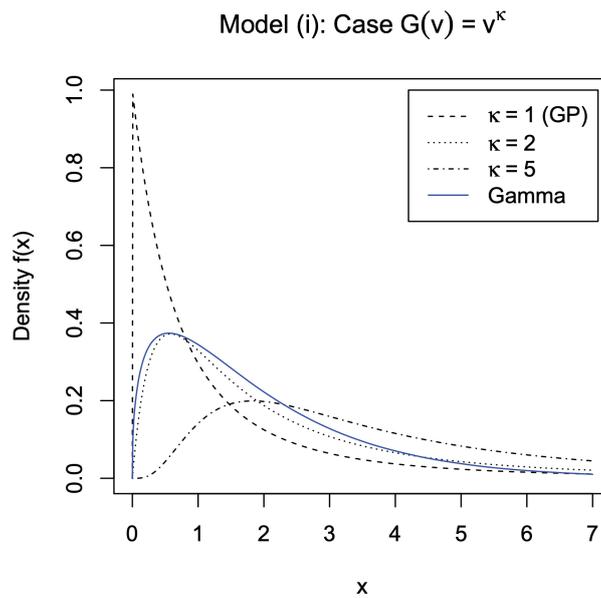


FIGURE 2.2 – Densité correspondant au modèle (2.3) avec  $\zeta(v) = v^\kappa$ , pour  $\sigma = 1$ ,  $\xi = 0.5$  et différentes valeurs du paramètre de la queue à gauche  $\kappa = 1, 2$  et  $5$  (en tirets, pointillés et tiret-pointillé respectivement). Le cas  $\kappa = 1$  correspond à la GPD. La ligne bleue représente la densité de la loi Gamma de paramètres  $(1.4, 1.4)$ .

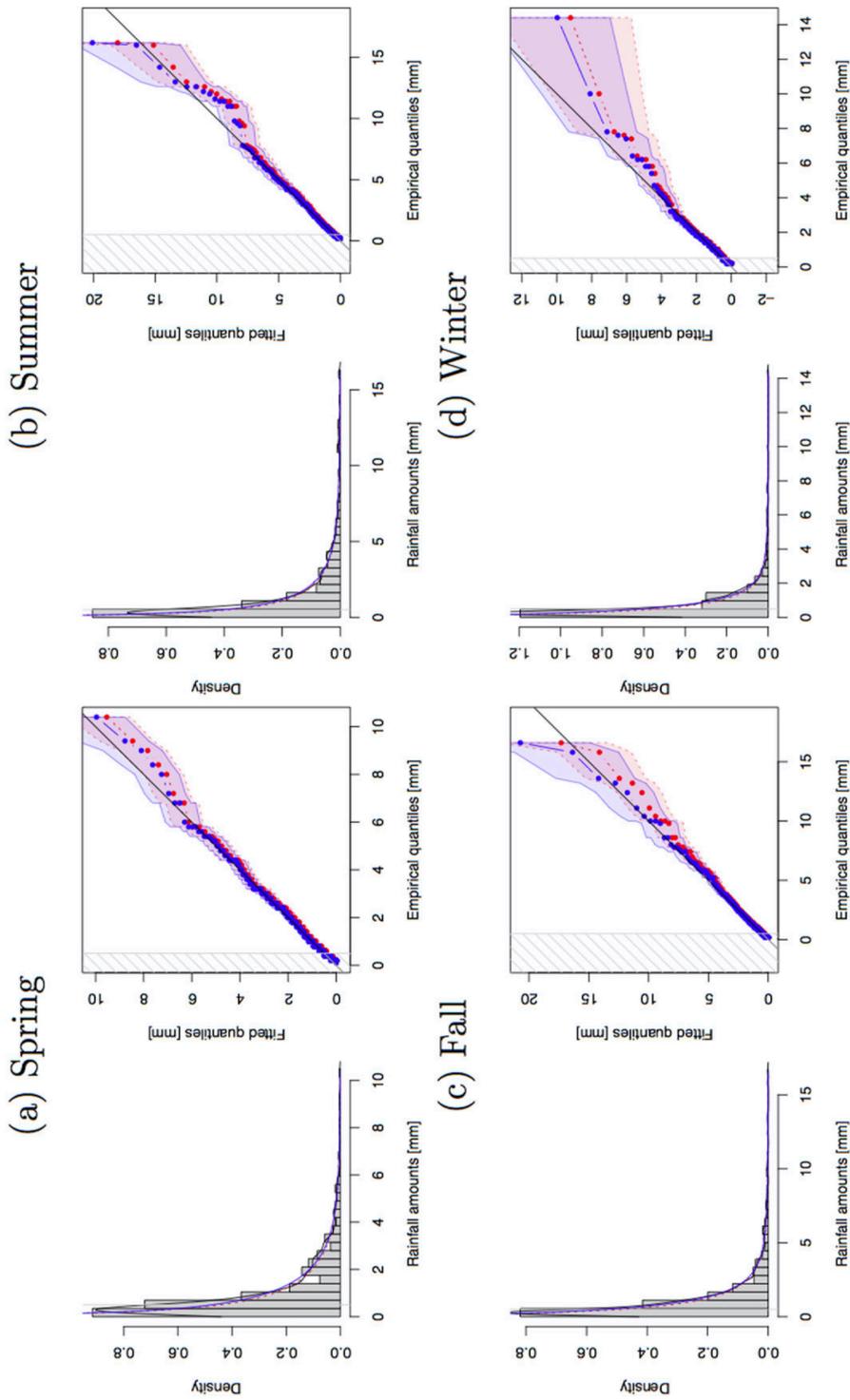


FIGURE 2.3 – Précipitations horaires à Lyon (1996-2011). Pour chaque saison, à gauche, histogramme des données, en noir une estimation de la densité par noyau et les densités des *extended GPD* (par maximum de vraisemblance en bleu et par PWM en tirets rouge). A droite quantile-quantile plot avec les intervalles à 95% basés sur 500 répliqués bootstrap.

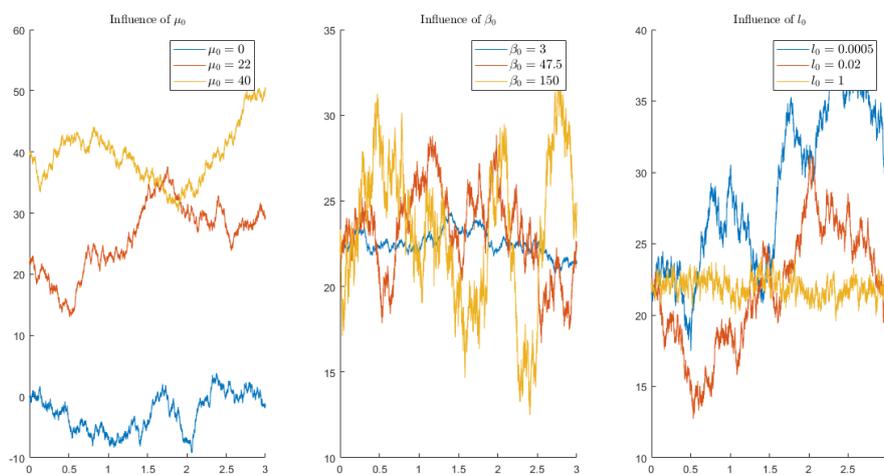


FIGURE 2.4 – Différentes trajectoires de réalisations de processus d’Ornstein-Uhlenbeck pour différents paramètres. Quand ce n’est pas précisé, les autres paramètres sont  $\mu_0 = 22$ ,  $l_0 = 0.02$  et  $\beta_0 = 47.5$ . L’échelle de temps est  $dt = 10^{-4}$  jours. Pour chaque trajectoire,  $X_0 = \mu_0$ .

# Chapitre 3

## Etude de processus max-stables et max-mélange

### 3.1 Estimation des paramètres d'un processus max-mélange

Dans Ahmed et al. (2021), nous proposons une méthode d'inférence pour les paramètres d'un modèle max-mélange basée sur une méthode des moindres carrés appliquée au madogramme.

Nous avons défini le  $F$ -madogramme dans la Définition 7. Nous allons considérer ici des processus isotropiques et on note  $h = \|\mathbf{h}\|$ . Nous calculons dans la proposition suivante  $\nu^F(h)$  pour un processus max-mélange présenté dans la Définition 11.

**Proposition 2** *Soit  $Z$  un processus max-mélange de paramètre  $a \in [0, 1]$ . Soit  $X$  la partie max-stable de fonction coefficient extrémal  $\theta_X(h)$ . Soit  $Y$  la partie inverse max-stable de coefficient  $\eta(h)$ . Alors le  $F$ -madogramme de  $Z$  est donné par*

$$\begin{aligned} \nu^F(h) = & \frac{a(\theta_X(h) - 1)}{a(\theta_X(h) - 1) + 2} - \frac{a\theta_X(h) - 1}{2a\theta_X(h) + 2} \\ & - \frac{1/\eta(h)}{a\theta_X(h) + (1 - a)/\eta(h) + 1} \beta\left(\frac{a\theta_X(h) + 1}{(1 - a)}, 1/\eta(h)\right), \end{aligned} \quad (3.1)$$

où  $\beta$  est la fonction Beta.

Contrairement aux fonctions  $\chi$  et  $\bar{\chi}$ , le  $F$ -madogramme combine les paramètres de la partie asymptotiquement dépendante et asymptotiquement indépendante.

On peut en déduire facilement en faisant tendre  $a$  vers 0 et vers 1 que le  $F$ -madogramme d'un processus max-stable est donné par

$$2\nu^F(h) = \frac{\theta(h) - 1}{\theta(h) + 1}$$

et celui d'un processus asymptotiquement indépendant est donné par

$$2\nu^F(h) = \frac{1 - \eta(h)}{1 + \eta(h)}.$$

Nous proposons d'estimer les paramètres du processus max-mélange par une méthode des moindres carrés non-linéaires en utilisant le  $F$ -madogramme. Cette procédure d'estimation est une alternative (ou un point de départ) à l'estimation par maximum de vraisemblance composite utilisée dans Bacro et al. (2016) et Wadsworth & Tawn (2012).

On considère  $(Z_t)_{t=1,\dots,T}$  des copies d'un processus isotropique max-mélange  $Z$  de loi marginale Fréchet standard (par la suite  $F_0$  désignera la fonction de répartition de la loi Fréchet standard). Ces copies peuvent être indépendantes ou alors elles doivent vérifier une propriété d' $\alpha$ -mélange (cf Bacro et al. (2016)). Soit  $\mathcal{H}$  un sous-ensemble de  $\mathcal{S}$ . On note

$$J(x, y) = \frac{1}{2}|x - y| \text{ et } Y_{h,t}(s) = J(F_0(Z_t(s)), F_0(Z_t(s+h))),$$

pour  $t = 1, \dots, T$  et  $h \in \mathcal{H}$ . Alors, pour  $t = 1, \dots, T$ , les vecteurs  $(Y_{h,t})_{h \in \mathcal{H}}$  ont la même loi et sont considérés soit indépendants soit  $\alpha$ -mélangeant en  $t$ . La principale motivation pour utiliser le  $F$ -madogramme est qu'il contient l'information sur la structure de dépendance de  $Y_{h,t}$  pour un  $h$  fixé (cf Section 3.2 dans Bacro et al. (2010)).

Par la suite on supposera que les vecteurs  $(Y_{h,t})_{h \in \mathcal{H}}$  sont indépendants et identiquement distribués. On remarque que par la définition du  $F$ -madogramme, on a  $\mathbb{E}[Y_{t,h}] = \nu^F(h, \psi)$  où  $\nu^F(h, \psi)$  est le  $F$ -madogramme de  $Z$  de paramètre  $\psi$  comme dans la Définition 7. Si  $Z$  est de paramètre inconnu  $\psi_0$  sur un ensemble compact  $\Psi \subset \mathbb{R}^d$ , on réécrit

$$Y_{h,t} = \nu^F(h, \psi_0) + \varepsilon_{h,t}. \quad (3.2)$$

Les vecteurs  $(\varepsilon_{h,t})_{h \in \mathcal{H}}$  sont des bruits indépendants et identiquement distribués tels que  $\mathbb{E}[\varepsilon_{h,t}] = 0$  et  $\mathbb{V}(\varepsilon_{h,t}) = \sigma_u^2 > 0$  est finie et inconnue.

On pose

$$\mathcal{L}(\psi) = \sum_{h \in \mathcal{H}} \frac{1}{T} \sum_{t=1,\dots,T} (Y_{t,h} - \nu^F(h, \psi))^2. \quad (3.3)$$

Tout vecteur  $\hat{\psi}_T$  dans  $\Psi$  qui minimise  $\mathcal{L}(\psi)$  est un estimateur des moindres carrés de  $\psi_0$ .

$$\hat{\psi}_T \in \arg \min_{\psi \in \Psi} \mathcal{L}(\psi). \quad (3.4)$$

Si on suppose que  $\Psi \subset \mathbb{R}^d$  est compact, que  $\psi \mapsto \nu^F(h, \psi)$  est continue pour tout  $h \in \mathcal{H}$  et que les vecteurs  $(Y_{h,t})_{h \in \mathcal{H}}$  sont indépendants et identiquement distribués, alors en notant  $(\hat{\psi}_T)_{T \in \mathbb{N}}$  l'estimateur des moindres carrés de  $\psi_0$ , on peut montrer que quelque soit le point limite (quand  $T$  tend vers l'infini)  $\psi$  de  $(\hat{\psi}_T)_{T \in \mathbb{N}}$  vérifie  $\nu(h, \psi) = \nu(h, \psi_0)$  pour tout  $h \in \mathcal{H}$ .

La normalité asymptotique de l'estimateur s'obtient en suivant Buhl et al. (2019) et en utilisant la normalité asymptotique du  $F$ -madogramme obtenue dans Cooley et al. (2006). Toutefois le calcul de la variance asymptotique est complexe puisqu'il nécessite de calculer les covariances entre  $\nu^F(h_i, \psi)$  et  $\nu^F(h_j, \psi)$  pour  $i \neq j$ .

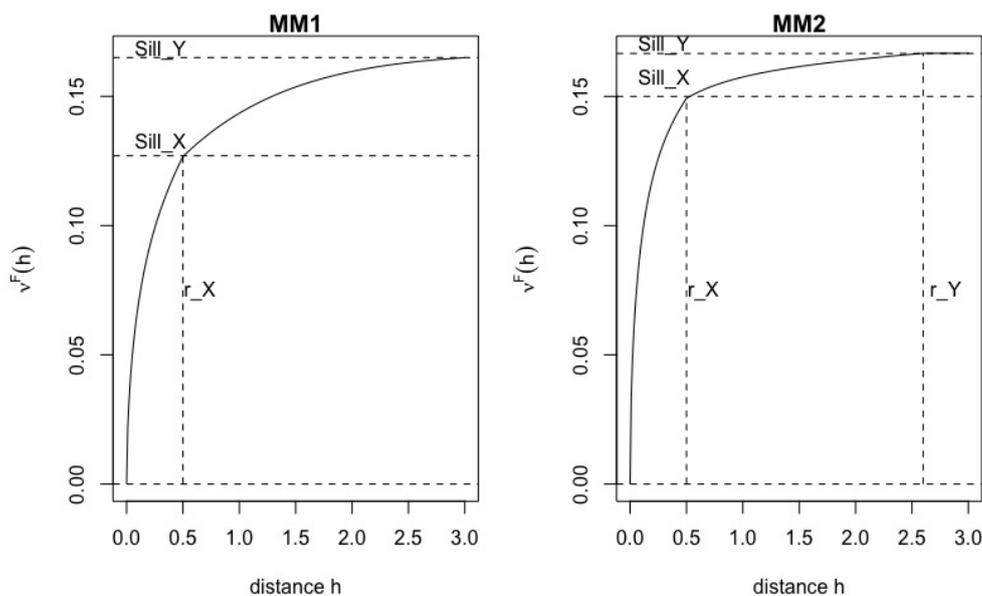


FIGURE 3.1 –  $F$ -madogramme pour les modèles max-mélanges MM1 et MM2. Le modèle MM1 a une fonction de corrélation  $\rho(h) = \exp(-h/\theta_X)$ ,  $r_X = 0.25$ ,  $\theta_X = 0.2$  et  $\sigma_Y = 0.6$ . Le modèle MM2 a une fonction de corrélation  $\rho_X(h) = \exp(-h/\theta_X)$  pour  $X$  et  $\rho_Y(h) = \exp(-h/\theta_Y)$  pour  $Y$  avec différentes valeurs  $\theta_X = 0.2$  et  $\theta_Y = 0.9$ . Les rayons  $r_X$  et  $r_Y$  sont représentés par les traits verticaux en tirets sur la figure. Pour les 2 modèles, on a  $a = 0.5$ .

Dans la Figure 3.1, nous avons tracé les  $F$ -madogrammes pour 2 modèles max-mélanges différents :

- **MM1** est un modèle max-mélange entre un processus max-stable TEG  $X$  (cf Table 1.1) de fonction de corrélation  $\rho_X(h) = \exp(-h/\theta_X)$  avec  $\theta_X = 0.2$  et  $\mathcal{A}$  un disque de rayon  $r_X = 0.25$ ;  $Y$  est un processus inverse-Smith (i.e., la transformation inverse d'un processus de Smith) de matrice de covariance  $\Sigma = \sigma_Y \mathbf{I}_d$  avec  $\sigma_Y = 0.6$ . On choisit pour paramètre de mélange  $a = 0.5$ . Dans ce modèle, les paires  $\{Z(s), Z(s+h)\}$  sont asymptotiquement dépendantes si  $h < 2r_X$  et asymptotiquement indépendantes si  $h \geq 2r_X$ .
- **MM2** est un modèle max-mélange entre un processus max-stable TEG  $X$  et un inverse TEG  $Y$ .  $X$  a pour fonction de corrélation  $\rho_X(h) = \exp(-h/\theta_X)$  avec  $\theta_X = 0.2$  et  $\mathcal{A}_X$  un disque de rayon  $r_X = 0.25$ .  $Y$  a pour fonction de corrélation  $\rho_Y(h) \exp(-h/\theta_Y)$  avec  $\theta_Y = 0.8$  et  $\mathcal{A}_Y$  un disque de rayon  $r_Y = 1.3$ . On choisit pour paramètre de mélange  $a = 0.5$ . Dans ce modèle, les paires  $\{Z(s), Z(s+h)\}$  sont asymptotiquement dépendantes si  $h < 2r_X$ , asymptotiquement indépendantes si  $2r_Y > h \geq 2r_X$  et indépendantes pour  $h \geq 2r_Y$ .

Sur la Figure 3.1, on voit que le  $F$ -madogramme a deux seuils qui correspondent pour l'un à  $X$  et pour l'autre à  $Y$ . On peut aussi remarquer que le  $F$ -madogramme s'exprime avec tous les paramètres du modèle ce qui le rend plus utile que la fonction  $\chi$  qui ne dépend que de la partie max-stable ou que  $\bar{\chi}$  qui ne dépend que de la partie asymptotiquement indépendante.

Nous avons effectué une étude de simulation (que l'on peut retrouver plus détaillée dans Ahmed et al. (2021)) pour comparer l'estimation des paramètres des modèles MM1 et MM2 pour différentes valeurs de  $a = 0, 0.25, 0.5, 0.75$  et  $1$  par la méthode des moindres carrés basée sur le  $F$ -madogramme (de l'Equation (3.4)) et par la méthode du maximum de vraisemblance composite censuré (cf Bacro et al. (2016)). Sur les résultats que l'on obtient, on remarque que généralement les estimateurs par les 2 méthodes fonctionnent correctement même si pour la partie asymptotiquement indépendante on obtient une assez grande variance d'estimation. De manière générale, en terme de RMSE, les estimateurs des moindres carrés sont meilleurs que ceux obtenus par vraisemblance composite pour  $a$  proche de  $0$  ( $a = 0, 0.25$ ) alors qu'on observe le phénomène inverse lorsque  $a$  est proche de  $1$  ( $a = 0.75, 1$ ). Les deux estimateurs montrent des performances similaires lorsque  $a = 0.5$ .

On peut aussi définir un critère pour la sélection de modèle, équivalent au CLIC

pour le maximum de vraisemblance composite :

$$MIC := \log \mathcal{L}(\hat{\psi}) + (2k(k+1)/(T-k-1)),$$

où  $k$  est le nombre de paramètres du modèle et  $T$  le nombre d'observations, i.e.,  $T = N \times |\mathcal{H}|$ , où  $|\mathcal{H}|$  est le nombre de paire observées. Sur des données de précipitations sur la côte Est Australienne, les deux critères (MIC et CLIC) retiennent le même modèle.

## 3.2 Estimation “non-paramétrique” du paramètre de mélange

Dans Abu-Awwad et al. (2019), nous proposons d'estimer le paramètre de mélange d'un processus max-mélange (cf Définition 11) sans spécifier les modèles pour la partie max-stable et la partie asymptotiquement indépendante en nous basant sur le  $F^\lambda$ -madogramme.

Le  $F^\lambda$ -madogramme, présenté dans la Définition 8, est une mesure de la dépendance qui apporte plus d'informations que le  $\chi$  ou  $\bar{\chi}$  dans le cadre des processus max-mélanges. En effet, l'écriture explicite obtenue dans l'Equation (3.5) du  $F^\lambda$ -madogramme dans ce cadre là dépend de tous les paramètres du modèle.

**Proposition 3** *Soit  $\{X(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus max-stable standard de fonction coefficient extrémal  $\theta_X$ , et  $\{Y(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  un processus inverse max-stable de fonction coefficient extrémal  $\theta_Y$ . Alors le  $F^\lambda$ -madogramme du processus max-mélange (cf Définition 11)  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  est donné par :*

$$\begin{aligned} \nu_{F^\lambda}(\mathbf{h}) = & \frac{\lambda}{1+\lambda} - \frac{2\lambda}{a(\theta_X(\mathbf{h})-1) + 1 + \lambda} + \frac{\lambda}{a\theta_X(\mathbf{h}) + \lambda} \\ & - \frac{\lambda\theta_Y(\mathbf{h})}{(1-a)\theta_Y(\mathbf{h}) + a\theta_X(\mathbf{h}) + \lambda} \beta\left(\frac{a\theta_X(\mathbf{h}) + \lambda}{1-a}, \theta_Y(\mathbf{h})\right), \end{aligned} \quad (3.5)$$

où  $\beta(.,.)$  est la fonction Beta.

En faisant tendre  $a$  vers 1, on obtient le  $F^\lambda$ -madogramme d'un processus max-stable  $X$  de fonction coefficient extrémal  $\theta_X$  :

$$\nu_{F^\lambda}(\mathbf{h}) = \frac{\lambda}{\lambda+1} \frac{\theta_X(\mathbf{h}) - 1}{\lambda + \theta_X(\mathbf{h})} \quad (3.6)$$

et on a  $\nu_{F^\lambda}(\mathbf{h}) \in [0, \frac{\lambda}{(1+\lambda)(2+\lambda)}]$  (cf Bel et al. (2008)). En posant  $a = 0$ , on obtient le  $F^\lambda$ -madogramme d'un processus inverse max-stable  $Y$  de fonction coefficient extrémal  $\theta_Y$  :

$$\nu_{F^\lambda}(\mathbf{h}) = \frac{1}{1+\lambda} - \frac{\lambda\theta_Y(\mathbf{h})}{\lambda + \theta_Y(\mathbf{h})}\beta(\lambda, \theta_Y(\mathbf{h})). \quad (3.7)$$

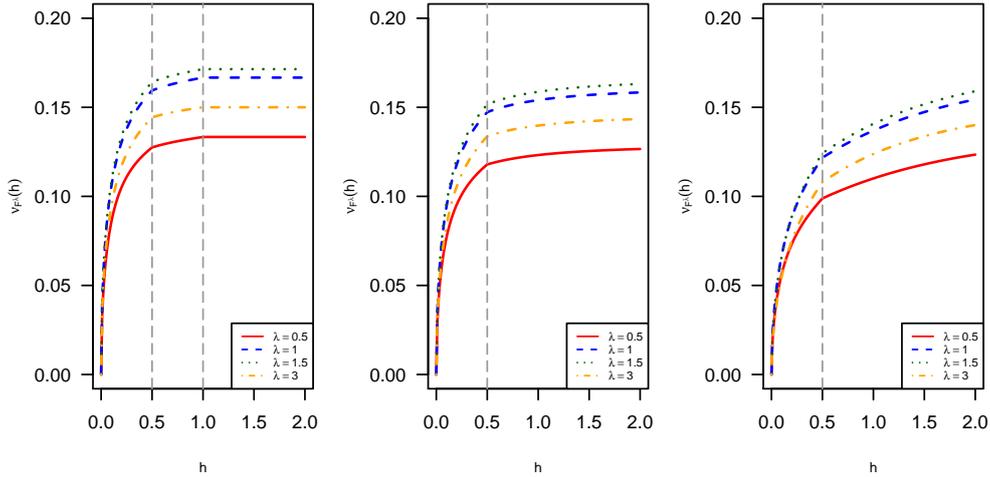


FIGURE 3.2 –  $F^\lambda$ -madogramme (Equation (3.5)) pour  $\lambda = \{0.5, 1, 1.5, 3\}$  pour divers modèles max-mélanges avec coefficient de mélange  $a = 0.5$ . A gauche, la partie max-stable  $X$  est un processus TEG avec  $\mathcal{A}_X$  un disque de rayon  $r_X = 0.25$  et  $\rho_X(\mathbf{h}) = \exp(-\|\mathbf{h}\|/0.2)$ . Le processus  $Y$  est un inverse TEG avec  $\mathcal{A}_Y$  un disque de rayon  $r_Y = 0.5$  et  $\rho_Y(\mathbf{h}) = \exp(-\|\mathbf{h}\|/0.4)$ . Au milieu,  $X$  est un TEG comme précédemment et  $Y$  est un inverse extremal- $t$  de degrés de liberté  $\nu = 2$  et  $\rho_Y(\mathbf{h}) = \exp(-\|\mathbf{h}\|)$ . A droite,  $X$  est un TEG comme précédemment et  $Y$  est un inverse Smith avec  $\Sigma = \mathbf{Id}_2$ . Les lignes verticales représentent les diamètres des disques pour les processus TEG.

Si on prend  $\lambda = 1$  on obtient la formule du  $F$ -madogramme (cf Proposition 2 ou Ahmed et al. (2021)). La Figure 3.2 montre le comportement du  $F^\lambda$ -madogramme pour différents processus max-mélanges en fonction de  $\mathbf{h}$  pour différentes valeurs de  $\lambda$ .

### Procédure d'estimation

Nous présentons la procédure d'estimation. Pour des raisons computationnelles, nous séparons la procédure en 2 étapes : d'abord nous estimons les coefficient

extrémaux  $\theta_X(\mathbf{h})$  et  $\theta_Y(\mathbf{h})$  pour un paramètre  $a$  fixé et ensuite nous estimons le paramètre  $a$  à partir des estimations précédentes. En partant de l'Equation (3.5), on peut écrire le  $F^\lambda$ -madogramme comme une fonction de  $a$ ,  $\lambda$ ,  $\theta_X(\mathbf{h})$  et  $\theta_Y(\mathbf{h})$ , i.e.,  $\nu_{F^\lambda}(\mathbf{h}) = g(a, \lambda, \theta_X(\mathbf{h}), \theta_Y(\mathbf{h}))$ . L'idée est d'estimer  $\theta_X(\mathbf{h})$  et  $\theta_Y(\mathbf{h})$  par moindres carrés sur  $g(a, \lambda, \theta_X(\mathbf{h}), \theta_Y(\mathbf{h}))$  et sa version empirique et d'ensuite estimer  $a$  par  $\hat{a}$  tel que la version empirique du  $F^{\lambda'}$ -madogramme est la plus proche de  $g(a, \lambda', \hat{\theta}_X(\mathbf{h}), \hat{\theta}_Y(\mathbf{h}))$ . Une approche similaire a été proposée par Bel et al. (2008) pour estimer la fonction coefficient extrémal dans un cadre max-stable.

Concrètement, on considère des copies indépendantes  $Z_i$ ,  $i = 1, \dots, T$  de  $Z$  et on pose

$$Q_i(\mathbf{h}, \lambda) = \frac{1}{2} |F^\lambda(Z_i(\mathbf{s})) - F^\lambda(Z_i(\mathbf{s} + \mathbf{h}))|,$$

où  $F$  est la fonction de répartition d'une loi Fréchet unitaire. Par la définition du  $F^\lambda$ -madogramme, on a  $\mathbb{E}[Q_i(\mathbf{h}, \lambda)] = \nu_{F^\lambda}(\mathbf{h})$ . Si les marginales du processus ne sont pas Fréchet, ce qui est souvent le cas dans la pratique, on les transforme via la fonction de répartition empirique  $\hat{F}$  au lieu de  $F$ . On a alors :

$$\begin{aligned} \widehat{Q}_i(\mathbf{h}, \lambda) &= \frac{1}{2} |\hat{F}^\lambda(Z_i(\mathbf{s})) - \hat{F}^\lambda(Z_i(\mathbf{s} + \mathbf{h}))|, \text{ et} \\ \widehat{\nu}_{F^\lambda}(\mathbf{h}) &= \frac{1}{2T} \sum_{i=1}^T |\hat{F}^\lambda\{Z_i(\mathbf{s})\} - \hat{F}^\lambda\{Z_i(\mathbf{s} + \mathbf{h})\}|, \end{aligned}$$

où  $\widehat{\nu}_{F^\lambda}(\mathbf{h})$  est la version empirique de  $\nu_{F^\lambda}(\mathbf{h})$ .

Si on note  $\Lambda \subset (0, \infty)$  un ensemble fini des valeurs possibles pour  $\lambda$ , alors pour une valeur donnée de  $a$ , on peut obtenir un estimateur semi-paramétrique par moindres carrés de  $\boldsymbol{\theta}(\mathbf{h}) = (\theta_X(\mathbf{h}), \theta_Y(\mathbf{h}))^t$  par

$$\widehat{\boldsymbol{\theta}}_{NLS}^a(\mathbf{h}) = \arg \min_{(\theta_X, \theta_Y) \in [1, 2]} T^{-1} \sum_{\lambda \in \Lambda} \sum_{i=1, \dots, T} \left[ \widehat{Q}_i(\mathbf{h}, \lambda) - g(a, \lambda, \theta_X(\mathbf{h}), \theta_Y(\mathbf{h})) \right]^2. \quad (3.8)$$

Ensuite pour estimer  $a$ , on suppose que les  $Z_i$  sont observés à des sites  $\mathbf{s}_1, \dots, \mathbf{s}_D$  et que  $\mathbf{h}$  sont les distances entre les sites  $\mathbf{s}_j$ . Pour un  $a$  fixé, on pose  $\widehat{\boldsymbol{\theta}}_{NLS}^a(\mathbf{h}) = (\widehat{\theta}_X^a(\mathbf{h}), \widehat{\theta}_Y^a(\mathbf{h}))^t$ , estimé par la Formule (3.8) pour différentes valeurs  $\lambda \in \Lambda$ . On prend  $\lambda' \notin \Lambda$  et on note par  $\mathcal{H} \subset [0, \infty)$  l'ensemble fini de distance spatiale  $\mathbf{h}$ , i.e.,

$$\mathcal{H} = \left\{ \mathbf{h}_k = \mathbf{s}_i - \mathbf{s}_j, i = 1, \dots, D-1, j = i+1, \dots, D, k = 1, \dots, \frac{D(D-1)}{2} \right\}.$$

De plus  $\widehat{\nu}_{F^{\lambda'}} = g(a, \lambda', \widehat{\theta}_X^a(\mathbf{h}), \widehat{\theta}_Y^a(\mathbf{h}))$ . On définit alors le critère de décision (DC) suivant :

$$\text{DC}(a) = \sum_{\mathbf{h} \in \mathcal{H}} \omega(\mathbf{h}) \left[ \frac{\hat{\nu}_{F\lambda'}(\mathbf{h})}{\tilde{\nu}_{F\lambda'}(\mathbf{h})} - 1 \right]^2, \quad (3.9)$$

où les  $\omega(\mathbf{h})$  sont des poids non négatifs qui peuvent être utilisés pour réduire le nombre de paires utilisées dans l'estimation (un peu dans l'esprit de la vraisemblance composite). Un choix simple de poids est  $\omega(\mathbf{h}) = \mathbb{1}_{\{\|\mathbf{h}\| \leq r\}}$  où la valeur  $r$  peut être choisie comme le quantile d'ordre  $q$  des distances  $h$  entre les sites,  $q \in (0, 1)$ . Finalement on estime  $a$  par le  $\hat{a}$  qui minimise  $\text{DC}(a)$ .

On peut montrer sous certaines hypothèses assez générales que l'estimateur  $\hat{a}$  est convergent (cf Abu-Awwad et al. (2019) pour plus de détails).

Sur une étude de simulation, nous avons trouvé que les différentes valeurs de  $\lambda$  qui fonctionnent le mieux sont  $\Lambda = \{1, 3\}$  et  $\lambda' = 1.5$ . Nous avons aussi remarqué que notre procédure donne de meilleurs résultats pour l'estimation de  $\theta_X(\mathbf{h})$  lorsque la valeur du paramètre de mélange  $a$  augmente et de façon analogue, l'estimation de  $\theta_Y(\mathbf{h})$  s'améliore lorsque le paramètre  $a$  diminue ce qui est cohérent avec les résultats obtenus dans la Section 3.1 et Ahmed et al. (2021). De même les plus petites valeurs du critère DC correspondent en général au vrai paramètre de mélange.

Bien entendu, on peut utiliser des méthodes basées sur le maximum de vraisemblance (composite par exemple) pour estimer l'ensemble des paramètres du modèle et pas seulement les mesures de dépendance bivariées. Toutefois la motivation ici est de ne pas poser de modèle sur les parties max-stable et asymptotiquement indépendante du modèle max-mélange. On réduit ainsi le risque de modèle. On peut donc voir cette méthode comme une étude préliminaire vers l'étude d'un modèle paramétrique.

### Application sur les données de précipitations Australiennes

Nous avons appliqué notre procédure à des données de précipitation en Australie. La Figure 3.3 montre la localisation des 41 stations considérées sur la côte Sud-Est de l'Australie. Nous n'avons considéré que la période hivernale (Avril à Septembre). L'analyse complète de ces données est disponible dans Abu-Awwad et al. (2019), ici, nous ne détaillerons que la partie qui concerne l'analyse exploratoire et donc la procédure que nous avons décrite.

La Figure 3.4 représente les fonctions  $\chi_u$  et  $\bar{\chi}_u$  avec  $u = 0.97$  dans 4 différentes

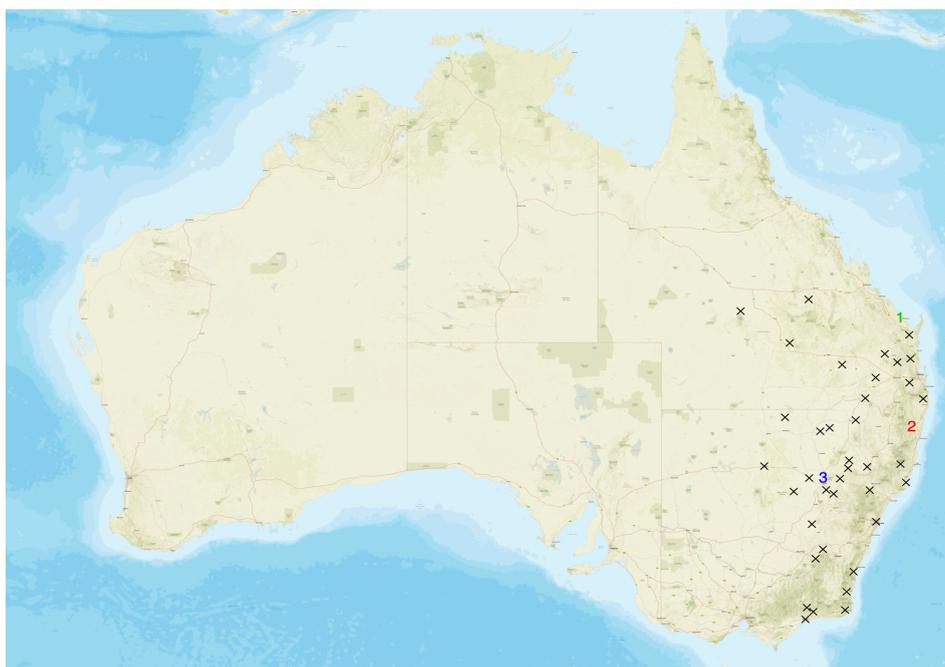


FIGURE 3.3 – Localisations des 41 stations météorologiques considérées sur la côte Est Asutralienne. Les 38 stations marquées d'une croix noire ont été utilisées pour l'inférence et les 3 autres numérotées en couleur ont été utilisées pour vérifier la qualité d'ajustement.

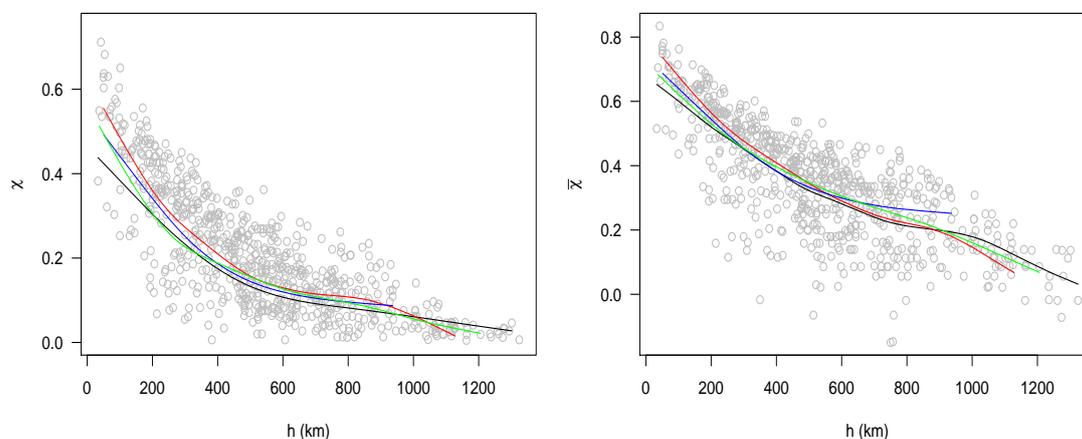


FIGURE 3.4 – Estimations empiriques de  $\chi(\cdot)$  (à gauche) et  $\bar{\chi}(\cdot)$  (à droite) en fonction de la distance  $h$  avec un seuil  $u = 0.970$ . Les points gris sont les estimations empiriques pour chaque paire de sites. Les lignes en couleur représentent des versions lissées des estimations empiriques dans chacune des direction : noire  $(-\pi/8, \pi/8]$ , rouge  $(\pi/8, 3\pi/8]$ , bleu  $(3\pi/8, 5\pi/8]$ , et verte  $(5\pi/8, 7\pi/8]$ .

directions  $(-\pi/8, \pi/8]$ ,  $(\pi/8, 3\pi/8]$ ,  $(3\pi/8, 5\pi/8]$ , et  $(5\pi/8, 7\pi/8]$ . D'après cette figure, il ne semble pas y avoir d'anisotropie sur les données. De plus, ces graphiques semblent indiquer une dépendance asymptotique entre stations distantes de moins de 500 km (i.e.,  $\widehat{\chi}_{0.97}(\mathbf{h}) > 0$  pour  $h \leq 500$  km). On peut supposer l'indépendance asymptotique pour des stations séparées d'une distance entre 500 et 1000 km ( $\widehat{\chi}_{0.97}(\mathbf{h}) < 1$  pour  $h \in (500, 1000)$  km). D'autre part les paires de sites distantes de plus de 1000 km semblent être (exactement) indépendantes puisque  $\widehat{\chi}_{0.97}(\mathbf{h}) \approx 0$  si  $h \geq 1000$  km.

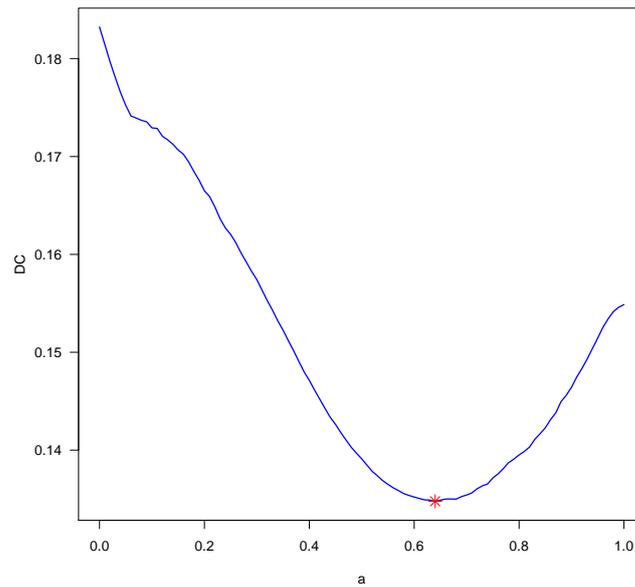


FIGURE 3.5 –  $a \mapsto DC(a)$  sur le jeu de données des précipitations Australiennes,  $a \in (0, 1)$  avec un pas de 0.01. La croix rouge symbolise le minimum du DC qui correspond à  $a = 0.64$ .

Après traitement des données, nous appliquons la procédure décrite précédemment pour estimer les paramètres du modèle max-mélange :

- (i) On estime d'abord  $a$ , où  $a$  varie de 0 à 1 par pas de 0.01. Nous fixons  $\Lambda = \{1, 3\}$  et  $\lambda' = 1.5$ . La fonction  $a \mapsto DC(a)$  est tracée dans la Figure 3.5. Le meilleur modèle max-mélange ajusté d'après le critère DC a un paramètre de mélange estimé  $\widehat{a} = 0.64$ .
- (ii) Ensuite, pour estimer les fonctions  $\theta_X(\mathbf{h})$  et  $\theta_Y(\mathbf{h})$ , on ajuste de nouveau

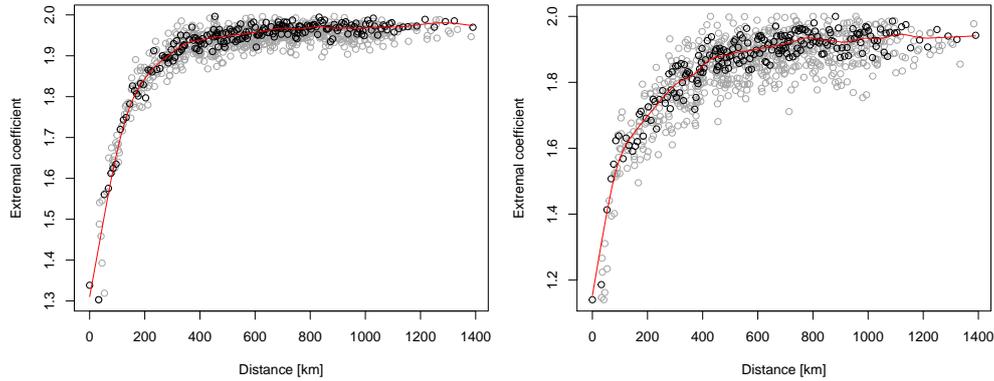


FIGURE 3.6 – Estimation non linéaire par moindres carrés de  $\theta_X(\mathbf{h})$  (gauche) et  $\theta_Y(\mathbf{h})$  (droite) en fonction de la distance  $\|\mathbf{h}\|$ . Les points gris sont les estimations par paires, les points noirs des estimations 'binned' et la courbe rouge une version lissée de l'estimation 'binned'.

les données par la procédure des moindres carrés non-linéaire de l'Equation (3.8) sur la base du paramètre de mélange estimé en (i). Les résultats se trouvent dans la Figure 3.6.

L'estimation de  $\theta_Y(\mathbf{h})$  a l'air d'avoir plus de variance ce qui pouvait être attendu puisque  $\hat{a} > 0.5$ . Cette approche peut servir comme outil pour faire ensuite de l'inférence par des méthodes basées sur la vraisemblance en particulier dans le choix de modèle.

Sur la partie gauche de la Figure 3.6, la dépendance asymptotique semble être présente jusqu'à, à peu près, 400km alors qu'on peut penser qu'après il y a indépendance asymptotique. Sur la partie de droite, jusqu'à 600km on garde la dépendance asymptotique. Au delà de cette distance, seules quelques paires de sites apparaissent avec des  $\hat{\theta}_X$  ou  $\hat{\theta}_Y$  proche de 2. Cela veut dire que même à longue distance il y a encore une faible dépendance extrême. Un tel processus pourrait être modélisé par un processus Brown-Resnick ou extrémal- $t$  et leurs versions inverses pour la partie asymptotiquement indépendante.

En plus de permettre un meilleur choix de modèle, notre procédure peut aussi servir pour améliorer les méthodes basées sur le maximum de vraisemblance en donnant un point de départ à la procédure d'optimisation et pour permettre ainsi d'améliorer le temps de calcul mais aussi l'efficacité statistique.

### 3.3 Estimation semi-paramétrique pour des processus max-stables spatio-temporels

Dans cette partie, nous proposons deux méthodes d'estimation des paramètres de processus max-stables spatio-temporels basées sur le madogramme, soit en estimant séparément les paramètres liés à l'espace et ceux liés au temps, soit en estimant conjointement tous les paramètres.

On se place ici dans le cadre de processus max-stables avec séparabilité spectrale, c'est-à-dire

$$X(\mathbf{s}, t) = \max\{\delta X(\mathbf{s} - \boldsymbol{\tau}, t - 1), (1 - \delta)H(\mathbf{s}, t)\}, \quad (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}. \quad (3.10)$$

Dans ce cas là, nous pouvons généraliser le  $\lambda$ -madogramme au cas spatio-temporel :

**Proposition 4** ( *$\lambda$ -madogramme pour processus max-stable spatio-temporel avec séparabilité spectrale*)

Dans le cas de l'Equation (3.10), pour  $\mathbf{h} \in \mathbb{R}^2$  et  $l \in \mathbb{R}$ , le  $\lambda$ -madogramme spatio-temporel pour tout  $\lambda \in (0, 1)$  est donné par

$$\nu_\lambda(\mathbf{h}, l) = \frac{(1 - \lambda)\{V_{\mathbf{0}, \mathbf{h} - l\boldsymbol{\tau}}(\lambda, (1 - \lambda)\delta^{-l})\} + 1 - \delta^l}{(1 - \lambda)\{1 + V_{\mathbf{0}, \mathbf{h} - l\boldsymbol{\tau}}(\lambda, (1 - \lambda)\delta^{-l})\} + 1 - \delta^l} - c(\lambda), \quad (3.11)$$

où  $c(\lambda) = \frac{3}{2(1 + \lambda)(2 - \lambda)}$ .

Nous allons maintenant décrire deux schémas pour l'estimation des paramètres des processus max-stables spatio-temporels en nous basant sur l'expression du  $F$ -madogramme spatio-temporel. On peut vérifier :

$$\nu_F(\mathbf{h}, l) = \frac{1}{2} - \frac{1}{\theta(\mathbf{h}, l) + 1}. \quad (3.12)$$

Nous considérons que le processus  $X := \{X(\mathbf{s}, t) : (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}\}$  est un processus stationnaire max-stable spatio-temporel. Il est observé à des stations sur une grille régulière

$$S_n = \{\mathbf{s}_i : i = 1, \dots, n^2\} = \{(x, y), x, y \in \{1, \dots, n\}\},$$

et à des temps équidistants donnés par  $\{t_1, \dots, t_T\} = \{1, \dots, T\}$ . Ce type d'observations a été adopté dans la littérature (cf Davis et al. (2013b), Buhl & Klüppelberg (2016) et Buhl et al. (2019)). Nous présentons brièvement les 2 procédures d'estimation des paramètres du processus  $X$ .

## Schéma 1

On note  $\boldsymbol{\psi} = (\boldsymbol{\psi}^{(s)}, \boldsymbol{\psi}^{(t)})$  le vecteur des paramètres à estimer du processus  $X$ , où  $\boldsymbol{\psi}^{(s)}$  et  $\boldsymbol{\psi}^{(t)}$  sont respectivement les paramètres spatiaux et temporels du processus. Dans ce schéma d'estimation, on considère séparément comment le processus évolue dans l'espace à un certain temps fixé et comment il évolue dans le temps à une station fixée. Donc,  $\boldsymbol{\psi}^{(s)}$  et  $\boldsymbol{\psi}^{(t)}$  peuvent être estimés séparément.

Plus précisément, on note  $\mathcal{H} \subset [0, \infty)$  et  $\mathcal{K} \subset [0, \infty)$  les ensembles finis de distances temporelles et spatiales sur lesquels l'estimation est basée. L'ensemble  $\mathcal{B}_h$  contient toutes les paires  $\mathcal{S}_n$  qui conduisent à la même distance  $h \in \mathcal{H}$ , i.e.,

$$\mathcal{B}_h = \{(\ell, p) \in \{1, \dots, n^2\}^2 : \|\mathbf{s}_\ell - \mathbf{s}_p\| = \|\mathbf{h}\| = h\}.$$

L'inférence se fait alors en 3 étapes :

1. Comme point de départ, nous estimons le  $F$ -madogramme spatial et celui temporel non paramétriquement par leurs versions empiriques. On note par  $\hat{\nu}_F^{(t)}(\mathbf{h})$ ,  $\|\mathbf{h}\| \in \mathcal{H}$  (respectivement  $\hat{\nu}_F^{(s)}(l')$ ,  $l' \in \mathcal{K}$ ) l'estimateur non paramétrique du  $F$  madogramme spatial (resp., temporel). Comme en géostatistique "classique", on calcule  $\hat{\nu}_F^{(t)}(\mathbf{h})$  à partir du  $F$ -madogramme spatio-temporel  $\hat{\nu}_F(\mathbf{h}, l)$  à des distances spatio-temporels  $(\mathbf{h}, 0)$ , c'est-à-dire pour  $\{t_1, \dots, t_T\}$ ,

$$\hat{\nu}_F^{(t)}(\mathbf{h}) = \hat{\nu}_F(\mathbf{h}, 0) = \frac{1}{2|\mathcal{B}_h|} \sum_{\substack{p=1 \\ \|\mathbf{s}_\ell - \mathbf{s}_p\| = \|\mathbf{h}\| = h}}^{n^2} \sum_{\ell=1}^{n^2} |F\{X(\mathbf{s}_\ell, t)\} - F\{X(\mathbf{s}_p, t)\}|, \quad h \in \mathcal{H},$$

où  $|\mathcal{B}_h|$  est le cardinal de l'ensemble  $\mathcal{B}_h$  et  $F$  est la fonction de répartition de la loi de Fréchet standard.

Par ailleurs,  $\hat{\nu}_F^{(s)}(l')$  est calculé à partir du  $F$ -madogramme spatio-temporel  $\hat{\nu}_F(\mathbf{h}, l')$  à des distances spatio-temporelles  $(\mathbf{0}, l')$ , c'est-à-dire pour tout  $\mathbf{s} \in \mathcal{S}_n$

$$\hat{\nu}_F^{(s)}(l') = \hat{\nu}_F(\mathbf{0}, l') = \frac{1}{2(T-l')} \sum_{k=1}^{T-l'} |F\{X(\mathbf{s}, t_k)\} - F\{X(\mathbf{s}, t_{k+l'})\}|, \quad l' \in \mathcal{K}.$$

2. Ensuite l'estimateur global du  $F$ -madogramme purement spatial (resp., temporel)  $\hat{\nu}_F(h)$  (resp.,  $\hat{\nu}_F(l')$ ) est calculé comme la moyenne sur les moments temporels (resp., les localisations spatiales), c'est-à-dire :

$$\hat{\nu}_F(h) = \frac{1}{T} \sum_{\substack{k=1 \\ \|h\|=h}}^T \hat{\nu}_F^{(t_k)}(h), \quad h \in \mathcal{H}. \quad (3.13)$$

$$\widehat{\nu}_F(l') = \frac{1}{n^2} \sum_{\ell=1}^{n^2} \widehat{\nu}_F^{(s_\ell)}(l'), \quad l' \in \mathcal{K}. \quad (3.14)$$

3. Finalement, on applique une procédure des moindres carrés pour estimer les paramètres d'intérêt :

$$\widehat{\boldsymbol{\psi}}^{(s)} = \arg \min_{\boldsymbol{\psi}^{(s)}} \sum_{\|\mathbf{h}\|=h \in \mathcal{H}} \omega^{\mathbf{h}} \left( \widehat{\nu}_F(\mathbf{h}) - \nu_F^{(s)}(\mathbf{h}; \boldsymbol{\psi}^{(s)}) \right)^2, \quad h \in \mathcal{H}, \quad (3.15)$$

$$\widehat{\boldsymbol{\psi}}^{(t)} = \arg \min_{\boldsymbol{\psi}^{(t)}} \sum_{l' \in \mathcal{K}} \omega^{l'} \left( \widehat{\nu}_F(l') - \nu_F^{(t)}(l'; \boldsymbol{\psi}^{(t)}) \right)^2, \quad l' \in \mathcal{K}, \quad (3.16)$$

où  $\nu_F^{(s)}(\mathbf{h}; \boldsymbol{\psi}^{(s)}) = \nu_F(\mathbf{h}, 0; \boldsymbol{\psi}^{(s)})$  et  $\nu_F^{(t)}(l'; \boldsymbol{\psi}^{(t)}) = \nu_F(\mathbf{0}, l'; \boldsymbol{\psi}^{(t)})$  sont respectivement le  $F$ -madogramme théorique spatial et temporel du modèle.  $\omega^{\mathbf{h}} \geq 0$  et  $\omega^{l'} \geq 0$  sont respectivement les poids spatiaux et temporels. Comme on s'attend à ce que les paires distantes spatialement et temporellement aient peu d'influence sur l'estimation des paramètres de dépendance, un choix simple pour que l'on ne considère que les paires ayant une distance spatio-temporelle maximale de  $(r, q)$  est de prendre  $\omega^{\mathbf{h}} = \mathbb{1}_{\{\|\mathbf{h}\| \leq r\}}$ ,  $\omega^{l'} = \mathbb{1}_{\{l' \leq q\}}$ .

A noter que cette approche est proche de celle de Buhl et al. (2019) où les auteurs utilisent l'extrémogramme spatio-temporel.

## Schéma 2

Nous généralisons maintenant le Schéma 1 pour estimer les paramètres spatiaux et temporels conjointement. On considère comment le processus  $X$  évolue conjointement dans le temps et l'espace. En géostatistique "classique", pour un processus stationnaire spatio-temporel  $\{X(\mathbf{s}, t) : (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}\}$ , le semi-variogramme spatio-temporel empirique est défini par

$$\widehat{\gamma}(\mathbf{h}, l) = \frac{1}{2|\mathcal{B}(\mathbf{h}, l)|} \sum_{\mathcal{B}(\mathbf{h}, l)} (X(\mathbf{s}_i, t_i) - X(\mathbf{s}_j, t_j))^2,$$

où  $\mathcal{B}(\mathbf{h}, l) = \{(\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j) : \mathbf{s}_i - \mathbf{s}_j = \mathbf{h} \text{ et } t_i - t_j = l\}$ . En adaptant cet estimateur à notre cadre, on considère la procédure d'estimation suivante

1. D'abord nous estimons non paramétriquement le  $F$ -madogramme spatio-temporel par sa version empirique. L'ensemble  $\mathcal{B}_{(h, l')}$  regroupe toutes les paires de  $\mathcal{S}_n$  qui sont séparées d'une distance spatiale  $h \in \mathcal{H}$  et d'une même distance temporelle  $l' \in \mathcal{K}$ , i.e.,

$$\mathcal{B}_{(h, l')} = \{(\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j) : |t_i - t_j| = l', \|\mathbf{s}_i - \mathbf{s}_j\| = h\}.$$

On estime  $\nu_F(\mathbf{h}, l')$  par

$$\hat{\nu}_F(\mathbf{h}, l') = \frac{1}{2|\mathcal{B}(h, l')|} \sum_{\substack{\mathcal{B}(h, l') \\ \|s_i - s_j\| = \|\mathbf{h}\| = h}} |F\{X(\mathbf{s}_i, t_i) - F\{X(\mathbf{s}_j, t_j)\}|, (h, l') \in \mathcal{H} \times \mathcal{K}, \quad (3.17)$$

où  $|\mathcal{B}(h, l')|$  est le cardinal de  $\mathcal{B}(h, l')$ .

2. Ensuite, on applique une méthode des moindres carrés pour obtenir les paramètres du processus  $\boldsymbol{\psi}$ , i.e.,

$$\hat{\boldsymbol{\psi}} = \arg \min_{\boldsymbol{\psi}} \sum_{l' \in \mathcal{K}} \sum_{\substack{h \in \mathcal{H} \\ \|\mathbf{h}\| = h}} \omega^{\mathbf{h}, l'} (\hat{\nu}_F(\mathbf{h}, l') - \nu_F(\mathbf{h}, l'; \boldsymbol{\psi}))^2, (h, l') \in \mathcal{H} \times \mathcal{K}, \quad (3.18)$$

où  $\omega^{\mathbf{h}, l'} \geq 0$  sont les poids spatio-temporels et  $\nu_F(\mathbf{h}, l'; \boldsymbol{\psi})$  est le  $F$ -madogramme spatio-temporel théorique du modèle.

L'idée derrière le Schéma 2 est que lorsque l'on modélise ou prédit un phénomène spatio-temporel, on peut gagner beaucoup à considérer comment le phénomène évolue dans le temps et l'espace conjointement contrairement au Schéma 1. Le prix à payer est un grand nombre de paramètres à estimer avec une seule procédure.

## Exemple d'estimation d'un processus spatio-temporel Brown-Resnick (sans séparabilité spectrale)

Considérons un processus spatio-temporel Brown-Resnick comme défini en (1.40) de fonction de répartition bivariable donnée en (1.41), où la structure de dépendance est donnée par le semi-variogramme spatio-temporel FBM (*Fractal Brownian Motion*) pour  $\gamma$  :

$$\gamma(h, l') = 2\phi_s h^{\kappa_s} + 2\phi_t l'^{\kappa_t}, \quad (3.19)$$

où  $\phi_t, \phi_s > 0$  déterminent les paramètres d'échelle spatiale et temporelle et  $\kappa_t, \kappa_s \in (0, 2]$  sont les paramètres de lissage spatial et temporel du processus Gaussien sous-jacent. Le  $F$ -madogramme spatio-temporel associé à ce processus est donné par :

$$\nu_F(h, l') = \frac{1}{2} - \frac{1}{2\Phi(\sqrt{\phi_s h^{\kappa_s} + \phi_t l'^{\kappa_t}}) + 1}, \quad (3.20)$$

où  $\theta(h, l') = 2\Phi(\sqrt{\phi_s h^{\kappa_s} + \phi_t l'^{\kappa_t}})$  est la fonction coefficient extrémal spatio-temporel associée. La Figure 3.7 donne une représentation du semi-variogramme

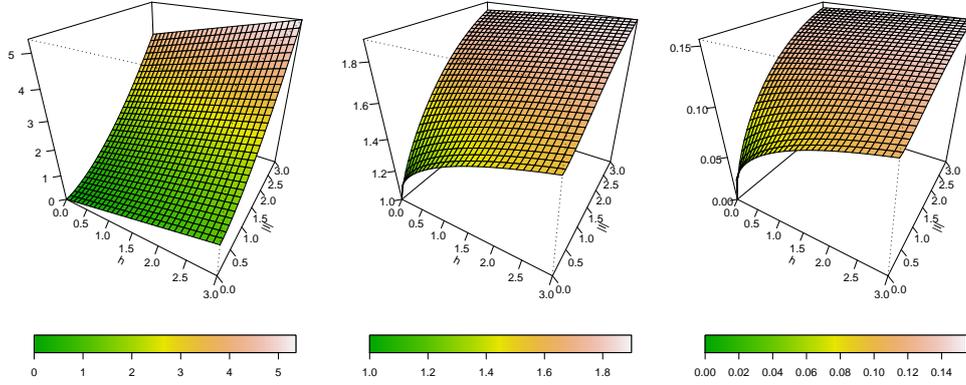


FIGURE 3.7 – Variogramme spatio-temporel FBM  $\gamma(h, l') = 0.8h^{1.5} + 0.4l'$  (gauche), la fonction coefficient extrémal spatio-temporelle associée (milieu) et le  $F$ -madogramme spatio-temporel associé (droite), où  $l' = |l|$  et  $h = \|\mathbf{h}\|$ .

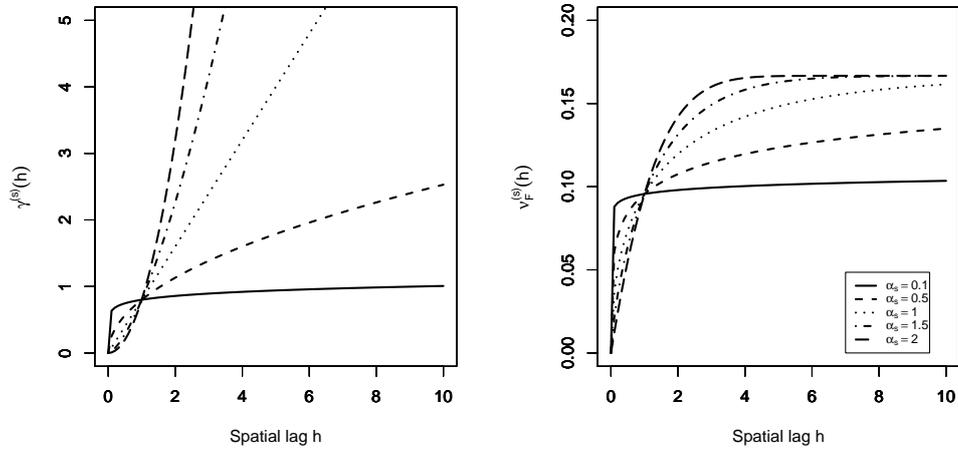


FIGURE 3.8 – Comportement théorique du variogramme FBM  $\gamma^{(s)}(h; \alpha_s) = 0.8h^{\alpha_s}$  et du  $F$ -madogramme spatial associé  $\nu_F^{(s)}(h; \alpha_s) = 0.5 - \{2\Phi(\sqrt{0.4h^{\alpha_s}}) + 1\}^{-1}$ , avec différentes valeurs du paramètre de lissage  $\alpha_s \in \{0.1, 0.5, 1, 1.5, 2\}$ ,  $\mathcal{H} = [0, 10]$  et  $h = \|\mathbf{h}\|$ .

spatio-temporel donné en (3.19) et les mesures de dépendance associées : la fonction  $\theta : \mathbb{R}^2 \times \mathbb{R}^+ \mapsto [1, 2]$  et le  $F$ -madogramme spatio-temporel  $\nu_F : \mathbb{R}^2 \times \mathbb{R}^+ \mapsto [0, 1/6]$ . La Figure 3.8 montre le comportement du semi-variogramme spatial FBM  $\gamma^{(s)}(h; \kappa_s)$  et du  $F$ -madogramme spatial  $\nu_F^{(s)}(h; \kappa_s)$ . Evidemment suivant la valeur du paramètre  $\kappa_s$ , ces mesures montrent une large variété de formes de dépendance.

Avec cette construction, basée sur le Schéma 1, le problème non-linéaire d'optimisation des moindres carrés donné en (3.15) et (3.16) s'écrit :

$$\begin{pmatrix} \widehat{\kappa}_s \\ \widehat{\phi}_s \end{pmatrix} = \arg \min_{\substack{\phi_s > 0 \\ \kappa_s \in (0, 2]}} \sum_{h \in \mathcal{H}} \omega^h \left( \widehat{\nu}_F(h) - \left\{ \frac{1}{2} - \frac{1}{2\Phi(\sqrt{\phi_s h^{\kappa_s}} + 1)} \right\} \right)^2, \quad h \in \mathcal{H}, \quad (3.21)$$

$$\begin{pmatrix} \widehat{\kappa}_t \\ \widehat{\phi}_t \end{pmatrix} = \arg \min_{\substack{\phi_t > 0 \\ \kappa_t \in (0, 2]}} \sum_{l' \in \mathcal{K}} \omega^{l'} \left( \widehat{\nu}_F(l') - \left\{ \frac{1}{2} - \frac{1}{2\Phi(\sqrt{\phi_t l'^{\kappa_t}} + 1)} \right\} \right)^2, \quad l' \in \mathcal{K}. \quad (3.22)$$

Finalement, avec  $(h, l') \in \mathcal{H} \times \mathcal{K}$  et sur la base du Schéma 2, le problème d'estimation des moindres carrés en (3.18) s'écrit sous la forme :

$$\begin{pmatrix} \widehat{\kappa}_s \\ \widehat{\phi}_s \\ \widehat{\kappa}_t \\ \widehat{\phi}_t \end{pmatrix} = \arg \min_{\substack{\phi_s, \phi_t > 0 \\ \kappa_s, \kappa_t \in (0, 2]}} \sum_{l' \in \mathcal{K}} \sum_{h \in \mathcal{H}} \omega^{h, l'} \left( \widehat{\nu}_F(h, l') - \left\{ \frac{1}{2} - \frac{1}{2\Phi(\sqrt{\phi_s h^{\kappa_s} + \phi_t l'^{\kappa_t}} + 1)} \right\} \right)^2. \quad (3.23)$$

## Exemple d'estimation d'un processus spatio-temporel de Smith (avec séparabilité spectrale)

Pour le processus de Smith spatio-temporel (avec séparabilité spectrale), la procédure d'estimation peut être simplifiée puisque les paramètres purement spatiaux et ceux purement temporels peuvent être estimés indépendamment. On considère donc le processus défini en (1.43) avec  $H := \{H(\mathbf{s}, t) : (\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T}\}$  une suite indépendante et identiquement distribuée de processus spatiaux de Smith de matrice de covariance

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}. \quad (3.24)$$

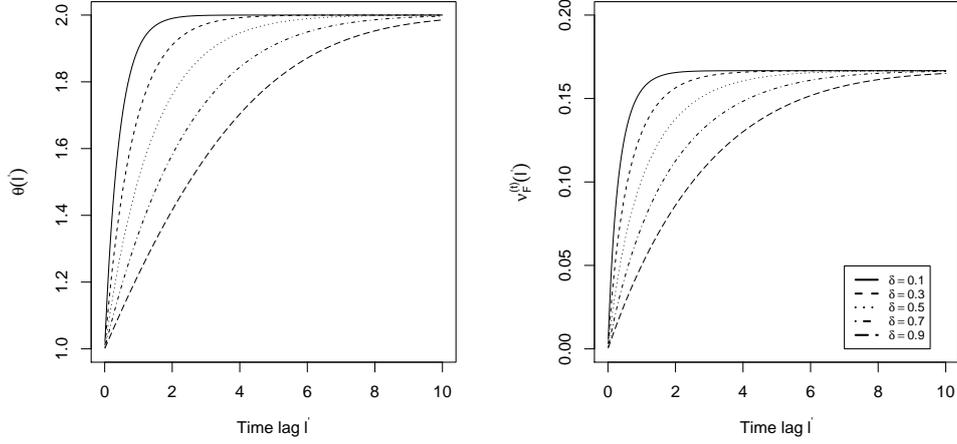


FIGURE 3.9 –  $\theta(l')$  et le  $\nu_F^{(t)}(l')$  associé pour  $\delta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$  basé sur le processus spatio-temporel (1.43) avec  $H$  une suite indépendante et identiquement distribuée de processus spatiaux de Smith.  $\mathcal{K} = [0, 10]$ .

On note par  $\boldsymbol{\psi}$  le vecteur de paramètres à estimer, i.e.,  $\boldsymbol{\psi} = (\sigma_{11}, \sigma_{12}, \sigma_{22}, \boldsymbol{\tau}^t, \delta)^t$ . Il est possible de séparer l'estimation des paramètres : d'abord les paramètres spatiaux  $\boldsymbol{\psi}^{(s)} = (\sigma_{11}, \sigma_{12}, \sigma_{22})^t$  et ensuite ceux temporels  $\boldsymbol{\psi}^{(t)} = (\boldsymbol{\tau}^t, \delta)^t$ . Par conséquence, les Equations (3.15) et (3.16) du Schéma d'estimation 1 peuvent s'écrire :

$$\begin{pmatrix} \hat{\sigma}_{11} \\ \hat{\sigma}_{12} \\ \hat{\sigma}_{22} \end{pmatrix} = \arg \min_{\substack{h \in \mathcal{H} \\ \|\mathbf{h}\|=h}} \sum \omega^{\mathbf{h}} \left( \hat{\nu}_F(\mathbf{h}) - \left\{ \frac{1}{2} - \frac{1}{2\Phi(\sqrt{\mathbf{h}^t \boldsymbol{\Sigma}^{-1} \mathbf{h}/2} + 1)} \right\} \right)^2, \quad h \in \mathcal{H}, \quad (3.25)$$

$$\begin{pmatrix} \hat{\delta} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \end{pmatrix} = \arg \min_{a \in (0,1)} \sum_{l' \in \mathcal{K}} \omega^{l'} \left( \hat{\nu}_F(l') - \left\{ \frac{1}{2} - \frac{1}{\theta(l') + 1} \right\} \right)^2, \quad l' \in \mathcal{K}, \quad (3.26)$$

où  $\theta(l') = \Phi\left(\frac{\beta^*}{2} + \frac{1}{\beta^*} \log(\delta^{-l'})\right) + \delta^{l'} \Phi\left(\frac{\beta^*}{2} + \frac{1}{\beta^*} \log(\delta^{l'})\right) + 1 - \delta^{l'}$ ,  
et  $\beta^* = \sqrt{(\mathbf{0} - l' \boldsymbol{\tau})^t \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{0} - l' \boldsymbol{\tau})}$ .

Pour comprendre le rôle du paramètre temporel  $\delta$ , nous avons tracé dans la Figure 3.9 pour un site fixé  $\mathbf{s} \in \mathcal{S}$ , la fonction coefficient extrémal temporel  $\theta(l')$

et le  $F$ -madogramme temporel associé  $\nu_F^{(t)}$  pour  $\delta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$  pour ce processus. On pose :

$$\Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix},$$

et  $\boldsymbol{\tau} = (1, 1)^t$ . Clairement, quand  $\delta$  augmente, l'indépendance (i.e.,  $\theta(l') \rightarrow 2$ ) apparaît à des distances temporelles plus grandes  $l'$ .

Finalement, basé sur le Schéma 2, l'estimateur des moindres carrés  $\hat{\boldsymbol{\psi}} = (\hat{\sigma}_{11}, \hat{\sigma}_{12}, \hat{\sigma}_{22}, \hat{\tau}_1, \hat{\tau}_2, \hat{\delta})^t$  est donné par :

$$\hat{\boldsymbol{\psi}} = \arg \min_{\boldsymbol{\psi}} \sum_{\substack{l' \in \mathcal{K} \\ \|\mathbf{h}\|=h}} \sum_{h \in \mathcal{H}} \omega^{\mathbf{h}, l'} \left( \hat{\nu}_F(\mathbf{h}, l') - \left\{ \frac{1}{2} - \frac{1}{\theta(\mathbf{h}, l') + 1} \right\} \right)^2, \quad (h, l') \in \mathcal{H} \times \mathcal{K},$$

$$\text{où } \theta(\mathbf{h}, l') = \Phi \left( \frac{\beta}{2} + \frac{1}{\beta} \log(\delta^{-l'}) \right) + \delta^{l'} \Phi \left( \frac{\beta}{2} + \frac{1}{\beta} \log(\delta^{l'}) \right) + 1 - \delta^{l'} \text{ et } \beta = \sqrt{(\mathbf{h} - l' \boldsymbol{\tau})^t \Sigma^{-1} (\mathbf{h} - l' \boldsymbol{\tau})}. \quad (3.27)$$

## Simulations

Dans Abu-Awwad et al. (2021), nous avons réalisé une étude de simulations pour vérifier la qualité d'estimation de la procédure sur le processus Brown-Resnick spatio-temporel ainsi que sur le processus spatio-temporel de Smith avec les différents schémas d'estimation. Dans les deux cas, on observe que l'estimation des paramètres spatiaux est meilleure avec le Schéma d'estimation 1 alors que le Schéma d'estimation 2 est légèrement plus performant pour les paramètres temporels.

# Chapitre 4

## Conclusion et Perspectives

Pour conclure ce document de synthèse, donnons quelques perspectives de recherche. Mes projets de recherche s'orientent autour de deux thèmes principaux : la Théorie des Valeurs Extrêmes univariée et les extrêmes spatiaux.

Le choix de la taille de bloc dans l'approche *Block Maxima* est un problème important, tout comme le choix du seuil dans l'approche PoT. La Formule (1.2) permet d'offrir une solution. En effet, si on se fixe une taille de bloc suffisamment grande, nous obtenons un échantillon de maxima approximativement distribué suivant une distribution GEV de paramètres  $(\mu_{GEV}, \sigma_{GEV}, \xi_{GEV})$ . Si l'approximation par une distribution GEV est valide alors, en considérant comme seuil pour l'approche PoT la valeur  $\mu_{GEV}$ , les excès au delà du seuil sont distribués suivant une GPD de paramètre  $(\sigma_{GPD}, \xi_{GPD})$  avec  $\xi_{GPD} \approx \xi_{GEV}$  et  $\sigma_{GPD} \approx \sigma_{GEV}$ . L'idée est donc, pour différentes tailles de bloc, de considérer une distance entre les paramètres obtenus pour l'ajustement des maxima par une GEV et l'ajustement des excès au delà du seuil  $\mu_{GEV}$  par une GPD. On cherche ensuite le minimum pour obtenir une taille de bloc adéquat.

La Figure 4.1 montre un exemple d'utilisation du critère appliqué à des données de précipitations à Avignon. Le graphique semble indiquer qu'un bon choix de taille de bloc serait 40. A noter que l'on peut aussi considérer comme critère la différence entre des niveaux de retours basés sur l'approche *Block Maxima* et sur l'approche PoT.

Le projet qui va m'occuper les prochaines années est l'utilisation des techniques d'apprentissage profond appliquées aux processus spatiaux. Nous avons commencé à travailler dessus avec Véronique Maume-Deschamps et Manaf Ahmed grâce à un financement de post-doctorat par le projet PAUSE. La première partie de ce projet a récemment été soumise. Il s'agit d'utiliser une modification d'un réseau de

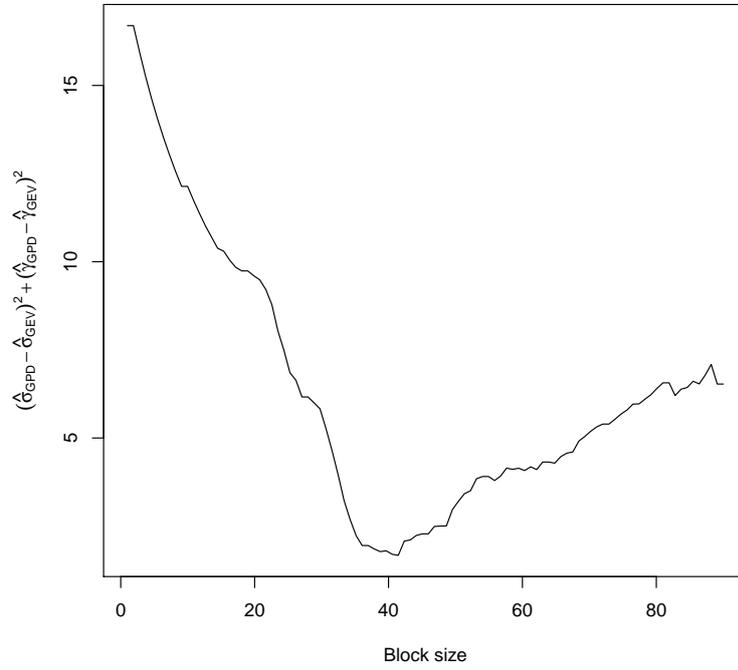


FIGURE 4.1 – Critère pour le choix de la taille de bloc appliqué aux données de précipitations à Avignon.

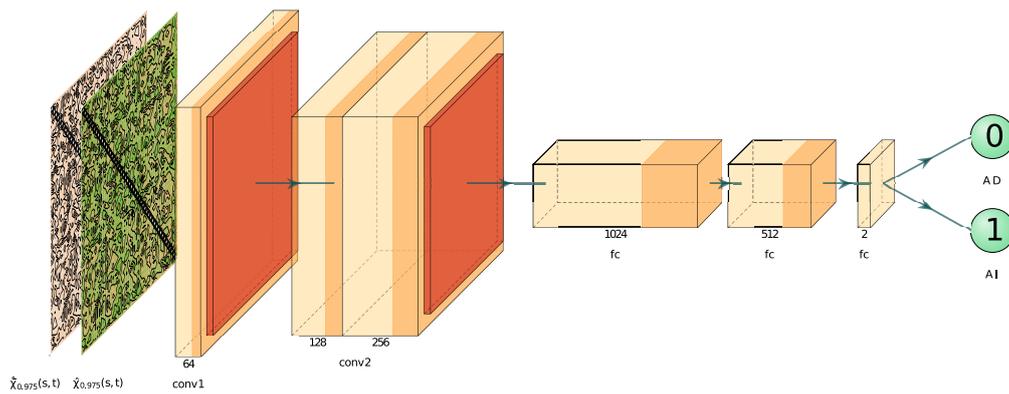


FIGURE 4.2 – Architecture du réseau de neurones artificiels pour détecter la dépendance/indépendance asymptotique.

neurones convolutifs pour détecter l'indépendance ou la dépendance asymptotique. L'architecture générale du réseau de neurones utilisée est représentée dans la Figure 4.2.

Nous simulons une grande quantité de processus suivant différents modèles. Pour chaque simulation, nous estimons les fonctions  $\chi(\cdot)$  et  $\bar{\chi}(\cdot)$ . Ce sont ces fonctions estimées qui entraînent l'algorithme. Les premiers résultats sont très encourageants. Sur les simulations tests, l'algorithme classe correctement 97% des modèles asymptotiquement indépendants mais aussi des modèles asymptotiquement dépendants même lorsque les simulations tests sont effectuées avec des modèles qui n'étaient pas parmi les modèles utilisés dans la partie entraînement. Nous travaillons aussi sur le même type d'algorithme pour reconnaître le modèle spatial.

Par la suite, nous pensons poursuivre en proposant une méthodologie qui travaille directement sur les données, sans passer par les fonctions  $\chi$  et  $\bar{\chi}$ , ainsi qu'une méthodologie qui puisse traiter la non-stationnarité.

Sur des thèmes annexes, nous avons deux projets (avec des financements de thèse CIFRE) liés à l'apprentissage statistique. Le premier avec la société Ginkyo fait suite à deux contrats de recherche pour la valorisation financière de joueurs de football. A partir des données de jeu d'un joueur, nous avons proposé plusieurs modélisations de son prix de vente. Le projet actuel vise à inclure une dimension temporelle et être capable de faire des projections ou des trajectoires pour les joueurs des centres de formation. Le deuxième projet avec la société Natixis vise à construire des méthodes basées sur l'apprentissage statistique pour capter des liens de causalité afin de dégager des prédictions sur des entreprises, des secteurs ou des variables macroéconomiques dans le cadre des propagations des crises. L'objectif est donc de développer de nouveaux modèles causaux pour les risques bancaires (risques de défaut, risque de contagion ...).

# Bibliographie

- Abu-Awwad, A., V. Maume-Deschamps, & P. Ribereau  
2019. Fitting spatial max-mixture processes with unknown extremal dependence class : an exploratory tool. *TEST*, 29(2) :479–522.
- Abu-Awwad, A., V. Maume-Deschamps, & P. Ribereau  
2020. Censored pairwise likelihood-based tests for mixing coefficient of spatial max-mixture models. *Revista de Investigacion Operacional*, 41(1) :1–26.
- Abu-Awwad, A., V. Maume-Deschamps, & P. Ribereau  
2021. Semiparametric estimation for space-time max-stable processes : F-madogram-based estimation approaches. *Accepté à Statistical Inference for Stochastic Processes*.
- Ahmed, M., V. Maume-Deschamps, P. Ribereau, & C. Vial  
2019. Spatial risk measure for max-stable and max-mixture processes. *Stochastics*, 92(7) :1005–1020.
- Ahmed, M., V. Maume-Deschamps, P. Ribereau, & C. Vial  
2021. A semi-parametric estimation for max-mixture spatial processes. *Soumis pour publication*.
- Azzalini, A.  
1985. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12 :171–178.
- Bacro, J.-N., L. Bel, & C. Lantuéjoul  
2010. Testing the independence of maxima : from bivariate vectors to spatial extreme fields. *Extremes*, 13(2) :155–175.
- Bacro, J.-N., C. Gaëtan, & G. Toulemonde  
2016. A flexible dependence model for spatial extremes. *Journal of Statistical Planning and Inference*, 172 :36–52.

- Beirlant, J., Y. Goegebeur, J. Segers, & J. Teugels  
2004. *Statistics of Extremes : Theory and Applications*. Wiley Series in Statistic.
- Bel, L., J.-N. Bacro, & C. Lantuéjoul  
2008. Assessing extremal dependence of environmental spatial fields. *Environmetrics*, 19(2) :163–182.
- Blanchet-Scaillet, C., D. Dorobantu, L. Gay, V. Maume-Deschamps, & P. Ribereau  
2018. Risk assessment using suprema data. *Stochastic Environmental Research and Risk Assessment*, 32 :2839–2848.
- Buhl, S., R. A. Davis, C. Klüppelberg, & C. Steinkohl  
2019. Semiparametric estimation for isotropic max-stable space-time processes. *Bernoulli*, 25(4A) :2508–2537.
- Buhl, S. & C. Klüppelberg  
2016. Anisotropic brown-resnick space-time processes : estimation and model assessment. *Extremes*, 19(4) :627–660.
- Cho, Y. B., R. A. Davis, & S. Ghosh  
2016. Asymptotic properties of the empirical spatial extremogram. *Scandinavian Journal of Statistics*, 43(3) :757–773.
- Coles, S.  
2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag.
- Coles, S., J. Heffernan, & J. Tawn  
1999. Dependence measures for extreme value analyses. *Extremes*, 2(4) :339–365.
- Cooley, D., P. Naveau, & P. Poncet  
2006. *Dependence in Probability and Statistics*, volume 187, chapter Variograms for spatial max-stable random fields, Pp. 373–390. Springer.
- Davis, R. A., C. Klüppelberg, & C. Steinkohl  
2013a. Max-stable processes for modeling extremes observed in space and time. *Journal of the Korean Statistical Society*, 42(3) :399–414.
- Davis, R. A., C. Klüppelberg, & C. Steinkohl  
2013b. Statistical inference for max-stable processes in space and time. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 75(5) :791–819.

- Davis, R. A. & T. Mikosch  
 2009. The extremogram : A correlogram for extreme events. *Bernoulli*, 15(4) :977–1009.
- Davis, R. A. & S. I. Resnick  
 1989. Basic properties and prediction of max-arma processes. *Advances in Applied Probability*, 21(4) :781–803.
- Davison, A., R. Huser, & E. Thibaud  
 2013. Geostatistics of dependent and asymptotically independent extremes. *Mathematical Geosciences*, 45(5) :511–529.
- Davison, A. C., S. A. Padoan, & M. Ribatet  
 2012. Statistical modeling of spatial extremes. *Statistical science*, 27(2) :161–186.
- De Haan, L.  
 1984. A spectral representation for max-stable processes. *The Annals of Probability*, 12(4) :1194–1204.
- Diebolt, J., A. Guillo, P. Naveau, & P. Ribereau  
 2008. Improving probability-weighted moment methods for the generalized extreme value distribution. *REVSTAT*, 6(1) :33–50.
- Diebolt, J., A. Guillo, & I. Rached  
 2007. Approximation of the distribution of excesses through a generalized probability weighted moments method. *Journal of Statistical Planning and Inference*, 137 :841–857.
- Dischel, B.  
 1998a. At least : a model for weather risk, weather risk special report. *Energy and Power risk management*, (March) :30–32.
- Dischel, B.  
 1998b. Black-scholes won't do it, weather special report. *Energy and Power risk management*, (October) :9–9.
- Embrechts, P., C. Klüppelbreg, & T. Mikosch  
 1997. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag.
- Embrechts, P., E. Koch, & C. Robert  
 2016. Space-time max-stable models with spectral separability. *Advances in Applied Probability*, 48(A) :77–97.

- Falk, M., J. Hüsler, & R. D. Reiss  
2010. *Laws of small numbers : Extremes and rare events, 3rd edition.*
- Gaetan, C. & X. Guyon  
2010. *Spatial Statistics and modeling.* Springer.
- Genton, M. G.  
2004. *Skew-Elliptical Distributions and Their applications : A journey beyond normality.*
- Greenwood, J. A., J. M. Landwehr, N. C. Matalas, & J. R. Wallis  
1979. Probability weighted moments : definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15 :1049–1054.
- Guillou, A., P. Naveau, & A. Schorgen  
2014. Madogram and asymptotic independence among maxima. *REVSTAT-Statistical Journal*, 12(2).
- Heffernan, J. E.  
2000. A directory of coefficients of tail dependence. *Extremes*, 3(3) :279–290.
- Hosking, J. R. M. & J. R. Wallis  
1987. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29 :339–349.
- Hosking, J. R. M., J. R. Wallis, & E. F. Wood  
1985. Estimation of the generalized extreme value distribution by the method of probability weighted moments. *Technometrics*, 27 :251–261.
- Huser, R. & A. C. Davison  
2014. Space-time modelling of extreme events. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 76(2) :439–461.
- Kabluchko, Z., M. Schlather, & L. De Haan  
2009. Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37(5) :2042–2065.
- Katz, M., M. Parlange, & P. Naveau  
2002. Extremes in hydrology. *Advances in Water Resources*, 25 :1287–1304.
- Katz, R. W.  
. Precipitation as a chain-dependent process. *Journal of Applied Meteorology*, 16 :671–676.

- Klein Tank, A. M. G., J. B. Wijngaard, G. P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A. F. V. van Engelen, E. Forland, M. Mielus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar, J. Antonio López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L. V. Alexander, & P. Petrovic 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the european climate assessment. *International Journal of Climatology*, 22(12) :1441–1453.
- Laaidi, K., M. Pascal, M. Ledrans, A. Le Tertre, S. Medina, C. Casério, J.-C. Cohen, J. Manach, P. Beaudeau, & P. Empereur Bissonnet 2004. Le système français d’alerte canicule et santé 2004 (sacs 2004) un dispositif intégré au plan national canicule. *Bulletin épidémiologique hebdomadaire*, 30-31 :134–136.
- Landwehr, J., N. Matalas, & J. Wallis 1979. Probability weighted moments compared with some traditional techniques in estimating gumbel parameters and quantiles. *Water Resources Research*, 15 :1055–1064.
- Ledford, A. W. & J. A. Tawn 1996. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1) :169–187.
- Ledford, A. W. & J. A. Tawn 1997. Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 59(2) :475–499.
- Naveau, P., A. Guillou, D. Cooley, & J. Diebolt 2009. Modelling pairwise dependence of maxima in space. *Biometrika*, 96(1) :1–17.
- Opitz, T. 2013. Extremal t processes : Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis*, 122 :409–413.
- Opitz, T., J. N. Bacro, & P. Ribereau 2015. The spectrogram : A threshold-based inferential tool for extremes of stachastic processes. *Electronic Journal of Statistics*, 9 :842–868.

- Papastathopoulos, I. & A. J. Tawn  
 2013. Extended generalised pareto models for tail estimation. *Journal of Statistical Planning and Inference*, 143 :131–143.
- Pickands, J.  
 1975. Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1) :119–131.
- Ribereau, P., E. Masiello, & P. Naveau  
 2016. Skew generalized extreme value distribution : Probability weighted moments estimation and application to block maxima procedure. *Communications in Statistics : Theory and Method*, 45(17) :5037–5052.
- Schlather, M.  
 2002. Models for stationary max-stable random fields. *Extremes*, 5(1) :33–44.
- Schlather, M. & J. A. Tawn  
 2003. A dependence measure for multivariate and spatial extreme values : Properties and inference. *Biometrika*, 90(1) :139–156.
- Smith, R. L.  
 1985. Maximum likelihood estimation in a class of non-regular cases. *Biometrika*, 72 :67–90.
- Tian, Y. & B. J. Reich  
 2020. A bayesian semi-parametric mixture model for bivariate extreme value analysis with application to precipitation forecasting. *to appear in Statistica Sinica*.
- Turkman, M. A., K. F. Turkman, P. de Zea Bermudez, P. S., P. Perira, & M. Carvalho  
 2020. Calibration methods for spatial data. *Preprint*.
- Vrac, M., M. Stein, & K. Hayhoe  
 2007. Statistical downscaling of precipitation through a nonhomogeneous stochastic weather typing approach. *Climate Research*, 34 :169–184.
- Wadsworth, J. L. & J. A. Tawn  
 2012. Dependence modelling for spatial extremes. *Biometrika*, 99(2) :253–272.
- Wilks, D.  
 2006. *Statistical methods in the atmospheric sciences (second edition)*. Elsevier.