STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **10**, March 2022, pp 505–518.
Published online in International Academic Press (www.IAPress.org)

A Different Way of Choosing a Threshold in a Bivariate Extreme Value Study

Andréhette Verster *, Nicholas Kwaramba

Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa

Abstract The choice of optimum threshold in Extreme Value Theory, peaks over threshold, has been a topic of discussion for decades. A threshold must be chosen high enough to control the bias of the extreme value index. On the other hand, if a threshold is chosen too high the variance becomes a problem. This is a very difficult trade-off and has been studied over the years from various viewpoints. More often these studies aim at methods for choosing the threshold in univariate settings. Not as many literature are available for choosing the threshold in a multivariate setting. In this paper we consider an approach for choosing the threshold when working with bivariate extreme values above a threshold. This approach makes use of Bayesian methodology. It adds value to the existing literature since it is also possible to use this approach without visual inspection.

Keywords Bayesian approach, bivariate extreme values, dependence parameter, threshold, Topp-Leone Pareto

AMS 2010 subject classifications 60G70, 62G32

DOI: 10.19139/soic-2310-5070-1318

1. Introduction

The choice of threshold plays an important role in Extreme Value Theory (EVT) where peak over threshold (POT) models are considered. POT models only model the observations above the chosen threshold. A threshold should be chosen high enough so that the bias of the extreme value index (EVI) is controlled, but on the other hand if it is chosen too high the variance becomes out of control. Various literature are devoted in trying to find a midway between these scenarios, see for example Coles [5], Guillou & Hall [8], Beirlant et al. [2], Thompson et al. [17], Wong & Li [21], Yang et al. [22] and Kiran & Srinivas [11] to name a few. A popular approach for choosing a threshold is to estimate the EVI at various thresholds. These estimates are then plotted against the various threshold values. Visual inspection is then required, an optimal threshold choice will be where the EVI estimates appears to be stable. Although this method is popular and rather simple to use, the drawback is that the plot is not always easy to inspect and a clear picture, of a stable EVI estimate area, is not always clearly visible, especially in real life situations. Verster & Raubenheimer [20] considered a generalized model, the Topp-Leone Pareto distribution, to assist in choosing a threshold in the γ positive domain. Their method also allows for choosing a threshold without visual inspection. Most of the above literature are concerned with the threshold choice in a univariate setting. The estimation of the dependence parameter plays a vital role in multivariate extreme values. It is essential to choose a suitable threshold because the threshold is important for estimating the dependence parameter effectively, see for example Joe et al. [10], Thibaud et al. [16] and Lee et al. [14]. It is therefore also essential to consider choosing an optimum threshold in a multivariate extreme value scenario. See for example Borsos [4] for a recent application of bivariate extreme models. In this paper we address the issue of choosing a threshold in a bivariate

*Correspondence to: Andréhette Verster(verstera@ufs.ac.za.)Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

extreme value setting. We introduce a different approach for choosing the threshold by incorporating the bivariate tail model of Ledford & Tawn [13] and the threshold approach of Verster & Raubenheimer [20]. A Bayesian approach is considered for modeling the dependence tail parameter. The rest of the paper is set out as follows: A brief summary on univariate POT models is given in Section 1. In Section 2 the theory of the Ledford & Tawn [13] bivariate tail model is explained as well as the Bayesian estimation of the tail parameter following the paper by Verster & Kwaramba [19]. Section 3 explains our method for choosing an optimum threshold by expanding on the work of Verster & Raubenheimer [20]. A small simulation study is conducted in Section 4 and Section 5 applies our threshold selection method on real data sets.

2. POT models

It is well known in EVT that the excesses of the conditional distribution, $X-u$, given $X>u$ can be modelled through a POT model known as the generalized Pareto distribution (GPD), see for example Beirlant et al. [2] and de Haan and Ferreira (2006). The GPD is given in (1), where u is the threshold

$$P\left(X-u>y\,|X>u\right)=\frac{\overline{F}\left(u+y\right)}{\overline{F}\left(u\right)}\rightarrow\overline{H}_{\gamma}\left(y\right)=\left(1+\gamma y\right)^{-\frac{1}{\gamma}},\quad x>u,\quad\gamma\in\mathbb{R}. \tag{1}$$

It often happens that one is only interested in Pareto-type scenarios where γ is positive. The survival function of Pareto-type distributions is given in (2) with slowly varying function l, satisfying $\frac{l(xu)}{l(u)}\rightarrow 1$ as $u\rightarrow\infty$ Beirlant et al. [2]. The POT model in (1) can then be simplified to a more convenient (simpler) model, given in (3), as as $u\rightarrow\infty$

$$\overline{F}\left(x\right)=x^{-\gamma}l\left(x\right),\quad x>1, \tag{2}$$

$$P\left(\frac{X}{u}>y\,|X>u\right)\rightarrow\overline{H}_{\gamma}\left(y\right)=y^{-\gamma},\quad y>1. \tag{3}$$

The density function of (3) is:

$$f\left(y,\gamma\right)=\gamma\left(y\right)^{-\left(\gamma+1\right)},\quad\gamma>1,y>\frac{x}{u}. \tag{4}$$

For convenience (4) is referred to as the probability density function of the bounded Pareto distribution. A threshold should be chosen high enough so that the bias of the EVI, $\frac{1}{\gamma}$, is controlled, but on the other hand if it is chosen too high the variance becomes out of control. In this paper we consider a Bayesian approach for estimating the EVI. If Jeffrey's prior is assigned to γ, $p\left(\gamma\right)\propto\frac{1}{\gamma}$, the posterior distribution (prior x likelihood function) is

$$p(\gamma|y)\&\propto\gamma^{n-1}e^{-\gamma(\sum_{i=1}^{n}\log y_i)}, \tag{5}$$

Beirlant et al. [2]. It is clear that the posterior of γ follows a Gamma distribution and that the EVI, $\frac{1}{\gamma}$, follows an Inverse Gamma distribution with parameters n and $\sum_{i=1}^{n}\log\left(y_i\right)$, where n is the number of observations above the threshold.

3. Bivariate tail model

In bivariate EVT the estimation of the dependence parameter plays an important role. Ledford and Tawn (1996) approaches the modeling of bivariate extreme values in the following way: The joint tail model is: $P(X>r,Y>r)\sim\mathcal{L}(r)\;r^{-1/\eta}$ as $r\rightarrow\infty$ for random vector $(X,\;Y)$ where $\mathcal{L}$ is a slowly varying function with $\eta\in(0,1]$.

η is known as the tail index or dependence parameter. If $\eta = 1$ it indicates that the two random variables are asymptotically dependent and $\eta < 1$ refer to asymptotic independence where η measures the amount of tail dependence within the asymptotical independence (Beirlant & VandeWalle [3]). Let $S = \min$ then

$$P(S > r) = P(X > r, Y > r) \sim \mathcal{L}(r)\, r^{-1/\eta}, \quad r \to \infty. \tag{6}$$

Equation (6) corresponds to (2), thus $P(\frac{S}{u} > r | S > u) \to \bar{H}_\eta(r) = r^{-\frac{1}{\eta}}$, $r > 1$ as $u \to \infty$ (see(3)). The random vector $(X,\ Y)$ has unit Fréchet or Pareto marginals (Draisma et al. [6]). In this paper we will transform the marginals to unit Pareto distributions as explained in Beirlant & VandeWalle [3] as follows: Suppose that $R_{X,i}$ is the rank of $X_i,\ i = 1,\ \ldots, n$ and $R_{Y,i}$ is the rank of $Y_i,\ i = 1,\ \ldots, n$ then

$$S_i = \min\left(-\frac{1}{log\left(\frac{R_{X,i}}{n+1}\right)}, -\frac{1}{log\left(\frac{R_{Y,i}}{n+1}\right)}\right). \tag{7}$$

The dependence parameter, η can be estimated through any well know EVI estimator, such as the Hill estimator, generalized Hill estimator and biased reduced Hill estimator to name a few. Refer to Beirlant et al. [2] and Albrecher et al. [1] for more information on these estimates. In this paper we consider the Bayesian estimate introduced by Verster & Kwaramba [19]. Since the dependence parameter η lies between 0 and 1, a Beta prior is chosen on η, $\pi(\eta) \propto \eta^{a-1}(1-\eta)^{b-1}$. The prior allows us to build prior information about η into the estimation process. The posterior is then

$$\pi(\eta|a,b,t) \propto\ \eta^{a-1}(1-\eta)^{b-1}\prod_{i=1}^{n}\frac{1}{\eta}t_i^{-1-\frac{1}{\eta}} \propto\ \eta^{a-n-1}(1-\eta)^{b-1}e^{-\frac{1}{\eta}\sum_{i=1}^{n}\log(t_i)},\ t_i > 1,\ 0 < \eta < 1, \tag{8}$$

where $t_i = \frac{s_i}{u}$ represents the relative excesses above the threshold and n the number of observations above the threshold Verster & Kwaramba [19]. It can be shown that the posterior is proper, refer to (Verster & Kwaramba [19]) for more information. In their paper they showed that when a large number of η's is simulated from the posterior the average over these simulated η's provides a good estimate for η. We assume that the hyper parameters, a and b, are known and fixed at each threshold. Since the excesses above a sufficient threshold, follows a Pareto-type distribution and according to (5) the posterior of $\frac{1}{\gamma}$ follows an Inverse Gamma distribution, we use an empirical Bayes approach to choose a and b accordingly. The expected value and variance of an Inverse Gamma are $E(\frac{1}{\gamma}|y) = \frac{\sum_{i=1}^{n}\log y_i}{n-1}$ and $Var(\frac{1}{\gamma}|y) = \frac{(\sum_{i=1}^{n}\log y_i)^2}{}(n-1)^2(n-2)$ respectively. One can then solve a and b by setting the expected value and variance of the Beta prior equal to $E(\frac{1}{\gamma}|y)$ and $Var(\frac{1}{\gamma}|y)$ respectively. The main question however remains, where should one choose the threshold, such that η is estimated properly and that the correct inference is made about the dependence structure between the random variables?

4. Top Leone generalization

Verster & Raubenheimer [20] showed that the Topp Leone Pareto distribution can successfully be used to indicate a suitable threshold when $\gamma > 0$. This generalization of the Topp & Leone [18] distribution given by Rezaei et al. [15] replaces the original $0 \le x \le 1$ with the CDF of any baseline distribution. Verster & Raubenheimer [20] showed that when the bounded Pareto CDF is considered as the baseline distribution the generalization parameter, α, assists in choosing the threshold. This is because $\alpha = 1$ refers back to the bounded Pareto distribution with $EVI = \frac{\gamma}{2}$ (see Verster & Raubenheimer [20]). In this paper we consider the generalized Topp-Leone distribution of Rezaei et al. [15] with CDF and pdf given in (9) and (10)

$$F(x;\alpha,\boldsymbol{\theta}) = \{G(x;\boldsymbol{\theta})[2-G(x;\boldsymbol{\theta})]\}^\alpha = \left\{1-[1-G(x;\boldsymbol{\theta})]^2\right\}^\alpha \tag{9}$$

and

$$f\left(x;\alpha,\boldsymbol{\theta}\right)=2\alpha g\left(x;\boldsymbol{\theta}\right)\left[1-G\left(x;\boldsymbol{\theta}\right)\right]\left\{1-\left[1-G\left(x;\boldsymbol{\theta}\right)\right]^{2}\right\}^{\alpha-1} \tag{10}$$

where $g\left(x;\boldsymbol{\theta}\right)$ and $G\left(x;\boldsymbol{\theta}\right)$ denote the probability density and CDF of the bounded Pareto from (6). Thus, the Top-Leone Pareto (TLPa) density becomes:

$$f\left(y;\ \alpha,\gamma\right)=\frac{2\alpha}{\eta}y^{-2/\eta-1}\left(1-y^{-2/\eta}\right)^{\alpha-1}, \tag{11}$$

where $y>1,\ \eta>0$ (the dependence parameter) and $\alpha>0$. If $\alpha=1$, (11) becomes the bounded Pareto distribution with $EVI=\xi=\frac{\eta}{2}$. The main idea in this study is the following: according to the POT theory from Section 1, the TLPa becomes the bounded Pareto (at a suitable high threshold) such that $\alpha\to1$ as $n\to\infty$. One would therefore inspect the behavior of α over different threshold values. At a sufficient threshold, α will be close to one. Again, we consider a Bayesian approach to estimate the two parameters, ξ and α. Since the dependence parameter lies between 0 and 1 a Beta prior is assigned to ξ, a Jeffreys prior is assigned to α, $p(\alpha)\propto1/\alpha$. The independent joint prior is then: $p(\xi,\alpha)\propto\alpha^{-1}\xi^{a-1}\left(1-\xi\right)^{b-1}$ with posterior:

$$p\left(\xi,\alpha|a,b,y\right)\propto\alpha^{n-1}\xi^{a-1-n}\left(1-\xi\right)^{b-1}e^{-\left(\frac{1}{\xi}+1\right)\sum_{i=1}^{n}\log y_i}e^{(\alpha-1)\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)}. \tag{12}$$

The conditional posteriors of α and ξ can be derived as:

$$p\left(\alpha|\xi,a,b,y\right)\propto\alpha^{n-1}e^{\alpha\left[\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)\right]}\sim Gamma\left(n,-\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)\right), \tag{13}$$

and

$$p\left(\xi|\alpha,a,b,y\right)\propto\xi^{a-1-n}\left(1-\xi\right)^{(b-1)}e^{-\frac{1}{\xi}\sum_{i=1}^{n}\log y_i}\,e^{(\alpha-1)\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)}, \tag{14}$$

respectively. It can be shown that (14) is proper, see the Appendix. The following section inspects some bivariate scenarios through a small simulation study to inspect the behavior of ξ and α.

5. Simulations

In this section, 300 observations are simulated from the following distributions:

i bivariate Normal distribution with $\rho=0.9$ and $\rho=0.6$. The true dependence parameter is $\frac{1+\rho}{2}$ [13].
ii bivariate Morgenstern distribution with α_{BM}= 0.75, 0 and -0.75 respectively. The true dependence parameter is 0.5, for all values of α_{BM} [13].
iii bivariate extreme value distribution with a logistic dependence structure with α_{BEV}= 0.6 and 0.8. The true dependence parameter is 1, for all values of α_{BEV} [13].

For each simulation scenario the dependence parameter, ξ, and α will be estimated over a range of thresholds, from the 5th largest observation to the 10th smallest observation. For each threshold 2000 ξ's and α's are simulated from the conditional posteriors in (13) and (14) through a Gibbs sampler technique. The simulation time is greatly reduced, since the conditional posterior of α can be expressed as a Gamma distribution, The mean over the 2000 simulated parameter values are taken as an estimate at each threshold. The whole simulation process is repeated 100 times. The average over the 100 parameter estimates is then taken at each threshold. The popular Hill estimate is also calculated at each threshold as a comparison to our estimate. The Hill estimate is

$$\hat{\eta}_{Hill}=\frac{1}{k}\sum_{j=1}^{k}\log S_{n-j+1,n}-\log S_{n-k,n}. \tag{15}$$

[9] where k represents the number of observations above the threshold. The aim of this study is to investigate a method for choosing an optimum threshold. Therefore, the focus is mainly of the performance of α (as it moves towards one) and not on the dependence parameter's performance, and the comparison thereof with other known dependence parameter estimators. The root mean square error (RMSE), given below are also calculated for the various simulations at each threshold. A small RMSE (close to zero) is preferred.

$$RMSE(\hat{\xi}) = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N=2000}\left(\hat{\xi}_j - \xi\right)^2\right)} \tag{16}$$

The left side of Figures (1- 7) shows the RMSE of ξ and the right side shows the estimate of α over the various thresholds. From Figures (1- 7) (right side) one can see that there are a range of k values for which α is close to one. This will be the region where the excesses follow the bounded Pareto distribution and where the optimum threshold will lie. For threshold values where α is not close to one the excesses do not yet follow a bounded Pareto distribution and should rather not be chosen as an optimum threshold. Figures(1- 7) (left side) shows that for most of the simulations (except for the bivariate Morgenstern, α = 0.75) the EVI estimate of our method outperforms the Hill, since the RMSE is smaller over the different thresholds. In the case of the bivariate Morgenstern, α= 0.75, the performance of the two estimators is similar. This shows that our proposed Beta prior and Bayesian approach is appropriate for estimating the EVI. The RMSE (in all the figures) also shows values close to zero in the range of thresholds that coincides with α values close to one, confirming an appropriate choice of threshold.

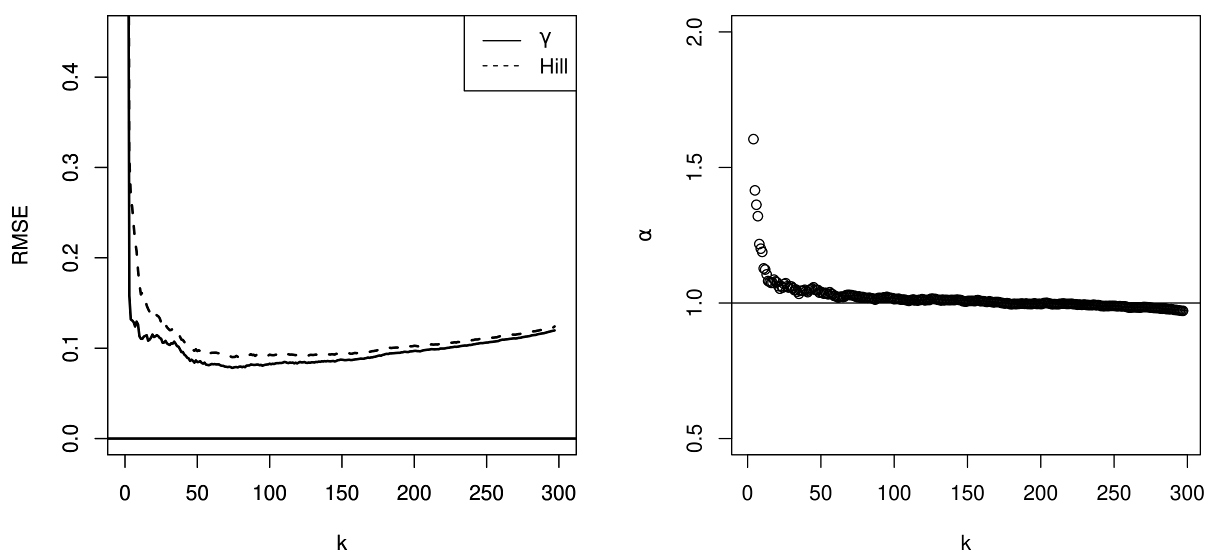


Figure 1. RMSE of ξ (left) and α estimates (right) from a bivariate Normal distribution with $\rho = 0.6$ (true EVI is 0.8).

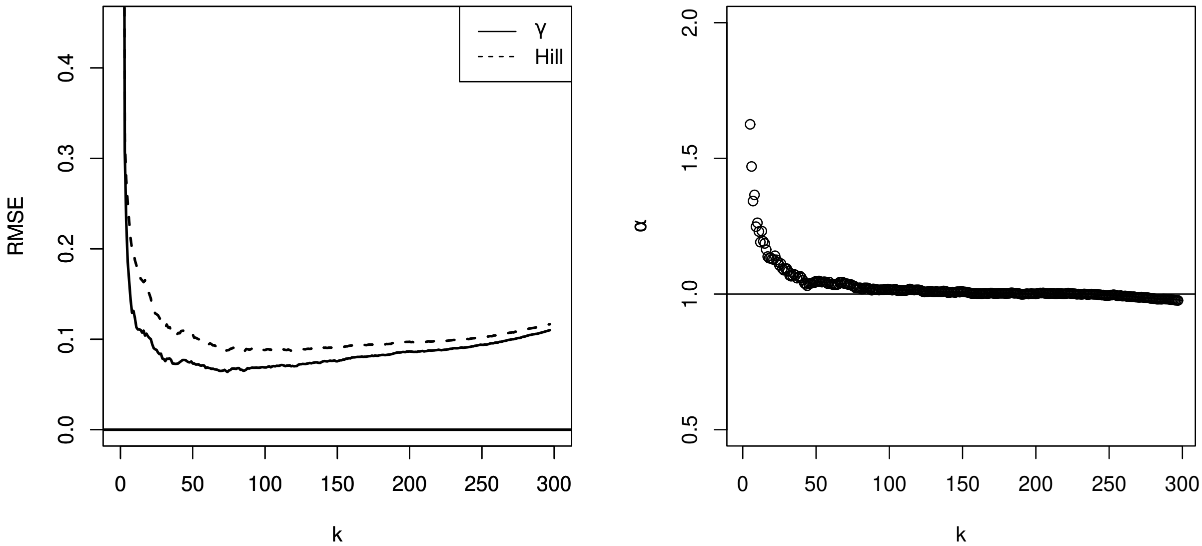


Figure 2. RMSE of ξ (left) and α estimates (right) from a bivariate Normal distribution with $\rho = 0.9$ (true EVI is 0.95).

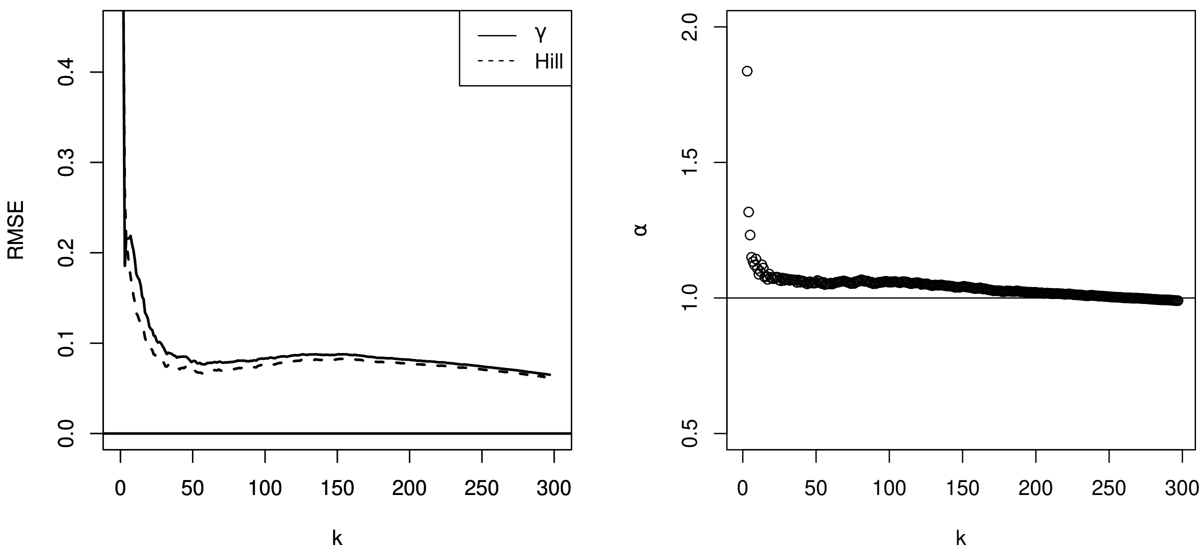


Figure 3. RMSE of ξ (left) and α estimates (right) from a bivariate Morgenstern distribution with α=0.75 (true EVI is 0.5).

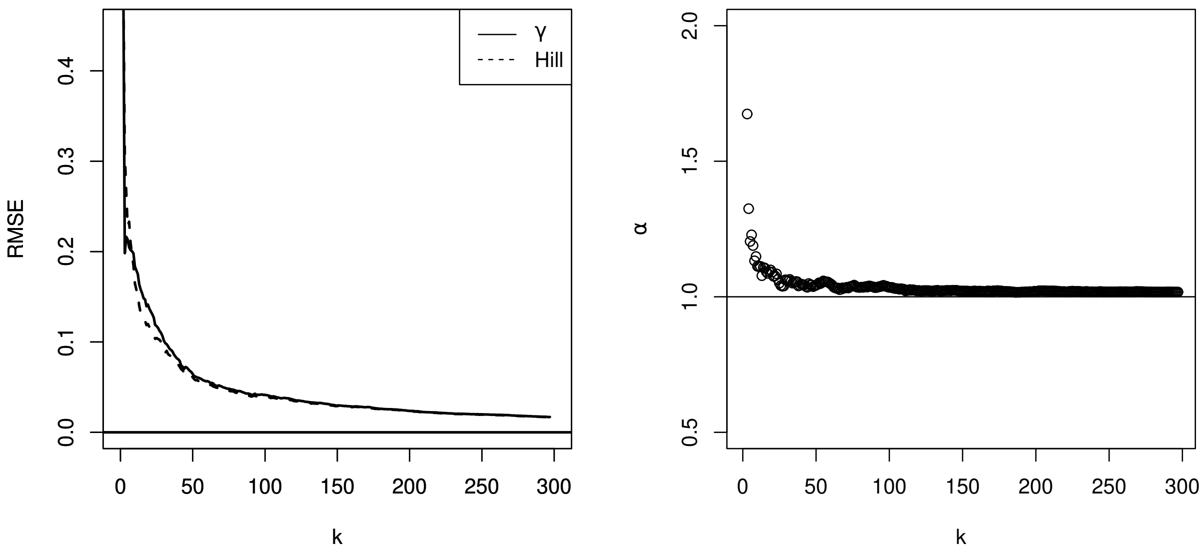
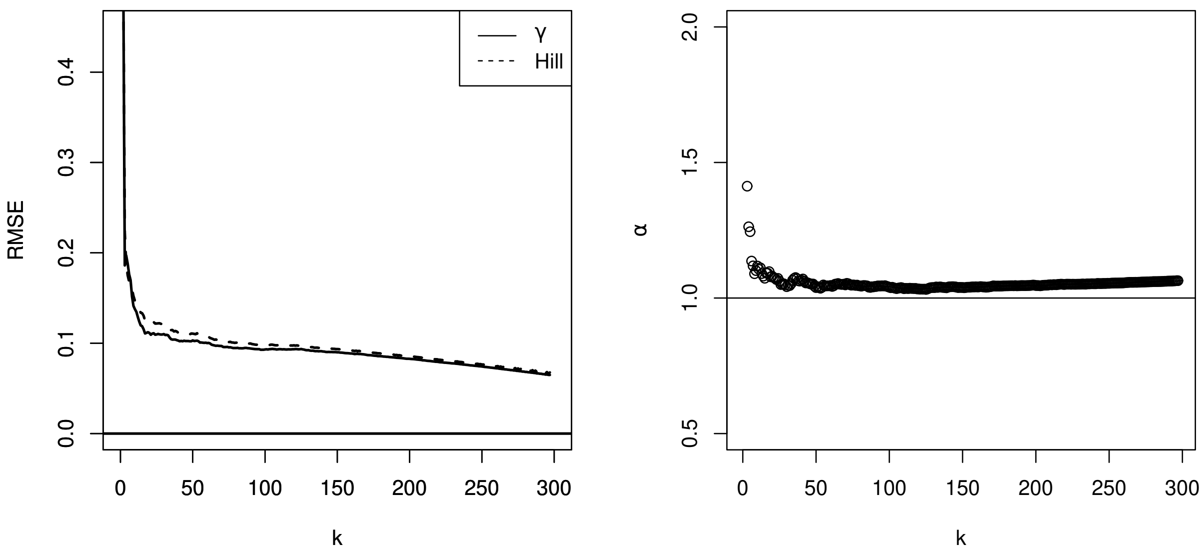


Figure 4. RMSE of ξ (left) and α estimates (right) from a bivariate Morgenstern distribution with α=0 (true EVI is 0.5).



fig

Figure 5. RMSE of ξ (left) and α estimates (right) from a bivariate Morgenstern distribution with α=-0.75 (true EVI is 0.5).

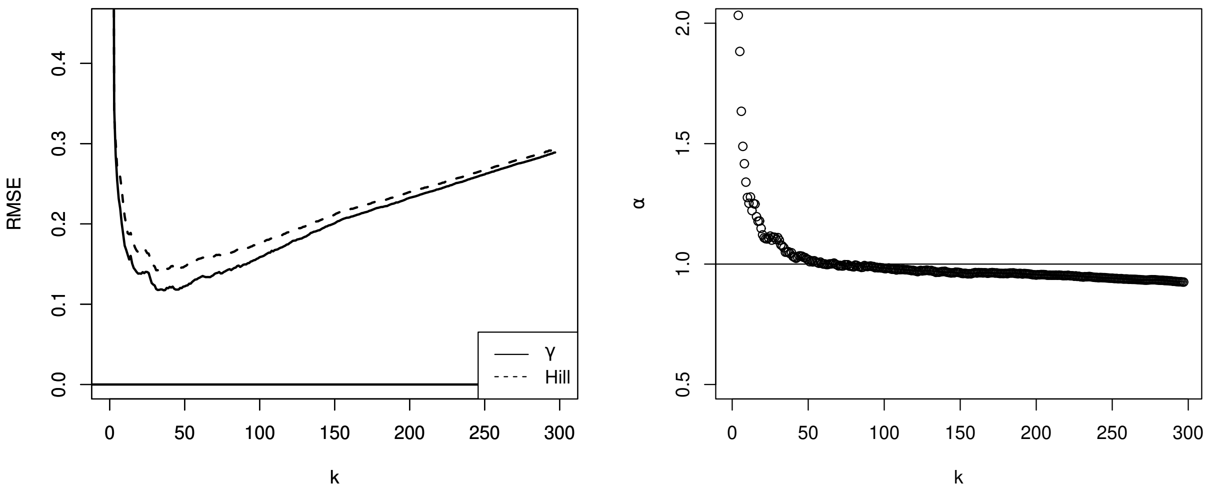


Figure 6. RMSE of ξ (left) and α estimates (right) from a bivariate extreme value distribution with α=0.6 (true EVI is 1).

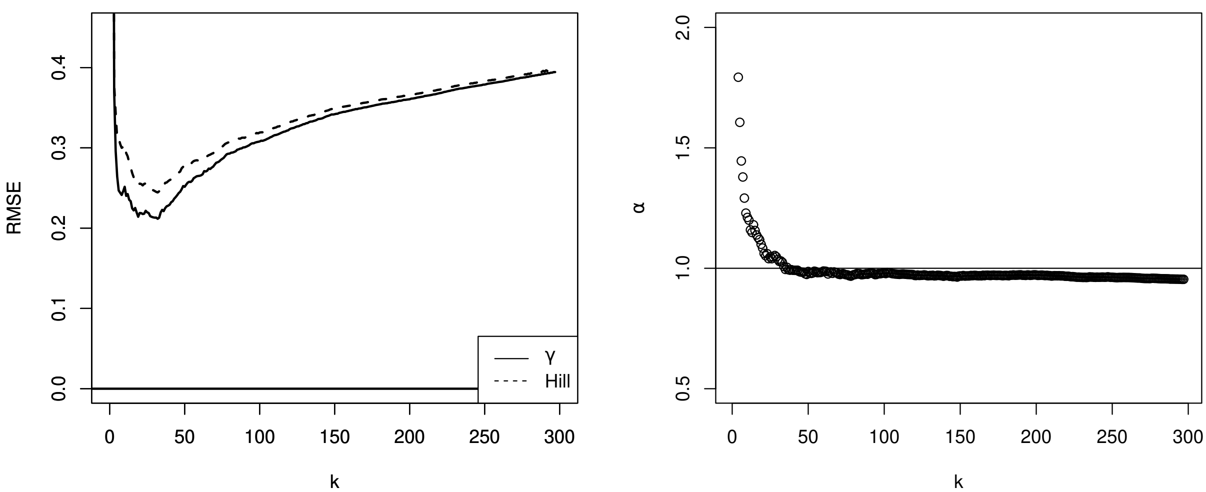


Figure 7. RMSE of ξ (left) and α estimates (right) from a bivariate extreme value distribution with α = 0.8 (true EVI is 1).

6. Real data Sets

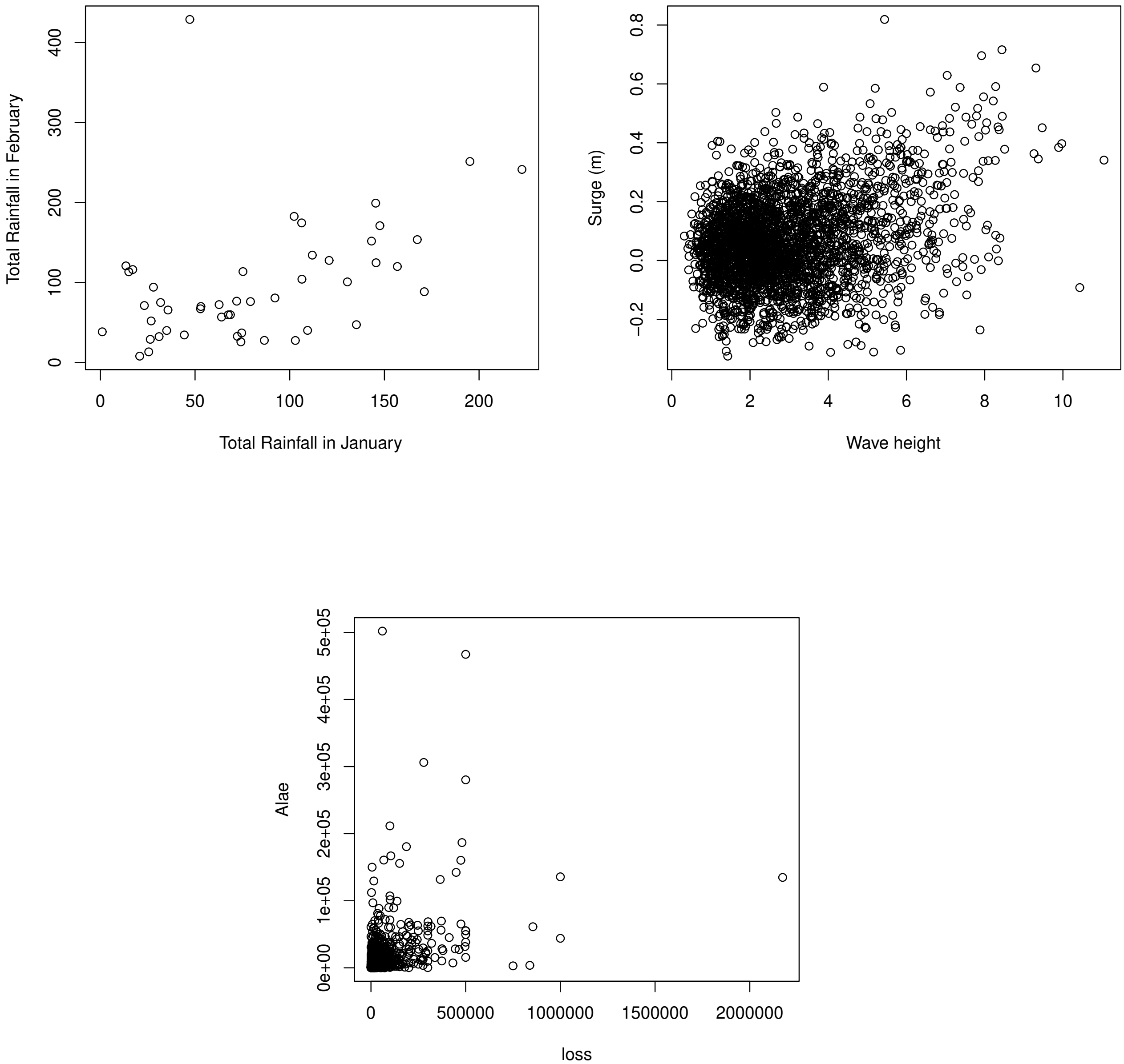


Figure 8. Scatterplots of the Bloemfontein rainfall (top-left), wave and surge height (top-right) and Loss-ALAE (bottom) datasets.

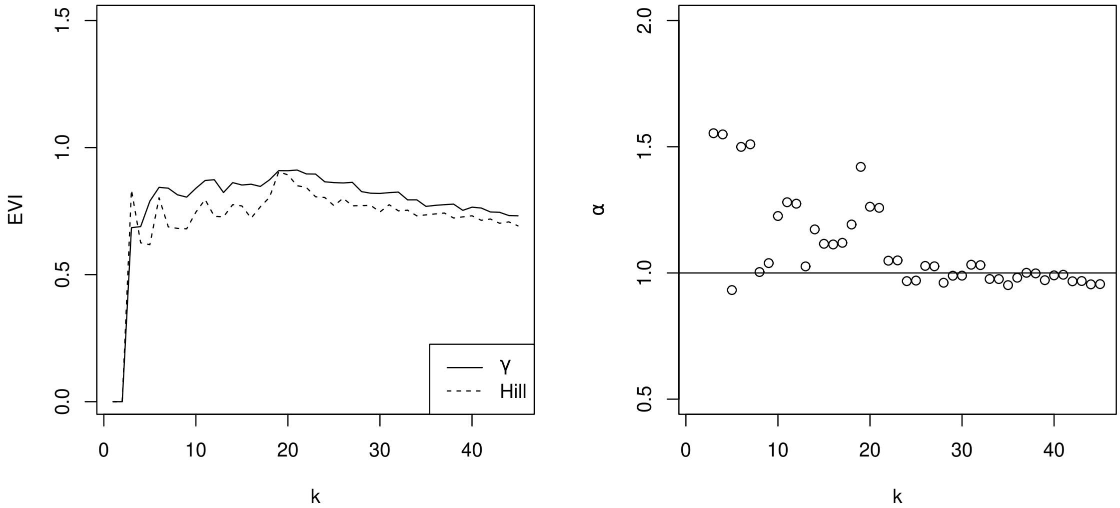


Figure 9. EVI estimate (left) and α estimate (right) for Bloemfontein rain

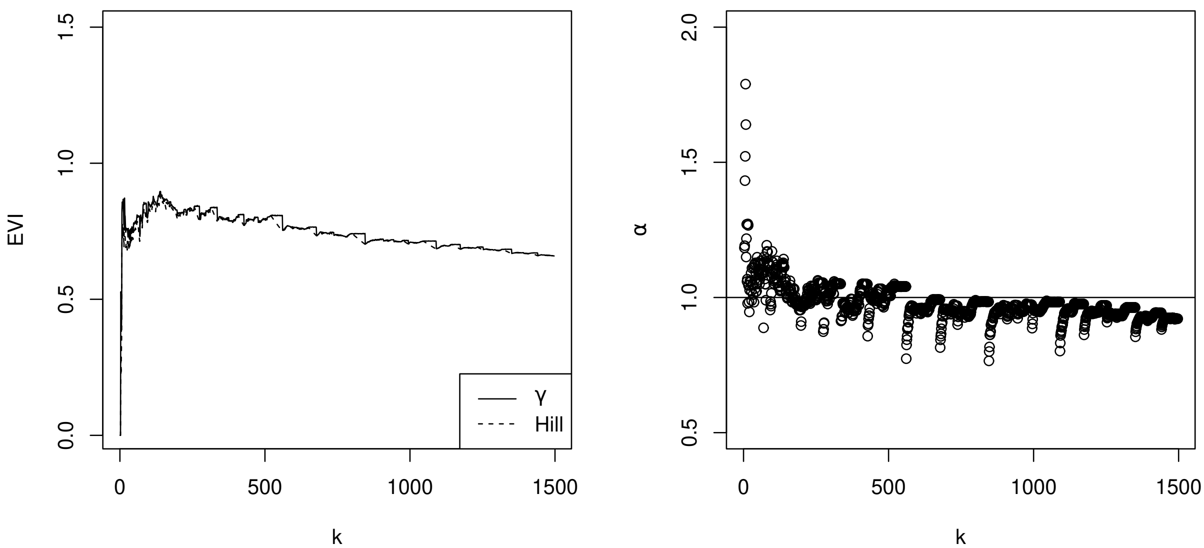


Figure 10. EVI estimate (left) and α estimate (right) for the wave and surge height.

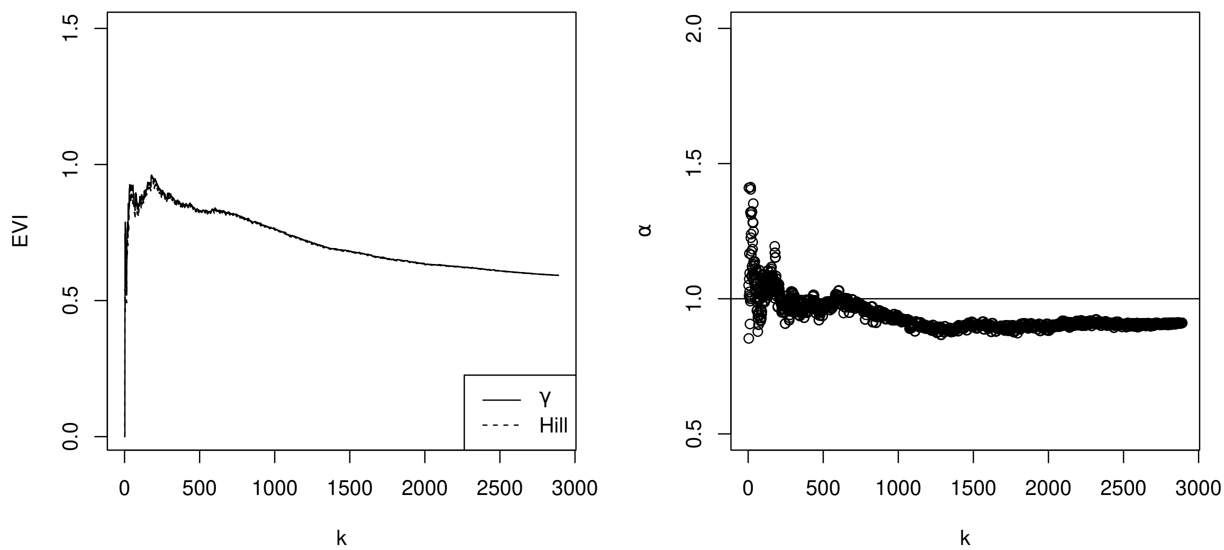


Figure 11. EVI estimate (left) and α estimate (right) for the Loss-ALAE.

Three datasets are considered: the wave and surge heights data, the Bloemfontein rainfall data and the Loss-ALAE data. The wave and surge heights data contain 2894 data pairs with a wave height and surge height measurement in metres at a single location of South-West England. The data frame can be found in the "ismev" package in R, https://cran.r-project.org/web/packages/ismev/ismev.pdf. The Bloemfontein rainfall dataset contains the total rainfall for the months of January and February from 1970 to 2017 (refer to Verster & Kwaramba [19]). The Loss-ALAE data contains 1500 data pairs with indemnity payment (loss) and the allocated loss adjustment expense (alae) observations in the USD. More information on this dataset can be found in Frees & Valdez [7] and Klugman & Parsa [12]. The data frame can be found in the "evd" package in R, https://cran.r-project.org/web/packages/evd/evd.pdf. Figure 8 shows the scatterplots of the three datasets, Bloemfontein rainfall (top, left), wave and surge heights (top, right) and Loss-ALAE (bottom). It is clear from the figures that dependence is present. Figures (9- 11) (left) shows the ξ estimate over the different thresholds and Figures (9- 11) (right) shows the α estimate. From Figures(9- 11) (left column) it is difficult to see a stable area over the thresholds. Choosing an optimum threshold (and an ξ estimate) from the left-side graphs is difficult. However, the right column of Figures (9- 11) assists in the choice of threshold. Through visual inspection one can see that there is an area where α is close to one. For the Bloemfontein rainfall, Figure 9 (right) α is close to one for $k \approx 20 - 40$ observations above the threshold. For the wave and surge heights, Figure 10 (right) α is close to one for $k \approx 250 - 750$ observations above the threshold. In the Loss-ALAE case, Figure 11 (right) α is close to one for $k \approx 250 - 500$ observations above the threshold. These will be the thresholds for which the excesses follow a bounded Pareto distribution and k should be chosen in this region. Up to now the threshold was chosen visually by inspecting the behavior of α. We will choose the threshold value in a region where α is close to one. It is also possible to choose a threshold without visual inspection by calculating the squared error between the simulated α's and 1 at different thresholds. This can easily be achieved by using (13). The conditional posterior of α, given ξ and the data above the threshold, follows a Gamma distribution. For a set of ξ values, α's can be simulated (from the Gamma distribution) at different thresholds. The threshold that matches the smallest squared error can then be chosen (together with the ξ estimate at that threshold). This approach was considered for the three datasets and the corresponding values were calculated as follows:

- Bloemfontein rainfall: Threshold: $k = 31$, EVI estimate = 0.85.

- Wave and surge heights: Threshold: k = 659, EVI estimate = 0.83011.
- Loss-ALAE: Threshold: k = 248, EVI estimate = 0.8288.

7. Conclusion

In this study we have shown that the TLPa distribution can be used successfully when modelling the minimum transformed bivariate extreme values excesses above a threshold. The TLPa was previously used in a univariate setting. In this paper the TLPa with dependance parameter, ξ, and generalization parameter, α, is fitted to bivariate extreme observations. The parameter, α plays a valuable role in choosing the optimum threshold. It is well known in EVT that above a suitable threshold the relative excesses follow a Pareto-type distribution. For this to hold, α in the TLPa (also a POT model) must go to one. A Bayesian approach is considered, and a Beta prior is assumed on the dependence parameter since we know the parameter lies between 0 and 1. The conditional distribution of α given the dependence parameter and the data was derived in Section 4 and it is was shown that this conditional distribution can be expressed as a Gamma distribution. This expression reduces the simulation time. The advantage of our method is that the conditional posterior (13) can be used to choose a threshold without visual inspection by considering the mean squared difference. This was shown in the real-life examples of Section 6.

Appendix

The conditional density of $\xi|a,\ b,\ \alpha,\ y$ is

$$p\left(\xi|\alpha,a,b,y\right)\propto\xi^{a-1-n}\left(1-\xi\right)^{(b-1)}e^{-\frac{1}{\xi\sum_{i=1}^{n}\log y_i}}\ e^{(\alpha-1)\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)}$$

where

$$e^{(\alpha-1)\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)}=\prod_{i=1}^{n}\left(1-y_i^{-\frac{1}{\xi}}\right)^{(\alpha-1)}=\prod_{i=1}^{n}\left(1-e^{-1/\xi\sum_{i=1}^{n}logy_i}\right)^{\alpha-1}.$$

Then Using Maclaurin expansion $e^{-1/\xi\sum_{i=1}^{n}logy_i}\approx1-1/\xi\sum_{i=1}^{n}logy_i$ and $e^{(\alpha-1)\sum_{i=1}^{n}\log\left(1-y_i^{-\frac{1}{\xi}}\right)}=\prod_{i=1}^{n}\left(\frac{1}{\xi}logy_i\right)^{\alpha-1}$ the posterior becomes

$$p(\xi|\alpha,a,b,y)\propto\xi^{a-1-\alpha n}(1-\xi)^{(b-1)}e^{-\frac{1}{\xi}\sum_{i=1}^{n}\log y_i}.$$

To show that $p\left(\xi|\alpha,a,b,y\right)$ is proper we have to show that $\int_0^1 p\left(\xi|\alpha,a,b,y\right)d\xi=c$, where c is a constant. Approximate $e^{-\frac{1}{\xi}\sum_{i=1}^{n}\log y_i}$ with Maclaurin expansion as $1-\frac{1}{\xi}\sum_{i=1}^{n}\log y_i$, then

$$\int_0^1 p\left(\xi|\alpha,a,b,y\right)d\xi=\int_0^1\xi^{a-1-\alpha n}\left(1-\xi\right)^{b-1}d\xi-\sum_{i=1}^{n}\log y_i\int_0^1\xi^{a-2-\alpha n}\left(1-\xi\right)^{b-1}d\xi=c$$

where $c=\left\{\frac{\Gamma(b)\Gamma(a-\alpha n)}{\Gamma(a+b-\alpha n)}-\frac{\sum_{i=1}^{n}\log y_i\ \Gamma(b)\Gamma(a-\alpha n-1)}{\Gamma(a+b-\alpha n-1)}\right\}^{-1}$

REFERENCES

[1] Albrecher, H., Beirlant, J., & Teugels, J. (2017). *Reinsurance: Actuarial and Statistical Aspects.* John Wiley and Sons.

[2] Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. (2004). *Statistics of Extremes: Theory and Applications.* Chichester: John Wiley and Sons.

[3] Beirlant, J. & VandeWalle, B. (2002). Some comments on the estimation of a dependence index in bivariate extreme value statistics. *Statistics & Probability Letters.*, 60(3), 265 – 278.

[4] Borsos, A. (2021). . application of bivariate extreme value models to describe the joint behavior of temporal and speed related surrogate measures of safety. *Accident Analysis and Prevention.*, 159.

[5] Coles, S. G. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.

[6] Draisma, G. H., Drees, A., Ferreira, A., & de Haan, L. (2001). Dependence in independence. *Technical report, Eurandom.*

[7] Frees, E. & Valdez, E. (1998). Understanding relationships using copulas. *North American Actuarial Journal.*, 2(1), 1–15.

[8] Guillou, A. & Hall, P. (2001). A diagnostic for selecting the threshold in extreme value analysis. *Journal of the Royal Statistical Society .*, 63, 293–305.

[9] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics.*, 3(5), 1163 – 1174.

[10] Joe, H., Smith, R. L., & Weissman, I. (1992). Bivariate threshold methods for extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology).*, 54, 171 – 183.

[11] Kiran, K. G. & Srinivas, V. (2021). A mahalanobis distance-based automatic threshold selection method for peaks over threshold model. *Water Resources Research,*, 57.

[12] Klugman, S. A. & Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics.*, 24, 139 – 148.

[13] Ledford, A. & Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika.*, 8, 168 – 187.

[14] Lee, J., Fan, Y., & Sisson, S. A. (2015). Bayesian threshold selection for extremal models using measures of surprise. *computaional Statistics and Data Analysis.*, 85, 84 – 99.

[15] Rezaei, S., Sadr, B. B., Alizadeh, M., & Nadarajah, S. (2017). Topp-Leone generated family of distributions: Properties and applications. *Communications in Statistics - Theory and Methods.*, 46(6), 2893 – 2909.

[16] Thibaud, E., Mutzner, R., & Davison, A. C. (2013). Threshold modeling of extreme spatial rainfall. *Water Resources Research.*, 49, 4633 – 4644.

[17] Thompson, K., Bernier, K., & Chan, P. (2013). Extreme sea levels, coastal flooding and climate change with a focus on atlantic canada. *Natural Hazards.*, 51, 139 –150.

[18] Topp, C. W. & Leone, F. C. (1955). A family of J-shaped frequency functions. *Journal of the American Statistical Association.*, 50, 209 – 219.

[19] Verster, A. & Kwaramba, N. (2021). Estimating the dependence parameter in bivariate extreme value statistics through a bayesian approach. *Under Review (Available on Request).*

[20] Verster, A. & Raubenheimer, L. (2020). A different approach for choosing a threshold in peaks over threshold. *Statistics, Optimization & Information Computing.*

[21] Wong, T. & Li, W. (2010). A threshold approach for peaks-over-threshold modeling using maximum product of spacings. *Statistica Sinica.*, 20, 1257 – 1272.

[22] Yang, X., Zhang, J., & Ren, W. X. (2010). Threshold selection for extreme value estimation of vehicle load effect on bridges. *International Journal of Distributed Sensor Networks.*, 14.