On some semi-parametric methods for extensions of spatial max-stable processes.

Véronique Maume-Deschamps
Joint works with
Manaf Ahmed, Abdul-Fattah Abu-Awwad, Pierre Ribereau, Céline Vial
Séminaire de Probabilité et statistique du LMB.

26 février 2018

Plan

1. Introduction
2. Max-stable and max-mixture processes
   - Extreme spatial processes
   - The λ madogram
3. Statistical methods based on the λ-madogram
   - Estimation of the parameters
   - Selection criterium for the mixing coefficient a
4. Conclusion

Plan

1. Introduction
2. Max-stable and max-mixture processes
3. Statistical methods based on the λ-madogram
4. Conclusion

Modeling environmental data

We are interested in the modelization of environmental data. e.g.

- precipitation,
- temperature,
- wind speed,
- ...

$\mathcal{S}$ is a region of interest. $X(s)$, $s \in \mathcal{S}$ random variable at each location $s \Rightarrow$ spatial process $(X(s))_{s\in\mathcal{S}}$,

Spatial processes

Stationary spatial processes:

$$(X(s_1), \ldots, X(s_k)) \stackrel{\mathcal{L}}{=} (X(s_1 + h), \ldots, X(s_k + h))$$

for any $s_i \in \mathcal{S}$, $i = 1, \ldots, k$ and h with $s_i + h \in \mathcal{S}$.

Spatial processes

Stationary spatial processes:

$$(X(s_1), \ldots, X(s_k)) \stackrel{\mathcal{L}}{=} (X(s_1 + h), \ldots, X(s_k + h))$$

for any $s_i \in \mathcal{S}$, $i = 1, \ldots, k$ and h with $s_i + h \in \mathcal{S}$.

In the Gaussian case, the dependence structure, is caracterised by the covariogram: $\mathrm{Cov}(X(s), X(s+h)) = \gamma(h)$, depends only on $\|h\|$ in the isotropic case.

Plan

1 Introduction

2 Max-stable and max-mixture processes
- Extreme spatial processes
- The λ madogram

3 Statistical methods based on the λ-madogram

4 Conclusion

Max-stable spatial processes

Gaussian processes not well suited for e.g. rainfall, wind... $\Rightarrow$ max-stable processes, unit Fréchet margins, dependence structure given by the exponent measure function V, that is:

$$\mathbb{P}(X(s) \leq x) = e^{-\frac{1}{x}},\ \mathbb{P}\big(X(s) \leq x_1, X(t) \leq x_2\big) = \exp(-V_{s,t}(x_1, x_2)).$$

V is homogeneous of degree -1.
The process is isotropic if $V_{s,t}(x_1, x_2)$ depends only on $h = \|t - s\|$.
Max-stable processes have been defined by De Haan (1984).

Max-stable spatial processes

Gaussian processes not well suited for e.g. rainfall, wind... $\Rightarrow$ max-stable processes, unit Fréchet margins, dependence structure given by the exponent measure function V, that is:

$$\mathbb{P}(X(s) \le x) = e^{-\frac{1}{x}}, \ \mathbb{P}\big(X(s) \le x_1, X(t) \le x_2\big) = \exp(-V_{s,t}(x_1, x_2)).$$

V is homogeneous of degree -1.
The process is isotropic if $V_{s,t}(x_1, x_2)$ depends only on $h = \|t - s\|$.
Max-stable processes have been defined by De Haan (1984).

Spectral representation (De Haan):

$$X(s) = \max_{i \ge 1} W_i(s)/\xi_i,$$

where $\{\xi_i, i \ge 1\}$ is an i.i.d unit rate Poisson point process on $(0, \infty)$ and $\{W_i, i \ge 1\}$ are i.i.d copies of a positive random field W, independent of ξ_i.

Max-stable spatial processes

Gaussian processes not well suited for e.g. rainfall, wind... $\Rightarrow$ max-stable processes, unit Fréchet margins, dependence structure given by the exponent measure function V, that is:

$$\mathbb{P}(X(s) \leq x) = e^{-\frac{1}{x}},\ \mathbb{P}\big(X(s) \leq x_1, X(t) \leq x_2\big) = \exp(-V_{s,t}(x_1, x_2)).$$

V is homogeneous of degree -1.
The process is isotropic if $V_{s,t}(x_1, x_2)$ depends only on $h = \|t - s\|$.
Max-stable processes have been defined by De Haan (1984).

Max-stable processes are Asymptotically Dependent in the sense that either $X(s)$ and $X(s+h)$ are independent or

$$\chi(h) = \lim_{u \to 1} \mathbb{P}\big(F(X(s)) > u | F(X(s+h)) > u\big) > 0.$$

Max-stable spatial processes

Gaussian processes not well suited for e.g. rainfall, wind... $\Rightarrow$ max-stable processes, unit Fréchet margins, dependence structure given by the exponent measure function V, that is:

$$\mathbb{P}(X(s) \leq x) = e^{-\frac{1}{x}}, \ \mathbb{P}\big(X(s) \leq x_1, X(t) \leq x_2\big) = \exp(-V_{s,t}(x_1, x_2)).$$

V is homogeneous of degree -1.
The process is isotropic if $V_{s,t}(x_1, x_2)$ depends only on $h = \|t - s\|$.
Max-stable processes have been defined by De Haan (1984).

Max-stable processes are Asymptotically Dependent in the sense that either $X(s)$ and $X(s+h)$ are independent or

$$\chi(h) = \lim_{u \to 1} \mathbb{P}\big(F(X(s)) > u | F(X(s+h)) > u\big) > 0.$$

Our purpose:
Semi / non-parametric estimations for models allowing various dependence structures.

Multivariate distribution function

The multivariate distribution function of a max-stable process X has following expression:

$$\mathbb{P}\big(X(s_1) \leq x_1, ..., X(s_k) \leq x_k\big) = \exp\{-V(x_1, ..., x_k)\},$$

where V is called the exponent measure and homogeneous of order -1.

The density function writes in terms of the derivatives of V.

Extreme coefficient

For any pair $(X(s), X(s+h))$, the bivariate distribution function satisfies for any $x > 0$:

$$\mathbb{P}\big(X(s) \leq x, X(s+h) \leq x\big) = \exp\{-\Theta(h)/x\},$$

where, $\Theta(h) = V(1,1) \in [1,2]$ is the Extremal coefficient function introduced in Schlather and Tawn (2002).
Θ is related to the χ function:

$$\chi(h) = 2 - \Theta(h).$$

Examples of max-stable processes I.

Smith (1990) Model

$$V_h(x_1, x_2) = \frac{1}{x_1}\Phi\left(\frac{\tau(h)}{2} + \frac{1}{\tau(h)}\log\frac{x_2}{x_1}\right) + \frac{1}{x_2}\Phi\left(\frac{\tau(h)}{2} + \frac{1}{\tau(h)}\log\frac{x_1}{x_2}\right);$$

$\tau(h) = \sqrt{h^T \Sigma^{-1} h}$ and $\Phi(\cdot)$ the standard normal cumulative distribution function.

Schlather (2002) Model

$$V_h(x_1, x_2) = \frac{1}{2}\left(\frac{1}{x_1} + \frac{1}{x_2}\right)\left[1 + \sqrt{1 - 2(\rho(h)+1)\frac{x_1 x_2}{(x_1 + x_2)^2}}\right].$$

$+$ parametric models for ρ.

Examples of max-stables processes II.

Extremal-t process proposed in Opitz (2013) and Ribatet & Sedki (2013)

$$V_h(x_1, x_2) = \frac{1}{x_1} T_{v+1}\left(\alpha\rho(h) + \alpha\left(\frac{x_2}{x_1}\right)^{1/v}\right) + \frac{1}{x_2} T_{v+1}\left(\alpha\rho(h) + \alpha\left(\frac{x_1}{x_2}\right)^{1/v}\right)$$

where T_v is the Student distribution with v degrees of freedom and $\alpha(h) = [v + 1/\{1 - \rho^2(h)\}]^{1/2}$.

Inverse max-stable processes

Let X' be a max-stable process as above, consider

$$X(s) = g(X'(s)) = -\frac{1}{\log\{1 - e^{-1/X'(s)}\}} \quad s \in \mathcal{S}.$$

X is called inverse max-stable process, defined by Ledford and Tawn (1996).It has unit Fréchet margin and its bivariate survivor function satisfies:

$$\mathbb{P}(X(s_1) > x_1, X(s+h) > x_2) = \exp\big(-V_h(g(x_1), g(x_2))\big).$$

Inverse max-stable processes are Asymptotically Independent in the sense that $\chi(h) = 0$ for any h.
The exponent measure of X' is called the exponent measure of X and denoted V_X. The extremal coefficient of X' is called the extremal coefficient of X and denoted Θ_X.

Max-mixture processes

Wadsworth and Tawn (1997) proposed to mix max-stable and inverse max-stable processes, studied also by Bacro *et al.* (2016): Let X be a max-stable process, with exponent measure function V_h^X. Let Y be an inverse max-stable process with and exponent measure function V_h^Y. Let $a \in [0,1]$ and define

$$Z(s) = \max\{aX(s), (1-a)Y(s)\}, \quad s \in \mathcal{S}.$$

Max-mixture processes

Wadsworth and Tawn (1997) proposed to mix max-stable and inverse max-stable processes, studied also by Bacro *et al.* (2016): Let X be a max-stable process, with exponent measure function V_h^X. Let Y be an inverse max-stable process with and exponent measure function V_h^Y. Let $a \in [0,1]$ and define

$$Z(s) = \max\{aX(s), (1-a)Y(s)\}, \quad s \in \mathcal{S}.$$

Z has unit Fréchet marginals. Its bivariate distribution function is given by $\mathbb{P}\big(Z(s) \leq z_1, Z(s+h) \leq z_2\big) =$

$$e^{-aV_h^X(z_1,z_2)}\left[e^{\frac{-(1-a)}{z_1}} + e^{\frac{-(1-a)}{z_2}} - 1 + e^{-V_h^Y(g_a(z_1),g_a(z_2))}\right],$$

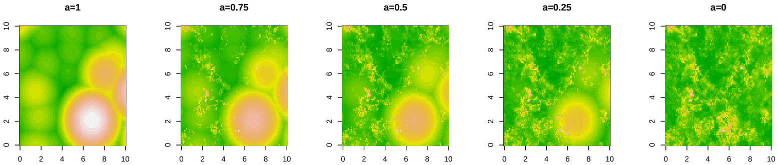
where $g_a(z) = g(\frac{z}{1-a})$.

Max-mixture processes

Wadsworth and Tawn (1997) proposed to mix max-stable and inverse max-stable processes, studied also by Bacro *et al.* (2016): Let X be a max-stable process, with exponent measure function V_h^X. Let Y be an inverse max-stable process with and exponent measure function V_h^Y. Let $a \in [0,1]$ and define

$$Z(s) = \max\{aX(s), (1-a)Y(s)\}, \quad s \in \mathcal{S}.$$

Examples: (Plots on the logarithm scale with different values of a. X is an isotropic Smith process and Y is an isotropic inverted extremal-t process)

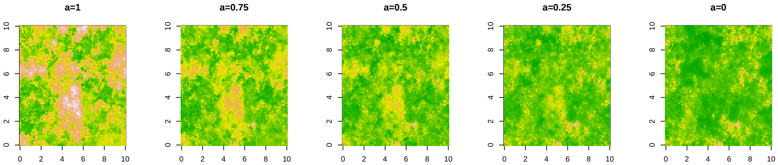


Max-mixture processes

Wadsworth and Tawn (1997) proposed to mix max-stable and inverse max-stable processes, studied also by Bacro *et al.* (2016): Let X be a max-stable process, with exponent measure function V_h^X. Let Y be an inverse max-stable process with and exponent measure function V_h^Y. Let $a \in [0,1]$ and define

$$Z(s) = \max\{aX(s), (1-a)Y(s)\}, \quad s \in \mathcal{S}.$$

Examples: (Plots on the logarithm scale according different values of mixing coefficient a. X is an isotropic extremal-t process and Y is an isotropic inverted extremal-t process)



Definition of the λ-madogram

When Gaussian processes are involved, the variogram is a useful and widely used tool:

$$\gamma(h) = \frac{1}{2}\operatorname{var}(X(s) - X(s+h)).$$

The processes that we are studying have Fréchet marginal laws $\Longrightarrow$ no finite variance. The λ-madogram, proposed e.g. in Cooley *et al.* is used instead: for $\lambda > 0$,

$$\nu_\lambda(h) = \frac{1}{2}\mathbb{E}(|F^\lambda(X(s+h)) - F^\lambda(X(s))|),$$

where F is the unit Fréchet distribution function (so that $F(X(s)) \sim \mathcal{U}([0,1])$).

λ-madogram for max-mixture processes

Property

Let X be a max-stable process, with extremal coefficient function $\Theta_X(h)$, and Y be an inverted max-stable process with extremal coefficient function $\Theta_Y(h)$. Let $a \in [0,1]$ and $Z = \max(aX, (1-a)Y)$. Then, the F^{λ}-madogram of the spatial max-mixture process $Z(s)$ is given by

$$\nu_\lambda(h) = \frac{\lambda}{1+\lambda} - \frac{2\lambda}{a(\Theta_X(h)-1)+1+\lambda} + \frac{\lambda}{a\Theta_X(h)+\lambda} - \frac{\lambda\Theta_Y(h)}{(1-a)\Theta_Y(h)+a\Theta_X(h)+\lambda} \times \beta\left(\frac{a\Theta_X(h)+\lambda}{1-a}, \Theta_Y(h)\right).$$

Plan

1 Introduction

2 Max-stable and max-mixture processes

3 Statistical methods based on the λ-madogram
- Estimation of the parameters
- Selection criterium for the mixing coefficient a

4 Conclusion

Least squared methods

We consider $Z_i, i = 1, ..., N$ copies of Z,

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

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

where $\widehat{F}$ denotes the empirical distribution function (or any consistent estimator of the distribution function F).
We have $\mathbb{E}[\widetilde{Q}_i(h,\lambda)] = \nu_\lambda(h)$. We shall estimate either

- the parameters of the max-mixture model (for a given model) or
- give non parametric estimations of $\Theta_X(h)$, $\Theta_Y(h)$ and provide a decision criterium for the parameter a.

It will be based on the minimization of

$$\sum_{i=1}^{N}(Q_i(h,\lambda) - \nu_\lambda(h))^2.$$

Parametric max-mixture models

$Z = \max(aX, (1-a)Y) \Longrightarrow$ chose a model for X and for Y.
Recall that the bivariate distribution function is given by

$$e^{-aV_h^X(z_1,z_2)}\left[e^{\frac{-(1-a)}{z_1}} + e^{\frac{-(1-a)}{z_2}} - 1 + e^{-V_h^Y(g_a(z_1),g_a(z_2))}\right].$$

You may also write a formula for all the finite dimensional joint distribution functions $\Longrightarrow$ theoretically you may compute the density function but it is practically untracktable, so that maximum likelihood estimation is not an option.

Estimation of parameters

Method usually used (developped by Padoan *et al.* (2010)) : Maximum Composite Likelihood Estimation. The composite likelihood is the product of the pairwise likelihood. Then the parameter estimation is done by maximizing

$$\ell_N = \sum_{(s_k, s_j)} \sum_{i=1}^{N} \log f(Z_i(s_k), Z_i(s_j)) \Longrightarrow \widehat{\psi}_L,$$

where $Z_i, i = 1, ..., N$ are independent (or α-mixing) copies of Z, observed at locations s_k, $k = 1, \ldots, M$.

Estimation of parameters

Method usually used (developped by Padoan *et al.* (2010)) : Maximum Composite Likelihood Estimation. The composite likelihood is the product of the pairwise likelihood. Then the parameter estimation is done by maximizing

$$\ell_N = \sum_{(s_k,s_j)} \sum_{i=1}^{N} \log f(Z_i(s_k), Z_i(s_j)) \Longrightarrow \widehat{\psi}_L,$$

where $Z_i, i = 1, ..., N$ are independent (or α-mixing) copies of Z, observed at locations s_k, $k = 1, \dots, M$.
Adjust several models and retain the one with the smallest CLIC:

$$\text{CLIC} = -2\left(\ell_N(\widehat{\psi}_L) - tr[\mathcal{J}(\widehat{\psi}_L)\mathcal{H}^{-1}(\widehat{\psi}_L)]\right)$$

where $\mathcal{H}$ is the sensitivity matrix and $\mathcal{J}$ is the variability matrix. Both intervene in the asymptotic variance of the estimator.

Estimation of parameters

Alternatively, we propose to minimize the squared madogram difference:

$$\mathcal{L}_N = \sum_{h} \sum_{\|s_k - s_j\| = h} \sum_{i=1}^{N} (Q_i(h,1) - \nu(h))^2 \Longrightarrow \widehat{\psi}_M .$$

Consistency of the estimators, under additional identifiability assumption.

Estimation of parameters

Alternatively, we propose to minimize the squared madogram difference:

$$\mathcal{L}_N = \sum_h \sum_{\|s_k - s_j\| = h} \sum_{i=1}^{N} (Q_i(h,1) - \nu(h))^2 \Longrightarrow \widehat{\psi}_M.$$

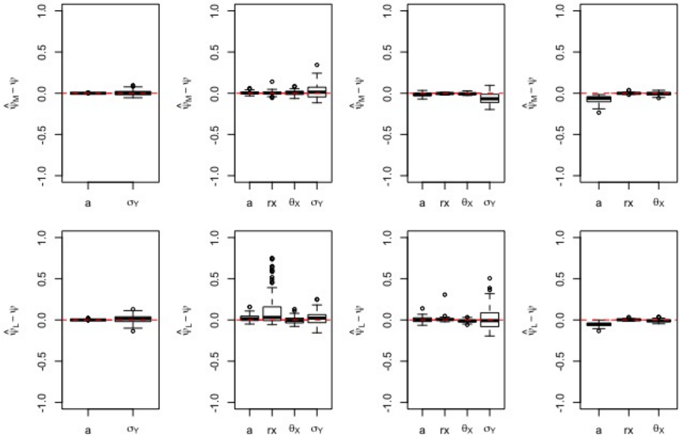
Consistency of the estimators, under additional identifiability assumption.
Adjust several models and retain the one with the smallest selection criterium:

$$\text{SC} = \log \mathcal{L}_N + \frac{2k(k+1)}{(N-k-1)}$$

where k is the number of parameters in the model.

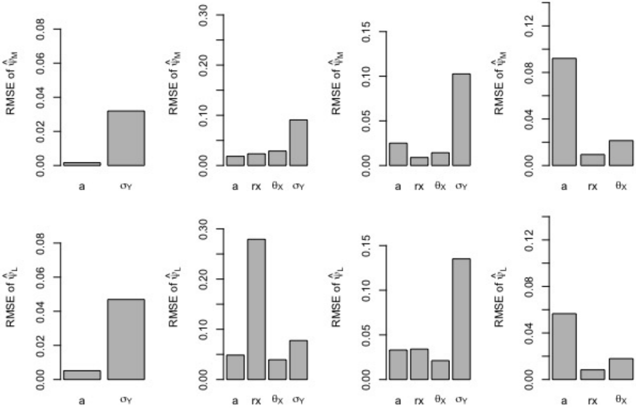
Simulation study

Simutation of a max-mixture between a truncated Schlater process X and an inverse Smith process Y. $N = 1000$ i.i.d observations on 50 sites. This experiment is replicated $J = 100$ times.



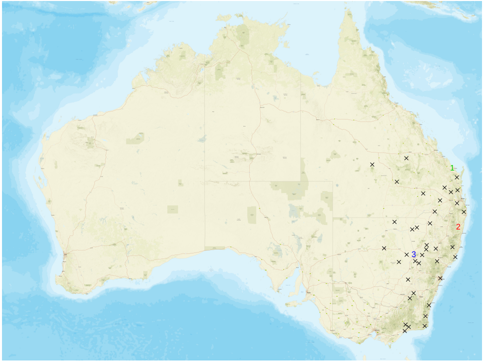
Simulation study

Simutation of a max-mixture between a truncated Schlater process X and an inverse Smith process Y. $N = 1000$ i.i.d observations on 50 sites. This experiment is replicated $J = 100$ times.



Real data example

Rainfall in east cost of Australia, also used in Bacro *et al.*



Daily rainfall amounts at 39 locations over years 1982-2016 occurring during April - September.

The data exploration shows no anisotopy nor temporal dependence.

Models.

Real data example

Model		a	θ_X	r_X	θ_Y	σ_Y	SC
MM1	CL	0.262	1217.3	1364.5	3102.4	3.457	**6807406**
	LS	0.259	1285.7	1390.0	5794.8	2.013	**1.917034**
MM2	CL	0.248	31.16	70.15	998.84		7924609
	LS	0.185	35.51	48.14	871.19		1.917234
		θ_X	r_X				
M1	CL	931	307.86				7926261
	LS	1270	255.64				1.945177
		θ_X	σ_X	θ_Y	σ_Y		
M2	CL	931.02	3.078663				7926261
	LS	361.36	1.90816				1.96165
M3	CL			1644.76	2.702282		7918643
	LS			1383.08	1.394928		1.924574
M4	CL		85.34				8016633
	LS		193.43				1.988753
M5	CL					256.39	7988838
	LS					334.60	1.929235

A model free procedure

First, for a fixed a, estimate non parametrically $\Theta_X(h)$ and $\Theta_Y(h)$ using the λ-madogram with two different values of λ.
We may write the λ-madogram as a function of a, λ, Θ_X and Θ_Y, that is $\nu_{F^\lambda}(h) = \Phi(a, \lambda, \Theta_X(h), \Theta_Y(h))$.

Madogram.

$$\hat{\mathbf{\Theta}}^{a}_{NLS}(h) = \arg\min_{\theta \in [1,2]^2} \sum_{i=1,\dots,N} [Q_i(h,\lambda_1) - \Phi(a,\lambda_1,\theta_1,\theta_2)]^2 + \left[Q_i(h,\lambda_1') - \Phi(a,\lambda_1',\theta_1,\theta_2)\right]^2.$$

A model free procedure

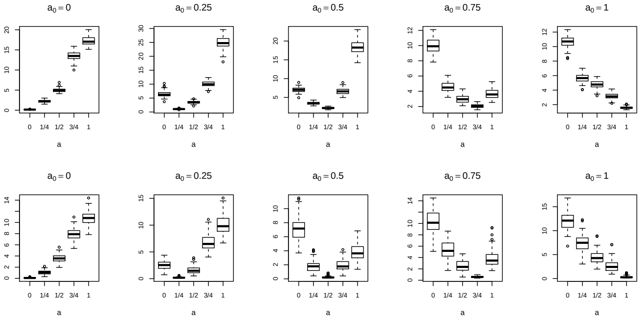
Secondly, chose a realizing the least squared difference between empirical and theoretical λ-madogram, with a third value of λ.
Assume that the Z_i's are observed at locations $s_1, \ldots, s_K$ and let h_j be the pairwise distances between the s_j's.

$$\hat{\nu}_\lambda(h_j) = \frac{1}{2N} \sum_{i=1}^{N} Q_i(h_j, \lambda).$$

$$\mathrm{DC}(a) = \sum_{h_j} \left[\frac{\hat{\nu}_{\lambda_2}(h_j)}{\Phi(a, \lambda_2, \widehat{\Theta}_X(h_j), \widehat{\Theta}_Y(h_j))} - 1 \right]^2$$

Estimate a as the argmin of $\mathrm{DC}(a)$. Consistent under additional identifiability assumption.

Simulation study



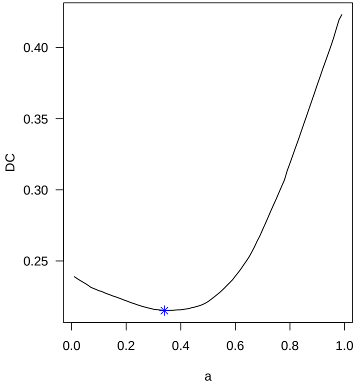
First line: max-mixture between a truncated Schlater and an inverted truncated Schlater.

Second line: max-mixture between a truncated Schlater and an inverted extremal-t process.

50 sites, $N = 2000$ independent replications. Each experiment is repeated 100 times.

Real data example

Rainfall data in the same region as before.

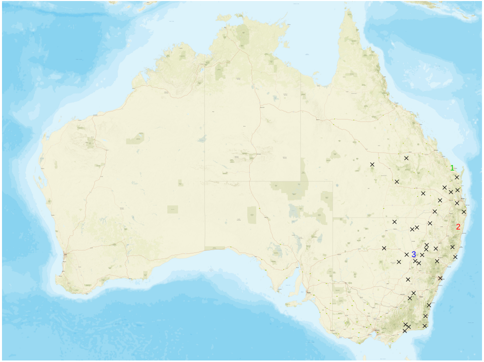


Daily rainfall data at 38 sites occuring during April - Spetember over the years 1972 - 2016.

$\widehat{a} = 0.34.$

Prediction with non parametric estimation

3 unused stations s^* in the estimation.



Prediction with non parametric estimation

Compare the estimations of $\mathbb{P}(Z(s^*) > z|Z(s) > z)$ by adjusting a parametric model vs the non parametric estimations of a, Θ_X and Θ_Y.

$$\mathbb{P}[Z(s_0^*) > z|Z(s_0) > z] = \frac{1 - 2e^{-\frac{1}{z}} + e^{-\frac{a\Theta_X(h_0)}{z}} \left\{ -1 + 2e^{-\frac{1-a}{z}} + \left[1 - e^{-\frac{1-a}{z}} \right]^{\Theta_Y(h_0)} \right\}}{1 - e^{-\frac{1}{z}}}.$$

Where $h_0 = \|s_0^* - s_0\|$.

Prediction with non parametric estimation

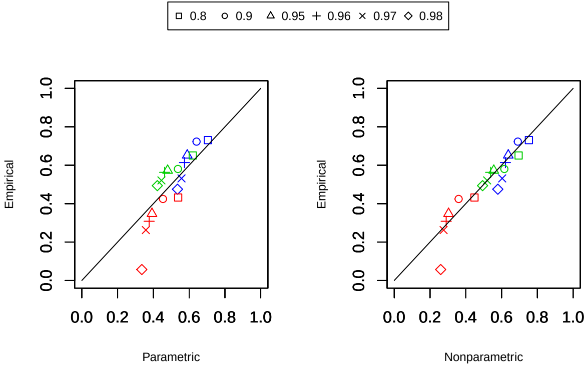


Figure: Diagnostic P-P plots for threshold excess conditional probabilities for the three unused sites obtained by both approaches; the best parametric model as judged by the CLIC and our nonparametric approach. Green: site 1; red: site 2; blue: site 3.

Plan

1 Introduction

2 Max-stable and max-mixture processes

3 Statistical methods based on the λ-madogram

4 Conclusion

Conclusion

- Importance of the dependence structure for spatial processes.
- The λ-madogram captures the main dependence informations of max-mixture processes.
- We have used it as an alternative to composite likelihood estimation.
- It is also useful in model-free estimation.
- Our estimations are consistent.

To Do Extension to spatio-temporal processes.

To Do Asymptotic normallity of the estimators.

References I

Bacro, J.-N., Gaëtan, C., and Toulemonde, G. (2016).
A flexible dependence model for spatial extremes.
Journal of Statistical Planning and Inference, 172:36–52.

Cooley, D., Naveau, P., and Poncet, P. (2006).
Variograms for spatial max-stable random fields.
Dependence in probability and statistics, pages 373–390.

De Haan, L. (1984).
A spectral representation for max-stable processes.
The annals of probability, pages 1194–1204.

References II

Ledford, A. W. and Tawn, J. A. (1996).
Statistics for near independence in multivariate extreme values.
Biometrika, 83(1):169–187.

Opitz, T. (2013).
Extremal t processes: Elliptical domain of attraction and a spectral representation.
Journal of Multivariate Analysis, 122:409–413.

Padoan, S. A., Ribatet, M., and Sisson, S. A. (2010).
Likelihood-based inference for max-stable processes.
Journal of the American Statistical Association, 105(489):263–277.

References III

Ribatet, M. and Sedki, M. (2013).
Extreme value copulas and max-stable processes.
Journal de la Société Française de Statistique, 154(1):138–150.

Schlather, M. (2002).
Models for stationary max-stable random fields.
Extremes, 5(1):33–44.

Schlather, M. and Tawn, J. (2002).
Inequalities for the extremal coefficients of multivariate extreme value distributions.
Extremes, 5(1):87–102.

References IV

- Schlather, M. and Tawn, J. A. (2003).
  A dependence measure for multivariate and spatial extreme values: Properties and inference.
  Biometrika, 90(1):139–156.

- Smith, R. L. (1990).
  Max-stable processes and spatial extremes.
  Unpublished manuscript, 205.

- Wadsworth, J. L. and Tawn, J. A. (2012).
  Dependence modelling for spatial extremes.
  Biometrika, 99(2):253.

Thank you

Merci pour votre attention.

N'oubliez pas qu'AMIES peut vous aider dans vos collaborations avec les entreprises

LES MATHS
VECTEUR D'INNOVATION

Models

Back.

- MM1: max-mixture between a truncated Schlater and a Brown-Resnik.
- MM2: max-mixture between a truncated Schlater and an inverted Smith.
- M1: a truncated Schlater.
- M2: a Brown-Resnik.
- M3: an inverted Brown-Resnik.
- M4: a Smith process.
- M5: an inverted Smith.