# Clustering and studying evolution of rainfall extremes in France


| | |
|---|---|
| Olivier Mestre | DCSC & CNRM, Météo-France |
| Philippe Naveau | LSCE & IPSL |
| Paola Marson | DCSC, Météo-France |

# Motivation

- **Is there an evolution of rainfall extremes (*i.e.* 10 years return level for example) over time ?**

- **A challenging problem**

  - very noisy data
  - short datasets (hourly rainfall recorded since 1992 only in France)
  - Generalized Extreme Values distribution (GEV) parameters estimation requires a lot of data
  - Peak over Threshold (POT) approaches are not adapted to the non-statonnary case

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

# Motivation

- Classic approach based on annual extremes of single station data
- Simple trend models for GEV position and scale parameters
- Proposed alternative
  - **sampling extremes in « homogeneous » climatic regions**
  - **semi-parametric approach using the R-largest values distribution within the defined regions**
- Data : 171 daily rainfall series starting in 1950 and 196 hourly rainfall series starting in 1992 (checking the homogeneity status)

RÉPUBLIQUE FRANÇAISE
*Liberté Égalité Fraternité*

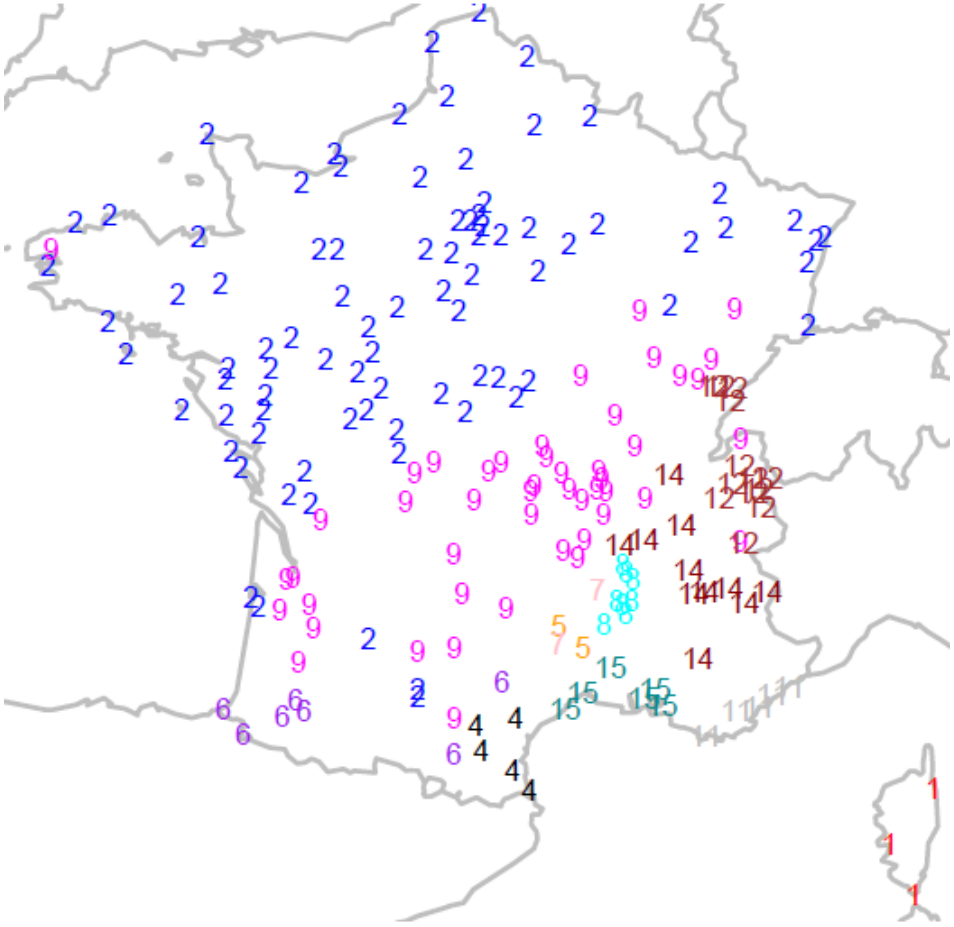
METEO FRANCE

# Defining homogeneous climatic regions

# Clustering

# Homogeneous climatic regions

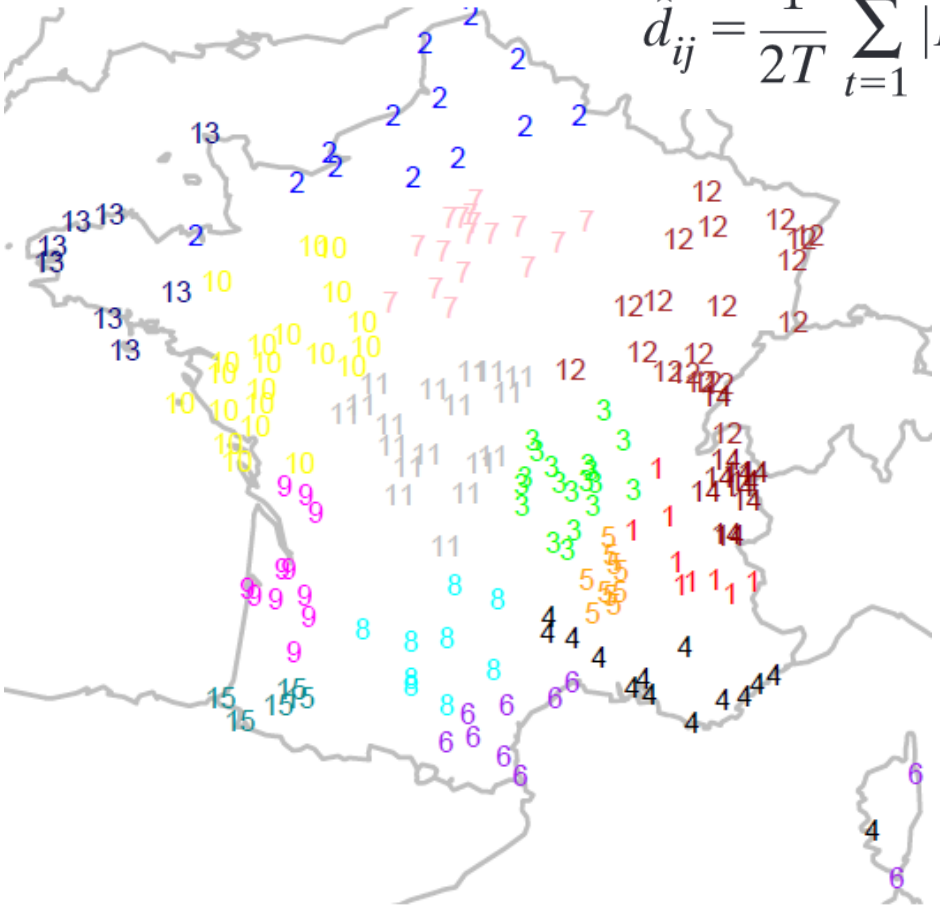
- **Clustering monthly maxima of daily rainfall**
  - classic approach : K-means



# Homogeneous climatic zones

- **Clustering monthly maxima of daily rainfall**
  - proposed approach : PAM + F-Madogram distance

$$\hat{d}_{ij} = \frac{1}{2T} \sum_{t=1}^{T} |\hat{F}_i(M_i^{(t)}) - \hat{F}_j(M_j^{(t)})|$$



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

# Extremes and evolution

# R-largest values distribution

- Closely related to the GEV distribution
- **GEV is the special case when R=1**
- Same parameters as the GEV
- Attraction domains remain the same (in theory, same shape parameter)

- Annual R-largest values retrieved from station data **within the same cluster**
- !!! R-largest values should be independant (not taken from the same rainfall episode)

# Vector Generalized Additive Models

- **WHY?**

Joint variations of non-independent parameters

Model location (μ), scale (σ) parameters of the R-largest values distribution as ***SMOOTH*** functions of time (spline). For computation reasons, the shape parametrer (ξ) remains constant in our study.

$$\mu_t = \mu_o + S_1(t) \qquad \sigma_t = \sigma_o + S_2(t) \qquad \xi = \xi_o$$

**Data driven approach rather than model driven approach**

# Vector Additive Modelling

- **HOW?**

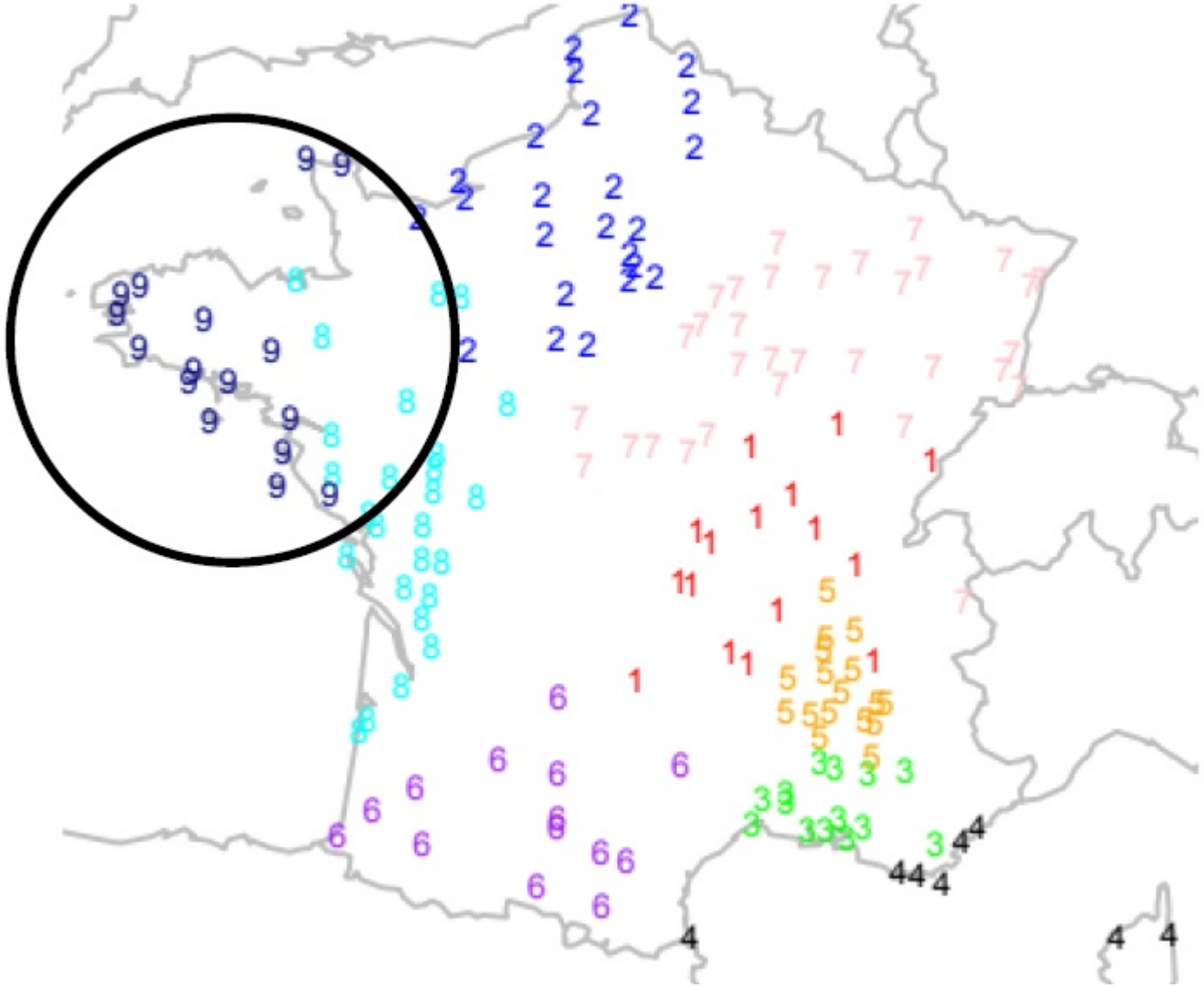
  Vector Generalized Additive Modelling technique (Yee & Wild, 1996) provides flexible smoothing via modified *vector backfitting algorithm* implementation and *vector spline*: VGAM package in R (Yee, 2015).

- **WARNINGS**
  - few predictors should be used
  - only pointwise standard error estimates are provided, not the full covariance matrix: full inferences should be obtained using linear techniques
  - Convergence may be hard to achieve – use link functions : $\log(\sigma)$

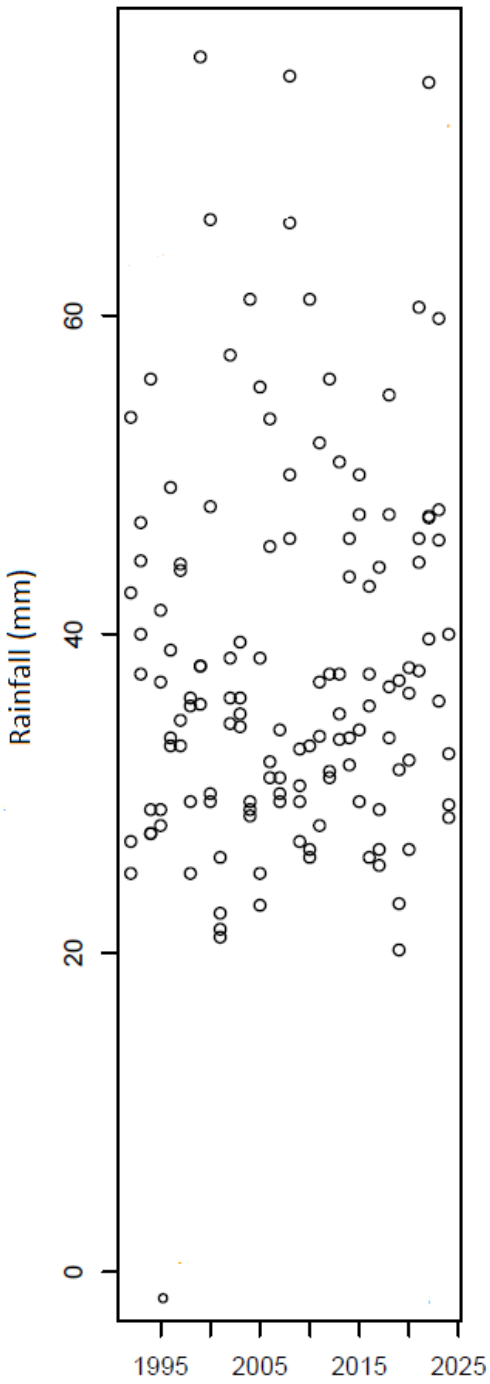
RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

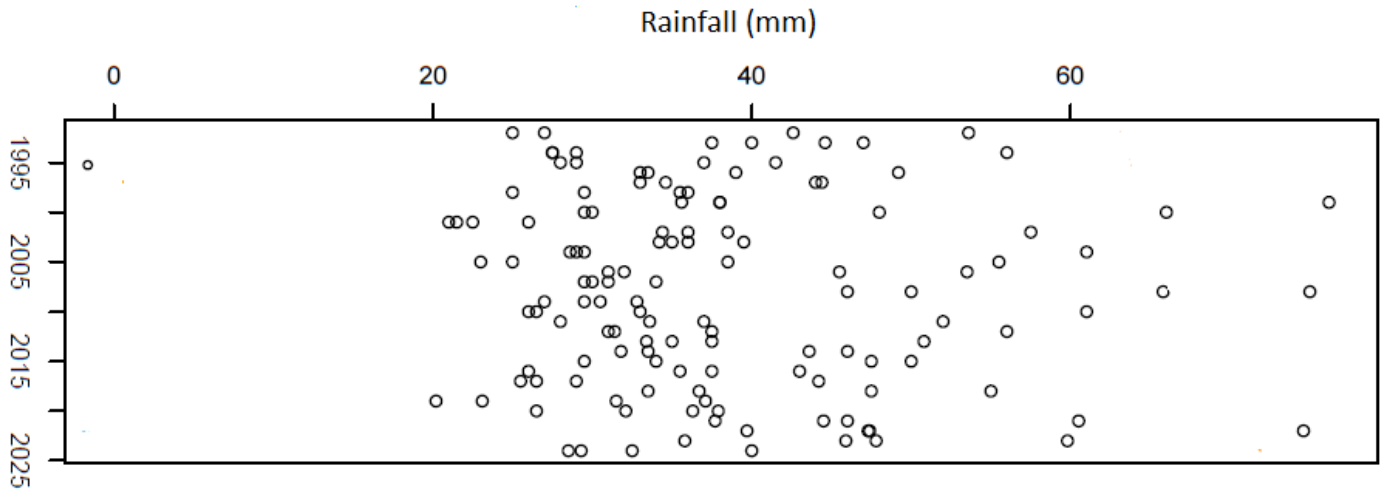
METEO FRANCE

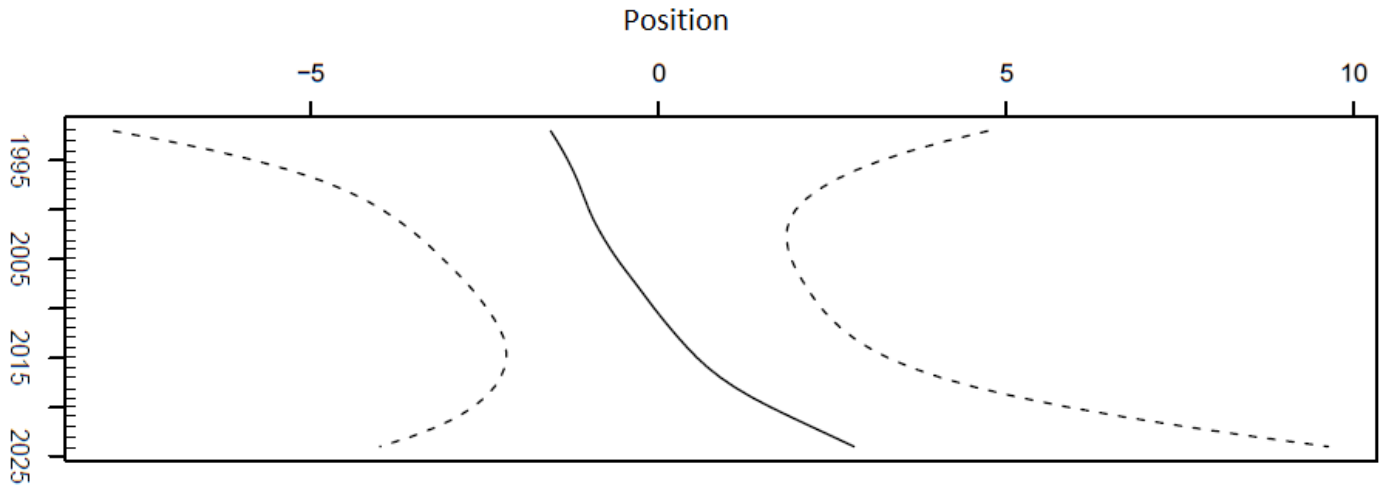
# Bretagne example

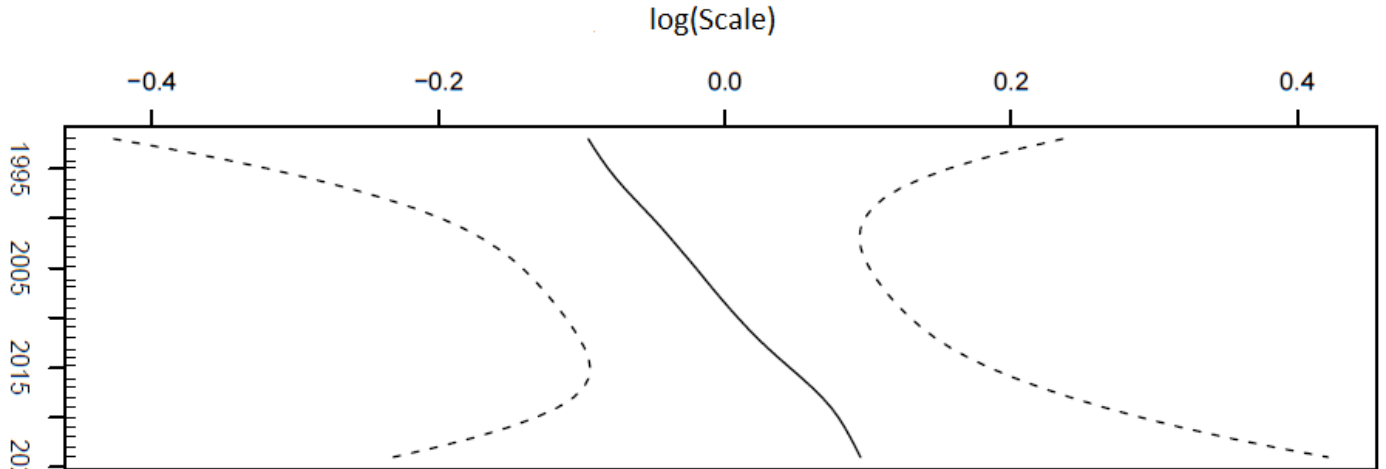


RÉPUBLIQUE FRANÇAIS
Liberté
Égalité
Fraternité

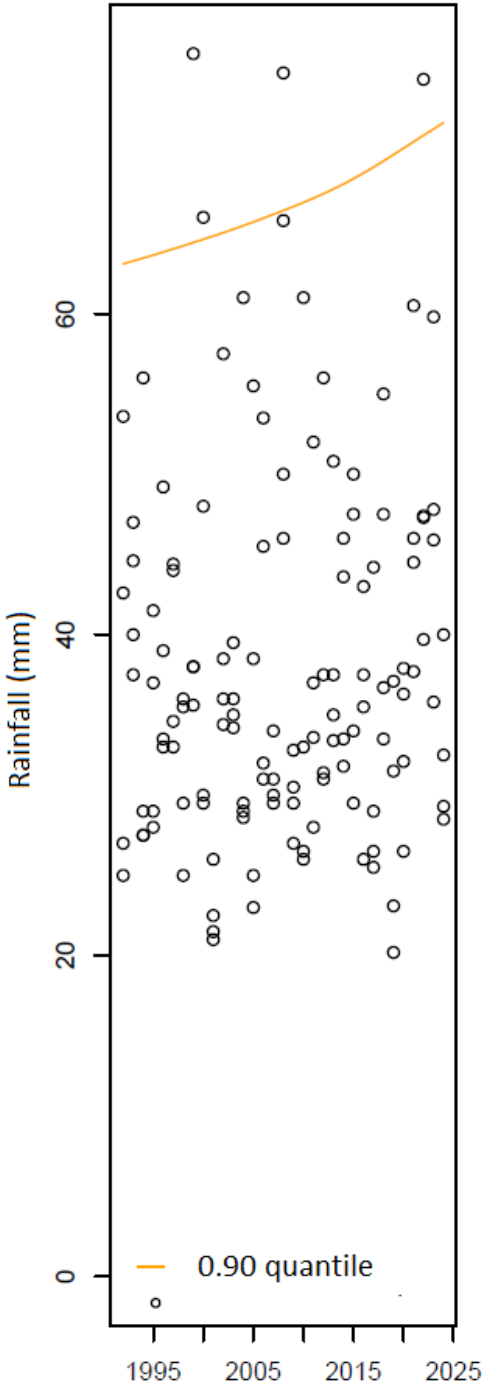


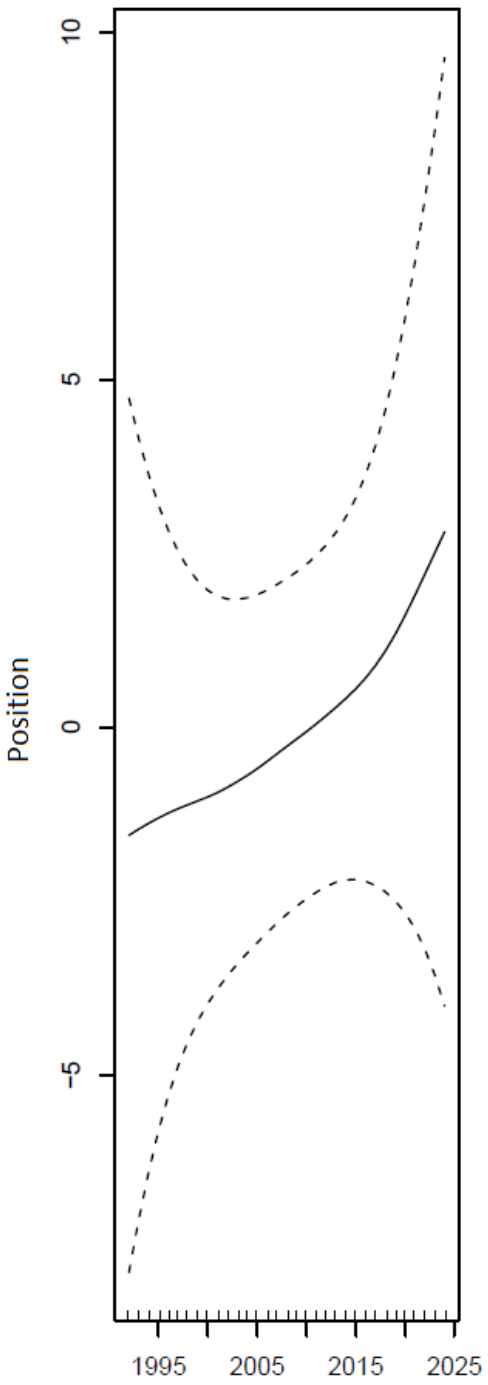


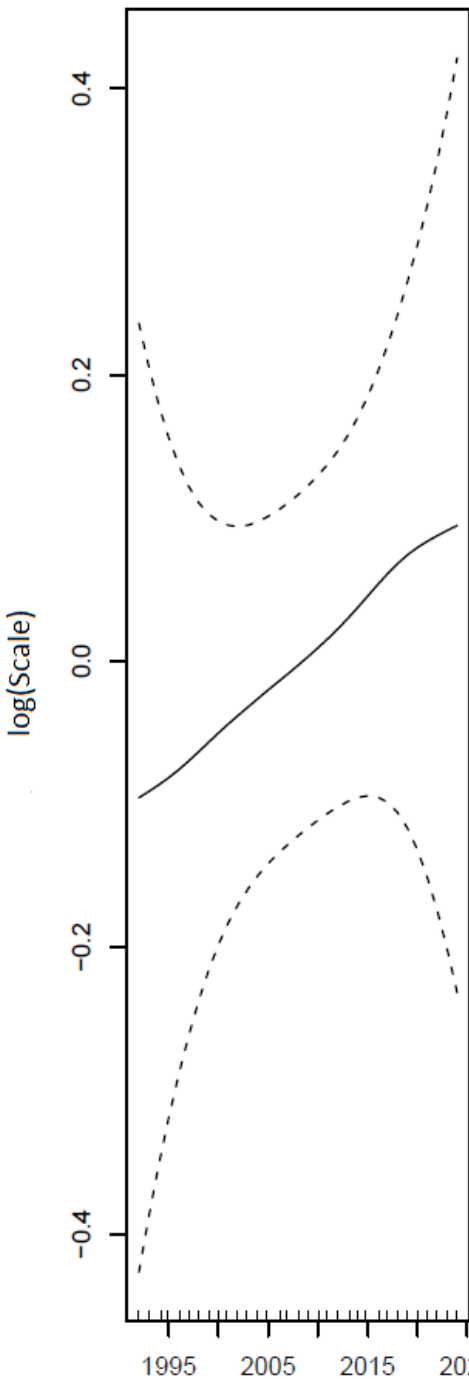



RÉPUBLIQUE
FRANÇAISE
*Liberté*
*Égalité*
*Fraternité*



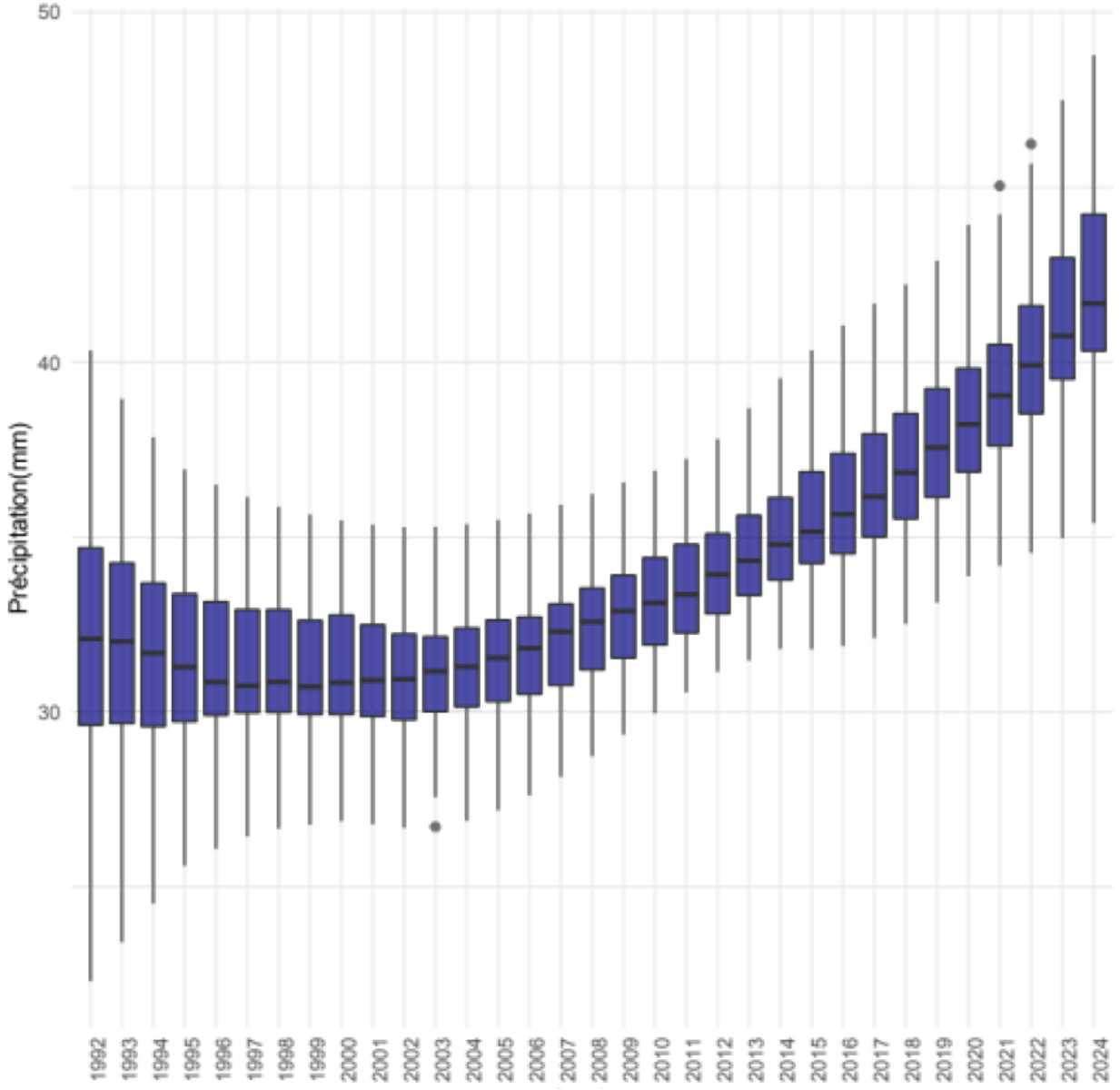



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

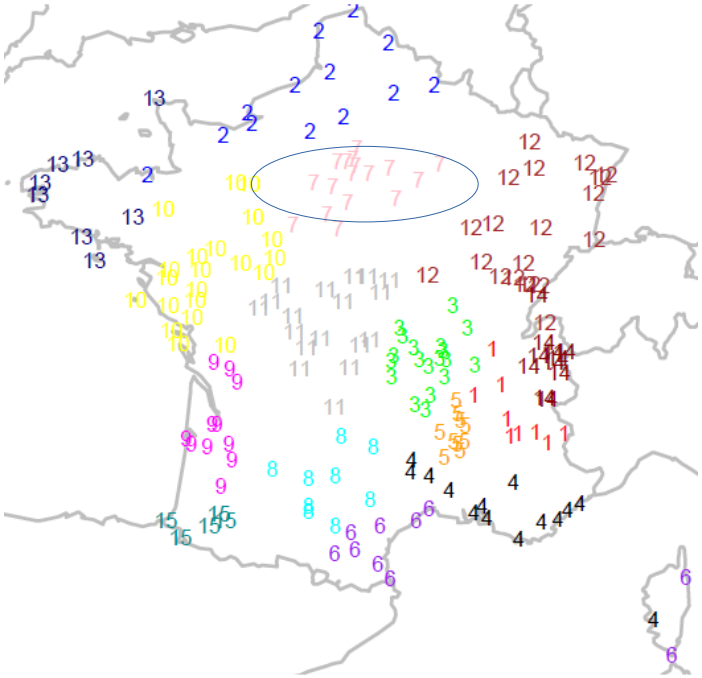
METEO FRANCE

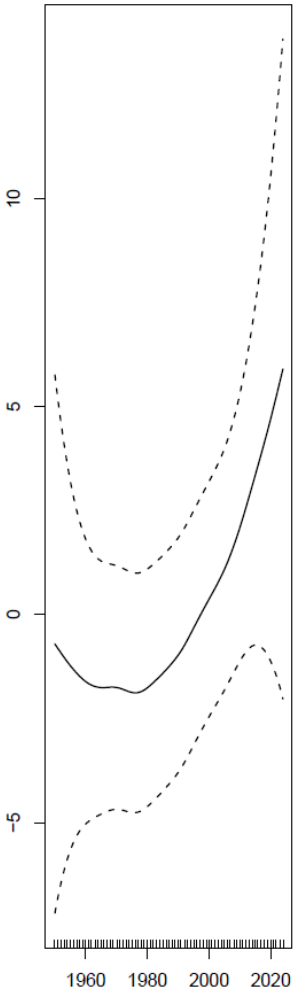
# Bretagne rainfall 10 year return level (Q90)

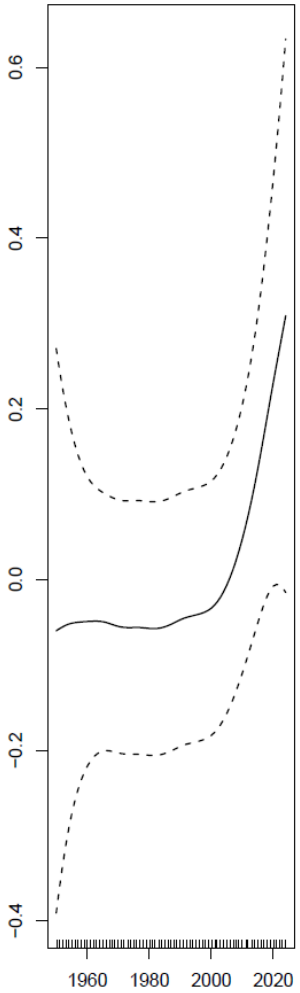


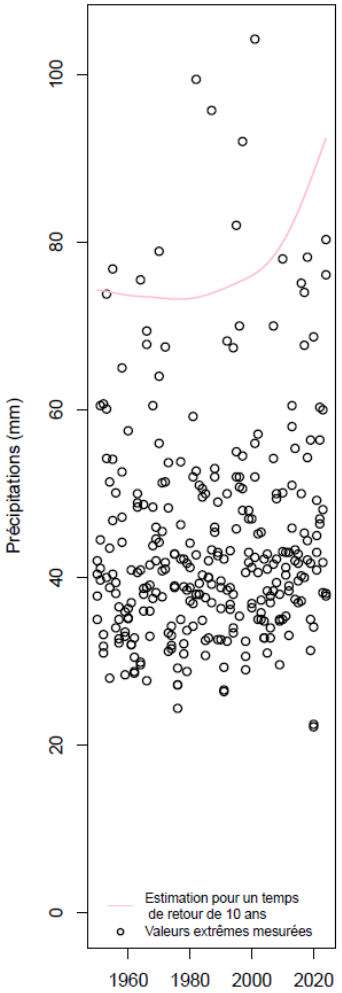
RÉPUBLIQUE FRANÇAISE
*Liberté Égalité Fraternité*

METEO FRANCE





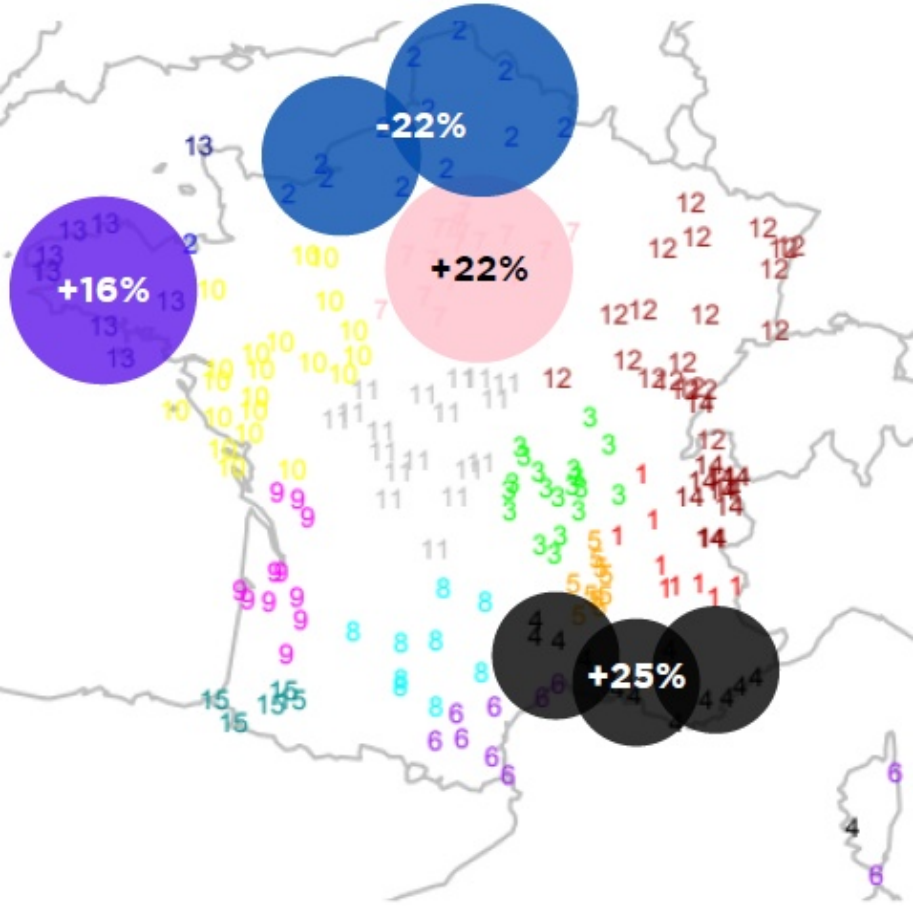



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

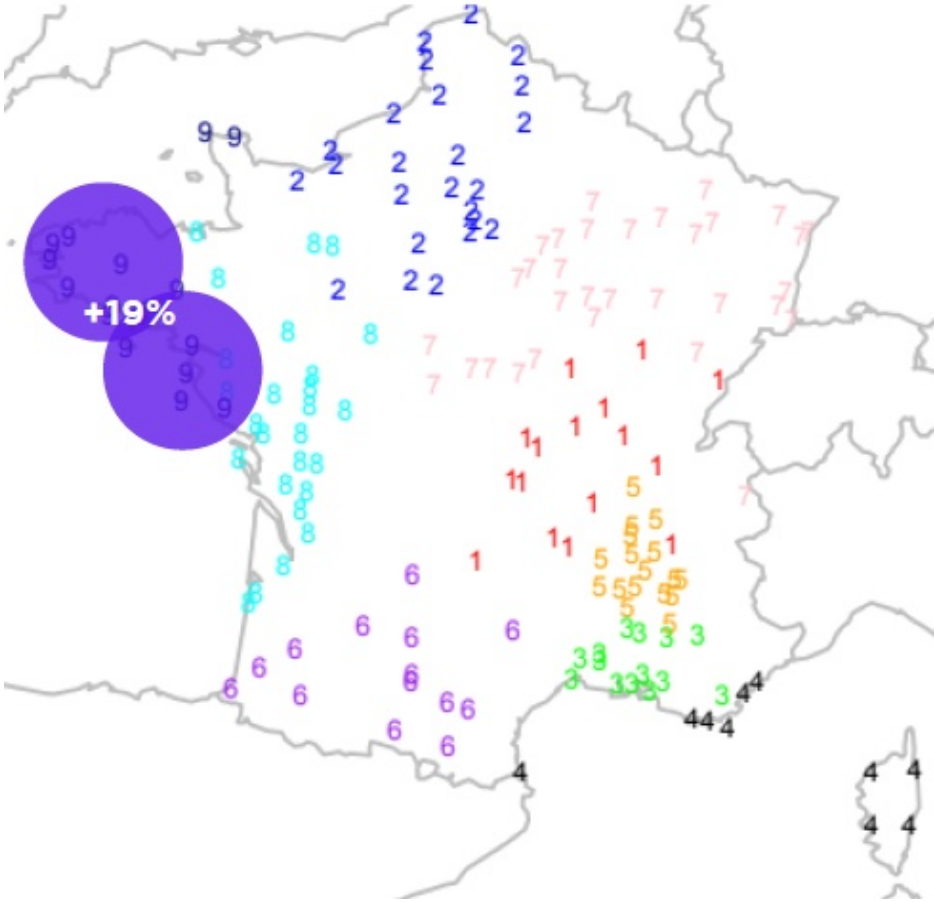
METEO FRANCE

# Preliminary results over France

## Daily rainfall




## Hourly rainfall



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

# CONCLUSION

- Adapted clustering method
- Semi-parametric approach : R-Largest values distribution + VGAM + bootstrap allows put into evidence changes in extremes
- Perspectives
  - put more efforts on the input data
  - clustering : other distances might be relevant
  - choice of R
  - test parametric models suggested by VGAM

# BIBLIOGRAPHY


- Bernard, Naveau, Vrac & Mestre (2013)
  *Clustering of Maxima : Spatial Dependencies among Heavy Rainfall in France*
  J. Climate, Vol. 26, pp. 7929-7937

- Yee & Wild, 1996
  *Vector Generalized Additive Models*
  JRSS series B, Vol. 58, n°3, pp. 481-493

- Yee, 2015. *Vector Generalized Linear and Additive Models: With an Implementation in R*.
  Springer, New York, USA.