PAULA GONZALEZ

DEVELOPMENT OF CAUSAL LEARNING AND EXPERT AGGREGATION ALGORITHMS

TO TAKE INTO ACCOUNT THE MULTI-MODEL ERROR IN STUDIES OF DETECTION AND ATTRIBUTION OF CLIMATE RECORDS

Encadrement : O. Thao, P. Naveau (LSCE), J. Worms (UVSQ)

Motivation

Motivation

- Climate models are affected by different types of biases. The Detection and Attribution (D&A) of climate anthropogenic signals involves reducing the uncertainty of climate multi-model projections. Improving D&A of extreme events, such as heatwaves and heavy rainfall, is important because they have high societal impacts.

Starting point of the work:

- *P.Naveau and S.Thao* (2020) -> Multi-model errors and emergence times in climate attribution studies
- *J.Worms and P.Naveau* (2020) -> Record events attribution in climate studies
- This internship -> Merge the two works and propose a framework that improves record events attribution analysis by taking in account multi-model error.

Factual and Counterfactual theory

We imagine two worlds, Counterfactual and Factual:

Counterfactual : World in which anthropogenic emissions have never occurred. It did not happen in reality.

Factual: World with todays forcing conditions, i.e. with anthropogenic forcing included.

Is the probability of occurrence of a extreme event in the factual world different from the probability of the same event but under a counterfactual scenario ?

Multi-model errors and emergence times in climate attribution studies.

P.Naveau and S.Thao (2020)

For the same climate variable, we consider $Z_t^{(m)}$ and $X_t^{(m)}$, factual an counterfactual trajectories from the CMIP model m, where t is a year. **Which one is bigger?**

Attribution statistic:

$q_t^{(m)} = P(Z_t^{(m)} > X_t^{(m)})$

- $q_t^{(m)} = 1/2$: Observing the event $\{Z_t^{(m)} > X_t^{(m)}\}$ is as likely as $\{X_t^{(m)} > Z_t^{(m)}\}$.
- $q_t^{(m)} \neq 1/2$: Pattern change due to anthropogenic forcing.

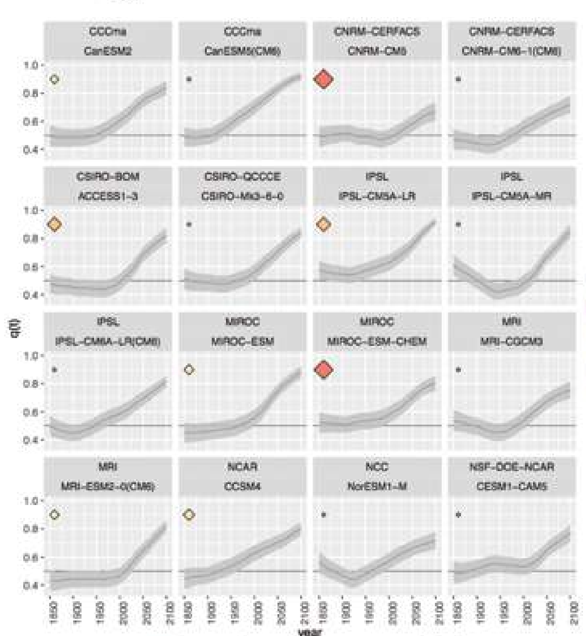


Figure 1: Analysis of yearly maxima of daily precipitation from 16 CMIP climate runs at the grid point near Oxford, Great Britain. y-axis represents the estimate of q_t.

Each model m estimates a different probability.

How do we decide which is the good estimation?

Multi-model aggregation:

Models $q_t^{(m)}$ estimations are weighed so as

$\hat{q}_t = 1/2$ in the pre-industrial period (1850-1900)

Theory of records

The observation Y_r is a record at time r if the latest recorded value of the last r years is the largest i.e.

$$Y_r \geq \max\ (Y_1, \ldots, Y_{r-1})$$

Record event attribution:

$$p_{1,r} = P(Z_r > \max\{X_1, \ldots X_{r-1}\})$$

Probability that Z_r is a record within the counterfactual world

The counterfactual world is considered stationary. If $p_{1,r} \gg \frac{1}{r}$, there is a strong probability of causality between the record and anthropogenic emissions.

We realize that $p_{1,2_t} = q_t$. **Record event attribution statistic** is essentially the **Attribution statistic for extremes** (large r).

Records events attribution in climate studies.

J.Worms and P.Naveau (2020)

The figure shows us the probability that the temperature sample, coming from the factual world, is superior than the counterfactual samples of the last 10 years.

The white zones are where $\hat{p}_{1,10} < \frac{1}{10}$

$\hat{p}_{1,10} > \frac{1}{10}$ in most of the surface: There is a strong probability of causality between the record near-surface temperatures and anthropogenic emissions in most of the surface of the earth.

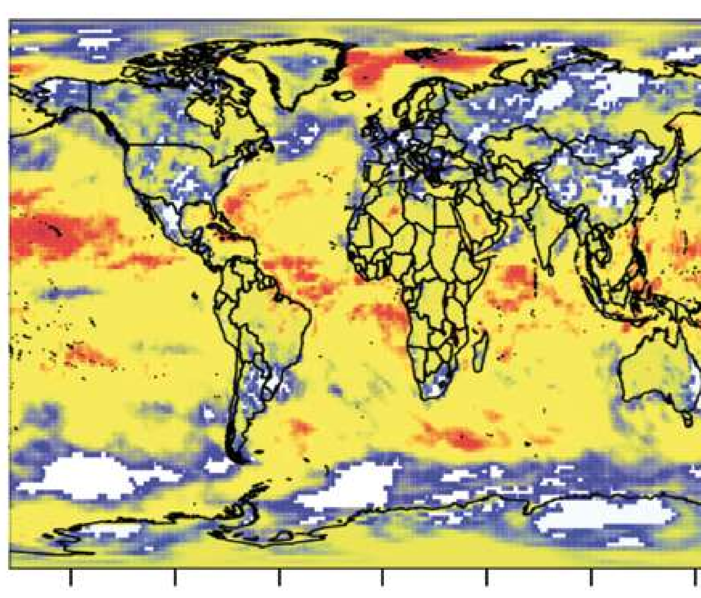
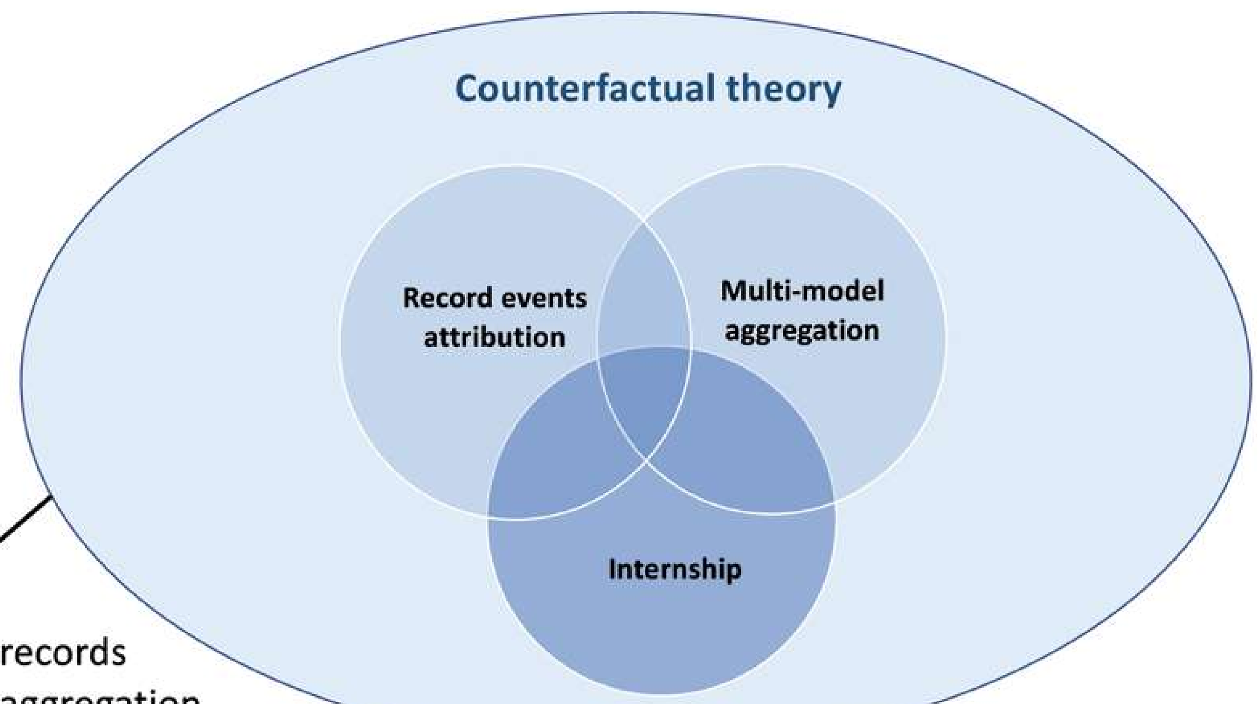


Figure 2: Record analysis of yearly maxima of daily maxima of near-surface air temperature issued from numerical climate model CNRM-CM5 of Météo-France.

Internship



Tools
- Theory of records
- Theory of aggregation

LSCE UVSQ université PARIS-SACLAY INVESTIR L'AVENIR IPSL