Températures maximales en France au 21ème siècle

Comité de suivi de première année - 2023

Encadrants: Philippe NAVEAU, Nathalie BERTRAND, Aurélien RIBES

Occitane Barbaux (occitane.barbaux@umr-cnrm.fr)

12 Juillet 2023

ÉDUCATION

ENSAE | INGÉNIEUR MASTER 2

Diplômée en Novembre 2020 | Paris, France

Diplômée de l'École Nationale de la Statistique et de l'Administration Économique (ENSAE), **voie Machine Learning, Statistiques et Apprentissages**.

EXPERIENCE

IRSN | SOUTIEN STATISTIQUE AU GROUPE DE TRAVAIL VENT ET NEIGE | INTÉRIM 18 MOIS

Nov 2020 – 22 Mai 2022 | BEHRIG | Fontenay-aux-roses, France

IRSN | STAGIAIRE MODÉLISATION DE VALEURS EXTRÊMES

Mai 2020 – Sep 2020 | BEHRIG | Fontenay-aux-roses, France

Stage de **modélisation des vents extrême**.

UNIVERSITY OF CAMBRIDGE | STAGIAIRE BIOSTATISTIQUES

Jan 2019 – Jul 2019 | Cardiovascular Epidemiology Unit | Cambridge, UK

Stage de recherche sur les facteurs de risques cardio-vasculaires.

Formations suivies :

- Stage ”Changement Climatique” de Météo France. Octobre 2022.
- École d’été ”Traitement des Données Massives et Apprentissage: Applications en Géophysique, Écologie et SHS” - Juin 2023
- Semaine des Doctorants de l’ED 129. (Science Ouverte, Dévelopment Durable, Ethique) - Avril 2023
- MOOC ”Ethique de la recherche et intégrité scientifique” - Janvier 2023

Présentation des travaux :

- Journées des doctorants IRSN (Poster) - Mars 2023
- Valpred 2023 (Poster) - Avril 2023
- Semaine des Doctorants de l'ED 129 (Poster) - Avril 2023
- Extreme Value Analysis 2023 (Poster) - Juin 2023

Enseignement :

- "Classification Non-supervisée" à l'École Nationale de la Météo (4h de cours et 9h de TD) - Janvier / Février 2023

Le réchauffement climatique:

- Rapport du GIEC sorti en 2022 : **Réchauffement climatique global avéré** d'environ 1°C.
- Une **augmentation de la fréquence et la sévérité** des périodes de températures extrêmes. Ex: 2003, 2006 et record national de 45°C en 2019.
- Une meilleur connaissance du phénomène : **Chroniques d'observations et modèles de climats**.
- Pour l'adaptation, besoin de **projections à l'échelle locale**.

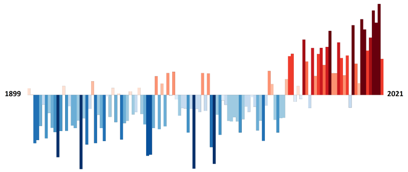


Figure: Anomalie de Température moyenne annuelle en France entre 1850 et 2022. D'après: #ShowYourStripes

Divers enjeux :

- **Tenue des équipements importants** pour la sûreté nucléaire.
- Impact sur la **santé humaine**.
- Certaines normes dont l'Eurocode utilisant le Q50, une température non-stationnaire peut conduire à des **problèmes lors des constructions**.

$\Longrightarrow$ Nécessité de **choisir un indicateur adapté**.

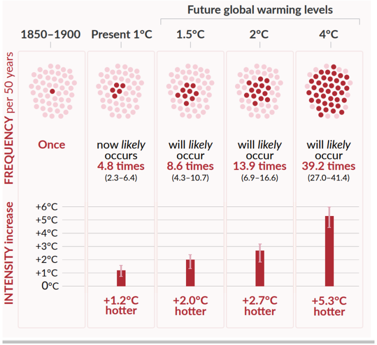


Figure: Projection de changements de l'intensité et de la fréquence des extremes chauds (GIEC 2021)

Sujet de thèse: Réaliser des **projections** en fin de siècle, de **températures extrêmes**, à l'**échelle locale**.

Objectifs:

- Choisir un indicateur adapté au contexte non stationnaire.
- L'estimer en utilisant une méthode qui intègre les modèles climatiques et les observations.
- L'adapter aux conditions locales .

Problème:

- Niveau de retour stationnaire: Non utilisable.
- Probabilité Annuelle de dépassement **non fixe dans le temps**.
- Besoin: Définir une quantité suivant sa **probabilité annuelle de dépassement** et **la période** à couvrir.
- Nécessité: **Logique similaire** au Niveau de retour.

Possibilités (S. Parey 2022):

- Design Life Level
- Equivalent Reliability
- Expected Waiting Time
- Borne Maximum
- Average Design Life Level
- Expected Number of Events
- Facteur d'amplification

Equivalent Reliability:

Pour $p = 1/T$ probabilité annuelle et T_1, T_2 la période d'intérêt:

$$P[\bigcap_{t=\mathbf{T_1}}^{\mathbf{T_2}} (Y_t < \mathbf{z}^{\mathbf{ER}}_{\mathbf{T_2}-\mathbf{T_1}})] = (1 - \frac{1}{\mathbf{T}})^{\mathbf{T_2}-\mathbf{T_1}+1}$$

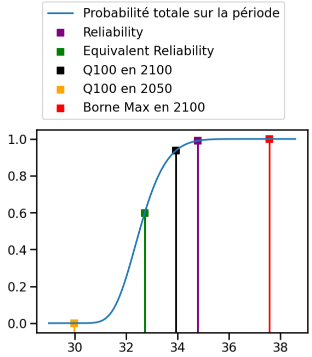


Figure: Probabilité totale de Non-dépassement sur 2050-2100

Le site du Tricastin:

- Dans la Vallée du Rhône (Relief)
- En service depuis 1980
- Altitude: 54m (Google Earth)

Figure: Situation du CNPE Tricastin

Données de modèles climatiques:

- 28 Modèles de climats **Globaux**, Génération CMIP6.
- Accès aux runs **historiques et scénarios** futurs, ici SSP 5-8.5.
- **Données locales:** TX (Température Maximale Quotidienne) au point le plus proche de Tricastin, fraction terre-mer $> 95\%$.
- **Données globales:** Valeurs moyennes annuelles sur l'Europe.
- Quelques **explosions numériques** (Exclues)

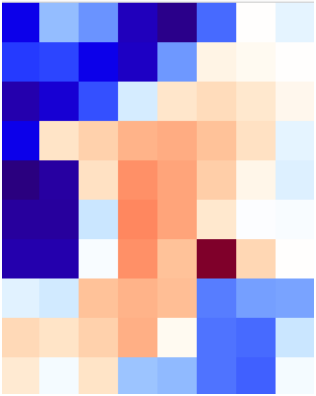


Figure: TX du 6 aout 2066 pour UKESM1-0-LL - France

Données d'observations météorologiques :

- Données **locales** observées (Pierrelatte)
- 37 ans d'observations
- **Qualité SQR** (Ruptures faibles uniquement), expertisée.
- **Alternatives:** Orange (60 ans) , Montélimar (60 ans), Tricastin (16 ans)
- Pas d'information sur l'évolution future

Température moyenne sur l'Europe: Crutem5 (Température grille)

⟹ Quelles données utiliser sans observations locales ?

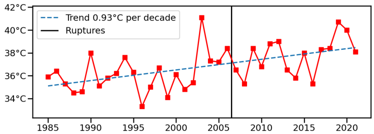


Figure: Maximums Annuels à Pierrelatte

Contraintes:

- Travail sur des **valeurs extrèmes**: Maxima Annuels, Distribution **GEV**.
- **Non-stationarité**: **Covariable** de température Globale (Européenne). Relation au temps et intégration des scénarios.

$$Y \sim \mathbb{P}_t = GEV(\mu_t, \sigma_t, \xi)$$

$$\mathbb{M}(t) = \begin{cases} \mu(t) & = \mu_0 + \mu_1 X_t \\ \sigma(t) & = exp(\sigma_0 + \sigma_1 X_t) \\ \xi(t) & = \xi_0 \end{cases}$$

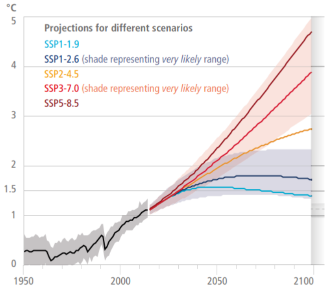


Figure: Différents scénarios d'évolution. (IPCC, 2022)

Méthode tirée de Robin et Ribes (2020):

1. Estimation des paramètres de la GEV pour chaque modèle climatique.
2. Multisynthèse: Création de la prior.
3. Contrainte par les observations.

Utilisée sur la **France entière** et pour l'attribution.
Utilisé aussi sur les **valeurs moyennes** de température.

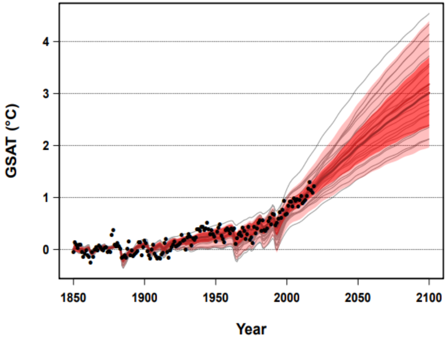


Figure: Illustration de la méthode par contrainte sur des températures moyennes.

Pour chaque modèle climatique:

- **Estimation de** $X_{1850} - X_{2100}$ avec un modèle EMB (Energy Balance Model, package NSSEA)
- **Estimation des paramètres de la GEV** θ_{GEV} par Maximum de Vraisemblance (Package SDFC).
- Choix: Travail en **anomalies** (Période de référence 1986- 2016)

$\Longrightarrow$ On obtient **28 tirages** différents de θ_{GEV} et $X_{1850} - X_{2100}$

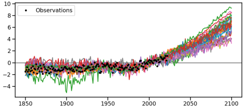


Figure: Anomalies de température Moyenne sur l'Europe (Covariable)

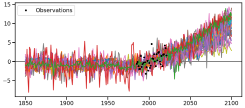


Figure: Anomalies de température Maximale Annuelle sur Tricastin

Création d'une **prior**:

- Incluant les 5 paramètres GEV θ_{GEV} et les 250 ans de la covariable $X_{1850} - X_{2100}$.
- **Gaussienne Multivariée** (Moyenne et covariance des jeux de paramètres)

Raisonnement:

- Chaque ajustement sur un modèle climatique est une réalisation possible.
- **Hypothèse:** Les modèles sont statistiquement indistinguables de la vérité. ($\neq$ modèles centrées sur la vérité.)
- **Attention:** la vérité peut être en dehors de la distribution **si les modèles ne sont pas assez dispersés.**

$\Longrightarrow$ Comment évaluer la qualité de la prior ?

Grâce au théoreme de Bayes:
Avec $\theta = \{X_{1850} - X_{2100}, \theta_{GEV}\}$, Y^0 les observations locales et X^0 celles de la covariable.

$$\mathbb{P}(\theta | Y^0 \cap X^0) = \frac{\mathbb{P}[Y^0 | (\theta | X^0)] \mathbb{P}(\theta | X^0)}{\mathbb{P}(Y^0)}$$

L'estimation se fait donc en **deux étapes**:

- **La contrainte par la covariable** $\mathbb{P}(\theta | X^0)$. Posterior obtenue par **loi conjuguée** grâce au Gaussian conditioning theorem (Eaton 1983).
- **La contrainte par les observations locales.** Pas de posterior explicite, donc estimation par **MCMC**.

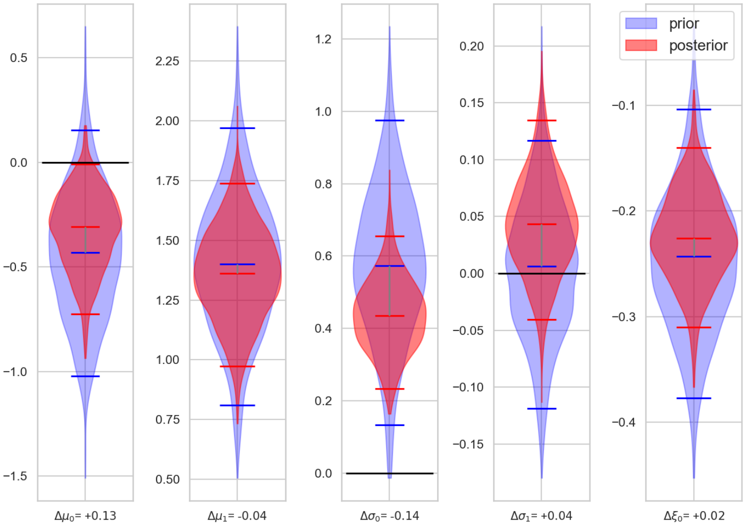
On suppose X_t **connu** (Tirage), on estime les paramètres de la GEV
$\theta_{GEV} = \{\mu_0, \mu_1, \sigma_0, \sigma_1, \xi\}$.

Estimation par **Metropolis Hasting.**
Concept:

- A chaque étape, **bruiter** les paramètres.
- **Probabilité d'accepter** la version bruitée suivant la vraisemblance de Y^0 et de la prior $\pi(\theta)$

Des **paramètres à choisir** (Forme de la fonction de transition).

$\Longrightarrow$ En cours de reprise (Analyse de sensibilité des paramètres, simulations, choix de la transition, etc)
$\Longrightarrow$ Alternativement, modèle Hybride MH-Gibbs ou méthodes alternatives (INLA).


prior
posterior
0.5
0.0
−0.5
−1.0
−1.5
2.25
2.00
1.75
1.50
1.25
1.00
0.75
0.50
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.1
−0.2
−0.3
−0.4
$\Delta\mu_0$= +0.13
$\Delta\mu_1$= -0.04
$\Delta\sigma_0$= -0.14
$\Delta\sigma_1$= +0.04
$\Delta\xi_0$= +0.02

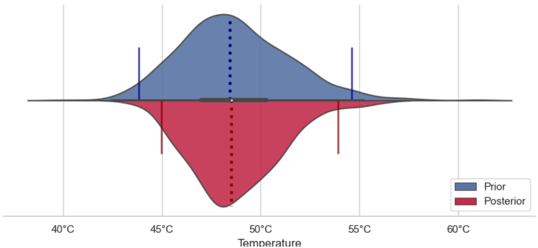


Figure: Equivalent Reliability, probabilité annuelle 0.01 sur 2050-2100

A posteriori: Distribution plus étroite (**Gain de confiance**)

- Quantile 5% de 43.7°C à 45.2°C
- Quantile 95% de 54.4°C à 53.3°C

Valeur médiane à 48.5°C : **Peu de différence** (Environ 0.1°C).

Sur une période de 50 ans, un évènement dont la probabilité annuelle est de 1/100 **n'est plus rare**. (41 % au total)

- **Choix** d’un indicateur: Equivalent Reliability
- **Prise en main** de la méthode de Robin et Ribes 2020.
- **Première Application** sur Tricastin.
- Nécessité de documenter et **adapter l’algorithme MCMC.**
- Ouverture de **nombreuses pistes** (Données, évaluation de la prior, information locale, etc)

Étape 1 (Juillet - Août -Septembre): **MCMC**
Évaluation du MCMC: simulation, évaluation, choix de la transition, poids des observations.
Alternatives (INLA, Hybride, Adaptative MCMC pour la transition)
Que conserver du tirage (Une médiane ? plusieurs tirage par run? Garder un tirage ?)
$\Longrightarrow$ **But:** Version Stable du MCMC + Confiance en ses estimations.

Étape 2 (Automne-Hiver suivant choix) : **Application**
Tester des Variantes (Scénarios, Modèles régionaux seuls, CMIP5 ?)
Application sur plusieurs points d'intérêt (Autres sites nucléaire ? Grande villes Françaises ? Choix zone d'intérêt ?)
Evaluation de la qualité d'une prior ?
$\Longrightarrow$ **But:** Obtenir une méthode pour évaluer la prior, et estimer la prise en compte de l'information locale.

Étape 3 (Incertain):
Globalement, **intégration de l'échelle locale**.
Suivant les résultats de l'étape 2, **plusieurs possibilités**:

- **Axe Modèle:** Intégrer GCM/ modèle régionaux (Hiérarchique, autre ?), statistical downscaling avant l'étape prior, etc.
- **Axe Observations locales** : Krigeage (Simulation d'observations locales quand la station proche est trop courte/ dans de mauvaises conditions. Modèles existants prenant en compte le relief, etc. Base de données existante), Modèle hiérarchique (Composante spatiale)

Utilisés pour cette présentation:

- ”Estimation d’extrêmes hydrométéorologiques en contexte de changement climatique” by S. Parey, HDR, 2022.
- Robin, Y. and Ribes, A.: Nonstationary extreme value analysis for event attribution combining climate models and observations, Adv. Stat. Clim. Meteorol. Oceanogr., 6, 205–221,https://doi.org/10.5194/ascmo-6-205-2020 , 2020.
- Packages python SDFC et NSSEA
- [en cours d’utilisation] ”Bayesian Computation Via Markov Chain Monte Carlo” by Radu V. Craiu and Jeffrey S. Rosenthal in Annual Review of Statistics and Its Application 2014 1:1, 179-201

Températures maximales en France au 21ème siècle

Merci d'avoir suivi cette présentation
Des questions?

Only need to be able to simulate from conditionnal distributions. (Maybe possible use of X_T)
Multivariate : $\psi = (\psi_1, \ldots, \psi_d)'$, full conditionnals are $\pi(\psi_i|\psi_{-i}) = \pi_i(\psi_i)$
Description of algorithm:

- Initialisation: k=1, initial state of chain $\psi^{(0)}$
- Boucle: For new value $\psi^{(k)}$:
  - $\psi_1^{(k)} \sim \pi(\psi_1|\psi_{-1}^{(k-1)})$
  - $\psi_2^{(k)} \sim \pi(\psi_2|\psi_{-1,2}^{(k-1)}, \psi_1^{(k)})$
  - ...
  - $\psi_d^{(k)} \sim \pi(\psi_d|\psi_{-d}^{(k)})$

$\pi(\psi)$ is still the density of interest. We now have a transition kernel $p(\psi_{i+1}, \psi_i)$, easy to simulate from, to get successive values.

- Initialisation : k=1,initial state of chain $\psi^{(0)}$
- Boucle: For new value $\psi^{(k)}$:
  - Generate new proposed value ψ' using the kernel transition function.
  - Calculate Acceptance Probability (ratio) $A(\psi^{(k-1)}, \psi')$ of the proposed changeof value:

$$A(\psi^{(k)}, \psi') = min\{1, \frac{\pi(\psi')L(\psi'|x)p(\psi', \psi^{(k-1)})}{\pi(\psi^{(k-1)})L(\psi^{(k-1)}|x)p(\psi^{(k-1)}, \psi')}\}$$

  - Accept $\psi^{(k)} = \psi'$ with probability $A(\psi^{(k)}, \psi')$ and keep $\psi^{(k)} = \psi^{(k-1)}$ otherwise.

Based on Bayesian Modelling of Extreme Rainfall Data from Elizabeth Smith
Gibbs concept (each parameter is updated in turn) and conditionals are MH (do we accept the new value produced by the transition function?)
Avantage: Each parameter has his own trajectory (One may not move much and another a lot) + varying transition kernel (proportionnal) (not hard to do for simple MH too)
$\rightarrow$ Less dependance than normal MH?.
Description of algorithm:

- Initialisation : k=1,initial state of chain $\theta^{(0)}$
- Boucle: For new value $\theta^{(k)}$:
    - In turn, for each parameter $\theta_j^{(k)}$
        - $\theta_j^{'} = \theta_j^{(k-1)} + \varepsilon_j$
        - Accept or refuse using $A(\theta_j^{(k-1)}, \theta_j^{'})$ with $\theta_{-j}^{(k)}$ seen as known.