

Comparison of Parameter Estimators for Generalized Pareto Distribution under Peak over Threshold

Wilhemina Adoma Pels*, Atinuke Olusola Adebajji, Sampson Twumasi-Ankrah

Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Kumasi, PMB, Ghana

Received September 19, 2020; Revised November 6, 2020; Accepted November 19, 2020

Cite This Paper in the following Citation Styles

(a): [1] Wilhemina Adoma Pels, Atinuke Olusola Adebajji, Sampson Twumasi-Ankrah, "Comparison of Parameter Estimators for Generalized Pareto Distribution under Peak over Threshold," *Mathematics and Statistics*, Vol. 8, No. 6, pp. 711 - 720, 2020. DOI: 10.13189/ms.2020.080612.

(b): Wilhemina Adoma Pels, Atinuke Olusola Adebajji, Sampson Twumasi-Ankrah (2020). Comparison of Parameter Estimators for Generalized Pareto Distribution under Peak over Threshold. *Mathematics and Statistics*, 8(6), 711 - 720. DOI: 10.13189/ms.2020.080612.

Copyright©2020 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The study focused on the Generalized Pareto Distribution (GPD) under the Peak Over Threshold approach (POT). Twenty-one estimation methods were considered for extreme value modeling and their performances were compared. Our goal is to identify the best method in various conditions by the use of a systematic simulation study. Some other estimators which were initially not created under the POT framework (NON-POT) were also compared concurrently with the ones under the POT framework. The simulation results under varying shape parameters showed the Zhang Estimator as "best" in performance for NON-POT in estimating both the shape and scale parameter for heavy-tailed cases. In the POT framework, the Zhang Estimator again performed "best" in estimating very heavy tails for the shape and very short tails for the scale regardless of the value of the scale parameter. When varying sample size, under the NON-POT framework, the Zhang estimator performed as "best" heavy-tailed while for the POT framework, the Pickands Estimator was "best" in performance at estimating the shape parameter for large sample sizes and the Zhang, small sample sizes.

Keywords Generalized Pareto Distribution, Peak over Threshold (POT), Shape Parameter, Scale Parameter

technique in statistical model building. A wrong choice in the estimation procedure would seriously affect the inferences drawn in the study, hence, making researchers more cautious in estimation. The three parameter Generalized Pareto Distribution (GPD) (location (μ), scale (σ), and shape (ξ)) is also prone to misspecification. However, the GPD is sometimes specified by only σ and ξ instead of the three parameters [1]. In the last four decades, a wide range of methods have been proposed to estimate these two parameters [2,3]. The most commonly used estimators are the Maximum Likelihood Method (MLE), the Probability Weighted Method (PWM), and the Method of Moments (MOM) [4]. However, these estimators all have their pitfalls.

De Zea Bermudez[4] posits that, the PWM method performs well for shape parameter between $0 \leq \xi \leq 1$ but performs very good when $\xi \leq 0.5$. Though, this method performs well in 0, 1 interval, the asymptotic properties cannot be derived when $\xi > 0.5$. Castillo[5] also showed that PWM performs well in small samples for $\xi \leq 0.5$ and research by Hosking[6] indicated that the method of moments (MOM) performs poorly when the shape parameter is greater than 1. This is due to the mean and variance not existing for values greater than 1.

In a review of GPD estimators by Deidda[7], the MLE which is based on the maximization of the likelihood function improves asymptotically and best when sample size is large. Macklay[8] and Giles[9] in their study further showed that the MLE in comparison to MOM and PWMs leads to biased estimates when sample size is small. In addition to that, Dupuis[10] and Ashkar[11] discussed that

1. Introduction

Parameter estimation is a significant and fundamental

these estimators in some cases produce estimates, which are inconsistent with observed data. That is, there are samples for which ξ is negative and $Z_n > \frac{-\hat{\sigma}}{\hat{\xi}}$ (where $\hat{\sigma}$ is the estimated scale parameter and $\hat{\xi}$ is the estimated shape parameter), violating the definition of Generalized Pareto Distribution. They showed that when $\xi = -0.5$ about 30% of samples results in non-feasible parameter estimates, but the occurrence decreases with increasing ξ . To fix these problems Zhang[12] and del Castillo[13] suggested a likelihood-based method of estimation which is computationally easy. Zhang[14] and Zhang[15] based their estimators on likelihood moment and Bayesian methodology. Additionally, a nonlinear least square (NLS) approach was suggested by Song[16] to estimate both ξ and σ of which the authors claimed that the NLS performs better than other estimators. However, their comparison was made unfairly because all the other approaches used in their work were based on the POT system, whereas the NLS was not though they claim of using the POT framework [17]. Park[17] then modified this approach by providing weights for the NLS.

Hence, there have been various new estimators proposed for estimating the parameters of the GPD. An extensive discussion of the various proposed methods have been presented by de Zea Bermudez[4]; Macklay[8]. Although de Zea Bermudez[4] gives some guidance on the relative merits of the a number of methods yet no quantitative comparison of the performances by means of a simulation study was made. Though Ashkar[11]; Deidda[7]; Macklay[8]; Kang[1] have compared the performance of estimators through simulation studies, however, few estimators were used for comparison, limited range of sample sizes, or one-sided GPD shape parameters. Most often, the performance of the GPD parameter estimators depends on both the sample size and the value of the GPD shape parameter [18,8,1]. Therefore, estimators that perform well in some situations may perform poorly in others. This study seeks to make a systematic comparison of twenty-one conventional estimators for estimating the parameters of the GPD over a wide range of sample sizes with varying shape parameters and scale parameters and make practical recommendations for the most suitable methods for estimation through simulation studies.

2. Methods and Materials

This section presents the theoretical review of statistical tools used in the study.

2.1. Generalized Pareto Distribution (GPD)

GPD is one of the fundamental distributions in extreme value modeling with distribution function defined as follows

Let Z be a random variable such that

$$F_{\mu,\sigma,\xi}(z) = \begin{cases} 1 - \left(1 + \xi \frac{z-\mu}{\sigma}\right)^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{z-\mu}{\sigma}\right), & \text{if } \xi = 0 \end{cases} \quad (1)$$

where the location, scale, and shape parameters as μ ($\mu \in \mathbb{R}$), σ ($\sigma > 0$), and ξ ($\xi \in \mathbb{R}$) respectively. Based on the value of the shape parameter, the distribution can be divided into three groups, thus $F_{\mu,\sigma,\xi}(z)$ is short-tailed when $\xi < 0$, medium tailed when $\xi = 0$, and lastly heavy-tailed when $\xi > 0$.

2.2. Peak Over Threshold (POT)

Let $Z_1 > Z_2 > \dots > Z_N$ be an ordered sequence of independent and identically distributed (iid) random variables with common continuous distribution function F and n be the number of observations over a threshold u . The excess distribution above a certain threshold u for a random variable Z is defined as

$$F_u(z) = P(Z - u \leq z | Z > u) \quad (2)$$

for $0 \leq z < z_F - u$ where z_F is the right endpoint of F . The variable z represents the exceedances over u . The excess distribution function can also be written as

$$F_u(z) = \frac{F(z+u) - F(u)}{1 - F(u)} \quad (3)$$

2.3. Error Accuracy Measure

2.3.1. The Root Mean Square Error (RMSE)

This error metric measure is the square root of the mean squared error which measures the differences between the estimate and the actual or true value

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (a - \hat{a})^2} \quad (4)$$

Where a is the actual value and \hat{a} is the estimated value of the parameters (ξ and σ)

2.3.2. The Mean Square Error (MSE)

The mean square error is a measure of the mean square deviation.

$$\frac{1}{n} \sum_{i=1}^n (a - \hat{a})^2 \quad (5)$$

2.3.3. The Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is one of the simplest error metric modelings. It takes the average of the absolute errors.

$$\frac{1}{n} \sum_{i=1}^n |a - \hat{a}| \quad (6)$$

2.3.4. The Median Absolute Error (MDAE)

This computes the median absolute difference between the estimate and true value

$$Med|a - \hat{a}| \quad (7)$$

Table 1. Parameter framework

Author	μ	σ	ξ
Kang[1]	0, 50, 100	1, 10	-3, -2, -1, -0.5, 0, 1, 2, 5.
Park[17]	0	1, 10	0.5, 1, 2, 4
Song[16]	0	1, 10, 100.	0, 0.5, 1, 2, 3, 4, 5
Deidda[7]	0	7,12	$\xi \in (-0.5,0.5)$
Zhang[15]	0	1	-6, -3, -1, -0.25, 0.5
Castillo[13]	0	1	-0.2,-0.1, 0.1, 0.2, 0.8, 0.9, 1.1, 1.2
Gharib[18]	Threshold value	Ranging from 20 to 1000	-0.5,-0.4,-0.3,-0.2,-0.1, 0, 0.1, 0.2, 0.3,0.4,0.5

2.3.5. Parameter Framework

Table 1 presents parameter framework used by some researchers. Generally, it can be seen that most of these authors used a location parameter of “0”, a scale parameter of “1” and “10”

2.4. Simulation Procedure

This section gives the algorithm for comparison on the performance of twenty-one estimators in parameter estimation with (POT Framework) and without the POT (NON-POT Framework) method. This is due to the fact that many authors have studied the parameter estimation for GPD not following the POT method [1]. The value of the shape parameter to be used are $\xi = -3, -2, -1, -0.5, 0, 1, 2, 3, 5$ and that of the scale parameter are 1, 10. Sample sizes to be considered ranges from small to large, that is $n = 10, 20, 30, 50, 100, 200, 500, 1000$, thus for NON-POT framework.

For POT framework with a quantile; $q = 0.9$, the sample sizes to be considered are 100, 200, 300, 500, 1000, 2000, 5000 and 10000 while with quantile; $q = 0.95$, the sample sizes to be considered are 200, 400, 600, 1000, 2000, 4000, 10000 and 20000. This is chosen in order to have the same number of observations as the NON-POT framework.

2.4.1. NON-POT Framework

The following algorithm is to be used for the simulation procedure not following the POT framework thus using all observations.

- I. Generate an n number of observations following GPD with the known shape and scale parameter.
- II. Using all the observations generated, estimate each of the shape and scale parameters with 21 different estimation techniques.
- III. Repeat step I and II 10,000 times
- IV. Compute and record the performance metric of each estimator of which the minimum value means better performance.
- V. Steps I-IV are to be repeated for all sample sizes.

- VI. To find the best estimator for varying the shape parameter at a particular value of scale parameter, compute the average value of both ξ and σ for each performance metric across all sample sizes using:

$$\begin{cases} B_{\xi} = \frac{\sum_{i=1}^m \delta_i}{m} \\ B_{\sigma} = \frac{\sum_{i=1}^m \omega_i}{m} \end{cases} \quad (8)$$

where δ is the value of $\hat{\xi}$ for each method, ω is the value of $\hat{\sigma}$ for each method and m is the number of sample size.

- Value VII. To find the best estimator for varying sample sizes, compute the average for each performance metric across all values of the shape parameter at a particular value of scale parameter.

$$\begin{cases} S_{\xi} = \frac{\sum_{i=1}^k \delta_i}{k} \\ S_{\sigma} = \frac{\sum_{i=1}^k \omega_i}{k} \end{cases} \quad (9)$$

where k is the total number of the shape parameter

2.4.2. POT Framework

The following algorithm is also to be used for the simulation procedure under the POT framework thus using observations above a certain threshold.

- I. Generate an N number of observations following GPD.
- II. Take the 100qth quantile as a threshold value. Thus using $q = 0.9$ or 0.95.
- III. Using all the observations above the threshold, estimate the shape and scale parameters with 21 different estimation techniques.
- IV. Repeat steps I, II and III 10,000 times
- V. Compute their performance metric for each of the estimators of which the minimum value means better performance.
- VI. Steps I-V are to be repeated for all the sample sizes.
- Value VII. To find the best estimator for varying the shape parameter at a particular value of scale parameter, compute the average value of both ξ and σ for each performance metric across all sample sizes using “(8)”

VIII. To find the best estimator for varying sample sizes, compute the average for each performance metric across all values of the shape parameter at a particular value of scale parameter with the use of “(9)”.

3. Results and Discussions

The main focus of this study is to compare the performance of 21 estimation techniques under different conditions. These conditions considered are 1. Use of 9 different values of the shape parameter with 2 values of the scale parameter 2. Use of 8 different sample sizes again with 2 values of the scale parameter. Table 2 shows the 21 parameters used in the study. From the table, item 10 to item 16 (AD, ADR, AD2R, ADL, AD2L, CM, and KS) are the Maximum Goodness of Fit (MGF) estimators used by Luceño [24]

R codes for MLEn; Zhang; NLS-S; NLS-P and WNLS-P are obtained from Kang [1] while the rest can be found in the R package POT created and maintained by Ribatet [25].

3.1. The “Best” Performance of Conventional Methods of the GPD with Varying Shape Parameter

3.1.1. Simulated Results for NON-POT Framework

Tables 3 shows the “best” and “worst” performing

methods for the NON-POT framework of 21 estimators with respect to varying shape parameters when $\sigma = 1$ and 10.

With regards to short-tailed cases (i.e. $\xi < 0$), the CM, ADR, MLE, and the Zhang estimators performed very well in the estimating concurrently both the shape and scale parameters at $\xi = -3, -2, -1$ and -0.5 for $\sigma = 1$ and 10 respectively. However, the MLEn estimator performed very poorly in the estimating concurrently both the shape and scale parameters for $\sigma = 1$ and 10 respectively.

More so, for a medium tailed case, that is with $\xi = 0$, at $\sigma = 1$ the LME performed well in estimating the shape parameter whiles Zhang’s performance was good in estimating the scale. Whereas $\sigma = 10$, the Zhang estimator was “best” for shape and the KS was “best” for scale. Yet, the MLEn estimator was still “worst” in performance for simultaneously estimating both the shape and scale parameters for $\sigma = 1$ and 10 respectively.

Also, for heavy-tailed cases (i.e. $\xi > 0$), the Zhang estimator outperformed all other estimators for both the shape and scale parameter except at $\xi = 5$ with $\sigma = 10$ where the MLEn performed “best”. Nonetheless, at both when $\sigma = 1$ and 10, the AD2L estimator performed “worse” than all other estimators for both the shape and scale parameters except at $\xi = 5$ where the MDPD and all the MGF estimators performance were “worst”.

Table 2. Estimating Methods used in the study

1. Maximum Likelihood Estimator (MLE)	12. Right tail Anderson Darling (2nd deg) (AD2R)
2. Biased Probability Weighted Moments (PWMB)	13. Left tail Anderson Darling (ADL)
3. Unbiased Probability Weighted Moments (PWMU)	14. Left tail Anderson Darling (2nd deg) (AD2L)
4. Mean power density divergence (MDPD)	15. Cramer von Mises (CM)
5. Likelihood moment estimator (LME)	16. Kolmogorov-Smirnov (KS)
6. Method of Moments (MOM)	17. Zhang 2010 (ZHANG)
7. Median (MED)	18. Nonlinear least square (NLS-S)
8. Pickands (PICKANDS)	19. Nonlinear least square by Park (NLS-P)
9. Maximum penalized likelihood (MPLE)	20. Weighted nonlinear least square (WNLS-P)
10. Anderson Darling (AD)	21. Modified Maximum Likelihood (MLEn)
11. Right tail Anderson Darling (ADR)	

Table 3. Summary results for NON-POT framework when varying shape parameter

True Shape ξ	"Best" performing method(s)				"Worst" performing method(s)			
	$\sigma = 1$		$\sigma = 10$		$\sigma = 1$		$\sigma = 10$	
	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$
-3	CM	CM	CM	CM	MLEn	MLEn	MLEn	MLEn
-2	ADR	ADR	ADR	ADR	MLEn	MLEn	MLEn	MLEn
-1	MLE	MLE	MLE	MLE	MLEn	MLEn	MLEn	MLEn
-0.5	ZHANG	ZHANG	ZHANG	ZHANG	MLEn	MLEn	MLEn	MLEn
0	LME	ZHANG	ZHANG	KS	MLEn	MLEn	MLEn	MLEn
1	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	AD2L	AD2L	KS
2	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	AD2L	AD2L	AD2L
3	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	AD2L	AD2L	AD2L
5	ZHANG	ZHANG	ZHANG	MLEn	AD2L	MDPD	AD2L	MDPD, ALL MGF

Table 4. Summary results for POT framework with $q = 0.9$ when varying shape parameter

True Shape ξ	"Best" performing method(s)				"Worst" performing method(s)			
	$\sigma = 1$		$\sigma = 10$		$\sigma = 1$		$\sigma = 10$	
	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$
-3	PICKANDS	ZHANG	PICKANDS	ZHANG	MLEn	MLEn	MLEn	MLEn
-2	ADL	ZHANG	CM	ZHANG	MLEn	MLEn	MLEn	MLEn
-1	AD2R	MOMENTS	MLE	AD2R	MLEn	MLEn	MLEn	ZHANG
-0.5	AD2R	MPLE	AD2R	MPLE	MLEn	MLEn	MLEn	ZHANG
0	LME	LME	KS	KS	MLEn	ZHANG	MLEn	ZHANG
1	LME	MLEn	PWMU	LME	AD2L	KS	MLEn	KS
2	MED	MLEn	PICKANDS	LME	AD2L	AD2L	AD2L	KS
3	ZHANG	MLEn	ZHANG	LME	AD2L	AD2R	AD2L	AD2R
5	ZHANG	MLEn	ZHANG	MLEn	AD2L	AD2R	MLEn	AD2R

3.1.2. Simulated Results for the POT Framework

Tables 4 shows results for the POT framework of which the 90th quantile of the observations is used in the estimation with $\sigma = 1$ and $\sigma = 10$ respectively. This table gives the "best" and "worst" performing estimators for all the shape and the scale parameters when the values of the shape parameter vary.

For $\xi < 0$ at $\sigma = 1$, the Pickands and the ADL was "best" in performance at when $\xi = -3$ and -2 respectively whiles AD2R was "best" for $\xi = -1$ and -0.5 . However, at $\sigma = 10$ for estimating the shape parameter Pickands, CM, MLE, AD2R were "best" in performance with increasing value of ξ . Whereas for estimating the scale parameter at $\xi = -3$ and -2 the Zhang method outperformed all other methods while the AD2R and the MPLE were "best" at $\xi = -1$ and -0.5 respectively. With regards to "worst" performance, the MLEn estimator performance was poor in estimating both the shape and scale parameters for $\sigma = 1$ and $\sigma = 10$. However, when $\sigma = 10$, in estimating the scale, the Zhang performed "worse" at $\xi = -1$ and -0.5 .

Furthermore, with $\xi = 0$, in estimating both the shape and scale parameter, the LME was "best" for $\sigma = 1$

while for $\sigma = 10$, the KS estimator was "best" as compared to all other estimating methods. Pertaining to "worst" performance, the MLEn and the Zhang estimators were still "worst" in performance for estimating both the shape and scale respectively at $\sigma = 1$ and 10 .

For $\xi > 0$, with $\sigma = 1$ in estimating for the shape parameter, the LME and Median method were "best" in performance when $\xi = 1$ and 2 respectively whilst the Zhang estimator was "best" in estimating values of ξ at 3 and 5 . The MLEn method outperforms all other estimators in estimating the scale parameter. With $\sigma = 10$ in estimating for the shape parameter, the PWMU and Pickands were "best" in performance when $\xi = 1$ and 2 respectively, whilst the Zhang estimator was best for ξ values of 3 and 5 . In estimating for the scale, the LME estimator was relatively excellent when $\xi = 1, 2, 3$ whereas the MLEn method was good in estimation when $\xi = 5$. In terms of "worst" performance, the AD2L estimator performed "worse" than all other estimators for the shape whilst for the scale KS, AD2L and AD2R performed poorly. At $\sigma = 10$, the MLEn and AD2L were the "worst" in estimating the shape. For scale, both the KS and AD2R were "worst" in performance.

In estimating using the POT framework of 95th quantile as the threshold value, Table 5 gives the summary of the “best” and “worst” performing estimators in that regard.

For short-tailed cases, $\sigma = 1$ the Pickands, PWMB, ADR and MDPD were “best” in performance with minimal error metric. Pickands, CM, MPLE, and ADR performed well in estimating each of the shapes whereas the Zhang, Moment and MPLE were “best” in estimating the scale at $\sigma = 10$.

For the medium-tailed case, the LME and the KS outperformed all other estimators for both the shape and the scale at $\sigma = 1$ and $\sigma = 10$ respectively.

For heavy-tailed, when estimating the scale, the MLEn was good at $\sigma = 1$ whilst for $\sigma = 10$ the LME was good in performance. With regards to the shape, the Zhang performed better compared to all other estimators used in the study regardless of the value of the scale parameter.

For both short-tailed and medium tailed cases in terms of “worst” performing estimators, performance was the same when threshold value was $q = 0.9$ at both $\sigma = 1$ and 10 respectively.

However, when $\xi > 0$ at $\sigma = 1$ and 10, the AD2L

and the MLEn estimators performed “worst” than all other estimators for the shape whilst for the scale KS, and AD2R performed poorly.

3.2. The “Best” Performance of Conventional Methods of the GPD with Varying Sample Size

3.2.1. Simulated Results for NON-POT Framework

In estimating shape at both scale values of 1 and 10, the Zhang estimator outperforms all the other estimators from small sample sizes to larger sample sizes (i.e. $20 \leq n \leq 200$) but when sample size was 500 and 1000, the Pickands estimator performed better. When checking “worst” performing estimating methods, the AD2L estimator performed “worse” than all the other estimators for all sample sizes.

With regards to the estimation of the scale, the Zhang estimator outperforms all the other estimators at all sample sizes varying from small to large (i.e. $20 \leq n \leq 1000$) whilst the MDPD and all the MGF estimators all gave the same estimating values for the error metric measure indicating “worst” performance. These results are presented in Table 6.

Table 5. Summary results for POT framework with $q = 0.95$ when varying shape parameter

True Shape ξ	"Best" performing method(s)				"Worst" performing method(s)			
	$\sigma = 1$		$\sigma = 10$		$\sigma = 1$		$\sigma = 10$	
	ξ	$\hat{\sigma}$	ξ	$\hat{\sigma}$	ξ	$\hat{\sigma}$	ξ	$\hat{\sigma}$
-3	PICKANDS	ZHANG	PICKANDS	ZHANG	MLEn	MLEn	MLEn	MLEn
-2	PWMB	ZHANG	CM	ZHANG	MLEn	MLEn	MLEn	MLEn
-1	ADR	MOMENTS	MPLE	MOMENTS	MLEn	MLEn	MLEn	ZHANG
-0.5	MDPD	MPLE	ADR	MPLE	MLEn	MLEn	MLEn	ZHANG
0	LME	LME	KS	KS	MLEn	ZHANG	MLEn	ZHANG
1	PWMU	MLEn	PWMU	LME	AD2L	KS	MLEn	KS
2	ZHANG	MLEn	ZHANG	LME	AD2L	KS	AD2L	KS
3	ZHANG	MLEn	ZHANG	LME	AD2L	AD2R	AD2L	AD2R
5	ZHANG	NLS-P&NLS-S	ZHANG	MLEn	MLEn	AD2R	MLEn	AD2R

Table 6. Summary results for NON-POT framework when varying sample size

Sample size	"Best" performing method(s)				"Worst" performing method(s)			
	$\sigma = 1$		$\sigma = 10$		$\sigma = 1$		$\sigma = 10$	
	ξ	$\hat{\sigma}$	ξ	$\hat{\sigma}$	ξ	$\hat{\sigma}$	ξ	$\hat{\sigma}$
10	ZHANG	NLS-S	ZHANG	NLS-S	AD2L	MDPD	AD2L	MDPD & MGF
20	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	MDPD & MGF	AD2L	MDPD & MGF
30	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	MDPD	AD2L	MDPD & MGF
50	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	MDPD & MGF	AD2L	MDPD & MGF
100	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	MDPD & MGF	AD2L	MDPD & MGF
200	ZHANG	ZHANG	ZHANG	ZHANG	AD2L	MDPD & MGF	AD2L	MDPD & MGF
500	PICKANDS	ZHANG	PICKANDS	ZHANG	AD2L	MDPD & MGF	AD2L	MDPD & MGF
1000	PICKANDS	ZHANG	PICKANDS	ZHANG	AD2L	MDPD & MGF	AD2L	MDPD & MGF

3.2.2. Simulated Results for the POT Framework

Table 7 shows results for the POT framework of which the 90th quantile of the observations is used in the estimation with $\sigma = 1$ and $\sigma = 10$ respectively. This table gives both the “best” and “worst” performing estimators for all the shape and the scale parameters when varying sample size.

It is obvious from the table that regardless of the value of σ , the Zhang method was “best” in performance for estimating the shape when sample size was small (i.e. $n = 10$ to 30). Nonetheless, when sample size(n) was from 50 to 1000, the Pickands performed better than other methods. Whereas, the AD2L estimator performed “worse” than all the other estimators for all sample sizes (n), except for $n = 10$ at $\sigma = 10$ when the MLEn performed “worse”.

In estimating the scale parameter, at $\sigma = 1$, the “best” performing method for varying sample sizes(n) was the MLEn at $20 \leq n \leq 1000$ while the MDPD and the AD2R performed poorly. For $\sigma = 10$, the LME method performed “best” for nearly all sample sizes. Whilst all the MGF and MDPD estimators all gave the same estimating values for

the error metric measure indicating “worst” performance for sample sizes(n) 10 to 200. At sample size(n) 500 and 1000 the AD2R gave the “worst” performance.

At a threshold value of the 95th quantile, with varying sample sizes, Table 8 gives the “best” and “worst” performing methods. It’s obvious from the table that the MLEn outperforms all the other estimators in estimating the scale parameter (i.e. $\sigma = 1$) with small (20, 30) and even large sample sizes (50 to 1000). For $\sigma = 10$, the LME outperformed all other methods for nearly all sample sizes(n). However, the MDPD and the AD2R performed poorly regardless of the value of σ . In estimating for the shape parameter at both $\sigma = 1$ and $\sigma = 10$, the Zhang estimator achieved good performance with small sample sizes(n) (i.e. 10 to 30) and when sample sizes are large (i.e. 50 to 1000), the Pickands estimator showed a very good performance. Nonetheless, at $\sigma = 1$, the AD2L estimator performed “worse” than all the other estimators for all sample sizes, except for $n = 10$ when the MLEn performed “worse”. More so, at $\sigma = 10$ the MLEn “worst” in performance for all sample sizes.

Table 7. Summary results for POT framework with $q = 0.9$ when varying sample size

Sample size N(n)	"Best" performing method(s)				"Worst" performing method(s)			
	$\sigma = 1$		$\sigma = 10$		$\sigma = 1$		$\sigma = 10$	
	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\sigma}$
100(10)	ZHANG	NLS-S	ZHANG	NLS-S	AD2L	MDPD	MLEn	MDPD & MGF
200(20)	ZHANG	MLEn	ZHANG	LME	AD2L	MDPD	AD2L	MDPD & MGF
300(30)	ZHANG	MLEn	ZHANG	LME	AD2L	MDPD	AD2L	MDPD & MGF
500(50)	PICKANDS	MLEn	PICKANDS	LME	AD2L	MDPD	AD2L	MDPD & MGF
1000(100)	PICKANDS	MLEn	PICKANDS	LME	AD2L	MDPD	AD2L	MDPD & MGF
2000(200)	PICKANDS	MLEn	PICKANDS	LME	AD2L	MDPD	AD2L	MDPD & MGF
5000(500)	PICKANDS	MLEn	PICKANDS	LME	AD2L	AD2R	AD2L	AD2R
10000(1000)	PICKANDS	MLEn	PICKANDS	LME	AD2L	AD2R	AD2L	AD2R

N = the sample size, (n) = the number of observations above the threshold

Table 8. Summary results for POT framework with $q = 0.95$ when varying sample size

Sample size N(n)	$\sigma = 1$		$\sigma = 10$		$\sigma = 1$		$\sigma = 10$	
	"Best" Method for shape	"Best" Method for scale	"Best" Method for shape	"Best" Method for scale	"Worst" Method for shape	"Worst" Method for scale	"Worst" Method for shape	"Worst" Method for scale
200(10)	ZHANG	NLS-S	ZHANG	NLS-S	MLEn	MDPD	MLEn	MDPD & MGF
400(20)	ZHANG	MLEn	ZHANG	LME	AD2L	MDPD	MLEn	MDPD & MGF
600(30)	ZHANG	MLEn	ZHANG	LME	AD2L	MDPD	MLEn	MDPD
1000(50)	PICKANDS	MLEn	PICKANDS	LME	AD2L	MDPD	MLEn	MDPD
2000(100)	PICKANDS	MLEn	PICKANDS	LME	AD2L	MDPD	MLEn	MDPD
4000(200)	PICKANDS	MLEn	PICKANDS	LME	AD2L	AD2R	MLEn	AD2R
10000(500)	PICKANDS	MLEn	PICKANDS	LME	AD2L	AD2R	MLEn	AD2R
20000(1000)	PICKANDS	MLEn	PICKANDS	LME	AD2L	AD2R	MLEn	AD2R

N = the sample size, (n) = the number of observations above the threshold

Table 9. Performance score for “best” estimator when varying shape parameter

For shape on both $\sigma = 1$ and $\sigma = 10$			For scale on both $\sigma = 1$ and $\sigma = 10$		
Method	Performance score	Rank	Method	Performance score	Rank
NON-POT					
CM	11.11	2nd	CM	11.11	2nd
ADR	11.11	2nd	ADR	11.11	2nd
MLE	11.11	2nd	MLE	11.11	2nd
ZHANG	61.11	1st	ZHANG	55.56	1st
LME	5.56	3rd	KS	5.56	3rd
			MLEn	5.56	3rd
POT					
PICKANDS	13.89	2nd	ZHANG	22.22	2nd
ADL	2.78	5th	MOMENTS	8.33	4th
AD2R	8.33	3rd	MPLE	11.11	3rd
LME	8.33	3rd	LME	22.22	2nd
MED	2.78	5th	MLEn	25.00	1st
ZHANG	27.78	1st	AD2R	2.78	6th
CM	5.56	4th	KS	5.56	5th
MLE	2.78	5th	NLS-P	2.78	6th
KS	5.56	4th	NLS-S	2.78	6th
PWMU	8.33	3rd			
PWMB	2.78	5th			
ADR	5.56	4th			
MDPD	2.78	5th			
MPLE	2.78	5th			

3.3. The Performance Score of Conventional Methods of the GPD

The overall performance of an estimator with respect to varying both the shape parameter and sample size was computed using a performance score and is given by:

$$PS = \frac{f_B}{\sum_{i=1}^p f_B} \quad (10)$$

where f_B is the number of times an estimator recorded the minimum error metric and p is the total number of values for shape parameter or sample size.

3.3.1. The Performance Score of Conventional Methods of the GPD Varying Shape Parameter

In Table 9 the performance score for the estimators by varying shape parameter for both the NON-POT and the POT framework are presented. All estimators reported in Table 9 are estimators that had at least a minimum error metric in comparison to other estimators on the scenario of the shape parameter being varied.

For the NON-POT framework, the Zhang estimator is

ranked as the overall best method for estimating both the shape and scale parameter. Also, for the POT framework, in estimating the shape parameter, the Zhang estimator is again the overall best while the MLEn is the overall best in estimating the scale parameter.

3.3.2. The Performance Score of Conventional Methods of the GPD Varying Sample Size

The performance score for the estimators by varying sample size for both the NON-POT and the POT framework is shown in Table 10. Again, all estimators reported in Table 10 are also estimators that had at least a minimum error metric in comparison to other estimators on the scenario of the sample size being varied. In this regard also, the Zhang estimator was ranked as the overall best method for estimating both the shape and scale parameter for the NON-POT framework. Under the POT framework, in estimating the shape parameter, the PICKANDS was ranked overall “best” while MLEn and the LME showed the highest performance score in estimating the scale parameter.

Table 10. Performance score for “best” estimator when varying sample size

For shape on both $\sigma = 1$ and $\sigma = 10$			For scale on both $\sigma = 1$ and $\sigma = 10$		
Method	Performance score	Rank	Method	Performance score	Rank
NON-POT					
ZHANG	75	1st	ZHANG	93.75	1st
PICKANDS	25	2nd	NLS-S	6.25	2nd
POT					
ZHANG	37.5	2nd	NLS-S	12.5	2nd
PICKANDS	62.5	1st	MLEn	43.75	1st
			LME	43.75	1st

4. Conclusions

Traditional methods of parameters estimation in the GPD have some identified defects. The results presented in the previous section have compared the accuracy of estimators under the assumption that the estimators are to perform better with increasing sample size and different shape parameters. Generally, none of the estimators consistently performed better with increasing shape parameters. Also, there is a large bias recorded when estimating the scale parameter with most of the estimators.

It was further observed from the results that with regards to the scale parameter being one (1), the existence of the MDPD estimator is restricted to $\xi = -3$ for all sample sizes under the POT framework with $q = 0.9$ and 0.95 . Thus, it is unable to estimate the value of shape parameter when $\xi = -3$. The MLE method with $\sigma = 10$, faces a convergence problem when $\xi = 5$ under the POT framework.

For NON-POT, the Zhang Estimator was "best" in performance in estimating both the shape and scale parameter for most heavy-tailed cases. Also, for the POT framework with $q = 0.9$ and $q = 0.95$, the Zhang Estimator was "best" in performance for estimating very heavy tails (i.e. $\xi = 3, 5$) for the shape and very short tails (i.e. $\xi = -3, -2$) for the scale regardless of the value of σ . Hence, the Zhang estimator was then ranked overall “best” in estimating both the shape and scale parameter for NON-POT case whilst again, the Zhang estimator the overall best in estimating the shape and the MLEn is the overall best in estimating the scale parameter for case of the POT framework.

Additionally, in estimating the shape parameter, under the NON-POT framework, the Zhang estimator performed “best” in all cases of scale parameter with increasing sample size. Under the POT framework, with $q=0.9$ and $q=0.95$, the Pickands Estimator was "best" in performance at estimating the shape parameter for large sample sizes while Zhang was for small sizes. However, in terms of overall “best” both the MLEn and the LME were ranked “best” in estimating the scale for the POT scenario.

When varying shape parameters, the MLEn was “worst” in estimating both the shape and scale parameter for most

short and medium tailed cases whereas for some heavy tails the AD2L and KS were “worst” in performance. Whilst when varying sample sizes, the MDPD, and the MGF estimators were “worst” in performance with at times having the same estimates.

In conclusion, the performance of estimators depends on sample size, the value of the shape parameter and also the value of scale parameter.

Appendix A. Supplementary Material

Supplementary material related to this article can be found online at <https://bit.ly/2EXd4UR>

REFERENCES

- [1] S. Kang, J. Song. Parameter and quantile estimation for the generalized Pareto distribution in peaks over threshold framework, Journal of the Korean Statistical Society, Vol.46, No.4, 487-501, 2017.
- [2] J. Beirlant, Y. Goegebeur, J. Segers, J. L. Teugels. Statistics of extremes: theory and applications, John Wiley & Sons, 2006.
- [3] L. De Hann, A. Ferreira. Extreme value theory: An introduction: Springer series in operations research, 2006.
- [4] P. de Zea Bermudez, S. Kotz. Parameter estimation of the generalized Pareto distribution – Part II. Journal of Statistical Planning and Inference, Vol.140, No.6, 1374-1388, 2010.
- [5] E. Castillo, A. S. Hadi. Fitting the generalized Pareto distribution to data. Journal of the American Statistical Association, Vol.92, No.440, 1609-1620, 1997.
- [6] J. R. Hosking, J. R. Wallis. Parameter and quantile estimation for the generalized pareto distribution. Technometrics, Vol.29, No.3, 339–349, 1987.
- [7] R. Deidda, M. Puliga. Performances of some parameter estimators of the generalized pareto distribution over rounded-off samples. Physics and Chemistry of the Earth, Parts A/B/C, Vol.34, No.10-12, 626– 634, 2009.
- [8] E. B. Mackay, P. G. Challenor, A. S. Bahaj. A comparison of

- estimators for the generalized Pareto distribution. *Ocean Engineering*, Vol.38, No.11-12, 1338-1346, 2011.
- [9] D. E. Giles, H. Feng, R. T. Godwin. Bias-corrected maximum likelihood estimation of the parameters of the generalized Pareto distribution. *Communications in Statistics-Theory and Methods*, Vol.5, No.8, 2465-2483, 2016.
- [10] D. J. Dupuis. Estimating the probability of obtaining nonfeasible parameter estimates of the generalized Pareto distribution. *Journal of Statistical Computation and Simulation*, Vol.54, No.1-3, 197-209, 1996.
- [11] F. Ashkar, C. N. Tatsambon. Revisiting some estimation methods for the generalized Pareto distribution. *Journal of Hydrology*, Vol.346, No.3-4, 136-143, 2007.
- [12] J. Zhang. Likelihood moment estimation for the generalized Pareto distribution. *Aust. NZ J. Stat.* Vol.49, No.1, 69-77, 2007.
- [13] J. del Castillo, I. Serra. Likelihood inference for generalized Pareto distribution. *Computational Statistics & Data Analysis*, Vol.83, 116-128, 2015.
- [14] J. Zhang, M. A. Stephens. A new and efficient estimation method for the generalized Pareto distribution. *Technometrics*, Vol.51, No.3, 316-325, 2009.
- [15] J. Zhang. Improving on estimation for the Generalized Pareto Distribution. *Technometrics*, Vol.52, No.3, 335-339, 2010.
- [16] J. Song, S. Song. A quantile estimation for massive data with generalized Pareto distribution. *Computational statistics & data analysis*, Vol.56, No.1, 143-150, 2012.
- [17] M. H. Park, J. H. Kim. Estimating extreme tail risk measures with generalized Pareto distribution. *Computational Statistics & Data Analysis*, Vol.98, 91-104, 2016.
- [18] A. Gharib, E. Davies, G. Goss, M. Faramarzi. Assessment of the combined effects of threshold selection and parameter estimation of generalized Pareto distribution with applications to flood frequency analysis. *Water*, Vol.9, No. 9, 692, 2017.
- [19] A. J. McNeil. *Statistical Analysis of Extreme Values: From Insurance, Finance, Hydrology and Other Fields*, 1516-1518, 1998.
- [20] J. Jockovic, P. Mladenovic. Coupon collector's problem and generalized Pareto distributions. *Journal of Statistical Planning and Inference*, Vol. 141, No.7, 2348-2352, 2011.
- [21] J. Jockovic. Quantile estimation for the generalized Pareto distribution with application to finance. *Yugoslav Journal of Operations Research*, Vol.22, No.2, 2016.
- [22] R. W. Katz, G. S. Brush, M. B. Parlange. Statistics of extremes: modeling ecological disturbances. *Ecology*, Vol 86, No.5, 1124-1134, 2005.
- [23] J. Kyselý, J. Píček, R. Beranová. Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold. *Global and Planetary Change*, Vol.72 No.1-2, 55-68, 2010.
- [24] A. Luceño. Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics & Data Analysis*, Vol.51, No.2, 904-917, 2006.
- [25] M. Ribatet. POT: modelling peaks over a threshold. *R. News* 7, 34-36, 2007.