



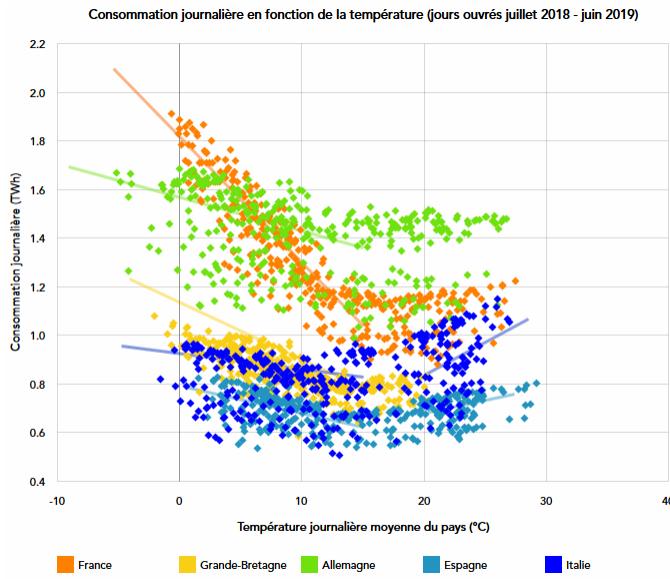
# STOCHASTIC MODELLING AND EXTREMES OF CLIMATE VARIABLES IN AN INDUSTRIAL CONTEXT



# EDF AND CLIMATE

Electricity generation, demand and transmission are linked to the meteorological conditions in many ways

## Daily electricity demand in Europe

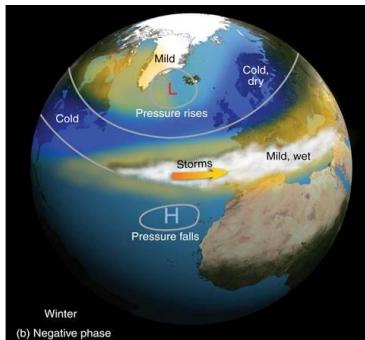
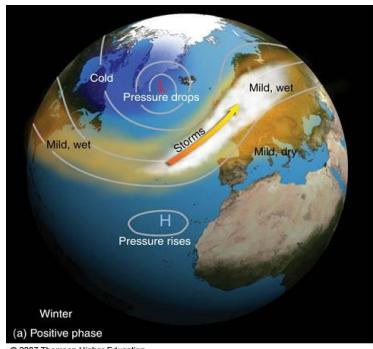


## Electricity generation

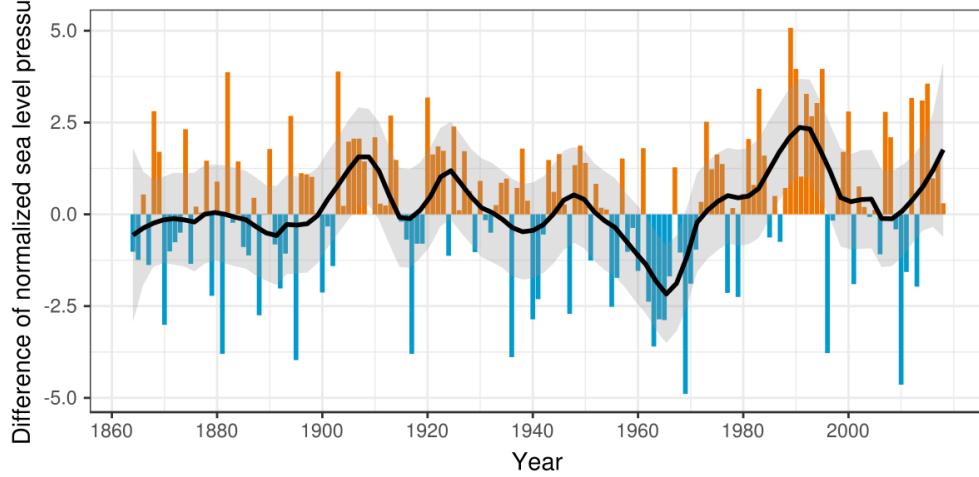


# EDF IS ADAPTED TO CURRENT CLIMATE VARIABILITY

NAO

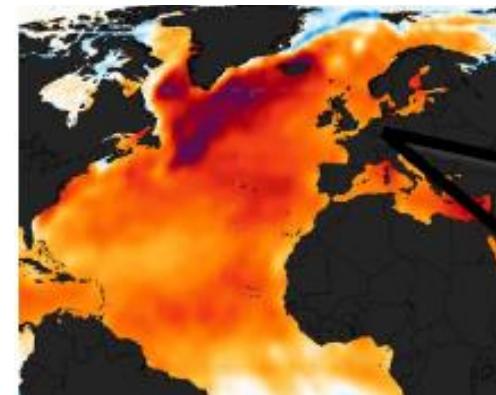


North Atlantic Oscillation (NAO) winter index  
Lisbon - Stykkishólmur/Reykjavík, December to March

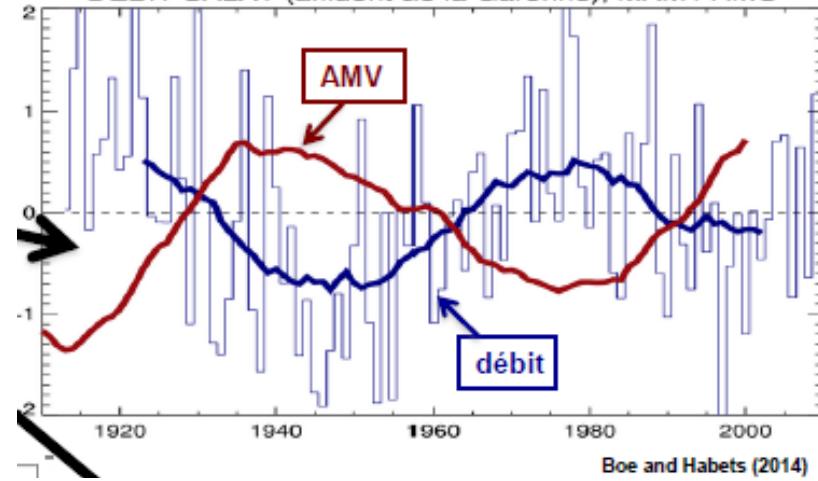


NAO Index Data provided by the Climate Analysis Section, NCAR, Boulder, USA, Hurrell (2003)  
Updated regularly. Accessed 2018-10-21

AMV

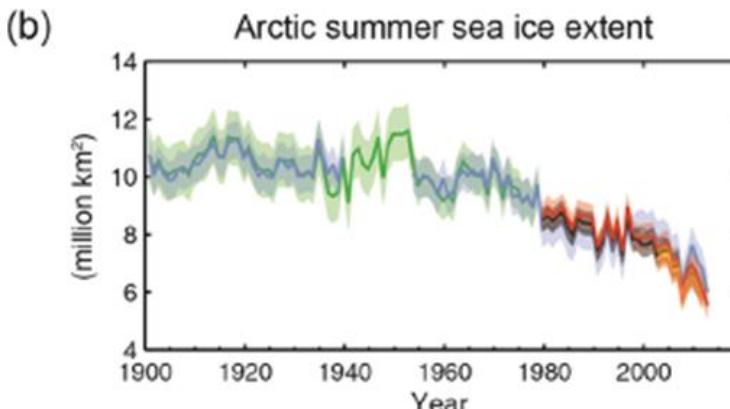
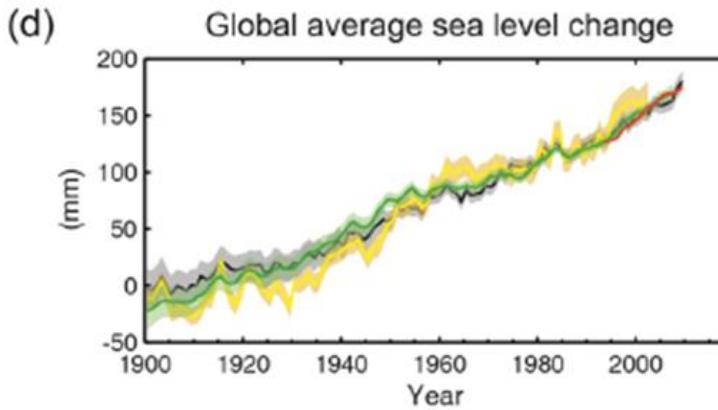
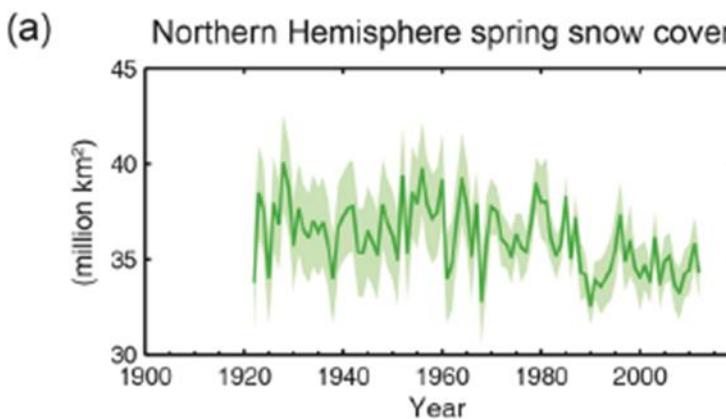
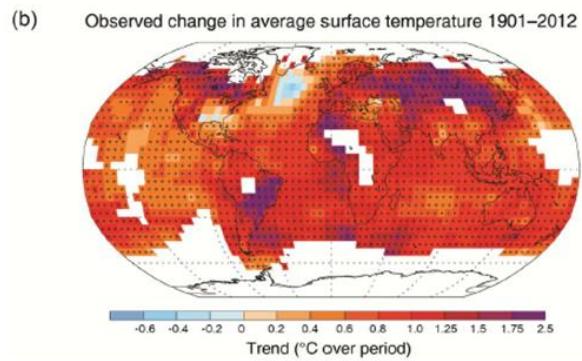
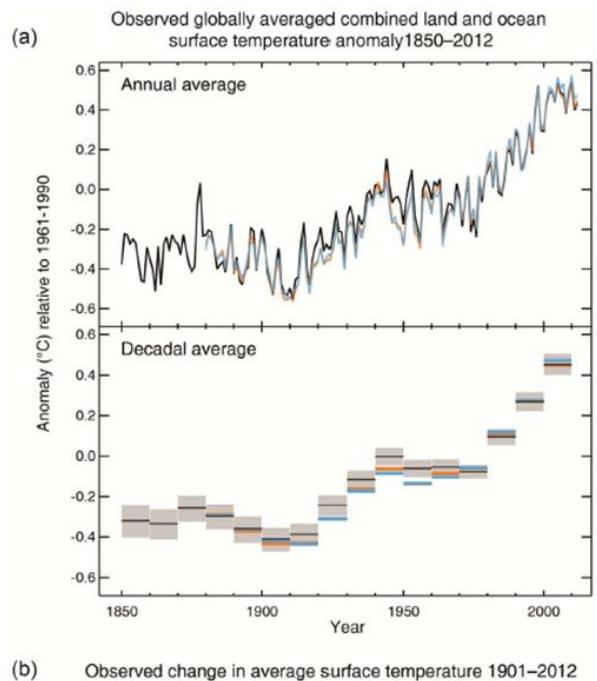


Atlantic Multidecadal Variability (AMV)  
DEBIT SALAT (affluent de la Garonne), MAM / AMO



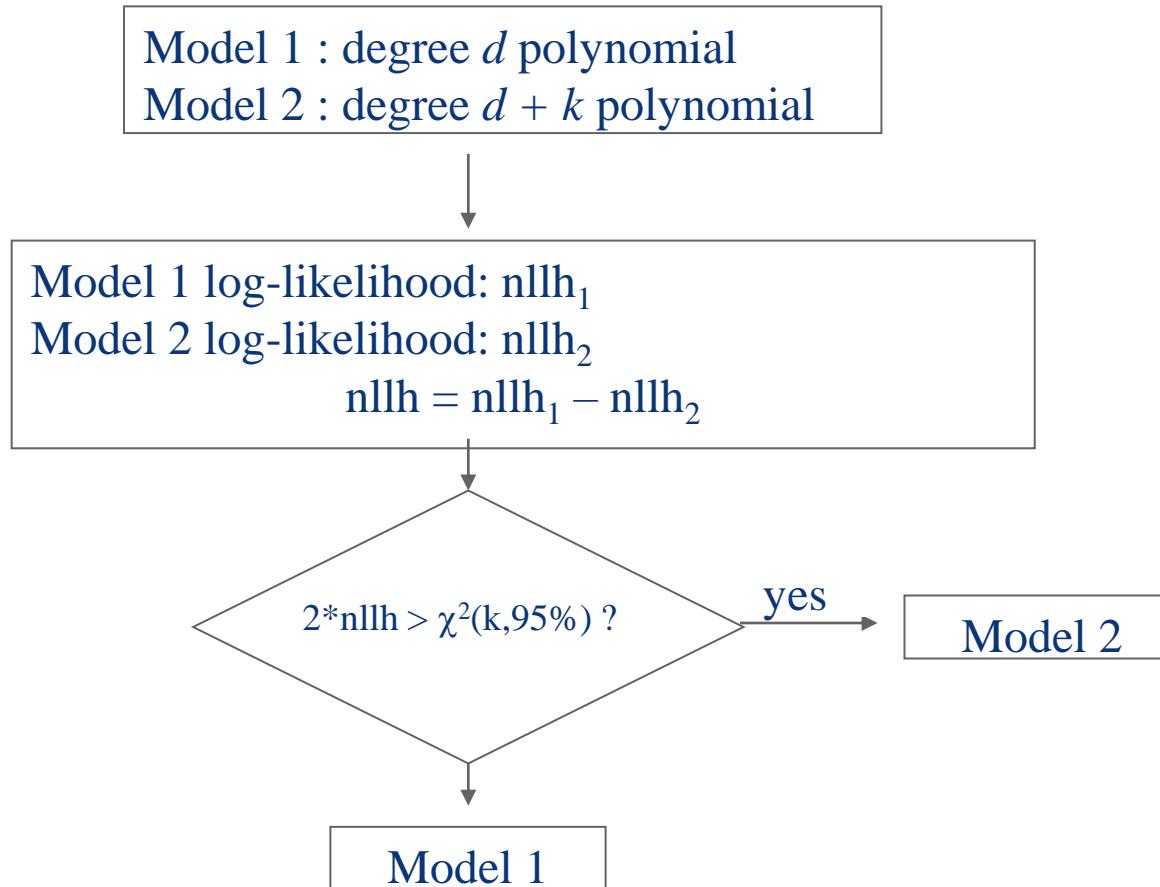
Boe and Habets (2014)

# AND IS ADAPTING TO CLIMATE CHANGE



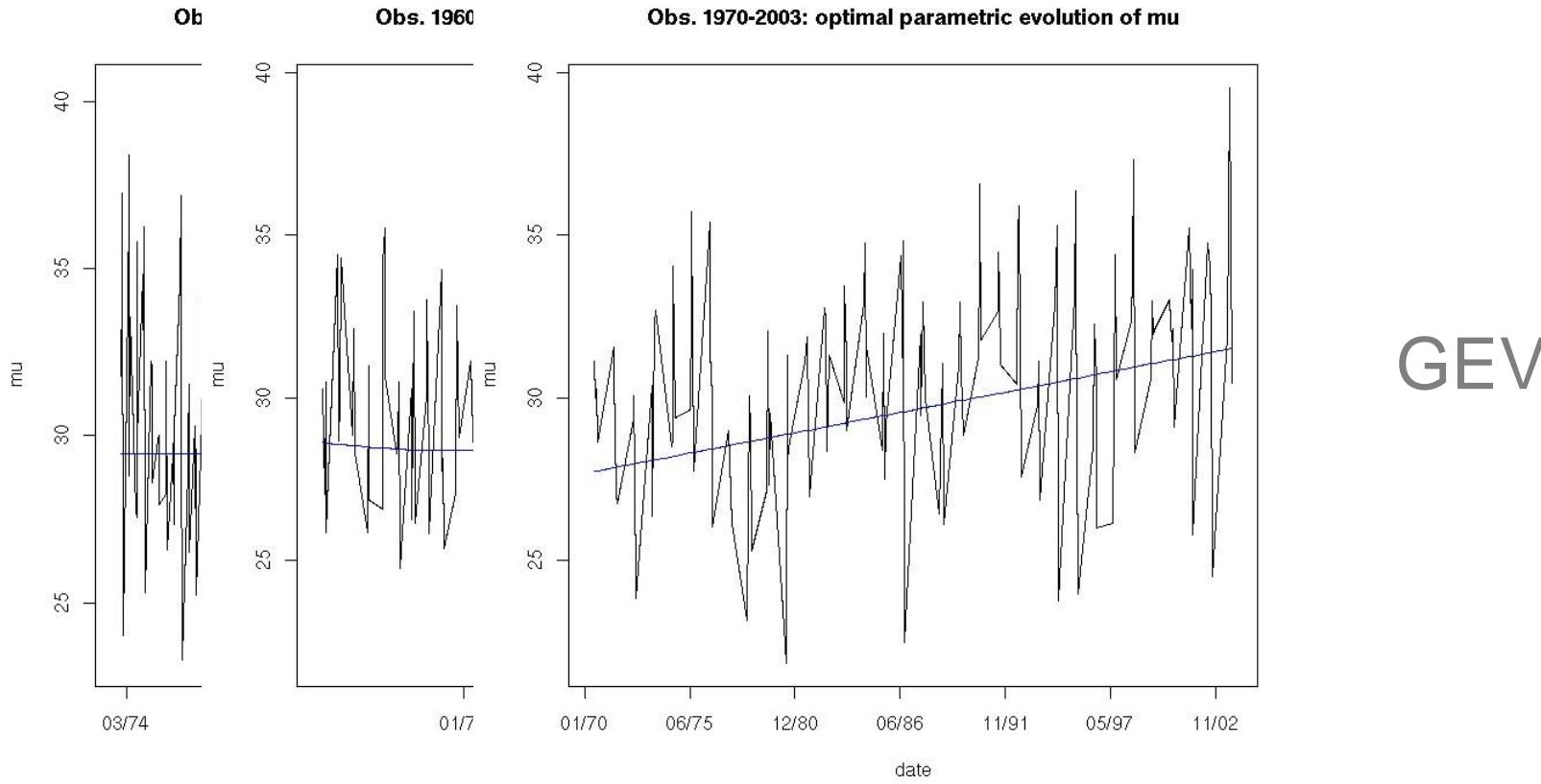
# EXAMPLE: EXTREME LEVEL ESTIMATION

- First studies in 2002 in collaboration with Professor Dacunha-Castelle (Paris 11 university)



S. Parey, F. Malek, C. Laurent, D. Dacunha-Castelle: Trends and climatic evolution: Statistical approach for very high temperatures in France, Climatic Change (2007) 81:331 - 352

# BUT THE TREND MAY DEPEND ON THE PERIOD

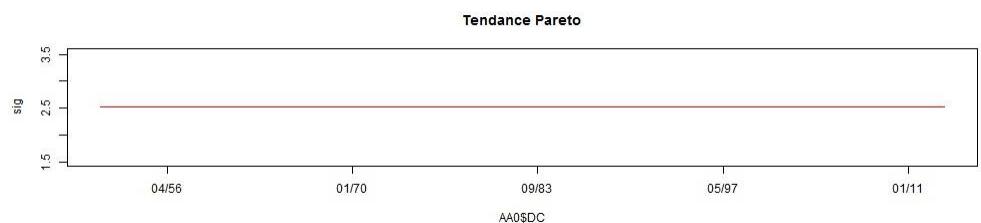
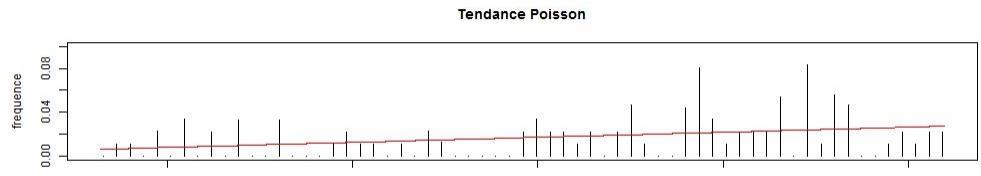
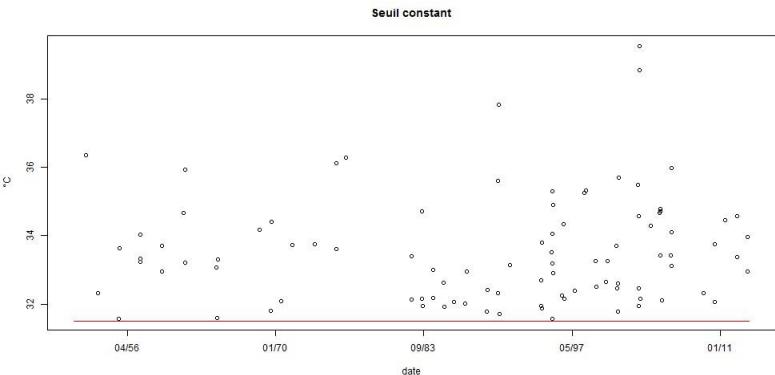


POT

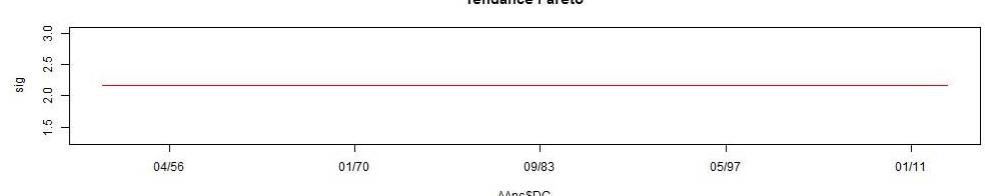
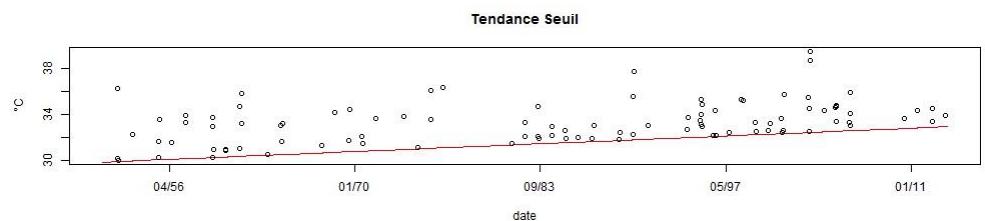
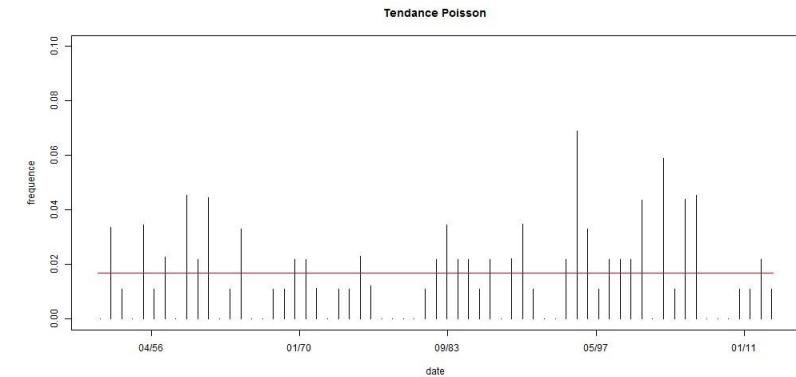
Period	P10	P21	P22	P31	P32	P33
$I(t)$ degree	5	0	2	0	0	2
$\sigma(t)$ degree	1	0	0	0	4	0

# DIFFERENT CHOICES FOR THE TREND DETECTION

## Constant threshold



## Time varying threshold



# THEN THE RETURN LEVEL MUST BE RE-DEFINED

- Value exceeded 1 every  $N$  years on average => value whose exceedance expectation in the next  $N$  years is 1

$$\frac{1}{nb} \sum_{t=t_0}^{t_0+365a} \left\{ 1 - \exp \left[ - \left( 1 + \frac{\xi}{\sigma(t)} (z_a - \mu(t)) \right)^{-\frac{1}{\xi}} \right] \right\} = 1 \quad \text{GEV}$$

POT

$$\sum_{t \in D(t_0, a)} \left( 1 + \frac{\xi}{\sigma(t)} (z_a - u) \right)^{-\frac{1}{\xi}} I(t) = 1$$

- depends on the identified trend
- implies the need to take the trend uncertainty into account in the confidence interval: bootstrap
- Other definitions in the literature:

- ENE:  $\sum_{t=T_1}^{T_1+m-1} \{1 - G_Z(z^{ENE}(m) | \theta_t)\} = 1$

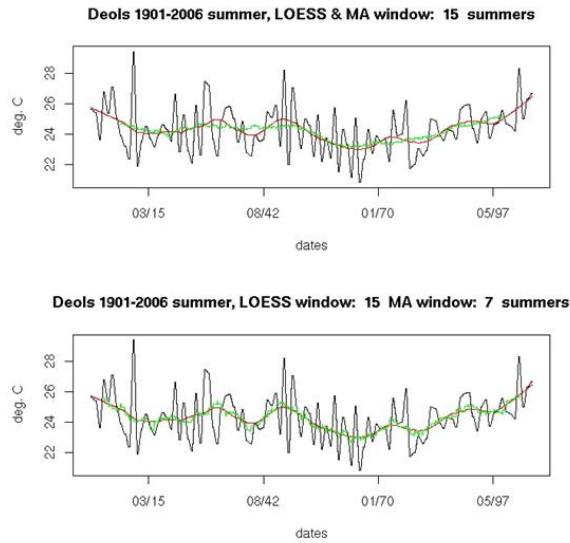
- RE:  $RE_{T_1-T_2}^{ns} = \prod_{t=T_1}^{T_2} (1 - p_t) = \prod_{t=T_1}^{T_2} G_Z(z_q | \theta_t)$

- DLL:  $z_{T_1-T_2}^{DLL}(m) = F_{T_1-T_2}^{-1}(1 - 1/m)$  with  $F_{T_1-T_2}(z) = \prod_{t=T_1}^{T_2} G_{Z,t}(z) = \prod_{t=T_1}^{T_2} G_Z(z | \theta_t)$

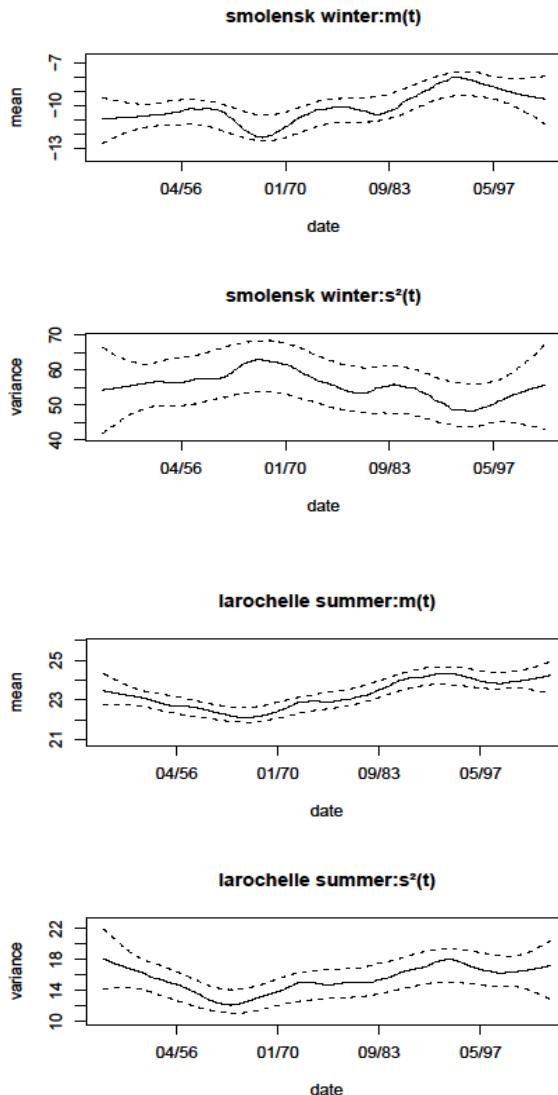
- ADLL:  $\frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} G_Z(z_{T_1-T_2}^{ADLL}(m) | \theta_t) = 1 - 1/m$

# LIMITATIONS OF THE APPROACH -> FURTHER RESEARCH

- Careful study of the links between mean, variance and extremes : Thi Thu Huong Hoang PhD (2010)
- Nonparametric trends
  - Cubic splines
  - Loess: local regression

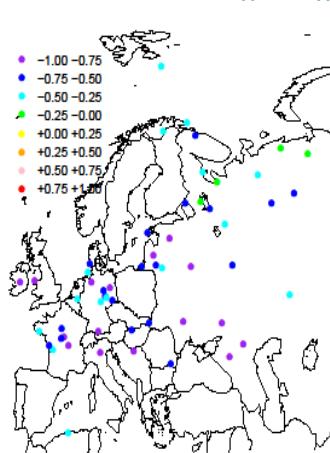


# TEMPERATURE MEAN AND VARIANCE TRENDS ARE LINKED

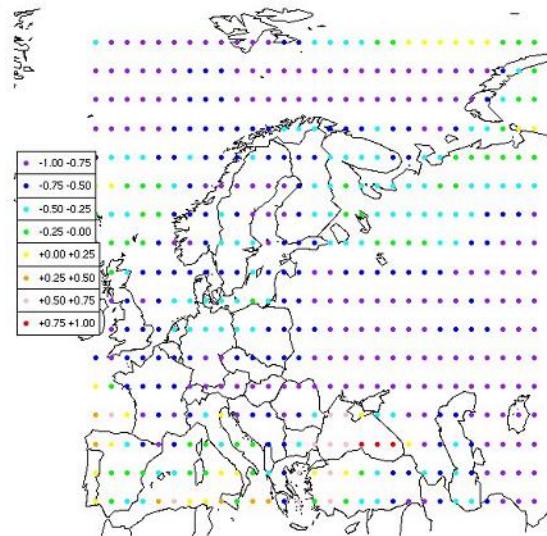


WINTER

55 ECA series: correlation between  $m(t)$  and  $s^2(t)$  in winter

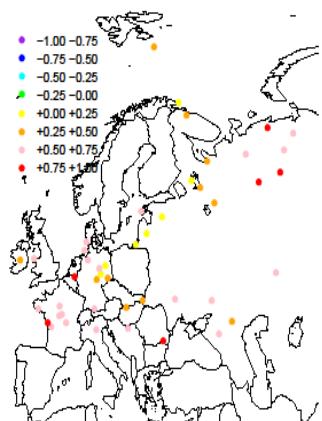


ERA40: correlations in winter

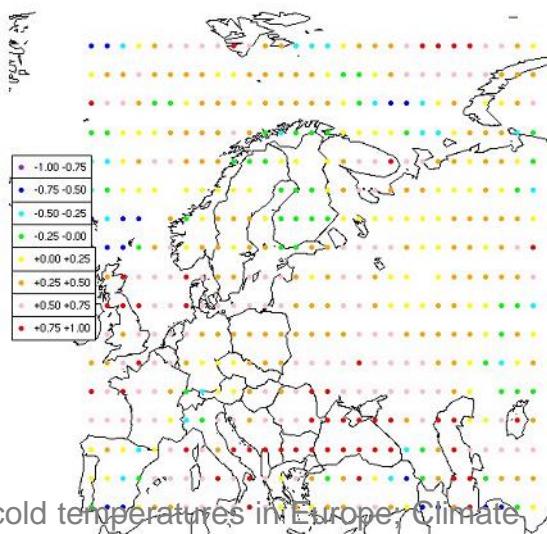


SUMMER

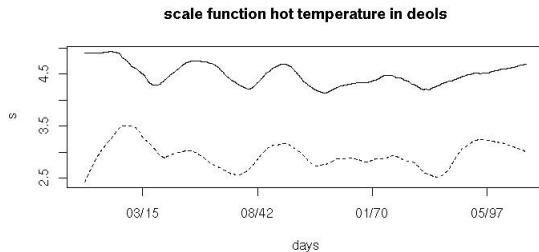
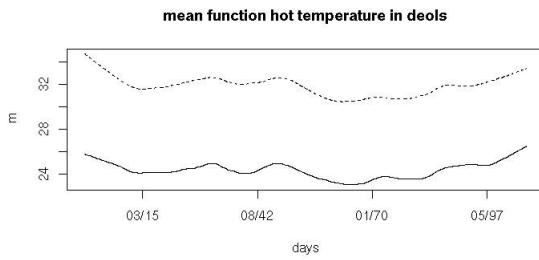
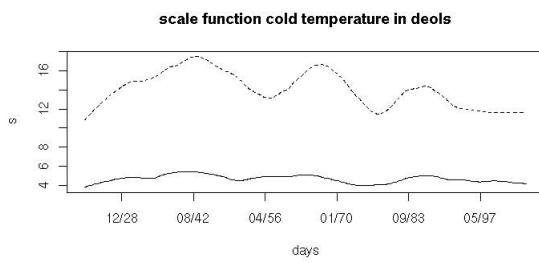
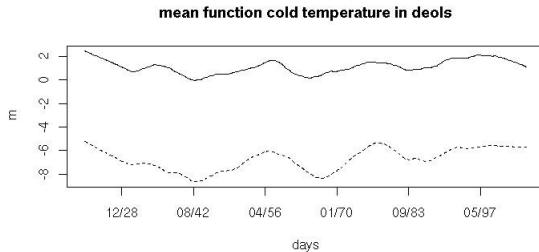
55 ECA series: correlation between  $m(t)$  and  $s^2(t)$  in summer



ERA40: correlations in summer

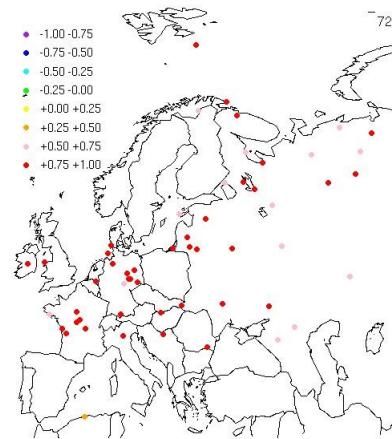


# LINK WITH THE TRENDS IN EXTREMES



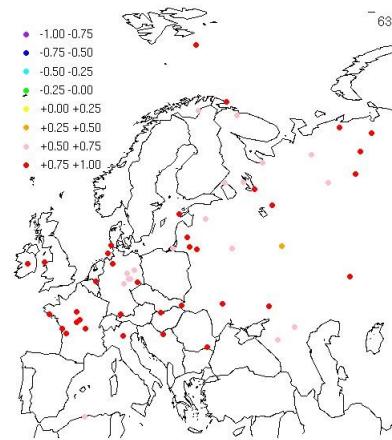
WINTER

55 ECA series: correlation between  $m(t)$  and  $mx(t)$  in winter



SUMMER

55 ECA series: correlation between  $m(t)$  and  $mx(t)$  in summer



# BUILD A VARIABLE WITH STATIONARY EXTREMES?

$$Y_t = \frac{X_t - m_t}{s_t}$$

Assumption :

the extremes of  $Y_t$  are stationary

Test

1) Define a distance between two functions:

$$\Delta(f, g) = \frac{1}{T} \left( \int (f(t) - g(t))^2 dt \right)^{1/2}$$

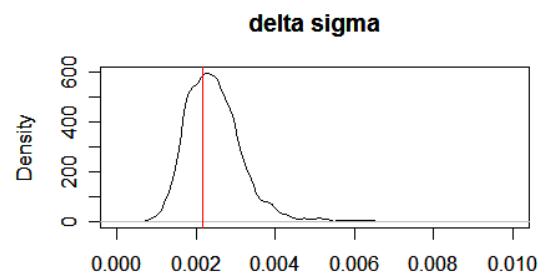
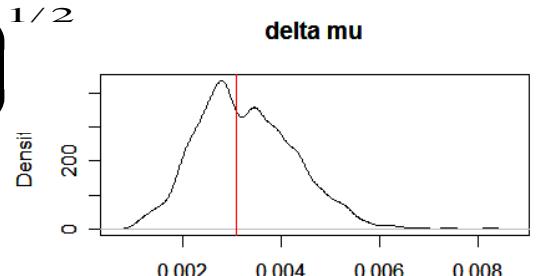
2) Compute a statistical table of these distances for a stationary distribution

- 1000 samples of a stationary GEV ( $\xi_Y, \mu_Y, \sigma_Y$ )

- Estimate the parameters:

- 1) As constants
- 2) As time varying

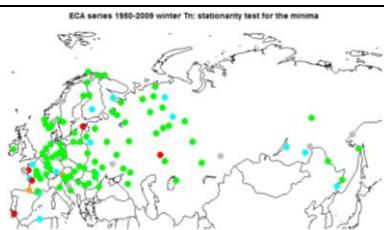
- Compute the distance  $\Delta$



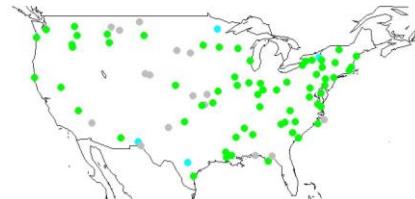
# RESULTS FOR DIFFERENT LOCATIONS

Minimum Winter TN

a)

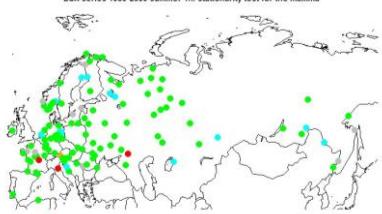


NOAA series 1950-2009 winter Tn: stationarity test for the minima

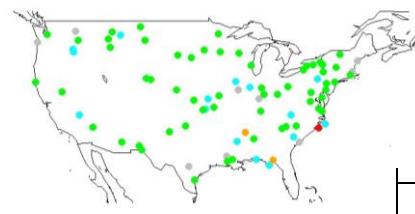


Maximum Summer TN

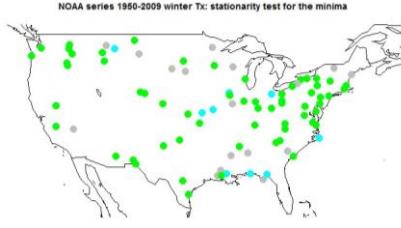
b)



NOAA series 1950-2009 summer Tn: stationarity test for the maxima

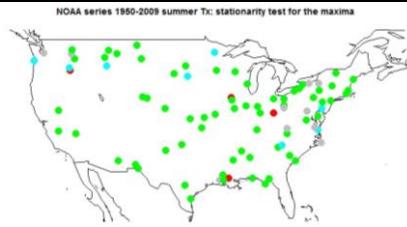


c)



Maximum Summer TX

d)



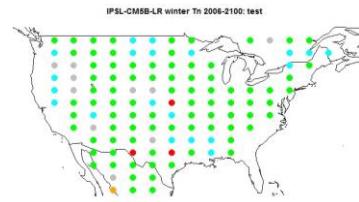
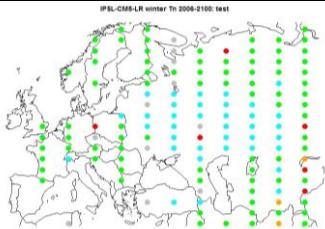
● TRUE for  $\mu$  &  $\sigma$  ● TRUE for  $\mu$  only ● TRUE for  $\sigma$  only ● FALSE ● non convergence



# AND FOR FUTURE CLIMATE

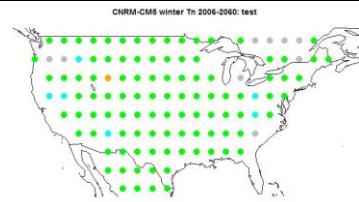
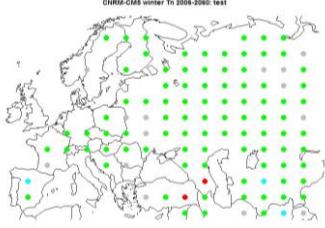
Minimum Winter TN: IPSL-CM5-LR

a)



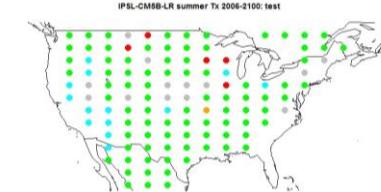
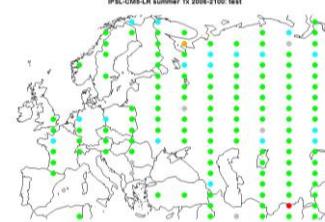
Minimum Winter TN: CNRM-CM5

b)



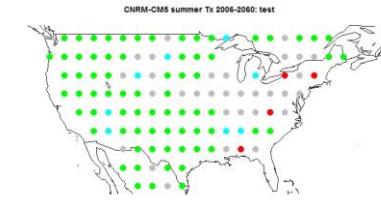
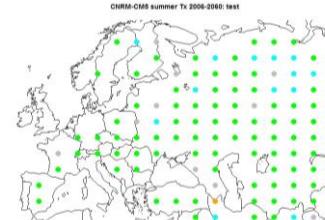
Maximum summer TX: IPSL-CM5-LR

c)



Maximum summer TX: CNRM-CM5

d)



# FUTURE RETURN LEVEL

- Hot extremes

- Estimation:  $Z_X = m_f + s_f Z_Y$ , with  $m_f$  and  $s_f$  in the desired future period obtained from climate simulations

- Cross validation:

Observation period splitted in 2 sub-periods

Estimation of  $Y$  over 1951-1986

Estimation of the RL over 1987-2016

model	Bourges	Perpignan	Caen	Clermont-Ferrand	Tarbes	Mont de Marsan	Belle Ile
Bcc-csm1-1-m	<b>40.2</b> [ 38.3 ; 42.1 ]	<b>40.3</b> [ 38.8 ; 41.9 ]	<b>38.8</b> [ 36.5 ; 41.1 ]	<b>40.7</b> [ 39.1 ; 42.4 ]	<b>39.2</b> [ 37.8 ; 40.6 ]	<b>41.1</b> [ 40.1 ; 42.1 ]	<b>35.2</b> [ 33.8 ; 36.6 ]
BNU-ESM	<b>39.7</b> [ 37.8 ; 41.5 ]	<b>40.2</b> [ 38.6 ; 41.7 ]	<b>38.4</b> [ 36.2 ; 40.7 ]	<b>40.4</b> [ 38.7 ; 42.0 ]	<b>39.3</b> [ 37.9 ; 40.7 ]	<b>40.9</b> [ 39.9 ; 41.9 ]	<b>34.0</b> [ 32.7 ; 35.3 ]
CNRM-CM5	<b>38.2</b> [ 36.4 ; 40.0 ]	<b>39.2</b> [ 37.6 ; 40.7 ]	<b>37.5</b> [ 35.3 ; 39.8 ]	<b>38.7</b> [ 37.1 ; 40.2 ]	<b>38.1</b> [ 36.7 ; 39.5 ]	<b>39.5</b> [ 38.5 ; 40.5 ]	<b>33.3</b> [ 32.1 ; 34.6 ]
EC-EARTH	<b>39.4</b> [ 37.5 ; 41.4 ]	<b>39.9</b> [ 38.3 ; 41.4 ]	<b>38.6</b> [ 36.3 ; 41.0 ]	<b>40.1</b> [ 38.4 ; 41.7 ]	<b>38.8</b> [ 37.4 ; 40.2 ]	<b>40.7</b> [ 39.7 ; 41.7 ]	<b>34.5</b> [ 33.2 ; 35.9 ]
IPSL-CM5A-MR	<b>39</b> [ 37.1 ; 40.9 ]	<b>40.6</b> [ 38.9 ; 42.2 ]	<b>37.5</b> [ 35.3 ; 39.8 ]	<b>40.0</b> [ 38.3 ; 41.7 ]	<b>38.8</b> [ 37.4 ; 40.3 ]	<b>40.4</b> [ 39.4 ; 41.4 ]	<b>33.4</b> [ 32.1 ; 34.6 ]
MIROC5	<b>38.6</b> [ 36.7 ; 40.4 ]	<b>40</b> [ 38.4 ; 41.5 ]	<b>36.1</b> [ 34.0 ; 38.2 ]	<b>39.3</b> [ 37.7 ; 40.9 ]	<b>38.1</b> [ 36.7 ; 39.5 ]	<b>39.6</b> [ 38.6 ; 40.6 ]	<b>33.1</b> [ 31.9 ; 34.4 ]
MIROC-ESM-CHEM	<b>37.4</b> [ 35.5 ; 39.2 ]	<b>38.0</b> [ 36.5 ; 39.5 ]	<b>35.7</b> [ 33.6 ; 37.9 ]	<b>38.0</b> [ 36.5 ; 39.6 ]	<b>37.3</b> [ 35.9 ; 38.7 ]	<b>39.0</b> [ 38.0 ; 39.9 ]	<b>33.5</b> [ 32.2 ; 34.8 ]
MIROC-ESM	<b>40.5</b> [ 38.5 ; 42.6 ]	<b>40.2</b> [ 38.5 ; 41.9 ]	<b>37.5</b> [ 35.2 ; 39.9 ]	<b>41.3</b> [ 39.5 ; 43.1 ]	<b>39.3</b> [ 37.7 ; 40.8 ]	<b>40.8</b> [ 39.7 ; 41.9 ]	<b>33.7</b> [ 32.3 ; 35.0 ]
MPI-ESM-MR	<b>38.9</b> [ 37.0 ; 40.8 ]	<b>39.8</b> [ 38.2 ; 41.4 ]	<b>37.2</b> [ 35.0 ; 39.4 ]	<b>39.5</b> [ 37.9 ; 41.2 ]	<b>38.4</b> [ 37.0 ; 39.8 ]	<b>40.1</b> [ 39.1 ; 41.1 ]	<b>33.5</b> [ 32.2 ; 34.8 ]
CCSM4	<b>39.9</b> [ 37.9 ; 41.9 ]	<b>40.0</b> [ 38.4 ; 41.6 ]	<b>38.1</b> [ 35.8 ; 40.4 ]	<b>40.8</b> [ 39.1 ; 42.6 ]	<b>39.0</b> [ 37.6 ; 40.5 ]	<b>41.1</b> [ 40.0 ; 42.2 ]	<b>34.3</b> [ 32.9 ; 35.6 ]
NorESM1-M	<b>40.7</b> [ 38.6 ; 42.7 ]	<b>40.7</b> [ 39.0 ; 42.4 ]	<b>40</b> [ 37.4 ; 42.5 ]	<b>41.4</b> [ 39.6 ; 43.2 ]	<b>39.8</b> [ 38.2 ; 41.3 ]	<b>41.4</b> [ 40.3 ; 42.5 ]	<b>35.1</b> [ 33.7 ; 36.5 ]
GFDL-CM3	<b>40.0</b> [ 38.0 ; 42.0 ]	<b>40.8</b> [ 39.1 ; 42.5 ]	<b>38.5</b> [ 36.2 ; 40.9 ]	<b>40.7</b> [ 38.9 ; 42.4 ]	<b>39.2</b> [ 37.7 ; 40.7 ]	<b>40.8</b> [ 39.8 ; 41.9 ]	<b>35.2</b> [ 33.8 ; 36.6 ]
CESM1-CAM5	<b>39.7</b> [ 37.7 ; 41.6 ]	<b>40.2</b> [ 38.7 ; 41.8 ]	<b>38.9</b> [ 36.5 ; 41.3 ]	<b>40.3</b> [ 38.6 ; 41.9 ]	<b>39.1</b> [ 37.6 ; 40.5 ]	<b>40.3</b> [ 39.3 ; 41.3 ]	<b>34.2</b> [ 32.9 ; 35.5 ]
observations	<b>40.2</b> [ 38.8 ; 41.5 ]	<b>39.5</b> [ 38.7 ; 40.3 ]	<b>38.4</b> [ 36.8 ; 40.0 ]	<b>40.9</b> [ 39.7 ; 42.0 ]	<b>39.5</b> [ 38.5 ; 40.5 ]	<b>41.2</b> [ 40.1 ; 42.3 ]	<b>33.9</b> [ 32.8 ; 34.9 ]

# STOCHASTIC MODELLING OF AIR TEMPERATURE

Hoang, T. T. H. (2010), Modélisation de séries chronologiques non stationnaires, non linéaires: application à la définition des tendances sur la moyenne, la variabilité et les extrêmes de la température de l'air en Europe <http://www.tel.archivesouvertes.fr/tel-00531549/fr/>

Dacunha-Castelle D., Hoang T.T.H., Parey S.: Modeling of air temperatures: preprocessing and trends, reduced stationary process, extremes, simulation, Journal de la Société Française de Statistique, 2013

Parey S., Hoang T.T.H., Dacunha-Castelle D. (2013): Validation of a stochastic temperature generator focusing on extremes and an example of use for climate change, Climate Research

# PRE-PROCESSING

- Remove deterministic parts to obtain a stationary process
- Based on both parametric and nonparametric estimations

$$X(t) = m(t) + S(t) + s(t)S_V(t)Z(t)$$

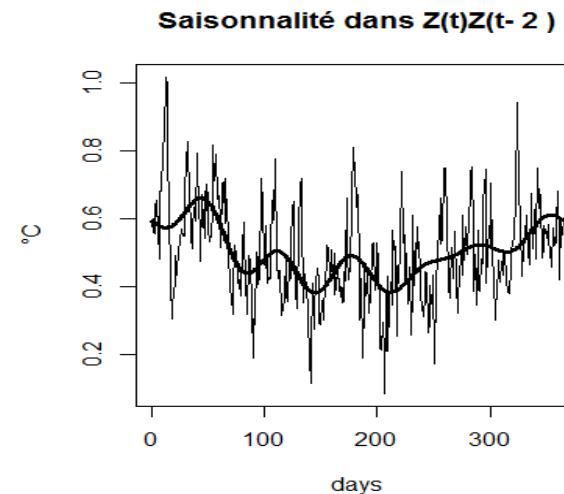
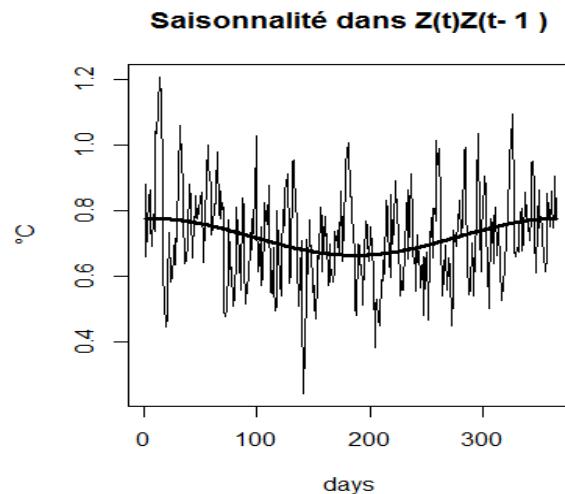
- $m(t), s(t)$  : trends in mean and standard deviation,  $S(t), S_V(t)$  : seasonality
- **Estimation :**
  - » estimate  $m(t)$  using loess,  $S(t)$  as a trigonometric function from  $X(t)$ , then  $s(t)$  using loess and  $S_V^2(t)$  as a trigonometric function from  $[X(t) - \hat{m}(t) - \hat{S}(t)]^2$
  - » For  $m(t), s(t)$ , modified partitioned cross-validation<sup>(1)</sup> is used, while for  $S(t), S_V^2(t)$  an Akaike criterion is considered
  - » Standardized timeseries:

$$Z_t = (X_t - \hat{m}_t - \hat{S}_t) / (\hat{s}_t \hat{S}_{Vt})$$

<sup>(1)</sup> Modified partitioned CV: new algorithm for correlated data (thesis of Hoang, 2010)

# CHARACTERISTICS OF THE REDUCED VARIABLE

- short memory
- seasonality remains in higher moments and autocorrelations
- cyclo-stationary
- tails are bounded
- nonlinear
- volatility depends on the state
- studies:
  - Trend tests<sup>(2)</sup> for basic statistics (mean, variance, skewness, kurtosis)
  - Cyclo-stationarity test<sup>(3)</sup> for the extremes
  - Analysis of trends and seasonality in  $Z_t Z_{t-k}$



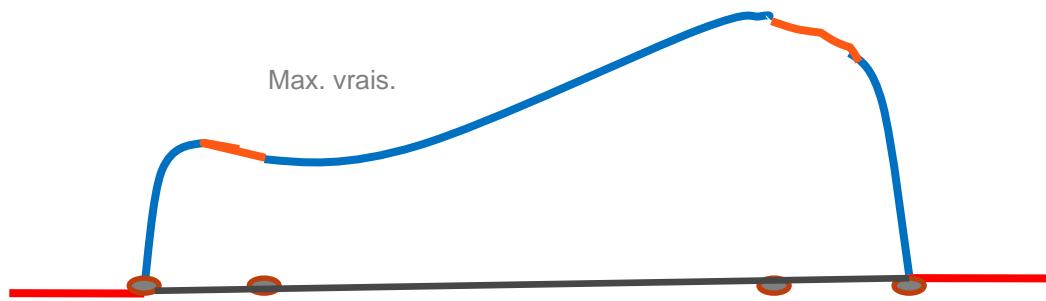
<sup>(2), (3)</sup>: A new test proposed in Hoang (2010), the principle of test is described in the next slide

# PRINCIPLE OF THE STATIONARITY TESTS

- Considered model:  $X(t) = \theta(t) + \varepsilon(t)$ , the distribution of  $\varepsilon$  is known or not
- **Hypotheses of the test:**  $\theta$  is constant /  $\theta$  is not constant
- If:
  - $\hat{c}_n$  is a constant estimator of  $\theta$  (through mlh if the distribution of  $\varepsilon$  is known, through least squares otherwise)
  - $\hat{\theta}_n$  is a nonparametric estimator of  $\theta$  (through splines if the distribution of  $\varepsilon$  is known, through loess otherwise)
- **goal** : compare those 2 estimators using an  $L^2$  distance:
$$\Delta = \|\hat{\theta}_n - \hat{c}_n\|$$
- **In practice:** test using  $\Delta$  (build an empirical distribution of  $\Delta$  under H0 by simulation if the distribution of  $\theta$  is known or by permutation (or block bootstrap) otherwise)
- **Trend tests for the moments:** mean, variance, skewness, kurtosis, correlation, and for the extremes

# MODEL TYPE

## Conditional variance form



**FARCH Model** (Functional AutoRegressive conditional Heteroscedastic)

Approximation (order 1 Euler scheme) of the discrete Markov chain given by the discrete observations of a continuous diffusion with the same coefficients

# THE SFHAR MODEL (SEASONAL FUNCTIONAL HETEROSCEDASTIC AUTOREGRESSIVE)

➤  $Z_t = b(Z_{t-1}) + a(Z_{t-1})\varepsilon_t, \quad \varepsilon_t \sim N(0,1)$

➤ Extension: SFHAR model

$$Z(t) = \left[ \theta_{0,k} + \sum_{j=1}^{p_1} \left( \theta_{1,k}^j \cos \frac{2j\pi t}{365} + \theta_{2,k}^j \sin \frac{2j\pi t}{365} \right) \right] Z(t-1) + a(t, Z_{t-1})\varepsilon_t$$
$$\varepsilon_t \sim N(0,1)$$

➤ Estimate  $a^2(t, Z_{t-1})$  with constraints:

- Zero outside the boundaries
- positive
- constraints C for the 1st derivatives of the continuous time process

$$(a^2)'(r_1) = \frac{2b(r_1, t)}{1 - 1/\xi_1} \quad \text{and} \quad (a^2)'(r_2) = \frac{2b(r_2, t)}{1 - 1/\xi_2}$$

- Which gives a

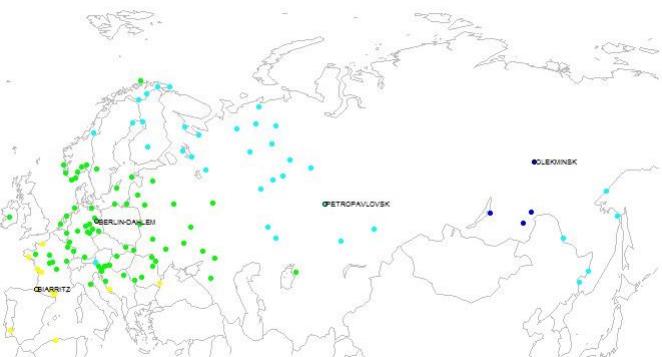
$$\begin{cases} \hat{a}^2(t, Z_{t-1}) = (\hat{r}_2 - t)(t - \hat{r}_1) \sum_{k=0}^5 \sum_{j=1}^{p_2} \left( \alpha_{1,k}^j \cos \frac{2j\pi t}{365} + \alpha_{2,k}^j \sin \frac{2j\pi t}{365} \right) Z_{t-1}^k \\ C(\hat{r}_1, t), C(\hat{r}_2, t) \\ \hat{a}^2(t) > 0 \quad \forall t \end{cases}$$

# ESTIMATION PROCEDURE

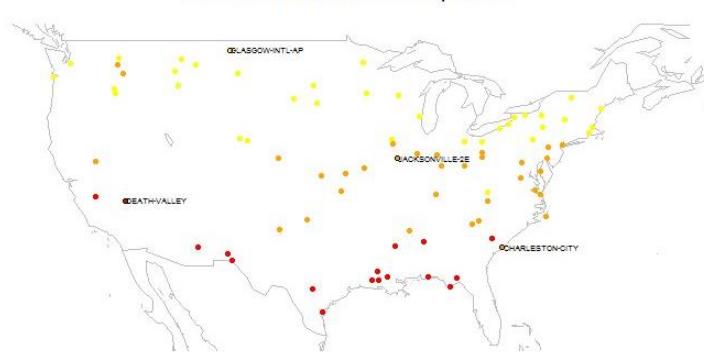
- Estimation of the autoregressive part (AR(1))
- Choice of the number of sines and cosines using Akaike criterion
- Volatility estimation by maximum likelihood with constraints
  - Initial values obtained by least squares
  - Maximum likelihood: estimation with constraints
- Simulation of  $Z(t)$
- Computation of  $X(t)$  by reintroducing trends and seasonality

# VALIDATION: CHOICE OF DIFFERENT CLIMATES

1950-2009 ECA series TN: annual mean temperature



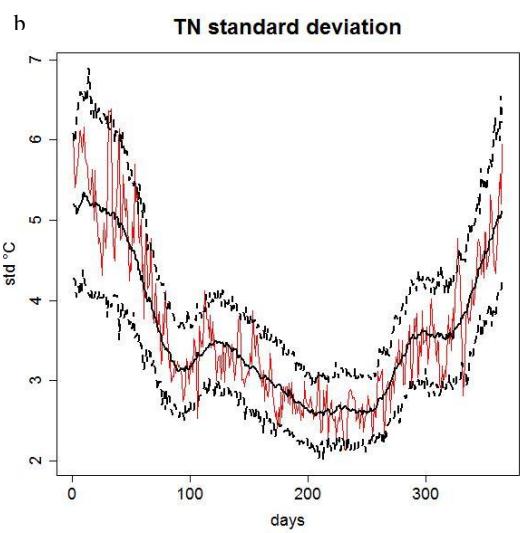
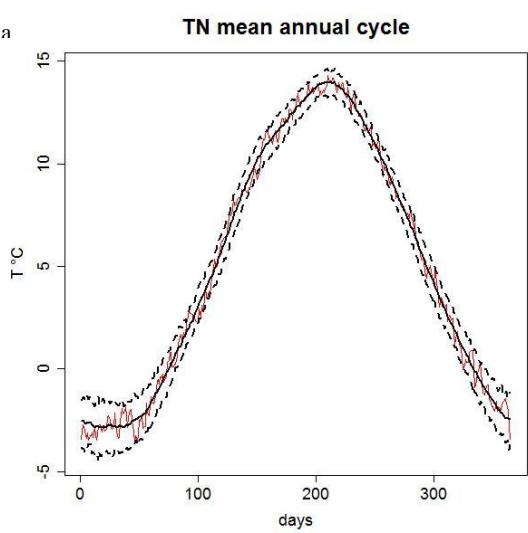
NOAA series TX annual mean temperature



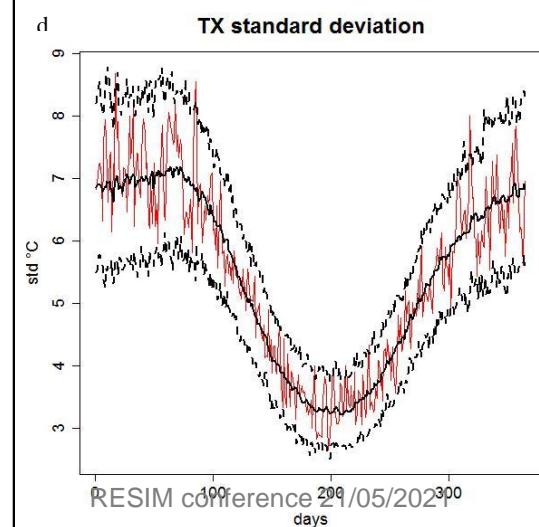
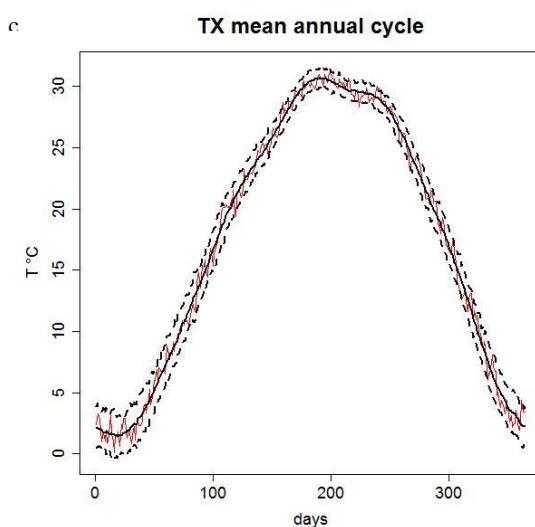
Weather station	Daily minimum temperature		Daily maximum temperature	
	period	Mean annual mean (°C)	period	Mean annual mean (°C)
Biarritz	1956-2009	10.1	1956-2009	17.7
Berlin	1950-2009	5.1	1950-2009	13.4
Petropavlovsk	1950-2009	-3.3	1950-2009	6.9
Olekmansk	1950-2009	-11.3	-	-
Death Valley	1962-2009	17.0	1962-2009	32.8
Charleston	1950-2009	15.4	1950-2009	23.0
Jacksonville	1950-2009	5.2	1950-2009	17.5
Glasgow	1950-2009	-0.7	1950-2009	12.5

# MEAN ANNUAL CYCLES

BERLIN: daily minimum temperature TN



JACKSONVILLE: daily maximum temperature TX



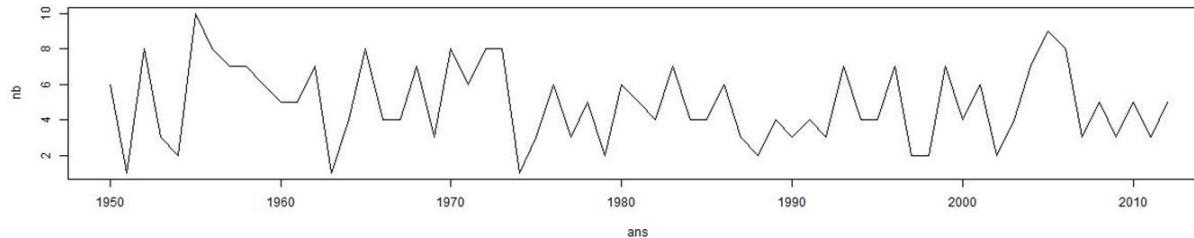
RESIM conference  
21/05/2021

# EXTREMES

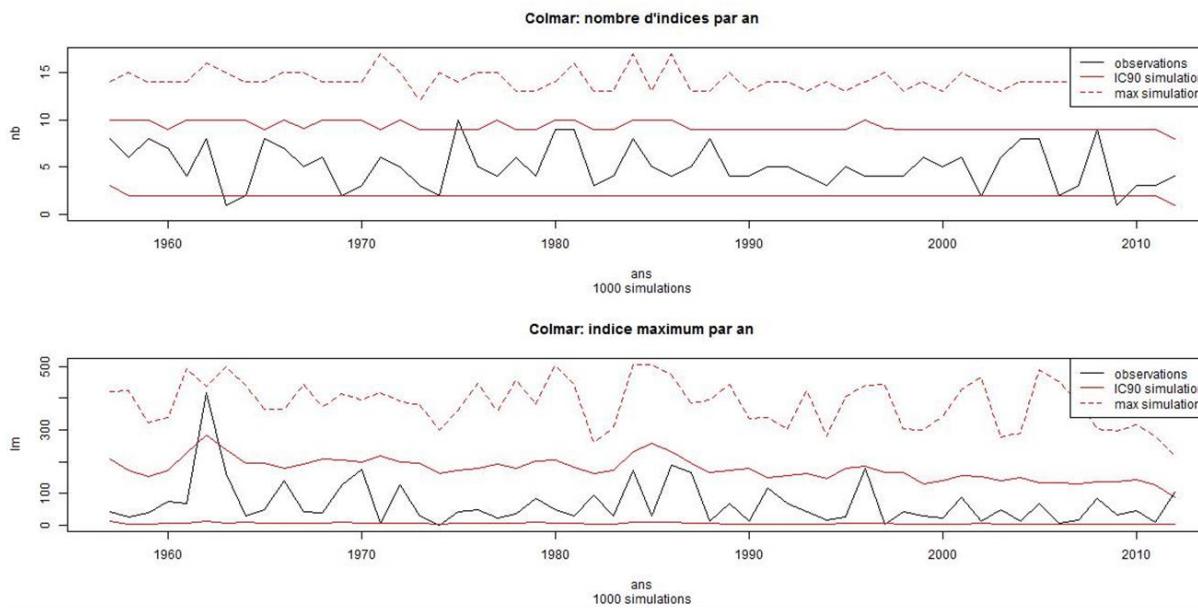
	TX		TN	
	observations	simulations	observations	simulations
Berlin	38.2 [37.1;39.2]	39.8 [38.8;41.0]	-23.4 [-25.5;-21.0]	-26.5 [-31.5;-22.9]
Biarritz	39.6 [38.8;40.4]	41.0 [39.0;43.5]	-9.4 [-12.2;-6.6]	-11.0 [-12.6;-9.7]
Petropavlovsk	38.5 [37.6;39.5]	41.5 [39.3;44.8]	-43.7 [-45.2;-42.1]	-48.7 [-52.5;-45.3]
Olekminsk	-	-	-56.3 [-57.8;-54.8]	-58.8 [-61.4;-56.2]
Death Valley	54.3 [53.5;55.1]	55.2 [54.3;56.1]	-6.4 [-7.5;-5.3]	-7.4 [-8.8;-6.0]
Jacksonville	41.8 [40.3;43.3]	43.1 [41.5;44.5]	-29.5 [-31.3;-27.7]	-33.8 [-38.5;-30.6]
Glasgow	42.0 [41.1;42.8]	45.5 [44.3;46.9]	-42.9 [-44.4;-41.4]	-46.9 [-50.4;-44.0]
Charleston	39.5 [38.6;40.4]	40.3 [39.5;41.2]	-11.3 [-13.7;-9.0]	-8.8 [-10.0;-7.5]

# EXAMPLE: FROST INDICES

- Frost index = cumulation of daily mean temperatures  $< 0^{\circ}\text{C}$   $\Rightarrow$  few events each year



- Simulation of a large number of temperature time series consistent with the observed one



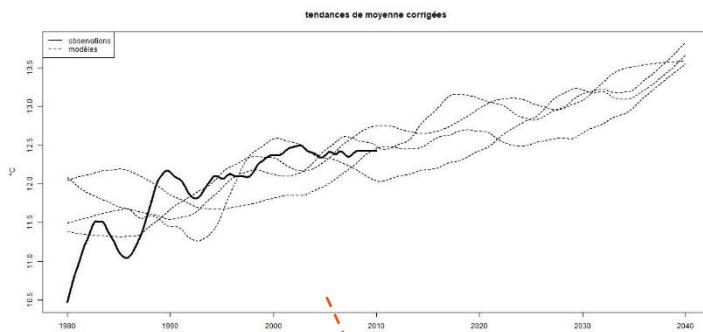
nb < nb5%: 2      nb > nb95%: 1  
Im < Im5%: 3      IM > IM95%: 2  
5% 56 years : 2,8

# BUILDING TIMESERIES COVERING RECENT PAST AND NEAR FUTURE

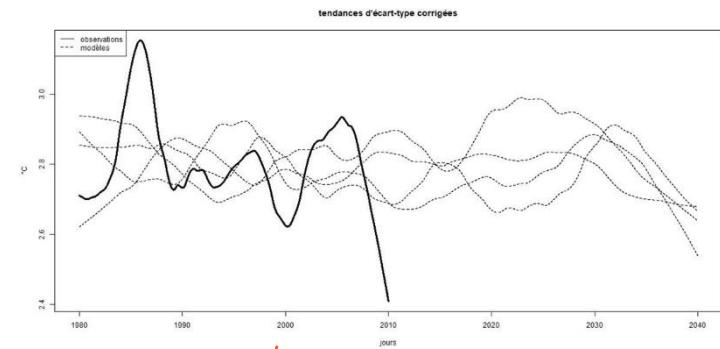
we use:

- The bias adjusted trends given by climate models

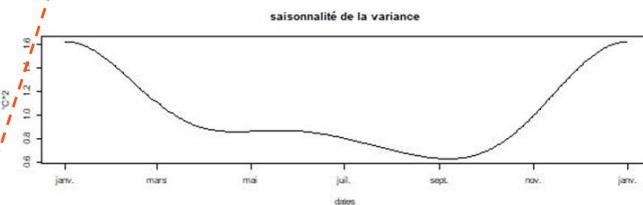
mean



variance



- The observed seasonality



- Residuals simulated with the stochastic model

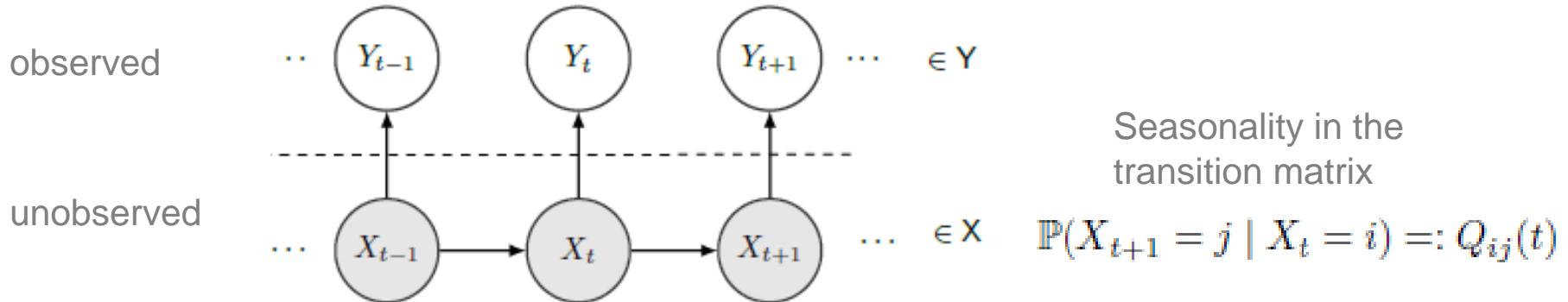
# MULTIVARIATE STOCHASTIC MODEL

Touron A. (2019): Modélisation multivariée de variables météorologiques. Statistiques [math.ST]. Université Paris-Saclay. Français. [NNT : 2019SACLSS264](#). [tel-02319170](#)

Touron, A. (2019): Consistency of the maximum likelihood estimator in seasonal hidden Markov models. *Stat Comput* **29**, 1055–1075, <https://doi.org/10.1007/s11222-019-09854-4>

# NON HOMOGENOUS HIDDEN MARKOV MODELS

- Model principle



- Trends and seasonality in the emission distributions

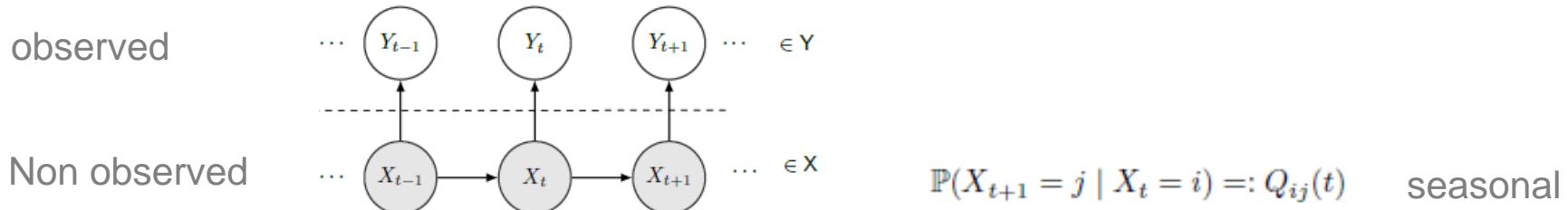
- Different models proposed

- Univariate: temperature, precipitations, wind
- Bivariate: temperature and precipitations
- Tri-variate: temperature, precipitations, wind

# EXAMPLE: BIVARIATE MODEL RAINFALL / TEMPERATURE

- **Bivariate weather generator: temperature and rainfall**

- Non-Homogeneous Hidden Markov Model (NH HMM)



- Emission law for  $Y(t)$  given  $X(t) = k$ :

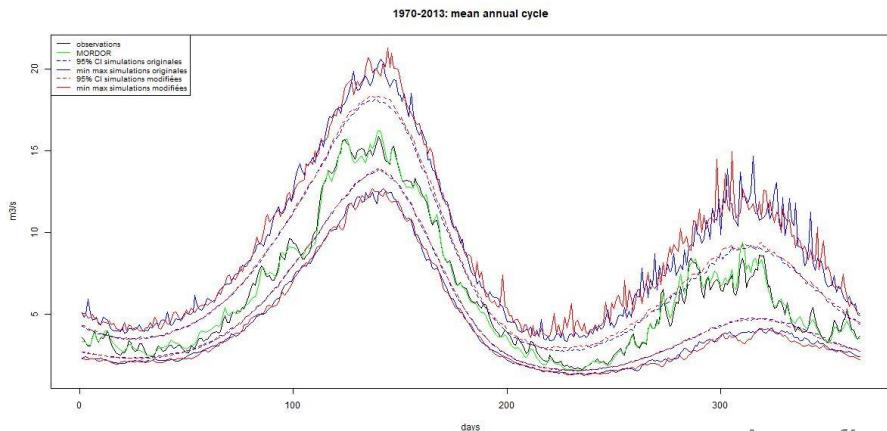
$$\begin{aligned}\nu_k(t) &= \sum_{m=1}^{M_1} p_{km} \delta_0 \otimes \mathcal{N}(T_k(t) + S_k(t) + \mu_{km}, \sigma_{km}^2) \\ &+ \sum_{m=M_1+1}^M p_{km} \mathcal{E} \left( \frac{\lambda_{km}}{1 + \sigma_k(t)} \right) \otimes \mathcal{N}(T_k(t) + S_k(t) + \mu_{km}, \sigma_{km}^2)\end{aligned}$$

- **Choice of the “meta-parameters”: 7 states,  $M1 = 2$ ,  $M = 4$**
- **1000 simulations for T & Precipitation 1970-2013 => 44 000 years**

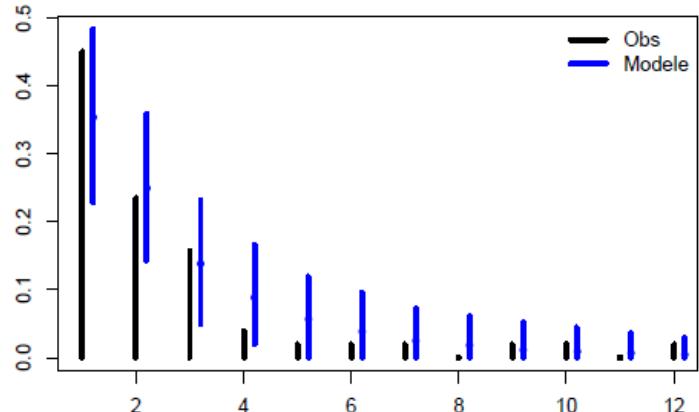
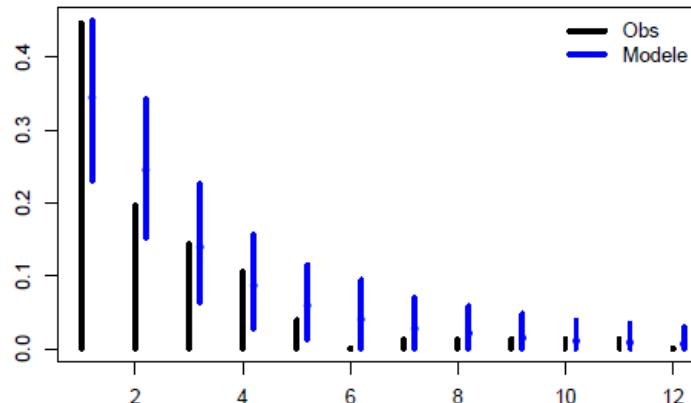
# USE TO SIMULATE RIVER FLOW

- Temperature and rainfall are used as inputs of a physically based river flow model
- Validation

Mean annual cycle

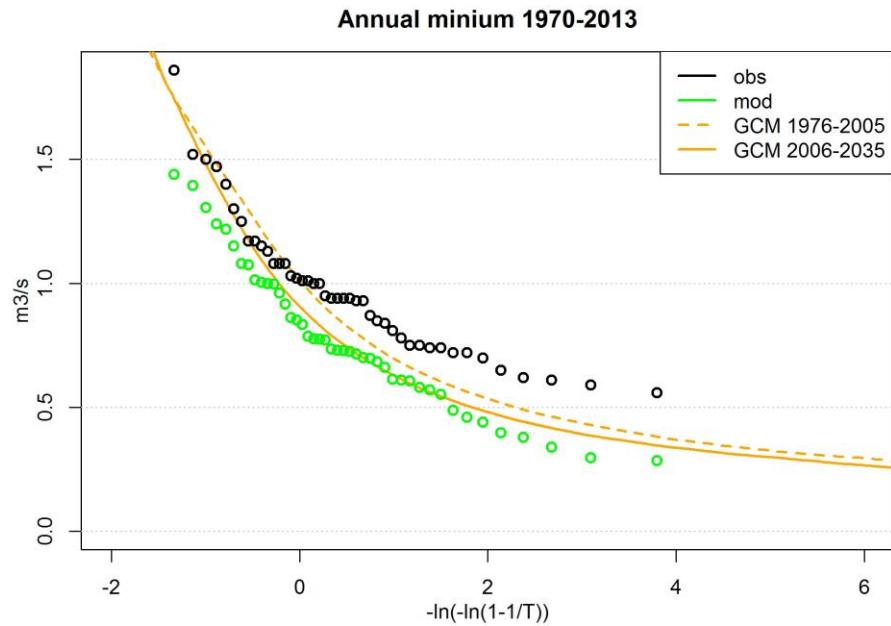


Low flows: streamflow < 5th and 3rd percentiles



# AND ESTIMATE FUTURE EXTREMES

- Calibration of the model using bias adjusted climate simulation
- Example: streamflow simulation over period 1976-2035



- Historical period 1976-2005
  - 100-year RL: 0.342 [0.340;0.344]
- Future period 2006-2035
  - 100-year RL: 0.312 [0.309;0.314]

# ADDITION OF THE SPATIAL DIMENSION

- Work in progress in collaboration with Emmanuel Gobet and David Métivier (CMAP, école Polytechnique)

- Spatial rainfall generator

- Based on Hidden Markov Models

- First step : rainfall occurrence

- 3 models tested and compared

$$-(H_1)_s^{(t)} = Y_s^{(t)} \in I_{C_1} = \{d, w\}$$

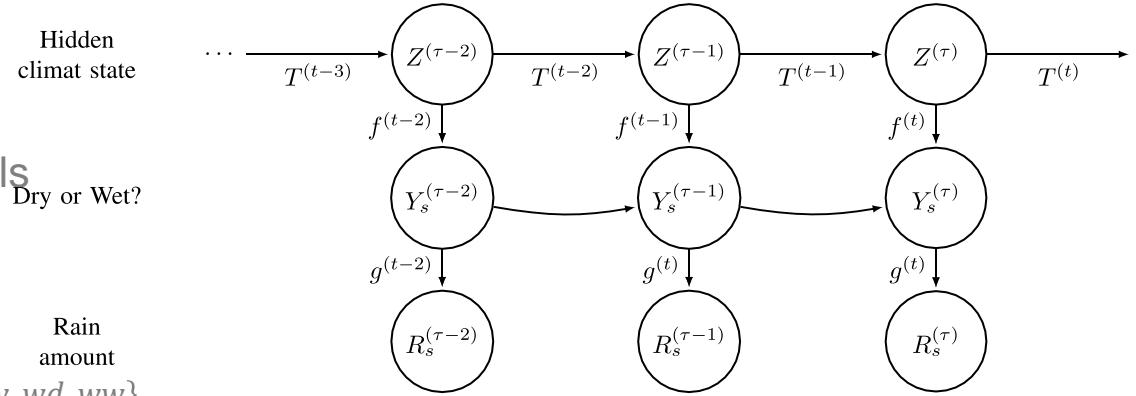
$$-(H_2)_s^{(t)} = (Y_s^{(t-1)}, Y_s^{(t)}) \in I_{C_2} = \{dd, dw, wd, ww\}$$

$$-(H_3)_s^{(t)} = (Y_s^{(t-2)}, Y_s^{(t-1)}, Y_s^{(t)}) \in I_{C_3} = \{ddd, wdd, dwd, ddw, wwd, wdw, dww, www\}$$

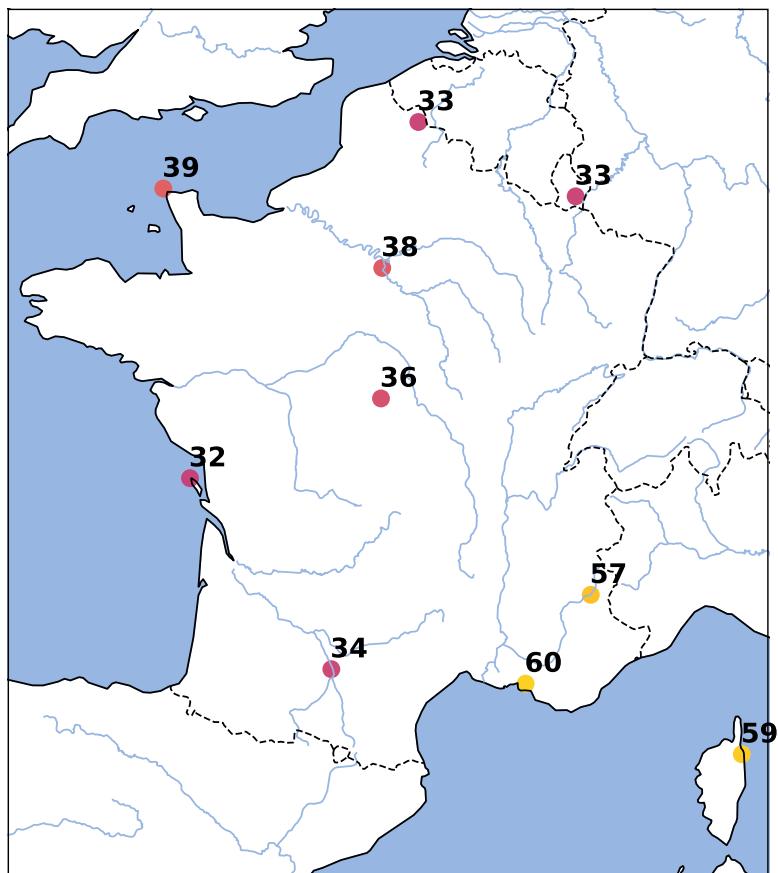
- Fit a mixture model as a probability product of each station:  $f(H) = \sum_{k=1}^K \pi_k \prod_{s=1}^S f_{s|k}(H_s)$
  - One mixture per day, with smoothing a posteriori
  - K = 4 and model C<sub>2</sub> are found optimal

- Rainfall amounts

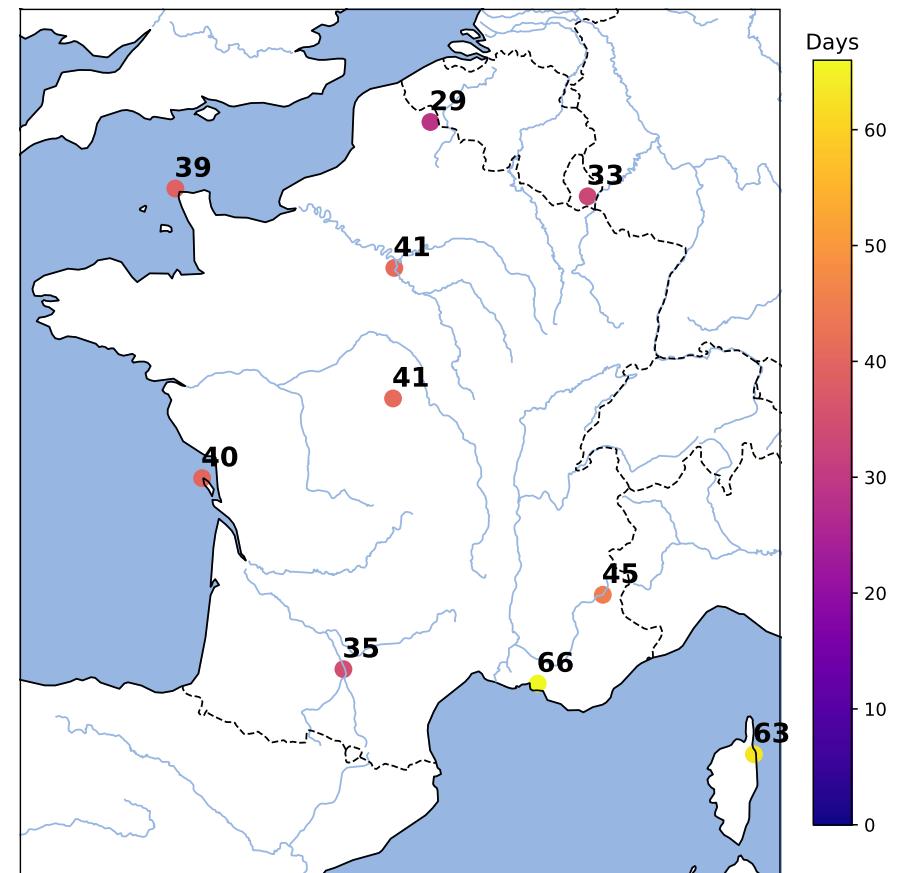
- For each station, each day and each state k: rainfall amount distribution = mixture of Gamma and exponential distributions
  - Station correlations introduced through Gaussian copula



# RECORD OF THE NUMBER CONSECUTIVE DAYS WITHOUT RAIN



Historical data 1956-2019



One 64 year simulation

# CONCLUSION AND PERSPECTIVES

- EDF activities are sensitive to the meteorological conditions
- This implies the need to
  - Estimate extreme weather conditions
  - Generate large samples of possible hydroclimatic variables evolutions to test the system
- Development of:
  - Methodologies to estimate extreme values in taking climate change into account
  - Different weather generators:
    - Local air temperature
    - Local multivariate timeseries: temperature, rainfall, wind
    - Spatial univariate generator
- Future work
  - Spatial (at least) bivariate generator: temperature and rainfall
  - Use to simulate streamflow and its possible changes with climate change to go further in the study of climate change impacts on nuclear generation

# Thank you