Multivariate generator at sub-daily resolution in a semi-arid Tunisian catchment

Farhani N.[1,3], Carreau J.[2], Boulet G.[3], Kassouk Z.[1] et al.

[1]LR17AGR01 / GREEN-TEAM (Institut National Agronomique de Tunis / Univ. Carthage), Tunisie - farhani.nesrine@gmail.com

[2]HydroSciences Montpellier (Univ. Montpellier/CNRS/IRD), France

[3]Centre d'Etudes Spatiales de la Biosphère (CNRS/UPS/CNES/IRD), France

LMI NAILA

gestio**N** des ressources en e**A**u dans les m**IL**ieux rur**A**ux tunisiens

water resources management in Tunisian rural areas

Bilateral research laboratory Tunisia / France

Co-directed : Insaf MEKKI, INRGREF Tunis & Frédéric JACOB, IRD / LISAH Montpellier

Co-funded : MESRS / IRESA / IRD

Study areas

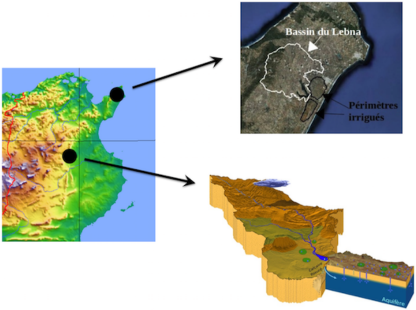
Two watersheds on the Tunisian mountain ridge

Lebna

- 200 km^2
- sub-humid climate

Merguellil

- 1200 km^2 / 400 km^2
- semi-arid climate



Water resources management

Evapotranspiration : key variable to monitor

1. actual water use
2. agricultural drought (hydric stress on plants)

Estimation :

- double source model based on energy balance equations
- relies on remote sensing data and observations

Forcing scenarios

Hydro-meteorological variables measured at sub-daily resolution :

Tair, Hr, Rg, Wave, Pa & Precip

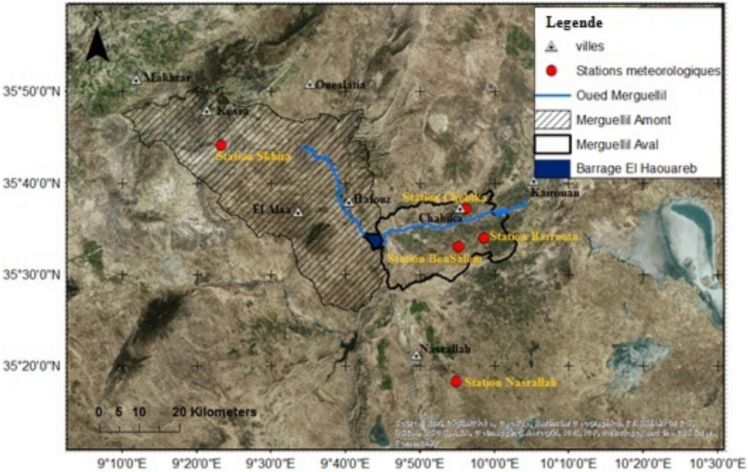
Goal : to provide spatio-temporal scenarios to extend observations in space and in time

Main issues : strong spatio-temporal variability, non-gaussianity, scarcely gauged networks

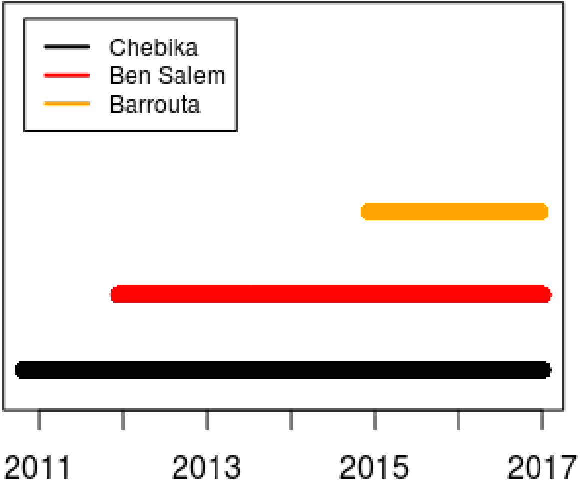
Project : Approches stochastiques et seMi-pAramétriques combinées à la télédétectioN pour l'étude du stress hyDriquE

Merguellil watershed

semi-arid climate : 200 - 650 mm / year   30 min temporal resolution



Three stations in the plains : geographical and meteorological proximity



➜ 1st step

Imputation : to fill in missing values at each of the 3 stations
Virtual station : to create a single representative station in the plains

Stochastic GLM : adaptation of RGlimClim from R. Chandler

→ 1st step

Imputation : to fill in missing values at each of the 3 stations
Virtual station : to create a single representative station in the plains

Stochastic GLM : adaptation of RGlimClim from R. Chandler

→ 2nd step

Temporal extension based on bias correction applied to Era-Interim's reanalyses
Spatial extension based on analogs applied to WRF's simulations

Could be compared with stochastic GLM approach

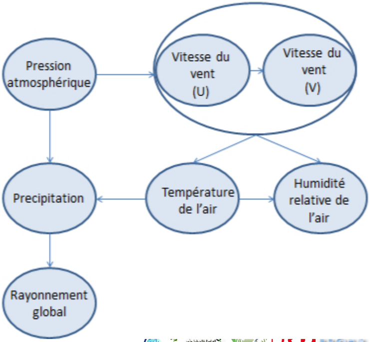
Stochastic GLM

- Inter-variable dependency
- GLM families / distributions
- Temporal and spatial effects - covariates
- Model selection
- Model validation
- *Spatial dependence*

Inter-variable dependency

Product rule : Y_i hydro-meteorological variable + *transformation*

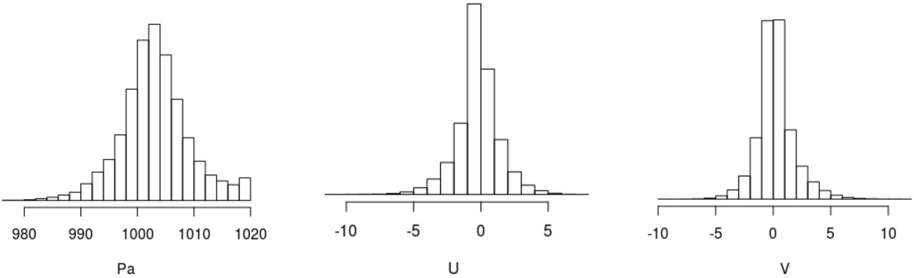
$$\mathbb{P}(Y_1, Y_2, \ldots, Y_p) = \mathbb{P}(Y_1) \prod_{i=2}^{p} \mathbb{P}(Y_i | Y_{i-1}, \ldots, Y_1)$$



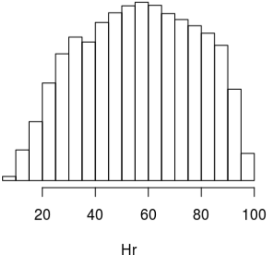
GLM families

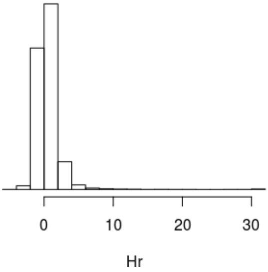
Normal distribution

$$Y|X = x \sim \mathcal{N}(\beta x, \sigma^2) \Longleftrightarrow Y = \beta X + \epsilon \quad \epsilon \sim \mathcal{N}(0, \sigma^2)$$



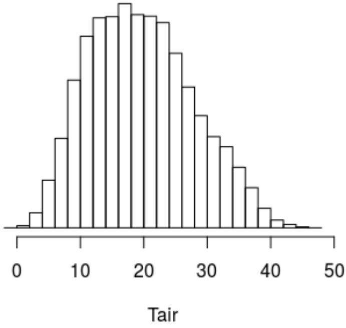
With a transformation $Y = \tan(\pi \widetilde{Y} - 0.5)$





Heteroscedastic Normal distribution

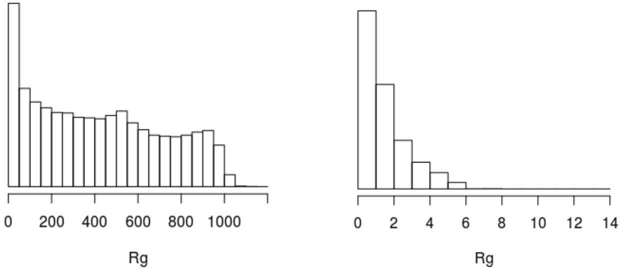
$$Y|X = x \sim \mathcal{N}(\beta x, \sigma(x)^2) \quad \sigma(x) = \exp(\eta x)$$



Gamma distribution

$$Y|X = x \sim Gamma(\mu(x), \nu) \quad \mu(x) = \exp(\eta x)$$

$$Y = \max(\log(\widetilde{Y})) - \log(\widetilde{Y})$$



Temporal effects

Seasonal effects

Covariates :

$$\cos\left(\frac{2\pi d}{k}\right), \quad \sin\left(\frac{2\pi d}{k}\right)$$

with d the day of the year and $k \in \{183, 365.25, \ldots\}$ the period

Diurnal effects

Covariates :

$$\cos\left(\frac{2\pi h}{k}\right), \quad \sin\left(\frac{2\pi h}{k}\right)$$

with h the hour of the day and $k \in \{12, 24, \ldots\}$ the period

Spatial effects

Covariates :

- geographical coordinates
- *landscape variables*

Memory effects

Lags of

- the variable itself Y_{t-k}
- a moving average

$$\frac{1}{m}\sum_{i=1}^{m} Y_{t-k-i+1}$$

- a spatial average

$$\frac{1}{s}\sum_{j=1}^{s} Y_{t-k}^{(j)}$$

- a moving spatial average

Model selection

With AIC, BIC and p-values

Number of selected memory effects

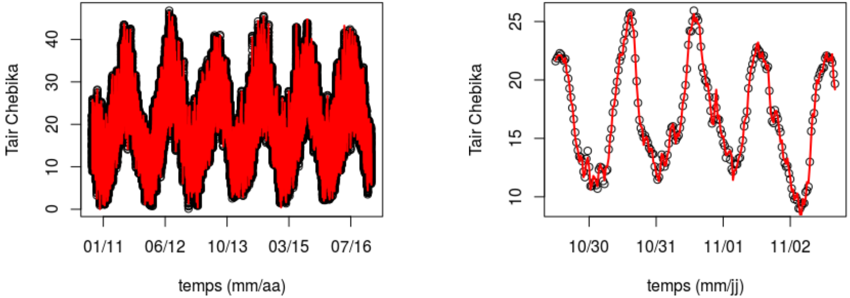
	itself	MA	SA	MSA
Pa	30	15	1	15
U	6	5	1	5
V	6	5	1	5
Tair	5	5	1	5
Hr	6	4	1	4
Rg	12	5	1	5

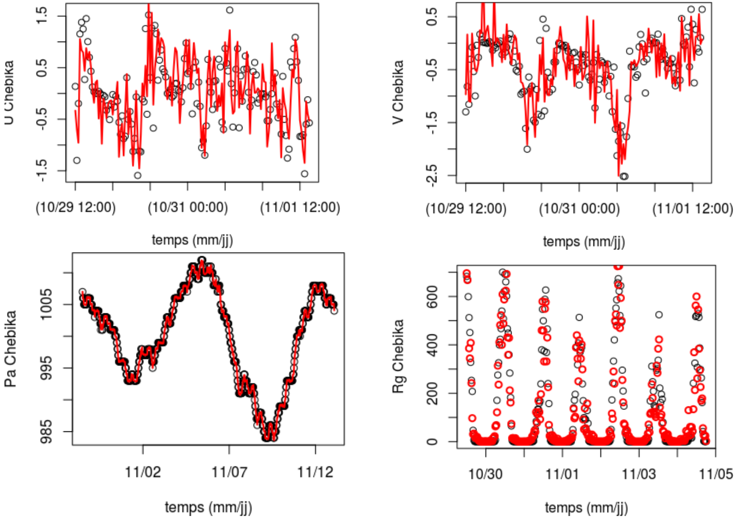
Model validation

Comparison of unconditional simulations from fitted models and observations

- general : temporal plots
- intensity : qq-plots
- dependence : pairwise scatter plots and Kendall's τ

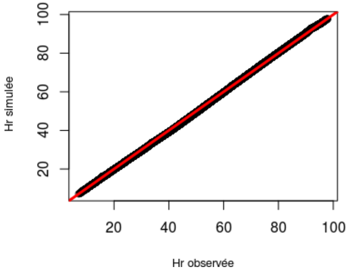
Air temperature at Chebika station

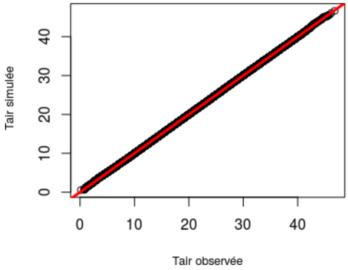


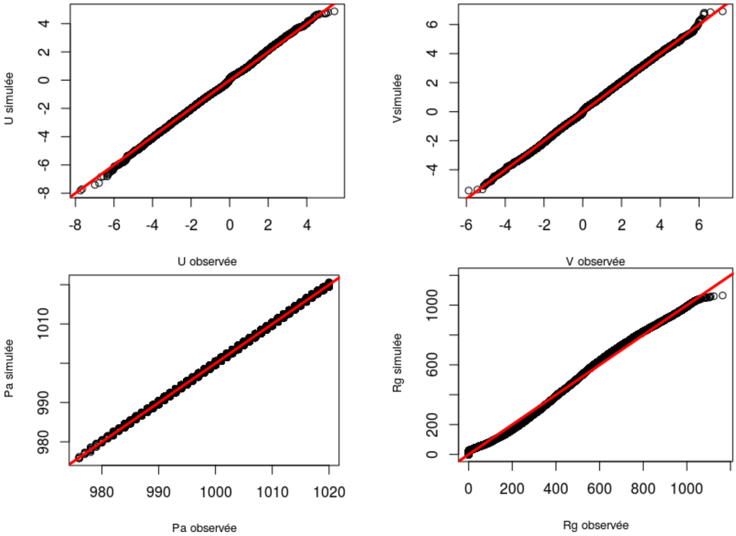

U Chebika
1.5
0.5
-0.5
-1.5
(10/29 12:00)
(10/31 00:00)
(11/01 12:00)
temps (mm/jj)
V Chebika
0.5
-0.5
-1.5
-2.5
(10/29 12:00)
(10/31 00:00)
(11/01 12:00)
temps (mm/jj)
Pa Chebika
1005
995
985
11/02
11/07
11/12
temps (mm/jj)
Rg Chebika
600
400
200
0
10/30
11/01
11/03
11/05
temps (mm/jj)

Intensity assessment

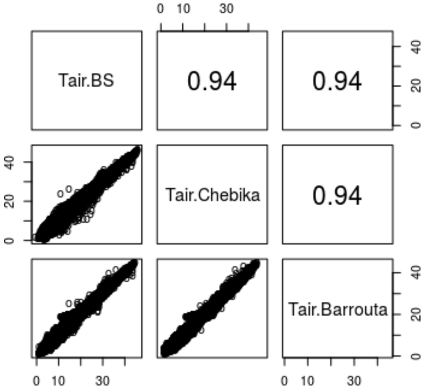
Chebika station

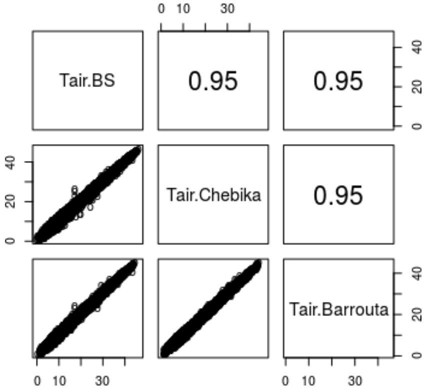


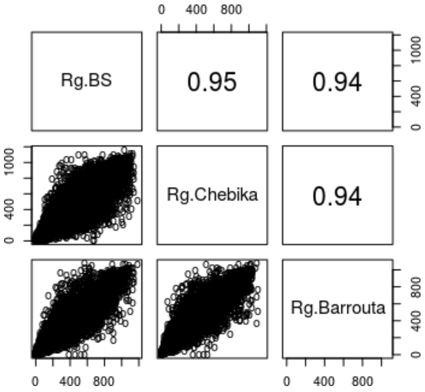


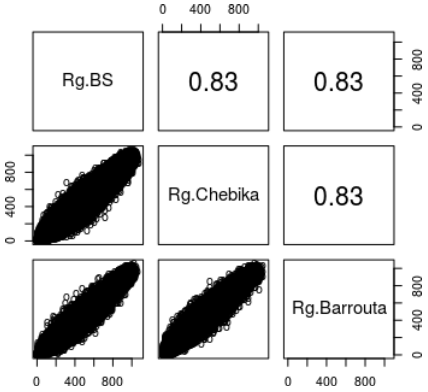

U simulée
4 2 0 -2 -4 -6 -8
-8 -6 -4 -2 0 2 4
U observée
Vsimulée
6 4 2 0 -2 -4
-6 -4 -2 0 2 4 6
V observée
Pa simulée
1010 990 980
980 990 1000 1010 1020
Pa observée
Rg simulée
1000 600 200 0
0 200 400 600 800 1000
Rg observée

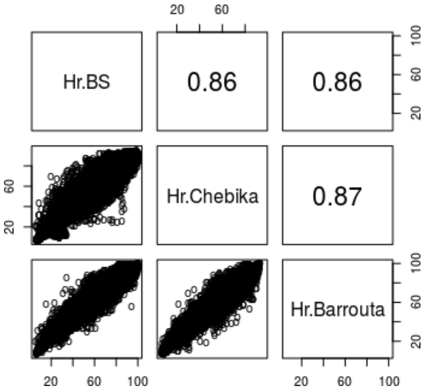
Inter-sites dependence

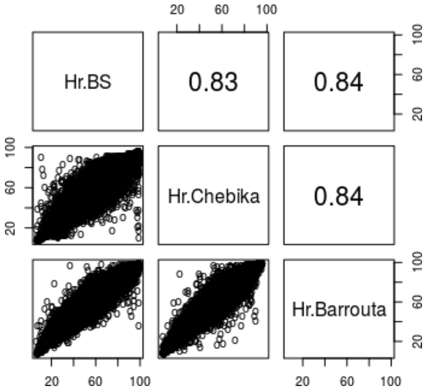




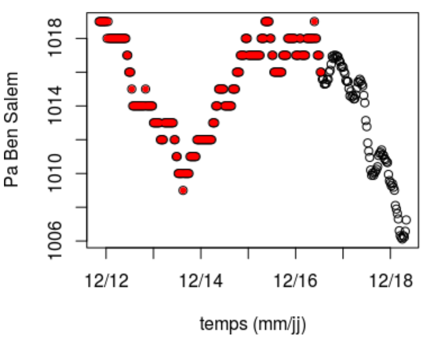

Rg.BS
0.95
0.94
Rg.Chebika
0.94
Rg.Barrouta

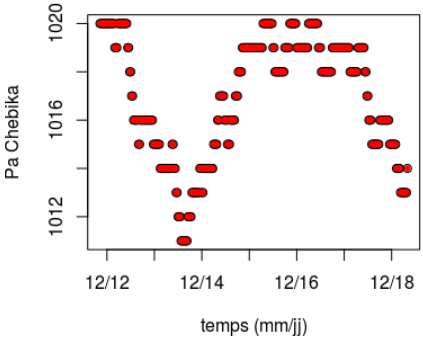

Rg.BS
0.83
0.83
Rg.Chebika
0.83
Rg.Barrouta

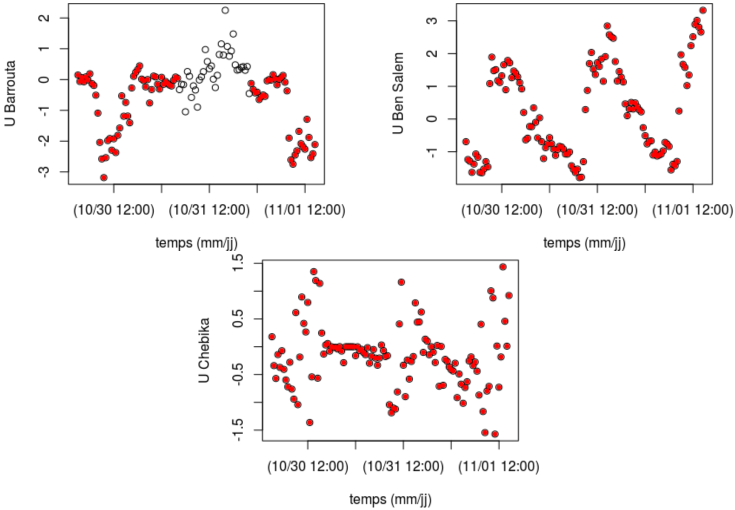

Hr.BS
0.86
0.86
Hr.Chebika
0.87
Hr.Barrouta
20
60
100


Hr.BS
0.83
0.84
Hr.Chebika
0.84
Hr.Barrouta
20
60
100

Conditional simulation : imputation






U Barrouta
temps (mm/jj)
(10/30 12:00)
(10/31 12:00)
(11/01 12:00)
U Ben Salem
temps (mm/jj)
U Chebika
temps (mm/jj)

Summary and outlook

A careful look at model selection, in particular for memory effects

More validation is needed regarding

- inter-variable dependencies
- temporal dependencies

Summary and outlook

A careful look at model selection, in particular for memory effects

More validation is needed regarding

- inter-variable dependencies
- temporal dependencies

➜ Need to introduce spatial dependencies

Summary and outlook

A careful look at model selection, in particular for memory effects

More validation is needed regarding

- inter-variable dependencies
- temporal dependencies

➜ Need to introduce spatial dependencies

➜ Spatial interpolation thanks to geographical covariates

Summary and outlook

A careful look at model selection, in particular for memory effects

More validation is needed regarding

- inter-variable dependencies
- temporal dependencies

➜ Need to introduce spatial dependencies

➜ Spatial interpolation thanks to geographical covariates

➜ Validation in terms of estimation of evapotranspiration and stress index

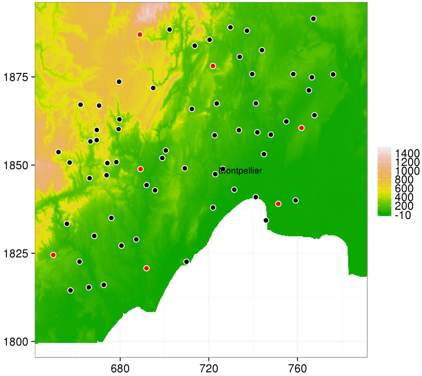
ABC for spatial max-stable model with potential asymmetry and anisotropy

Carreau J.[1], Toulemonde G.[2]

[1]HydroSciences Montpellier (Univ. Montpellier/CNRS/IRD), France - julie.carreau@ird.fr

[2]Institut Montpellierain Alexander Grothendiek, (Univ. Montpellier/CNRS) & INRIA / LEMON, France - gwladys.toulemonde@umontpellier.fr

Daily precipitation data



60 stations for training (black)

7 stations for validation (red)

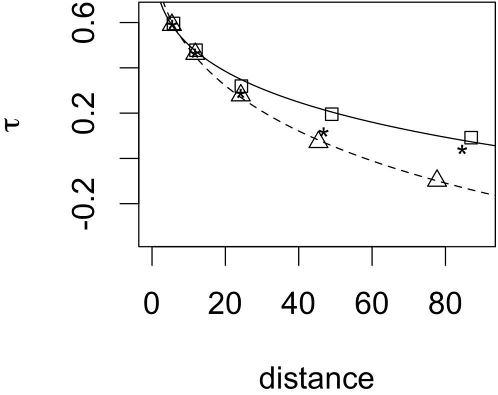
1958 - 2014 (57 years)

Our goal

Model the joint distribution of annual maxima at d sites $Y_1, \ldots, Y_d$

Anisotropy in the strength of the dependence

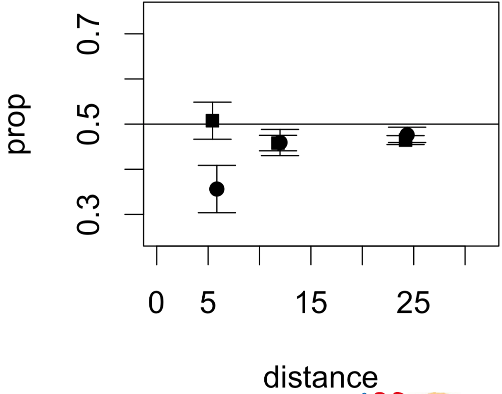
Correlograms of Kendall's τ for classes of distances
➜ along two directions 0 and $\pi/2$



$$R(h; \beta, \alpha) = 2 \exp\left\{-\left(\frac{h}{\beta}\right)^{\alpha}\right\} - 1$$

Asymmetry in the dependence

For a pair of stations : proportion of values in $[0,1]^2$ under the diagonal
Asymmogram : Proportions for classes of distances
➜ along two directions 0 and $\pi/2$



Extra-parametrized Gumbel (XGumbel)

The Gumbel copula in d-dimension and parameter $\beta \geq 1$ is given as :

$$C_\beta(\boldsymbol{u}) = \exp\left\{ - \left[\sum_{i=1}^{d} (-\ln u_i)^\beta \right]^{1/\beta} \right\}$$

Let $\boldsymbol{a} = (a_1, \ldots, a_d) \in \boldsymbol{I}^d$ be extra-parameters.

The **extra-parametrized Gumbel copula** is defined as, with $\psi = (\boldsymbol{a}, \beta_A, \beta_B)$, $\beta_A \leq \beta_B$:

$$\boldsymbol{C}_\psi(\boldsymbol{u}) = \boldsymbol{C}_{\beta_A}(\boldsymbol{u}^{\boldsymbol{a}})\ \boldsymbol{C}_{\beta_B}(\boldsymbol{u}^{1-\boldsymbol{a}})$$

Equivalently, let $\boldsymbol{U} \sim \boldsymbol{C}_{\beta_A}$ and $\boldsymbol{V} \sim \boldsymbol{C}_{\beta_B}$. Then

$$\boldsymbol{Z} = \max(\boldsymbol{U}^{1/a}, \boldsymbol{V}^{1/(1-a)}) \sim \boldsymbol{C}_\psi$$

Extra-parameter mapping

Max-stable spatial process defined by its finite-dimensional distributions

$\forall(s_1, \ldots, s_d)$ with $s_i \in \mathcal{S}$, let $\boldsymbol{x}_{s_i} \in \mathbb{R}^p$ be covariates

Define extra-parameters thanks to a mapping with parameters θ :

$$a_{s_i} = a(\boldsymbol{x}_{s_i}; \theta)$$

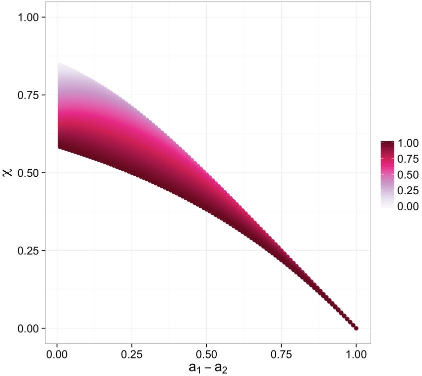
With $\psi = (a(\boldsymbol{x}_{s_1}; \theta), \ldots, a(\boldsymbol{x}_{s_d}; \theta), \beta_A, \beta_B)$, let

$$(\boldsymbol{Z}_{s_1}, \ldots, \boldsymbol{Z}_{s_d}) \sim \boldsymbol{C}_\psi$$

Since $C_{\beta.}(\cdot)$ is max-stable ➜ $C_\psi(\cdot)$ is max-stable as well

Strength of dependence

$$\chi(s_1, s_2) = \lim_{u \uparrow 1} \mathbb{P}(Z_{s_2} > u | Z_{s_1} > u)$$
$$= 2 - [(a_1^{\beta_A} + a_2^{\beta_A})^{1/\beta_A} + ((1 - a_1)^{\beta_B} + (1 - a_2)^{\beta_B})^{1/\beta_B}]$$



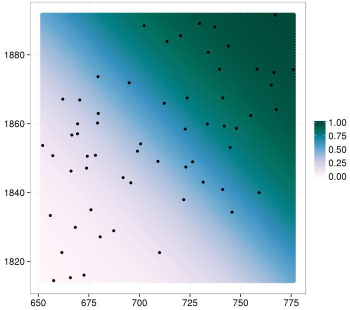
$a_1 = a_2 = 1 \rightarrow \boldsymbol{C}_{\psi}(u_1, u_2) = \boldsymbol{C}_{\beta_A}(u_1, u_2)$

$a_1 = a_2 = 0 \rightarrow \boldsymbol{C}_{\psi}(u_1, u_2) = \boldsymbol{C}_{\beta_B}(u_1, u_2)$

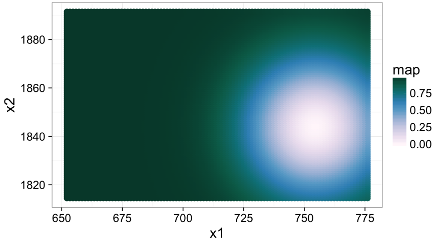
How to define $a(\cdot; \theta)$ so that

- dependence strength decreases with distance ?
- potential anisotropy ?

Linear mapping ?

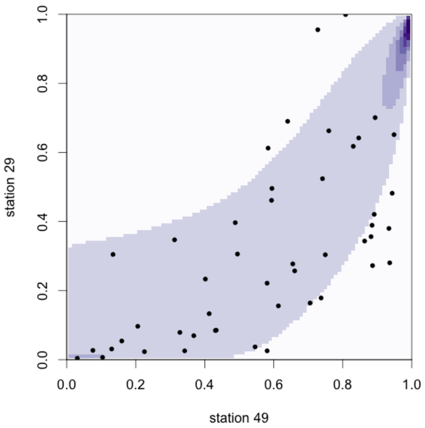


Spherical mapping ?

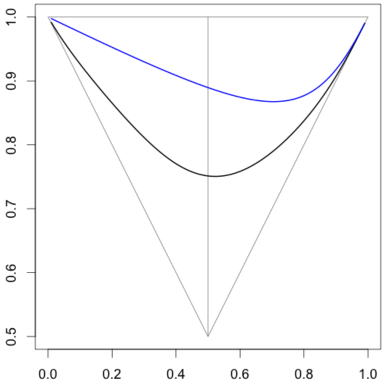


➜ Need of a stochastic mapping changing with time

Asymmetry



station 29
station 49
0.0
0.2
0.4
0.6
0.8
1.0



0.5
0.6
0.7
0.8
0.9
1.0
0.0
0.2
0.4
0.6
0.8
1.0

Latent XGumbel

Spherical mapping whose center changes each year t

$$a(\boldsymbol{x}; \boldsymbol{c}_t, \sigma) = 1 - \exp\left\{-\frac{||\boldsymbol{x} - \boldsymbol{c}_t||^2}{2\sigma^2}\right\} \quad \boldsymbol{c}_t \in \mathbb{R}^p,\ \sigma \in \mathbb{R}$$

The centers $\boldsymbol{c}_t$ are sampled from

$$\sum_{j=1}^{m} \pi_j \mathcal{N}(\boldsymbol{\mu}_j, \sigma_j \boldsymbol{I}_p)$$

Then $\theta = (\pi_1, \ldots, \pi_m, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m, \sigma_1, \ldots, \sigma_m)$

Bivariate density

$$\begin{aligned} c_{a_1 a_2}(u_1,u_2) &= a_1 a_2 u_1^{a_1-1} u_2^{a_2-1} C_{\beta_B}(u_1^{1-a_1},u_2^{1-a_2}) c_{\beta_A}(u_1^{a_1},u_2^{a_2}) + \\ & (1-a_1)(1-a_2) u_1^{-a_1} u_2^{-a_2} C_{\beta_A}(u_1^{a_1},u_2^{a_2}) c_{\beta_B}(u_1^{1-a_1},u_2^{1-a_2}) + \\ & a_1(1-a_2) u_1^{a_1-1} u_2^{-a_2} \frac{\partial C_{\beta_A}(u_1^{a_1},u_2^{a_2})}{\partial u_1^{a_1}} \frac{\partial C_{\beta_B}(u_1^{1-a_1},u_2^{1-a_2})}{\partial u_2^{1-a_2}} + \\ & (1-a_1) a_2 u_1^{-a_1} u_2^{a_2-1} \frac{\partial C_{\beta_A}(u_1^{a_1},u_2^{a_2})}{\partial u_2^{a_2}} \frac{\partial C_{\beta_B}(u_1^{1-a_1},u_2^{1-a_2})}{\partial u_1^{1-a_1}} \end{aligned}$$

➜ Pairwise log-likelihood works fine for standard XGumbel
Not possible for latent version

Let $\boldsymbol{Z} = (Z_{s_1}, \ldots, Z_{s_d})$ the observed sample

Let $S(\boldsymbol{Z})$ be a set of statistics

Rejection ABC

1. Draw (ψ_i, θ_i) from prior
2. Simulate $\boldsymbol{Z}^{(i)} = (Z_{s_1}^{(i)}, \ldots, Z_{s_d}^{(i)})$ from latent XGumbel with parameters (ψ_i, θ_i)
3. Reject (ψ_i, θ_i) if $d(S(\boldsymbol{Z}^{(i)}), S(\boldsymbol{Z})) > \epsilon$

$$S(\boldsymbol{Z}) = R(h; \hat{\beta}, \hat{\alpha}) \quad S(\boldsymbol{Z}^{(i)}) = R(h; \hat{\beta}_i, \hat{\alpha}_i)$$

➜ the entire ρ-correlogram curves fitted to the observations / simulations

Define

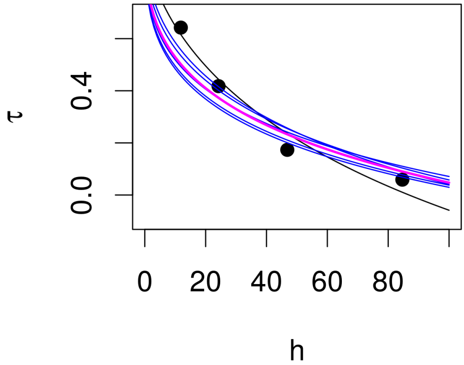
$$d(S(\boldsymbol{Z}^{(i)}), S(\boldsymbol{Z})) = \int |R(h; \hat{\beta}, \hat{\alpha}) - R(h; \hat{\beta}_i, \hat{\alpha}_i)| dh$$

Anisotropy : distances between directional ρ-correlogram curves

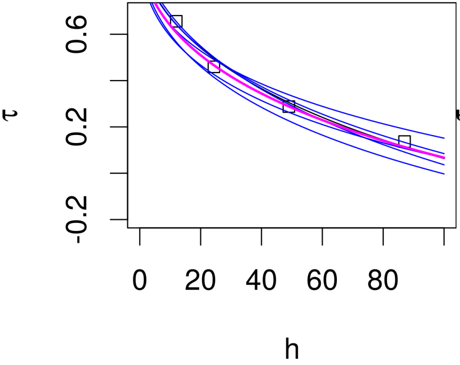
Erhardt, R.J. & Smith, R.L. (2012) *Approximate Bayesian computing for spatial extremes*, CSDA, 56(6), 1468-1481

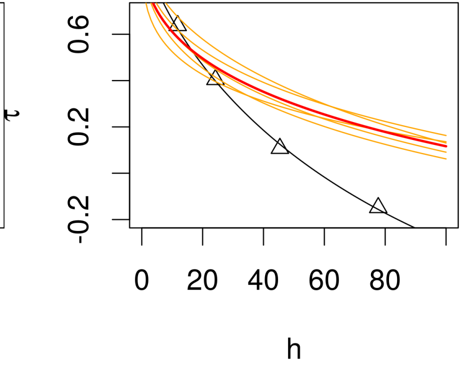
First results

Omni-directional ρ correlogram



Bi-directional ρ correlogram





Summary and outlook

Spatial latent XGumbel provides a flexible model for spatial maxima with potential asymmetry and anisotropy

Need to better understand how to design the mapping of the extra-parameters

ABC seems promising to fit models for spatial extremes

Need to assess if asymmetry can be reproduced as a summary statistics

➜ Perform a general validation of the spatial latent XGumbel model