Research Training Day Part I

Recap in Extreme Value Theory & Recent Developments on the Block Maxima Approach

Jonas Schröter[1], Felix Fauer[2], **Erik Haufs**[3]

ClimXtreme II Module C Meeting, November 13, 2024

[1]Freie Universität Berlin, Institute of Meteorology
[2]Deutscher Wetterdienst, Regionales Klimabüro Potsdam
[3]Ruhr University Bochum, Faculty of Mathematics

Bundesministerium für Bildung und Forschung

FONA Research for sustainability

CLIMXTREME Climate Change and Extreme Events

Motivation: Climate Extremes

- Extreme weather events are central object of study in ClimXtreme I+II
- Goals:
  - How likely was an extreme event to occur?
  - Did occurrence probabilities change with climate?
  - ...

Motivation: Climate Extremes

- Extreme weather events are central object of study in ClimXtreme I+II
- Goals:
  - How likely was an extreme event to occur?
  - Did occurrence probabilities change with climate?
  - ...
- All questions demand statistical inference for extreme events!
- Extreme Value Theory provides methods for this

Setting:

- $X_1, ..., X_r$ are idependent & identically distributed (iid)
- $M_r := \max\{X_1, ..., X_r\}$

Setting:

- $X_1, \ldots, X_r$ are idependent & identically distributed (iid)
- $M_r := \max\{X_1, \ldots, X_r\}$

Theorem (Fisher and Tippett 1928, Gnedenko 1943).

If there exist sequences $(a_r)_{r\in\mathbb{N}} \subset (0,\infty)$ and $(b_r)_{r\in\mathbb{N}} \subset \mathbb{R}$ such that $\mathbb{P}\left(\frac{M_r - b_r}{a_r}\right) \xrightarrow{\mathcal{D}} G(x)$ non-degenerate, then G is one of

Setting:

- $X_1, \ldots, X_r$ are idependent & identically distributed (iid)
- $M_r := \max\{X_1, \ldots, X_r\}$

Theorem (Fisher and Tippett 1928, Gnedenko 1943).

If there exist sequences $(a_r)_{r\in\mathbb{N}} \subset (0,\infty)$ and $(b_r)_{r\in\mathbb{N}} \subset \mathbb{R}$ such that $\mathbb{P}\left(\frac{M_r - b_r}{a_r}\right) \xrightarrow{\mathcal{D}} G(x)$ non-degenerate, then G is one of

$$G(x) = \begin{cases} \exp\left(-e^{-x}\right) & x \in \mathbb{R} \quad \text{(Gumbel distribution)}, \\ \exp\left(-x^{-\alpha}\right) & \text{if } x > 0 \quad \text{(Fréchet distribution)}, \\ \exp\left(-(-x)^{\alpha}\right) & \text{if } x \leq 0 \quad \text{(Weibull distribution)}. \end{cases}$$

The GEV distribution

- Three cases (Gumbel, Fréchet, Weibull) summarizable by Generalized Extreme Value (GEV) distribution

$$G_{\mathrm{GEV}(\vartheta)}(x) = \exp\left(- \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi} \right)$$

- $\theta = (\mu, \sigma, \xi)$
  - μ: location parameter
  - σ: scale parameter
  - ξ: shape parameter

The GEV distribution

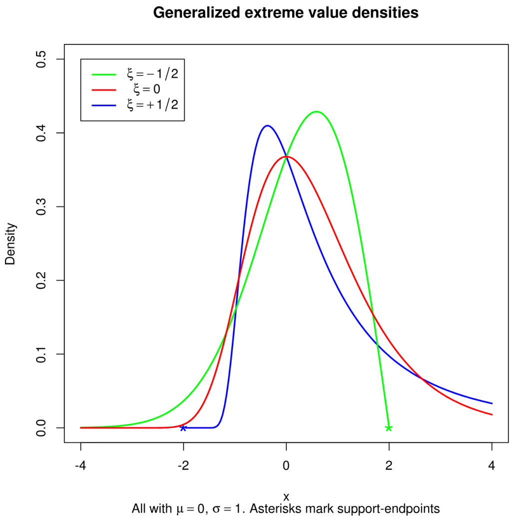


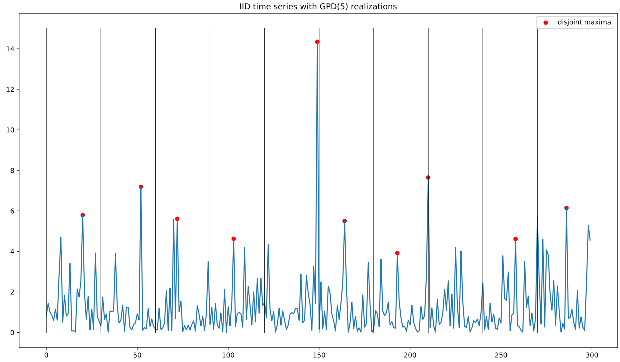
Figure 1: GEV densities

Block Maxima Modeling

- Given a time series $(X_n)_n$, e.g. daily precipitation, we take annual/monthly maxima M and approximate $\mathbb{P}(M \leq x) \approx G_{\mathrm{GEV}(\vartheta)}(x)$

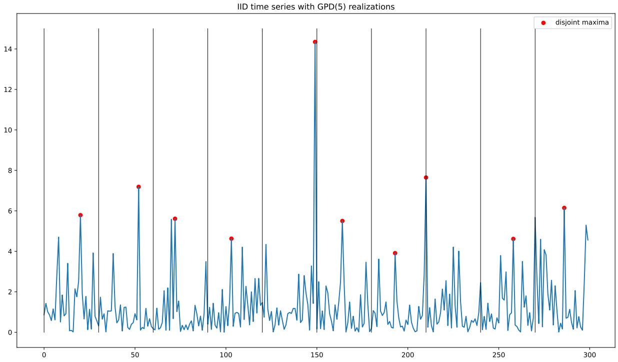
Block Maxima Modeling

- Given a time series $(X_n)_n$, e.g. daily precipitation, we take annual/monthly maxima M and approximate $\mathbb{P}(M \leq x) \approx G_{\mathrm{GEV}(\vartheta)}(x)$
  - Divide times series into blocks (e.g. months, years)
  - Take sample of block maxima to estimate the parameters $\theta = (\mu, \sigma, \xi)$
  - Use them to compute statistical targets (e.g. return levels/periods)



Block Maxima Modeling

- Given a time series $(X_n)_n$, e.g. daily precipitation, we take annual/monthly maxima M and approximate $\mathbb{P}(M \leq x) \approx G_{\mathrm{GEV}(\vartheta)}(x)$
  - Divide times series into blocks (e.g. months, years)
  - Take sample of block maxima to estimate the parameters $\theta = (\mu, \sigma, \xi)$
  - Use them to compute statistical targets (e.g. return levels/periods)
  - This is the **Block Maxima Method**



Drawbacks of Block Maxima $\rightsquigarrow$ POT

Drawbacks:

- Taking Block Maxima vastly reduces sample size
- Seldom more than 50 years of observations available $\rightsquigarrow$ high uncertainty in estimation
- Block Maxima misses other important extremes in the same block

Drawbacks of Block Maxima $\rightsquigarrow$ POT

Drawbacks:

- Taking Block Maxima vastly reduces sample size
- Seldom more than 50 years of observations available $\rightsquigarrow$ high uncertainty in estimation
- Block Maxima misses other important extremes in the same block

Peak Over Threshold Approach

- Focuses on exceedances over a high threshold u
- The excesses $X - u \mid X > u$ follow a Generalized Pareto Distribution (GPD):

$$G_u(y) = 1 - \left(1 + \xi \frac{y}{\sigma}\right)^{-1/\xi}, \quad y > 0$$

- Parameters: σ (scale), ξ (shape)
- Key advantage: More efficient use of data (all exceedances over threshold are included)

Drawbacks of POT $\rightsquigarrow$ SBM

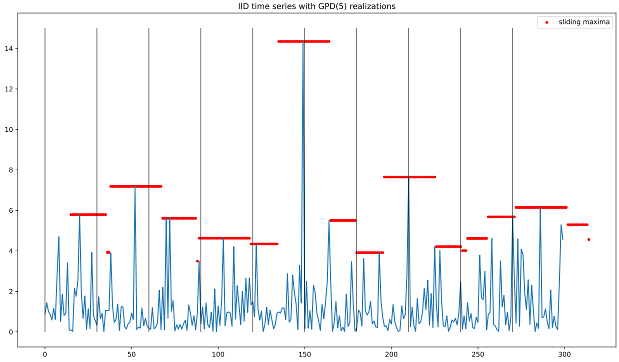
- POT has large disadvantage of how to choose the threshold

Drawbacks of POT $\rightsquigarrow$ SBM

- POT has large disadvantage of how to choose the threshold
- Way out: **Sliding Block Maxima (SBM)**
  - no need of threshold; captures more extremes than just Disjoint BM (DBM)

Drawbacks of POT $\rightsquigarrow$ SBM

- POT has large disadvantage of how to choose the threshold
- Way out: **Sliding Block Maxima (SBM)**
  - no need of threshold; captures more extremes than just Disjoint BM (DBM)



Given: Sample maxima $(M_1, \ldots, M_t)$ of t months [either disjoint or sliding blocks]

Probability Weighted Moments

- $\beta_\ell := \mathbb{E}[MF^\ell(M)]$

$$\hat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} M_{(i)} \qquad \hat{\beta}_1 = \frac{1}{n}\sum_{i=1}^{n} \frac{i-1}{n-1} M_{(i)} \qquad \hat{\beta}_2 = \frac{1}{n}\sum_{i=1}^{n} \frac{(i-1)(i-2)}{(n-1)(n-2)} M_{(i)}$$

- Possible to translate into GEV parameters (Hosking, Wallis, and Wood 1985)
- Robust alternative to MLE in EVT setting
- For Sliding Blocks: Bücher and Zanger 2023 (ClimXtreme I)
- **Takeaway: Sliding improves Disjoint for PWM estimators**

Maximum Likelihood Estimation

- Theory for ML estimation is difficult due to parameter-dependent support of GEV distribution
- $\rightsquigarrow$ focusing on two-parametric Frechét special case
- Recall Frechét distribution: $\mathbb{P}(M \leq x) = \exp\left(-\left(\frac{x}{\sigma}\right)^{-\alpha}\right)$
  - DBM estimator derived in Bücher and Segers 2018b
  - SBM estimator derived in Bücher and Segers 2018a

Maximum Likelihood Estimation

- Theory for ML estimation is difficult due to parameter-dependent support of GEV distribution
- $\rightsquigarrow$ focusing on two-parametric Frechét special case
- Recall Frechét distribution: $\mathbb{P}(M \leq x) = \exp\Big(-\Big(\frac{x}{\sigma}\Big)^{-\alpha}\Big)$
  - DBM estimator derived in Bücher and Segers 2018b
  - SBM estimator derived in Bücher and Segers 2018a

Asymptotic normality of MLE

Under some conditions, with number of blocks $k_n \to \infty$ and block size $r_n \to \infty$:

$$\sqrt{k_n}\begin{pmatrix}\hat{\alpha}_n^{\mathrm{dis}} - \alpha_0 \\ \hat{\sigma}_n^{\mathrm{dis}}/\sigma_n - 1\end{pmatrix} \rightsquigarrow \mathcal{N}_2(B, \Sigma^{\mathrm{disj}}) \qquad \text{and} \qquad \sqrt{k_n}\begin{pmatrix}\hat{\alpha}_n^{\mathrm{sl}} - \alpha_0 \\ \hat{\sigma}_n^{\mathrm{sl}}/\sigma_n - 1\end{pmatrix} \rightsquigarrow \mathcal{N}_2(B, \Sigma^{\mathrm{sl}})$$

and $\Sigma^{\mathrm{sl}} <_{\mathbb{L}} \Sigma^{\mathrm{disj}}$ in the Loewner-ordering.

- **Objective**: Estimate variability & confidence intervals of parameters
- **Bootstrap Process**:
  1. **Resampling**: Randomly resample block maxima (with replacement) from original dataset
  2. **Parameter Estimation**: For each resample, estimate GEV parameters (μ, σ, ξ)
  3. **Repetition**: Repeat process multiple times (e.g., 1000 iterations) to generate distribution of parameter estimates.
  4. **Uncertainty Quantification**: Analyze the distribution of estimates to compute standard errors, confidence intervals, and other measures of uncertainty.

Accessing Parameter Uncertainty: The Bootstrap

- **Objective**: Estimate variability & confidence intervals of parameters
- **Bootstrap Process**:
  1. **Resampling**: Randomly resample block maxima (with replacement) from original dataset
  2. **Parameter Estimation**: For each resample, estimate GEV parameters (μ, σ, ξ)
  3. **Repetition**: Repeat process multiple times (e.g., 1000 iterations) to generate distribution of parameter estimates.
  4. **Uncertainty Quantification**: Analyze the distribution of estimates to compute standard errors, confidence intervals, and other measures of uncertainty.
- **Advantages**:
  - Non-parametric: Does not assume a particular distribution for the data.
  - Flexible: Can handle small sample sizes or complex dependence structures.
  - Improves the robustness of the parameter estimates in climate extreme value modeling.

Bootstrapping Estimators based on SBM

Objective: Bootstrap Sliding Blocks

Problem: Naive block-bootstrap methods are inconsistent, even for i.i.d. cases.

Objective: Bootstrap Sliding Blocks

Problem: Naive block-bootstrap methods are inconsistent, even for i.i.d. cases. **Proposed solution**

(Bücher and Staud 2024):

- **Circular Block Maxima**: Introduces circular block maxima as a robust alternative, offering the same asymptotic variance as sliding block maxima.
- **Consistency**: Desired consistency results available there

Ongoing Research

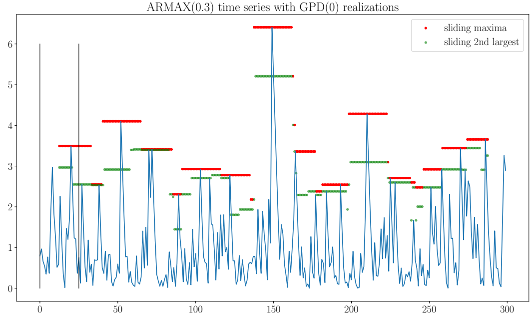
- Question: *How about utilizing other high order statistics of a block*?
- Currently working on a Top-Two Maximum Likelihood Estimator

Ongoing Research

- Question: *How about utilizing other high order statistics of a block*?
- Currently working on a Top-Two Maximum Likelihood Estimator
- problem: No joint likelihood for general time series $\rightsquigarrow$ Pseudo-MLE
- Due to pseudo estimation: consistency is lost in general $\rightsquigarrow$ handle with care!

Ongoing Research

- Question: *How about utilizing other high order statistics of a block*?
- Currently working on a Top-Two Maximum Likelihood Estimator
- problem: No joint likelihood for general time series $\rightsquigarrow$ Pseudo-MLE
- Due to pseudo estimation: consistency is lost in general $\rightsquigarrow$ handle with care!
- However: asymptotic variance is reduced significantly



Keep in mind:

- Real-world data often do not follow the i.i.d. assumption:
  - Weather and climate data are time series
  - Dependence structure, autocorrelation
- Theory is/may be extended to non-stationarity
  - e.g. use GMST as covariate, $\mu = \mu(\mathrm{GMST})$ or $\sigma = \sigma(\mathrm{GMST})$

Keep in mind:

- Real-world data often do not follow the i.i.d. assumption:
  - Weather and climate data are time series
  - Dependence structure, autocorrelation
- Theory is/may be extended to non-stationarity
  - e.g. use GMST as covariate, $\mu = \mu(\mathrm{GMST})$ or $\sigma = \sigma(\mathrm{GMST})$
- Theory holds just asymptotically!
- Real maxima distributions are just in the *domain of attraction (DOA)* of a GEV limit
  - $\rightsquigarrow$ don't expect $\max\{X_1, X_2, X_3\}$ close to any GEV!

Considerations in Real Data Applications

Keep in mind:

- Real-world data often do not follow the i.i.d. assumption:
  - Weather and climate data are time series
  - Dependence structure, autocorrelation
- Theory is/may be extended to non-stationarity
  - e.g. use GMST as covariate, $\mu = \mu(\mathrm{GMST})$ or $\sigma = \sigma(\mathrm{GMST})$
- Theory holds just asymptotically!
- Real maxima distributions are just in the *domain of attraction (DOA)* of a GEV limit
  - $\rightsquigarrow$ don't expect $\max\{X_1, X_2, X_3\}$ close to any GEV!
- Always validate models and assumptions with diagnostics

Key Takeaways

- EVT provides a robust framework for modeling extremes, but **assumptions must be checked carefully**
- **Block Maxima and POT** are two core approaches, each with advantages and drawbacks
- **Sliding Block Maxima** often better than Disjoint Block Maxima

References

- Bücher, Axel and Johan Segers (2018a). **"Inference for heavy tailed stationary time series based on sliding blocks".** In: *Electron. J. Stat.* 12.1, pp. 1098–1125. ISSN: 1935-7524. DOI: 10.1214/18-EJS1415.
- — (2018b). **"Maximum likelihood estimation for the Fréchet distribution based on block maxima extracted from a time series".** In: *Bernoulli* 24.2, pp. 1427–1462. ISSN: 1350-7265,1573-9759. DOI: 10.3150/16-BEJ903. URL: https://doi.org/10.3150/16-BEJ903.
- Bücher, Axel and Torben Staud (2024). **"Bootstrapping Estimators based on the Block Maxima Method".** In: *arXiv preprint arXiv:2409.05529*.
- Bücher, Axel and Leandra Zanger (2023). **"On the disjoint and sliding block maxima method for piecewise stationary time series".** In: *Ann. Statist.* 51.2, pp. 573–598. ISSN: 0090-5364,2168-8966. DOI: 10.1214/23-aos2260.
- Fisher, Ronald Aylmer and Leonard Henry Caleb Tippett (1928). **"Limiting forms of the frequency distribution of the largest or smallest member of a sample".** In: *Mathematical proceedings of the Cambridge philosophical society*. Vol. 24. 2. Cambridge University Press, pp. 180–190.
- Gnedenko, B. (1943). **"Sur la distribution limite du terme maximum d'une série aléatoire".** In: *Ann. of Math. (2)* 44, pp. 423–453. ISSN: 0003-486X. DOI: 10.2307/1968974. URL: https://doi.org/10.2307/1968974.
- Hosking, J. R. M., J. R. Wallis, and E. F. Wood (1985). **"Estimation of the generalized extreme-value distribution by the method of probability-weighted moments".** In: *Technometrics* 27.3, pp. 251–261. ISSN: 0040-1706,1537-2723. DOI: 10.2307/1269706. URL: https://doi.org/10.2307/1269706.