Feature and Quantile Selection for the Actuarial Climate Index: Everything, Everywhere, All at Once

Insurance Data Science Conference - Bayes Business School, London - June 20, 2025

Mathias Valla José Garrido

Chaire DIALog, Fondation du Risque, Institut Louis Bachelier
I2M, Aix-Marseille University
Concordia University, Montreal, Canada

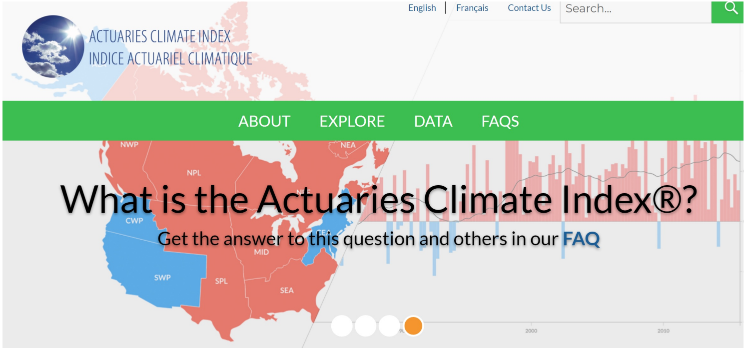
Outline

- Climate change challenges some fundamental principles of insurance : risk insurability, pooling, diversification, and risk transfer (Charpentier, 2008; Thistlethwaite and Wood, 2018; Courbage and Golnaraghi, 2022).
- Optimistic perspectives suggest that the insurance business could find in it an opportunity, through the development of new technical solutions (Rao and Li, 2023; Savitz and Gavriletea, 2019; Wagner, 2022).
- Pessimistically, climate change has forced already the strategic withdrawal of insurers in certain US markets, such as in California (Blood, 2023, APNews).

- In the Property & Casualty insurance sector, (Holzheu et al., 2021, SwissRe) forecasts increased frequency and severity of events due to climate change that will cost 30% to 63% more in insured catastrophe losses by 2040.
- This cost increase could even reach 90%–120% in specific markets, such as China, the UK, France and Germany.
- In all cases, it is important to add an actuarial perspective on the study of climate change and its impact on the insurance industry.

- The actuaries climate index™ (ACI) is linked to climate risks, a bit like the Consumer Price Index (CPI), which tracks the price variations of a basket of goods and services over time.
- Actuaries quantify and manage different types of risk. The ACI measures climate risk on the basis of a basket of extreme climate events and on variations of sea level.
- An index increase points to an increased number of extreme climate events.

Sponsors: American Academy of Actuaries (AAA), Canadian Institute of Actuaries (CIA), Casualty Actuarial Society (CAS) and Society of Actuaries.



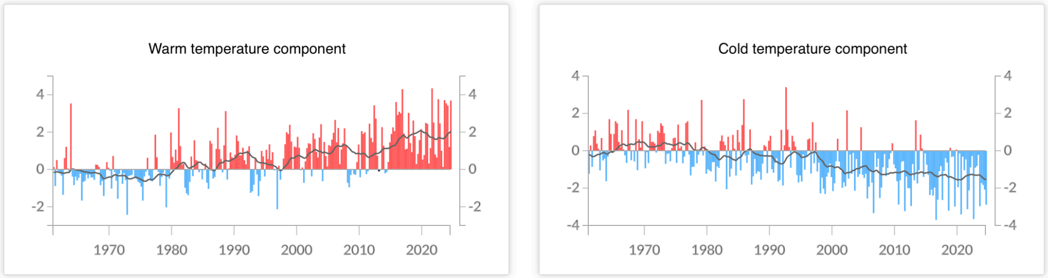
Source : https://actuariesclimateindex.org

The six Actuaries Climate Index components are:

1. Frequency of temperatures above the 90th percentile (*T90*);
2. Frequency of temperatures below the 10th percentile (*T10*);
3. Maximum rainfall per month in five consecutive days (*P*);
4. Annual maximum consecutive dry days (*D*);
5. Frequency of wind speed above the 90th percentile (*W*); and
6. Sea level changes (*S*).

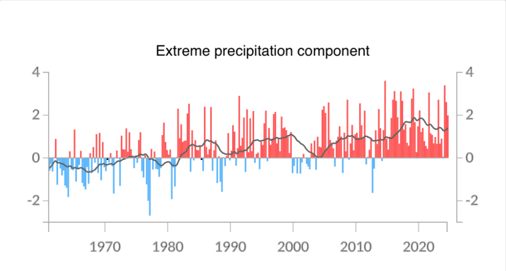
$$ACI = \frac{1}{6}(T90 - T10 + TP + CDD + W + S)$$

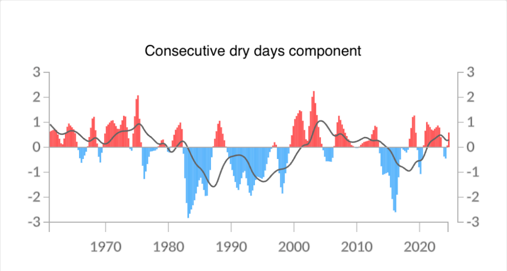
The ACI is traditionally computed as the unweighted average of standardized anomalies across multiple climatic variables.



The North-american temperature components

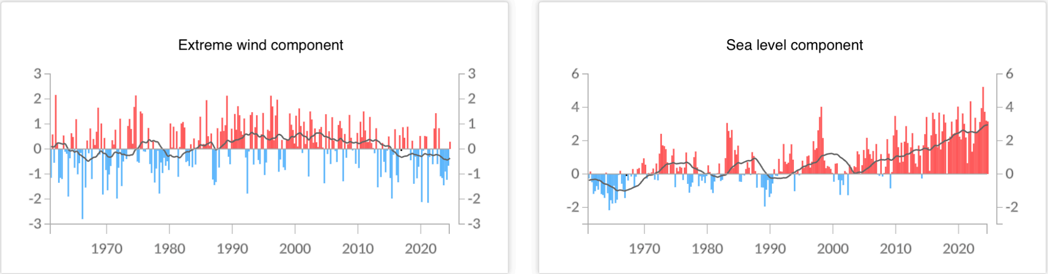
The ACI is traditionally computed as the unweighted average of standardized anomalies across multiple climatic variables.





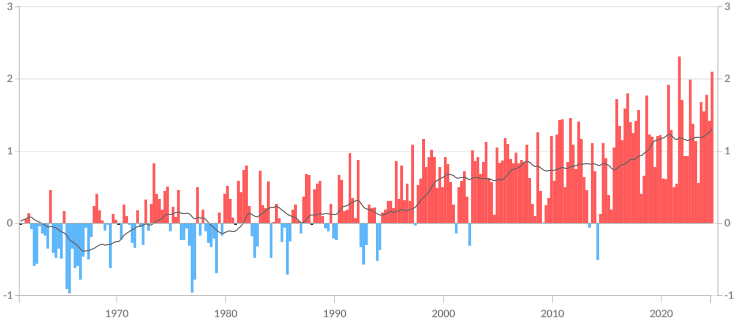
The North-american precipitation components

The ACI is traditionally computed as the unweighted average of standardized anomalies across multiple climatic variables.



The wind and sea level components

The ACI is traditionally computed as the unweighted average of standardized anomalies across multiple climatic variables.



The North-american Actuarial Climate Index

The ACI is traditionally computed as the unweighted average of standardized anomalies across multiple climatic variables.

KEY CHARACTERISTICS

- *Everything:* tries to encompass all insured risks.
- *Everywhere:* tries to be a consistent risk measure for any region of the world.
- *All at Once:* gives a global measurement of all risks all together at each time step.

The ACI is traditionally computed as the unweighted average of standardized anomalies across multiple climatic variables.

KEY GAPS

- *Everything ?* $\rightarrow$ do not adapt easily to specific risks.
- *Everywhere ?* $\rightarrow$ should be adapted to local characteristics.
- *All at Once ?* $\rightarrow$ do not account for compound events not interdependencies between climate anomalies.

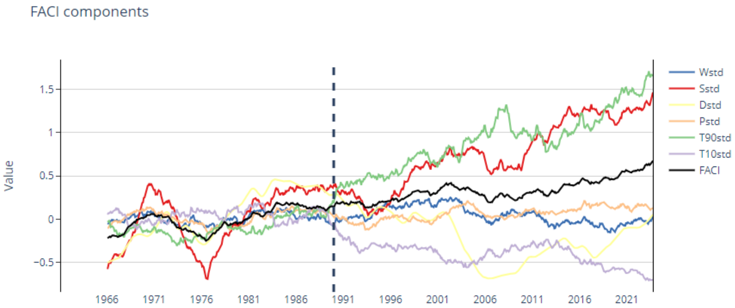
Limitations of the standard ACI

- Uniform feature weighting $\Rightarrow$ dilutes signals from key predictors.
- Ignores inter-feature dependencies $\Rightarrow$ over- or under-counts correlated signals.
- No feature and quantile selection $\Rightarrow$ we might miss crucial information on extremes.

Our research question for today

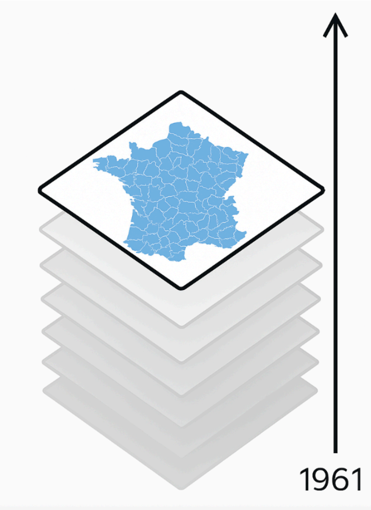
- Are the features and quantiles, as selected in the original ACI methodology, well chosen for specific risks, at specific places ?

Applying the ACI metodology over Metropolitan France (see (Garrido et al., 2023)), using x-aci (see our Python and R packages)



The French Actuarial Climate Index

- Climate features and standardized anomalies computed at various quantiles
  $\rightarrow$ from ERA5 data and standardized using the x-ACI package
- A risk feature: such as the excess of mortality computed using death counts from INSEE (French National Institute of Statistics and Economic Studies) and the methodology from Santé Publique France (the french National Public Health Agency)
- Monthly measurements, for all 96 departments of Metropolitan France



We chose to challenge the selected features and quantiles by giving the risk feature as a target variable to a model, fed with all climate features defined at various quantiles. We implemented that in 2 approaches:

LASSO

- Basic approach for regression/prediction
- For all time points t, we fit a LASSO on all data < t
- every LASSO goes through a 50 cross-validation process
- at each time, we count the number of time a given feature has been selected by the model

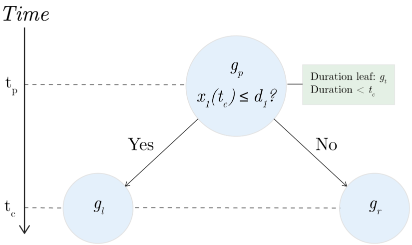
TpT

- A tree-based model for longitudinal data, for subgroup analysis
- Fit a TpT on the data.
- Use split importance to weight features over time.
- Assess optimal quantile thresholds by time-penalised gain.

Quickly: what is à TpT ?

A tree-based model for longitudinal covariates, or when the predictors and target are time series.
$\Rightarrow$ see (Valla, 2024; Valla and Milhaud, 2025)

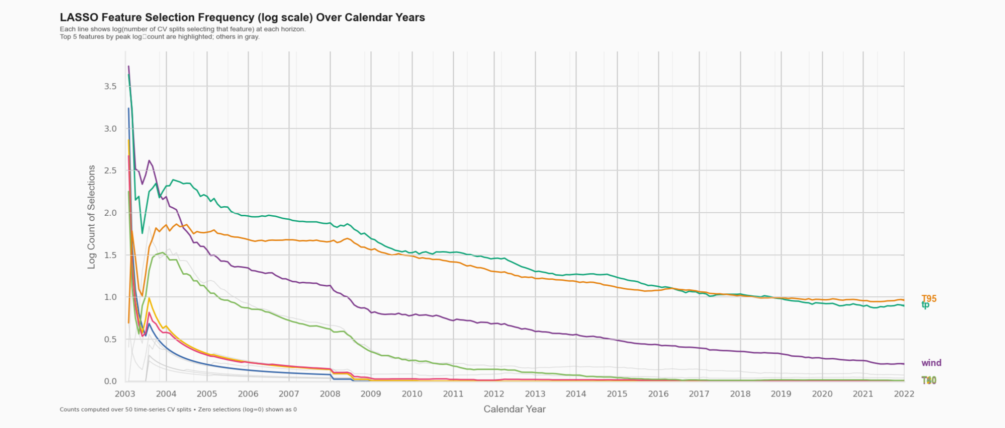
- Splits in $\mathcal{X} \times \mathcal{T}$: joint covariate-time partitions.
- Exponential penalty $e^{-\gamma \Delta t}$ applied to gain: discourages late splits.
- Each subject follows a single path maintains temporal coherence.



$$G_\gamma(g_p; g_l, g_r, g_t) = \left[I(g_p) - \left(\frac{\mathcal{N}(g_l)}{\mathcal{N}(g_p)} I(g_l) + \frac{\mathcal{N}(g_r)}{\mathcal{N}(g_p)} I(g_r) + \frac{\mathcal{N}(g_t)}{\mathcal{N}(g_p)} I(g_t) \right) \right] \cdot e^{-\gamma \cdot (t_c - t_p)}$$

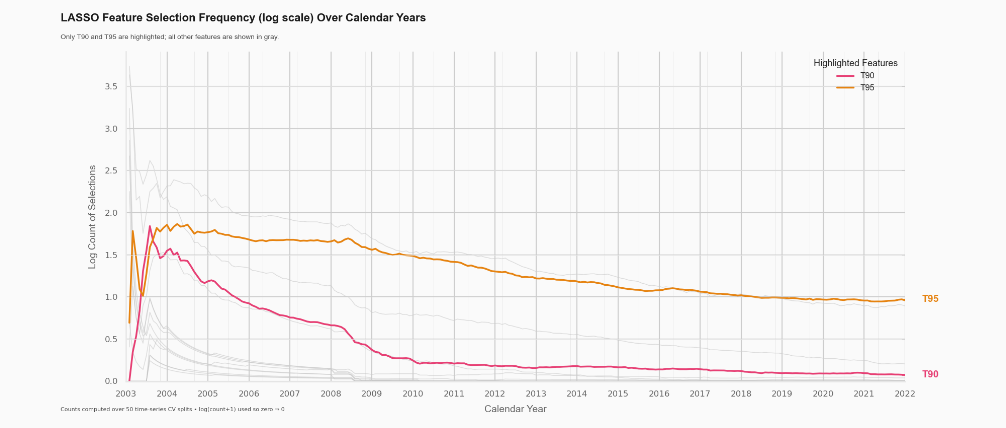
Single TpT split (from Valla & Milhaud, 2025)

- **LASSO:** T95 and precipitation covariates dominate. As we update the feature importance, the number of selected features decreases: the model gets simpler.



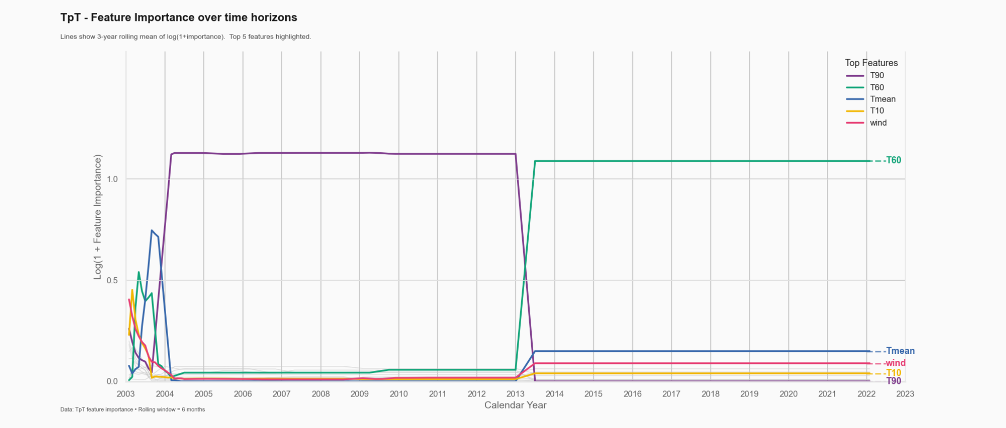
Chosen quantiles by TpT split gains

- **LASSO:** zoom on temperature quantiles. T90 had a dominating effect during the summer of 2003 but after that, T95 is consistently selected.



Chosen quantiles by TpT split gains

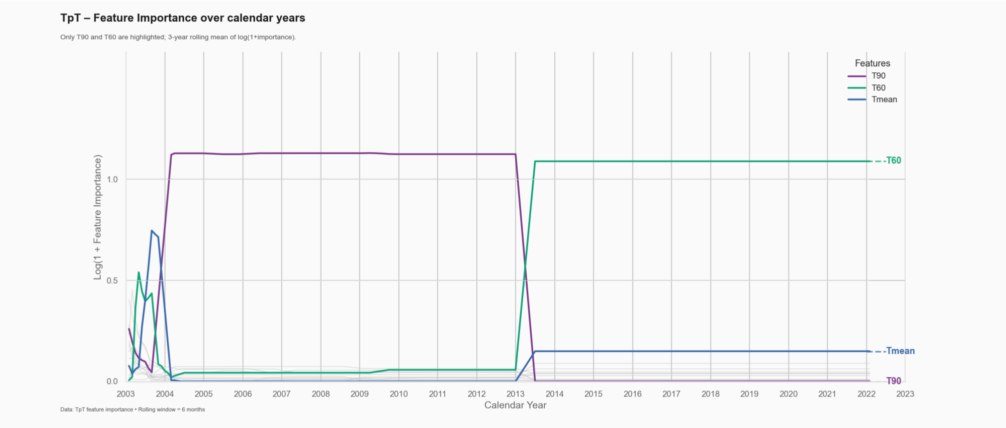
- **TpT:** The feature importance are much more stable with time. Temperature feature dominate in terms of importance.



Chosen quantiles by TpT split gains

Results for Excess Mortality: Quantile Selection

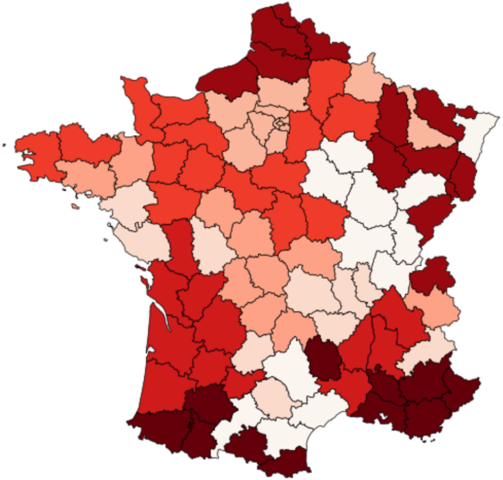
- **TpT:** zoom on temperature quantiles. T90 dominates until 2013, then a combination of T60 and Tmean is selected.



Chosen quantiles by TpT split gains

Results for Excess Mortality: Quantile Selection

- In both cases, the choice of analysing the frequency of events above the 90th quantile is far from clear to assess the excess of mortality.
- For regression purposes, the LASSO procedure shows that higher quantiles (T95) may be an even better threshold.
- For classification and subgroup analysis, the TpT procedure shows that lower quantiles (T60, Tmean) allows for a better separation into homogenous subgroups. If we force TpT to make clusters of at least 10 departments, T95 becomes the most informative covariate and it yield the following clustering:



What are our contributions:

- R and Python packages to replicate the ACI methodology on any region of the world (applied to Metropolitan France here)
- a new tree-based model methodology to analyse longitudinal/time series data

In this application, we simply wanted to challenge the choices of the original ACI methodology to make it more flexible:

- a better definition of the kind of anomaly that matter (quantile selection)
- a more climate science-driven choice of feature

In the future:

- we'd like to tackle all other identified gaps in the ACI methodology, namely the interdependencies/compounding effect of the features, the ability to be adjusted to any region and to be projection in the future.

Toward a true *Everything, everywhere, all at once*...

Blood, M. R. (2023). California insurance market rattled by withdrawal of major companies. *AP News*.

Charpentier, A. (2008). Insurability of climate risks. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 33:91–109.

Courbage, C. and Golnaraghi, M. (2022). Extreme events, climate risks and insurance. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 47:1–4.

Garrido, J., Milhaud, X., and Olympio, A. (2023). The definition of a French actuarial climate index; one more step towards a European index. working paper or preprint.

Holzheu, T., Lechner, R., Vischer, A., Bevere, L., Staib, D., Finucane, J., Belgibayeva, A., and Fan, I. (2021). More risk: the changing nature of P&C insurance opportunities to 2040. Technical report, Swiss Re Institute.

Rao, S. and Li, X. (2023). *Chinas Experiences in Climate Risk Insurance and Suggestions for Its Future Development*, pages 157–171.

Savitz, R. and Gavriletea, M. D. (2019). Climate change and insurance [klimato kaita ir draudimas]. *Transformations in Business and Economics*, 18:21–43.

References II

Thistlethwaite, J. and Wood, M. O. (2018). Insurance and climate change risk management: Rescaling to look beyond the horizon. *British Journal of Management*, 29:279–298.

Valla, M. (2024). Time-penalised trees (tpt): introducing a new tree-based data mining algorithm for time-varying covariates. *Annals of Mathematics and Artificial Intelligence*.

Valla, M. and Milhaud, X. (2025). Time-penalized trees: consistency results and simulations. working paper or preprint.

Wagner, K. R. H. (2022). Designing insurance for climate change. *Nature Climate Change*, 12:1070–1072.