Critical review and hydrologic application of threshold detection methods for the generalized Pareto (GP) distribution

Antonios Mamalakis[1] , Andreas Langousis[1] and Roberto Deidda[2]

[1]Department of Civil Engineering University of Patras, Greece (antonis_mamalakis@hotmail.com).
[2]Dipartimento di Ingegneria Civile, Ambientale e Architettura, University of Cagliari, Cagliari, Italy.

1. Introduction

One of the most crucial issues in statistical hydrology is that of estimation of extreme rainfall. Based on asymptotic arguments from Extreme Excess (EE) theory, numerous studies have shown that the Generalized Pareto (GP) distribution can efficiently model rainfall excesses above sufficiently high thresholds. Yet, because no theoretical threshold value exists under pre-asymptotic conditions (i.e. from finite samples), several methods (or conservative assumptions, e.g. maintaining less than 5% of the empirical observations), have been applied to determine a proper threshold level u above which the GP assumption holds.

In this study, we review the most representative approaches from different types of threshold detection methods, and compare their efficiency when applied to 1714 over-centennial daily rainfall records from the NOAA-NCDC database. This is the first time that a **detailed intercomparison** of GP threshold detection methods is presented, followed by an **application** to rainfall records collected **worldwide**.

➡ **Optimum threshold u^*:** the lowest possible threshold that ensures validity of the GP assumption.

- *include the maximum number of data points* ⇨ reduce uncertainties in parameter estimation;
- *minimize biases due to the approximate validity of the GP assumption under pre-asymptotic conditions* (i.e. finite samples).

2. Review of approaches

2.1 Non Parametric methods: *Gerstengarbe and Werner plot*

Procedure: 1) Apply a sequential version of Mann-Kendall test to the differences of a sorted sample. 2) Delimit the extreme region of the data, by detecting the point where an ascending trend starts.

Limitations:

- In general (exception: uniform distribution), the null Hypothesis of randomness (no trend) is invalid, from the beginning of the differentiated series. ⇨ **Results with no statistical meaning**
- The adjacent differences of a sorted quantized sample is (mostly) a sequence of zeros. ⇨ **Inconclusive results** in hydrologic applications

2.2 Graphical Methods: *Mean Residual Life Plot (MRLP)*

Basis: Consider a random variable (RV) X that follows a GP distribution above threshold u^*. RV $Y = [X\text{-}u|\ X > u]\ \forall\ u > u^*$ is also GP distributed with the same shape parameter ξ, and scale parameter that increases linearly with u.

Procedure: Plot the mean value of the excesses above different threshold levels u, and locate the starting point of approximate linearity.

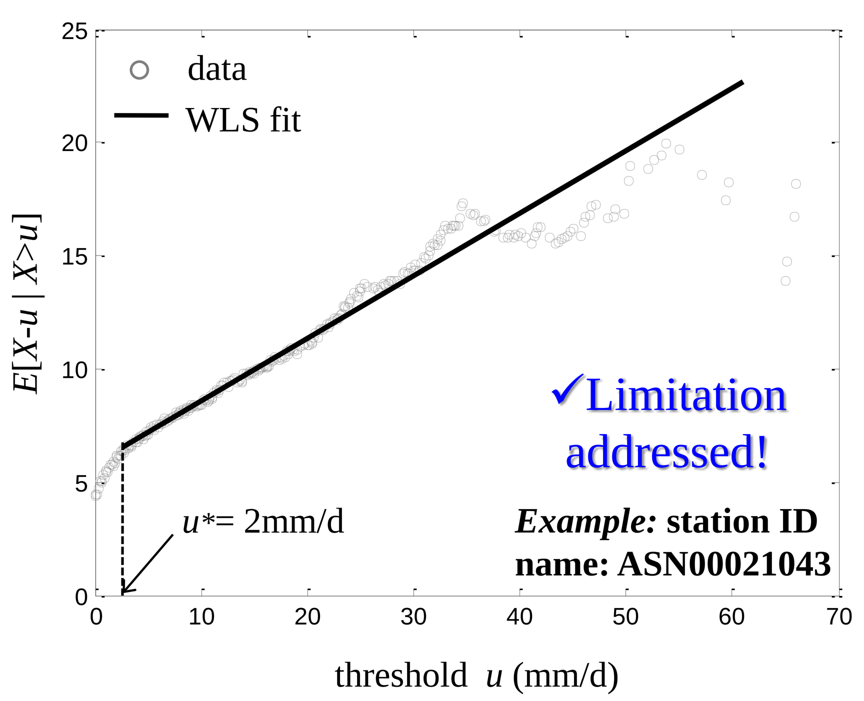


Figure 1: Application of MRLP to the positive rainrates extracted from a 126-year record of daily rainfall observations from Australia.

Limitation: No objective criterion to detect approximate linearity.
⇨ **Not suited to analyze large data sets**

2.3 Goodness-of-Fit-Based Methods: *Failure-to-reject method*

Basis: Cramér-von Mises (W^2) and Anderson-Darling (A^2) statistics quantify the deviation between the empirical distribution and the selected theoretical model.

Procedure: Start with the lowest possible threshold, and check the GP assumption (null Hypothesis) above increasing threshold levels u using the W^2 and/or A^2 statistics. Select the lowest threshold that the null hypothesis is not rejected.

Limitation: The asymptotic distributions of W^2 and A^2 are affected by the existence of **quantization** in the data, **leading to biased threshold estimates**:

▪ ratio $= \dfrac{\text{quantization level } \Delta}{\text{scale parameter } a_u}$ ▪ sample size

Table 1: Quantiles of W^2 calculated using 10,000 synthetic realizations of GP samples (threshold u = 0, and scale parameter a_0 = 10) with different sizes, shape parameter values and quantization levels.

■ 500 (black) / ■ 5000 (blue)		probability 0.90	probability 0.95	probability 0.99	type of data
shape parameter	0	0.121 = 0.121	0.150 = 0.151	0.220 = 0.221	*continuous (Δ=0)*
	0.1	0.115 = 0.116	0.143 = 0.143	0.209 = 0.210	
	0.2	0.112 = 0.111	0.137 = 0.137	0.200 = 0.199	
	0	0.128 < 0.165	0.158 < 0.199	0.229 < 0.288	*Δ=0.1 scale=10*
	0.1	0.118 < 0.156	0.143 < 0.187	0.211 < 0.262	
	0.2	0.116 < 0.151	0.141 < 0.181	0.201 < 0.264	
	0	0.218 < 1.000	0.256 < 1.074	0.349 < 1.231	*Δ=0.5 scale=10*
	0.1	0.212 < 0.947	0.248 < 1.016	0.343 < 1.169	
	0.2	0.199 < 0.895	0.236 < 0.968	0.323 < 1.109	

$\dfrac{\Delta}{a_u}$ ↓

✓ Similar findings for A^2

✓ Same results as MRLP

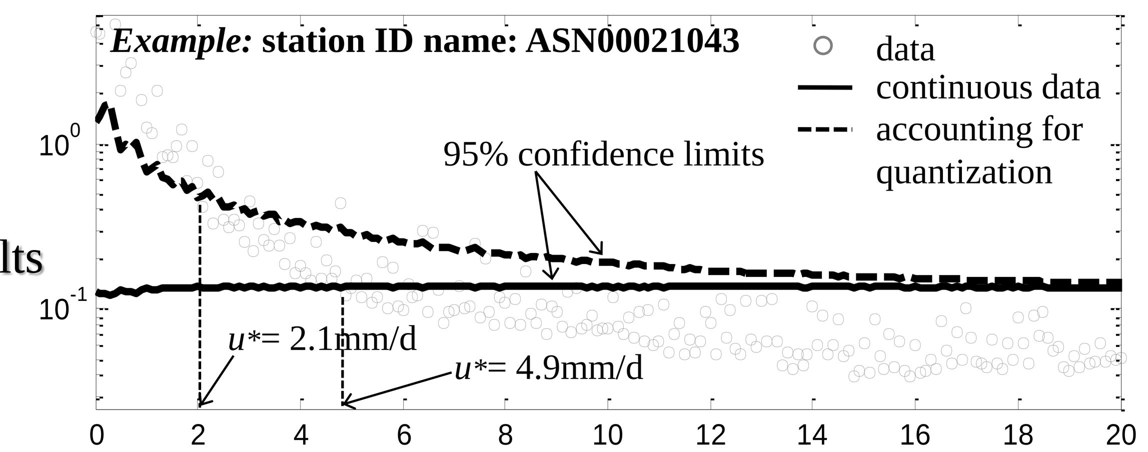


Figure 2: Application of the "failure-to-reject" method to the positive rainrates in Figure 1, using the W^2 Cramér-von Mises statistic.

2.4 Hill-based methods: *Jackson and Lewis modified kernel statistics*

Basis: Asymptotically, as $x \to \infty$, a log-transformed Pareto type variable is attracted to an exponential distribution with rate parameter equal to the Pareto tail index ⇨ *log-log linear CCDF plot (Pareto-quantile plot) with slope equal to the Pareto index.*

Procedure: Use the Jackson and Lewis modified kernel statistics to detect the lowest threshold level above which the Pareto-quantile plot displays approximate linearity. Use Hill's estimator to obtain the Pareto index (i.e. GP shape parameter).

Limitations:

- Overestimates the optimum threshold, as the outcome ensures asymptotic linearity in a Pareto-quantile plot: a much stricter condition than the GP assumption.
- Weak convergence to the asymptotic behavior for the case of GP samples with low shape parameter values (i.e. on the order of 0.1-0.2, as is the case of rainfall). ⇨ **Biases in shape parameter estimates**

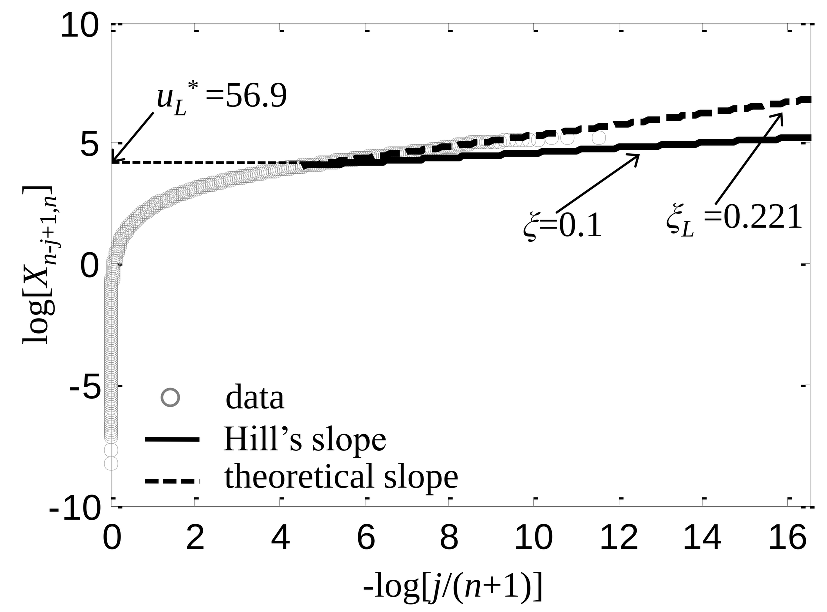


Figure 3: Pareto-quantile plot of a synthetic GP sample with u^*=0, ξ=0.1, a_u=10 and sample size 10^5.

➢ **threshold overestimation**

➢ **Significant biases of shape parameter estimates** even for considerably large sample sizes. Similar issues indicated in financial applications.

Convergence indicator for GP samples as a function of the probability level. ➡

$$l_p = 1 + \frac{\log[1-(1-p)^{\xi}]}{\log[(1-p)^{-\xi}]}$$

3. Application to the NOAA-NCDC Daily Rainfall Dataset

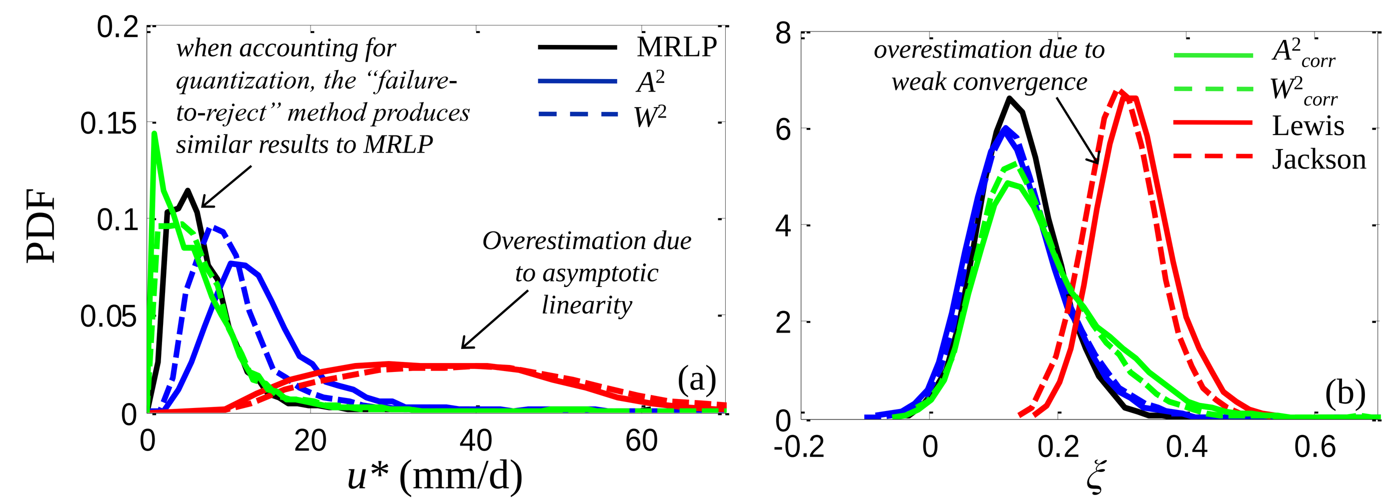
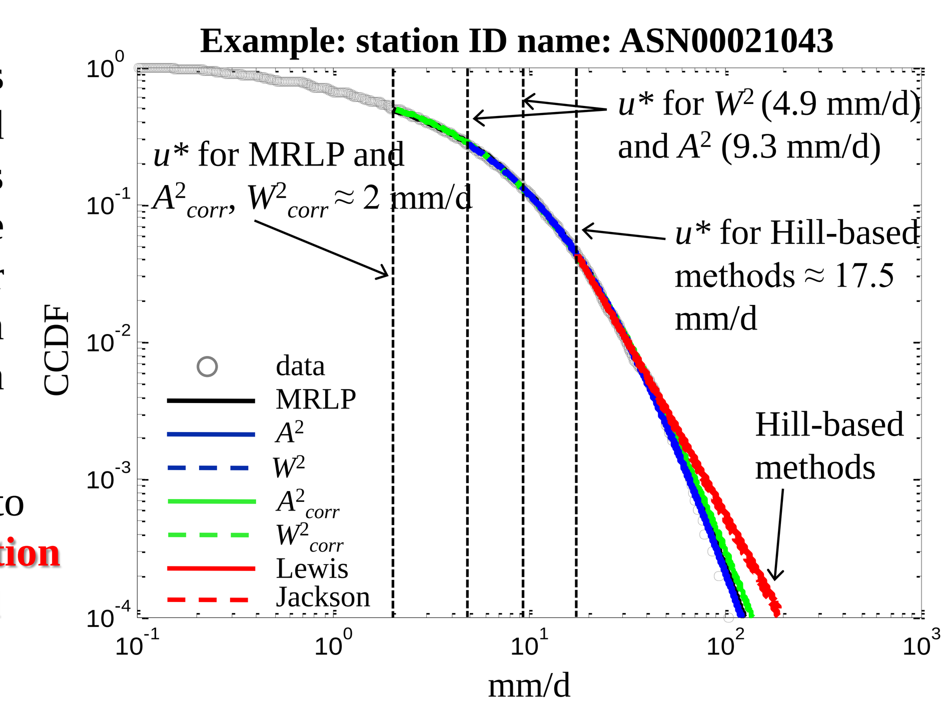


Figure 4: (a) Probability density function of thresholds u^*, estimated by applying the reviewed methods to 1714 rainfall records from NOAA-NCDC database, with more than 110 years of available observations; (b) Same as Figure 5a but for the shape parameter ξ.

Figure 5: Log-log plots of the empirical and theoretical CCDFs obtained for the positive rainrates of the 126-year daily rainfall record from Australia, used also in Figures 1 and 2.



➢ Hill-based methods lead to considerable **overestimation of the empirical rainfall quantiles.**

Table 2: Ensemble mean value and standard deviation of GP model parameters, estimated by applying the reviewed methods to 1714 rainfall records from NOAA-NCDC database, with more than 110 years of available observations (black). Values in blue have been obtained similarly, **but after randomly eliminating 70 years from each analyzed station.**

■ ≥ 100 (black) / ■ ≥ 40 (blue)			MRLP	A^2	W^2	MC A^2	MC W^2	LWS	JCK
threshold	mean	≥ 100	6.68	13.91	10.11	6.50	6.37	34.68	39.43
		≥ 40	6.63	7.53 ↓	5.53 ↓	3.67 ↓	3.69 ↓	27.96 ↓	31.33 ↓
	st. dev.	≥ 100	6.69	8.14	7.25	7.47	6.59	16.71	20.01
		≥ 40	6.70	4.22	3.74	3.29	3.32	13.49	15.00
shape	mean	≥ 100	0.140	0.138	0.144	0.177	0.167	0.320	0.301
		≥ 40	0.137	0.150	0.159	0.192	0.186	0.359 ↑	0.338 ↑
	st. dev.	≥ 100	0.060	0.088	0.096	0.096	0.092	0.054	0.053
		≥ 40	0.073	0.096	0.094	0.103	0.103	0.068	0.066

✓ MRLP is the most robust approach, producing similar results.

4. Conclusions

Intercomparison

- Gerstengarbe and Werner plot was proved **theoretically inconsistent**, while **not applicable** to **quantized samples**.
- **Hill's assumption** based methods lead to **unrealistically high threshold estimates** ⇨increased parameter estimation uncertainty.
- **Hill's** shape parameter estimates exhibit **considerable biases** due to the **slow convergence** of the log-transformed GP quantiles to those of an exponential distribution. **This is particularly the case for rainfall: ξ on the order of 0.1-0.2.**
- Although the "failure-to-reject" method is suited to the GP distribution, its **sensitivity** even to small levels of **data quantization** and to **sample length variations**, does not allow for routine applications.
- MRLP is the **most promising** method, as based on GP distribution properties valid also **under pre-asymptotic conditions**, while demonstrating reduced sensitivity to the length of the available data and to low levels of data quantization. ➘
  Room for improvements to include statistical arguments.

General comments

- The **existence of quantization** in rainfall records, along with **variations in their length**, constitute the two most important factors that may significantly affect the accuracy of the obtained results.
- For **daily rainfall applications**, GP threshold estimates range between **2-12 mm/d**, with a mean value around 6.5 mm/d.
- While several studies have used **the 95% (or higher) empirical quantiles** of the data, **much lower threshold values (i.e. in our case maintaining more than 10% of the empirical observations)** are also effective leading to reduced estimation variance of GP distribution parameters.

Reference

Langousis, A., A. Mamalakis, M. Puliga, and R. Deidda (2016) Threshold detection for the generalized Pareto distribution: Review of representative methods and application to the NOAA NCDC daily rainfall database, *Water Resour. Res.*, **52**, doi:10.1002/2015WR018502.

Acknowledgements

Andreas Langousis is sponsored by the Onassis Foundation under the "Special Grant and Support Program for Scholars' Association Members". The work conducted by Roberto Deidda was funded under the Sardinian Regional Law 7/2007 (funding call 2013).