SCaHA: Analysis of serial clustering and homogeneity

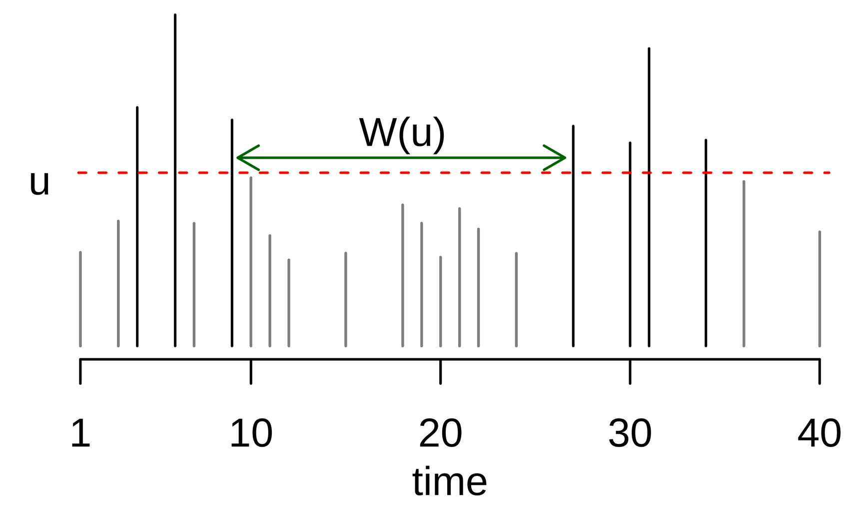
Merle Mendel, Roland Fried (TU Dortmund), Erik Haufs, Axel Bücher (Ruhr Uni Bochum)

Module B.3.3

Part I:

Models for Return Times

Objective: Identify suitable models for return times of different weather extremes.
Question: Which distributions for the return times between peaks over a high threshold u?



- Traditionally, POT often assumes a Poisson process (PP), where return times $W(u)$ of exceedances are exponentially distributed:

$$W \sim \text{Exp}(1/\sigma)$$

- In practice: temporal clustering of extreme events, i.e., short periods with several extreme events and long periods without

Return Times of Extratropical Cyclones

- Blender et al. [1] suggest modeling midlatitude winter cyclones with the fractional Poisson process (FPP), which assumes Mittag-Leffler distributed return times; $W \sim \text{ML}(\beta, \sigma)$

↳ β: tail parameter, controls the heavy-tailedness, σ: scale parameter

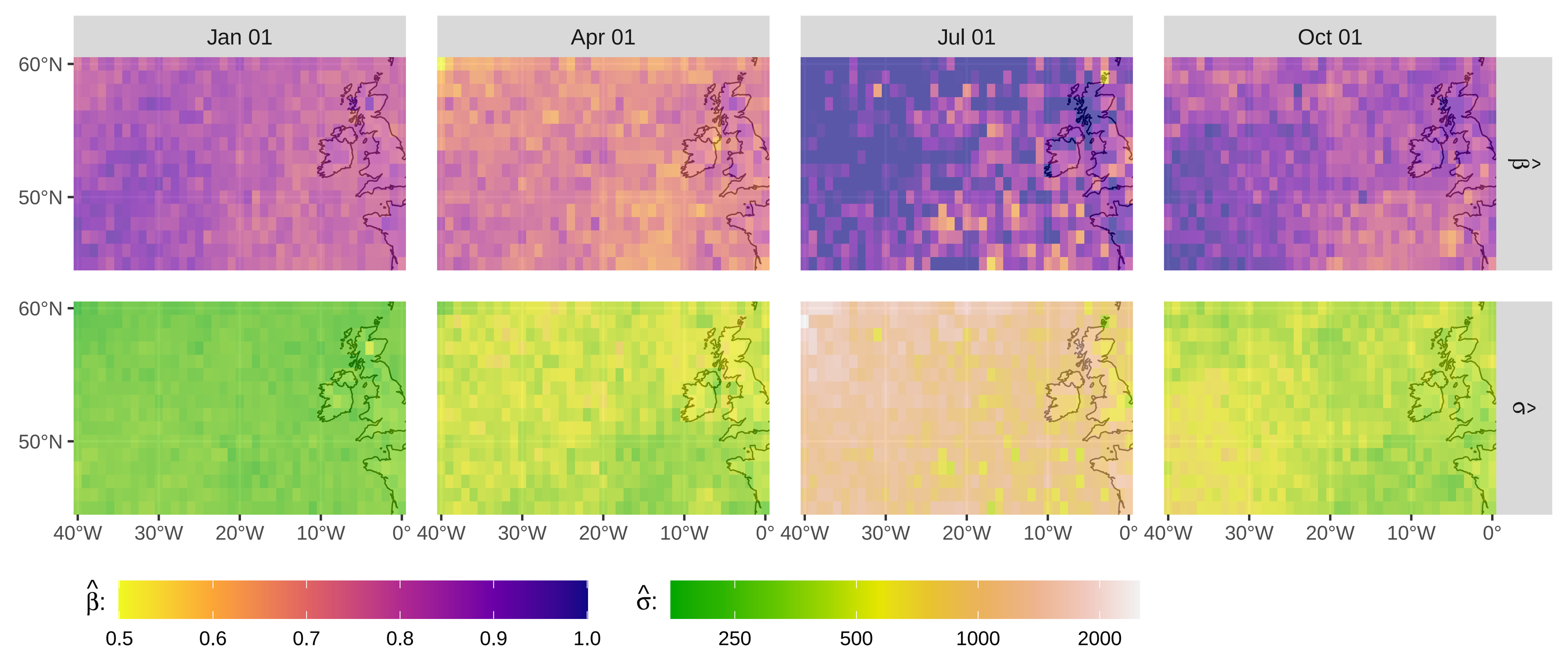
- Ferro & Segers [2] suggest Compound Poisson process (CPP), a mixture between the Dirac measure at zero and the exponential distribution: $W \sim (1-\theta)\cdot\varepsilon_0 + \theta\cdot\text{Exp}(1/\sigma)$
- Mathieu et al. [3] generalizes these approaches to the fractional compound Poisson process (FCPP); $W \sim (1-\theta)\cdot\varepsilon_0 + \theta\cdot\text{ML}(\beta, \sigma)$

↳ These approaches take clustering into account but do not regard seasonality or trend

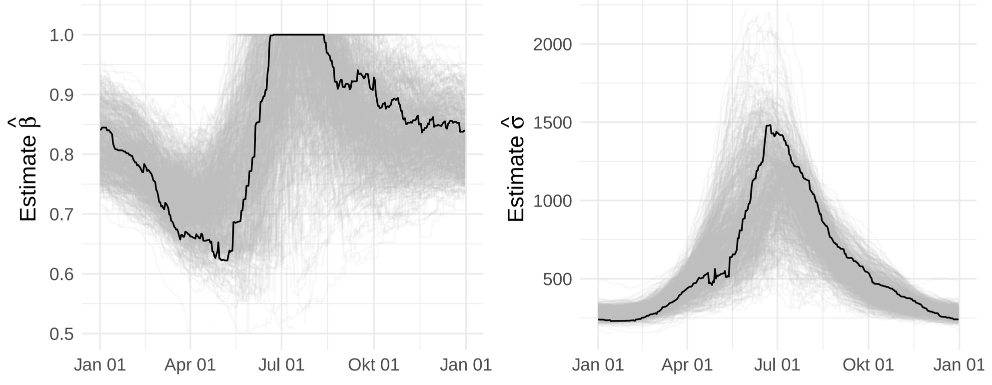
Objective: Incorporate seasonality using time dependent parameters $\beta(t)$, $\sigma(t)$
Approach: Weighted versions of the estimators for β and σ with distance-based weighting of the observations using an Epanechnikov kernel
Example: Return times of midlatitude cyclones

Scan for paper:

- Dataset: Relative vorticity of ERA5, 1940-2023, 1°×1° coordinate grid
- Threshold: 0.99 quantile calculated for each location separately
- Calculate daily parameter estimates $\hat{\beta}(t)$ and $\hat{\sigma}(t)$, $t = 1, \ldots, 365$
- Estimates of β (top) and σ (bottom) on January 1, April 1, July 1 and October 1 (from left to right)



Annual course of $\hat{\beta}$ (left) and scale parameter $\hat{\sigma}$ (right) of all coordinate points; estimates for 56° N, 40° W are shown in black.



- Large fluctuations of $\hat{\beta}$ and $\hat{\sigma}$ within a year, rather similar annual courses at the different locations
- For most coordinate points: minimum of $\hat{\beta}$ in April or May and maximum in July or August
- $\beta = 1$: FPP reduces to PP
- For most coordinate points: minimum of $\hat{\sigma}$ in December and maximum in July

Ongoing Research

1. **Extreme heat days:** Return times are described well by a zero-inflated discrete Weibull (ZIDW) distribution

   Currently: Analyze the type of trend and seasonality needed for appropriate modeling
2. Searching homogeneous parameter regions for extremes in relative vorticity and heat to improve parameter estimation

References

[1] Blender, R., Raible, C. C. & Lunkeit F. (2014): Non-exponential return time distributions for vorticity extremes explained by fractional Poisson processes. *Quarterly Journal of the Royal Meteorological Society*, 141 (686)

[2] Ferro C.A.T. & Segers J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**: 545-556.

[3] Mathieu, C., Mendel, M., Hees, K. & Fried, R. (2026): Models for temporal clustering of extreme events with applications to mid-latitude winter cyclones. https://arxiv.org/abs/2308.14625

[4] Mendel, M. & Fried, R. (2026): Parameter Estimation and Seasonal Modification of the Fractional Poisson Process with Application to Vorticity Extremes over the North Atlantic. *Data Science in Science*, 5 (1)

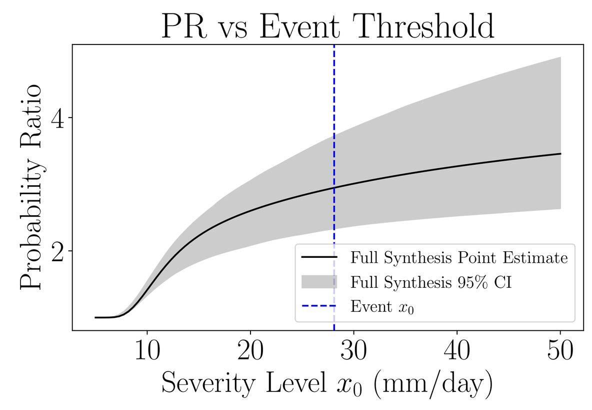
Part II: Synthesis Strategies & Sliding Blocks

Synthesis for Extreme Event Attribution

Situation: Different observational / model data products provide estimates on: *How did an extreme event change under global warming?*
Question: How to combine results from different data products?
Contribution: We provide **parameter-level** synthesis and contrast it to existing methods (World Weather Attrribution, WWA)

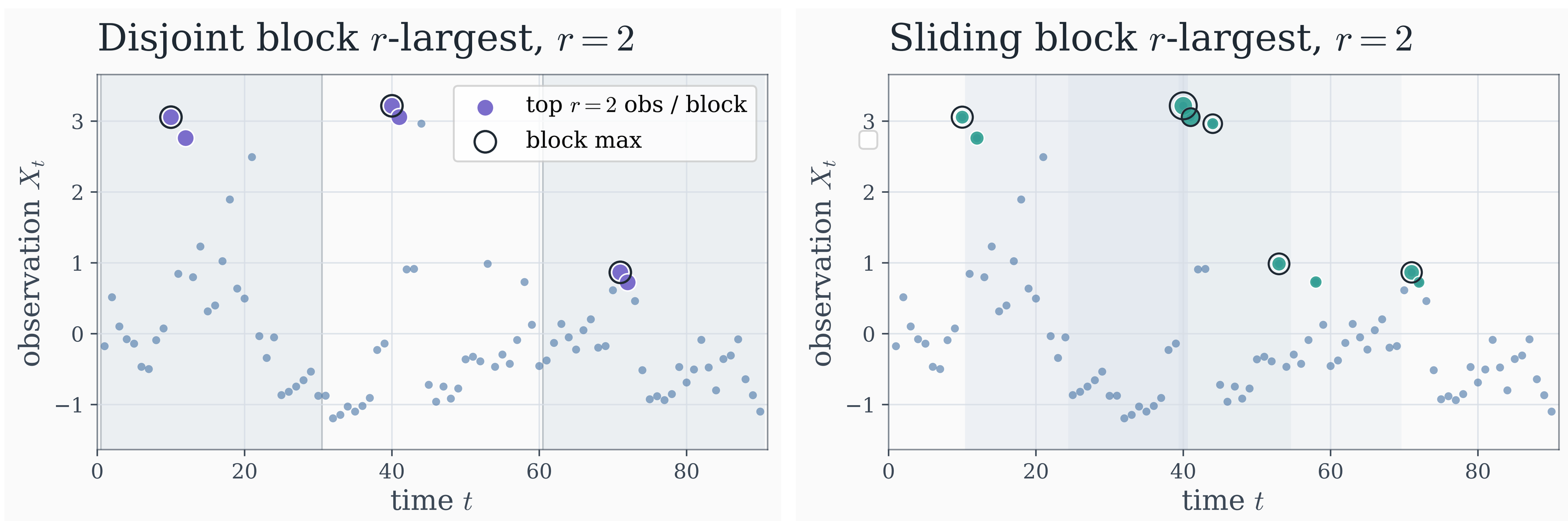
Results:

Method	Var.	Bias^2	MSE	Cover.	Length	Score
Parametric	1.30	0.02	1.32	0.9520	0.5219	0.6181
WWA	1.17	0.50	1.67	0.9977	2.6064	2.6227



Top-Two of Stationary Time Series

Situation: The *block maxima method* usually studies *just the block maximum*
Question: Can we utilitize *more than one extreme per block?*
Contribution: We develop estimators based on **Top-Two values**



Results:

- Top-Two estimators typically outperform max-only
- Sliding block top-two outperform disjoint blocks

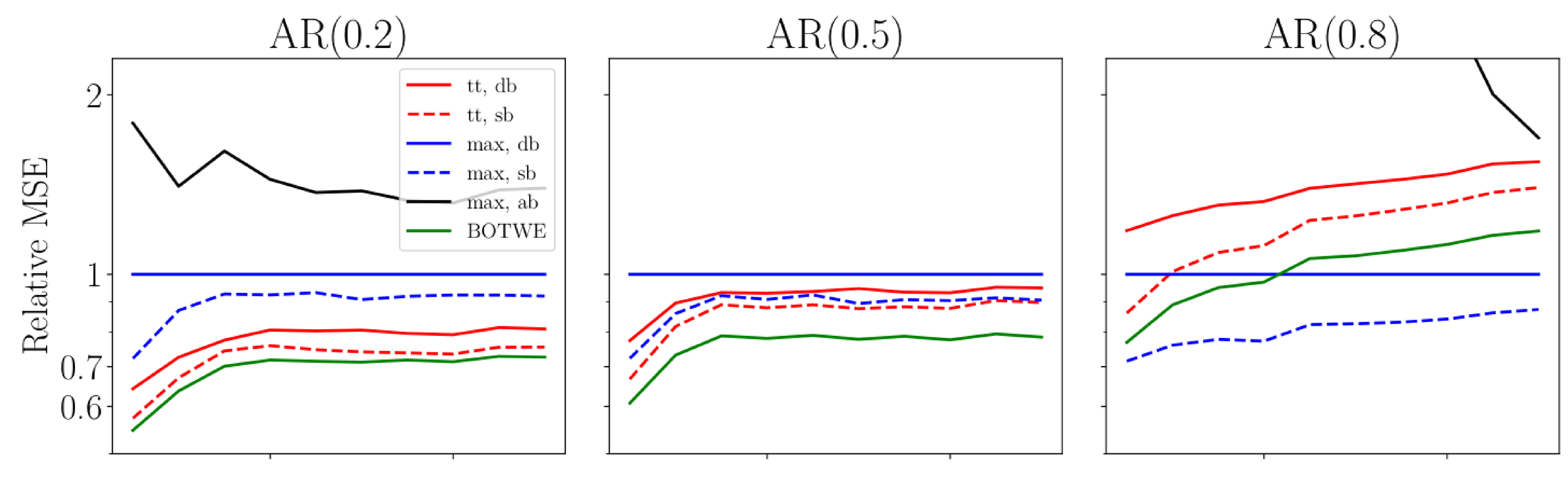
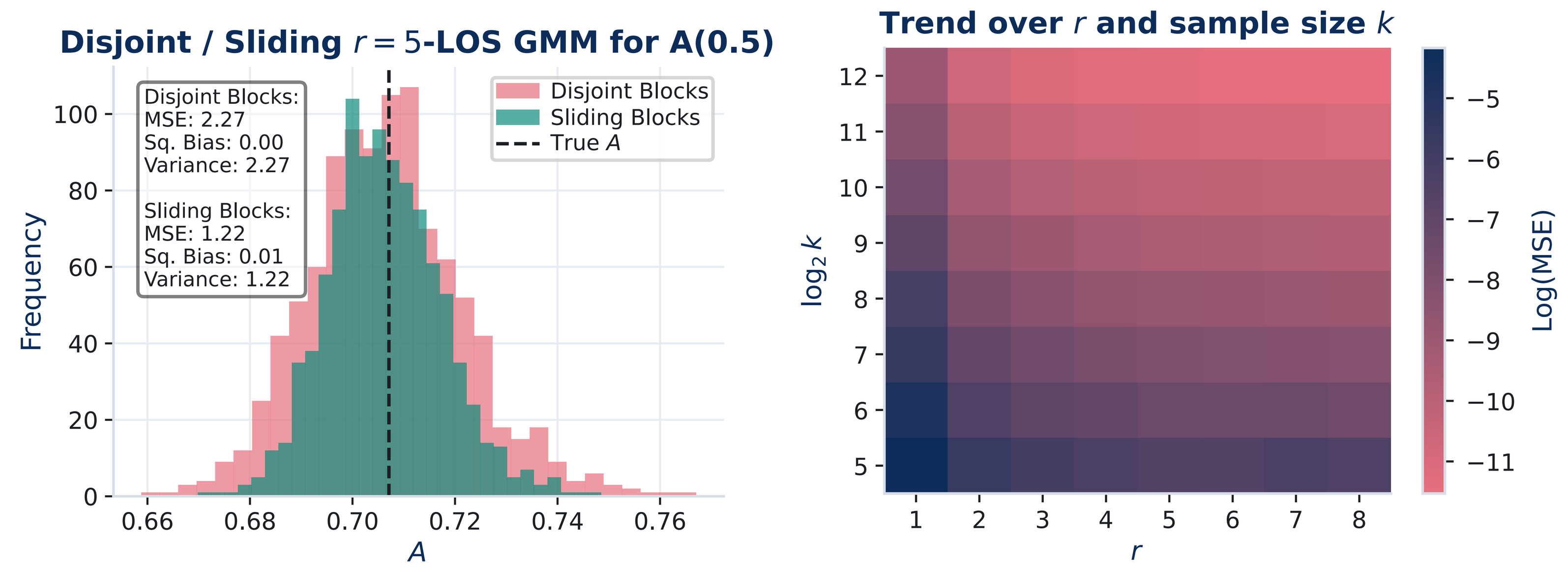


Figure 1: MSE for different Return Level Estimators: max-only vs. Top-Two vs. 'best compromise'

Multivariate r-Largest Order Statistics

Situation: The *multivariate* block maxima method usually studies *just the block maximum*
Question: Can we generalize previous results *to the multivariate?*
Contribution: We develop estimators of extremal dependence based on **r largest values**



Results:

- r-largest values carry information about extremal dependence
  ⇝ the larger r, the smaller the variance
- Sliding blocks improve disjoint blocks even for multivariate: variance decreases, while bias remains unchanged

References

[1] Bücher, A. & Haufs, E. (2026): Extreme Value Analysis based on Blockwise Top-Two Order Statistics. *Accepted for publication in Bernoulli.* https://arxiv.org/abs/2502.15036

[2] Bücher, A.; Haufs, E. & Schröter, J. (2026+): Synthesis for Extreme Event Attribution. *In preparation.*

[3] Bücher, A.; Haufs, E. *et al.* (2026+): Multivariate r-Largest Order Statistics. *Manuscript in preparation.*

Gefördert durch:
Bundesministerium für Forschung, Technologie und Raumfahrt

FONA Research for sustainability
CLIMXTREME Climate Change and Extreme Events