Spatially coherent generation of low, moderate and extreme Fire Weather Index in Portugal

Lorenzo Dell'Oro[1], Carlo Gaetan[2], Thomas Opitz[3]

[1]Università di Padova, Italy
[2]Università Ca' Foscari di Venezia, Italy
[3]INRAE, Avignon, France

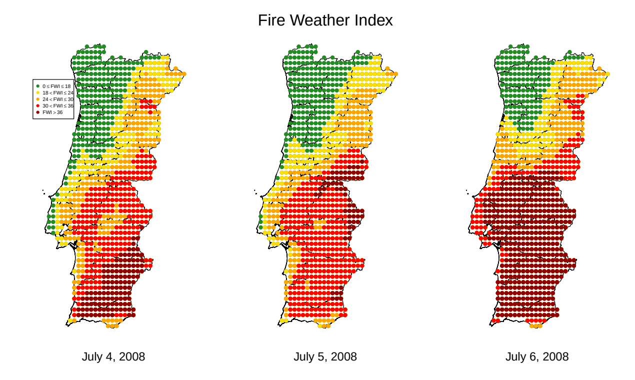
SWGEN 2025

Grenoble, 2-4 December 2025

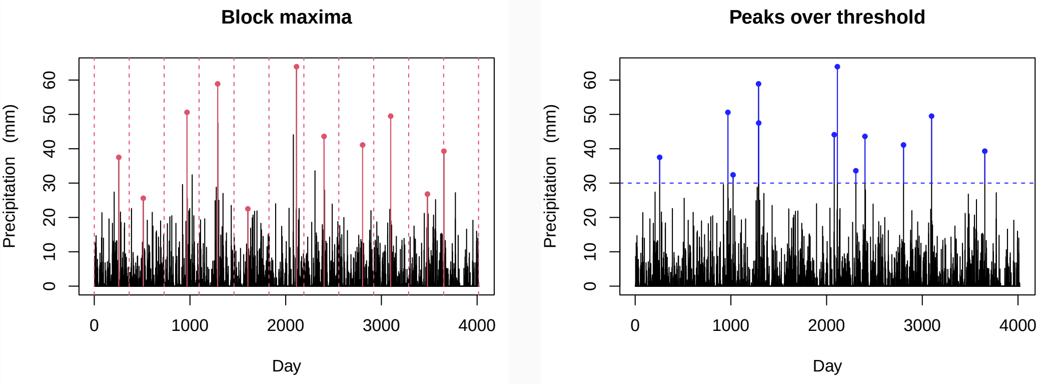
Motivating example

Fire Weather Index (FWI) computed on 947 spatial locations in Portugal

- Spatial dependence
- Interest in extreme and non-extreme observations

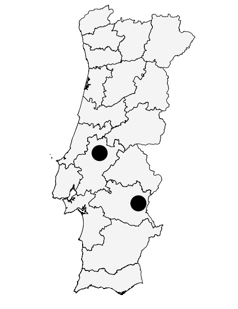


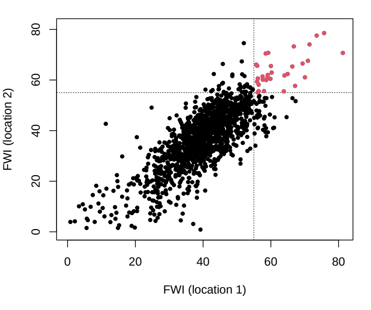
Statistical modelling of extreme values



- Models for block maxima (annual maxima)
- Models for tail behaviour by fixing a high threshold (exceedances)

Modelling of spatial extremes





- Study the risk of joint occurrence of extreme events
  $\Longrightarrow$ Spatial aggregation of risk
- Applied interest in climatic and environmental sciences
- Notion of extremal dependence $\Longrightarrow$ coefficient χ

Asymptotic Dependence/Independence

Let $Y_1 \sim F$ and $Y_2 \sim F$ be random variables with support $(0, +\infty)$. Defining

$$\chi(y) = \Pr\left(Y_1 > y \mid Y_2 > y\right),$$

bivariate **tail dependence coefficient**:

$$\chi = \lim_{y \to \infty} \chi(y).$$

- $0 \leq \chi \leq 1$
- If $\chi > 0$, Y_1 and Y_2 are **Asymptotically Dependent (AD)**
- If $\chi = 0$, Y_1 and Y_2 are **Asymptotically Independent (AI)**

Gaussian processes

$$\boldsymbol{W} = \{W(s)\} \sim \mathcal{GP}$$

- Parametrized by a correlation function $\rho(\cdot)$
- Easy to simulate from

Limitations:

- Symmetry, light tails
- Weak tail dependence: always AI unless $\rho(s_1, s_2) = 1$

Gaussian location and scale mixtures

If $W(s)$ is a Gaussian process:

Gaussian scale mixtures (Huser et al., 2017; Engelke et al., 2019)

$$X(s) = R \times W(s), \quad R \geq 0 \text{ indep. of } W(s)$$

- Can be heavy-tailed, R can induce asymptotic dependence (AD)

Gaussian location mixtures (Krupskii et al., 2018)

$$X(s) = S + W(s), \quad S \text{ indep. of } W(s)$$

- Can also be asymmetric

Gaussian location-scale mixtures

Our general framework:

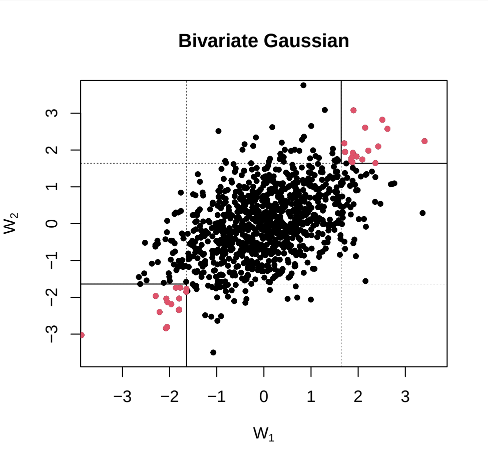
$$X(s) = S + R \times W(s), \quad R \geq 0,\ S \text{ indep. of } W(s)$$

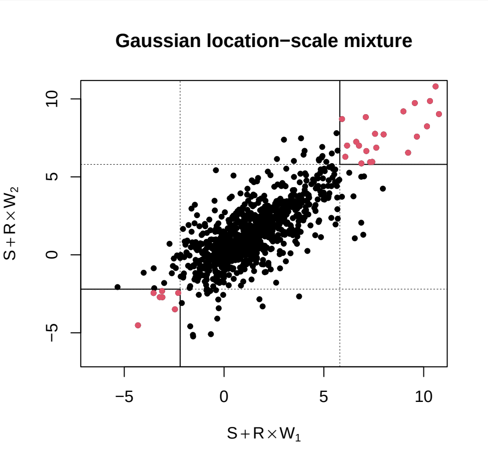
- More flexibility in model construction
- S can induce asymmetry in the tails
- Can be heavy-tailed and AD
- **Example**: $S \sim$ Asymmetric Laplace(λ_1, λ_2), $R^2 \sim$ Exp$(1/2)$
  Asymmetric, tail dependence controlled by (λ_1, λ_2): both tails can be AI or AD

Gaussian location-scale mixture (example)

$W(s)$ is Gaussian $\implies$ AI in both tails

$X(s) = S + R \times W(s)$ is AD in the upper tail, AI in the lower tail





New method for inference

Inference

$$X(s) = S + R \times W(s)$$

In few, well-known cases, **closed-form densities** are available, e.g.:

- $X(s) = R \times W(s)$, with $R = 1/\sqrt{G}$, $G \sim \text{Gamma}(\nu/2, \nu/2)$
  $\implies (X(s_0), \ldots, X(s_m)) \sim$ Multivariate Student's t with ν d.f.
- $X(s) = R \times W(s)$, with $R^2 \sim \text{Exp}(1/2)$
  $\implies (X(s_0), \ldots, X(s_m)) \sim$ Multivariate Laplace

In general, we don't know the multivariate distribution of $X(s)$

New solutions for inference

$$X(s) = S + R \times W(s)$$

Problem: Latent variables S and R to be integrated out

$\Longrightarrow$ High numerical cost of likelihood evaluations

Two steps solution:

1. Transform the data to estimate the parameters of $W(s)$
2. Estimate the parameters of S and R

Step 1

We apply the following data transformations:

- Location mixtures: $X(s) = S + W(s)$,

$$\begin{aligned} \boldsymbol{Z} &= (X(s_1) - X(s_0), \dots, X(s_m) - X(s_0)) \\ &= (W(s_1) - W(s_0), \dots, W(s_m) - W(s_0)) \end{aligned}$$

- Scale mixtures: $X(s) = R\,W(s)$

$$\boldsymbol{Z} = \left(\frac{X(s_1)}{X(s_0)}, \dots, \frac{X(s_m)}{X(s_0)} \right) = \left(\frac{W(s_1)}{W(s_0)}, \dots, \frac{W(s_m)}{W(s_0)} \right)$$

Step 1

- Location-scale mixtures: $X(s) = S + R\,W(s)$

$$\boldsymbol{Z} = \left(\frac{X(s_2) - X(s_0)}{X(s_1) - X(s_0)}, \ldots, \frac{X(s_m) - X(s_0)}{X(s_1) - X(s_0)}\right)$$
$$= \left(\frac{W(s_2) - W(s_0)}{W(s_1) - W(s_0)}, \ldots, \frac{W(s_m) - W(s_0)}{W(s_1) - W(s_0)}\right)$$

We can get the analytical distribution of $\boldsymbol{Z}$

$\implies$ use it to estimate the parameters of $\rho(\cdot)$ of the Gaussian process $W(s)$

Step 2

We still need to estimate the parameters of S and R.

Let

$$\bar{X} = \frac{1}{m+1}\sum_{j=0}^{m} X(s_j) \quad \text{and} \quad \bar{W} = \frac{1}{m+1}\sum_{j=0}^{m} W(s_j).$$

The distribution of

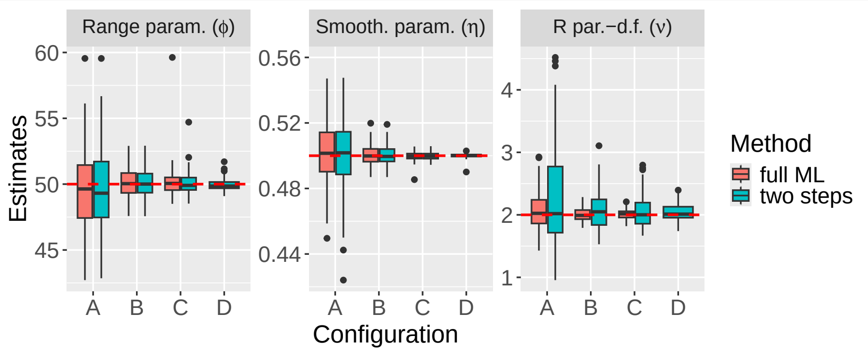
$$\bar{X} = S + R\,\bar{W}$$

depends now only on the parameters of S and R.

We can estimate them by minimizing a goodness-of-fit criterion for the edf of $\bar{X}$.

Simulation study

$X(s) = R\,W(s) \sim$ Multivariate Student's t with ν d.f. (known distribution):
Classical MLE for $X(s)$ vs Two-steps estimate for $W(s)$ and R



A: $m = 50,\ n = 100;$ B: $m = 100,\ n = 500;$
C: $m = 200,\ n = 1000;$ D: $m = 400,\ n = 2000.$

Marginal modelling

- Scale-location mixture $X(s)$ is a **copula model** for the spatial dependence
- Data observed in a **marginal scale** $Y(s)$
- Marginally, $X(s) \sim G$ and $Y(s) \sim F$:

$$X(s) = G^{-1}\{F(Y(s))\}$$

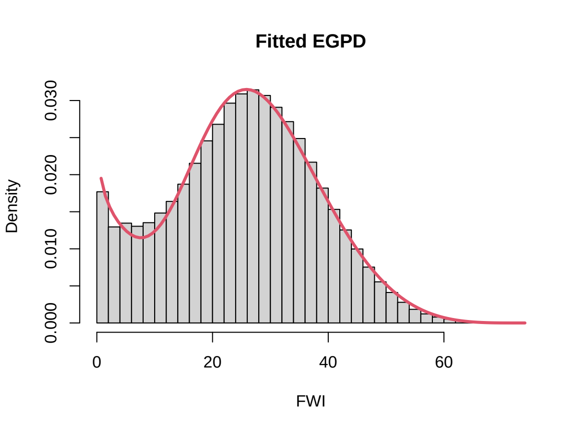
- Flexible F to model both the bulk and the tail of the data

Marginal modelling: EGPD

Extended Generalized Pareto Distribution (EGPD) from Naveau et al. (2016):

$$F(y) = B\left\{H_\xi\left(\frac{y}{\sigma}\right)\right\}$$

- $H_\xi(\cdot)$ is the cdf of the GPD
- $B(u) = p\,u^{k_1} + (1-p)\,u^{k_2}$
- Avoids threshold selection



Conditional simulation

Conditional simulation of Gaussian location-scale mixtures

Algorithm

Goal: simulate $\mathbf{X_2} \mid (\mathbf{X_1} = \mathbf{x_1})$ in a Gaussian location-scale mixture $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2})$

Inputs: Conditioning components $\mathbf{x_1}$; Gaussian variance-covariance matrix Σ

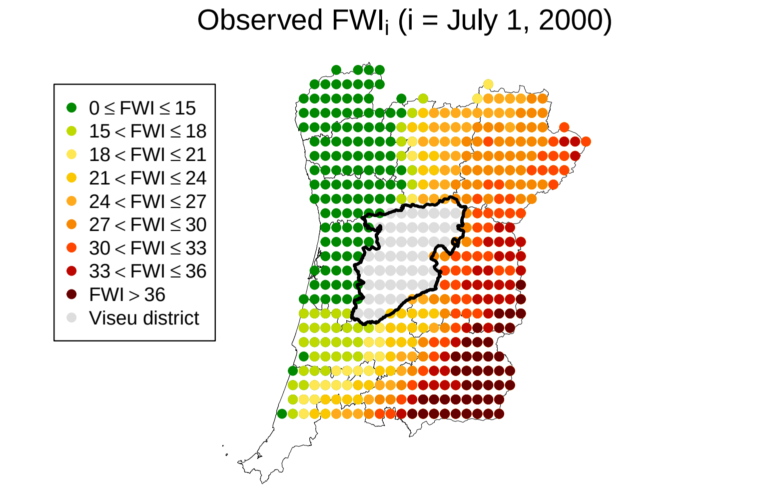
1. Simulate (R, S) from the conditional distribution of $(R, S) \mid \mathbf{X}_1$

$$f_{(R,S)|\mathbf{X}_1=\mathbf{x}_1}(r, s) = \frac{f_{(R,S,\mathbf{X}_1)}(r, s, \mathbf{x}_1)}{f_{\mathbf{X}_1}(\mathbf{x}_1)} = \frac{f_{(R,S,\mathbf{W}_1)}(r, s, (\mathbf{x}_1 - s)/r)r^{-m}}{f_{\mathbf{X}_1}(\mathbf{x}_1)}$$
$$\propto f_R(r) f_S(s) f_{\mathbf{W}_1}((\mathbf{x}_1 - s)/r) r^{-m}$$

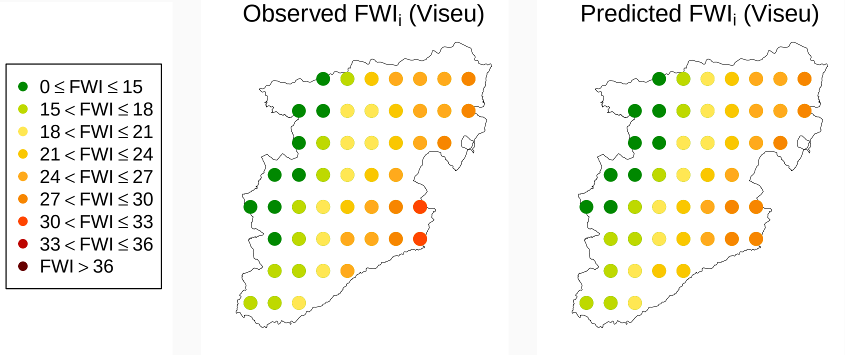
   through a random walk Metropolis-Hastings algorithm;

2. Compute $\mathbf{W}_1 = (\mathbf{X}_1 - S)/R$;
3. Simulate $\mathbf{W}_2$ from $\mathbf{W}_2 \mid \mathbf{W}_1 \sim N(\mu_{2|1}, \Sigma_{2|1})$;
4. Compute $\mathbf{X}_2 = S + R\,\mathbf{W}_2$.

Conditional simulation - example (I)

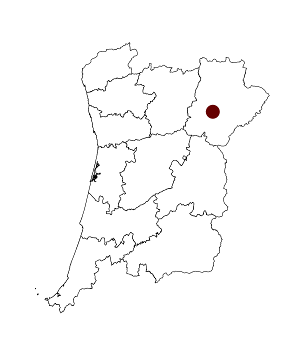


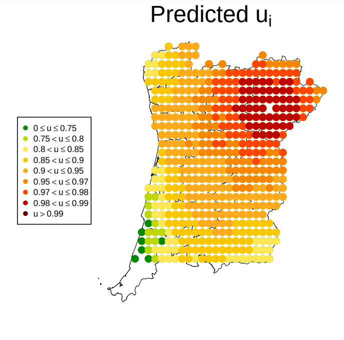
Conditional simulation - example (II)

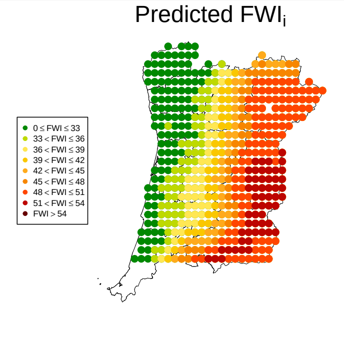


Ongoing work...

From an observed high value ($u = 0.99$), conditional simulation in the other locations







$\Longrightarrow$ use as input in models that simulate wildfire propagation

References

Engelke, S., Opitz, T., & Wadsworth, J. (2019). Extremal dependence of random scale constructions. *Extremes*, 22(4), 623–666.

Huser, R., Opitz, T., & Thibaud, E. (2017). Bridging asymptotic independence and dependence in spatial extremes using gaussian scale mixtures. *Spatial Statistics*, 21, 166–186.

Krupskii, P., Huser, R., & Genton, M. G. (2018). Factor copula models for replicated spatial data. *Journal of the American Statistical Association*, 113(521), 467–479.

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*, 52(4), 2753–2769.

Backup slide: conditional simulation

Plot in the uniform scale $\Longrightarrow$ no marginal non-stationarity effect