Joint Extremes of Precipitation and Wind Speed

Peiman Asadi (IAG), Morris Chen (UNSW), and Qihe Tang (UNSW)

All Actuaries Summit, Gold Coast, Australia, 1-3 May 2024

Outline

Outline

Current practice for insurance pricing

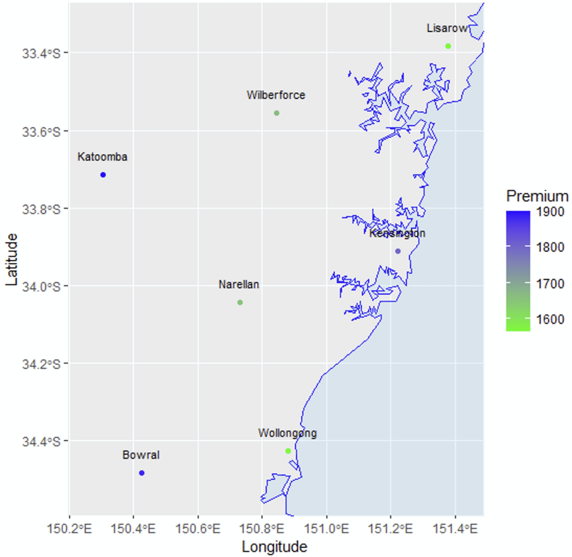
- Generalized linear models (GLMs) are widely applied in general insurance pricing (e.g., house insurance premium).
- A typical GLM is of the form

$$g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}.$$

  Such a GLM can be used to estimate both severity and frequency.
- One of the most important factors is region. Two identical houses merely separated by a street could be charged significantly different premiums if they fall under different regions.

Actuaries Institute.

Quotes of house insurance premium



- Except for storm risk, all other variables, including sum insured, building materials, etc., are set to be identical for properties.
- Other risk factors including bushfire and flood are excluded.
- Some observations ...

Motivation

- In this example, storm is a significant determinant of house insurance premiums.
- Primary risk factors for losses include precipitation and wind speed, both of which are region-specific.
- However, the levels of precipitation and wind speed cannot differ significantly for locations that are close to each other.
- Therefore, we aim to create a smoothed pricing surface to mitigate the issue of sudden variations in premiums.

Actuaries Institute.

Why climate data

- In current insurance practice, climate risks are not priced.
- Due to the variability and uncertainty of storms, there is no advanced storm model to utilize in the insurance market.
- For modeling purposes, insurers mainly rely on claim histories.
- We explore whether the use of abundant climate data could improve insurance pricing.
- Our study helps to understand the impact of climate change on insurance pricing.

Actuaries Institute.

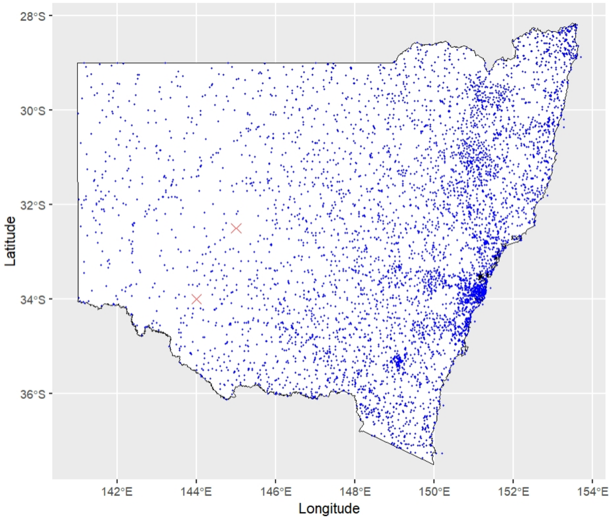
Outline

Description of data

- Time series data from the Australian Bureau of Meteorology (BOM)
- $N > 5,000$ climate stations
- Records are heterogeneous in nature: differences observed in recorded variables, record frequency, and the years of records. Luckily, both precipitation and wind speed are generally available.
- Non-negligible portions of missing data, due to manual records and station closures

Actuaries Institute.

Plot of available stations



- The stations are distributed across NSW, with a higher concentration along coastal areas but lower in inland regions.
- How to price a contract in areas marked X?

Actuaries Institute.

Thoughts to utilize the data

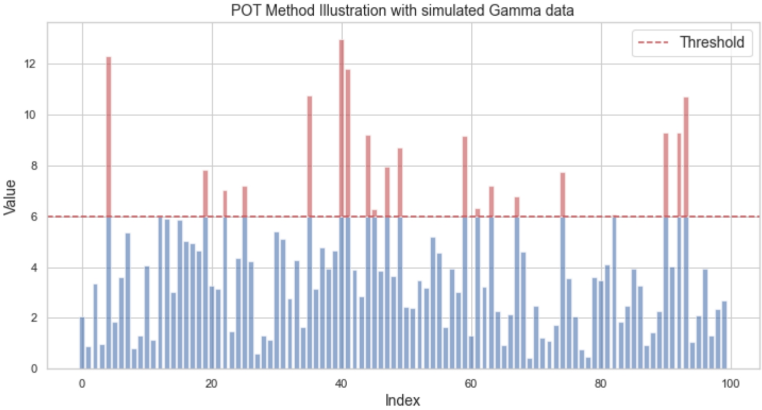
- We utilize daily records for total precipitation and maximum wind speed because:
  - they are the two most damaging factors to buildings;
  - there are also most data points available for analysis.
- We are concerned about tail risk.
  - The devil is in the tails (Donnelly and Embrechts, 2010)
  - The sting is in the tail (Talk by Alan Greenfield, Sarah Wood, and Simon Bradshaw yesterday)
- We employ extreme value theory (EVT) to analyze the data.

Actuaries Institute.

Outline

The peaks-over-threshold (POT) approach

The POT approach involves fitting a parametric model to the exceedances over a certain high threshold.



Generalized Pareto distribution (GPD)

- For a loss X, the exceedance $Y = X - u$ over a certain high threshold u approximately follows a GPD.
- The standard form of the GPD:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi},$$

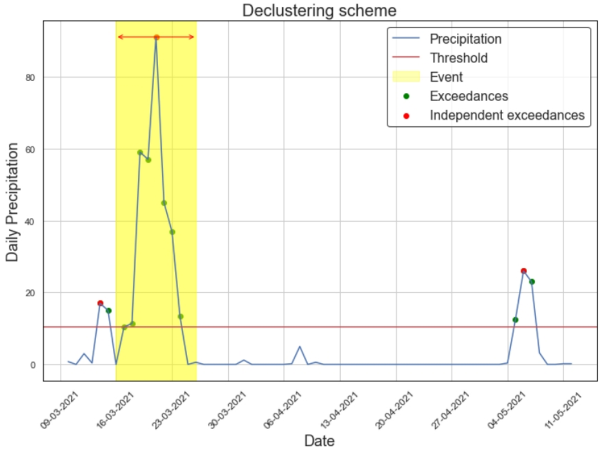
  where ξ is the shape parameter and $\sigma > 0$ is the scale parameter.
- The shape parameter ξ decides the tail behavior:
  - $\xi > 0$: fat tail
  - $\xi = 0$: exponential
  - $\xi < 0$: thin tail

Actuaries Institute.

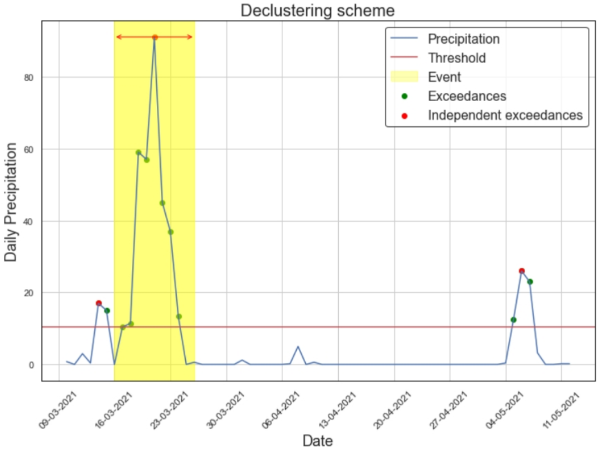
Declustering

- Extreme precipitations tend to occur in clusters as a result of strong temporal dependence within weather observations.
- This dependence between exceedances must be removed prior to fitting threshold models.
- Declustering is a common practice in dealing with climate data.

The automated declustering scheme



The automated declustering scheme



It is common to assume a running window with a fixed width p for all stations. However, . . .

Visualizing the shape parameters

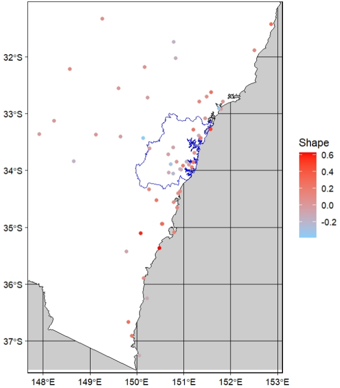


Figure: Shape for precipitation

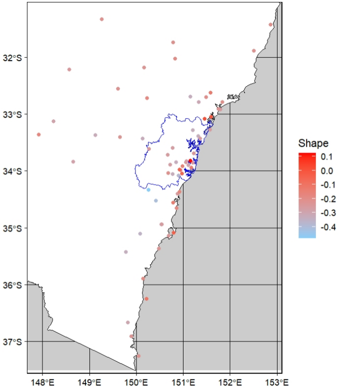


Figure: Shape for wind speed

Actuaries Institute.

Visualizing the 50-year return levels

The 50-year return level is defined as the level that is exceeded on average once every 50 years.

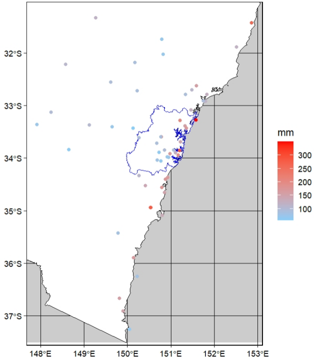


Figure: Return level of precipitation

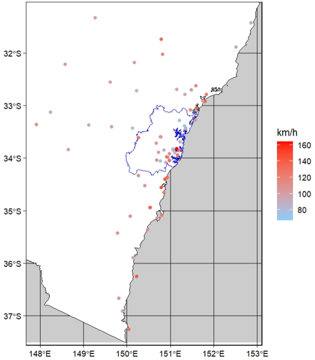


Figure: Return level of wind speed

Actuaries Institute.

The smoothing approach

The R-INLA package is applied to create a smoothed surface.

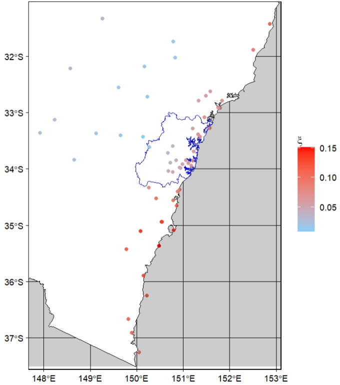


Figure: Shape for precipitation

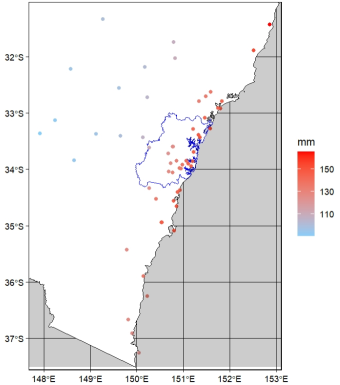


Figure: Return level of precipitation

Outline

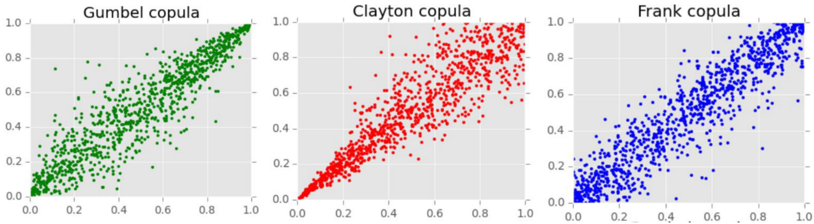
Multivariate EVT

- Large losses are driven by individual risk factors and their tail dependence. The latter leads to the so-called compounding effect, which is often overlooked.
- Tail dependence is much harder to perceive and quantify.
- Multivariate EVT = Univariate EVT + Tail Dependence
- Insurance pricing should take account of both individual tail risks and their tail dependence.

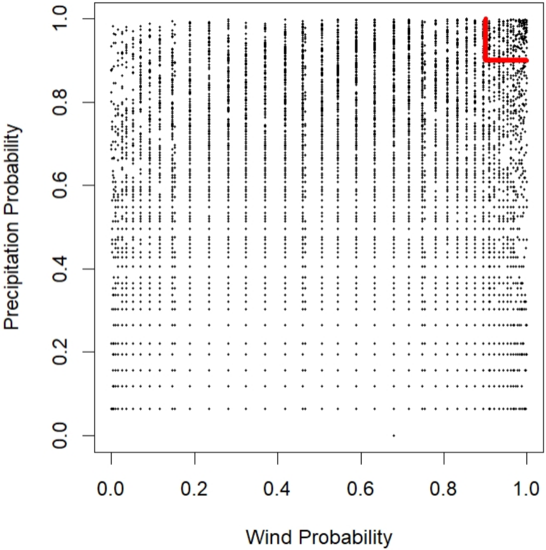
Tail dependence

For two losses X and Y with distributions F and G, respectively, their tail dependence coefficient is defined as

$$\chi = \lim_{u \to 1} P\left(F(X) > u \middle| G(Y) > u\right).$$



Plot of the tail dependence



Map of tail dependence

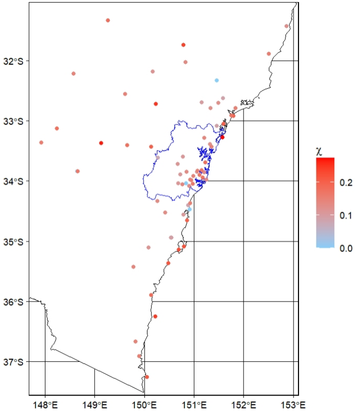


Figure: Tail dependence

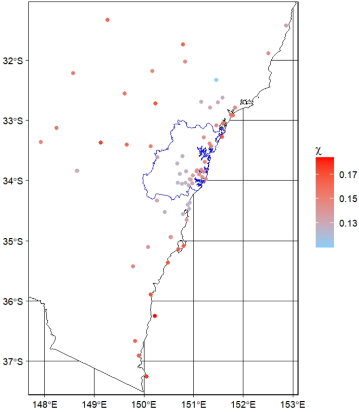


Figure: Smoothed tail dependence

Actuaries Institute.

Outline

Takeaways

- The use of climate data could potentially improve insurance pricing.
- A smoothed surface can help to mitigate the issue of sudden variations in premiums.
- Insurance pricing should take account of both individual tail risks and their tail dependence.
- EVT is a powerful tool for pricing climate risks. In doing so, we often need to do declustering.

Thank you very much!!

Actuaries Institute.