ICA CIA BERLIN 2018

# The Inter-Arrival Time approach to Clustering

S. Reimann; D. Baumgartner, Swiss Re

# Agenda

The inter-arrival time approach to clustering

Clustering in empirical data: Atlantic Hurricanes

A model for clustering

Implications and further research

Appendix

# The inter-arrival time approach to clustering



- Intuitively, we characterise clustering by the following property: "An event makes the next one more probable."
- In general, (re)insurance contracts are sensitive to clustering of loss generating events.
- The degree of clustering affects risk assessment as well as risk capital allocation in a 1-year view.

# Risk estimation based on the distribution of annual losses

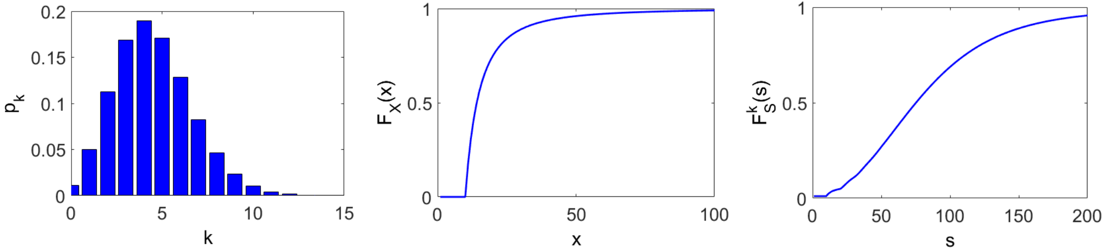


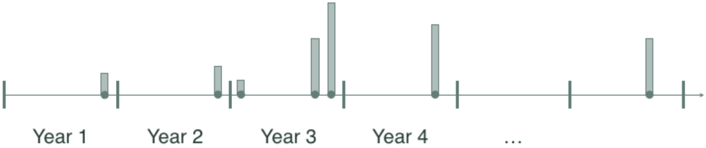

Figure: frequency distribution & severity distribution = annual loss distribution

# Explicit time makes the difference



Generate timestamps for events on the real time axis


Allocate a specific event loss to each timestamp



## How to construct a sequence of event times from empirical data?

ICA CIA BERLIN 2018

- The IAT approach is based on inter-arrival times (IAT) $\tau$.
- The time series of events (arrivals) is generated from the empirical distribution of inter-arrival times $\varphi_\tau$.
- The process admits clustering if the distribution of IAT's is heavy-tailed.
- Let $\tau$ be the time till the next event and $t$ the time since the last. Then

" If $\varphi_\tau$ *is heavy − tailed*, then $E[\tau|t]$ increases in $t$."

(Sornette & Knopoff, 1997)

# Atlantic Hurricanes

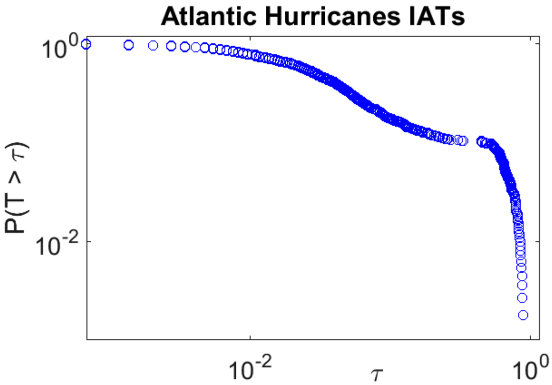
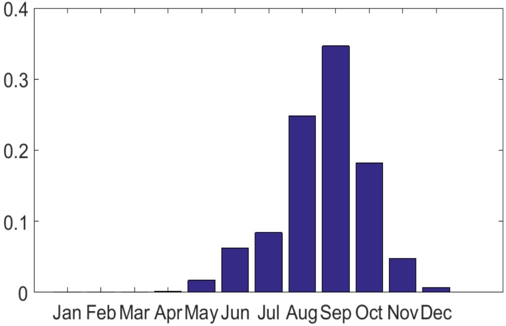
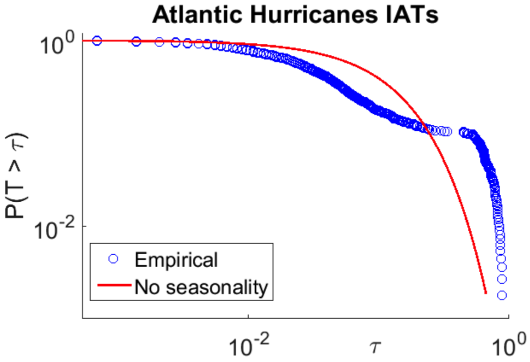


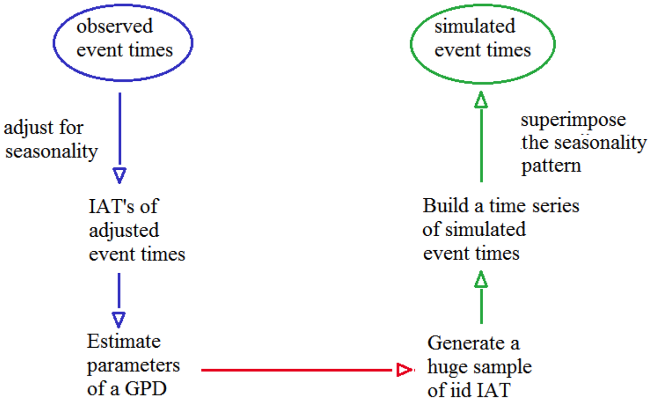

Figure: [HURDAT, 1891 - 2008]

# Adjusting for seasonality

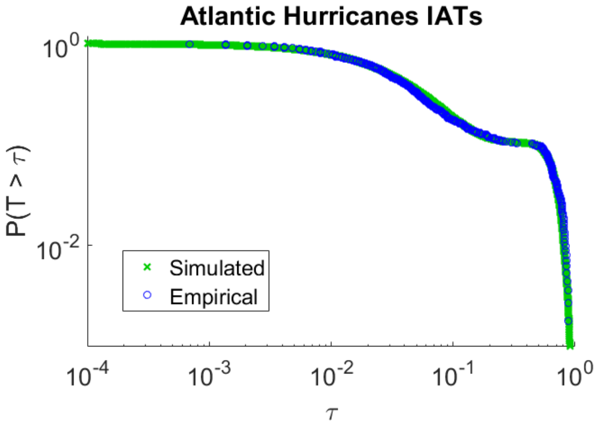


# From empirical to modelled IATs



# Modelled data matches empirical data



# The general Pareto distribution is a generic model

ICA CIA BERLIN 2018

- Empirical data provides statistical evidence that the *generalised Pareto Distribution* (gPD) is a superior model.

$$\text{if } \tau \sim \text{Exp}(\lambda) \,\Big|\, \lambda \sim \Gamma(\alpha, \beta) \text{ then } \tau \sim \text{gPD}(\xi, \sigma, 0).$$

  - shape parameter $\xi = 1/\alpha$, $0 \leq \xi < 1$;
  - scale parameter $0 < \sigma = \beta/\alpha$.

# Implied properties

ICA CIA BERLIN 2018

- *"The longer it has been since the last event 't' the longer is the expected time τ till the next."*
- The gPD admits clustering

$$E[\tau|t] = \frac{\sigma}{1-\xi} + t\,\frac{\xi}{1-\xi}$$

- The gPD admits over-dispersion due to

$$\delta_\xi := \frac{VAR(\tau)}{MEAN(\tau)^2} = \frac{1}{1-2\xi} \geq 1, \qquad 0 \leq \xi < \frac{1}{2}$$

# Possible model extensions



- Introduce truncation of IAT distribution
  - Lower bound due to temporal resolution in empirical data
  - Upper bound, e.g. due to periodicity
- Include peril specific external drivers
  - Atlantic Multidecadal Oscillation, El-Niño Southern Oscillation
  - Business activity cycle
  - Trends

# The negBin as a frequency model

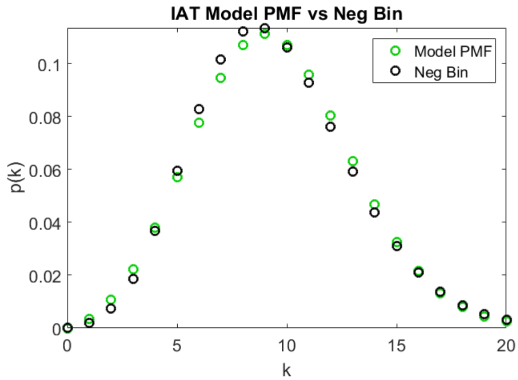
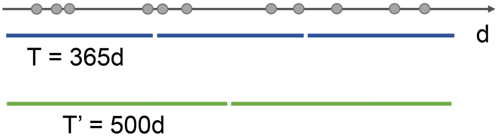



Figure: The PMF of the 1-year truncated gPD model approximates a negBin.

# Memory: Some implications

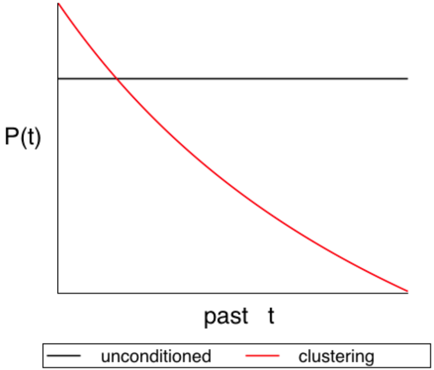
- Risk assessment based on samples drawn from 1 year disregards inter-year correlation.
- The economic 1 year view imposes an external time scale.




- Should the 1-year view on risk be conditioned on the past?

$$P(t) = E[\tau \leq 1 \mid t]$$

# Conditioning on the past alters expectations about future arrivals



# Further research

- Relation between the IAT approach and cluster analysis
- Scale dependence: In contrast to Poisson processes, processes with memory are susceptible to externally imposed time scales. Thus, results in general depend on the observation period.
- Inter-arrival times might depend on event severities: "Do strong storm cluster more than weak ones?"
- Multiperils: How to describe inter-dependencies between different perils in the IAT frame work?

## Thank you very much for your attention!

Contact details:
**Stefan Reimann** *)
**David Baumgartner** +)

mail*): reimannst@gmail.com
mail+): David_Baumgartner@swissre.com

# References



- D. Sornette, L. Knopoff: "The Paradox of the expected Time until the Next Earthquake", Bull. Seism. Soc. Am., Vol. 87 (1997), pp. 789-798
- R. Winkelmann, "Duration Dependence and Dispersion in Count-Data Models", Journal of Business & Economic Statistics, 1995, vol. 13, issue 4, 467-74