

# SPATIO-TEMPORAL MODELING OF URBAN EXTREME RAINFALL EVENTS AT HIGH RESOLUTION

KIT Seminar — May 19, 2026

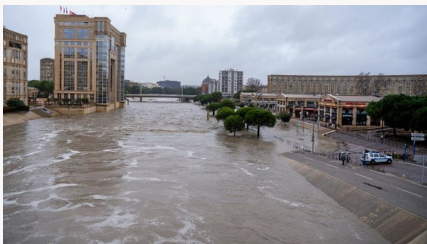
Chloé **SERRE-COMBE**<sup>1</sup>   Nicolas MEYER<sup>1</sup>   Thomas OPITZ<sup>2</sup>   Gwladys TOULEMONDE<sup>1</sup>

<sup>1</sup>IMAG, Université de Montpellier, LEMON Inria

<sup>2</sup>INRAE, BioSP, Avignon

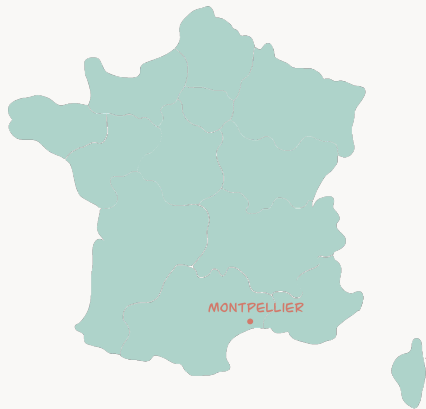


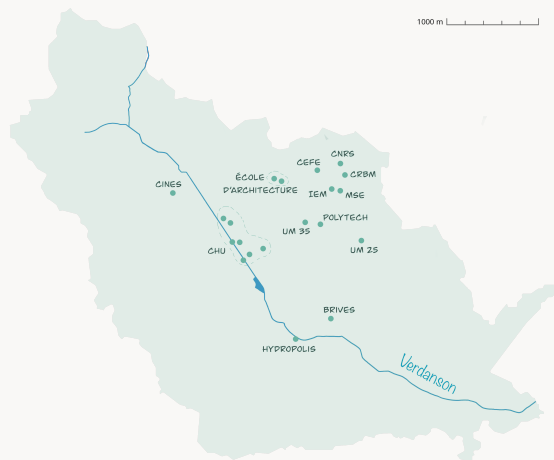
# GENERAL CONTEXT - MONTPELLIER, FRANCE



Floods in Montpellier, August 2015 and December 2025 (*Midi Libre*)

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



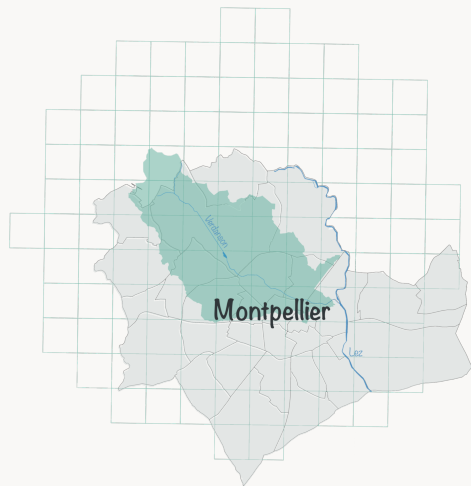


- ▶ **Study area:** Verdanson water catchment
- ▶ **Source:** Urban observatory of HydroScience Montpellier (HSM)<sup>2</sup>
- ▶ **Time period:** [Sept.2019, Jan.2025[
- ▶ **Temporal resolution:** 5 minutes
- ▶ **Spatial resolution:** 77 m to 2259 m

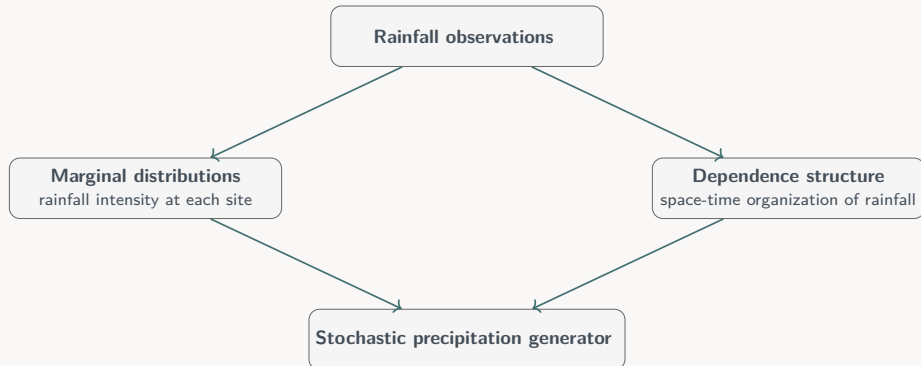
<sup>1</sup>Observ. Montpellierain et au Sud de l'Eau dans la Ville

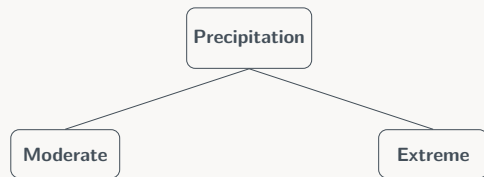
<sup>2</sup>FINAUD-GUYOT et al., 2023

$$\mathcal{S} = \{20 \text{ rain gauges}\} \subset \mathbb{R}^2 \text{ and } \mathcal{T} \subset \mathbb{R}_+$$



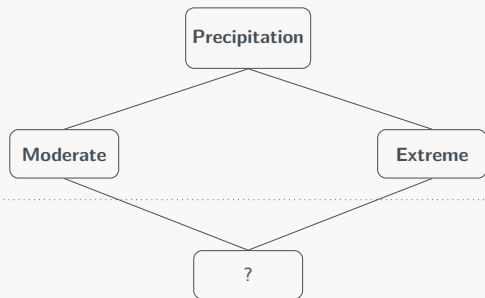
- ▶ **Source:** COMEPHORE, Météo France
- ▶ **Time period:** [1997, 2025]
- ▶ **Temporal resolution:** Every hour
- ▶ **Spatial resolution:** 1 km<sup>2</sup>



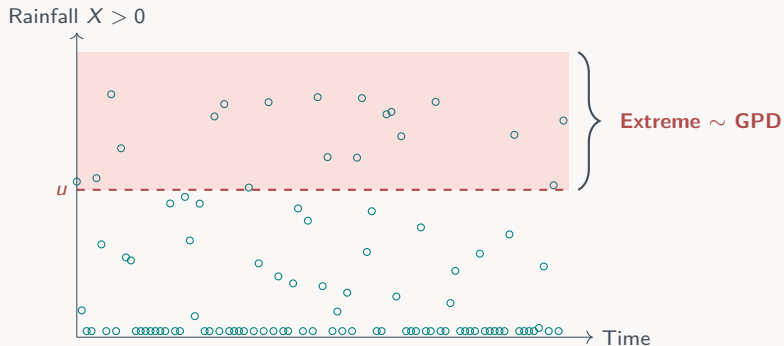


---

Univariate



Univariate

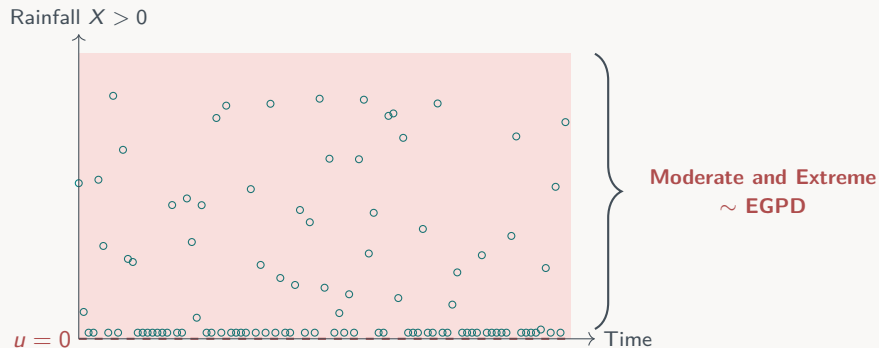


## ► Generalized Pareto Distribution (GPD)<sup>1</sup>

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

$\xi$ : shape (tail behavior),  $\sigma$ : scale,  $u$ : threshold

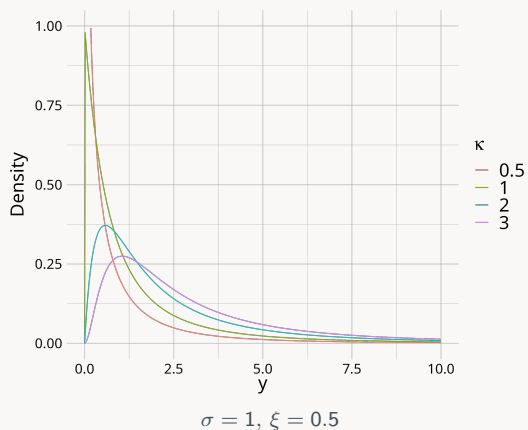
<sup>1</sup>PICKANDS III, 1975



► **Extended Generalized Pareto Distribution (EGPD)<sup>2</sup>**

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

<sup>2</sup>NAVEAU et al., 2016

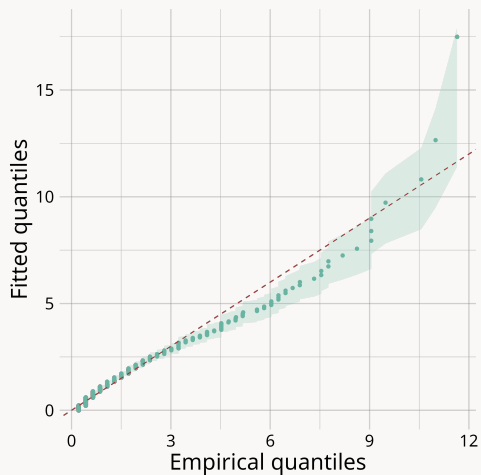
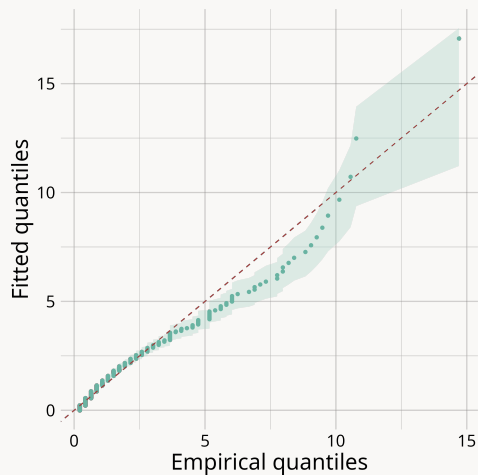


## Extended Generalized Pareto Distribution

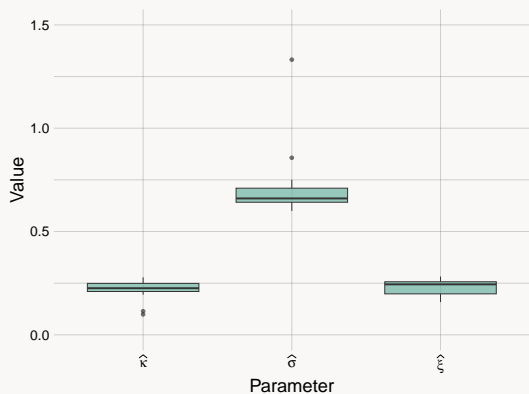
$$F(x) = G \left( H_{\xi} \left( \frac{x}{\sigma} \right) \right),$$

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

- ▶  $\kappa$ : controls the bulk of the distribution
- ▶ Models **moderate and extreme** precipitation
- ▶ Avoids a threshold choice

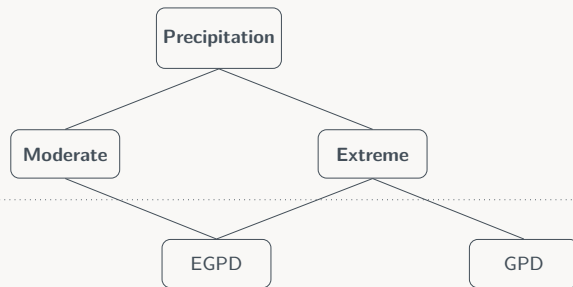


EGPD fitting for two rain gauges, CRBM (left) and UM (right) with left-censoring and 95% CI



Estimated parameters across rain gauges

- ▶  $\hat{\xi} \approx 0.25$ : extreme rainfall tail
- ▶  $\hat{\kappa} \approx 0.25$ : convective rainfall
- ▶  $\hat{\sigma} \approx 0.6$ : short-duration scale



Univariate

Dependence

**Rainfall random field:**  $\mathbf{X} = \{X_{s,t}, (s, t) \in \mathcal{S} \times \mathcal{T}\}$

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:**  $\mathbf{X} = \{X_{s,t}, (s, t) \in \mathcal{S} \times \mathcal{T}\}$

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)

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

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

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

**Rainfall random field:**  $\mathbf{X} = \{X_{s,t}, (s, t) \in \mathcal{S} \times \mathcal{T}\}$

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)

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

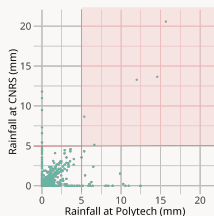
- Quantifies variability
- Higher  $\gamma(\mathbf{h}, \tau) \rightarrow$  weaker dependence

$\mathbf{h} \in \Lambda_{\mathcal{S}}, \tau \in \Lambda_{\mathcal{T}}, X_{s,t}^*$  uniform margins.

## Extremogram (DAVIS and MIKOSCH, 2009)

$$\chi(\mathbf{h}, \tau) = \lim_{q \rightarrow 1} \mathbb{P}(X_{s,t}^* > q \mid X_{s+\mathbf{h}, t+\tau}^* > q)$$

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

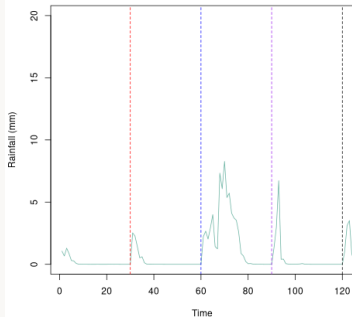


## Definition (DE FONDEVILLE and DAVISON, 2018)

For all  $s \in \mathcal{S}$  and  $t \in \mathcal{T}$ , a risk function  $r(\mathbf{X}) = X_{s_0, t_0}$ ,

$$u^{-1}X_{s,t} | X_{s_0, t_0} > u \xrightarrow{d} Y_{s,t} \quad \text{with} \quad Y_{s,t} = Re^{W_{s,t} - W_{s_0, t_0} - \gamma(s - s_0, t - t_0)},$$

where  $(s_0, t_0)$  is a space-time location,  $u$  is a high threshold,  $R \sim \text{Pareto}(1)$ ,  $W_{s,t}$  is a Gaussian process.



►  $r$ -extremogram

$$\chi_r(\mathbf{h}, \tau) = \lim_{q \rightarrow 1} \mathbb{P}(X_{s_0 + \mathbf{h}, t_0 + \tau}^* > q | X_{s_0, t_0}^* > q)$$

► Brown-Resnick dependence structure

## Spatio-temporal extremogram with a Brown-Resnick dependence

Let  $\mathbf{h} \in \Lambda_S$  and  $\tau \in \Lambda_T$ . We have

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

with  $\phi$  the std normal c.d.f. and  $\gamma$  the variogram of  $\mathbf{W}$ .

## Spatio-temporal extremogram with a Brown-Resnick dependence

Let  $\mathbf{h} \in \Lambda_S$  and  $\tau \in \Lambda_T$ . We have

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

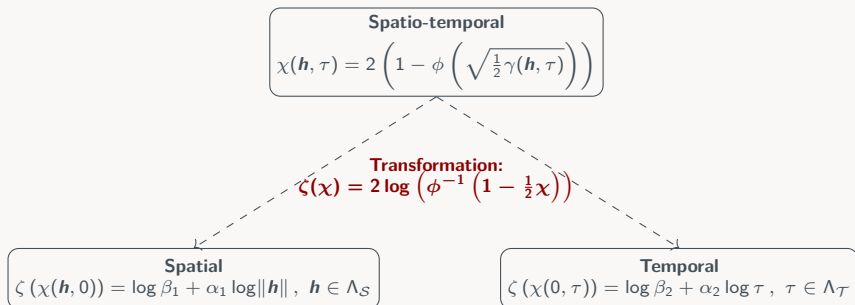
with  $\phi$  the std normal c.d.f. and  $\gamma$  the variogram of  $\mathbf{W}$ .

**Separable model:** Fractional Brownian motion with additive separability.

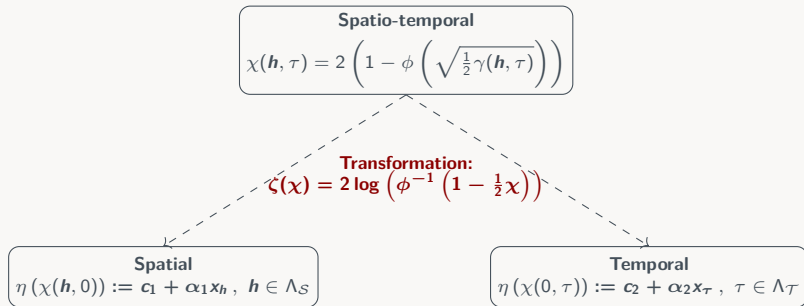
$$\frac{\gamma(\mathbf{h}, \tau)}{2} = \beta_1 \|\mathbf{h}\|^{\alpha_1} + \beta_2 |\tau|^{\alpha_2}$$

with  $0 < \alpha_1, \alpha_2 \leq 2$ ,  $\beta_1, \beta_2 > 0$  (BUHL et al., 2019).

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



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



**Weighted Least Squares Estimation (WLSE)**

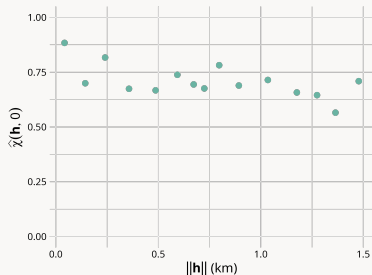
$$\begin{pmatrix} \hat{c}_i \\ \hat{\alpha}_j \end{pmatrix} = \operatorname{argmin}_{c_i, \alpha_j} \sum_x w_x (\zeta(\hat{x}) - (c_i + \alpha_j x))^2$$

## Empirical spatial extremogram

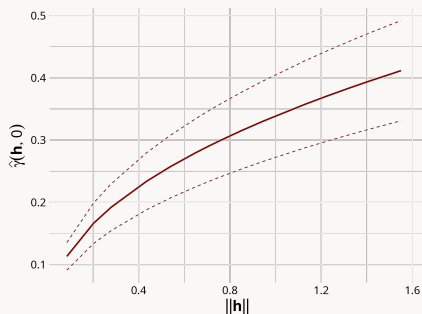
For a fixed  $t \in \mathcal{T}$  and  $q$  a high quantile,

$$\hat{\chi}_q^{(t)}(\mathbf{h}, 0) = \frac{\frac{1}{|N_{\mathbf{h}}|} \sum_{i,j | (s_i, s_j) \in N_{\mathbf{h}}} \mathbb{1}_{\{X_{s_i, t}^* > q, X_{s_j, t}^* > q\}}}{\frac{1}{|S|} \sum_{i=1}^{|S|} \mathbb{1}_{\{X_{s_i, t}^* > q\}}},$$

where  $C_{\mathbf{h}}$  are equifrequent distance classes and  $N_{\mathbf{h}} = \{(s_i, s_j) \in S^2 \mid \|s_i - s_j\| \in C_{\mathbf{h}}\}$ .



## Transformation and WLSE

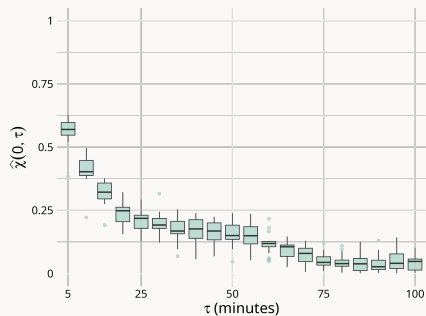


Spatial variogram  $\hat{\gamma}(\mathbf{h}, 0) = 2\hat{\beta}_1 \|\mathbf{h}\|^{\hat{\alpha}_1}$

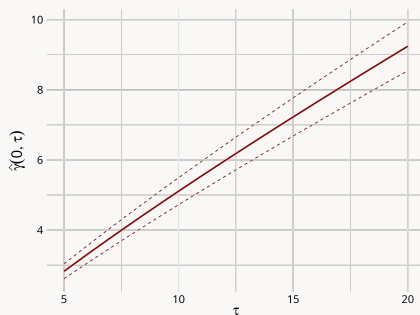
## Empirical temporal extremogram

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

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

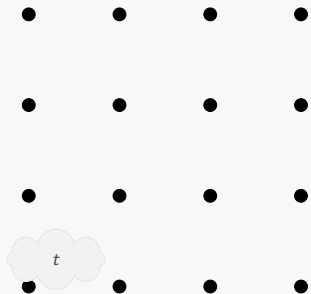


## Transformation and WLSE

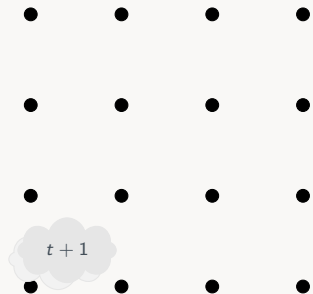


Temporal variogram  $\hat{\gamma}(\mathbf{0}, \tau) = 2\hat{\beta}_2|\tau|^{\hat{\alpha}_2}$

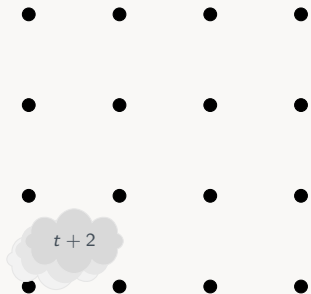
## Separable model



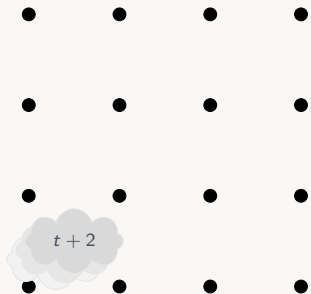
## Separable model



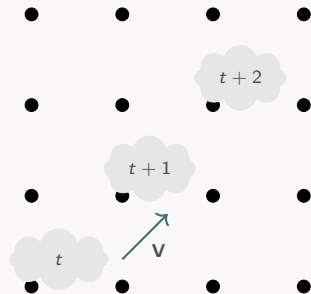
## Separable model



Separable model



Non-separable model



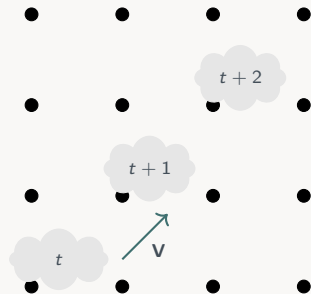
**Towards more realistic modeling:** introduce advection  $\mathbf{V}$  to relax separability

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

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

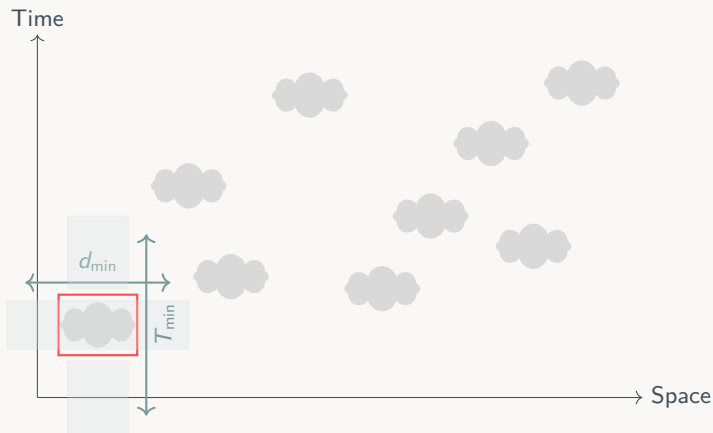
► Parameters:  $\Theta = (\beta_1, \beta_2, \alpha_1, \alpha_2, \mathbf{V})$

**Non-separable model**





Extreme episodes:  
Each episode is characterized  
by  $(s_0, t_0)$  for which  $X_{s_0, t_0} > u$

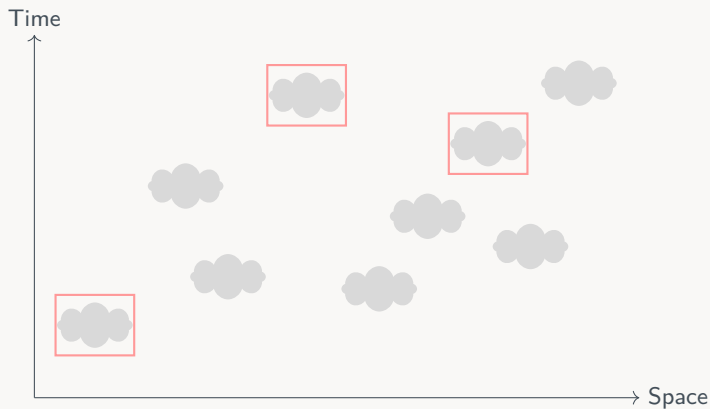
**Episode selection:**

Only episodes separated by

- spatial distance  $\geq d_{\min}$
- OR

- temporal gap  $\geq T_{\min}$

$\Rightarrow$  reduces dependence  
between **selected episodes**  $\in \mathcal{E}$ .

**Episode selection:**

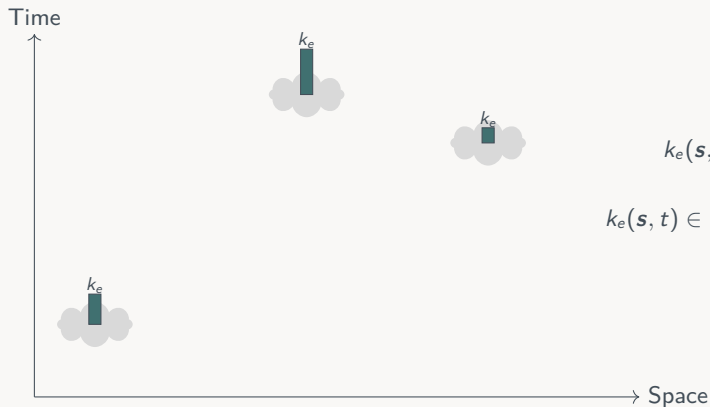
Only episodes separated by

- spatial distance  $\geq d_{\min}$

OR

- temporal gap  $\geq T_{\min}$

$\Rightarrow$  reduces dependence  
between **selected episodes**  $\in \mathcal{E}$ .



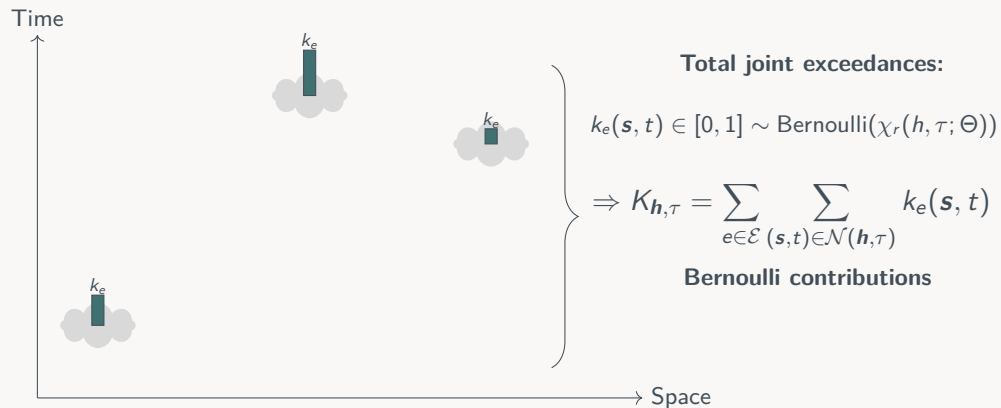
## Joint exceedances

$$k_e(\mathbf{s}, t) = \mathbb{1}_{\{X_{s_0, t_0} > u, X_{\mathbf{s}, t} > u\}},$$

with  $(\mathbf{s}, t) \in \mathcal{N}(\mathbf{h}, \tau)$

$$k_e(\mathbf{s}, t) \in [0, 1] \sim \text{Bernoulli}(\chi_r(\mathbf{h}, \tau; \Theta))$$

$$\text{with } \mathcal{N}(\mathbf{h}, \tau) = \{(\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T} \mid \|\mathbf{s} - \mathbf{s}_0\| = \|\mathbf{h}\|, |t - t_0| = \tau\}$$



$$\text{with } \mathcal{N}(\mathbf{h}, \tau) = \{(\mathbf{s}, t) \in \mathcal{S} \times \mathcal{T} \mid \|\mathbf{s} - \mathbf{s}_0\| = \|\mathbf{h}\|, |t - t_0| = \tau\}$$

**Bernoulli contributions (large  $u$ ):**

$$k_e(\mathbf{s}, t) = \mathbf{1}_{\{X_{s_0, t_0} > u, X_{\mathbf{s}, t} > u\}}, \quad (\mathbf{s}, t) \in \mathcal{N}(\mathbf{h}, \tau),$$

each treated as

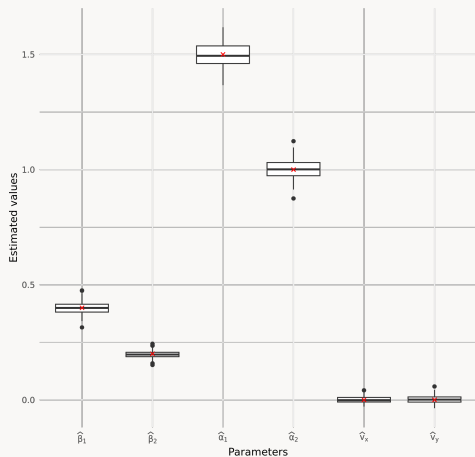
$$k_e(\mathbf{s}, t) \sim \text{Bernoulli}(\chi_r(\mathbf{h}, \tau; \Theta)).$$

### Composite log-likelihood

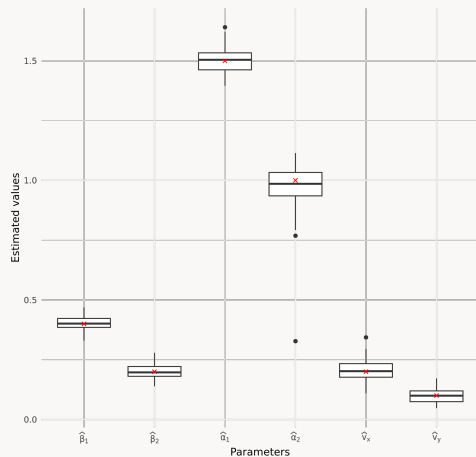
$$l_C(\Theta) \propto \sum_{e \in \mathcal{E}} \sum_{(\mathbf{h}, \tau) \in \Lambda_S \times \Lambda_T} \sum_{(\mathbf{s}, t) \in \mathcal{N}(\mathbf{h}, \tau)} k_e(\mathbf{s}, t) \log \chi_r(\mathbf{h}, \tau; \Theta) + (1 - k_e(\mathbf{s}, t)) \log(1 - \chi_r(\mathbf{h}, \tau; \Theta)).$$

- ▶ Optimization via maximization of  $l_C(\Theta)$
- ▶ Initial parameters: from WLSE of the separable model (i.e.,  $\mathbf{V} = \mathbf{0}$ )

# VALIDATION: CONSTANT ADVECTION



Without advection



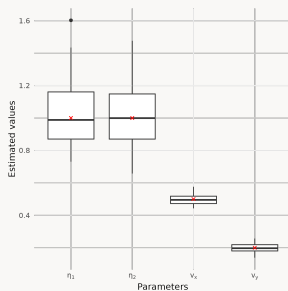
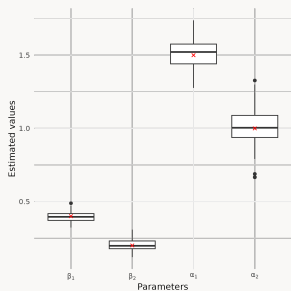
With advection

50 simulations, 500 replicates, 25 sites, 30 time steps

# EPISODE-WISE ADVECTION TO GLOBAL MODEL

$$\mathbf{v}^{\text{emp}} = \begin{bmatrix} \mathbf{v}^{(1)} \\ \mathbf{v}^{(2)} \\ \vdots \\ \mathbf{v}^{(N)} \end{bmatrix} \in \mathbb{R}^{N \times 2} \xrightarrow{\text{global model}} \mathbf{v}^{\text{final}} = \underbrace{\eta_1 \|\mathbf{v}^{\text{emp}}\|^{\eta_2} \frac{\mathbf{v}^{\text{emp}}}{\|\mathbf{v}^{\text{emp}}\|}}_{\text{transformation} \Rightarrow \eta \text{ to estimate}}, \eta_1 > 0, \eta_2 > 0$$

one component = one episode



50 simulations, 500 replicates, 25 sites, 30 time steps

$$\mathbf{v}^{\text{emp}} = \begin{bmatrix} \mathbf{v}^{(1)} \\ \mathbf{v}^{(2)} \\ \vdots \\ \mathbf{v}^{(N)} \end{bmatrix} \in \mathbb{R}^{N \times 2} \xrightarrow{\text{global model}} \mathbf{v}^{\text{final}} = \underbrace{\eta_1 \|\mathbf{v}^{\text{emp}}\|^{\eta_2} \frac{\mathbf{v}^{\text{emp}}}{\|\mathbf{v}^{\text{emp}}\|}}_{\text{transformation} \Rightarrow \eta \text{ to estimate}}, \eta_1 > 0, \eta_2 > 0$$

one component = one episode

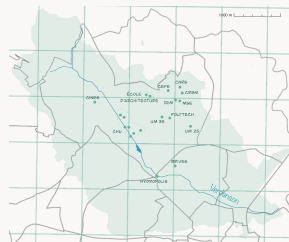
**Parameters:**  $\Theta = (\beta_1, \beta_2, \alpha_1, \alpha_2, \eta_1, \eta_2)$

**Challenge:**

OMSEV data  $\Rightarrow$  limited information for advection

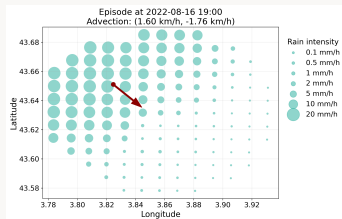
**Approach:**

Fusion COMEPHORE–OMSEV  $\rightarrow$  more reliable advection

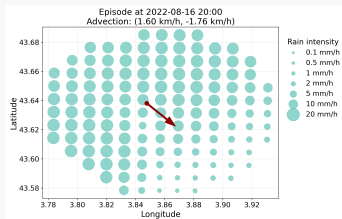


COMEPHORE pixels  
Météo France

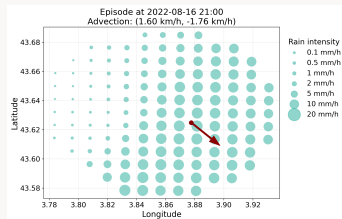
For one episode, we have:



$t_0 - 1$



$t_0$

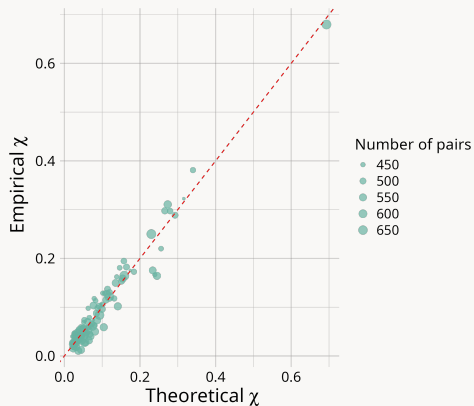


$t_0 + 1$

$V^{\text{emp}}$  is estimated from **rain storm barycenter displacement** within a time window.

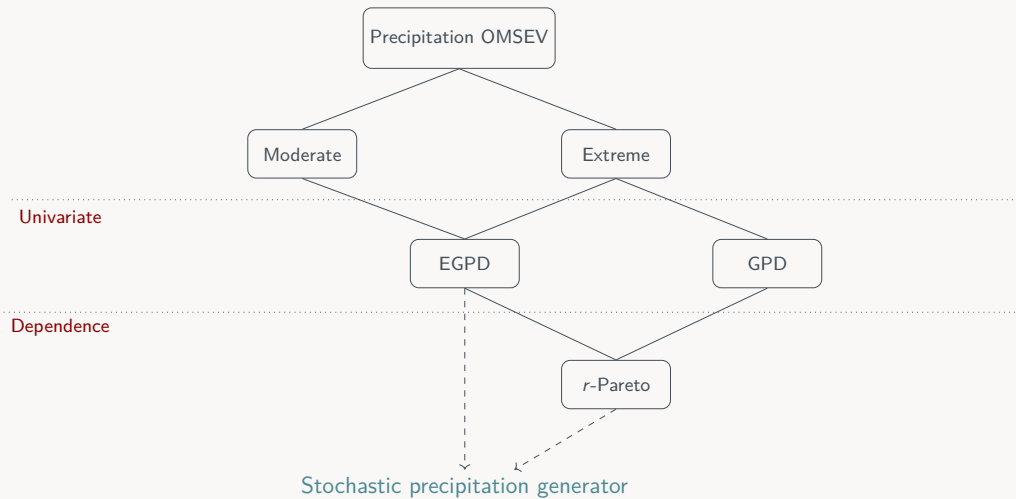
Optimization on COMEPHORE episodes  $\Rightarrow \hat{\eta}_1 = 3.896$  and  $\hat{\eta}_2 = 2.221$ .

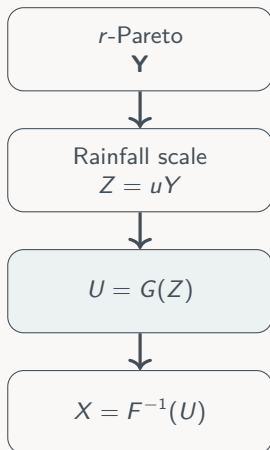
Parameter estimates over  $|\mathcal{E}| = 333$  episodes  
 $\hat{\eta}_1 = 3.896$  and  $\hat{\eta}_2 = 2.221$  fixed (COMEPHORE)



	m/5 min	
$\hat{\beta}_1$	0.231	[0.047, 1.060]
$\hat{\beta}_2$	0.807	[0.368, 1.655]
$\hat{\alpha}_1$	0.250	[0.064, 0.436]
$\hat{\alpha}_2$	0.666	[0.469, 0.863]

# STOCHASTIC SIMULATION OF PRECIPITATION EVENTS





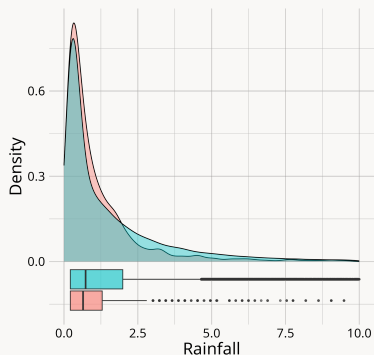
**Marginals: zero and EGPD**

$$F(x) = \begin{cases} 0, & x < 0, \\ \rho_0, & x = 0, \\ \rho_0 + (1 - \rho_0)F_{\text{EGPD}}(x), & x > 0. \end{cases}$$

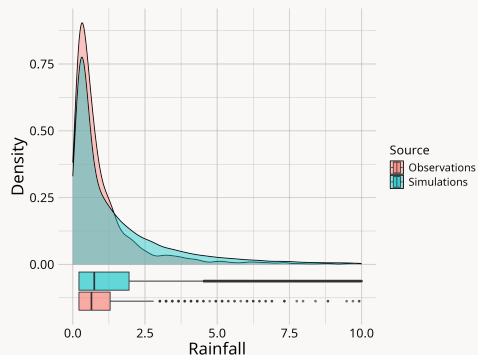
**Role of  $G$ :**

- ▶ maps to unit interval
- ▶ introduces mass at zero
- ▶ preserves Pareto tail

100 simulations by episode



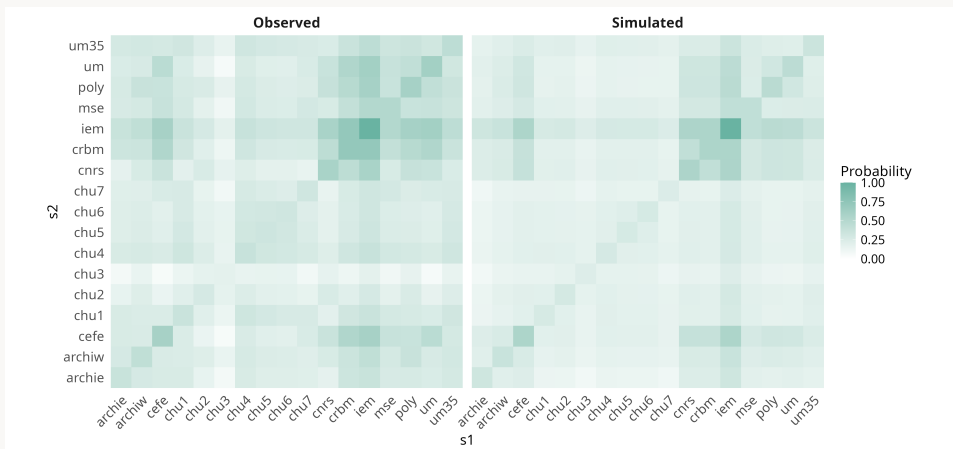
UM density



IEM density

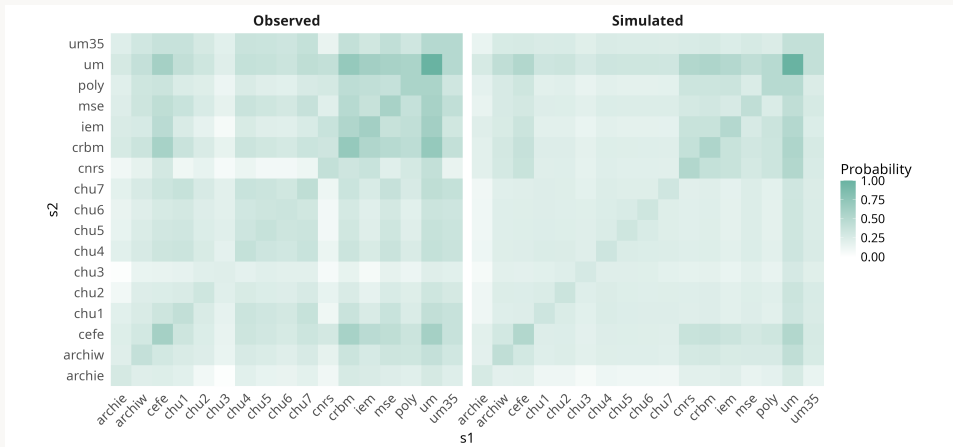
Comparison of marginal distributions (above-zero rainfall)

100 simulations by episode



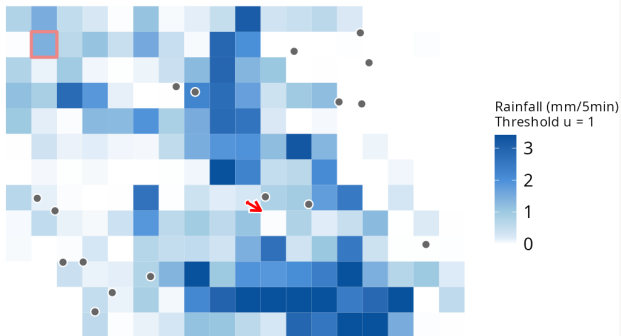
$$\mathbb{P}(X_{s_1} > u, X_{s_2} > u | X_{IEM} > u)$$

100 simulations by episode



$$\mathbb{P}(X_{s_1} > u, X_{s_2} > u | X_{UM} > u)$$

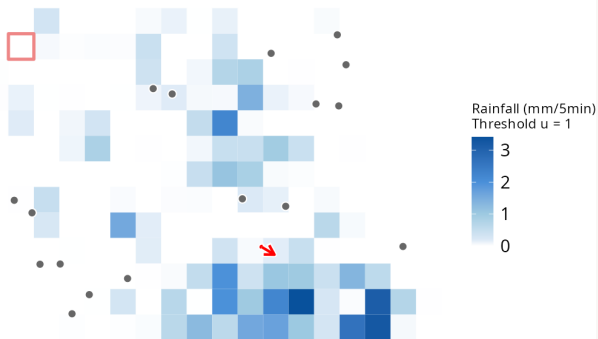
t = 1 | advection per 5 min = (50, -30) m



Over the OMSEV network area

- ▶ Spatial resolution: 100 m  $\times$  100 m
- ▶ Temporal resolution: 5 min
- ▶ Stochastic rainfall simulation

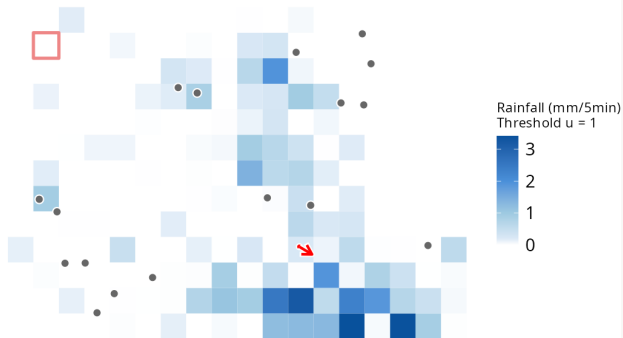
t = 2 | advection per 5 min = (50, -30) m



Over the OMSEV network area

- ▶ Spatial resolution: 100 m  $\times$  100 m
- ▶ Temporal resolution: 5 min
- ▶ Stochastic rainfall simulation

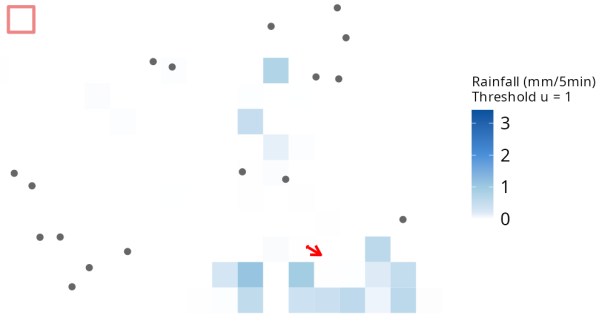
t = 3 | advection per 5 min = (50, -30) m



Over the OMSEV network area

- ▶ Spatial resolution: 100 m  $\times$  100 m
- ▶ Temporal resolution: 5 min
- ▶ Stochastic rainfall simulation

t = 4 | advection per 5 min = (50, -30) m

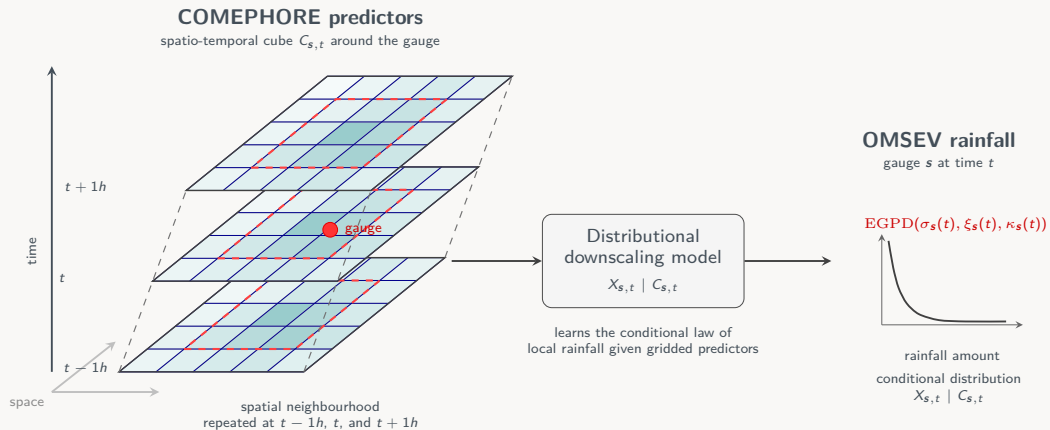


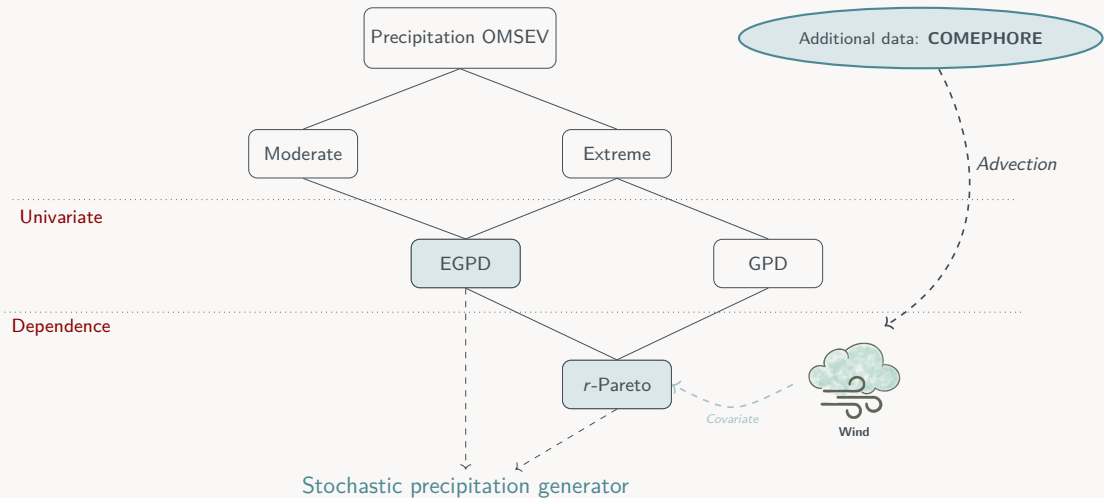
Rainfall (mm/5min)  
Threshold  $u = 1$

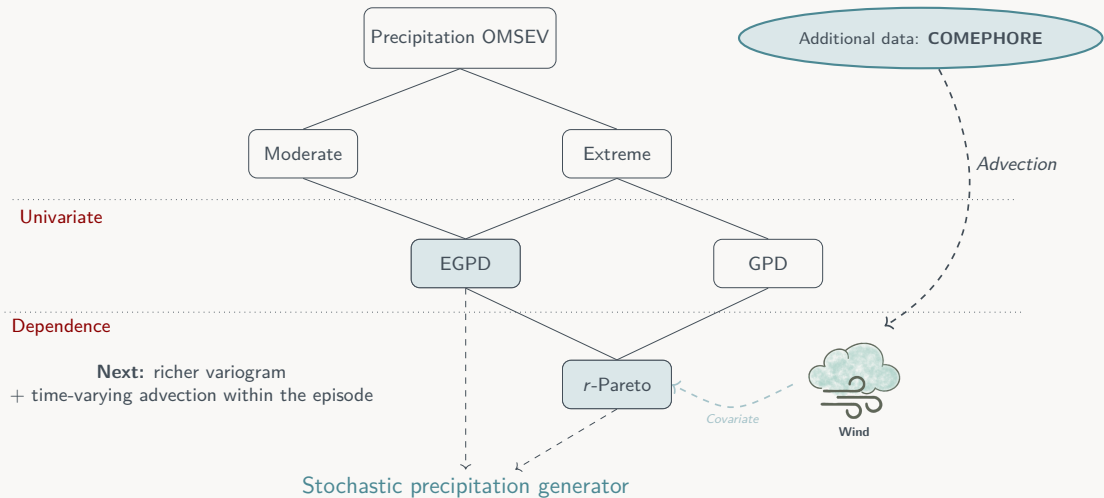
3  
2  
1  
0

## Over the OMSEV network area

- ▶ Spatial resolution: 100 m  $\times$  100 m
- ▶ Temporal resolution: 5 min
- ▶ Stochastic rainfall simulation

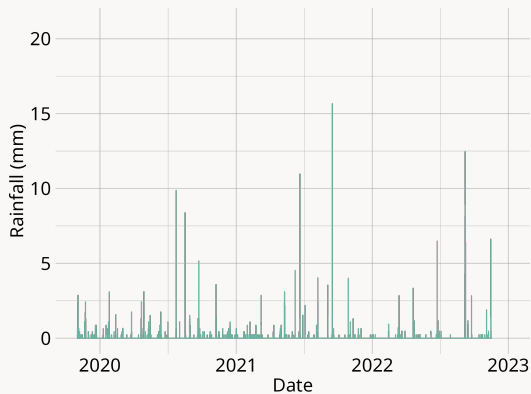
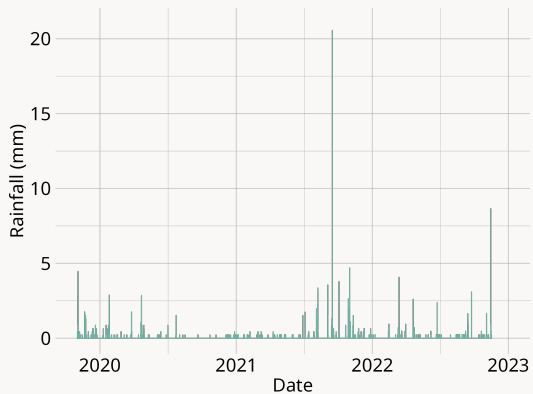






- BUHL, S., DAVIS, R. A., KLÜPPELBERG, C., & STEINKOHL, C. (2019). Semiparametric estimation for isotropic max-stable space-time processes. *Bernoulli*, 25, 2508–2537.
- DAVIS, R. A., & MIKOSCH, T. (2009). The extremogram: A correlogram for extreme events.
- DE FONDEVILLE, R., & DAVISON, A. C. (2018). 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 - OMSEV



Rainfall amounts on CNRS and Polytech rain gauges

## Generalized Pareto Distribution



## Extended GPD<sup>1</sup>

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

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

- ▶ Models extreme precipitation
- ▶ Depends on a threshold choice

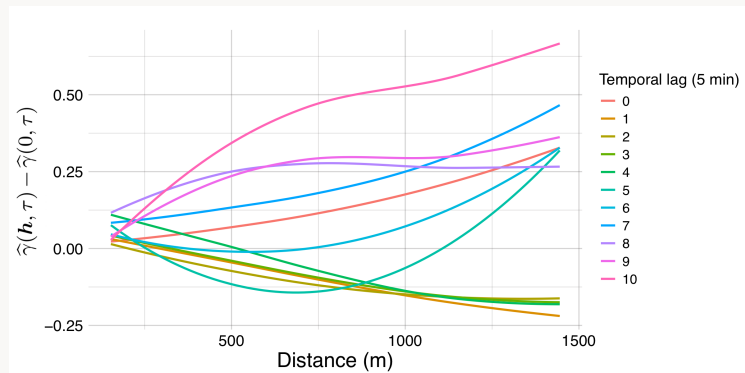
$$F(x) = G \left( H_\xi \left( \frac{x}{\sigma} \right) \right),$$

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

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

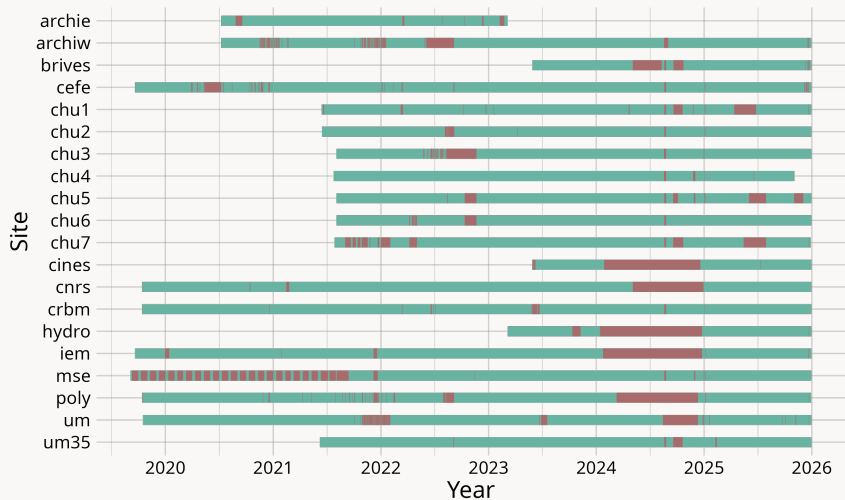
---

<sup>1</sup>NAVEAU et al., 2016



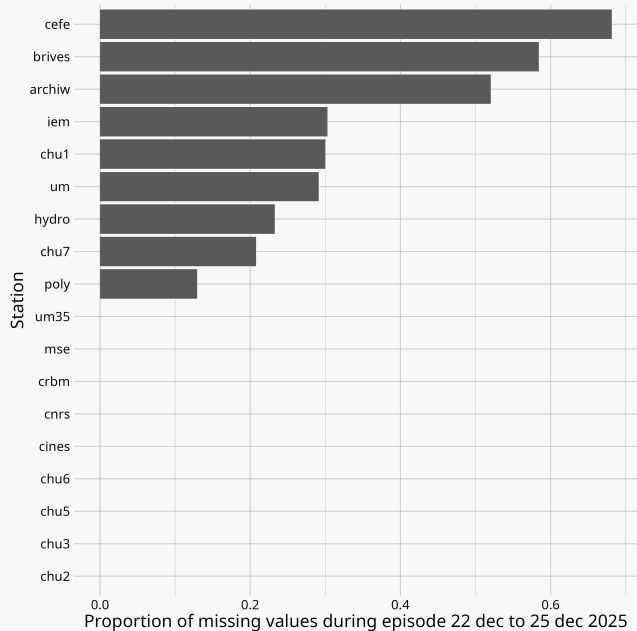
Empirical difference between the spatio-temporal variogram  $\gamma(\mathbf{h}, \tau)$  and the only temporal variogram on OMSEV data

# DATA ISSUE: MISSING VALUES



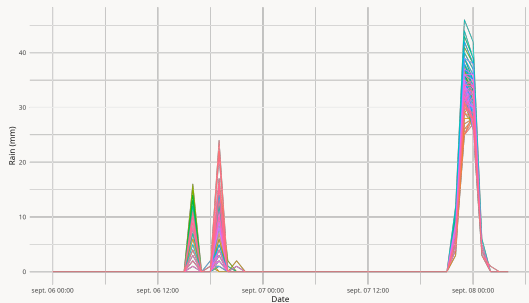
Missing values according to site activation periods (OMSEV dataset)

# DATA ISSUE: MISSING VALUES

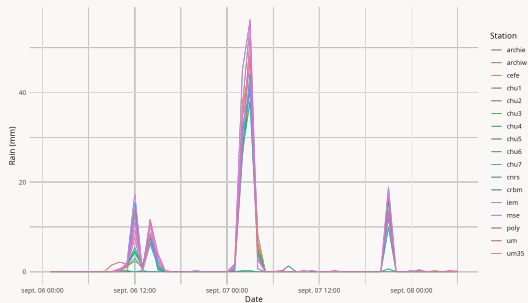


# DATA COMPARISON: RADAR VS RAIN GAUGES

## COMEPHORE (radar)



## OMSEV (gauges)



Mismatch between radar and gauge data