Methods for quantifying record probabilities in a changing climate

Paula GONZÁLEZ

Jury:

Pierre RIBEREAU, Université Claude Bernard Lyon 1 (reviewer)
Henning RUST, Freie Universität Berlin (reviewer)
Juliette BLANCHET, Université Grenoble Alpes (examiner)
Daniela CASTRO-CAMILO, University of Glasgow (examiner)
Ester MARIUCCI, Université Paris-Saclay, UVSQ (examiner)
Mathieu RIBATET, Université de Nantes (examiner)
Gwladys TOULEMONDE, Université de Montpellier (examiner)

Supervisors:
Philippe NAVEAU (LSCE), Soulivanh THAO (LSCE) and Julien WORMS (UVSQ)

We observe a record when the last recorded value is the largest

Copernicus Climate Change Service

About Us What we do Data

MONTHLY CLIMATE BULLETIN

The year 2024 set to end up as the warmest on record

Home / News

7th November 2024

It is now virtually certain that the year 2024 will be the warmest in the ERA5 reanalysis dataset, going back to 1940, based on the data available through October. The month was the second-warmest October globally, after October 2023, with an average surface air temperature of 15.25°C, 0.80°C above the 1991-2002 average for the month. October 2024 was 1.65°C above pre-industrial level, marking the 15th month in a 16-month period with average temperatures above the 1.5°C threshold set by the Paris Agreement.

FURTHER READING

Climate Bulletin Home

Climate Bulletin - About the

Le Monde

Se connecter

Planète | Comprendre le réchauffement climatique | 9 indicateurs de l'urgence climatique

Record mondial de températures moyennes pour un mois de juin, selon le service européen Copernicus

Si le thermomètre était proche ou inférieur aux normales de saison en Europe de l'Ouest, notamment en France, une grande partie du monde a subi des températures supérieures aux normales, voire exceptionnelles.

Le Monde avec AFP

Publié le 08 juillet 2024 à 06h04, modifié le 08 juillet 2024 à 07h29 · Lecture 1 min.

Lire plus tard

The Guardian

Eur

Football Tech Business Obituaries

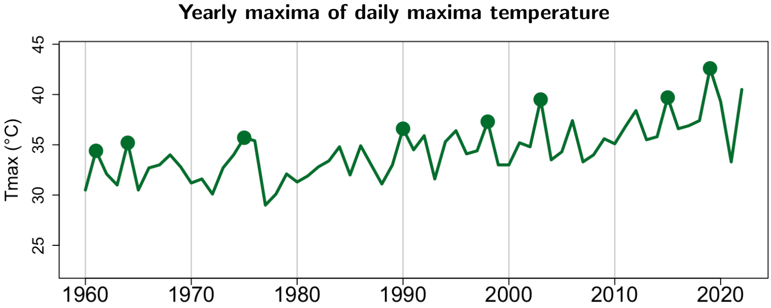
Unprecedented number of heat records broken around world this year

Jonathan Watts
Global environment editor

Wed 14 Aug 2024 12.01 CEST

Share

Records in Paris



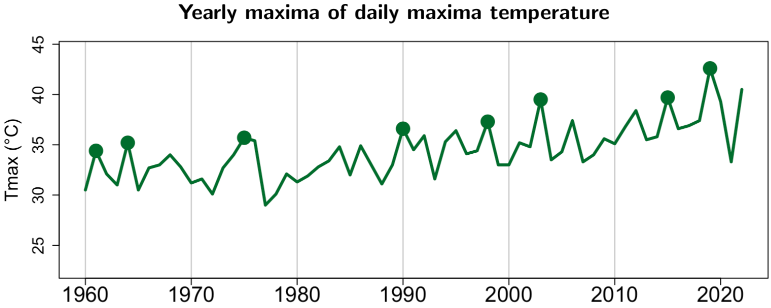
Annual maxima of daily temperature maxima measured at Paris-Montsouris (Paris) station from 1960 to 2022

[1] Chandler (1952)

Records in Paris

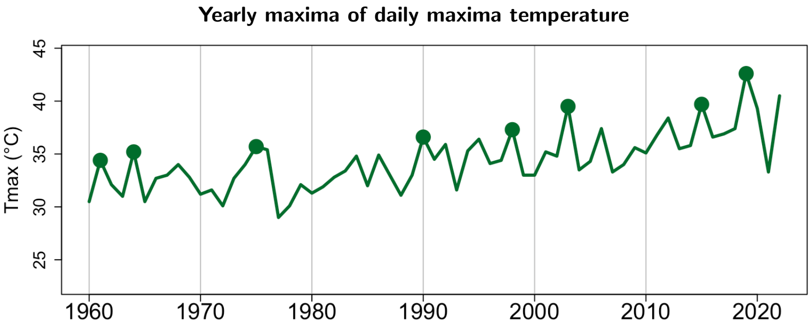
What is a (high) record [1]

$$\{Y_T > \max(Y_1, Y_2, \ldots, Y_{T-1})\}$$



Annual maxima of daily temperature maxima measured at Paris-Montsouris (Paris) station from 1960 to 2022

[1] Chandler (1952)

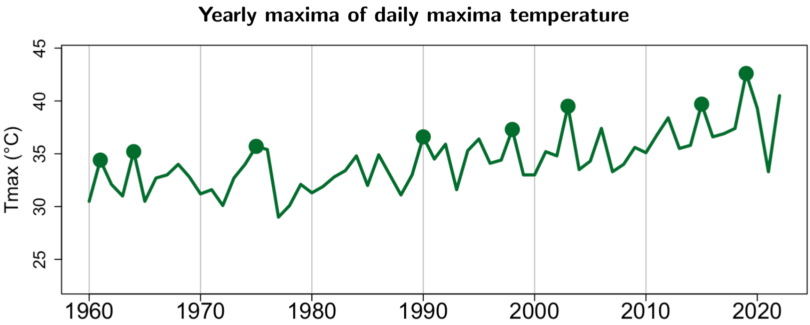


Annual maxima of daily temperature maxima measured at Paris-Montsouris (Paris) station from 1960 to 2022

Questions of study

Q1 : What are the odds, that a given year, e.g., 2030, could have beaten a 50-year counterfactual record?

Q2 :

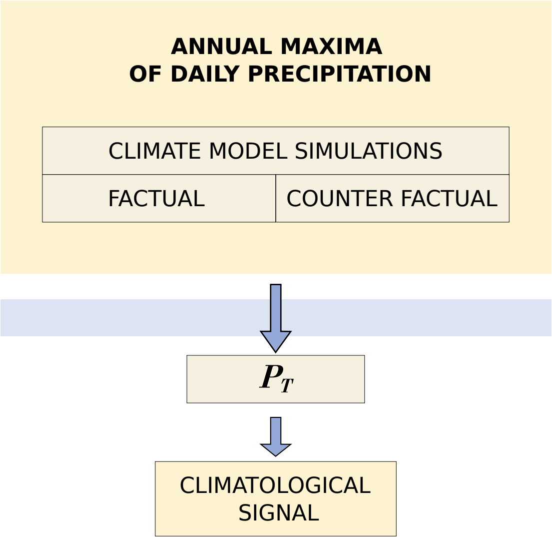


Annual maxima of daily temperature maxima measured at Paris-Montsouris (Paris) station from 1960 to 2022

Questions of study

Q1 : What are the odds, that a given year, e.g., 2030, could have beaten a 50-year counterfactual record?
Q2 : What is the probability of observing a record in 2023 in the center of Paris given measurement around Paris recorded up to 2022?

Q1 : What are the odds, that a given year, e.g., 2030, could have beaten a 50-year counterfactual record?



Q1 : What are the odds, that a given year, e.g., 2030, could have beaten a 50-year counterfactual record?

ANNUAL MAXIMA OF DAILY PRECIPITATION

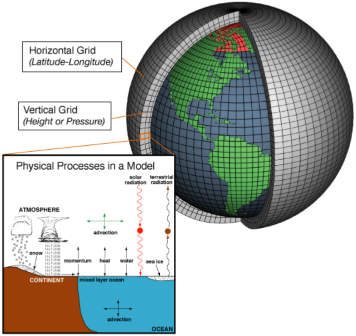
CLIMATE MODEL SIMULATIONS	
FACTUAL	COUNTER FACTUAL

↓

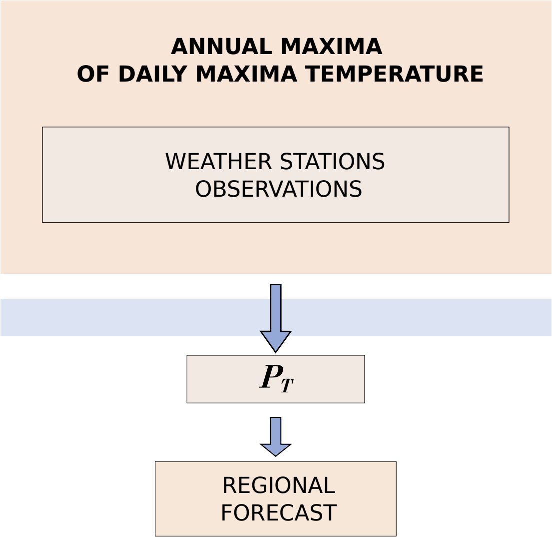
P_T

↓

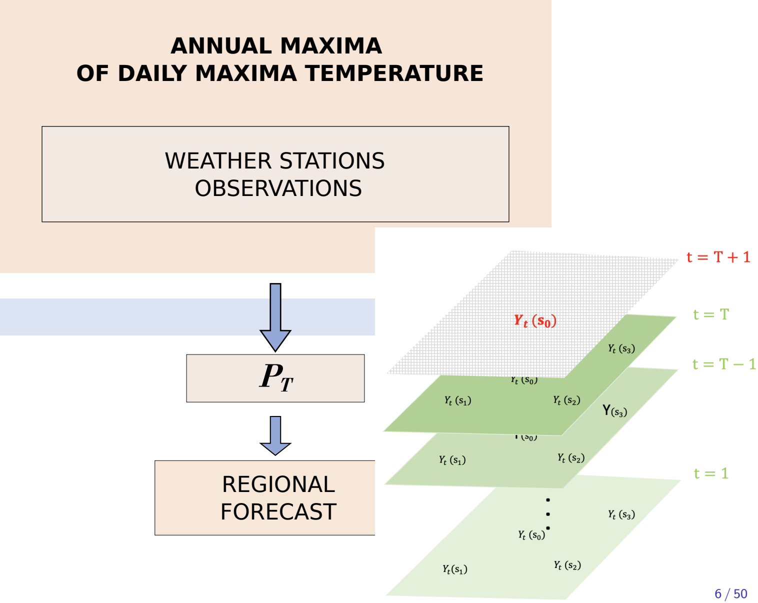
CLIMATOLOGICAL SIGNAL

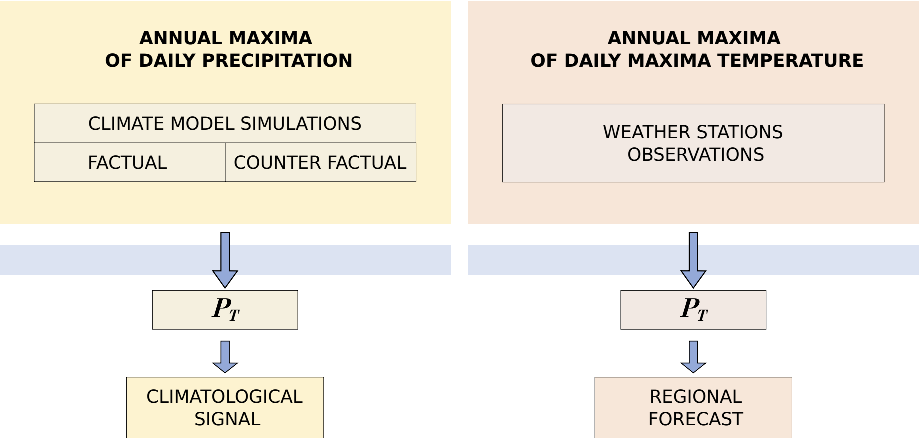


Q2 : What is the probability of observing a record in 2023 in the center of Paris given measurement around Paris recorded up to 2022?



Q2 : What is the probability of observing a record in 2023 in the center of Paris given measurement around Paris recorded up to 2022?

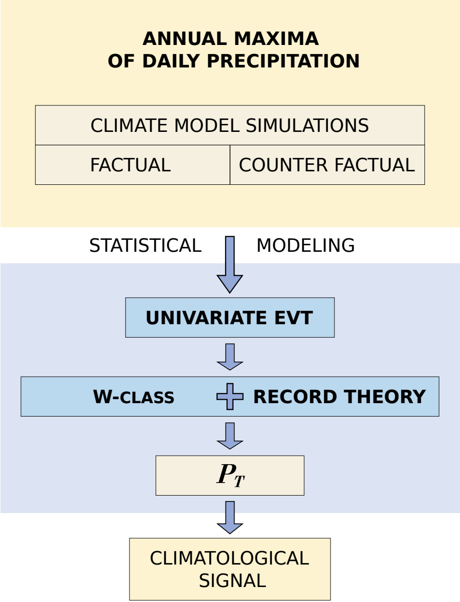



ANNUAL MAXIMA
OF DAILY PRECIPITATION
CLIMATE MODEL SIMULATIONS
FACTUAL
COUNTER FACTUAL
P_T
CLIMATOLOGICAL
SIGNAL
ANNUAL MAXIMA
OF DAILY MAXIMA TEMPERATURE
WEATHER STATIONS
OBSERVATIONS
P_T
REGIONAL
FORECAST

Limitations that we need to address

- Statistical modeling of record probability from yearly maxima and its evolution through time.
- Inference of a climate signal on record probability at a specific location.
- Incorporation of spatial information through the proper modeling of the dependencies among time series.

Q1 : What are the odds, that a given year, e.g., 2030, could have beaten a 50-year counterfactual record?



Extreme value theory for annual maxima [2]

Annual maxima distribution

$M_n = \max\{X_1, \ldots, X_n\}$ of n i.i.d. R.V

$$(M_n - b_n)/a_n \xrightarrow{d} GEV(\mu, \sigma, \xi) \quad \text{when} \quad n \to \infty$$

[2] see, e.g. Statistics of Extremes, Davison and Huser Annual Review of Statistics and Its Application 2015 2:1, 203-235

[3] Cooley et al. (2007), Kharin and Zwiers (2005)

Extreme value theory for annual maxima [2]

Annual maxima distribution

$M_n = \max\{X_1, \ldots, X_n\}$ of n i.i.d. R.V

$$(M_n - b_n)/a_n \xrightarrow{d} GEV(\mu, \sigma, \xi) \quad \text{when} \quad n \to \infty$$

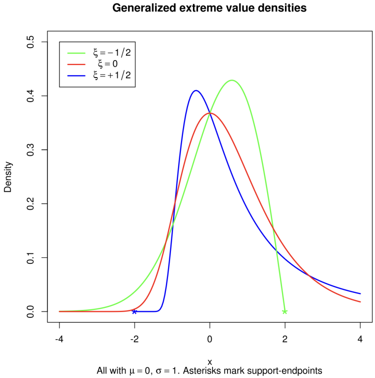
- $\xi > 0$

  e.g. Yearly maxima of daily precipitation [3]

- $\xi = 0$

- $\xi < 0$

  e.g. Yearly maxima of daily maxima temperature [3]



[2] see, e.g. Statistics of Extremes, Davison and Huser Annual Review of Statistics and Its Application 2015 2:1, 203-235

[3] Cooley et al. (2007), Kharin and Zwiers (2005)

W-class

Characterizing records by the relative behavior [4]

$Z \perp X$ random variables with $G(x) = P(X \leq x)$

$$P(Z > \max\{X_1, X_2, \dots, X_r\}) = \mathbb{E}(exp(-rW)).$$

$$\text{with} \quad W = -\log G(Z)$$

[4] Naveau et al. (2018)

[5] Worms and Naveau (2022)

W-class

Characterizing records by the relative behavior [4]

$Z \perp X$ random variables with $G(x) = P(X \leq x)$

$$P(Z > \max\{X_1, X_2, \ldots, X_r\}) = \mathbb{E}(exp(-rW)).$$

$$\text{with} \quad W = -\log G(Z)$$

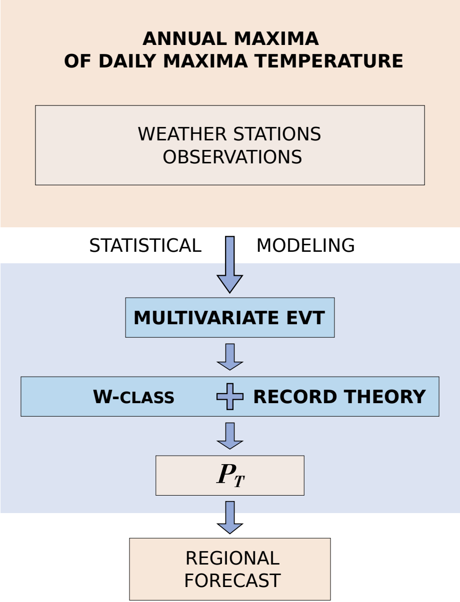
Definition (W-class) [5]

] Let $X \sim G$ and $Z \sim F$ be two random variables with the same support. We say that the couple (X, Z) belongs to the W-class if the random variables $W = -\log G(Z)$ is Weibull distributed with parameters (k, λ), i.e,

$$P(W > w) = \exp(-(w/\lambda)^k) \quad \text{for any} \quad w > 0$$

[4] Naveau et al. (2018)
[5] Worms and Naveau (2022)

Q2 : What is the probability of observing a record in 2023 in the center of Paris given measurement around Paris recorded up to 2022?



EVT for componentwise maxima[6]

If we consider the componentewise maximum $\mathbf{M}_n = (M_{n;1}, \ldots, M_{n;d})^T$ of n i.i.d. d-dimensional random vectors $\mathbf{X}$

Multivariate block maxima distribution (with unit-Fréchet marginals)

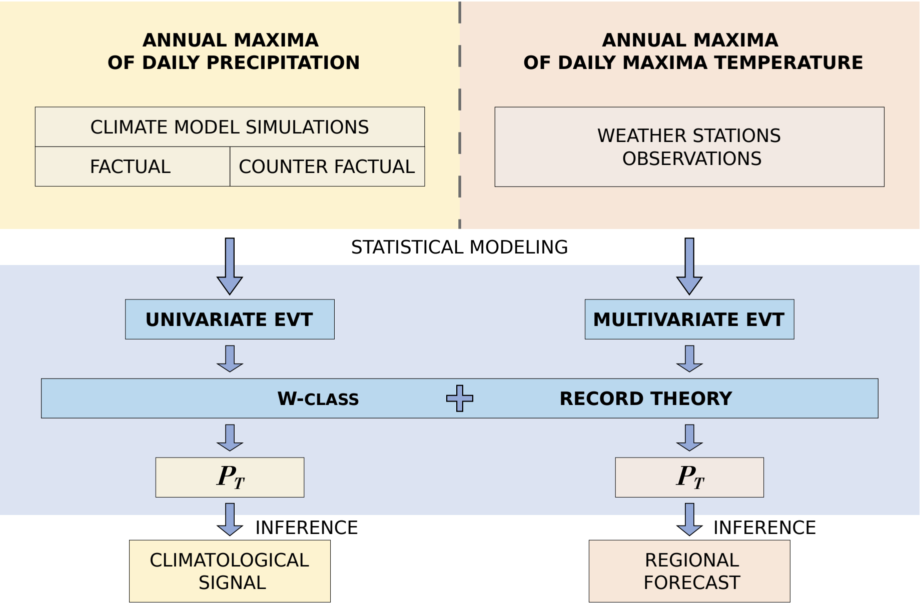
$$\frac{\mathbf{M}_n - b_n}{a_n} \xrightarrow{d} MGEV(\mathbf{x}) = \exp\{-V(\mathbf{x})\} \quad \text{when} \quad n \to \infty$$

with $\mathbf{x} = (x_1, \ldots, x_d)^T$ and
$V(\mathbf{x})$ a positive homogenous function of order -1, i.e. $V(c\mathbf{x}) = c^{-1}V(\mathbf{x})$

[6] *see, e.g. Fougères (2004), SpatialExtremes package (R)*

Q1 : What are the odds, that a given year, e.g., 2030, could have beaten a 50-year counterfactual record?

Q2 : What is the probability of observing a record in 2023 in the center of Paris given measurement around Paris recorded up to 2022?

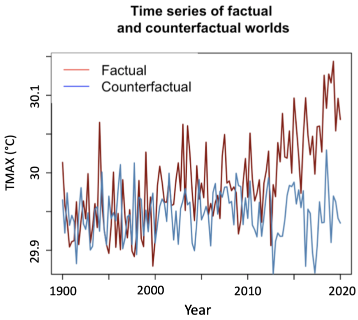


A statistical method to model non-stationarity in precipitation records changes

(Part 1)

Accepted in Geophysical Research Letters (Gonzalez et al., 2024).

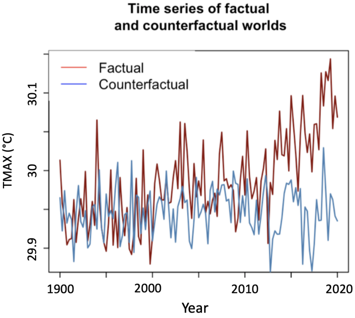
Preamble on Extreme Event Attribution (EEA) [7]



Counterfactual [8] X
Factual Z

[7] Scott et al.(2016); Allen (2003)
[8] Hannart et al. (2016), Angélil et al. (2017), Hegerl and Zwiers (2011)

Preamble on Extreme Event Attribution (EEA) [7]

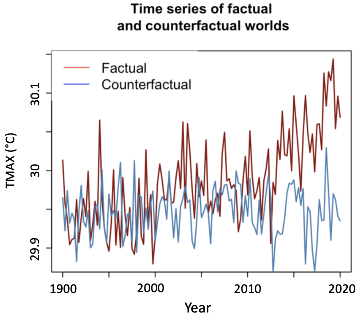


Counterfactual [8] X
Factual Z

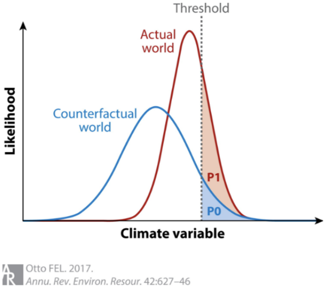
$$p_0 = P(X > u)$$
$$p_1 = P(Z > u)$$

[7] Scott et al.(2016); Allen (2003)
[8] Hannart et al. (2016), Angélil et al. (2017), Hegerl and Zwiers (2011)

Preamble on Extreme Event Attribution (EEA) [7]



Counterfactual [8] X
Factual Z



$$p_0 = P(X > u)$$
$$p_1 = P(Z > u)$$

[7] Scott et al.(2016); Allen (2003)
[8] Hannart et al. (2016), Angélil et al. (2017), Hegerl and Zwiers (2011)

Literature on EEA for records

- EEA of records based on counting record occurrences, e.g King (2017)
- Statistical models in a stationary framework, e.g **Worms and Naveau (2022)**, Naveau et al. (2018)
- Very sparse for non-stationary times series.

Q1

What are the odds that a given year, e.g 2050, could have beaten a r-year counterfactual record?

e.g. $r = 50$

EEA of records

Question

What are the odds that a given year, e.g 2050, could have beaten a r-year counterfactual record?

Event of interest : beat a r-year counterfactual record

being larger that the maxima of the $r-1$ previous counterfactual realizations

EEA of records

Question

What are the odds that a given year, e.g 2050, could have beaten a r-year counterfactual record?

Event of interest : beat a r-year counterfactual record

being larger that the maxima of the $r-1$ previous counterfactual realizations

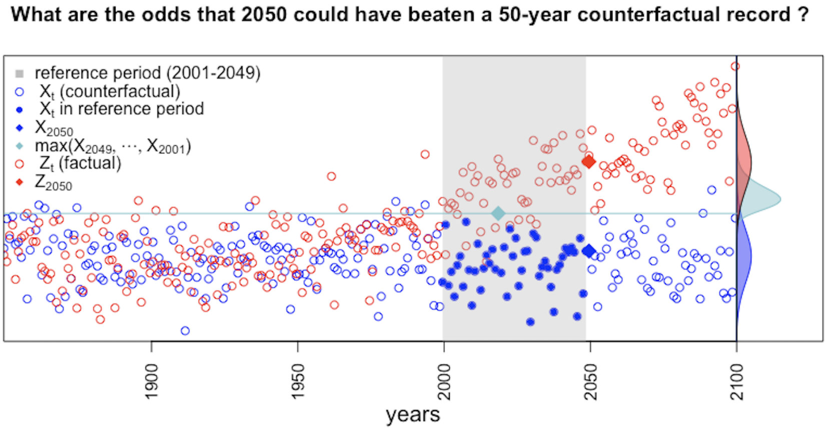
Mathematical definition

$$p_{0,r}(t) = P(X_t > \max(X_{t-1}, \ldots, X_{t-r+1}))$$
$$p_{1,r}(t) = P(Z_t > \max(X_{t-1}, \ldots, X_{t-r+1}))$$

EEA of records

Question

What are the odds that a given year, e.g 2050, could have beaten a r-year counterfactual record?



Record probability in a stationary climate [9]

Record probability in the counterfactual world

$X_{r-1}, \ldots, X_1$ are exchangeable

$$p_{0,r}(t) = P(X_r > \max(X_{r-1}, \ldots, X_1)) = \frac{1}{r}$$

[9] Arnold (1998), Glick(1978)

Attribute changes in records in a transient setup

Inferential objective

$$p_{1,r}(t) = P(Z_t > \max(X_{t-1}, \dots, X_{t-r+1}))$$

Assumptions

- $(X, Z) \in W$-class
- $X \perp Z$, $X_i \perp X_j$, $Z_i \perp Z_j$
- X i.i.d

Attribute changes in records in a transient setup

Inferential objective

$$p_{1,r}(t) = P(Z_t > \max(X_{t-1}, \ldots, X_{t-r+1}))$$

Assumptions

- $(X, Z) \in W$-class
- $X \perp Z$, $X_i \perp X_j$, $Z_i \perp Z_j$
- X i.i.d

Record probability under W-class

$$p_{1,r}(t) = \int_0^1 \exp(-(r-1)\lambda_t(-\log x)^{1/k_t})$$

with $k_t = \xi_{z_t}/\xi_{x_t}$ and $\lambda_t = (k_t \times \sigma_{z_t}/\sigma_{x_t})^{-1/\xi_{x_t}}$

Inference

A plug-in strategy

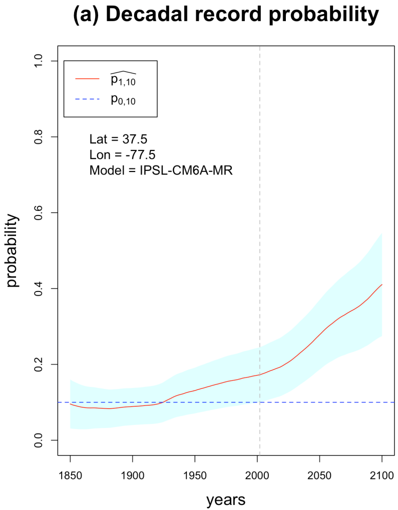
$$\hat{p}_{1,r}(t) = \int_0^1 \exp(-(r-1)\hat{\lambda}_t(-\log x)^{1/\hat{k}_t})$$

- $\hat{k}_t$, $\hat{\lambda}_t$ $\forall t$: Non-parametric Method-of-Moments
- $\hat{p}_{1,r}(t)$ for a chosen r : plug-in

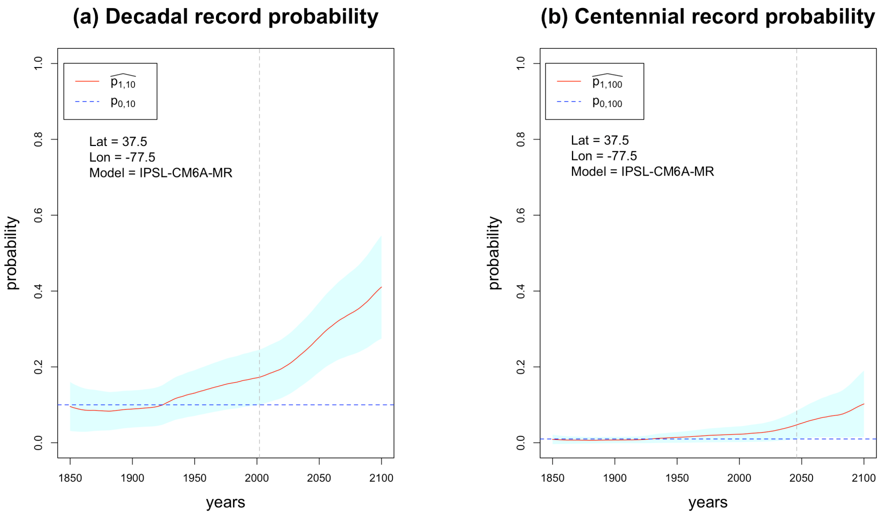
Application study

Yearly maxima of daily precipitation on
CMIP6 IPSL-CM6A-LR and SSP5-8.5

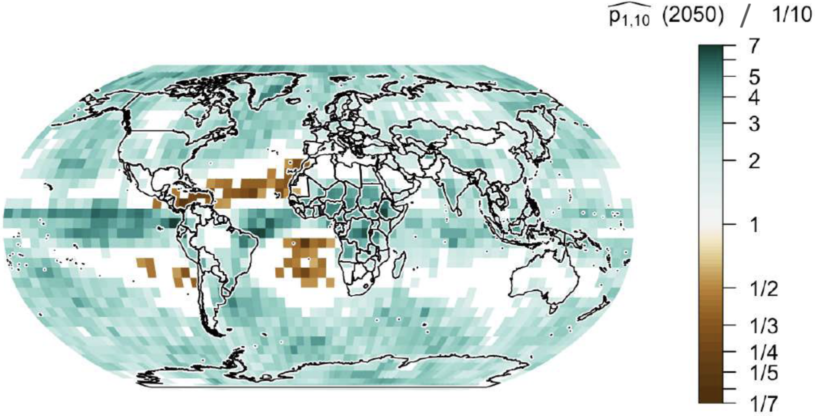
Richmond, Virginia, USA

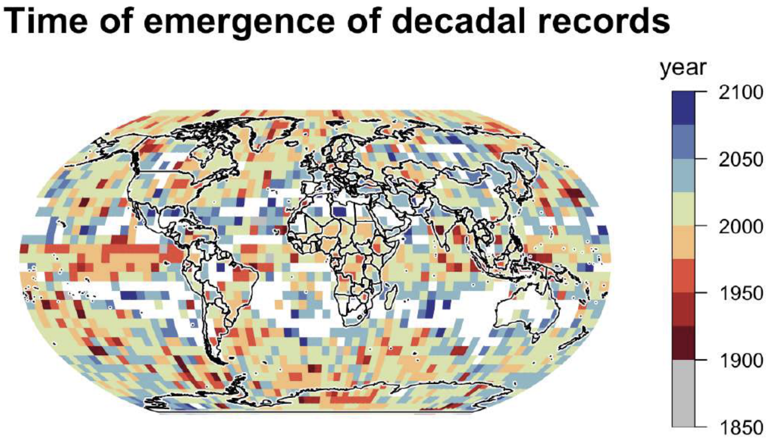


Richmond, Virginia, USA



Probability ratio of decadal records in 2050




Time of emergence of decadal records
year
2100
2050
2000
1950
1900
1850

Summary of part 1

Noticeable statistical features

- **Transient EEA of records**
- Bypassing the estimation of the GEV parameters
- Only 2 parameters need to be estimated
- Allow different tail parameters ξ in marginals
- Kernel smoothing Method-of moments full captures non-linear trends

Climatological records

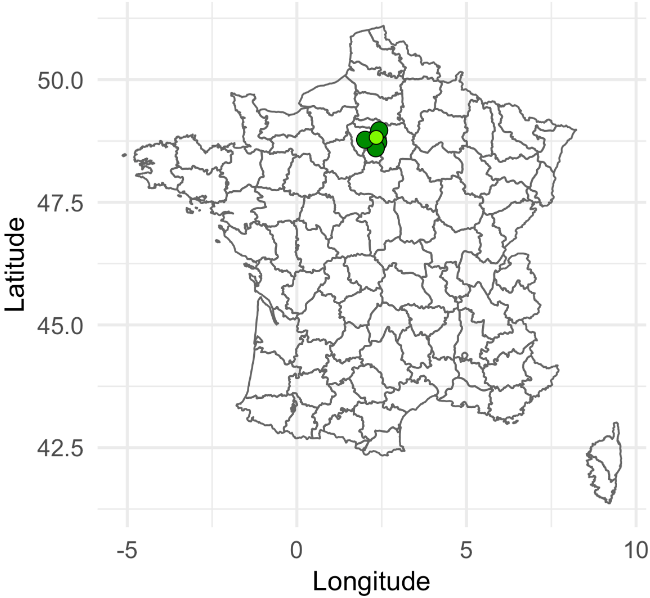
- Precipitation records are affected all over the world
- Consistent with previous studies on changes of precipitation.

Forecasting yearly record probabilities from historical measurements

(Part 2)

Shortly deposited on HAL (Gonzalez et al., 2025).

Records in Paris over a region



Definition of the event of study

Local records over a given region

$$\left\{ Y_{T+1}(s_o) > \max_{s \in \mathbb{S}} \max_{t=1,\dots,T} Y_t(s) \right\}$$

Definition of the event of study

Local records over a given region

$$\left\{ Y_{T+1}(s_o) > \max_{s \in \mathbb{S}} \max_{t=1,\dots,T} Y_t(s) \right\}$$

Notation

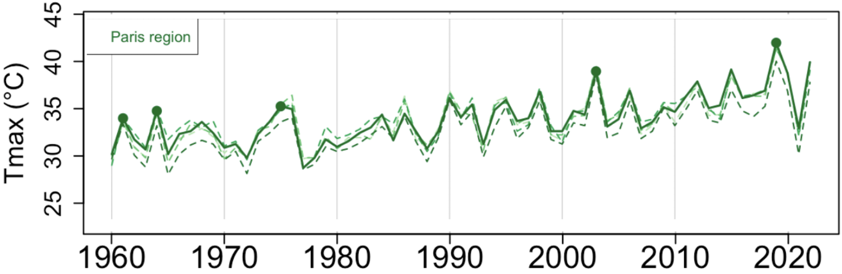
$\mathbb{S}$: ensemble of sites

s_o: site of interest

T: extent of the reference period

$Y_t(s)$: R.V associated to time time t and site s

Records in Paris over a region



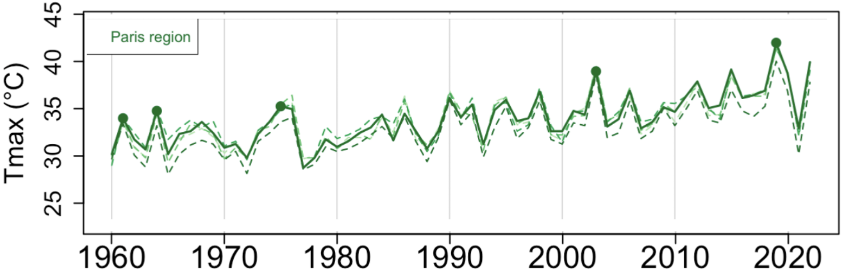
Annual maxima of daily temperature maxima measured at five stations around Paris from 1960 to 2022

Records in Paris over a region

i) Non-linear warming trend

ii) Correlated time-series

iii) Similar range of values



Annual maxima of daily temperature maxima measured at five stations around Paris from 1960 to 2022

Q2

What is the probability, at a given site, of observing a record over a given region in the upcoming year?

e.g. What is the probability of observing a record in 2023 in the center of Paris (Montsouris station) given measurement around Paris recorded up to 2022 ?

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s\in\mathbb{S}} \max_{t=1,\ldots,T} Y_t(s)\right)$$

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s \in \mathbb{S}} \max_{t=1,\dots,T} Y_t(s)\right)$$

Record probability for i.i.d. R.V

$$P_{T+1}^{(i.i.d.)} = \frac{1}{1 + T \times S} = \frac{1}{\#\text{ of random variables being compared}}$$

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s \in \mathbb{S}} \max_{t=1,\dots,T} Y_t(s)\right)$$

Record probability for i.i.d. R.V

$$P_{T+1}^{(i.i.d.)} = \frac{1}{1 + T \times S} = \frac{1}{\text{\# of random variables being compared}}$$

Literature on statistical models for record probabilities

- Sparse for non-stationary times series, e.g. Ballerini and Resnick (1987), Wergen and Krug (2010)
- Very sparse on multivariate times series, e.g. Godreche and Luck (2023)
- Very very sparse on forecasting of record probabilities in a non-linear multivariate context

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s \in \mathbb{S}} \max_{t=1,\ldots,T} Y_t(s)\right)$$

Our four assumptions

(i) All marginals of $Y_t(s)$ follow GEV distributions

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s \in \mathbb{S}} \max_{t=1,\dots,T} Y_t(s)\right)$$

Our four assumptions

(i) All marginals of $Y_t(s)$ follow GEV distributions

(ii) The region $\mathbb{S}$ is *homogeneous*, i.e. both the tail index and the support of any $Y_t(s)$ are constant[a] in time and over $\mathbb{S}$.

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s \in \mathbb{S}} \max_{t=1,\dots,T} Y_t(s)\right)$$

Our four assumptions

(i) All marginals of $Y_t(s)$ follow GEV distributions

(ii) The region $\mathbb{S}$ is *homogeneous*, i.e. both the tail index and the support of any $Y_t(s)$ are constant[a] in time and over $\mathbb{S}$.

(iii) For any $s \in \mathbb{S}$, all variables $\{Y_t(s)\}$ independent in time

Inferential objective

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s\in\mathbb{S}} \max_{t=1,\dots,T} Y_t(s)\right)$$

Our four assumptions

- **(i)** All marginals of $Y_t(s)$ follow GEV distributions
- **(ii)** The region $\mathbb{S}$ is *homogeneous*, i.e. both the tail index and the support of any $Y_t(s)$ are constant[a] in time and over $\mathbb{S}$.
- **(iii)** For any $s \in \mathbb{S}$, all variables $\{Y_t(s)\}$ independent in time
- **(iv)** The spatial random vector $\{Y_t(s)\}_{s\in\mathbb{S}}$ follows a max-stable multivariate distribution with time-invariant dependence $V(.)$

[a] Location $\mu_t(s)$ and scale $\sigma_t(s)$ parameters can vary in space and time

Statistical model for $P_{T+1}(s_o)$

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s\in\mathbb{S}} \max_{t=1,\dots,T} Y_t(s)\right)$$

Main result

Under assumptions **(i)**-**(iv)**,

$$P_{T+1}(s_0) \quad = \quad \frac{1}{1+\sum_{t\leq T} V\left(u_{t,T+1}(s_0, s_1), \dots, u_{t,T+1}(s_0, s_S)\right)}$$

where $V(.)$ is the spatial max-stable dependence function and

$$u_{t,T+1}(s_0, s) = \left(\frac{\sigma_{T+1}(s_0)}{\sigma_t(s)}\right)^{1/\xi}$$

Statistical model for $P_{T+1}(s_o)$

$$P_{T+1}(s_o) = P\left(Y_{T+1}(s_o) \geq \max_{s \in \mathbb{S}} \max_{t=1,\ldots,T} Y_t(s)\right)$$

Main result

Under assumptions **(i)**-**(iv)**,

$$P_{T+1}(s_0) \quad = \quad \frac{1}{1 + \sum_{t \leq T} V\left(u_{t,T+1}(s_0, s_1), \ldots, u_{t,T+1}(s_0, s_S)\right)}$$

where $V(.)$ is the spatial max-stable dependence function and

$$u_{t,T+1}(s_0, s) = \left(\frac{\sigma_{T+1}(s_0)}{\sigma_t(s)}\right)^{1/\xi}$$

Special cases

If all same margins are equal in time (no warming trends),

$$P_{T+1}^{V(\mathbf{1})} \quad = \quad 1/(1 + T \times V(\mathbf{1})),$$

with $V(\mathbf{1}) = S$ (the i.i.d. case) and $V(\mathbf{1}) = 1$ (the perfect dependence)

Illustration example

$$Y_t(s) \sim GEV(\mu_t = 26 + t(2^{-3}) + t^2(10^{-3}), \sigma_t, \xi = -0.2) \ \& V = \text{logistic}(\alpha = .3)$$

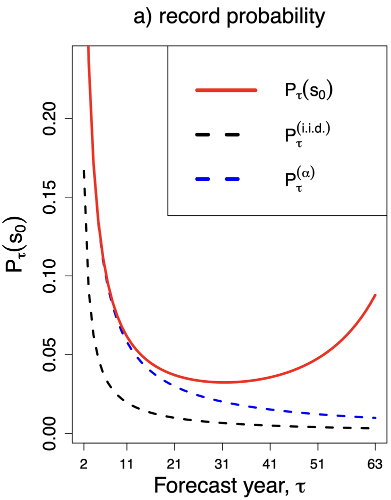
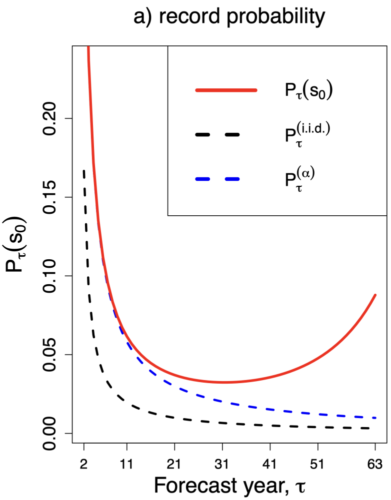
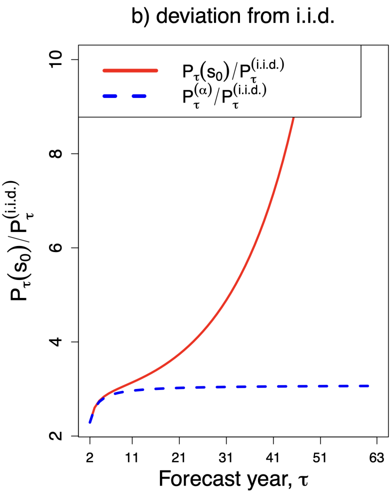


Illustration example

$$Y_t(s) \sim GEV(\mu_t = 26 + t(2^{-3}) + t^2(10^{-3}), \sigma_t, \xi = -0.2) \ \& V = \text{logistic}(\alpha = .3)$$





Non-parametric Inference

A plug-in strategy

(I) For every $\tau \leq T$

$$\hat{P}_\tau(s_0) = \frac{1}{1 + \sum_{t \leq T} \hat{V}\left(\hat{u}_{t,\tau}(s_0, s_1), \ldots, \hat{u}_{t,\tau}(s_0, s_S)\right)}$$

- $\hat{V}$: Madogram based approach[10]
- $\hat{u}_{t,\tau}(s_0, s_1)$: Method-Of-Moments as $F_t^{(s)}(Y_\tau(s_0)) \overset{d}{=} Beta(u_{t,\tau}(s_0, s), 1)$

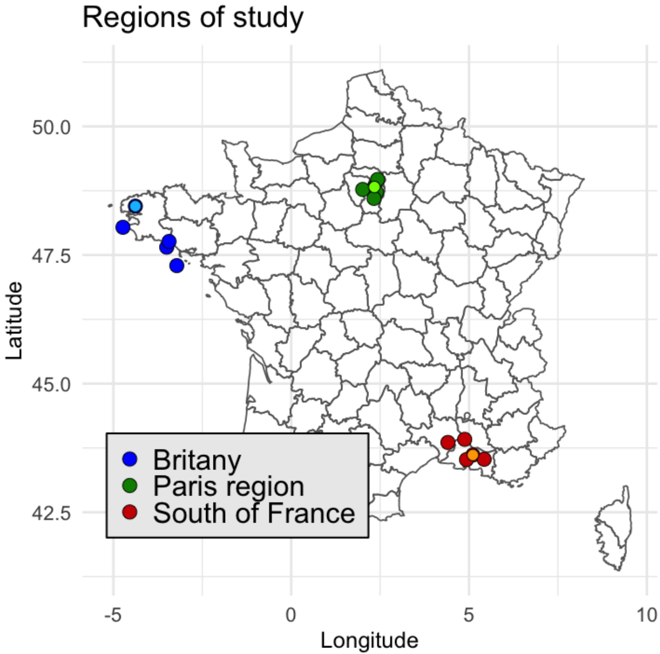
(II) Forecast[11] $\hat{P}_{T+1}(s_0)$ from $\{\hat{P}_\tau(s_0)\}_{\tau \leq T}$

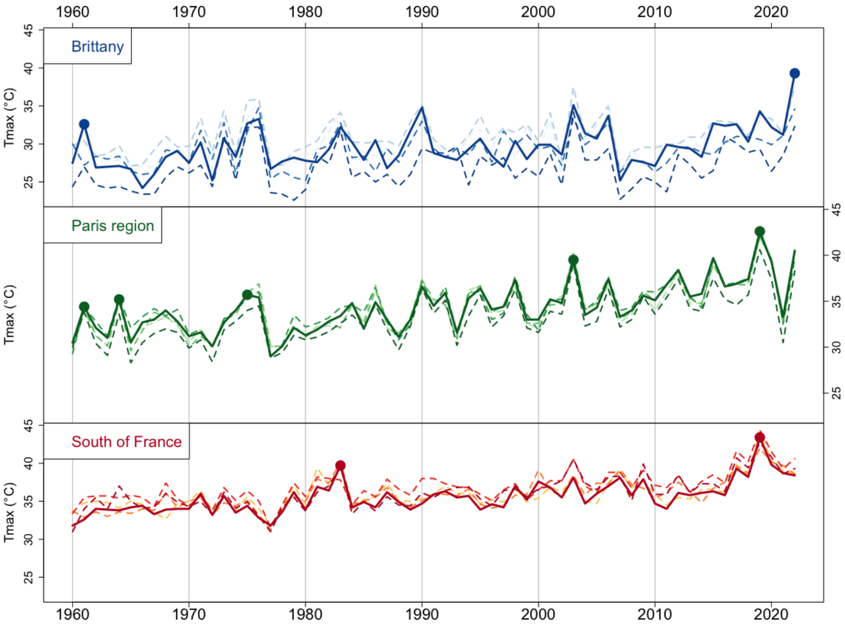
[10] Marcon et al. (2017)

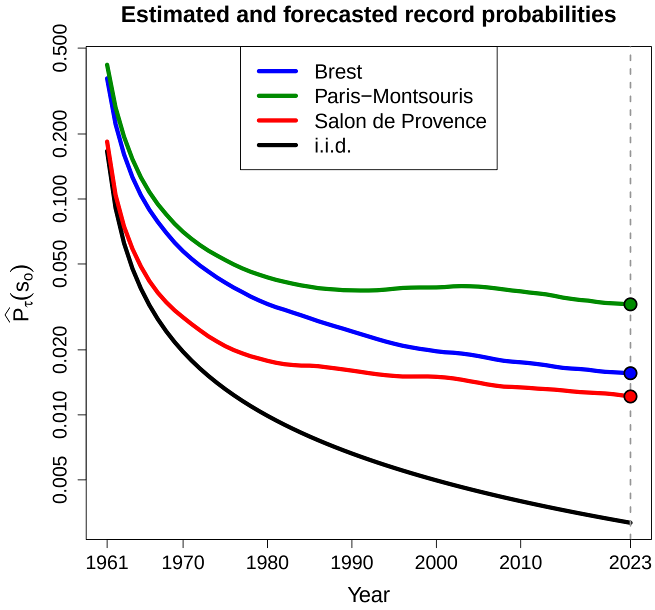
[11] e.g. second order exponential smoothing, see Winters (1960), Gardner and Dannenbring (1980)

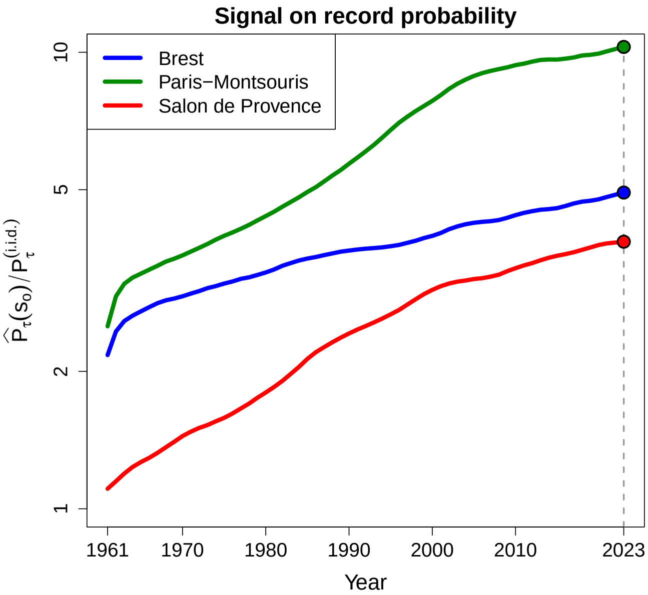
Application study

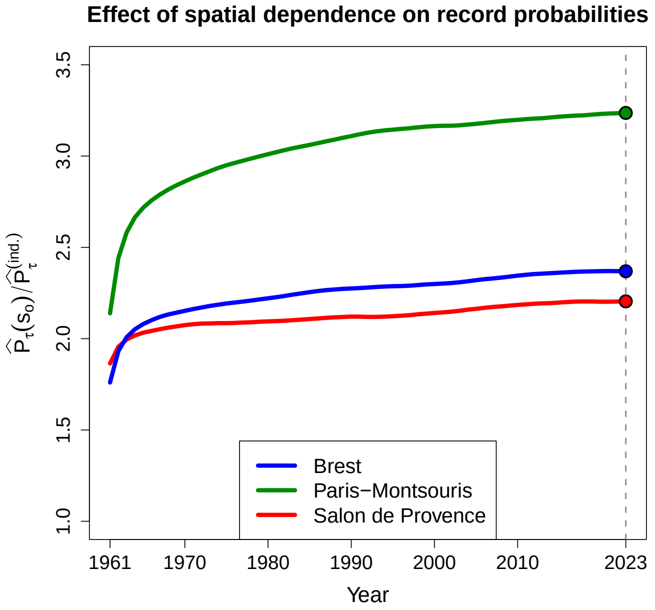
Yearly maxima of daily maxima temperature
from weather stations measurements in France


Regions of study
50.0
47.5
45.0
42.5
Latitude
-5
0
5
10
Longitude
Britany
Paris region
South of France


Brittany
Paris region
South of France
Tmax (°C)
1960
1970
1980
1990
2000
2010
2020
25
30
35
40
45


Estimated and forecasted record probabilities
Brest
Paris−Montsouris
Salon de Provence
i.i.d.
$\widehat{P}_\tau(s_0)$
0.500
0.200
0.100
0.050
0.020
0.010
0.005
1961
1970
1980
1990
2000
2010
2023
Year


Signal on record probability
Brest
Paris−Montsouris
Salon de Provence
10
5
2
1
$\widehat{P}_\tau(s_0)/P_\tau^{(i.i.d.)}$
1961
1970
1980
1990
2000
2010
2023
Year


Effect of spatial dependence on record probabilities
$\widehat{P}_\tau(s_0)/\widehat{P}_\tau^{(ind.)}$
3.5
3.0
2.5
2.0
1.5
1.0
1961
1970
1980
1990
2000
2010
2023
Year
Brest
Paris−Montsouris
Salon de Provence

Summary of part 2

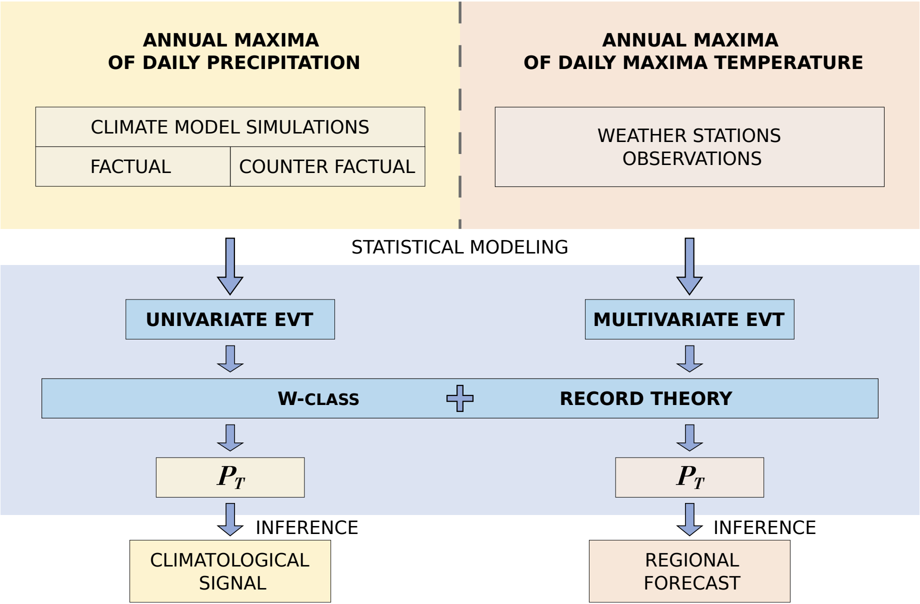
Noticeable statistical features

- **Records modeling in a non-stationary and multivariate context**
- Bypassing the estimation of the three GEV parameters
- Non-parametric max-stable structure
- Fast implementations

Climatological records

- Inter-annual forecasting of record probabilities seems possible
- Both trends and dependence strength play a role.
- Negligible records under a stationary climate becomes significant (up to 8 times)

Summary of the thesis


ANNUAL MAXIMA
OF DAILY PRECIPITATION
CLIMATE MODEL SIMULATIONS
FACTUAL
COUNTER FACTUAL
ANNUAL MAXIMA
OF DAILY MAXIMA TEMPERATURE
WEATHER STATIONS
OBSERVATIONS
STATISTICAL MODELING
UNIVARIATE EVT
MULTIVARIATE EVT
W-CLASS
RECORD THEORY
P_T
P_T
INFERENCE
INFERENCE
CLIMATOLOGICAL
SIGNAL
REGIONAL
FORECAST

Summary of the thesis

Two methodologies for the quantification of record probabilities from non-stationary yearly maxima

Summary of the thesis

Two methodologies for the quantification of record probabilities from non-stationary yearly maxima

Global statistical features

- Take advantage of the relative nature of records
- Comprehensively account for non-linear trends
- Capture signals and their inter-annual changes from relatively short sample sizes

Summary of the thesis

Two methodologies for the quantification of record probabilities from non-stationary yearly maxima

Global statistical features

- Take advantage of the relative nature of records
- Comprehensively account for non-linear trends
- Capture signals and their inter-annual changes from relatively short sample sizes

Limitations and perspectives (Q1 & Q2)

- No information about the magnitude of records

- Guidelines of the World Weather Attribution (WWA) protocol :
  - Multi-model aggregation.
  - Bias correction.
  - Assimilation of observations.

- Uncertainty of the estimator.
- Move from inter-annual to multi-decadal forecast

Thank you

Additional material

Additional material

Part 1

- Estimation algorithm for $p_{1,r}(t)$
- Validity of the Weibullity assumption
- Asymptotic confidence intervals
- Application on centennial records

Part 2

- Estimation of the dependence structure $V(.)$
- Estimation of parameters $u_{t,\tau}(s_0, s)$
- Validity of the Homogeneous region assumption

Under Wclass assumption: $p_{1,r}(t) = \int_0^1 \exp(-(r-1)\lambda_t(-\log x)^{1/k_t})\, dx$

Estimation algorithm:

Step 1[12]: $\forall\, t$

$$\widehat{p}_{1,2}(t) = \mathbb{E}(\mathbb{G}(Z_t)) = \sum_{j=1}^{J} \frac{K_h(t-t_j)}{\sum_{l=1}^{J} K_h(t-t_l)} \mathbb{G}(Z_{t_j})$$

$$\widehat{p}_{1,3}(t) = \mathbb{E}(\mathbb{G}^2(Z_t)) = \sum_{j=1}^{J} \frac{K_h(t-t_j)}{\sum_{l=1}^{J} K_h(t-t_l)} \mathbb{G}^2(Z_{t_j})$$

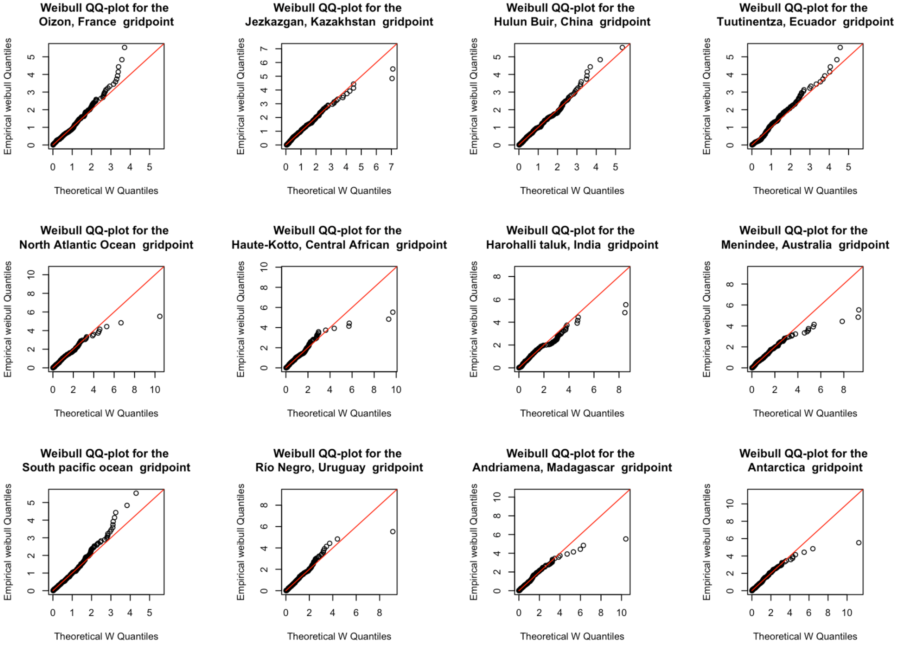
Step 2: for a chosen t

$$\begin{cases} \widehat{p}_{1,2}(t) = \int_0^1 \exp(-\hat{\lambda}_t(-\log x)^{1/\hat{k}_t})\, dx \\ \widehat{p}_{1,3}(t) = \int_0^1 \exp(-2\hat{\lambda}_t(-\log x)^{1/\hat{k}_t})\, dx\,. \end{cases} \rightsquigarrow (\hat{k}_t, \hat{\lambda}_t)$$

Step 3: $\widehat{p}_{1,r}(t) \leftarrow \int_0^1 \exp(-(r-1)\hat{\lambda}_t(-\log x)^{1/\hat{k}_t})\, dx$

[12] Naveau and Thao (2022)

Applying the transformation $T_t(w) = (w/\lambda_t)^{k_t}$



Asymptotic confidence intervals

When I and J go to infinity, if $\sqrt{J/I}$ converges to some finite constant, then for any X_t and Z_t belonging to the W-class and any fixed $r \geq 3$, the parametric estimator $\widehat{p}_{1,r}(t)$ satisfies

$$\sqrt{J}\,\frac{\widehat{p}_{1,r}(t) - p_{1,r}(t)}{\sigma_{rt}} \overset{d}{\approx} \mathcal{N}(0,1)$$

Then, we can compute the confidence intervals as follows

$$[\widehat{p}_{1,r}(t) \pm z_\alpha \widehat{\sigma}_{rt}]$$

with $\widehat{\sigma}_{rt}$ the estimation of the asymptotic standard deviation σ_{rt} and z_α the significance level.

Asymptotic confidence intervals

When I and J go to infinity, if $\sqrt{J/I}$ converges to some finite constant, then for any X_t and Z_t belonging to the W-class and any fixed $r \geq 3$, the parametric estimator $\widehat{p}_{1,r}(t)$ satisfies

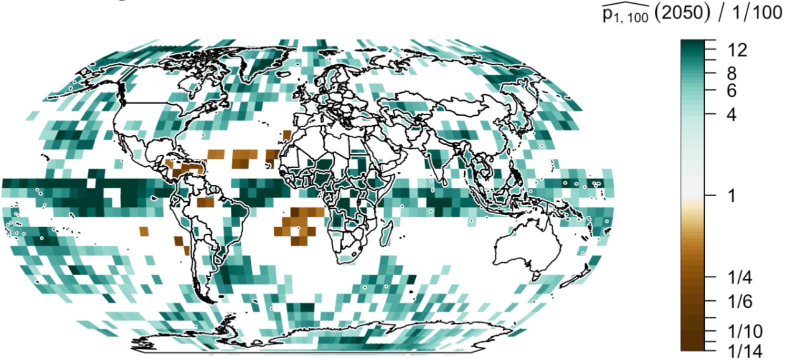
$$\sqrt{J}\,\frac{\widehat{p}_{1,r}(t) - p_{1,r}(t)}{\sigma_{rt}} \overset{d}{\approx} \mathcal{N}(0,1)$$

with

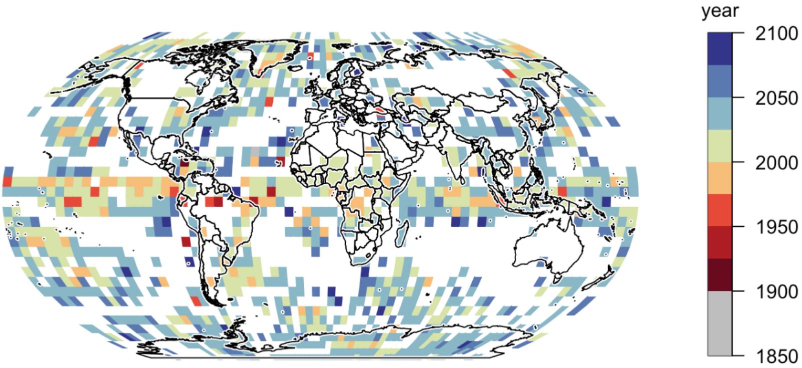
$$\sigma_{rt} = \sqrt{J_{r-1}(\theta_t)(J_{1,2}(\theta_t))^{-1}\Sigma_t\,(J_{1,2}^T(\theta_t))^{-1}(J_{r-1}(\theta_t))^T},$$

- $\theta_t = (\lambda_t, k_t)$: vector containing the parameters of the Weibull distribution at time t.
- $J_i(\theta_t)$: Jacobian matrix of $g_i(\theta_t) = \int_0^1 \exp(-j\lambda_t(-\log x)^{1/k_t})dx$ at time t for any integer $j \geq 1$
- $J_{1,2}(\theta_t)$: Jacobian matrix associated to $\theta_t \mapsto (g_1(\theta_t), g_2(\theta_t))^T$ at time t.
- Σ_t : Asymptotic covariance matrix of $\begin{pmatrix} \widehat{p}_{1,2}(t) - p_{1,2}(t) \\ \widehat{p}_{1,3}(t) - p_{1,3}(t) \end{pmatrix}$.

a) Probability ratio of centennial records in 2050



b) Time of emergence of centennial records



Non-parametric Inference

A madogram based approach (Marcon et al. 2017)

$$\hat{V}(\mathbf{x}) = \left(\sum_{s=1}^{S} x_s^{-1}\right) \times \frac{\hat{\nu}(\mathbf{w}) + c(\mathbf{w})}{1 - \hat{\nu}(\mathbf{w}) - c(\mathbf{w})},$$

with

$$w_s = x_s^{-1} / \sum_{s=1}^{S} x_s^{-1}$$

where

$$\hat{\nu}(\mathbf{w}) = \frac{1}{T} \sum_{t=1}^{T} \left(\max_{s \in \mathbb{S}} \left\{ \mathbb{F}_t^{(s)}(Y_t(s)) \right\}^{1/w_s} - \frac{1}{S} \sum_{s=1}^{S} \left\{ \mathbb{F}_t^{(s)}(Y_t(s)) \right\}^{1/w_s} \right)$$

$$c(\mathbf{w}) = S^{-1} \sum_{s=1}^{S} w_s / (1 + w_s)$$

and

$$\mathbb{F}_t^{(s)}(y) = \frac{1}{\sum_{k=1}^{T} K_h(t-k)} \sum_{j=1}^{T} K_h(t-j) 1_{Y_j(s) \leq y}.$$

Non-parametric Inference

A plug-in strategy

$$\hat{P}_\tau(s_0) \quad = \quad \frac{1}{1 + \sum_{t \leq T} \hat{V}\left(\hat{u}_{t,\tau}(s_0, s_1), \ldots, \hat{u}_{t,\tau}(s_0, s_S)\right)}$$

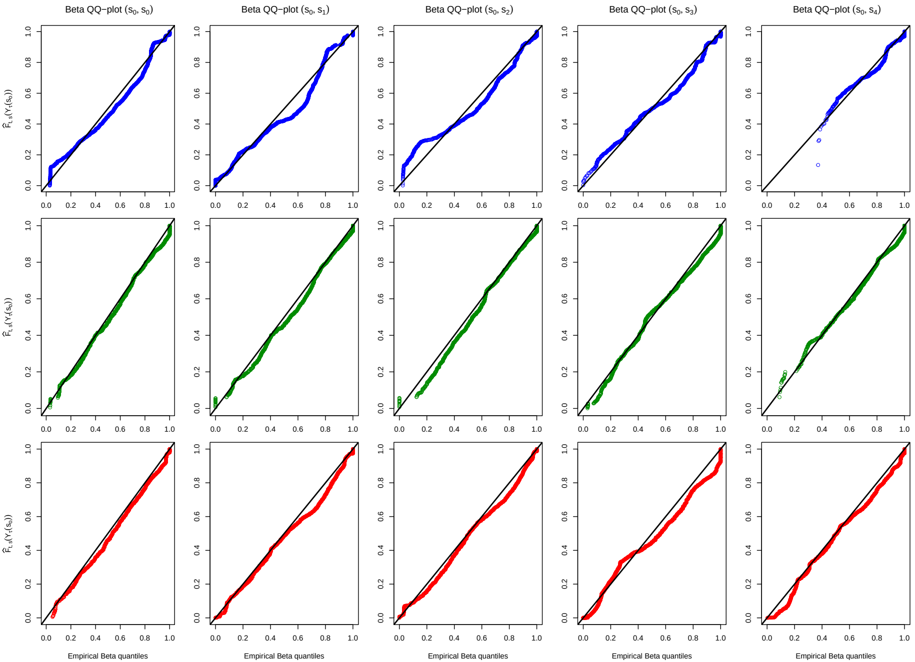
Method-Of-Moments as $F_t^{(s)}(Y_\tau(s_0)) \stackrel{d}{=} Beta(u_{t,\tau}(s_0, s), 1)$

$$\hat{u}_{t,\tau}(s_o, s) = \frac{\widehat{E}_{t,\tau}(s_0, s)}{1 - \widehat{E}_{t,\tau}(s_0, s)},$$

with

$$\widehat{E}_{t,\tau}(s_0, s) = \frac{1}{\sum_{l=1}^{T} K_{h'}(\tau - l)} \sum_{i=1}^{T} K_{h'}(\tau - i)\, \mathbb{F}_t^{(s)}(Y_i(s_o)), \text{ and}$$

$$\mathbb{F}_t^{(s)}(y) = \frac{1}{\sum_{k=1}^{T} K_h(t - k)} \sum_{j=1}^{T} K_h(t - j) 1_{Y_j(s) \leq y},$$


Beta QQ−plot (s0, s0)
Beta QQ−plot (s0, s1)
Beta QQ−plot (s0, s2)
Beta QQ−plot (s0, s3)
Beta QQ−plot (s0, s4)
F̂t,s(Yt(s0))
Empirical Beta quantiles