Spatio-temporal modeling of urban extreme rainfall events at high resolution

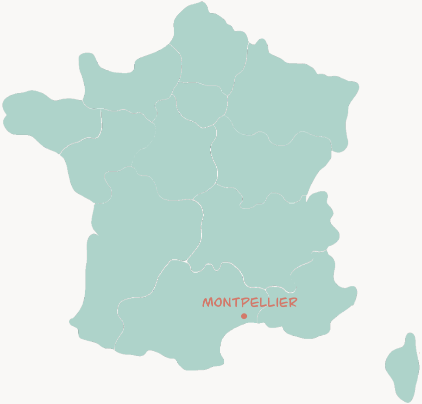
EGU, Vienna, April 28, 2025

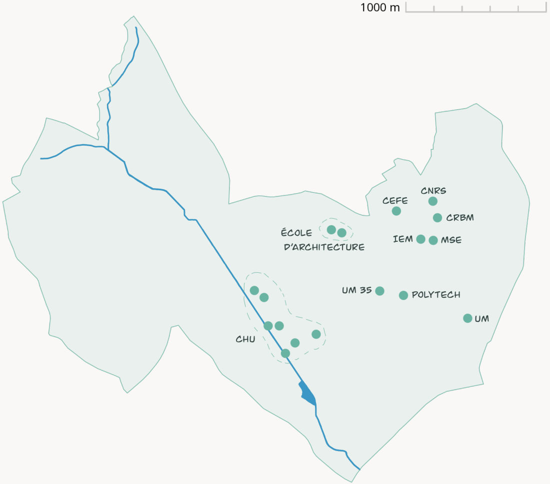
Chloé SERRE-COMBE [1] Nicolas MEYER [1] Thomas OPITZ [2] Gwladys TOULEMONDE [1]

[1] *IMAG, Université de Montpellier, LEMON Inria*
[2] *INRAE, BioSP, Avignon*

Floods in Montpellier, September 2022 and August 2015 (*Midi Libre*)

- Mediterranean events, localized rainfall
- Urban area, flood risks

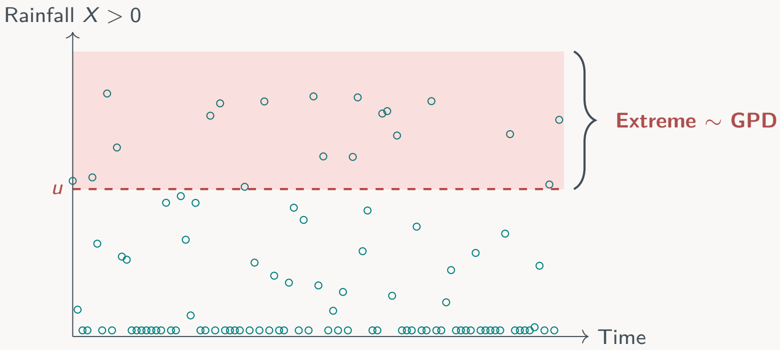




- **Study area:** Verdanson water catchment
- **Source:** Urban observatory of HydroScience Montpellier (HSM)[1]
- **Time period:** $[2019, 2023[$
- **Temporal resolution:** 5 minutes
- **Spatial resolution:** 77 m to 1531 m

$\mathcal{S} = \{\textbf{17 rain gauges}\} \subset \mathbb{R}^2$ and $\mathcal{T} \subset \mathbb{R}_+$

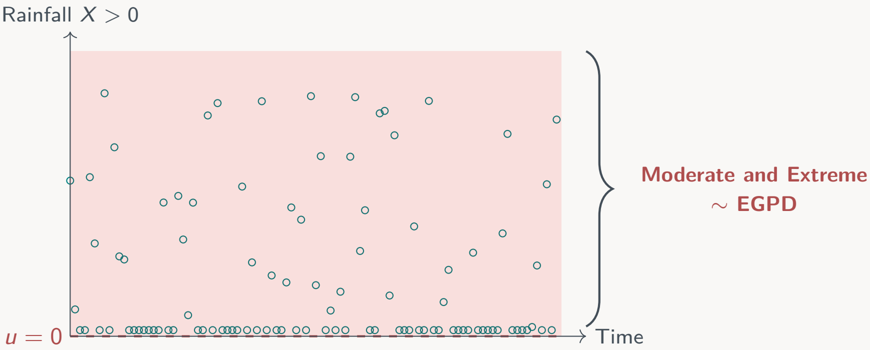
[1] Finaud-Guyot et al., 2023



- **Generalized Pareto Distribution (GPD)**[2]

$$X \mid X > u \sim \underbrace{H_\xi}_{\mathrm{GPD}(\xi,\sigma,u)} \text{ with } \xi \in \mathbb{R},\ \sigma > 0.$$

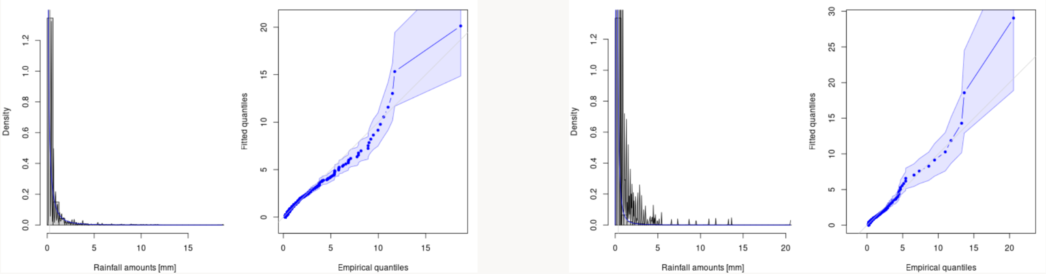
[2] PICKANDS III, 1975



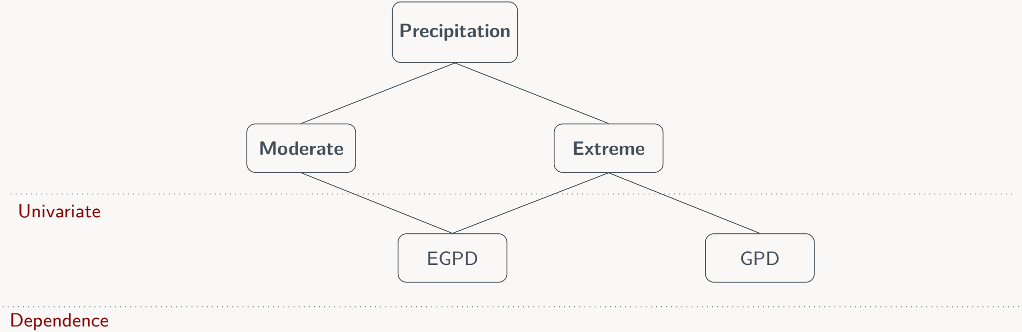
- **E**xtended **G**eneralized **P**areto **D**istribution (**EGPD**)[3]

$$X \sim \underbrace{G(H_\xi)}_{\text{EGPD}(\xi,\sigma,\kappa)} \text{ with } G(x) = x^\kappa,\ \kappa > 0$$

[3] NAVEAU et al., 2016



EGPD fitting for two rain gauges, CRBM (left) and CNRS (right) with left-censoring and 95% CI


Precipitation
Moderate
Extreme
Univariate
EGPD
GPD
Dependence

Rainfall random field: $\boldsymbol{X} = \{X_{\boldsymbol{s},t},\ (\boldsymbol{s},t) \in \mathcal{S} \times \mathcal{T}\}$, stationary and isotropic.

Let $\Lambda_{\mathcal{S}} \subset \mathbb{R}_+^2$ and $\Lambda_{\mathcal{T}} \subset \mathbb{R}_+$ be sets of spatial and temporal lags respectively.

Rainfall random field: $\boldsymbol{X} = \{X_{\boldsymbol{s},t},\ (\boldsymbol{s}, t) \in \mathcal{S} \times \mathcal{T}\}$, stationary and isotropic.

Let $\Lambda_{\mathcal{S}} \subset \mathbb{R}_+^2$ and $\Lambda_{\mathcal{T}} \subset \mathbb{R}_+$ be sets of spatial and temporal lags respectively.

Variogram (Matheron, 1963)

$$\gamma(\boldsymbol{h}, \tau) = \frac{1}{2}\mathrm{Var}(X_{\boldsymbol{s},t} - X_{\boldsymbol{s}+\boldsymbol{h},t+\tau})$$

- Quantifies variability
- Higher $\gamma(\boldsymbol{h}, \tau) \to$ weaker dependence

$\boldsymbol{h} \in \Lambda_{\mathcal{S}},\ \tau \in \Lambda_{\mathcal{T}}$

Rainfall random field: $\boldsymbol{X} = \{X_{\boldsymbol{s},t},\ (\boldsymbol{s},t) \in \mathcal{S} \times \mathcal{T}\}$, stationary and isotropic.

Let $\Lambda_{\mathcal{S}} \subset \mathbb{R}_+^2$ and $\Lambda_{\mathcal{T}} \subset \mathbb{R}_+$ be sets of spatial and temporal lags respectively.

Variogram (Matheron, 1963)

$$\gamma(\boldsymbol{h},\tau) = \frac{1}{2}\mathrm{Var}(X_{\boldsymbol{s},t} - X_{\boldsymbol{s}+\boldsymbol{h},t+\tau})$$

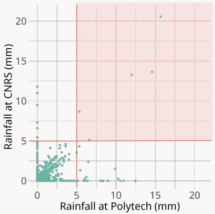
- Quantifies variability
- Higher $\gamma(\boldsymbol{h},\tau) \rightarrow$ weaker dependence

Extremogram (Davis and Mikosch, 2009)

$$\chi(\boldsymbol{h},\tau) = \lim_{q\to 1} \mathbb{P}(X^*_{\boldsymbol{s},t} > q \mid X^*_{\boldsymbol{s}+\boldsymbol{h},t+\tau} > q)$$

- Measures tail dependence
- Higher $\chi(\boldsymbol{h},\tau) \rightarrow$ stronger dependence

$\boldsymbol{h} \in \Lambda_{\mathcal{S}}$, $\tau \in \Lambda_{\mathcal{T}}$, $X^*_{\boldsymbol{s},t}$ uniform margins.

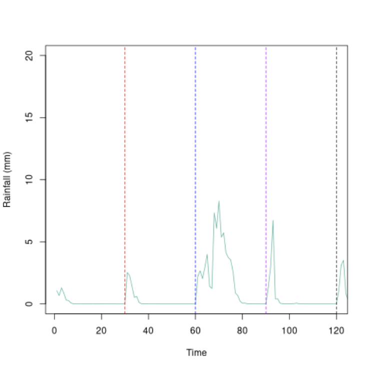


Definition (de Fondeville and Davison, 2018)

For all $\boldsymbol{s} \in \mathcal{S}$ and $t \in \mathcal{T}$, a risk function $r(\boldsymbol{X}) = X_{\boldsymbol{s}_0,t_0}$,

$$u^{-1} X_{\boldsymbol{s},t} \mid X_{\boldsymbol{s}_0,t_0} > u \xrightarrow{d} Y_{\boldsymbol{s},t} \quad \text{with} \quad Y_{\boldsymbol{s},t} = R_{\boldsymbol{s},t} e^{W_{\boldsymbol{s},t} - W_{\boldsymbol{s}_0,t_0} - \gamma(\boldsymbol{s}-\boldsymbol{s}_0, t-t_0)},$$

where $(\boldsymbol{s}_0, t_0)$ is a space-time location, u is a high threshold, $R_{\boldsymbol{s},t} \sim \text{Pareto}(1)$, $W_{\boldsymbol{s},t}$ is a Gaussian process.



Dependence measures

- **r-variogram:**

$$\gamma_r(\boldsymbol{h}, \tau) = \frac{1}{2} Var\left(W_{\boldsymbol{s}_0,t_0} - W_{\boldsymbol{s}_0+\boldsymbol{h},t_0+\tau}\right)$$

- **r-extremogram:**

$$\chi_r(\boldsymbol{h}, \tau) = \lim_{q \to 1} \mathbb{P}(X^*_{\boldsymbol{s}_0+\boldsymbol{h},t_0+\tau} > q)$$

Spatio-temporal extremogram with a Brown-Resnick dependence

Let $\boldsymbol{h} \in \Lambda_{\mathcal{S}}$ and $\tau \in \Lambda_{\mathcal{T}}$. We have

$$\chi(\boldsymbol{h}, \tau) = 2\left(1 - \phi\left(\sqrt{\frac{1}{2}\gamma(\boldsymbol{h}, \tau)}\right)\right)$$

with ϕ the std normal c.d.f. and γ the variogram of $\boldsymbol{W}$.

Spatio-temporal extremogram with a Brown-Resnick dependence

Let $\boldsymbol{h} \in \Lambda_{\mathcal{S}}$ and $\tau \in \Lambda_{\mathcal{T}}$. We have

$$\chi(\boldsymbol{h}, \tau) = 2\left(1 - \phi\left(\sqrt{\frac{1}{2}\gamma(\boldsymbol{h}, \tau)}\right)\right)$$

with ϕ the std normal c.d.f. and γ the variogram of $\boldsymbol{W}$.

Separable model: Fractional Brownian motion with additive separability.

$$\frac{\gamma(\boldsymbol{h}, \tau)}{2} = \beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}$$

Spatio-temporal extremogram with a Brown-Resnick dependence

Let $\boldsymbol{h} \in \Lambda_{\mathcal{S}}$ and $\tau \in \Lambda_{\mathcal{T}}$. We have

$$\chi(\boldsymbol{h},\tau) = 2\left(1 - \phi\left(\sqrt{\frac{1}{2}\gamma(\boldsymbol{h},\tau)}\right)\right)$$

with ϕ the std normal c.d.f. and γ the variogram of $\boldsymbol{W}$.

Separable model: Fractional Brownian motion with additive separability.

$$\frac{\gamma(\boldsymbol{h},\tau)}{2} = \beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}$$

- Efficient estimation using WLSE (Buhl et al., 2019)

Spatio-temporal extremogram with a Brown-Resnick dependence

Let $\boldsymbol{h} \in \Lambda_{\mathcal{S}}$ and $\tau \in \Lambda_{\mathcal{T}}$. We have

$$\chi(\boldsymbol{h}, \tau) = 2\left(1 - \phi\left(\sqrt{\frac{1}{2}\gamma(\boldsymbol{h}, \tau)}\right)\right)$$

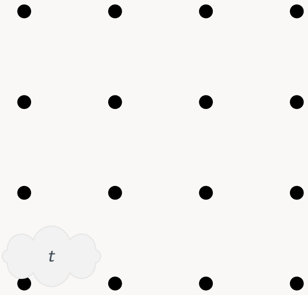
with ϕ the std normal c.d.f. and γ the variogram of $\boldsymbol{W}$.

Separable model: Fractional Brownian motion with additive separability.

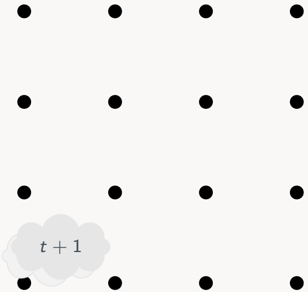
$$\frac{\gamma(\boldsymbol{h}, \tau)}{2} = \beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}$$

- Efficient estimation using WLSE (Buhl et al., 2019)
- **Not realistic**

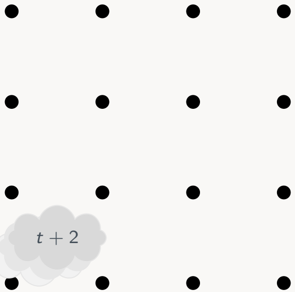
Separable model



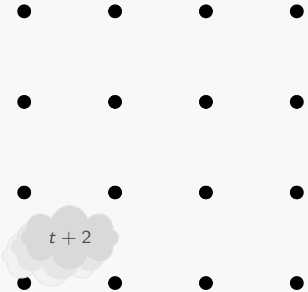
Separable model



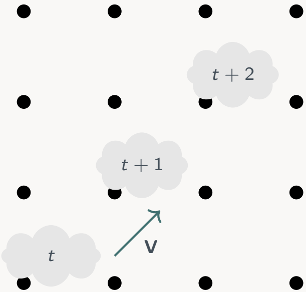
Separable model



Separable model



Non-separable model



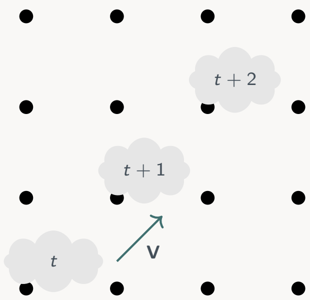
Towards more realistic modeling: introduce advection $\boldsymbol{V}$ to relax separability

$$\gamma_L(\boldsymbol{h}, \tau) = \gamma(\boldsymbol{h} - \tau \boldsymbol{V}, \tau)$$

$$\Rightarrow \quad \frac{1}{2}\gamma_L(\boldsymbol{h}, \tau) = \beta_1 \|\boldsymbol{h} - \tau \boldsymbol{V}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}$$

- Parameters: $\boldsymbol{\Theta} = (\beta_1, \beta_2, \alpha_1, \alpha_2, \boldsymbol{V})$

Non-separable model



Goal: Estimate variogram parameters $\boldsymbol{\Theta} = (\beta_1, \beta_2, \alpha_1, \alpha_2, \boldsymbol{V})$

Joint exceedances count

$$K_{\boldsymbol{h},\tau}(u) = \sum_{(\boldsymbol{s}_i,\boldsymbol{s}_j)\in\mathcal{N}(\boldsymbol{h})} \sum_{(t_i,t_j)\in\mathcal{N}(\tau)} \mathbb{1}_{\{X_{\boldsymbol{s}_i,t_i}>u,\, X_{\boldsymbol{s}_j,t_j}>u\}}$$

Asymptotic distribution (large u) on r-Pareto:

$$K_{\boldsymbol{h},\tau} \sim \mathcal{B}\left(\sum_{(\boldsymbol{s}_0,t_0)\in\mathcal{R}} \#\{\boldsymbol{s} : (\boldsymbol{s},\boldsymbol{s}_0) \in \mathcal{N}(\boldsymbol{h})\},\ \chi_{r,\boldsymbol{\Theta}}(\boldsymbol{h},\tau) \right)$$

with

- $\mathcal{R}$ is a subset of spatio-temporal conditioning points
- $\mathcal{N}(\boldsymbol{h})$ the set of pairs of sites separated by $\boldsymbol{h}$
- $\mathcal{N}(\tau)$ the set of pairs of times separated by τ

Asymptotic distribution (large u) on r-Pareto:

$$K_{\boldsymbol{h},\tau} \sim \mathcal{B}\left(\sum_{(\boldsymbol{s}_0,t_0)\in\mathcal{R}} \#\{\boldsymbol{s} : (\boldsymbol{s},\boldsymbol{s}_0) \in \mathcal{N}(\boldsymbol{h})\},\ \chi_{r,\boldsymbol{\Theta}}(\boldsymbol{h},\tau)\right)$$

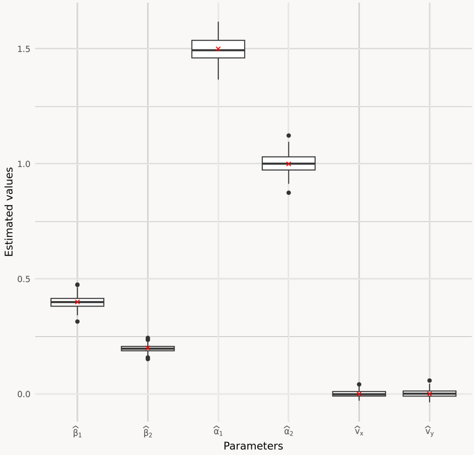
Composite log-likelihood maximization

$$\ell_C(\boldsymbol{\Theta}) \propto \sum_{(\boldsymbol{s}_0,t_0)\in\mathcal{R}} \sum_{(\boldsymbol{s},\boldsymbol{s}_0)\in\mathcal{N}(\boldsymbol{h})} \sum_{(t,t_0)\in\mathcal{N}(\tau)} k_{\boldsymbol{s},t} \log \chi_{r,\boldsymbol{\Theta}}(\boldsymbol{h},\tau) + (1-k_{\boldsymbol{s},t}) \log(1-\chi_{r,\boldsymbol{\Theta}}(\boldsymbol{h},\tau))$$

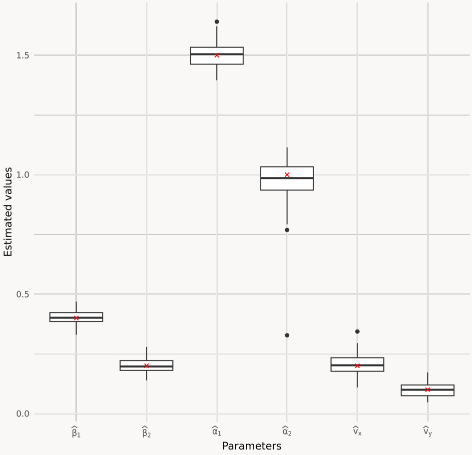
with $k_{\boldsymbol{s},t} = \mathbb{1}_{\{X_{\boldsymbol{s},t}>u,\, X_{\boldsymbol{s}_0,t_0}>u\}}$, for large u

- Optimization: maximization of $\ell_C(\boldsymbol{\Theta})$ with respect to $\boldsymbol{\Theta}$

Validation on simulations

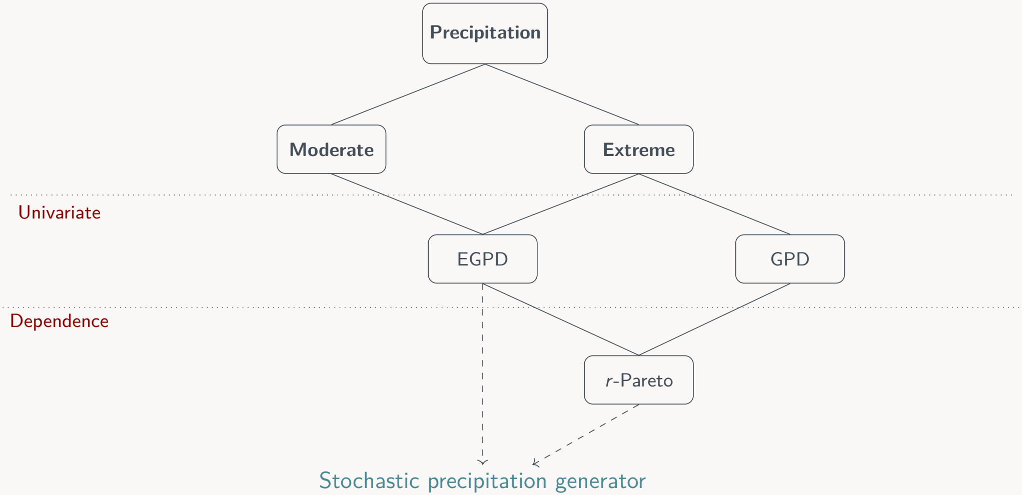


Without advection

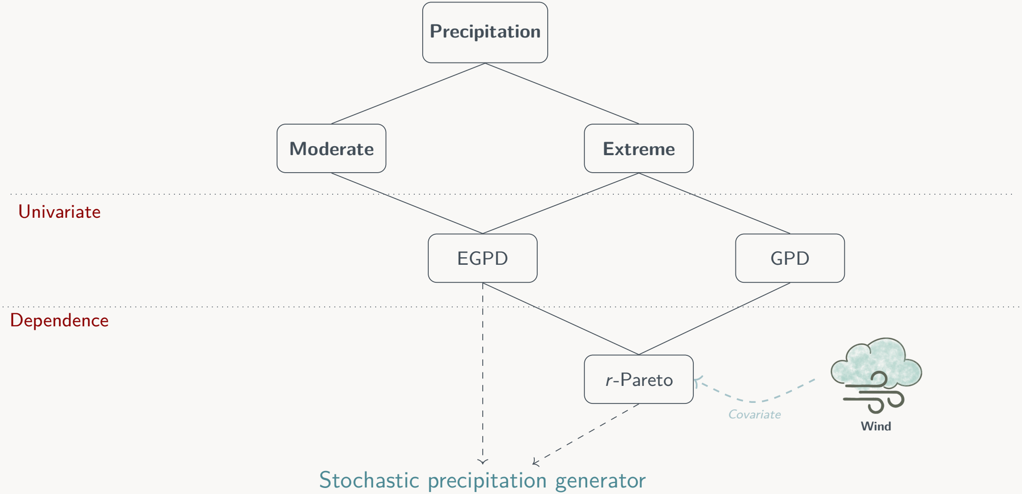


With advection

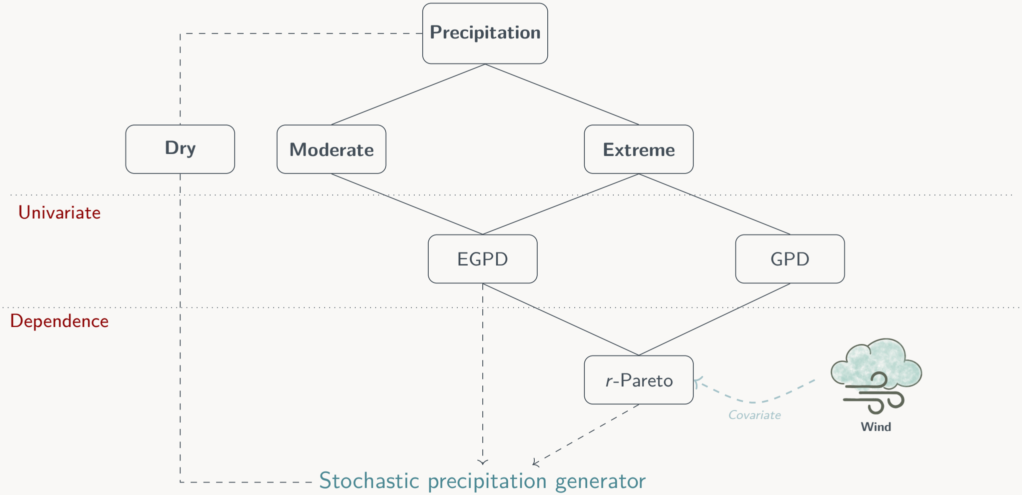
Parameter estimation on 100 simulations of 1000 replicates of r-Pareto processes with 25 sites and 30 time observations



Precipitation
Moderate
Extreme
Univariate
EGPD
GPD
Dependence
r-Pareto
Stochastic precipitation generator


Precipitation
Moderate
Extreme
Univariate
EGPD
GPD
Dependence
r-Pareto
Covariate
Wind
Stochastic precipitation generator

Conclusion


Precipitation
Dry
Moderate
Extreme
Univariate
EGPD
GPD
Dependence
r-Pareto
Covariate
Wind
Stochastic precipitation generator

Buhl, S., Davis, R. A., Klüppelberg, C., & Steinkohl, C. (2019).Semiparametric estimation for isotropic max-stable space-time processes.

Davis, R. A., & Mikosch, T. (2009).The extremogram: A correlogram for extreme events.

de Fondeville, R., & Davison, A. C. (2018).High-dimensional peaks-over-threshold inference. *Biometrika*, *105*(3), 575–592.

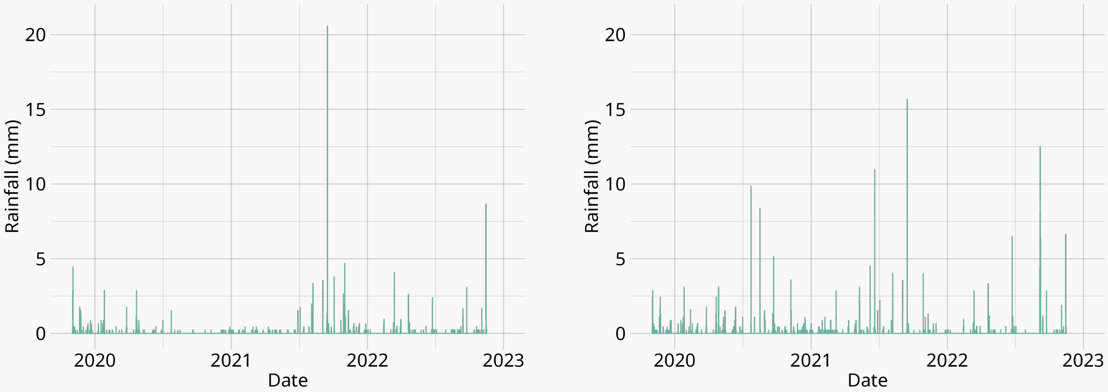
Finaud-Guyot, P., Guinot, V., Marchand, P., Neppel, L., Salles, C., & Toulemonde, G. (2023). Rainfall data collected by the HSM urban observatory (OMSEV).

Matheron, G. (1963).Principles of geostatistics. *Economic geology*, *58*(8), 1246–1266.

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016).Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*.

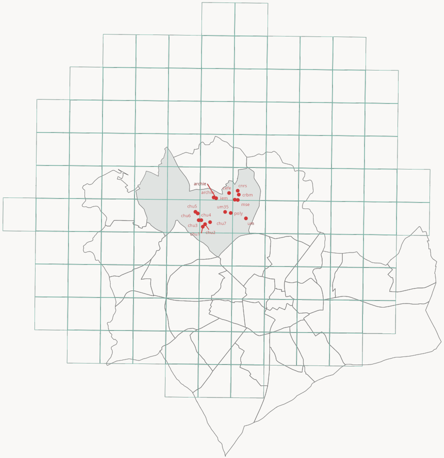
Pickands III, J. (1975).Statistical inference using extreme order statistics. *the Annals of Statistics*, 119–131.

Rainfall data - HSM



Rainfall amounts on CNRS and Polytech rain gauges

Pixels of 1km x 1km



- **Source:** COMEPHORE, Météo France
- **Time period:** $[1997, 2023[$
- **Temporal resolution:** Every hour
- **Spatial resolution:** 1 km^2

More consistent data: Both datasets + Neural Network Downscaling.

Generalized Pareto Distribution ⟶ **Extended GPD**[4]

$$\overline{H}_\xi\left(\frac{x-u}{\sigma}\right) = \begin{cases} \left(1+\xi\frac{x-u}{\sigma}\right)_+^{-1/\xi} & \text{if } \xi \neq 0\,, \\ e^{-\frac{x-u}{\sigma}} & \text{if } \xi = 0\,, \end{cases}$$

where $a_+ = \max(a, 0)$, $\sigma > 0$, $x - u > 0$

- Models extreme precipitation
- Depends on a threshold choice

$$F(x) = G\left(H_\xi\left(\frac{x}{\sigma}\right)\right),$$

where $G(x) = x^\kappa$, $\kappa > 0$

- Models **moderate and extreme** precipitation
- Avoids a threshold choice

[4] Naveau et al., 2016

Estimation of the separable variogram parameters

Case of additive separability: $\frac{\gamma(\boldsymbol{h},\tau)}{2} = \beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}, \quad 0 < \alpha_1, \alpha_2 \leq 2\,, \; \beta_1, \beta_2 > 0$

Spatio-temporal

$$\chi(\boldsymbol{h},\tau) = 2\left(1 - \phi\left(\sqrt{\tfrac{1}{2}\gamma(\boldsymbol{h},\tau)}\right)\right)$$

Transformation:

$$\eta(\chi) = 2\log\left(\phi^{-1}\left(1 - \tfrac{1}{2}\chi\right)\right)$$

Spatial

$$\eta\left(\chi(\boldsymbol{h},0)\right) = \log\beta_1 + \alpha_1 \log\|\boldsymbol{h}\|\,, \;\; \boldsymbol{h} \in \Lambda_{\mathcal{S}}$$

Temporal

$$\eta\left(\chi(0,\tau)\right) = \log\beta_2 + \alpha_2 \log\tau\,, \;\; \tau \in \Lambda_{\mathcal{T}}$$

Case of additive separability: $\frac{\gamma(\boldsymbol{h},\tau)}{2} = \beta_1 \|\boldsymbol{h}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}, \quad 0 < \alpha_1, \alpha_2 \leq 2\,, \ \beta_1, \beta_2 > 0$

Spatio-temporal

$$\chi(\boldsymbol{h}, \tau) = 2\left(1 - \phi\left(\sqrt{\tfrac{1}{2}\gamma(\boldsymbol{h}, \tau)}\right)\right)$$

Transformation:

$$\eta(\chi) = 2 \log\left(\phi^{-1}\left(1 - \tfrac{1}{2}\chi\right)\right)$$

Spatial

$$\eta\left(\chi(\boldsymbol{h}, 0)\right) := c_1 + \alpha_1 x_h\,, \ \ \boldsymbol{h} \in \Lambda_{\mathcal{S}}$$

Temporal

$$\eta\left(\chi(0, \tau)\right) := c_2 + \alpha_2 x_\tau\,, \ \ \tau \in \Lambda_{\mathcal{T}}$$

Weighted Least Squares Estimation (WLSE)

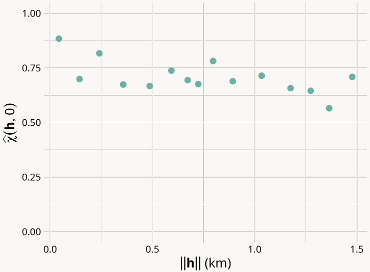
$$\begin{pmatrix} \widehat{c_i} \\ \widehat{\alpha_i} \end{pmatrix} = \text{argmin}_{c_i, \alpha_i} \sum_x w_x \left(\eta\left(\widehat{\chi}\right) - (c_i + \alpha_i x)\right)^2$$

Empirical spatial extremogram

For a fixed $t \in \mathcal{T}$ and q a high quantile,

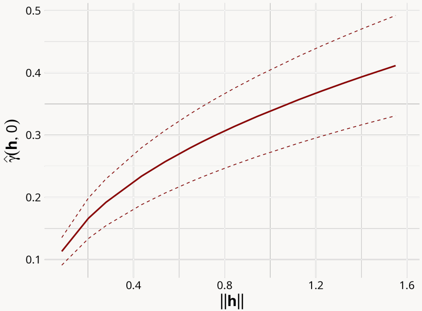
$$\widehat{\chi}_q^{(t)}(\boldsymbol{h}, 0) = \frac{\frac{1}{|N_{\boldsymbol{h}}|} \sum_{i,j \,|\, (\boldsymbol{s}_i, \boldsymbol{s}_j) \in N_{\boldsymbol{h}}} \mathbb{1}_{\{X^*_{\boldsymbol{s}_i,t} > q, X^*_{\boldsymbol{s}_j,t} > q\}}}{\frac{1}{|\mathcal{S}|} \sum_{i=1}^{|\mathcal{S}|} \mathbb{1}_{\{X^*_{\boldsymbol{s}_i,t} > q\}}},$$

where $C_{\boldsymbol{h}}$ are equifrequent distance classes and
$N_{\boldsymbol{h}} = \left\{ (\boldsymbol{s}_i, \boldsymbol{s}_j) \in \mathcal{S}^2 \;\middle|\; \|\boldsymbol{s}_i - \boldsymbol{s}_j\| \in C_{\boldsymbol{h}} \right\}$.



Transformation and WLSE

$\Longrightarrow$

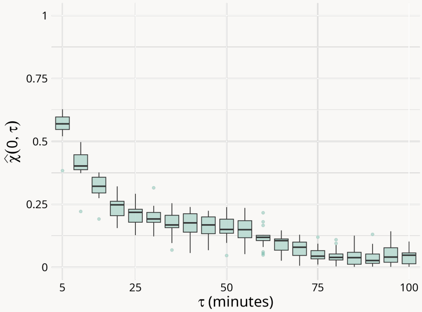


Spatial variogram $\widehat{\gamma}(\boldsymbol{h}, 0) = 2\widehat{\beta}_1 \|\boldsymbol{h}\|^{\widehat{\alpha}_1}$

Empirical temporal extremogram

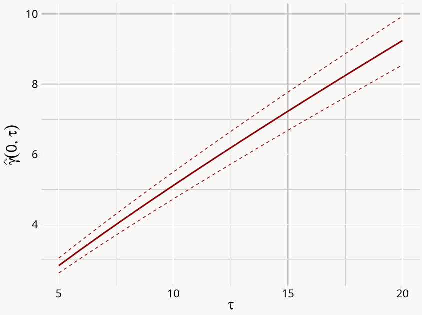
For a location $\boldsymbol{s} \in \mathcal{S}$, a high quantile q and $t_k \in \{t_1, \ldots, t_T\}$,

$$\widehat{\chi}_q^{(\boldsymbol{s})}(\mathbf{0}, \tau) = \frac{\frac{1}{T-\tau} \sum_{k=1}^{T-\tau} \mathbb{1}_{\{X^*_{\boldsymbol{s},t_k} > q, X^*_{\boldsymbol{s},t_k+\tau} > q\}}}{\frac{1}{T} \sum_{k=1}^{T} \mathbb{1}_{\{X^*_{\boldsymbol{s},t_k} > q\}}}$$



Transformation and WLSE

$\Longrightarrow$

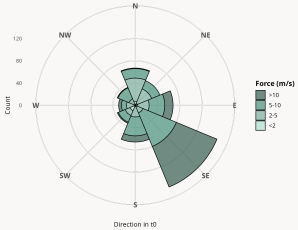


Temporal variogram $\widehat{\gamma}(\mathbf{0}, \tau) = 2\widehat{\beta}_2|\tau|^{\widehat{\alpha}_2}$

Composite log-likelihood maximization

$$\ell_C(\boldsymbol{\Theta}) \propto \sum_{(\boldsymbol{s}_0,t_0)\in\mathcal{R}} \sum_{(\boldsymbol{s},\boldsymbol{s}_0)\in\mathcal{N}(\boldsymbol{h})} \sum_{(t,t_0)\in\mathcal{N}(\tau)} k_{\boldsymbol{s},t} \log \chi_{r,\boldsymbol{\Theta}}(\boldsymbol{h},\tau) + (1-k_{\boldsymbol{s},t}) \log(1-\chi_{r,\boldsymbol{\Theta}}(\boldsymbol{h},\tau))$$

with $k_{\boldsymbol{s},t} = \mathbb{1}_{\{X_{\boldsymbol{s},t}>u,\, X_{\boldsymbol{s}_0,t_0}>u\}}$, for large u



Wind in t_0

Wind-informed advection:

$$\boldsymbol{V} = \eta_1 \cdot |\boldsymbol{V}_w|^{\eta_2} \cdot \mathrm{sign}(\boldsymbol{V}_w), \quad \eta_1, \eta_2 > 0$$

$$\Rightarrow \boldsymbol{\Theta} = (\beta_1, \beta_2, \alpha_1, \alpha_2, \eta_1, \eta_2)$$

with $\boldsymbol{V}_w$ real wind data.