

STOCHASTIC GENERATOR OF EXTREME PRECIPITATION AT HIGH SPATIO-TEMPORAL RESOLUTION

Chloé SERRE-COMBE — PhD defense, September 16, 2026

IMAG, Université de Montpellier, LEMON Inria

Supervisors

Nicolas MEYER
Thomas OPITZ
Gwladys TOULEMONDE

PhD committee

Juliette BLANCHET
Denis ALLARD
Laure RAYNAUD
Thomas OPITZ

Anna KIRILIOUK
Carlo GAETAN
Nicolas MEYER
Gwladys TOULEMONDE



IMAG
INSTITUT MONTPELLIERAIN
ALEXANDER GRÖTHENDIECK



GENERAL CONTEXT - MONTPELLIER, FRANCE



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

- ▶ Mediterranean events: intense and localized
- ▶ Urbanization increases flood risk



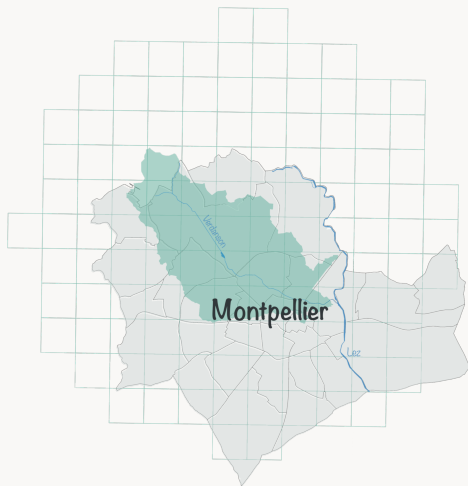


- **Study area:** Verdanson water catchment
- **Source:** OMSEV¹ (FINAUD-GUYOT et al., 2023)
- **Time period:** [Sept.2019, Dec.2025]
- **Temporal resolution:** 5 minutes
- **Spatial resolution:** 77 m to 2259 m

¹Observ. Montpellierain et au Sud de l'Eau dans la Ville

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

ADDITIONAL DATA AT A COARSER RESOLUTION - COMEPHORE



- **Source:** COMEPHORE, Météo France¹
- **Time period:** [2008, 2025]
- **Temporal resolution:** Every hour
- **Spatial resolution:** 1 km²

¹TABARY et al., 2012

How can rainfall be reproduced at high spatio-temporal resolution for hydrological modeling?

Question 1

How to generate extreme rainfall events
while preserving their spatio-temporal structure?



AXIS 1 – SPG

Stochastic generator of extreme rainfall
with space-time dependence and advection

Question 2

How to recover local rainfall distributions
from coarser spatio-temporal information?



AXIS 2 – DOWNSCALING

Conditional probabilistic model
for local rainfall distributions

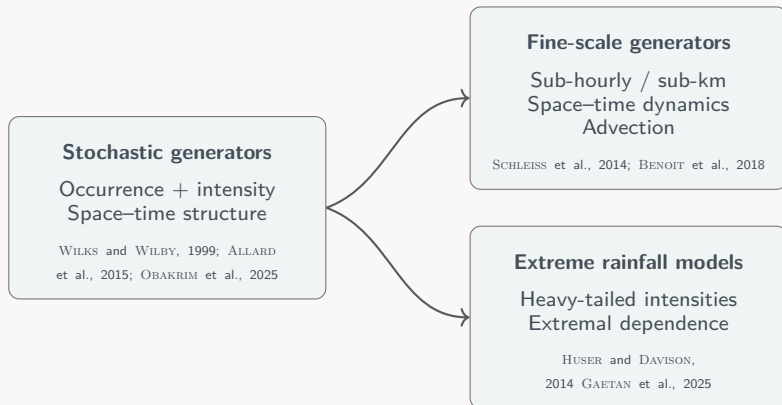
Transverse contribution – High-resolution rainfall data

OMSEV network of 20 rain gauges, 5-min resolution, 2019–2025

AXIS 1

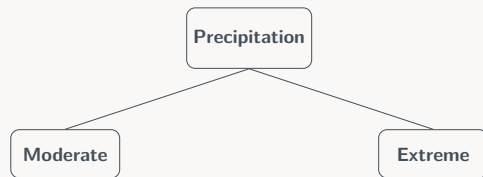
Stochastic precipitation generator

Modeling and generating extreme rainfall episodes
in space and time

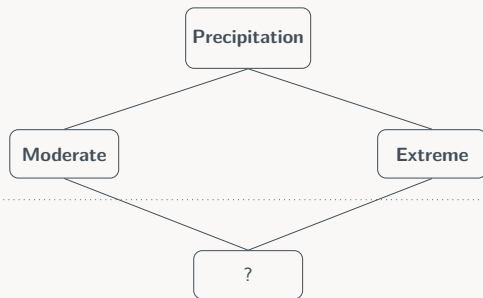


This thesis

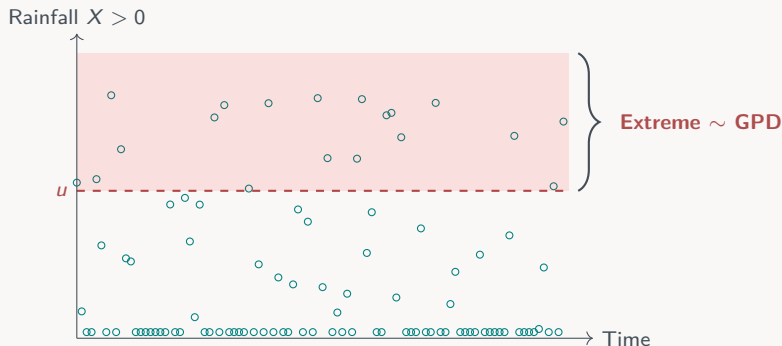
Extreme rainfall episodes at very high spatio-temporal resolution
Heavy-tailed margins + space-time extremal dependence + advection



Univariate



Univariate

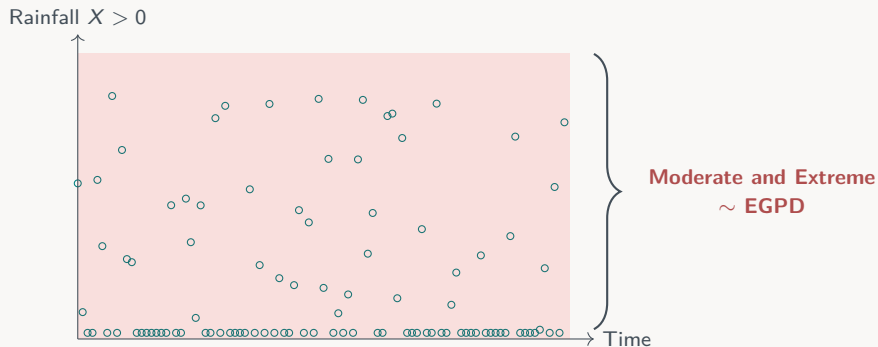


► Generalized Pareto Distribution (GPD)¹

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

ξ : shape (tail behavior), σ : scale, u : threshold

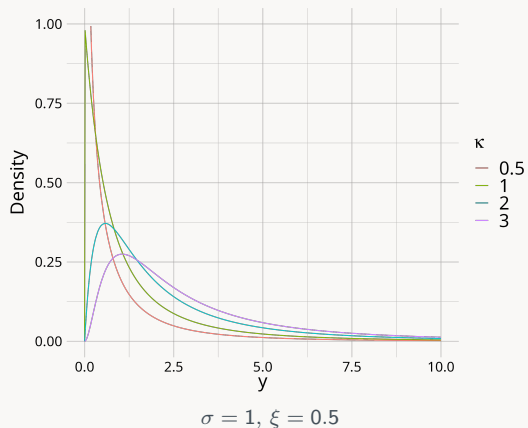
¹PICKANDS III, 1975



► Extended Generalized Pareto Distribution (EGPD)²

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

²NAVEAU et al., 2016

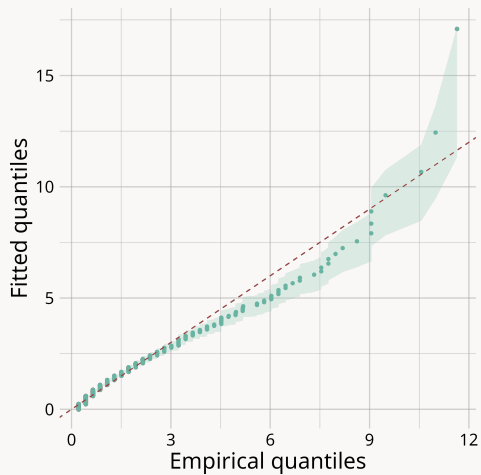
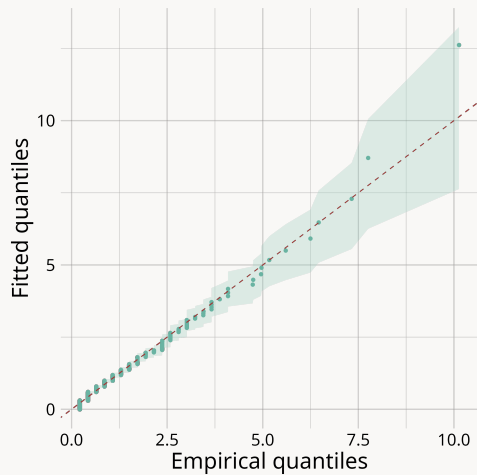


Extended Generalized Pareto Distribution

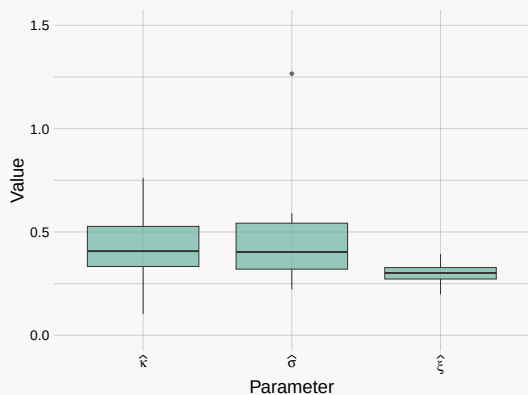
$$F_{\text{EGPD}}(x; \xi, \sigma, \kappa) = P\left(H_{\xi}\left(\frac{x}{\sigma}\right)\right),$$

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

- κ : controls the bulk of the distribution
- Models **moderate and extreme** precipitation
- Avoids a threshold choice



EGPD fitting for two rain gauges, Brives (left) and UM (right) with **left-censoring**

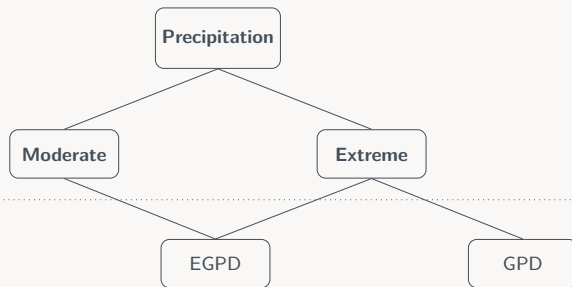


Full marginal distribution

$$F(x) = \begin{cases} 0, & x < 0, \\ p_0, & x = 0, \\ p_0 + (1 - p_0)F_{\text{EGPD}}(x; \xi, \sigma, \kappa), & x > 0. \end{cases}$$

Common marginal parameters

Parameter	$\hat{p}_0$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\kappa}$
Estimate	0.989	0.244	0.536	0.308



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)

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

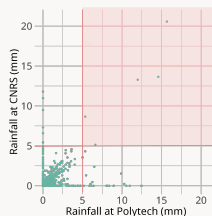
- Quantifies variability
- Higher $\gamma(\mathbf{h}, \tau) \rightarrow$ weaker dependence

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

Extremogram (DAVIS and MIKOSCH, 2009)

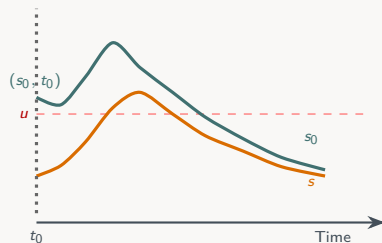
$$\chi(\mathbf{h}, \tau) = \lim_{q \rightarrow 1} \mathbb{P}(X_{s,t}^* > q \mid X_{s+\mathbf{h}, t+\tau}^* > q)$$

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



r -Pareto process with risk function $r(\mathbf{X}) = X_{s_0, t_0}$

One selected episode



Condition on $X_{s_0, t_0} > u$

r -Pareto process (DE FONDEVILLE and DAVISON, 2018)

Asymptotic model : $u^{-1}X_{s,t} \mid X_{s_0, t_0} > u \xrightarrow{d} Y_{s,t}$

$$Y_{s,t} = \underbrace{R}_{\text{Pareto(1) magnitude}} \exp \left[\underbrace{W_{s,t} - W_{s_0, t_0}}_{\text{Gaussian space-time field}} - \underbrace{\gamma(\mathbf{s} - \mathbf{s}_0, t - t_0)}_{\text{dependence}} \right].$$

W is a Gaussian process and γ is its space-time variogram.

r -extremogram: pairwise extremal dependence

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

How do we model the space-time dependence $\gamma(\mathbf{h}, \tau)$?

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

How do we model the space-time dependence $\gamma(\mathbf{h}, \tau)$?

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 ϕ the std normal c.d.f. and γ 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).

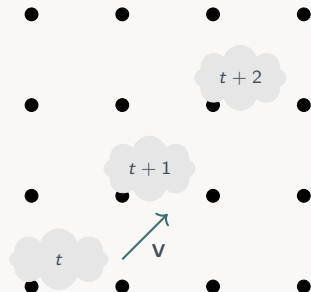
Towards more realistic modeling: introduce advection $\mathbf{V}$ to relax separability

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

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

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

Non-separable model



$\mathcal{E}$: set of selected extreme rainfall episodes

Joint exceedance

for episode $e \in \mathcal{E}$

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

$$k_e(\mathbf{s}, t) \in \{0, 1\}$$

$\mathcal{E}$: set of selected extreme rainfall episodes

Joint exceedance

for episode $e \in \mathcal{E}$

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

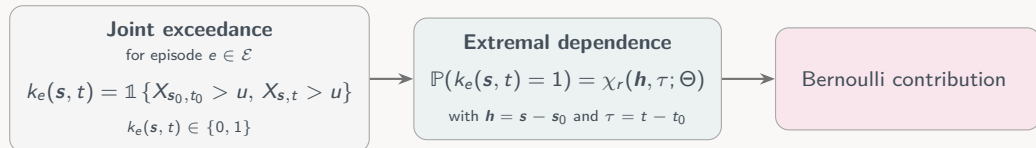
$$k_e(\mathbf{s}, t) \in \{0, 1\}$$

Extremal dependence

$$\mathbb{P}(k_e(\mathbf{s}, t) = 1) = \chi_r(\mathbf{h}, \tau; \Theta)$$

with $\mathbf{h} = \mathbf{s} - \mathbf{s}_0$ and $\tau = t - t_0$

$\mathcal{E}$: set of selected extreme rainfall episodes



Composite log-likelihood

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

$$\boxed{\hat{\Theta} = \arg \max_{\Theta} \ell_C(\Theta)} \quad \text{with} \quad \Theta = (\beta_1, \beta_2, \alpha_1, \alpha_2, \mathbf{V})$$

Rainfall episodes have different propagation patterns



$$\mathcal{A}(\mathbf{v}^{(i)}) = \eta_1 \left\| \mathbf{v}^{(i)} \right\|^{\eta_2} \frac{\mathbf{v}^{(i)}}{\left\| \mathbf{v}^{(i)} \right\|}$$

Episode-specific advection information + 2 global parameters η_1, η_2

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

OMSEV limitation

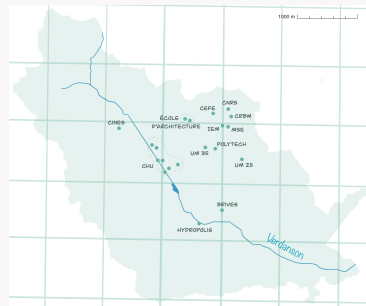
Local rain-gauge network
⇒ limited information on rainfall displacement

COMEPHORE contribution

Spatial rainfall fields over a larger area
⇒ more reliable **episode-wise advection**

Use in the model

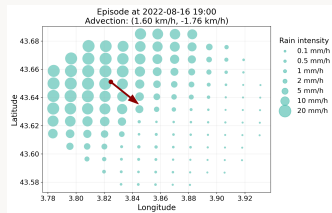
COMEPHORE ⇒ estimate η_1, η_2
 η_1, η_2 **fixed** for OMSEV inference



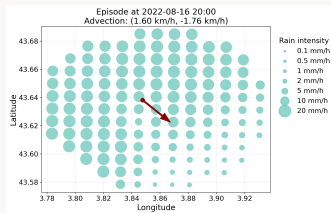
COMEPHORE pixels around
the OMSEV network

ADVECTION ESTIMATION FROM COMEPHORE

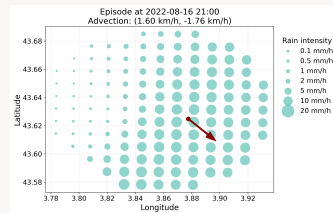
Episode-wise empirical advection



$t_0 - 1$



t_0



$t_0 + 1$

Rainfall barycenter displacement

$\Rightarrow$

$\mathbf{V}^{\text{emp}}$

Advection contribution:

$$\mathcal{A}(\mathbf{V}^{\text{emp}}) = \eta_1 \|\mathbf{V}^{\text{emp}}\|^{\eta_2} \frac{\mathbf{V}^{\text{emp}}}{\|\mathbf{V}^{\text{emp}}\|}$$

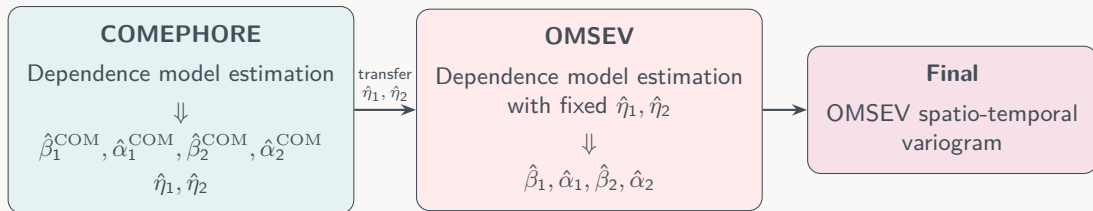
Optimization over COMEPHORE episodes

$\Rightarrow$

$\hat{\eta}_1, \hat{\eta}_2$

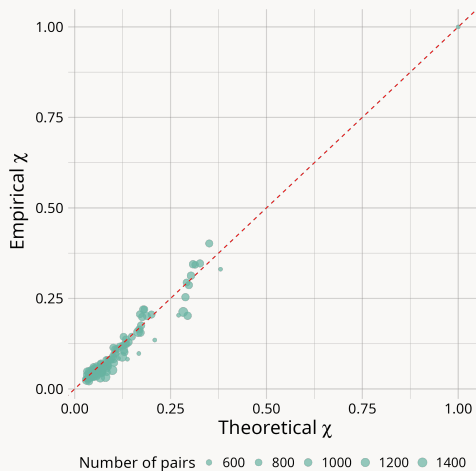
$\hat{\eta}_1 = 3.896$

$\hat{\eta}_2 = 2.221$



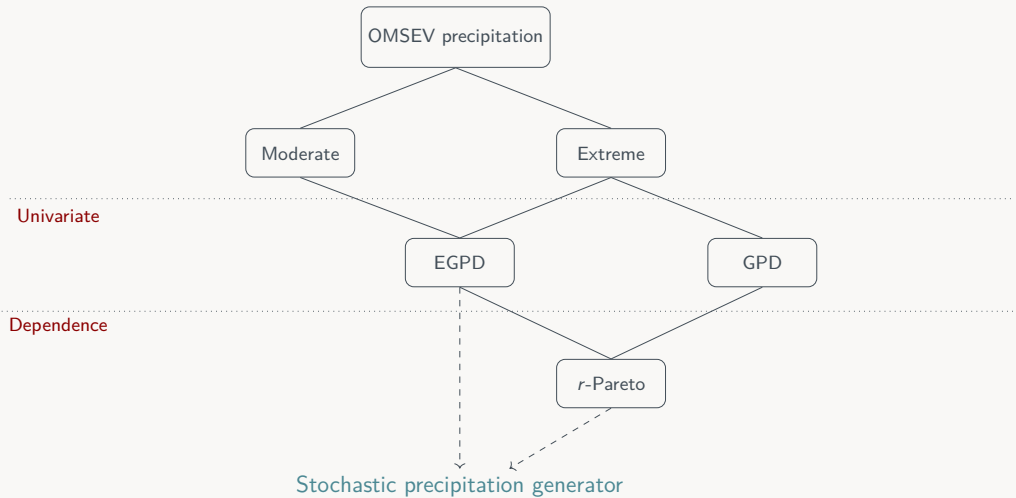
COMEPHORE estimates the advection contribution
OMSEV provides the final dependence structure

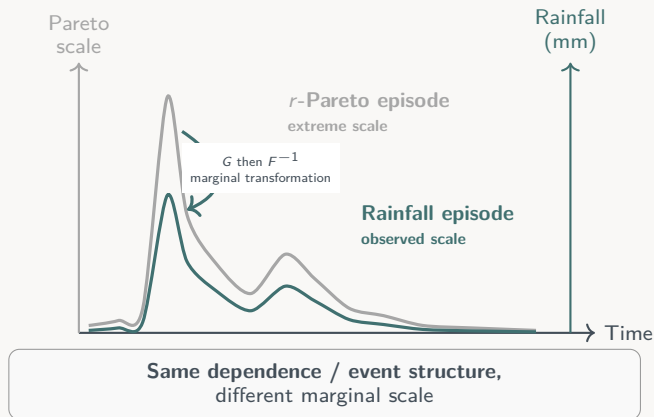
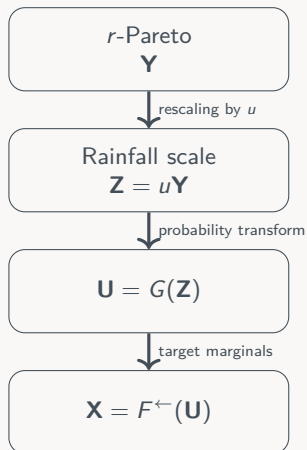
Parameter estimates over $|\mathcal{E}| = 511$ episodes
 $\hat{\eta}_1 = 3.896$ and $\hat{\eta}_2 = 2.221$ fixed (COMEPHORE)



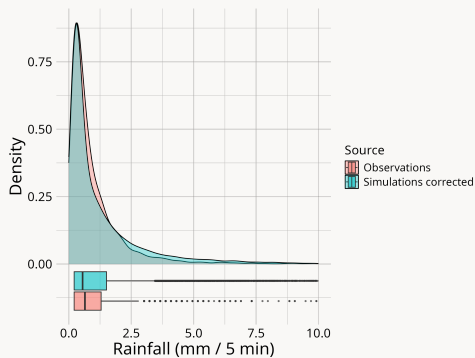
	Parameter	Estimate (JK CI)
Spatial	$\hat{\beta}_1$	1.076 [0.807, 1.344]
	$\hat{\alpha}_1$	0.116 [0.000, 0.263]
Temporal	$\hat{\beta}_2$	3.911 [2.675, 5.148]
	$\hat{\alpha}_2$	0.654 [0.485, 0.824]

STOCHASTIC SIMULATION OF PRECIPITATION EVENTS

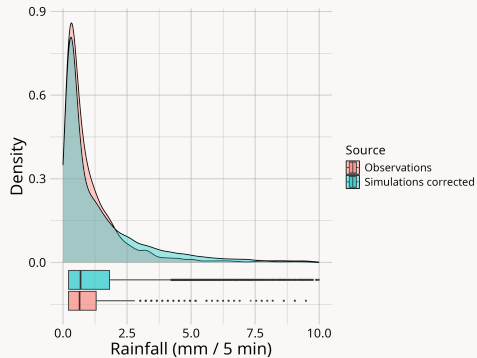




100 simulations by episode



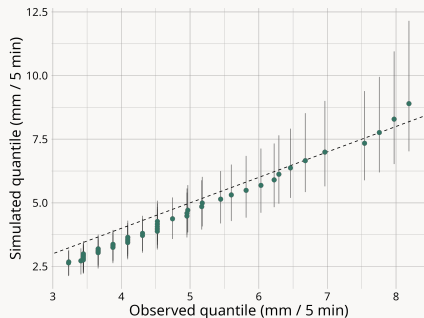
UM density



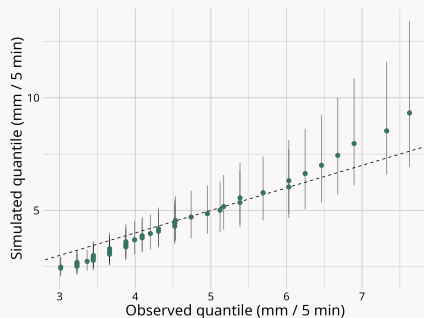
IEM density

Comparison of marginal distributions (above-zero rainfall)

100 simulations by episode

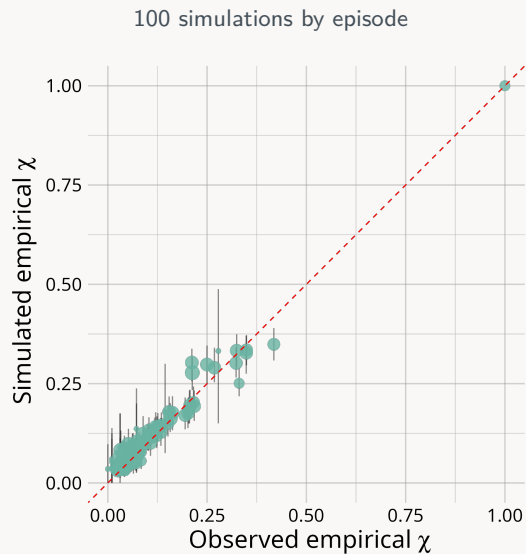


UM



IEM

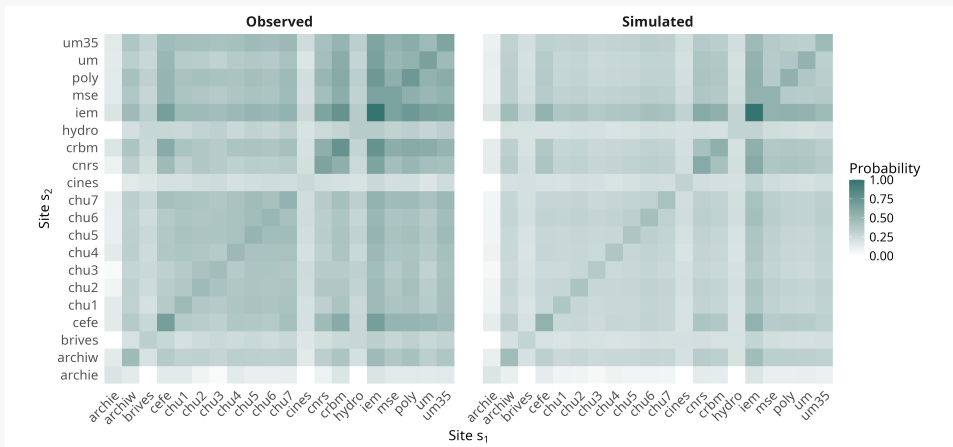
Comparison of extreme quantiles (above 95th percentile)



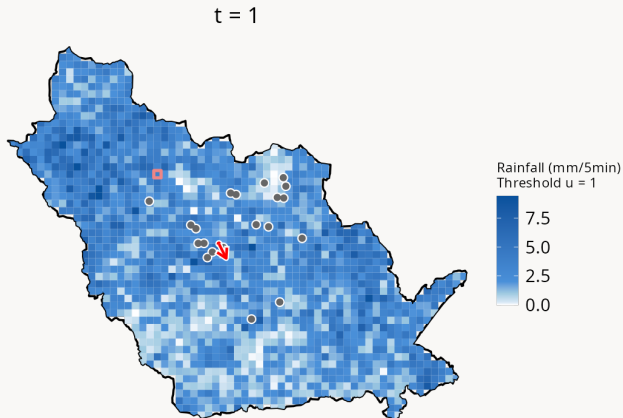
Spatio-temporal r -extremal dependence

DEPENDENCE VALIDATION OF SIMULATED PRECIPITATION

100 simulations by episode

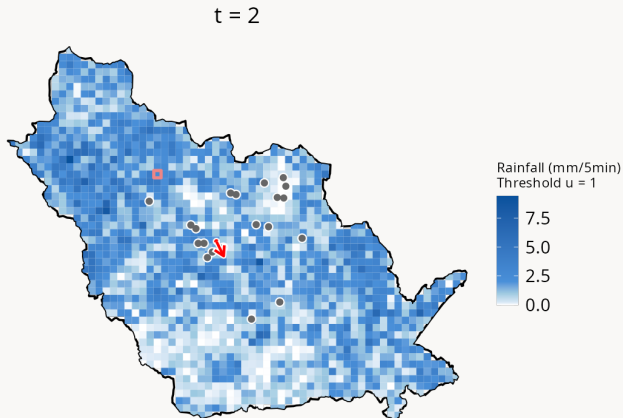


$$\mathbb{P}(X_{s_1,t}^* > q, X_{s_2,t}^* > q | X_{\text{IEM},t}^* > q, X_{s_0,t_0}^* > q)$$



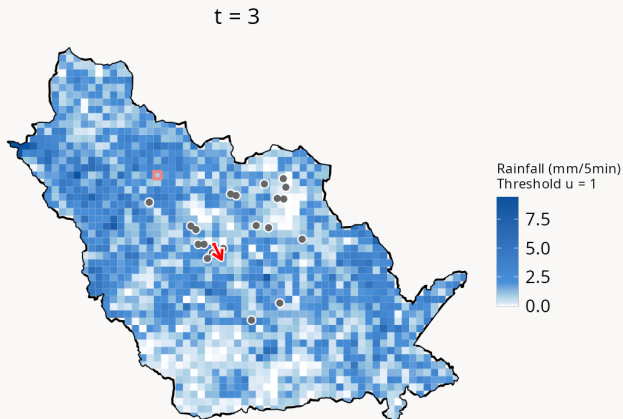
Over the OMSEV network area

- Spatial resolution: $100 \text{ m} \times 100 \text{ m}$
- Temporal resolution: 5 min
- Stochastic rainfall simulation



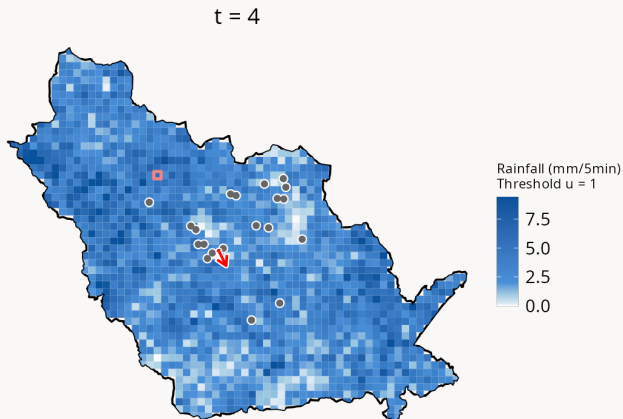
Over the OMSEV network area

- Spatial resolution: $100 \text{ m} \times 100 \text{ m}$
- Temporal resolution: 5 min
- Stochastic rainfall simulation



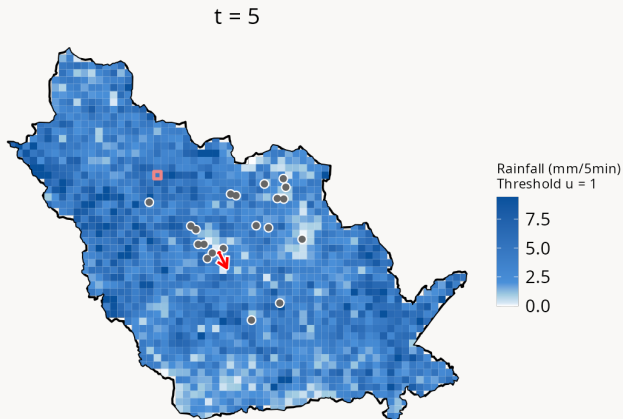
Over the OMSEV network area

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



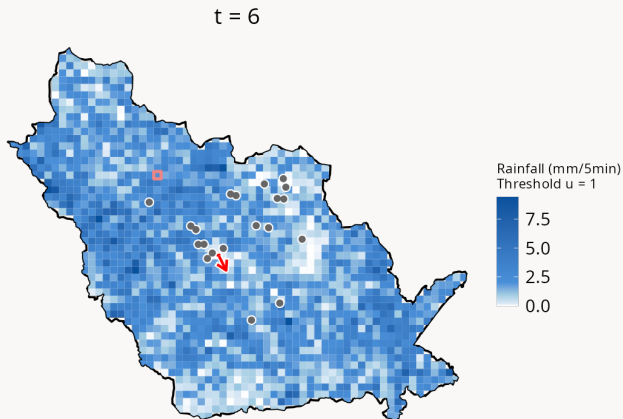
Over the OMSEV network area

- Spatial resolution: $100 \text{ m} \times 100 \text{ m}$
- Temporal resolution: 5 min
- Stochastic rainfall simulation



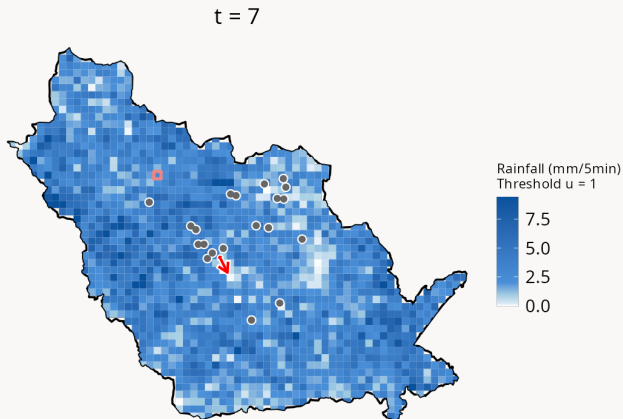
Over the OMSEV network area

- Spatial resolution: $100 \text{ m} \times 100 \text{ m}$
- Temporal resolution: 5 min
- Stochastic rainfall simulation



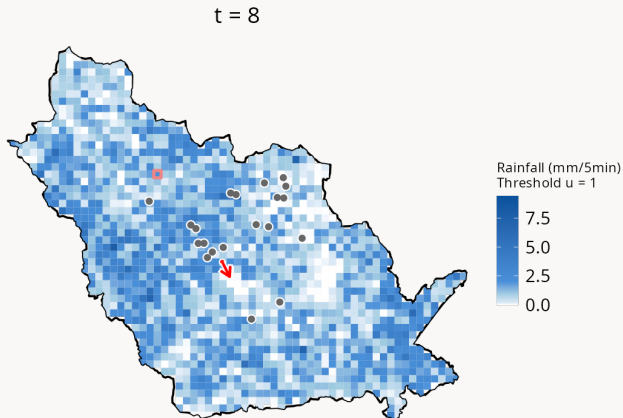
Over the OMSEV network area

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



Over the OMSEV network area

- Spatial resolution: $100 \text{ m} \times 100 \text{ m}$
- Temporal resolution: 5 min
- Stochastic rainfall simulation



Over the OMSEV network area

- Spatial resolution: $100 \text{ m} \times 100 \text{ m}$
- Temporal resolution: 5 min
- Stochastic rainfall simulation

AXIS 2

Statistical downscaling

From coarse-resolution rainfall information
to local high-resolution distributions

Learning relationships between coarse-scale information and local precipitation

Classical statistical models

Regression, GLM/GAM, analogues,
neural networks

MARAUN et al., 2010; OBAKRIM, 2022

Extreme precipitation

EVT-based models,
Quantile regression

KALLACHE et al., 2011;
TAREGHIAN and RASMUSSEN, 2013

This thesis: downscaling

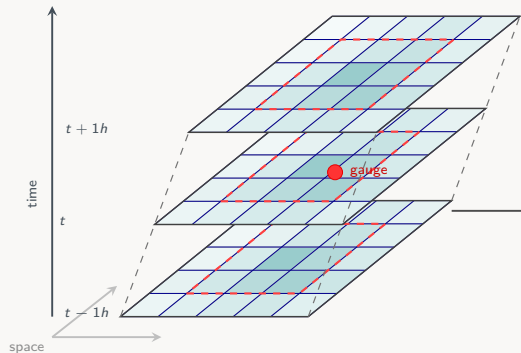
COMEPHORE (hourly/km) → **local 5-min rainfall distribution**

Flexible EGPD-based model - Neural networks

CISNEROS et al., 2024; RICHARDS, 2024

COMEPHORE predictors

spatio-temporal cube around the gauge



43 predictors $C_{s,t}$:
 cube (27) +
 spatial neighbourhood (4) +
 temporal context (6) +
 cube summaries (6)

OMSEV rainfall distribution

Occurrence

$$p_{s,t} = \mathbb{P}(X_{s,t} > 0 \mid C_{s,t})$$

Logistic regression

Positive rainfall intensity

$$X_{s,t} \mid X_{s,t} > 0, C_{s,t} \sim \text{EGPD}(\sigma_{s,t}, \kappa_{s,t}, \xi)$$

MODEL COMPARISON - POSITIVE RAINFALL INTENSITY

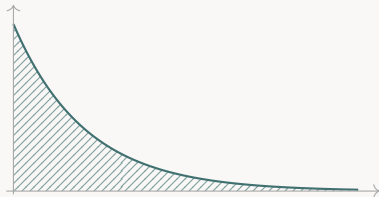
Stationary EGPD	σ, κ constant	no predictors
GLM	$\log \sigma_{s,t}, \log \kappa_{s,t}$ linear	central radar pixel
GAM	$\log \sigma_{s,t}, \log \kappa_{s,t}$ smooth	central radar pixel
Neural network	$\log \sigma_{s,t}, \log \kappa_{s,t}$ nonlinear	43 predictors

Neural network tuning

(8, 4) hidden units $c = 0.3$ mm / 5 min $\xi = 0.18$
network architecture *left-censoring threshold* *tail parameter*

Validation strategy: leave-one-site-out (LOSO)

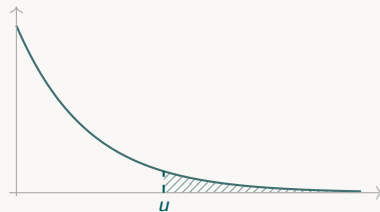
Overall distribution



$$\text{CRPS}(F, x) = \int_{-\infty}^{+\infty} [F(v) - \mathbb{1}\{x \leq v\}]^2 dv$$

Performance over the
whole predictive distribution

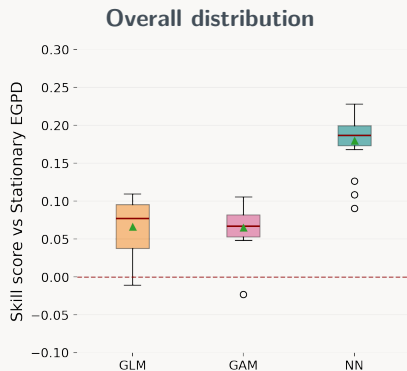
Upper-tail distribution



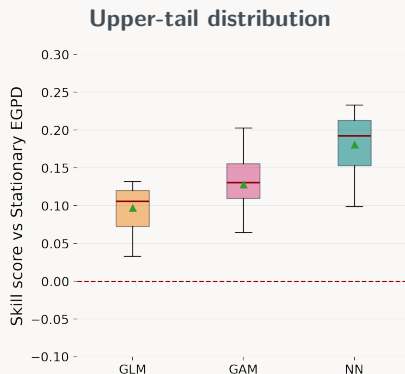
$$\text{twCRPS}(F, x) = \int_{-\infty}^{+\infty} w(v) [F(v) - \mathbb{1}\{x \leq v\}]^2 dv$$

$$w(v) = \mathbb{1}\{v \geq u\}$$

Performance focused on the
upper tail



SS_{CRPS}



SS_{twCRPS}

Skill score relative to the stationary EGPD:

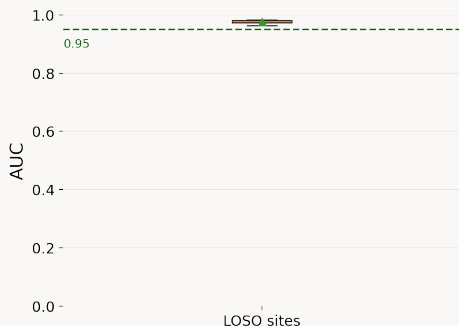
$$SS = 1 - \frac{S_{\text{model}}}{S_{\text{stationary}}}, \quad SS > 0 \Rightarrow \text{improvement}$$

All models outperform the stationary EGPD, with the neural network being the best.

Occurrence: $p_{\text{occ}}(\mathbf{s}, t) = \mathbb{P}(X_{\mathbf{s},t} > 0 \mid C_{\mathbf{s},t})$
Logistic regression – 16 predictors

Discrimination wet/dry events

AUC: higher is better



AUC > 0.95 at all held-out gauges

Probabilistic skill

Brier Skill Score: > 0 = better than reference

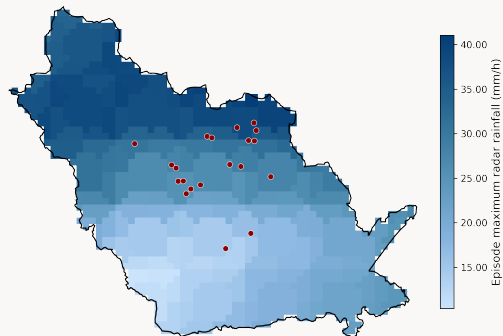


Mean Brier Skill Score: 0.215

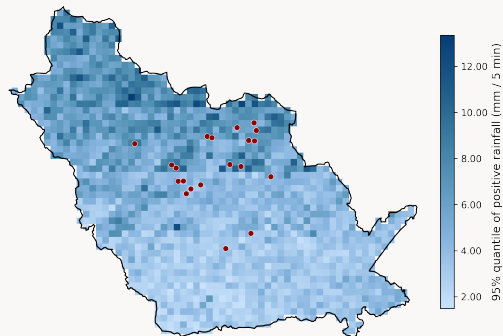
Rainfall occurrence can be accurately transferred to locations not used for training.

Example of spatial reconstruction – 6-8 September 2022

Based on a 5-min conditional simulation

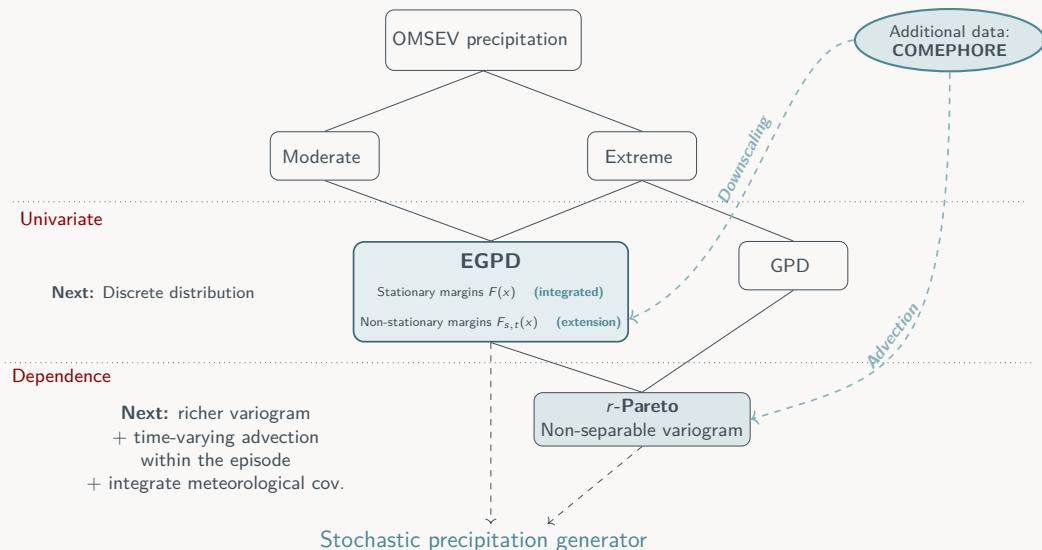


Maximum COMEPHORE rainfall over the episode



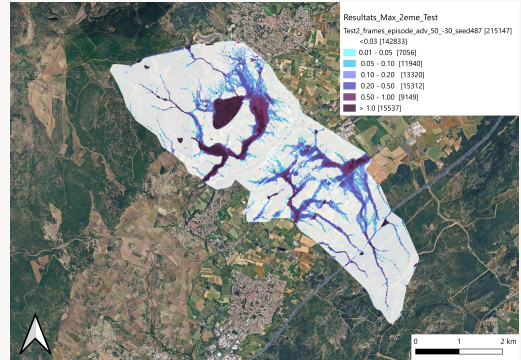
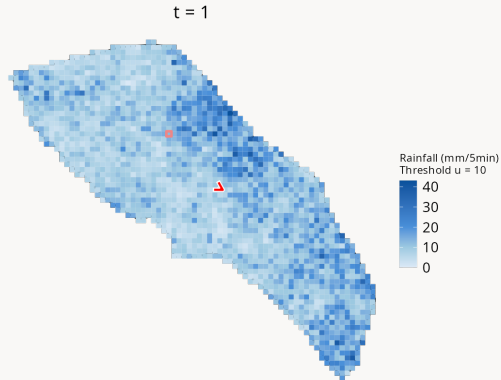
95th percentile of downscaled rainfall over the episode

CONCLUSION & OUTLOOK



Use: Shallow-water flow model (SW2D – Inria)

→ Agustin Bastida's internship (summer 2026)



Fabregues catchment (Hérault, France) – 2D shallow-water flow model (SW2D – Inria)

How can rainfall be reproduced at high spatio-temporal resolution for hydrological modeling?

Question 1

How to generate extreme rainfall events
while preserving their spatio-temporal structure?

AXIS 1 – SPG

Scientific contribution

Preprint – under revision

generain – R package

 chloesrcb/generain

Question 2

How to recover local rainfall distributions
from coarser spatio-temporal information?

AXIS 2 – DOWNSCALING

Scientific contribution

Preprint – in preparation

Python implementation

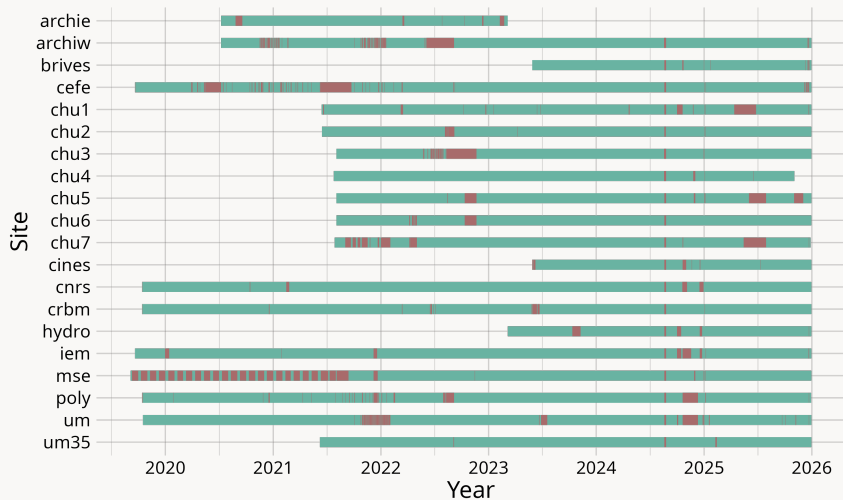
 chloesrcb/downscaling

- ALLARD, D., AILLIOT, P., MONBET, V., & NAVEAU, P. (2015, March 20). *Stochastic weather generators: An overview of weather type models* (Vol. 156).
- BENOIT, L., ALLARD, D., & MARIETHOZ, G. (2018). Stochastic Rainfall Modeling at Sub-kilometer Scale. *Water Resources Research*, 54(6), 4108–4130.
- BUHL, S., DAVIS, R. A., KLÜPPELBERG, C., & STEINKOHL, C. (2019). Semiparametric estimation for isotropic max-stable space-time processes. *Bernoulli*, 25, 2508–2537.
- CISNEROS, D., RICHARDS, J., DAHAL, A., LOMBARDO, L., & HUSER, R. (2024). Deep graphical regression for jointly moderate and extreme Australian wildfires. *Spatial Statistics*, 59, 100811.
- 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).
- GAETAN, C., OPITZ, T., TOULEMONDE, G., et al. (2025). Statistical modeling of extreme precipitation. In *Handbook on statistics of extremes*. Chapman; Hall.
- HUSER, R., & DAVISON, A. C. (2014). Space–time modelling of extreme events. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(2), 439–461.
- KALLACHE, M., VRAC, M., NAVEAU, P., & MICHELANGELI, P.-A. (2011). Nonstationary probabilistic downscaling of extreme precipitation. *Journal of Geophysical Research: Atmospheres*, 116(D5).

- MARAUN, D., WETTERHALL, F., IRESON, A. M., CHANDLER, R. E., KENDON, E. J., WIDMANN, M., BRIENEN, S., RUST, H. W., SAUTER, T., THEMESSEL, M., et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of geophysics*, 48(3).
- 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*.
- OBAKRIM, S. (2022). *Statistical downscaling and climate change in the coastal zone* [Doctoral dissertation, Université de Rennes].
- OBAKRIM, S., BENOIT, L., & ALLARD, D. (2025). A multivariate and space-time stochastic weather generator using a latent gaussian framework. *Stochastic Environmental Research and Risk Assessment*, 39(9), 3677–3701.
- PICKANDS III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, 119–131.
- RICHARDS, J. (2024). pinnEV: Partially Interpretable Neural Networks for Extreme Values [R package].
- SCHLEISS, M., CHAMOUN, S., & BERNE, A. (2014). Stochastic simulation of intermittent rainfall using the concept of “dry drift”. *Water Resources Research*, 50(3), 2329–2349.
- TABARY, P., DUPUY, P., L’HENAFF, G., GUEGUEN, C., MOULIN, L., & LAURANTIN, O. (2012). A 10-year (1997–2006) reanalysis of Quantitative Precipitation Estimation over France: Methodology and first results. *IAHS-AISH publication*, (351), 255–260.

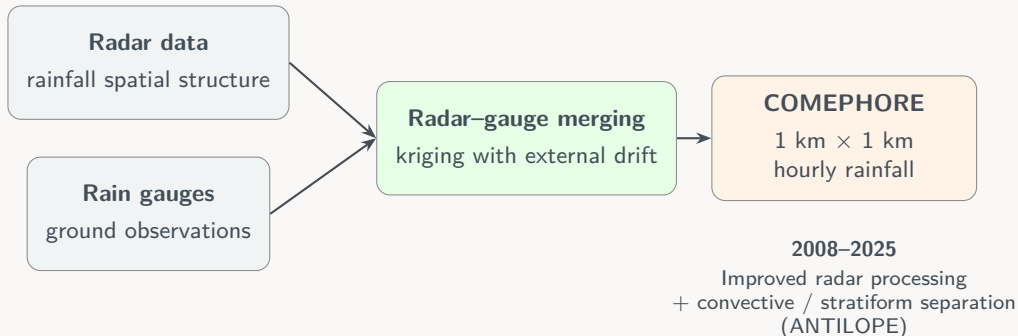
- TAREGHIAN, R., & RASMUSSEN, P. (2013). Statistical downscaling of precipitation using quantile regression. *Journal of Hydrology*, 487, 122–135.
- WILKS, D. S., & WILBY, R. L. (1999). The weather generation game: A review of stochastic weather models. *Progress in physical geography*, 23(3), 329–357.

DATA ISSUE: MISSING VALUES



Missing values according to site activation periods (OMSEV dataset)

HOW ARE COMEPHORE DATA PRODUCED?



Météo-France precipitation reanalysis combining radar and rain-gauge information.

Generalized Pareto Distribution



Extended GPD³

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

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

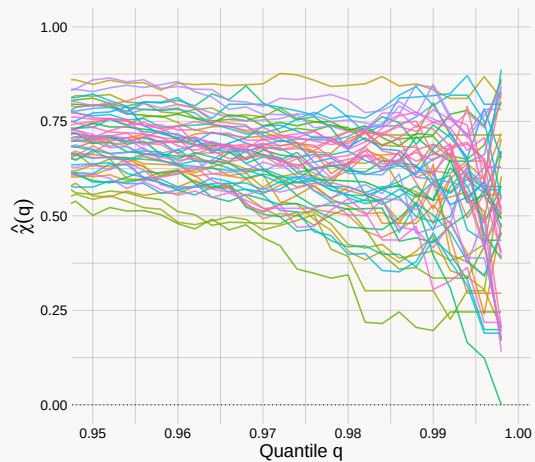
- Models extreme precipitation
- Depends on a threshold choice

$$F_{\text{EGPD}}(x) = P \left(H_\xi \left(\frac{x}{\sigma} \right) \right),$$

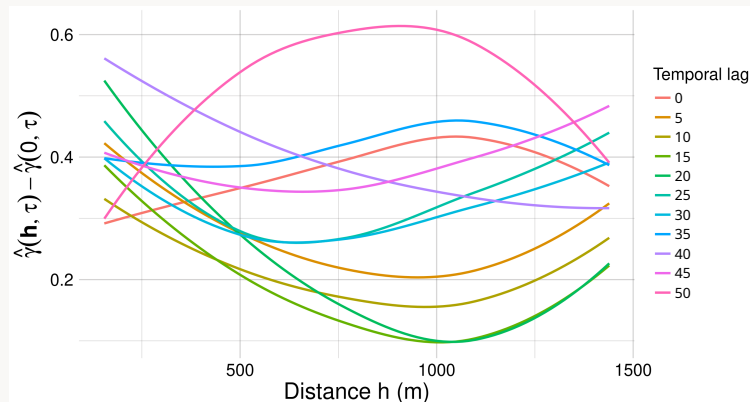
where $P(x) = x^\kappa$, $\kappa > 0$

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

³NAVEAU et al., 2016

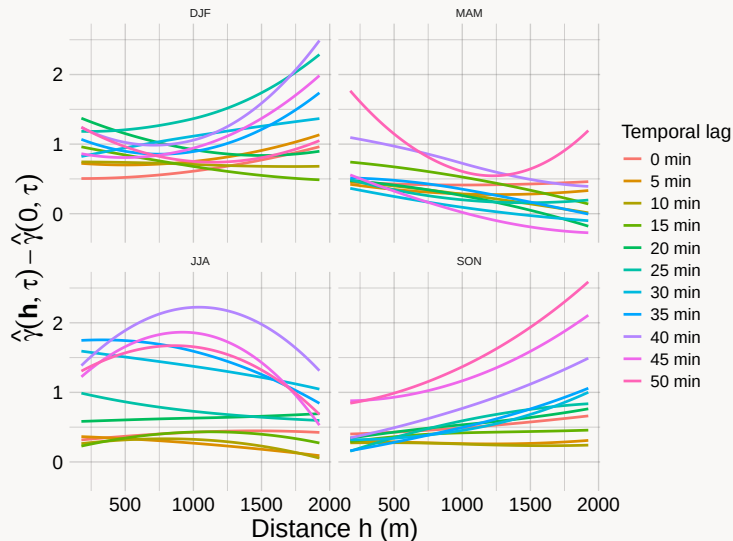


Pairwise χ for the OMSEV dataset



Empirical difference $\gamma(\mathbf{h}, \tau) - \gamma(\mathbf{h}, 0)$ for the OMSEV dataset

NON-SEPARABILITY



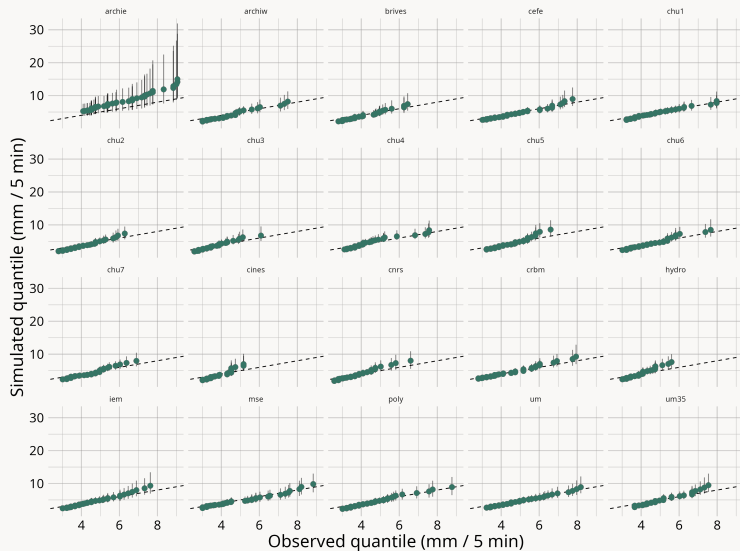
Empirical difference $\gamma(\mathbf{h}, \tau) - \gamma(\mathbf{h}, 0)$ for different seasons (OMSEV dataset)

FROM EXCEEDANCES TO EXTREME EPISODES



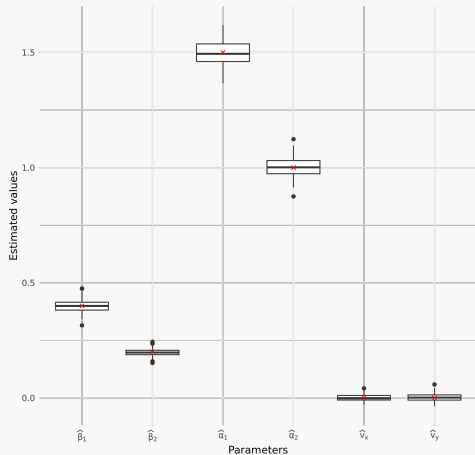
Only episodes separated by a minimum spatial and temporal distance are selected to reduce dependence.

MARGINAL VALIDATION OF SIMULATED PRECIPITATION

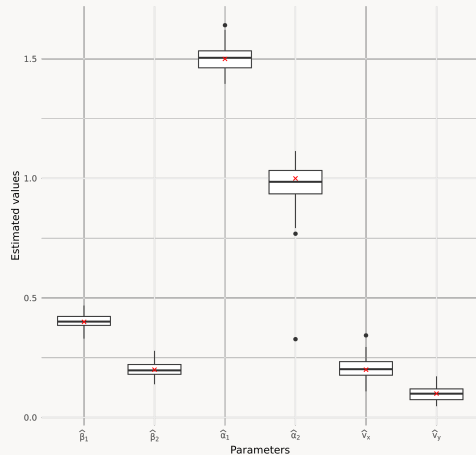


Comparison of extreme quantiles (above 95th percentile)

VALIDATION: CONSTANT ADVECTION



Without advection



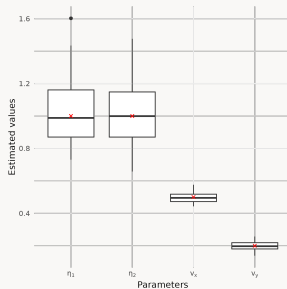
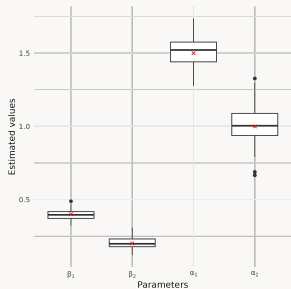
With advection

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

EPISODE-WISE ADVECTION TO GLOBAL MODEL

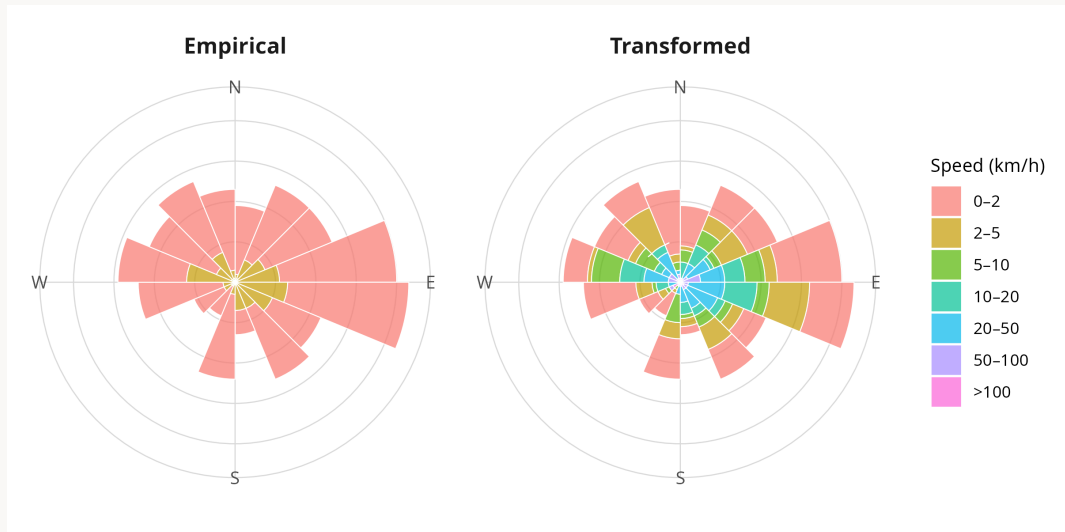
$$\mathbf{V}^{\text{emp}} = \begin{bmatrix} \mathbf{V}^{(1)} \\ \mathbf{V}^{(2)} \\ \vdots \\ \mathbf{V}^{(N)} \end{bmatrix} \in \mathbb{R}^{N \times 2} \xrightarrow{\text{global model}} \underbrace{\mathcal{A}(\mathbf{V}^{(i)}) = \eta_1 \|\mathbf{V}^{(i)}\|^{\eta_2} \frac{\mathbf{V}^{(i)}}{\|\mathbf{V}^{(i)}\|},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

one component = one episode



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

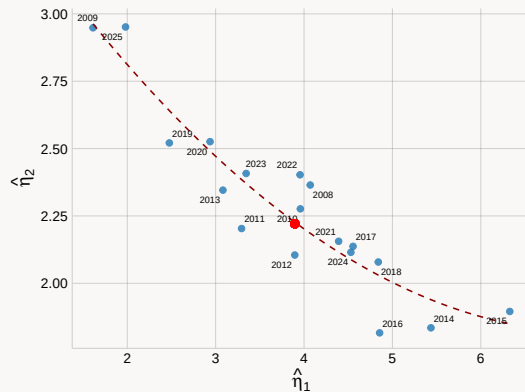
ADVECTION ESTIMATION ON REAL DATA



Stricly positive advection vector wind rose for the OMSEV dataset

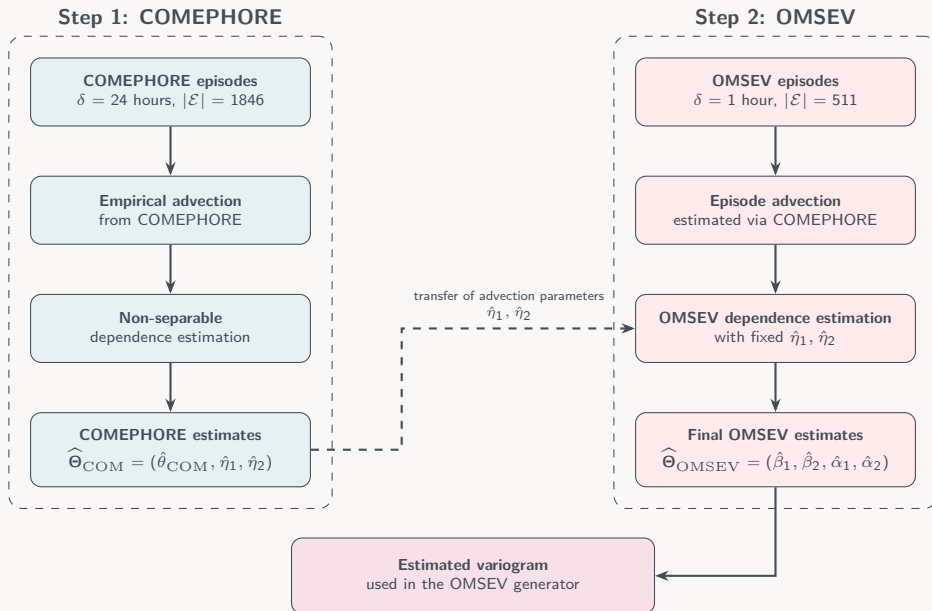
Number of null advection: 26/511 episodes (5%)

JACKKNIFE CONFIDENCE INTERVALS



Jackknife estimates of η_1 and η_2 for the COMEPHORE dataset

ESTIMATION STRATEGY FOR THE EXTREMAL DEPENDENCE STRUCTURE



Advection transformation estimated from COMEPHORE

$$\mathcal{A}(\mathbf{V}_{\text{emp}}^{(i)}) = \eta_1 \|\mathbf{V}_{\text{emp}}^{(i)}\|^{\eta_2} \frac{\mathbf{V}_{\text{emp}}^{(i)}}{\|\mathbf{V}_{\text{emp}}^{(i)}\|}$$

Transformation	η_1	η_2	NCL	CLIC
Nonlinear transformation	3.90	2.22	620496.8	1242562
Linear transformation	9.08	1 (fixed)	621623.6	1244470

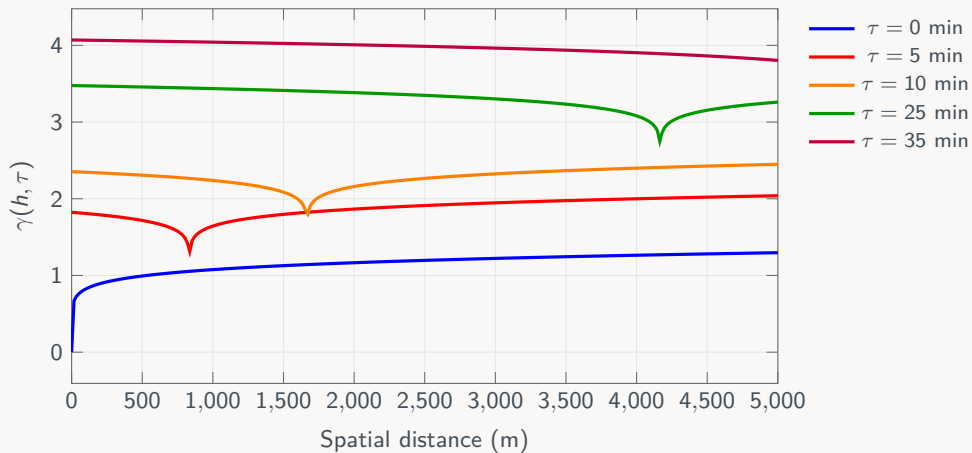
$$\Delta\text{NCL} = 620496.8 - 621623.6 = -1126.8$$

$$\Delta\text{CLIC} = 1242562 - 1244470 = -1908$$

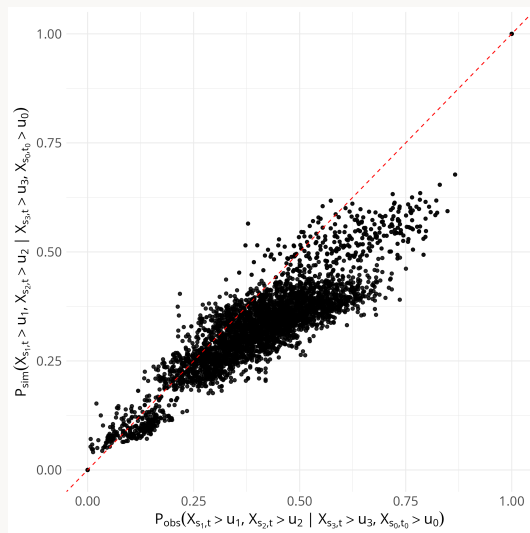
The nonlinear transformation remains preferred after accounting for model complexity.

η_2 is estimated rather than fixed to 1

SPATIO-TEMPORAL VARIOGRAM WITH ADVECTION



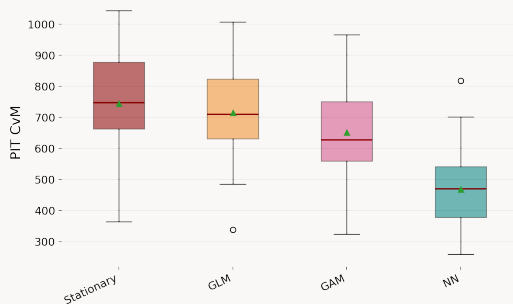
100 simulations by episode



Trivariate conditional probabilities

Downscaling - Calibration of Predictive Distributions

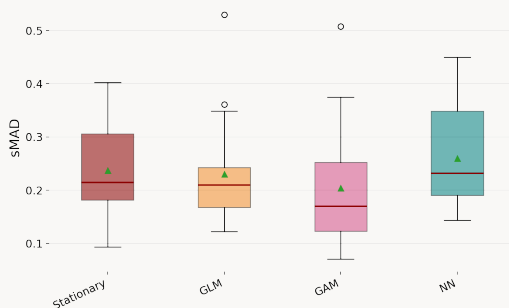
Overall calibration



PIT Cramér-von Mises

Deviation of PIT values from uniformity
lower is better

Upper-tail calibration



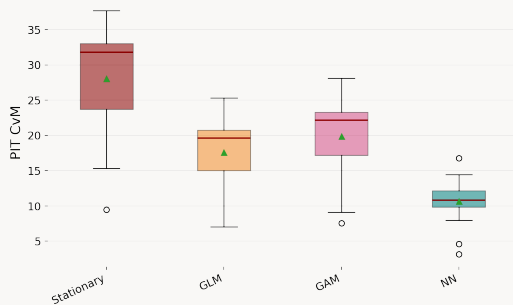
sMAD on exponential margins

Departure from the theoretical upper tail
lower is better

The neural network provides the **best overall compromise between predictive skill and calibration**, while the GAM remains competitive for upper-tail calibration.

Downscaling - Calibration of Predictive Distributions

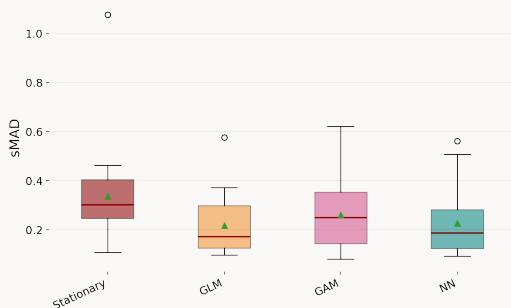
Overall calibration



PIT Cramér-von Mises

Deviation of PIT values from uniformity
lower is better

Upper-tail calibration

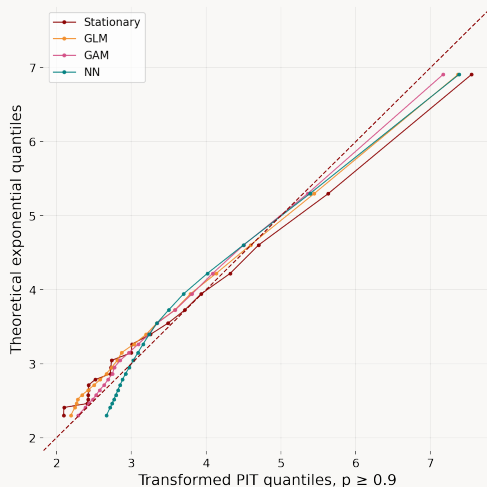


sMAD on exponential margins

Departure from the theoretical upper tail
lower is better

The neural network provides the **best overall compromise between predictive skill and calibration with censored data**, while the GAM remains competitive for upper-tail calibration.

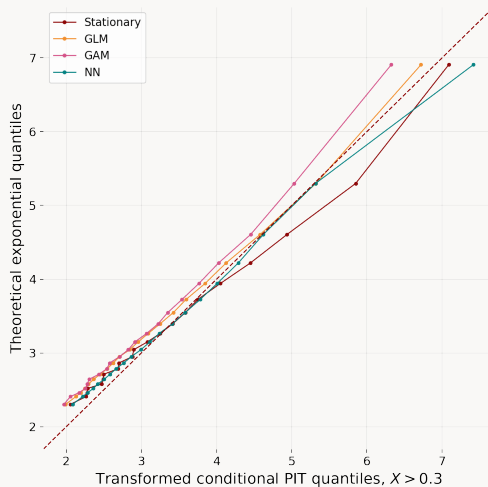
DOWNSCALING - UPPER-TAIL CALIBRATION DIAGNOSTICS



Exponential QQ plots of transformed PIT values for quantile levels above 0.90.

All models remain reasonably close to the theoretical exponential tail; differences between conditional models are moderate.

DOWNSCALING - UPPER-TAIL CALIBRATION DIAGNOSTICS



Exponential QQ plots of transformed PIT values for quantile levels above 0.90 with **left-censoring**.

All models remain reasonably close to the theoretical exponential tail; differences between conditional models are moderate.

Predictor groups	Number	Description
Temporal	6	Sine/cosine of time of day, day of year, month
Spatial	4	Latitude, longitude for site location and central radar location
Cube radar	27	$3 \times 3 \times 3$ cube of radar values
Radar summaries	6	Max, mean, sum over the cube and for the current time step

⇒ 43 predictors in total

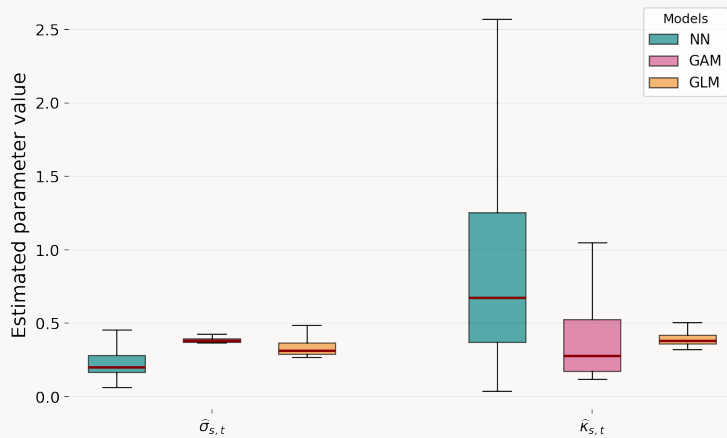
Comparison of predictive performance for positive rainfall intensities.

Model	$\Delta \text{twCRPS}_{\text{sum}}$	ΔtwCRPS ($\times 10^{-3}$)	$\Delta \text{CRPS}_{\text{sum}}$	ΔCRPS ($\times 10^{-2}$)	ΔsMAD ($\times 10^{-2}$)	ΔPIT CvM
NN	0.00	0.00	0.00	0.00	5.56	0.00
GAM	292.27	2.00	170.29	2.75	0.00	183.05
GLM	462.21	3.16	173.94	2.72	2.66	246.49
Stationary	1027.78	6.86	270.67	4.33	3.31	275.68

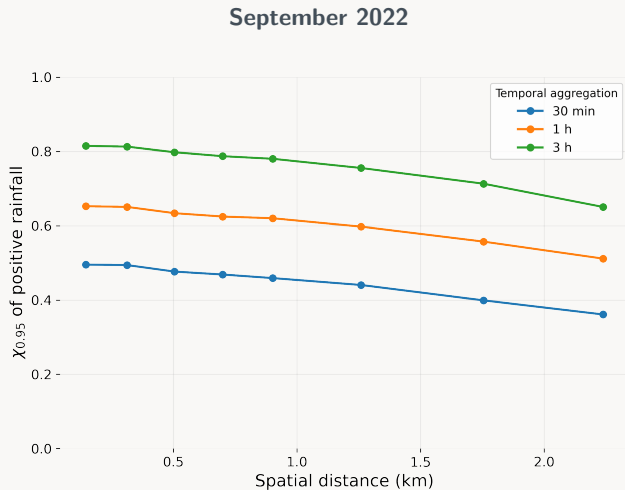
Comparison of predictive performance for positive rainfall intensities.

Model	$\Delta\text{twCRPS}_{\text{sum}}$	ΔtwCRPS ($\times 10^{-3}$)	$\Delta\text{CRPS}_{\text{sum}}$	ΔCRPS ($\times 10^{-2}$)	ΔsMAD ($\times 10^{-2}$)	ΔPIT CvM
NN	0.00	0.00	0.00	0.00	0.86	0.00
GAM	292.27	2.00	170.29	2.75	0.43	9.20
GLM	462.21	3.16	173.94	2.72	0.00	6.92
Stationary	1027.78	6.86	270.67	4.33	1.18	17.43

DOWNSCALED EGPD PARAMETERS



Boxplots of downscaled EGPD parameters for the neural network, GAM and GLM models.



Empirical χ at different temporal aggregation scales, using the 95th percentile of positive rainfall