



Données & Risques Climatiques

Arthur Charpentier (UQAM & Univ. Rennes)

SoScience & Allianz, Juin 2022

June 9, 2022



Arthur Charpentier

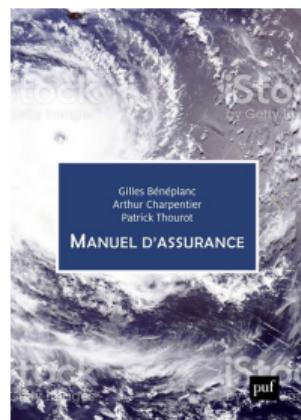
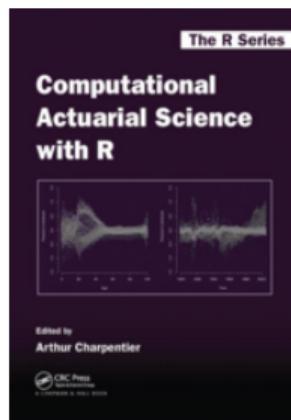
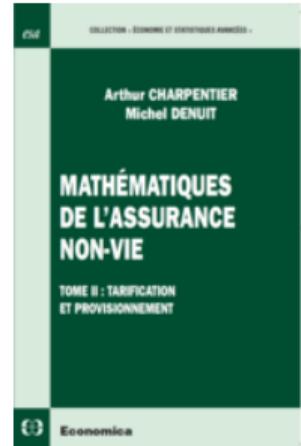
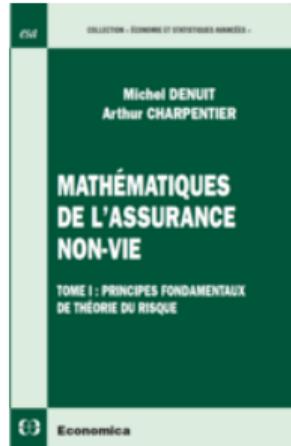
Université du Québec à Montréal

 @freakonometrics

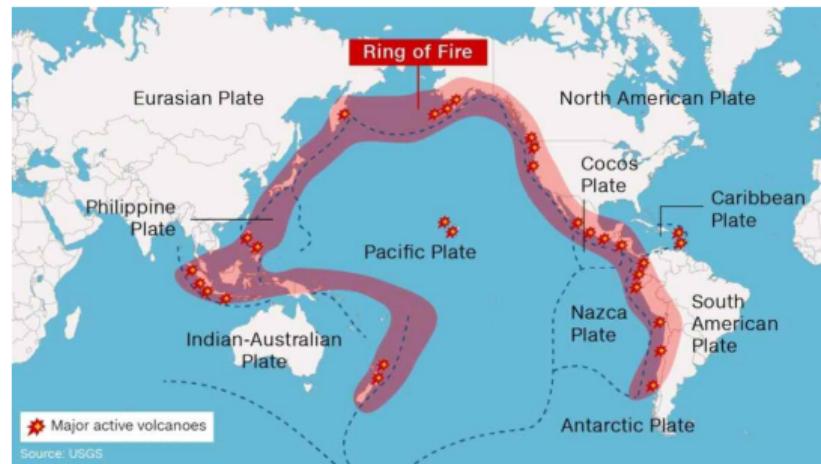
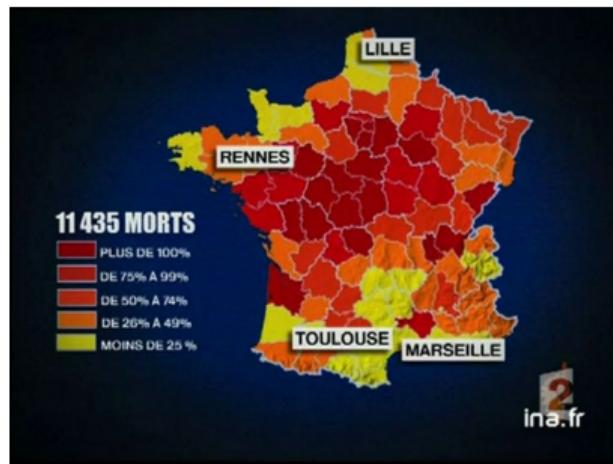
 freakonometrics

 freakonometrics.hypotheses.org

Predictive Modeling, Actuarial Science,
Mathematical Economics, Risk, Inequalities,
Econometrics, Statistics, Machine Learning
Climate Modeling, Extremes, Fairness



Dynamics of Natural Catastrophes (incl. Climate) [1]



- [1] A. Charpentier. "On the return period of the 2003 heat wave". In: *Climatic change* 109.3 (2011), pp. 245–260.
- [2] A. Charpentier and D. Sibaï. "Dynamic flood modeling: combining Hurst and Gumbel's approach". In: *Environmetrics* 20.1 (2009), pp. 32–52.

Dynamics of Natural Catastrophes (incl. Climate) [2]



IRIS is a university research consortium dedicated to monitoring the seismic and regional tectonic activity through the collection and distribution of geophysical data.

IRIS programs contribute to federally funded research, education, earthquake hazard mitigation, and the verification of the Comprehensive Test Ban Treaty.

Support for IRIS comes from the National Science Foundation, other federal agencies, universities, and private foundations.

This figure was produced in cooperation with University of Arizona, University of California, Berkeley, University of California, San Diego, Purdue University, and the US Geological Survey.

1200 New York Ave, NW #800
Washington, DC 20004
Phone: (202) 467-0200
Fax: (202) 467-0444
www.IRIS.edu

Education & Outreach Series

No. 6

How are Earthquakes Located?

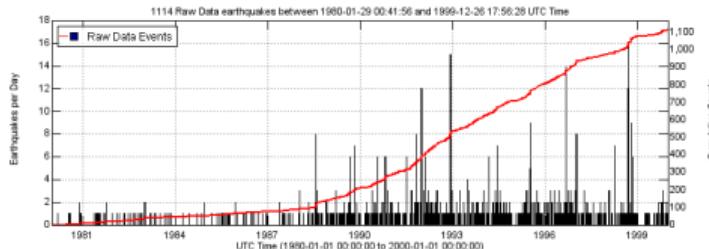
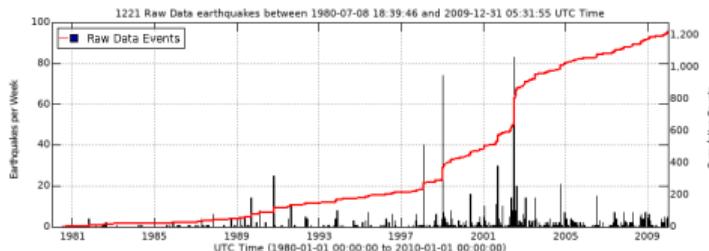
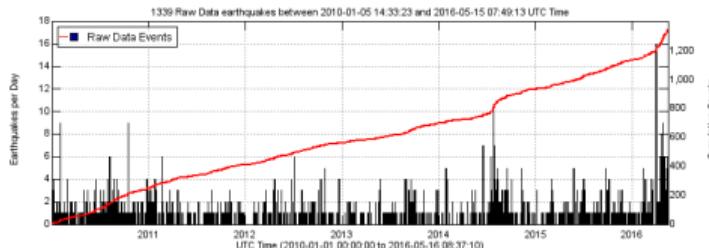
We can locate earthquakes using a simple fact: an earthquake creates different seismic waves (P waves, S waves, etc.) The different waves each travel at different speeds and therefore arrive at a seismic station at different times. P waves travel the fastest, so they arrive first. S waves, which travel at about half the speed of P waves, arrive later. A seismic station close to the earthquake records P waves and S waves in quick succession. With increasing distance from the earthquake the time difference between the arrival of the P waves and the arrival of the S waves increases.

Although modern techniques are more complex, we have illustrated the basic concept using an example of an earthquake near Mexico and seismic stations in North America. The following two steps show how we determine distance from the seismograms and estimate the location using three stations.

Step 1. The time between the arrival of the P wave and the arrival of the S wave (S-P time) is measured at each station. The S-P time indicates the distance to the earthquake similar to how the time interval between the flash of light and the sound of thunder indicates the distance to a thunderstorm. In our example, station TEIG (with an S-P time of 1.5 minutes) is closest to the earthquake, and station SSPA (with an S-P time of 5 minutes) is farthest away.

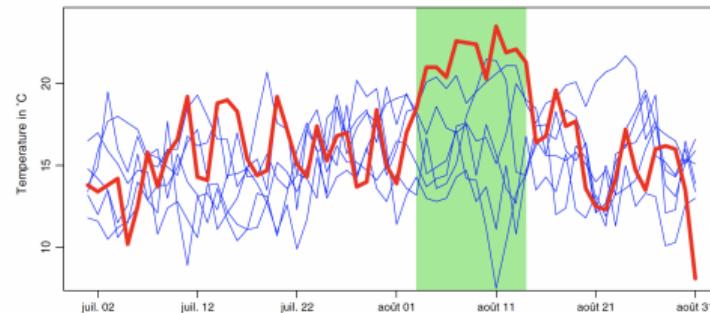
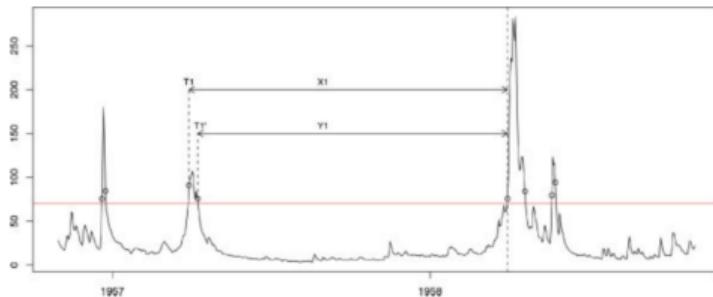
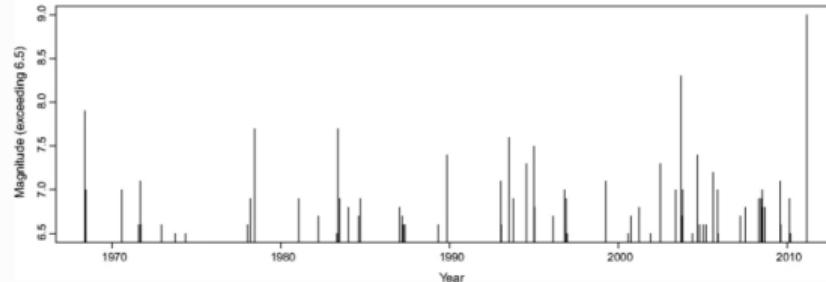
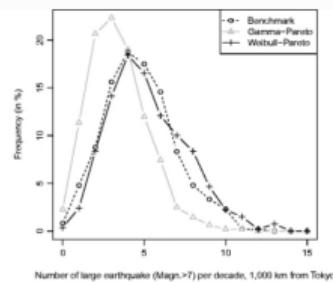
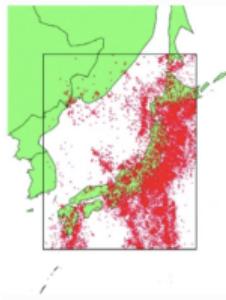
From observing and analyzing many earthquakes, we know the relationship between the S-P time and the distance between the station and the earthquake. We can therefore convert each measured S-P time to distance. A 1 minute interval of 1.5 minutes corresponds to a distance of 900 kilometers, 3 minutes to 1800 kilometers, and 5 minutes to 3300 kilometers.

Step 2. Once we know the distance to the earthquake for three stations, we can determine the location of the earthquake. For each station we draw a circle around the station with a radius equal to its distance from the earthquake. The earthquake occurred at the point where all three circles intersect.



Dynamics of Natural Catastrophes (incl. Climate) [3]

“seismic gap hypothesis” / dynamic of flood events / heat wave persistence



Flood Risk in France [1]



- [1] A. Charpentier, L. Barry, and M. James. "Insurance against Natural Catastrophes: Balancing Actuarial Fairness and Social Solidarity". In: *Geneva Papers on Risk & Insurance* (2021). DOI: [10.1057/s41288-021-00233-7](https://doi.org/10.1057/s41288-021-00233-7).
- [2] France Info. "Seine-et-Marne : la ville de Thoméry se remet progressivement des inondations". In: (2018).

Flood Risk in France [2]

- [1] A. Charpentier, L. Barry, and M. James. "Insurance against Natural Catastrophes: Balancing Actuarial Fairness and Social Solidarity". In: *Geneva Papers on Risk & Insurance* (2021). doi: 10.1057/s41288-021-00233-7.

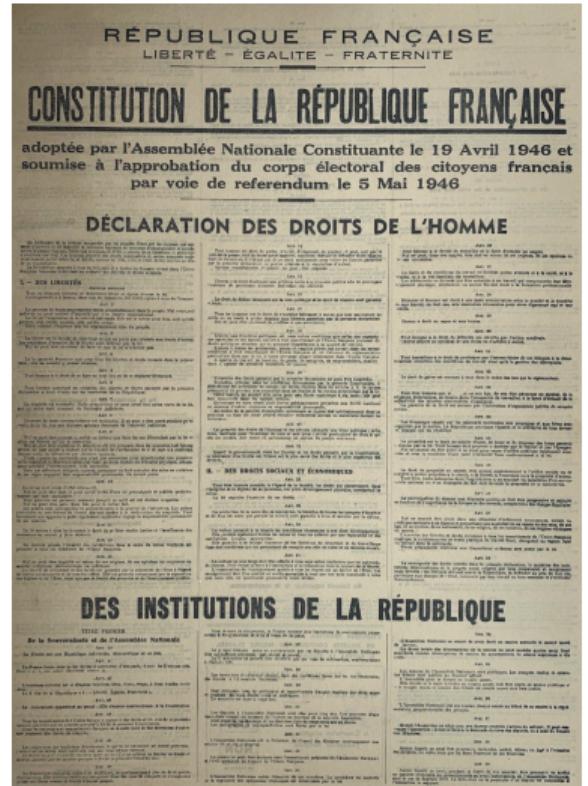
On fairness & solidarity

➤ French Constitution (1946)

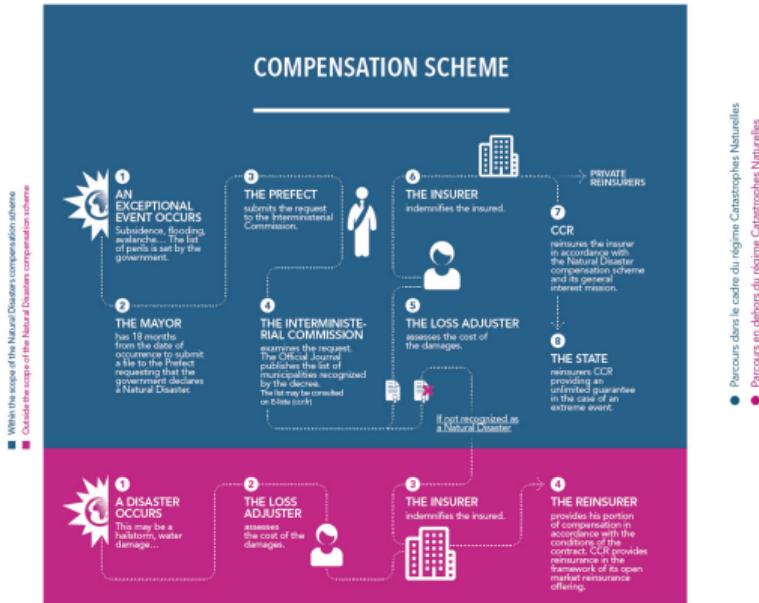
12. *La Nation proclame la solidarité et l'égalité de tous les Français devant les charges qui résultent des calamités nationales.*

➤ 82-600 Law (1982)

régime d'indemnisation des catastrophes naturelles

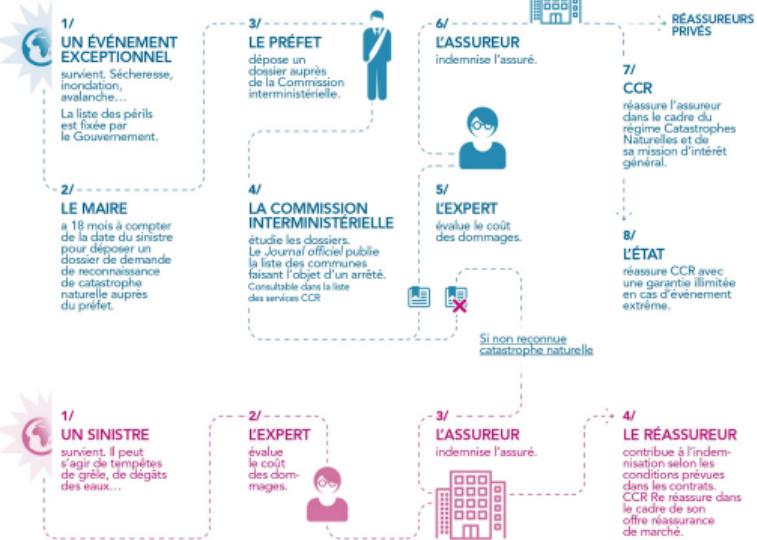


Flood Risk in France [3]



Parcours dans le cadre du régime Catastrophes Naturelles

Parcours en dehors du régime Catastrophes Naturelles

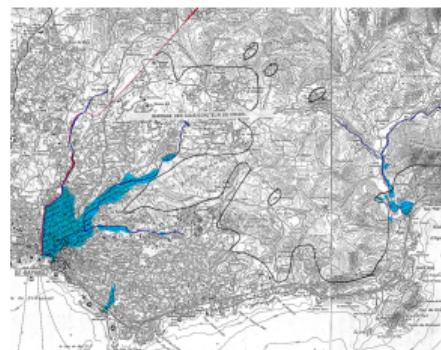
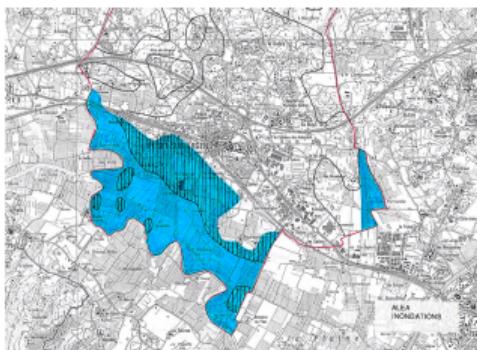
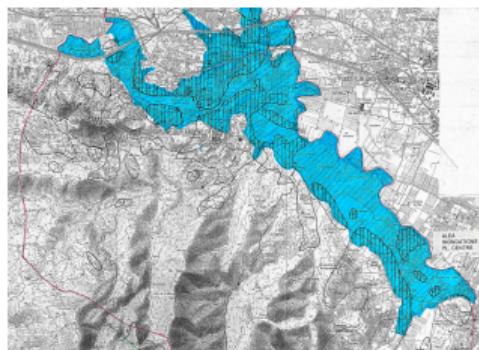


source: <https://www.ccr.fr/en/-/indemnisation-des-catastrophes-naturelles-en-france>

Flood Risk in France [4]

Two different flood perils : overflow vs. coastal

PPRIs ([plan de prévention du risque inondation](#)) in Roquebrune-sur-Argens, Puget and Saint-Raphaël. The plain area (in blue) is the risky area.



Areas clearly identified as risky, from documented (historical) floods.

Flood Risk in France [5]



Procès tempête Xynthia : deux élus condamnés à des peines de prison



- [1] France 3. "Tempête Xynthia: la responsabilité de La Faute-sur-Mer pourrait être alourdie". In: (2019).

Flood Risk in France [6]

PPRLs ([plan de prévention des risques littoraux](#)) in Vendée. The dashed area is the risky area. Areas with possible coastal risk.

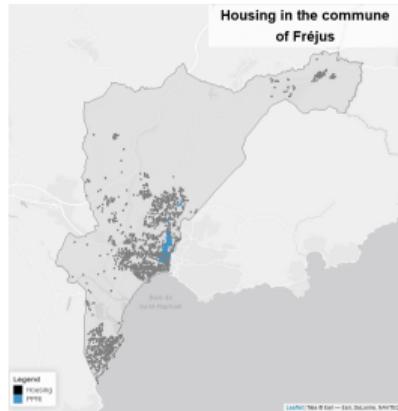
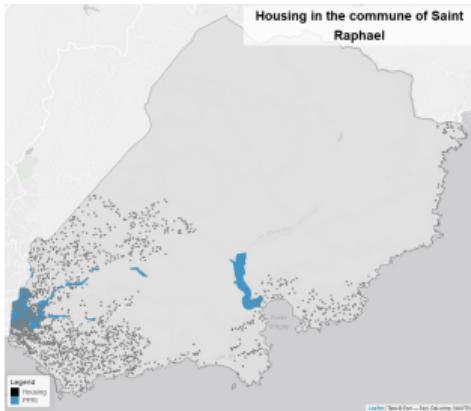
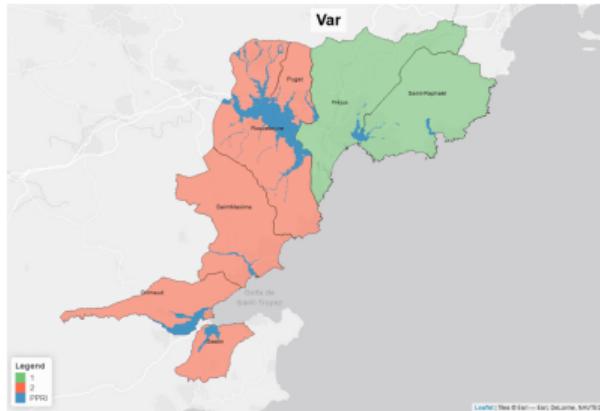


See <https://github.com/freakonometrics/floods>

Flood Risk in France [7]

10% of households represent 73.6% of the losses... who lives in those risky areas ?

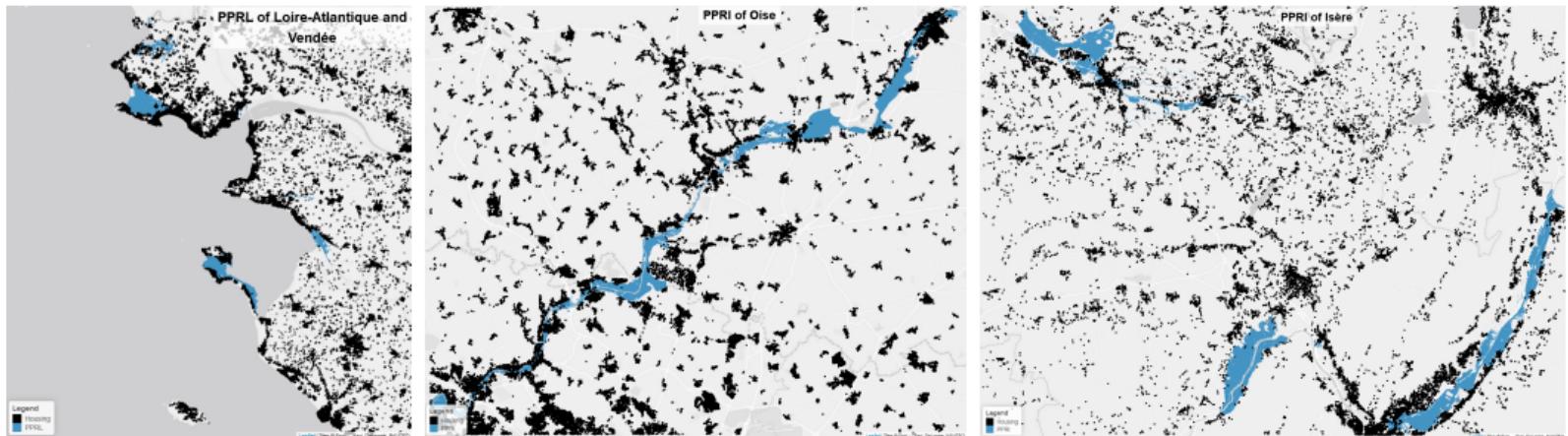
“ventes de biens fonciers” dataset, <https://cadastre.data.gouv.fr/dvf>, 2014-2018,



(possible bias on those 5-year notarial transactions...)

Flood Risk in France [8]

E.g. in 4 “departements” (Loire-Atlantique, Vendée, Oise, Isère)



- sold houses / apartments, ■ PPRI-PPRL areas

Flood Risk in France [9]

Table 1: coastal risk areas vs. Table 2: overflow / non-costal risk areas

			Average Price	Difference (%)	Maximum Price	Number	Proportion (%)	Welch <i>t</i> test
Vendée	Non risky	Apartments	4293		21840	329	9%	
		Houses	2928		65909	2795	74%	
	Risky	Apartments	3302	-23%	9773	39	1%	1.0
		Houses	10253	+250%	71483	637	17%	-60.1
Pays-Loire	Non risky	Apartments	4399		79913	8411	37%	
		Houses	3019		75472	12678	55%	
	Risky	Apartments	6784	+54%	68478	1001	4%	-8.6
		Houses	3245	+7%	22895	765	3%	-2.7

Table 1: Prices (€ per m^2) of houses sold (2014-2018) for Vendée - Western part of France, with PPRL (**coastal risk**). The *Difference* is the relative difference between average prices (per m^2) between the risky and the non-risky zones, either for apartments or houses.

Flood Risk in France [10]

			Average Price	Difference (%)	Maximum Price	Number	Proportion (%)	Welch t value
Var	Non risky	Apartments	5392			9874	53%	
		Houses	5957			6913	37%	
	Risky	Apartments	4190	-22%		1471	8%	6.4
		Houses	4172	-30%		226	1%	5.2
Haute Loire	Non risky	Apartments	2399		38333	3403	27%	
		Houses	1314		20625	8857	69%	
	Risky	Apartments	2163	-11%	28125	319	2%	1.6
		Houses	1247	-5%	7432	272	2%	0.9
Seine et Marne	Non risky	Apartments	6260		79710	82133	44%	
		Houses	3356		79167	98824	53%	
	Risky	Apartments	4333	-30%	40000	2177	1%	8.0
		Houses	2693	-20%	54096	1784	1%	7.5

Flood Risk in France [11]

		Average Price	Difference (%)	Maximum Price	Number	Proportion (%)	Welch t value
Isère	Non risky	Apartments	4960	79800	27982	52%	
		Houses	2429	69375	24600	45%	
	Risky	Apartments	3252	-3%	35714	885	2%
		Houses	2543	+5%	14067	435	1% -1.2
Oise	Non risky	Apartments	6170	79963	24613	34%	
		Houses	3126	78214	44737	62%	
	Risky	Apartments	5725	-7%	50000	1385	2% 2.1
		Houses	2866	-8%	62184	1640	2% 4.6

Table 2: Prices (€ per m^2) of houses sold (2000-2020) for several départements in France, with PPRI (overflow risk, or non-costal).

Subsidence Risk in France [1]



- [1] A. Charpentier, M. R. James, and H. Ali. "Predicting Drought and Subsidence Risks in France". In: *Natural Hazards and Earth System Sciences Discussions* (2021), pp. 1–27. DOI: [10.5194/nhess-2021-214](https://doi.org/10.5194/nhess-2021-214).
- [2] France Info. "Seine-et-Marne : la ville de Thoméry se remet progressivement des inondations". In: (2018).

Subsidence Risk in France [2]

Joint work with [Hani Ali](#) (Willis Re) and [Molly James](#) (EURIA / ACPR).

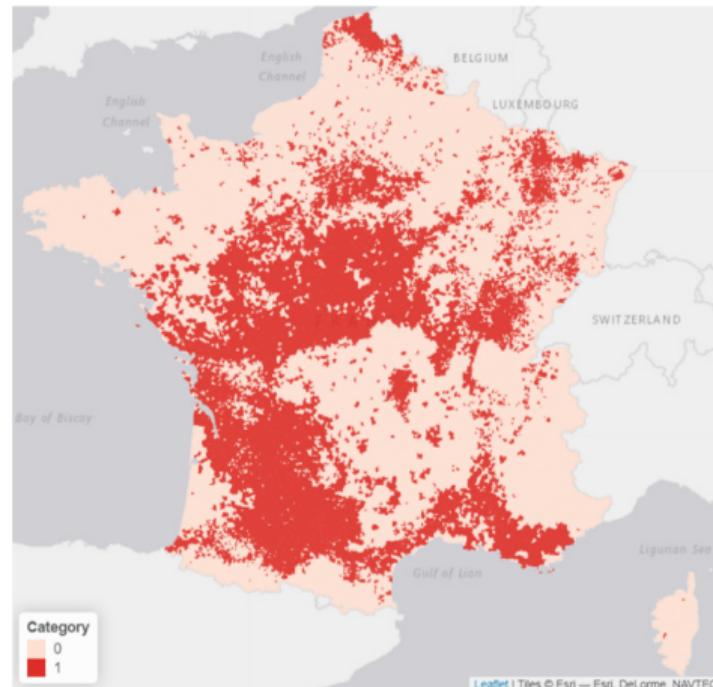
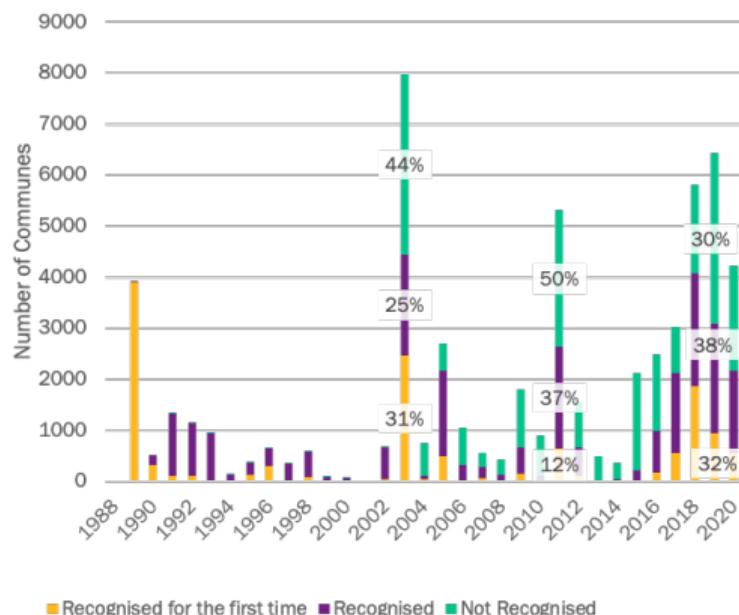
- [1] A. Charpentier, M. R. James, and H. Ali. "Predicting Drought and Subsidence Risks in France". In: *Natural Hazards and Earth System Sciences Discussions* (2021), pp. 1–27. DOI: [10.5194/nhess-2021-214](https://doi.org/10.5194/nhess-2021-214).

"Subsidence is caused by the shrinkage and swelling of clay soils"

- **Geotechnical factor:** Area of municipalities at medium or high risk $> 3\%$.
(categories based on clay concentration in the soil and historical statements)
- **Meteorological factor:** Standardized soil moisture index (SSWI), if an indicator of the season is lower than a return period of 25 years, then the whole season is eligible for the commune concerned.

Subsidence Risk in France [3]

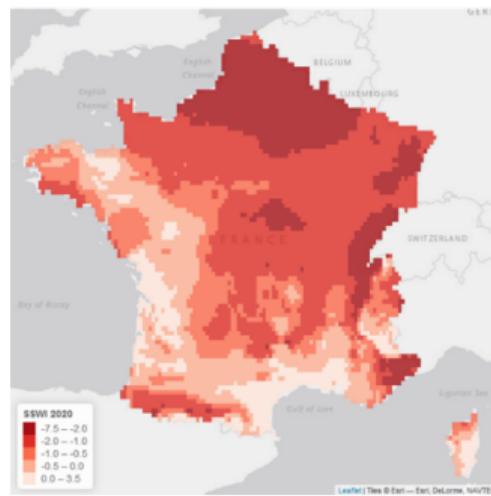
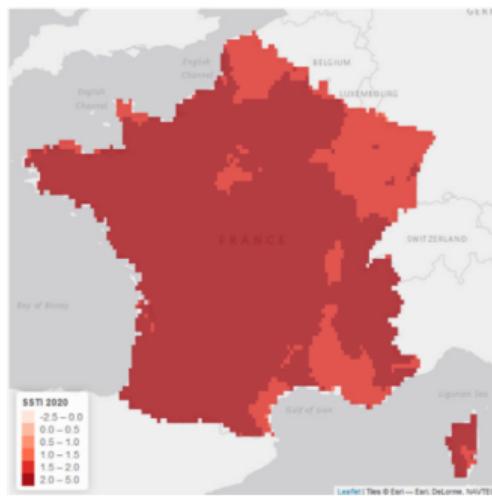
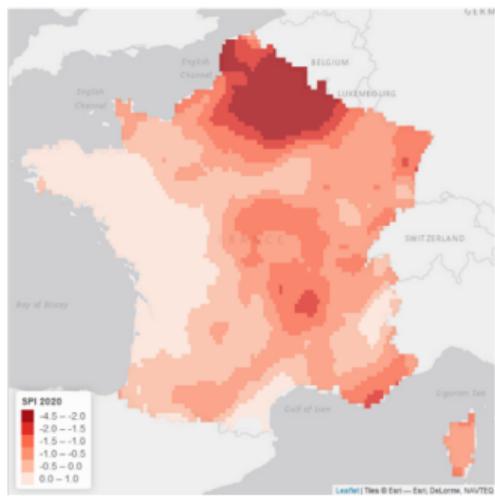
Data 1989-2018



Leaflet | Tiles © Esri — Esri, DeLorme, NAVTEQ

Subsidence Risk in France [4]

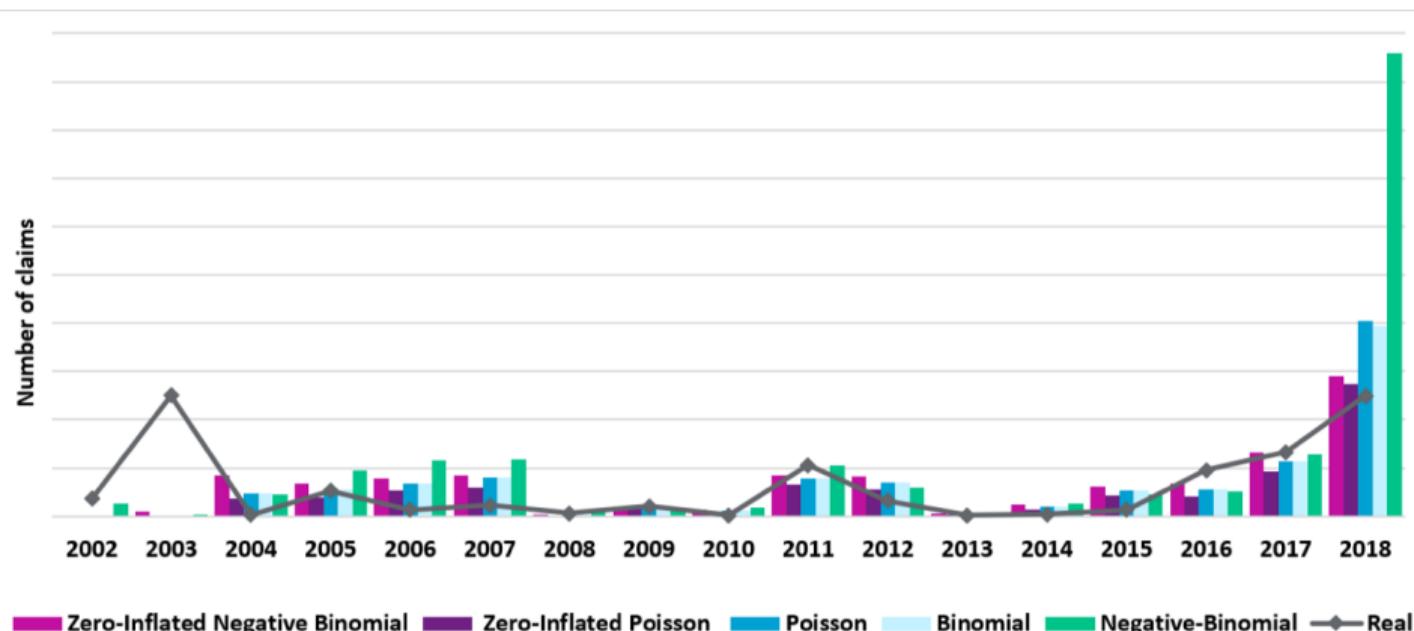
Indicators 2020: **ESPI** (precipitation) **ESSTI** (soil temperature) & **ESSWI** (soil humidity), ERA5-Land 9 km × 9 km



(via ESDAC (European Soil Data Centre) for soil concentration)

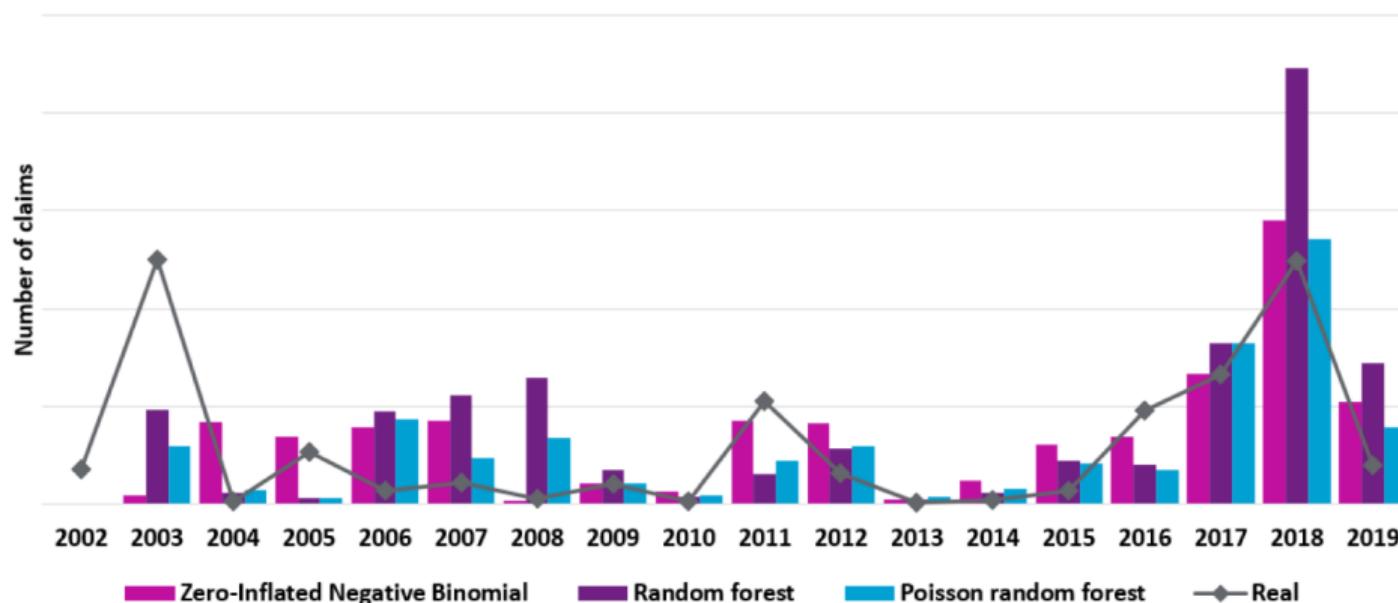
Subsidence Risk in France [5]

Regression models for frequencies: binomial, Poisson, negative binomial & zero-inflated Poisson, zero-inflated negative binomial,



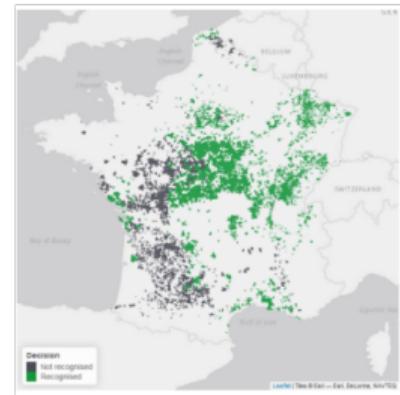
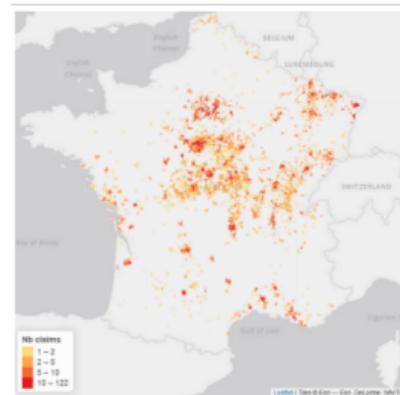
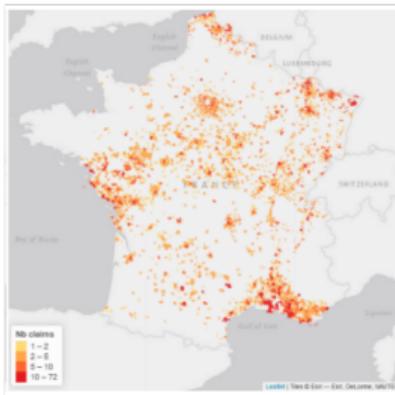
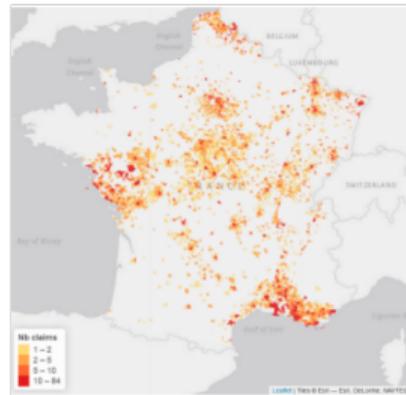
Subsidence Risk in France [6]

Random forest models for frequencies



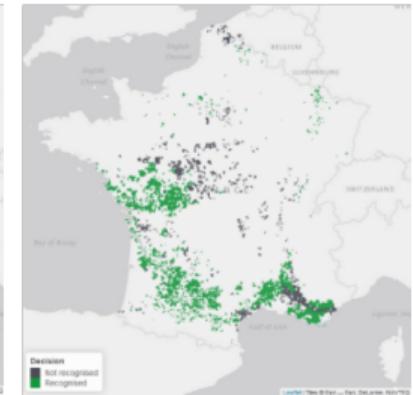
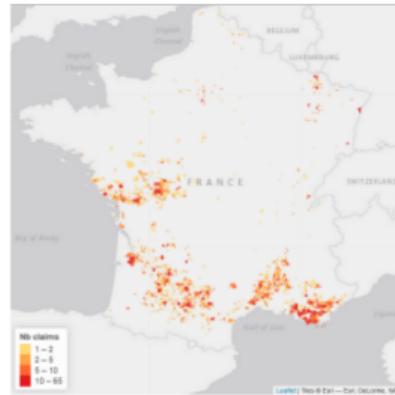
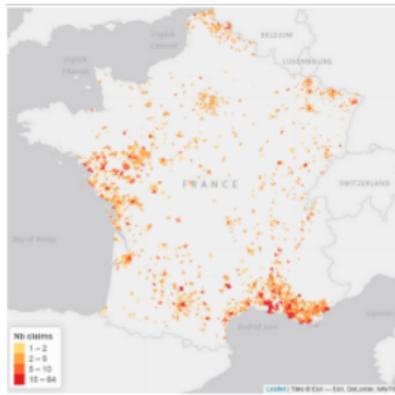
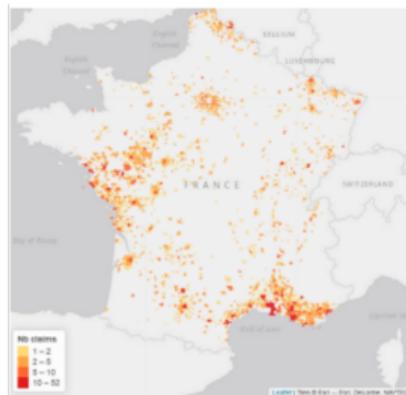
Subsidence Risk in France [7]

2017, random forest Poisson, zero inflated, observed, Nat Cat recognition



Subsidence Risk in France [8]

2018, random forest Poisson, zero inflated, observed, Nat Cat recognition

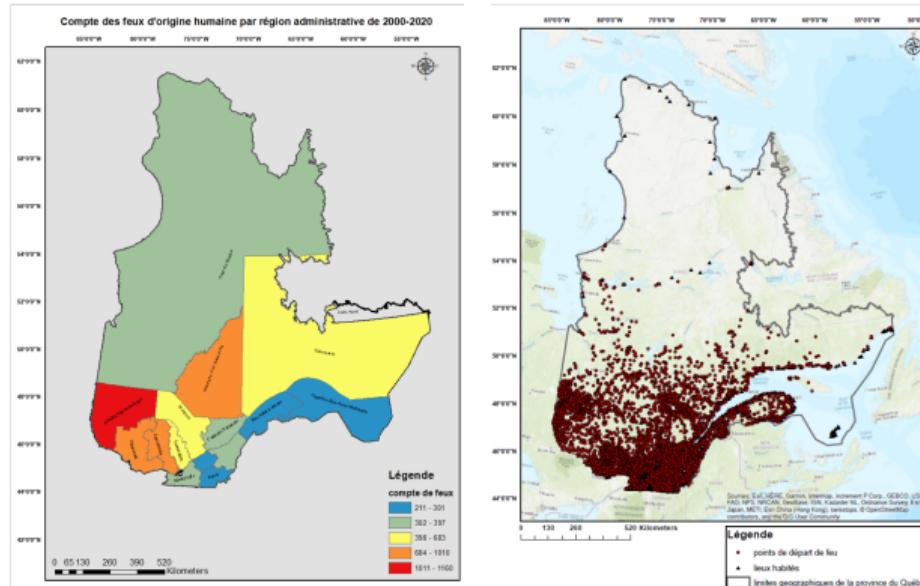


Wildfire Canada [1]



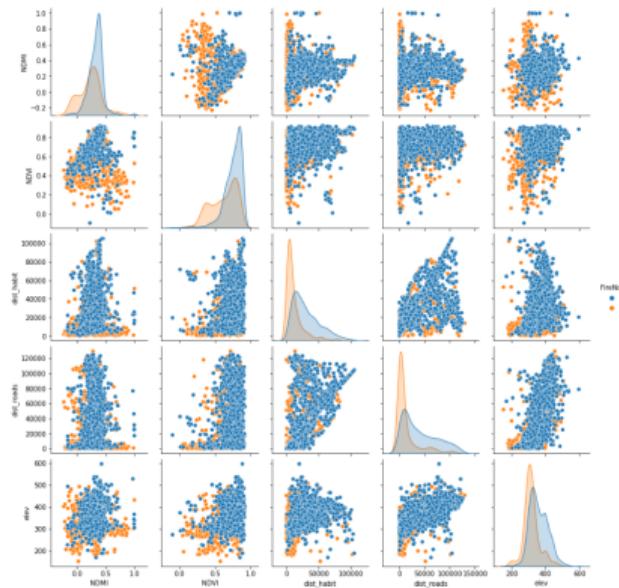
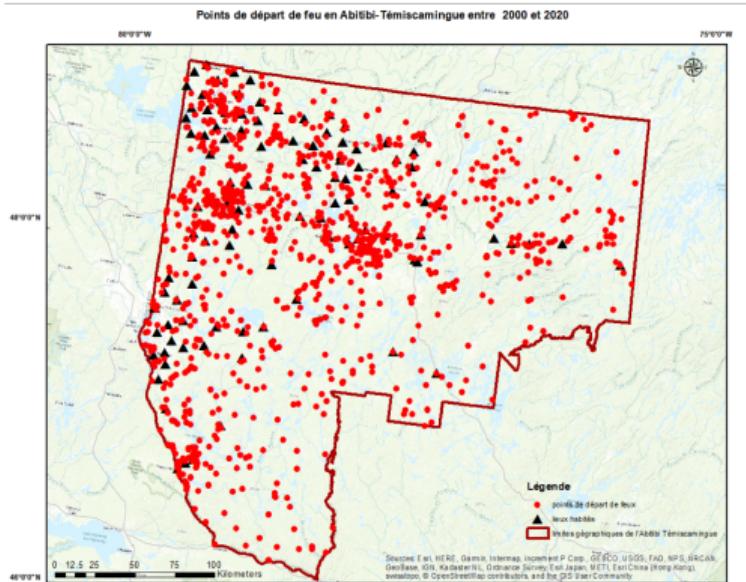
- [1] A. Benchallal, Y. Bouroubi, and A. Charpentier. "Predicting Wildfires in Québec". In: *Remote Sensing* (*To be submitted*) (2022).
- [2] France Info. "Canada : le "dôme de chaleur" provoque de violents incendies à Lytton". In: (2021).

Wildfire Canada [2]



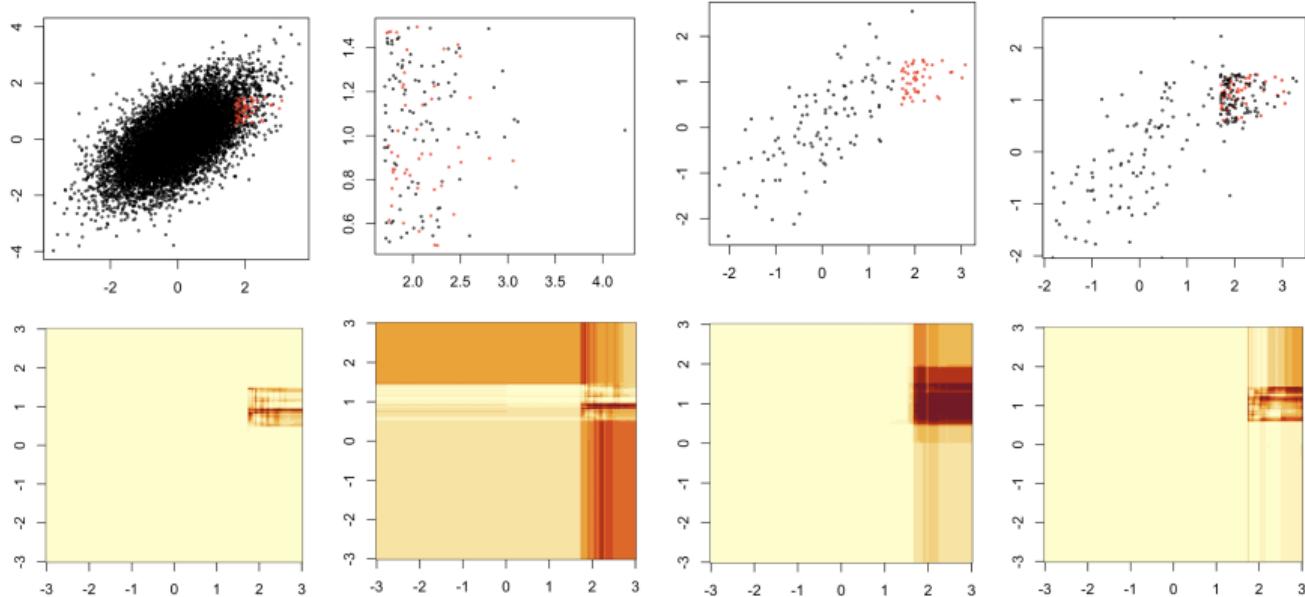
- [1] A. Benchallal, Y. Bouroubi, and A. Charpentier. "Predicting Wildfires in Québec". In: Remote Sensing (*To be submitted*) (2022).

Wildfire Canada [3]



Wildfire Canada [4]

Theoretical issues on creating a balanced from largely unbalanced data



[1] David Hand. *Dark Data: Why What You Don't Know Matters*. Princeton University Press, 2020.

Why do we need (more) detailed data ? [1]

➤ Central Limit Theorem

The sum of independent random variables, properly normalized, tends toward a normal distribution even if the original variables themselves are not normally distributed (with finite variance)

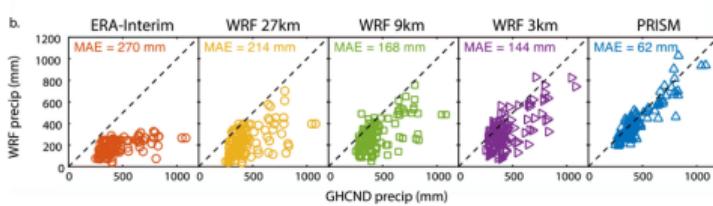
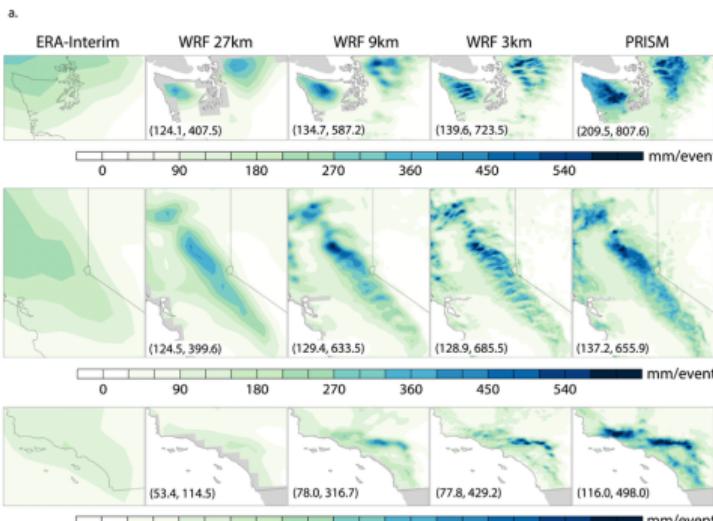
if X_i 's are i.i.d. with mean μ and variance σ^2 ,
$$\frac{(X_1 + \cdots + X_n)/n - \mu}{\sigma/\sqrt{n}} \approx \mathcal{N}(0, 1).$$

➤ Extreme Value Theory (Fisher & Tippett (1928))

there are a_n and b_n such that ,
$$\frac{\max\{X_1 + \cdots + X_n\} - a_n}{b_n} \approx GEV_\xi.$$

where ξ depends on the distributions of X_i 's ($\xi < 0$ Weibull domain of attraction: bounded X_i 's; $\xi = 0$ Gumbel domain of attraction: light tailed X_i 's; $\xi > 0$ Fréchet domain of attraction: heavy tailed X_i 's). E.g. $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $\xi \leq 0$.

Why do we need (more) detailed data ? [2]

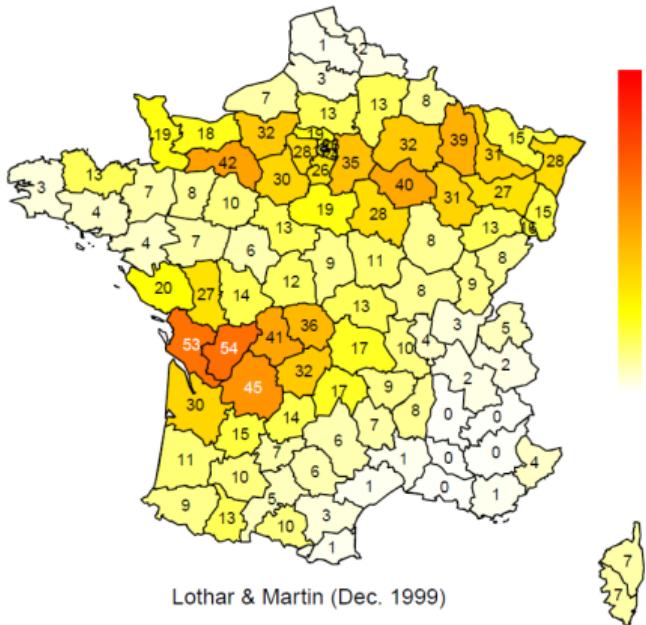


source picture:

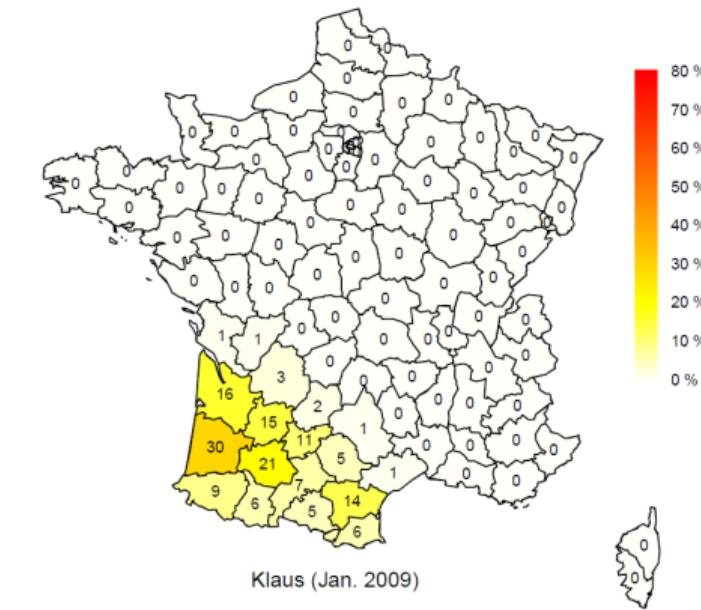
- [1] Xingying Huang et al. "Simulating and evaluating atmospheric river-induced precipitation extremes along the US Pacific Coast: Case studies from 1980–2017". In: *Journal of Geophysical Research: Atmospheres* 125.4 (2020), e2019JD031554.

"Some light-tailed atmospheric variables like temperatures are often averaged in space, time or both. The central limit theorem makes them very close to Gaussian variables. So, don't be surprised to find negative shape parameters for maxima of averaged values" (Philippe Naveau)

How to go further ?



Lothar & Martin (Dec. 1999)



Klaus (Jan. 2009)