Bayesian POT analysis of drought extremes

Quantifying the impact of drought index definition

Paolo Besana, Hugues Goosse, Francesco Ragone, Johan Segers

EARTH & LIFE INSTITUTE
FÉDÉRATION WALLONIE-BRUXELLES
EXALT
UCLouvain

The recipe for a drought: the ingredients

LACK OF WATER

How to measure it?

Water balance = Precipitation - Potential Evapotranspiration

TEMPORAL ASPECT

The lacking must persist

WHERE?

The lack can be in the soil, in the atmosphere, underground…

Focus on **meteorological droughts** → lack in the **atmosphere**

The recipe for a drought: how to mix

Drought indices → scalar number to quantify the magnitude

Most widely used : the Standard Precipitation Evapotranspiration Index (Vicente-Serrano et al., 2010)

01

Accumulate Water Balance (WB) **over N months**

02

Choose a **parametric distribution** for the WB → usually **Log-logistic**

03

Fit the distribution **for each month** and transform the **data into uniform** values

04

Map the uniform values in a **standard Gaussian scale**

A non-parametric choice : a z-score S (inspired by Burke et al., 2010)

01

Accumulate Water Balance (WB) **over N months**

02

Compute the **climatological mean** and **standard deviation** of the temporal series **of each month**

03

$$S3(\mathrm{t}, \mathrm{month}) = (WB(\mathrm{t}) - \mu_{\mathrm{month}})/\sigma_{\mathrm{month}}$$

The scientific question

WHAT?

What happens if we use **different indices** for extreme droughts?

HOW?

We use EVT techniques to **compare return times** estimated with S3 and SPEI3

UNDER WHICH CONDITIONS?

Given **common water-balance levels** (equivalent physical conditions)

The model : marginal fit

For each point of the European domain:

1. **Negate the temporal series** and select a **high** threshold to focus on extremes
2. Extract **Peaks-over-Threshold** exceedances and **de-cluster** them to reduce temporal dependence
3. Fit a **Generalized Pareto** (GP) distribution to the exceedances

We obtain **two spatial fields** over Europe for the GP distribution **parameters** $\xi_{i,u}$ and $\phi_{i,u} := \log[\sigma_{i,u}(1+\xi_{i,u})]$.

This parametrization ensures **parameter orthogonality** ↔ **diagonal Fisher information** (Chavez-Demoulin & Davison, 2005).

The model : Bayesian pooling

- Given a threshold u, for each location i, $\hat{\theta}_{i,u} = (\hat{\xi}_{i,u}, \hat{\phi}_{i,u})$ as noisy observations of latent spatial fields modelled with a **Gaussian Markov Random Field** (Rue et al., 2005)

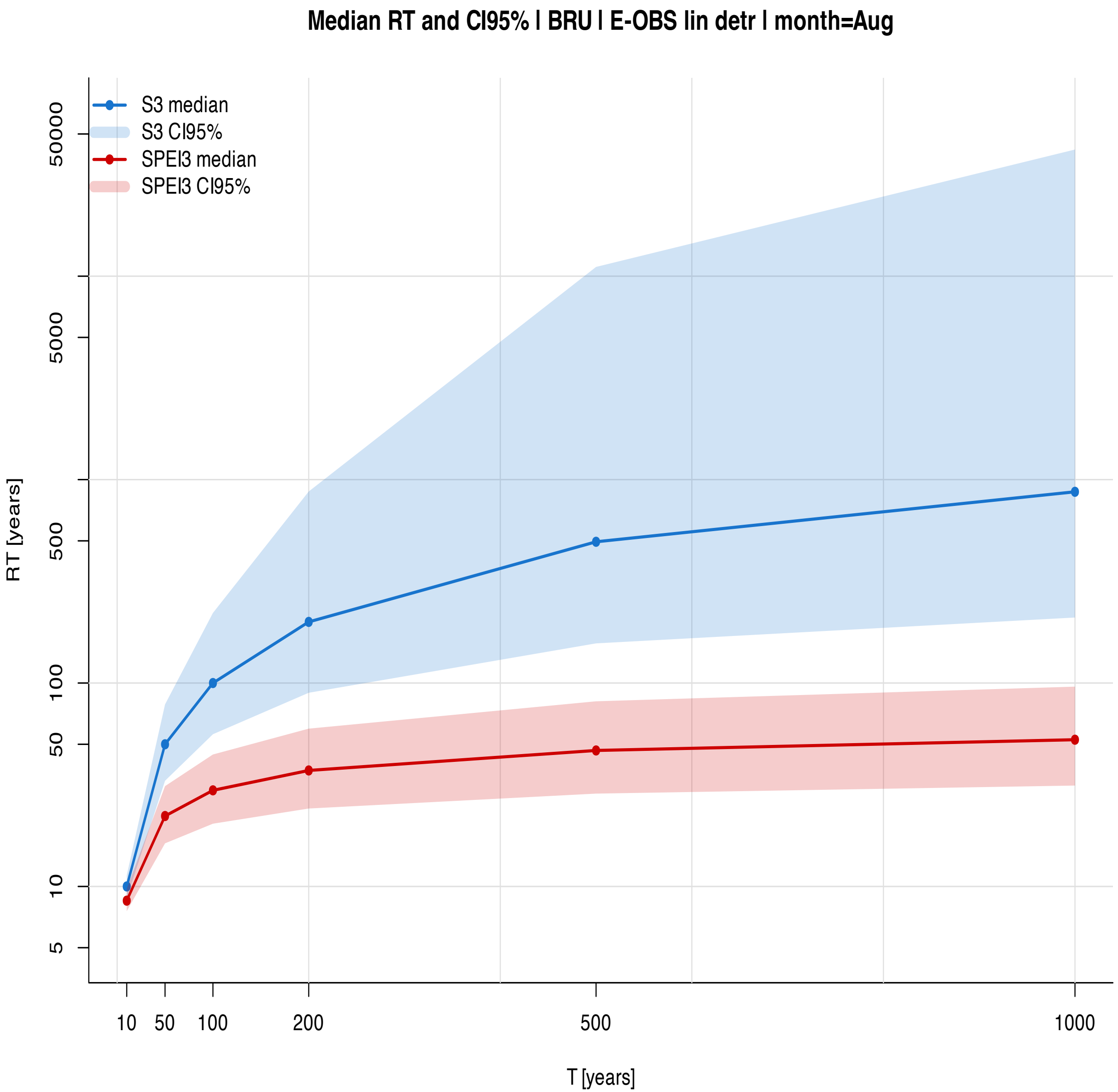
$$\hat{\theta}_{i,u} \mid \theta_{i,u}, \hat{\Sigma}_{i,u} \sim \mathcal{N}_2(\theta_{i,u}, \hat{\Sigma}_{i,u})$$

- The **latent model** is specified as

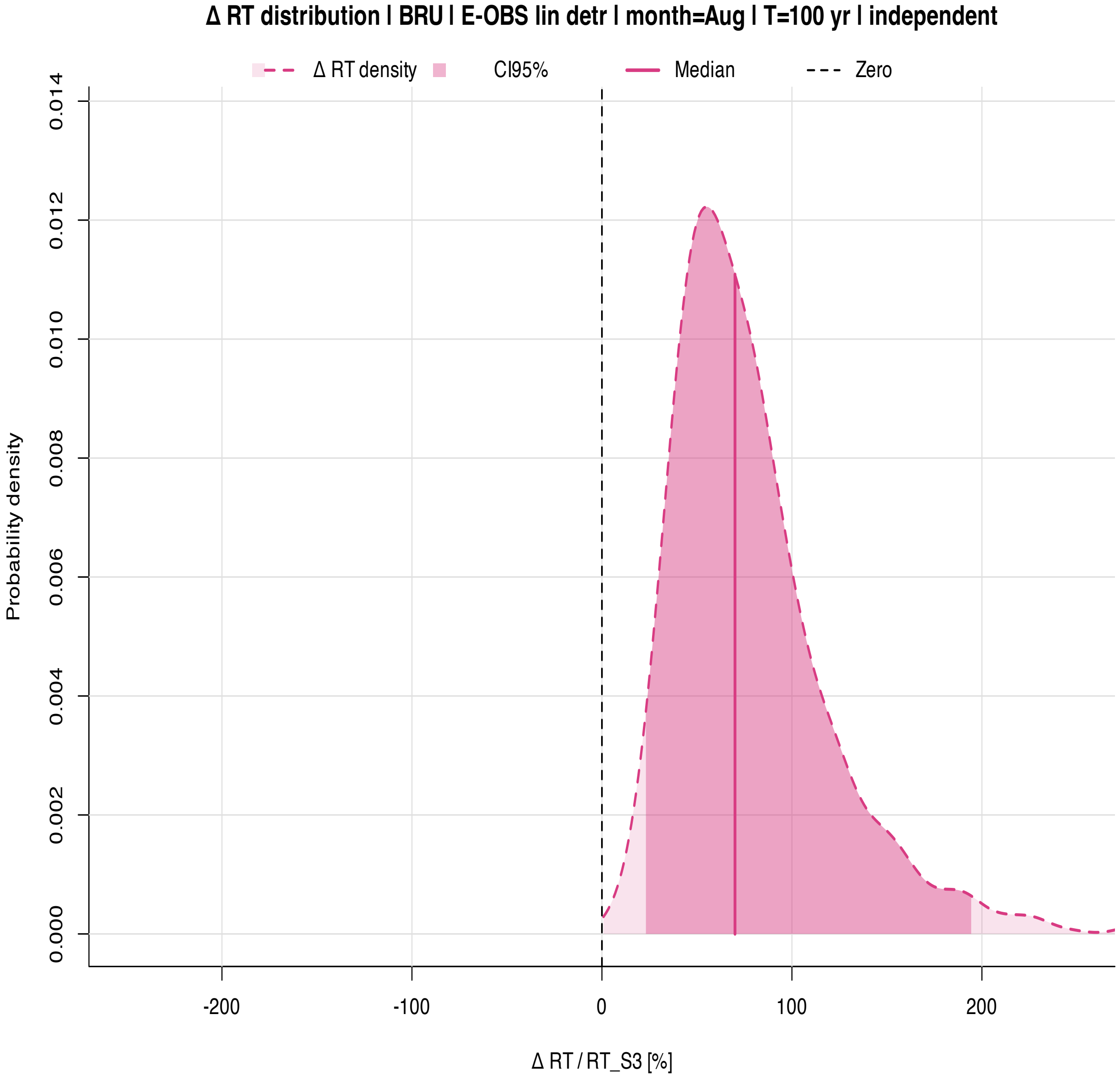
$$\xi_{i,u} = \beta_{\xi,u} + v_{i,u}^{(\xi)},$$
$$\phi_{i,u} = \beta_{\eta,u} + v_{i,u}^{(\eta)}$$

And the intercepts β, together with the latent fields v, **are estimated** under a **Gaussian** likelihood using the **Integrated Nested Laplace Approximation** (INLA) framework (Rue et al., 2009).

Computing the differences



- For **each location**, we obtain the joint **posterior distribution** of $(\xi_{i,u}, \phi_{i,u})$
- Assuming positive dependence between S3 and SPEI3, we **consider two limiting coupling cases** based on copula theory: **independence**, comonotonicity.
- This allows the construction of $\Delta\mathrm{RT} = \mathrm{RT}_{S3} - \mathrm{RT}_{SPEI3}$
- The **sign of ΔRT** is assessed by comparing the bilateral **95% credibility interval** with 0.



Computing the differences

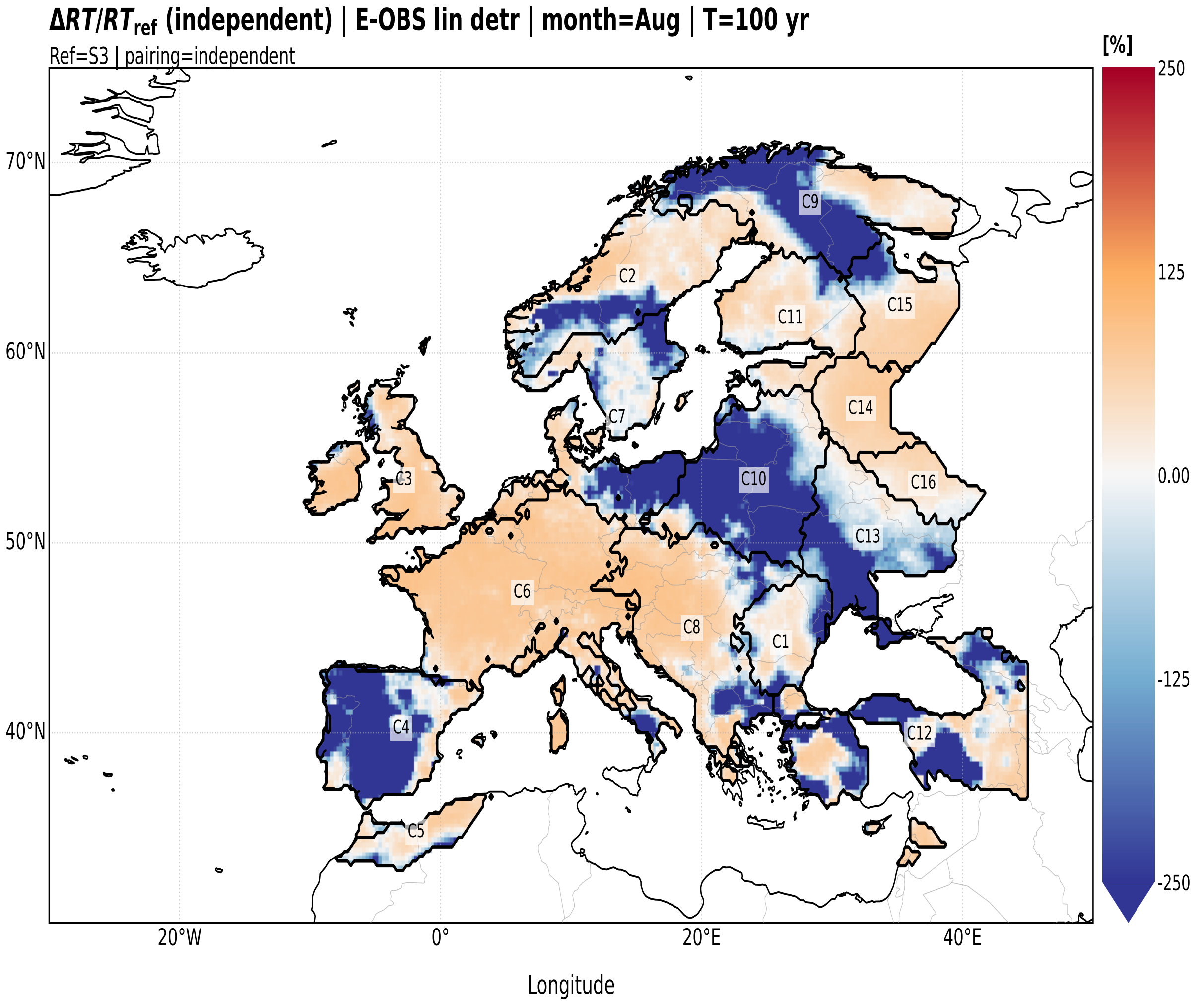
- For **each location**, we obtain the joint **posterior distribution** of $(\xi_{i,u}, \phi_{i,u})$

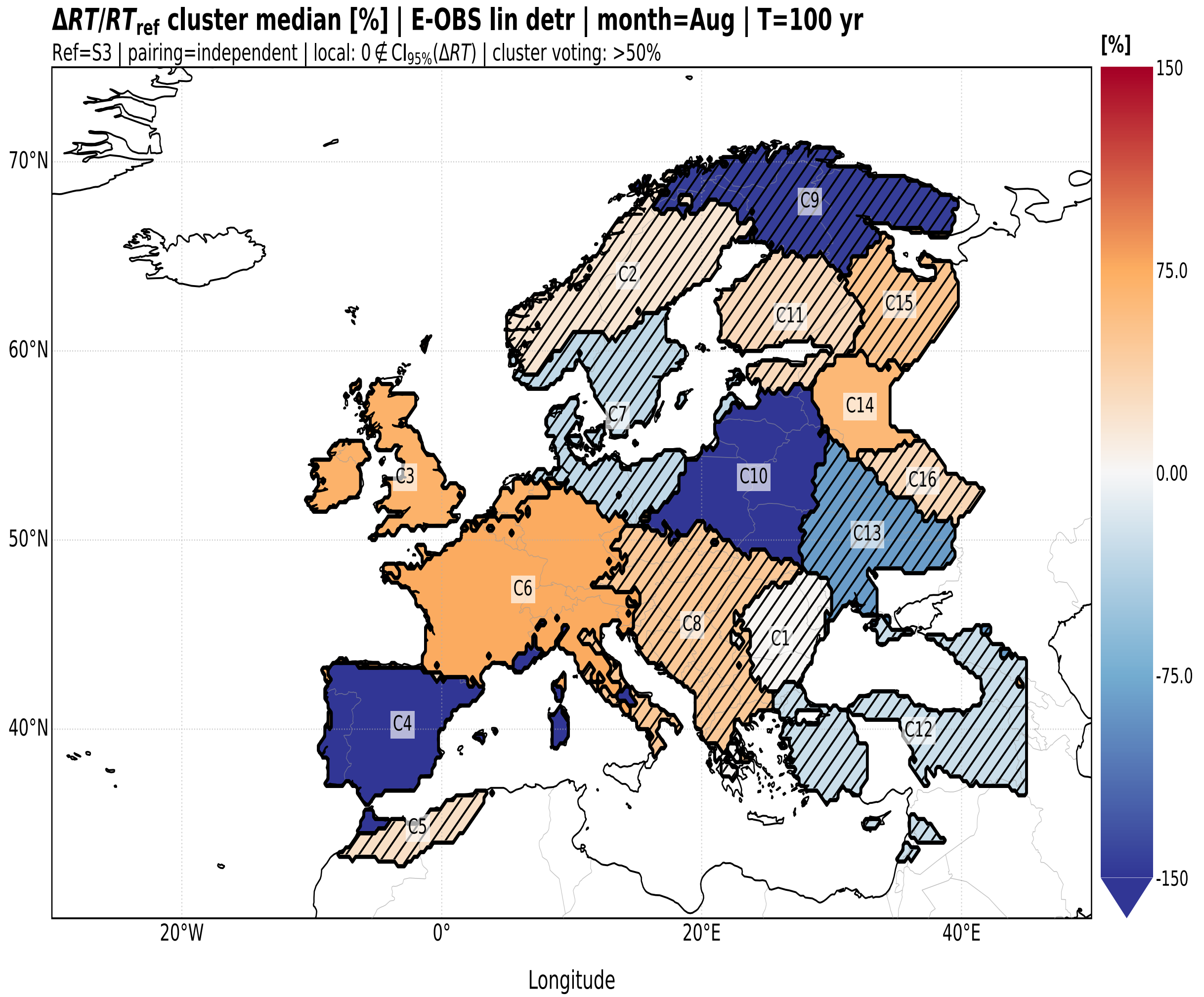
- Assuming positive dependence between S3 and SPEI3, we **consider two limiting coupling cases** based on copula theory: **independence**, comonotonicity.

- This allows the construction of $\Delta \mathrm{RT} = \mathrm{RT}_{\mathrm{S3}} - \mathrm{RT}_{\mathrm{SPEI3}}$

- The **sign of ΔRT** is assessed by comparing the bilateral **95% credibility interval** with 0.

Spatial patterns of ΔRT between S3 and SPEI3

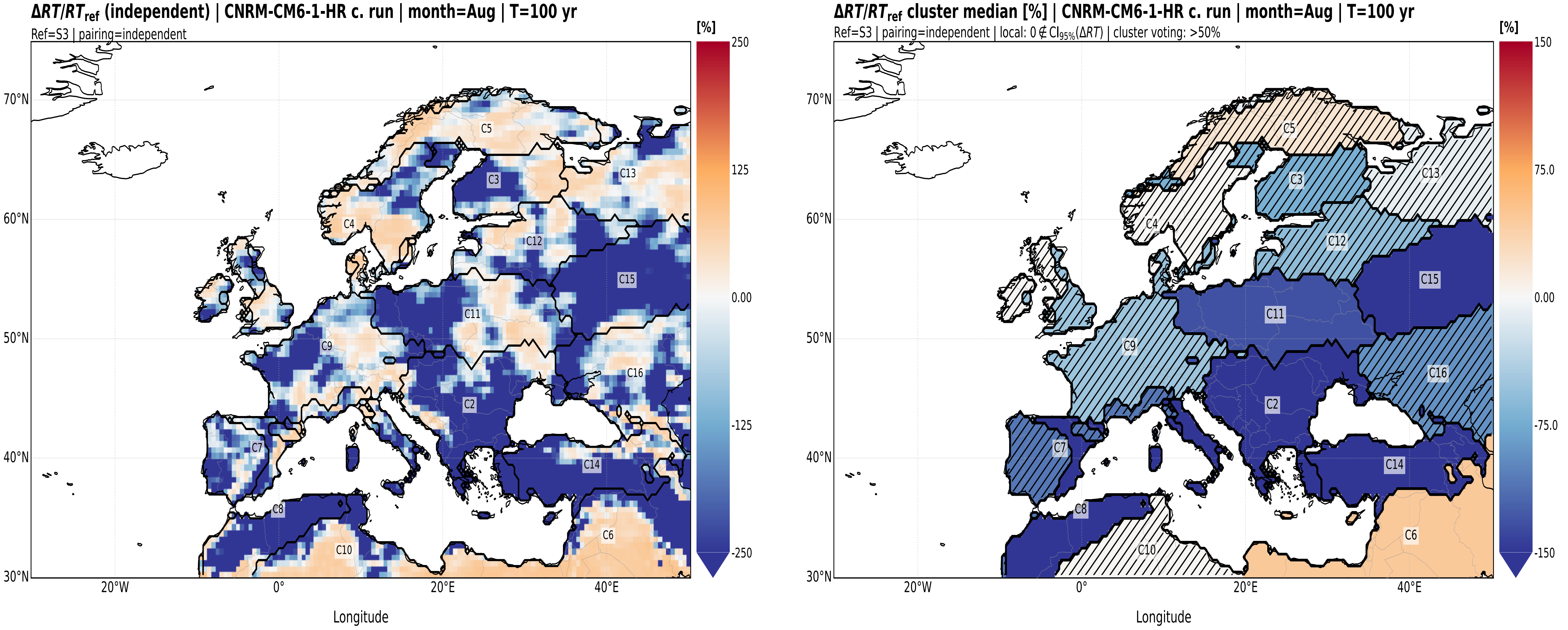




Dataset: E-OBS linear detrended observations

Spatial aggregation is performed using Partition Around Medoids (**PAM**) and the **tail-dependence**-based distance χ proposed by Kiriliouk & Naveau (2020).

Spatial patterns of ΔRT between S3 and SPEI3



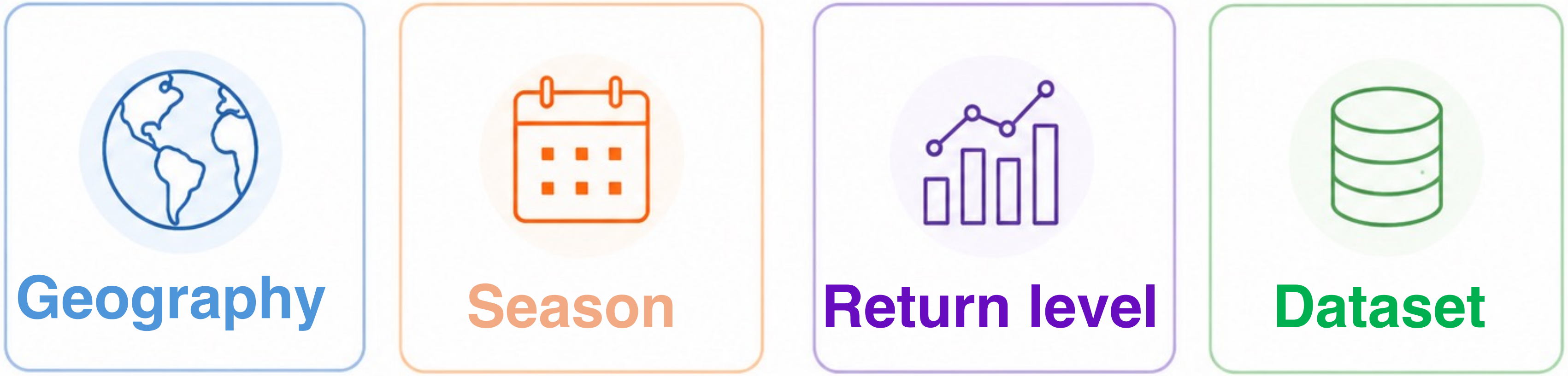
Dataset: CNRM-CM6-1-HR control run (stationary) simulation

Spatial aggregation is performed using Partition Around Medoids (**PAM**) and the **tail-dependence**-based distance χ proposed by Kiriliouk & Naveau (2020).

Conclusions

Using the **same Extreme Value model** with **different drought indices** can lead to **substantial differences** in return-time estimates.

The **log-logistic** fit used in SPEI **may not** optimally **represent** the **tail** behaviour, potentially **inducing distortions** in EVT-based return-time estimates depending on



Bibliography

- Burke, E. J., Perry, R. H., & Brown, S. J. (2010). An extreme value analysis of UK drought and projections of change in the future. Journal of Hydrology, 388(1-2), 131–143.
- Chavez-Demoulin, V., & Davison, A. C. (2005). Generalized additive modelling of sample extremes. Journal of the Royal Statistical Society Series C: Applied Statistics, 54(1),207–222.
- Kiriliouk, A. and Naveau, P. (2020). *Climate extreme event attribution using multivariate peaks-over-thresholds modeling and counterfactual theory*. The Annals of Applied Statistics, 14(3), 1342–1358.
- Rue, H. (2005). Gaussian Markov random fields: Theory and applications. Chapman & Hall/CRC.
- Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian Inference for Latent Gaussian models by using Integrated Nested Laplace Approximations. Journal of the Royal Statistical Society Series B: Statistical Methodology, 71(2), 319–392.
- Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. Journal of Climate, 23(7), 1696–1718.

Backup slides

Paolo Besana, Hugues Goosse, Francesco Ragone, Johan Segers

EARTH & LIFE INSTITUTE
FÉDÉRATION WALLONIE-BRUXELLES
EXALT
UCLouvain

Indices : Potential Evapotranspiration

Potential Evapotranspiration is computed from temperature and latitude following Thornthwaite (1948). For the location i, time t of the month m and year y:

$$\mathrm{PET}_{t,i} = 16\, K_{m,i} \left(\frac{10\, T_{t,i}^{+}}{I_{y,i}} \right)^{a_{y,i}},$$

- $T_{t,i}^{+}$ positive monthly mean temperature,
- $I_{y,i}$ annual heat index,
- $a_{y,i}$ empirical exponent,
- $K_{m,i}$ daylight and month-length correction factor.

The model : Run declustering procedure

1. Identify exceedances above the selected threshold

2. **Group consecutive exceedances** into the same cluster when they are **separated by less than** the **run length r**

3. Start a **new cluster** whenever the **temporal separation exceeds r**

4. Retain **only the maximum** value within each cluster

The model : latent spatial GMRF

Generalized Pareto parameter **estimates** ($\tilde{\theta}_{i,u}$) are treated as noisy observations of **latent spatial fields** (β, v) under a **Gaussian observation model** ($\epsilon_{i,u}$) , accounting for the **estimated covariance** structure ($\hat{\Sigma}$)

$$\tilde{\theta}_{i,u} = \hat{\Sigma}_{i,u}^{-1/2} \begin{pmatrix} \beta_{\xi,u} + v_{i,u}^{(\xi)} \\ \beta_{\phi,u} + v_{i,u}^{(\phi)} \end{pmatrix} + \epsilon_{i,u}, \qquad \epsilon_{i,u} \sim \mathcal{N}_2(0, I_2).$$

Spatial smoothness is imposed through a **Gaussian Markov Random Field prior** penalizing abrupt neighbour-to-neighbour ($i \sim j$) parameter variations.

$$p(v) \propto \exp\left[-\frac{\tau}{2}\sum_{i\sim j}\left(v_i - v_j\right)^2\right]$$

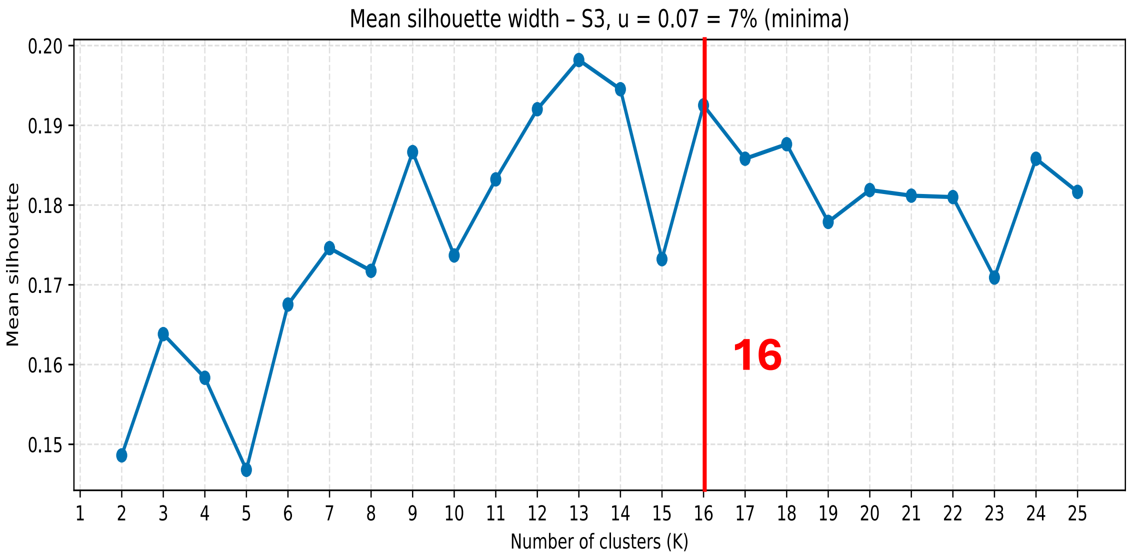
Clustering : PAM with χ

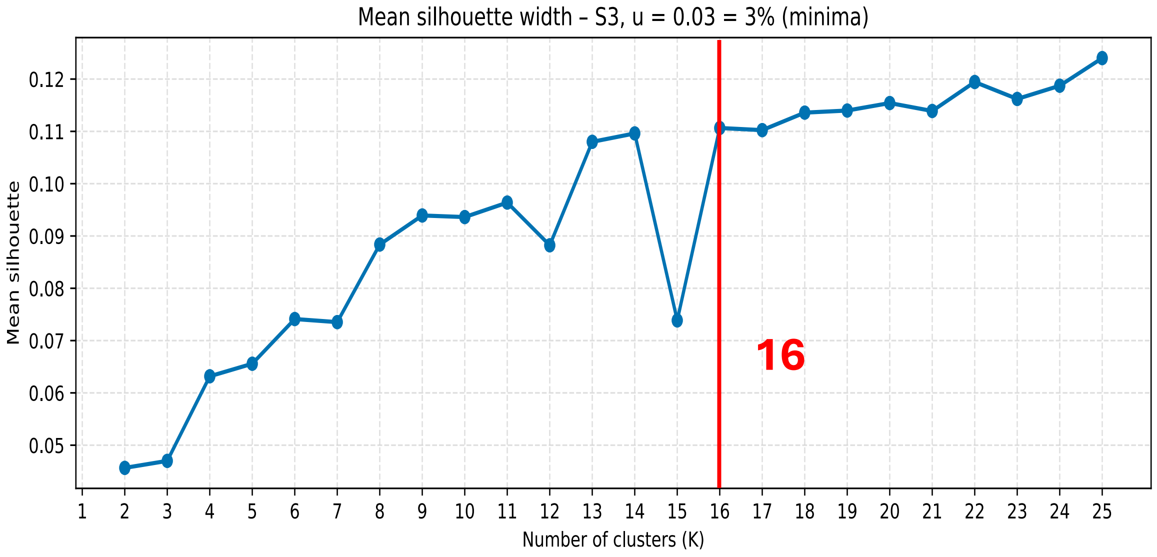
- We use **Partition Around Medoids** (PAM) to identify **regions** sharing **similar extreme-dependence** structures.
- Clustering is performed using a **tail-dependence-based dissimilarity** matrix computed **from pairwise extremal dependence** (Kiriliouk & Naveau, 2020). For locations i and j,

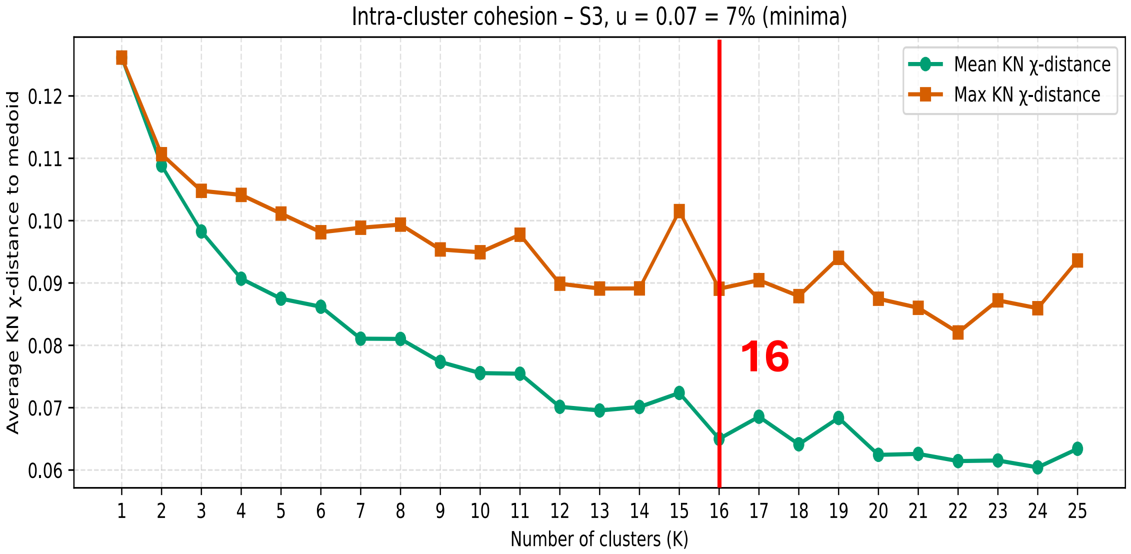
$$d_\chi(i,j) = \frac{1-\chi(i,j)}{2\big(3-\chi(i,j)\big)}.$$

- Locations with **stronger extremal dependence** are therefore **grouped** within the **same** spatial **cluster**.

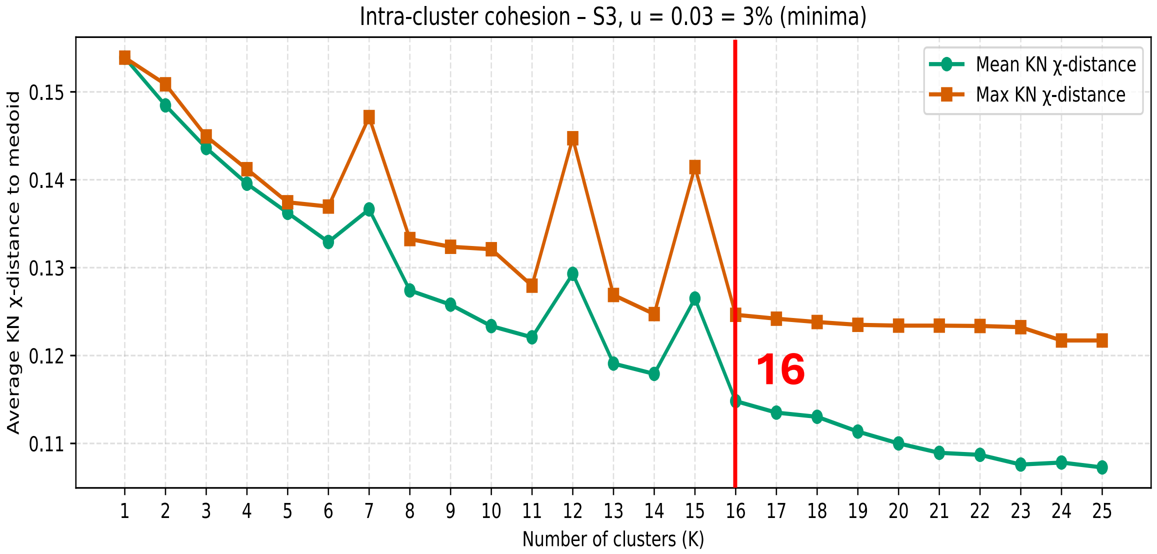
Clustering : choice of the number of clusters







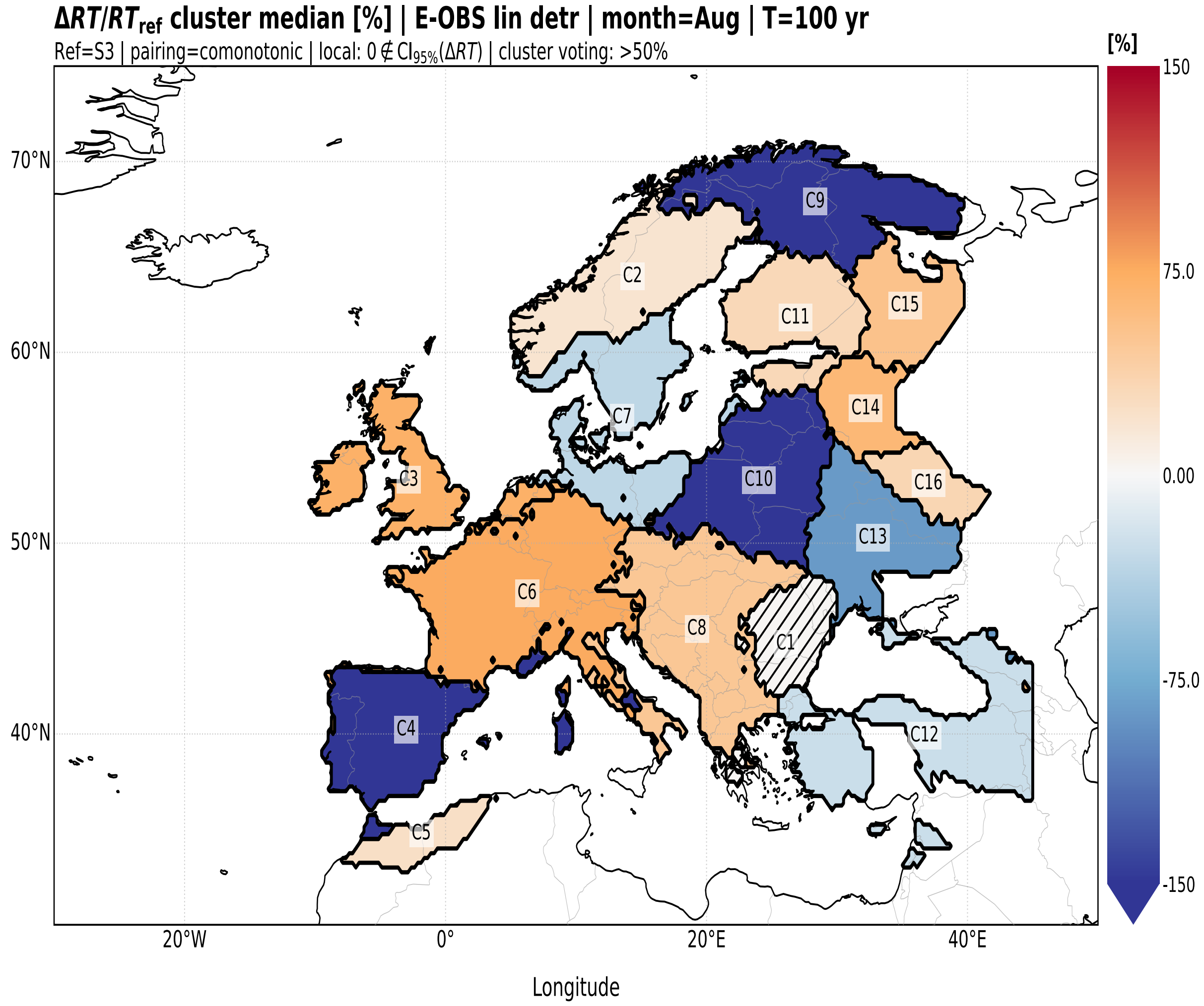
Dataset : E-OBS linear detrended observations



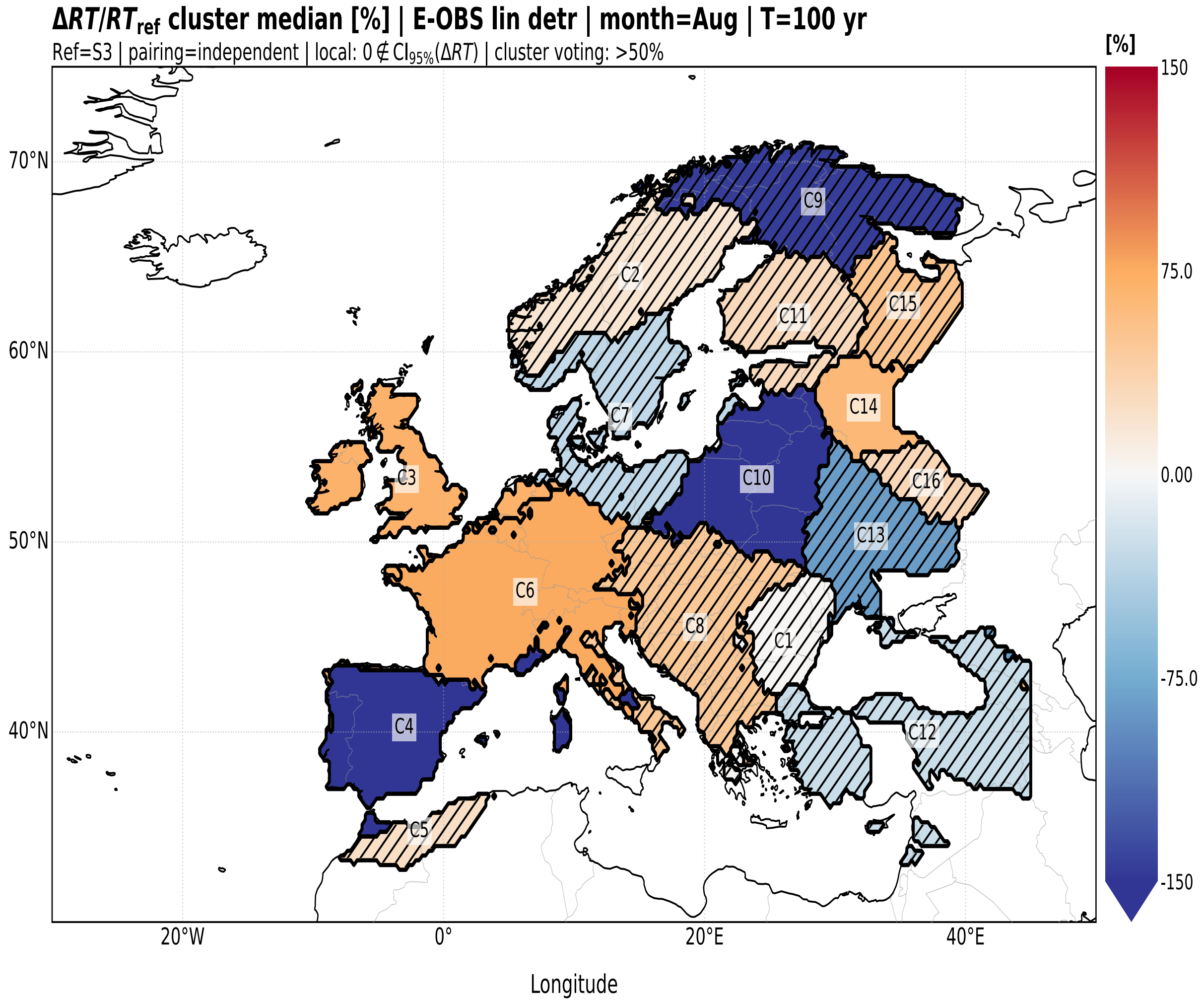
Dataset CNRM-CM6-1-HR control run (stationary) simulation

Spatial patterns of ΔRT : Comonotonic vs Independent

Comonotonic



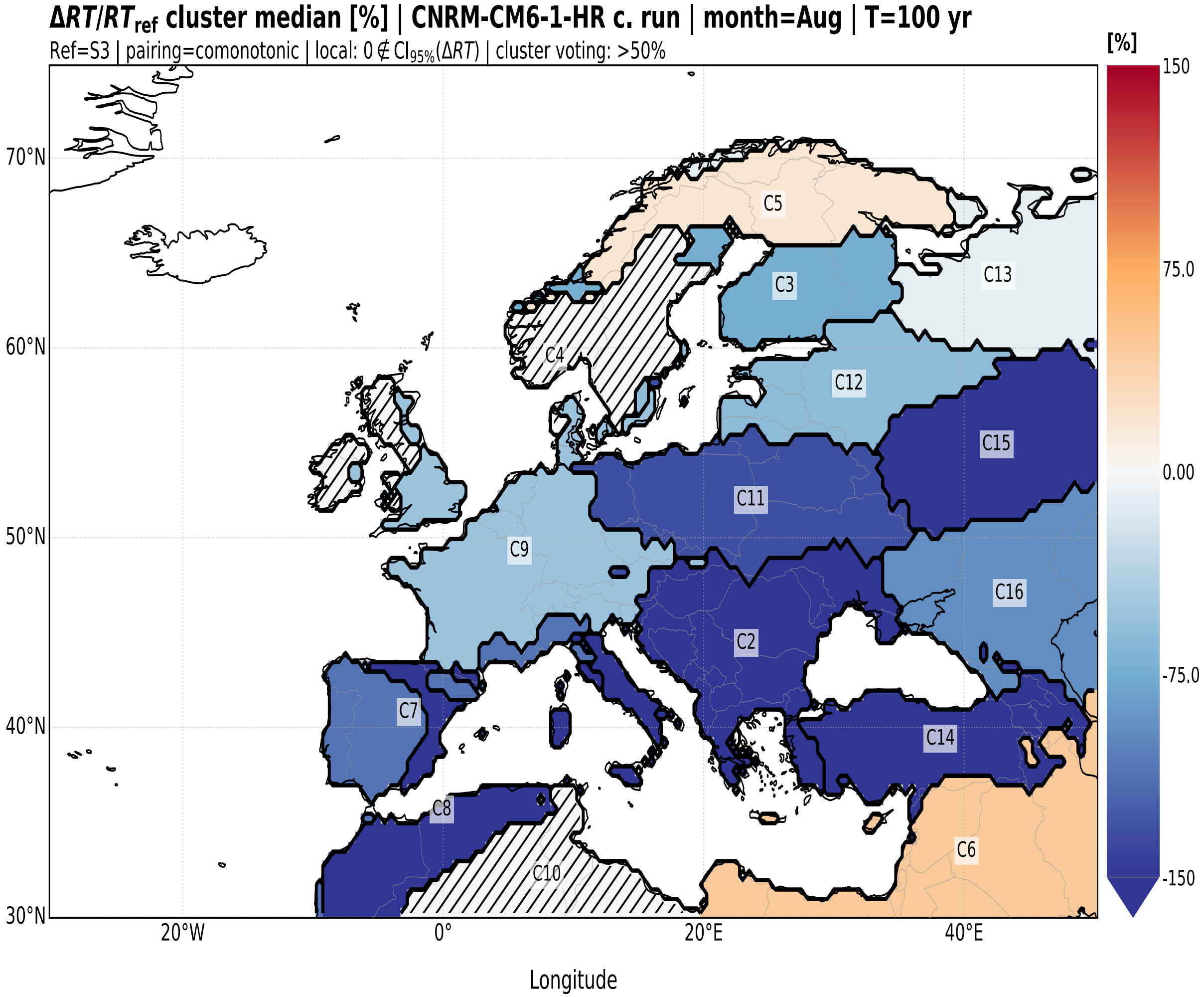
Independent



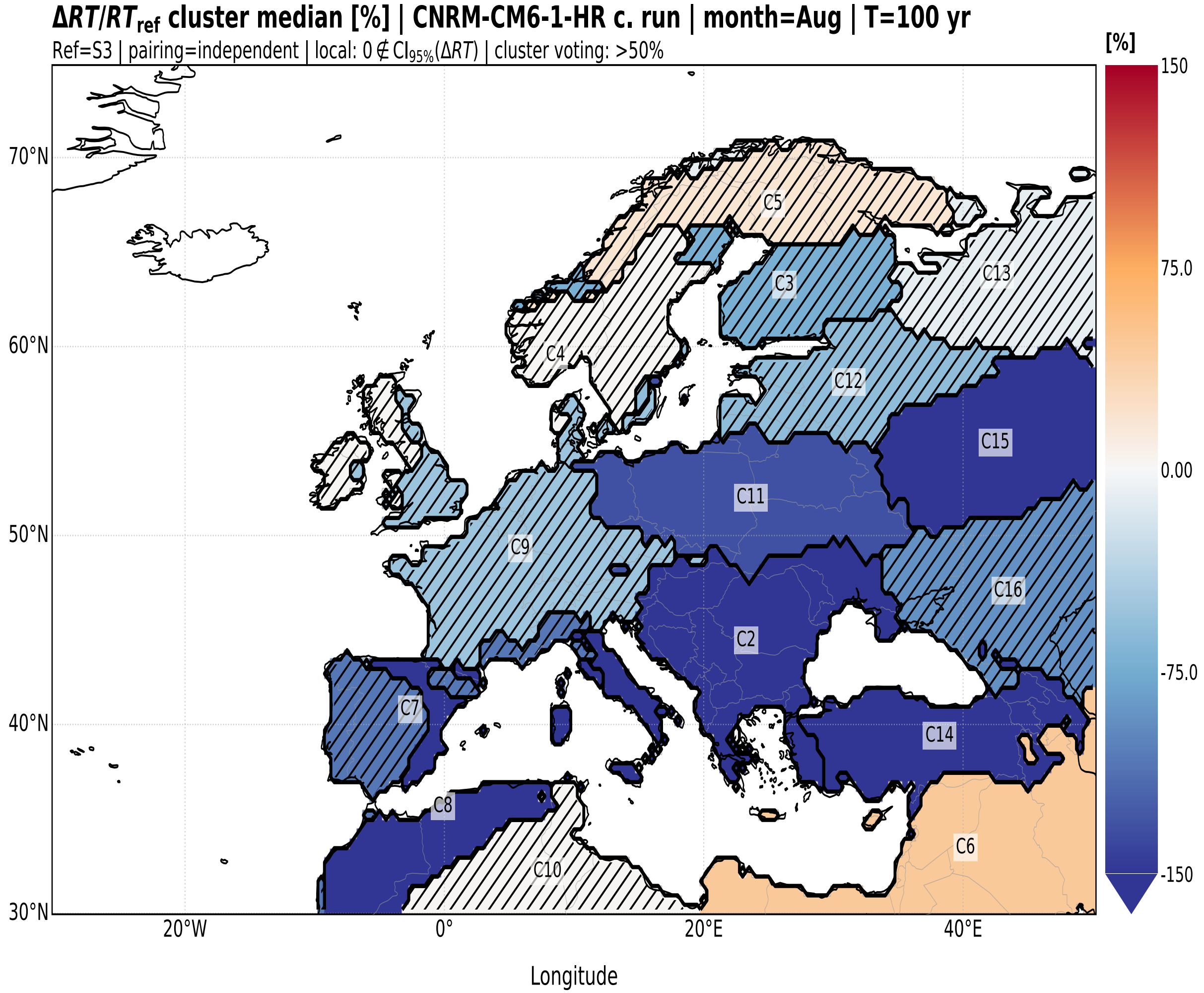
Dataset: E-OBS linear detrended observations

Spatial patterns of ΔRT : Comonotonic vs Independent

Comonotonic



Independent



Dataset: CNRM-CM6-1-HR control run (stationary) simulation