

# Default priors based on pseudo-likelihoods for the Poisson-GPD model

Stefano Cabras

**Abstract** Extreme values are usually modeled with the peaks over the threshold approach by means of the Poisson-Generalized Pareto Distribution (Poisson-GPD). This model is governed by three parameters: the Poisson rate, the scale and the shape of the GPD. The quantity of interest, in many applications, is the mean return level which is a function of Poisson-GPD parameters. Moreover, the shape parameter of GPD is itself of interest in order to gain more insights on the underlying extremal process. For a suitable orthogonal parametrization, we derive matching priors for shape, scale and Poisson rate parameters based on an approximate conditional pseudo-likelihood. The formal rule, used here to obtain such priors, in some cases leads to the same priors obtained with Jeffreys' and Reference procedures. Moreover, we can provide a formal proof that each marginal prior for shape and scale parameters, respectively, are second order matching priors. We estimate the coverages of the corresponding posterior credible intervals and apply our approach to an example from hydrology.

**Key words:** Approximate conditional likelihood, Bayesian statistics, Extreme values, Matching priors, Orthogonal parameters

## 1 Introduction

In the past two decades there has been an increasing interest for statistical models of extreme events in numerous disciplines such as environmental sciences, finance and insurance, among others (see for instance [12] and [32]).

We consider the class of models arising from extreme value theory ([21] and [29]) with particular attention to model the exceedances of a random variable  $Z \in \mathcal{Z} \subset \mathbb{R}$

---

Stefano Cabras

Department of Mathematics and Informatics - University of Cagliari, Via Ospedale, 72 09124-Cagliari (Italy), e-mail: s.cabras@unica.it

over a fixed high threshold,  $u$  ([16]). Let  $Z \sim H(z)$  represent the observable variable (i.e. rainfall level or return of a financial asset) and  $Y = Z - u > 0 | Z > u$  represent the exceedance over  $u$ . [29] shows that, under regularity conditions on  $H(z)$ , the limiting distribution of  $Y$ , given  $Z > u$ , when  $u \rightarrow z^* = \sup \mathcal{Z}$ , is the Generalized Pareto Distribution (GPD) with shape  $\xi$  and scale  $\sigma > 0$ ,

$$f(y | \xi, \sigma) = \begin{cases} \sigma^{-1} (1 + y\xi/\sigma)^{-(1+\xi)/\xi}, & \text{if } (1 + y\xi/\sigma) > 0 \\ \sigma^{-1} \exp(-y/\sigma), & \xi = 0 \end{cases}$$

In most applications the shape parameter is of interest because:

- a)  $\xi$  regulates the existence of an upper limit in the  $Y$  variation. In fact, when  $\xi < 0$ ,  $z^* = u - \sigma/\xi < \infty$ , while  $\xi \geq 0$  implies  $z^* = \infty$ .
- b)  $\xi$  regulates tail thickness of  $H(z)$  and, in particular,  $1/\xi$  determines the number of finite moments of  $Y$ , i.e., for  $\xi = 1$ ,  $E(Y) < \infty$  and  $E(Y^2) = \infty$ .

Because of a), model  $f(\cdot)$  is irregular: for  $\xi < 0$  the support of  $Y$  depends on  $\sigma$  and, in general, it does not satisfy usual regularity conditions for the maximum likelihood estimator (MLE) ([33]). In particular, the MLE does not exist when  $\xi < -1$  and they are not regular for  $-1 < \xi < -0.5$ , in order to avoid such irregularities, we assume  $\xi > -0.5$  in the rest of the paper. In practical data analyses,  $\xi > 0$  is usually encountered.

The main quantity of interest in extreme value analysis is the  $k$ -year mean return level,  $R_k$ , that is the value of  $Y$  that will be exceeded once every  $k$  years. In order to model this quantity we need to resort to a point process approach and use the Poisson-GPD model ([32]). Let  $\mathbf{y} = (y_1, \dots, y_n)$  be  $n$  exceedances over a fixed threshold  $u$  that have been observed during  $T$  years. The corresponding likelihood function for the Poisson-GPD model is

$$L(\theta) = p(\mathbf{y} | \theta) \propto \gamma^n \exp\{-\gamma T\} \times \prod_{i=1}^n f(y_i | \xi, \sigma), \quad (1)$$

where  $\theta = (\xi, \sigma, \gamma)$  and

$$R_k = u + (\sigma/\xi) \left[ (\gamma k)^\xi - 1 \right].$$

Several papers deal with point and interval estimation of  $\theta$  from a frequentist perspective; see e.g.: [15], [33], [18], [25], [23] and [9]. Frequentist approach, in extreme value analysis, pose difficulties not only because of model irregularities, but also because  $n$  cannot be large as extreme values are rare events. Such difficulties induced several authors ([16], [32]) to resort to a Bayesian approach where  $L(\theta)$  is combined with a suitable prior  $\pi(\theta)$  that mitigates the odd behavior of  $L(\theta)$ . Bayesian methods however, introduce the problem of eliciting such prior  $\pi(\theta)$ . In particular, elicitation is problematic because of the lack of a physical meaning of  $\theta$ . In this sense, paper [3] proposes an elicitation procedure on  $\theta$ , while [17] use a separate elicitation for  $\xi < 0$  and  $\xi > 0$ , leading to a prior which is discontinuous at  $\xi = 0$ . The use of vague priors is further complicated by the need of a sensitivity

analysis of the posterior with respect to the prior and this has to be done for each  $\mathbf{y}$ . This is not needed under a default Bayesian perspective, where  $\pi(\theta)$  is determined by a formal rule equal for all possible  $\mathbf{y}$ . The default Bayesian perspective is that used here along with other papers such as [1] that proposes a non informative prior for the regular case of  $\xi > 0$  and [8] that discussed the use of the joint improper Jeffreys' prior for  $(\xi, \sigma)$ ,

$$\pi^J(\xi, \sigma) \propto \sigma^{-1}(1 + \xi)^{-1}(1 + 2\xi)^{-1/2}, \quad \xi > -1/2, \sigma > 0,$$

which leads to a proper posterior distribution [8, Theorem 1]. The Jeffreys' rule is found to produce accurate inference on  $(\xi, \sigma)$  mitigating the odd behavior of  $L(\theta)$ . Finally, a recent work of [24] studies a matching prior for the quantiles of the GPD. However, differently from this work, none of the above papers study matching priors for parameters of the Poisson-GPD model in equation (1).

The main contribution of this paper is to propose matching priors  $\pi^*(\cdot)$ , for the parameters of model (1), based on the approximate conditional (AC) likelihood function in [13]. The prior  $\pi^*(\cdot)$  is derived according to the formal rule developed in [36]. We thus have a formal proof that the matching priors  $\pi^*(\xi)$  and  $\pi^*(\sigma)$  are second order matching priors, while for the joint prior  $\pi^*(\xi, \sigma) \propto \pi^J(\xi, \sigma)$  we have heuristic arguments and numerical evidence suggesting that  $\pi^*(\xi, \sigma)$  is jointly matching, according to the definition in [14], only for  $\xi > 0$ .

The rest of the paper is organized as follows: Section 2 recalls the use of pseudo-likelihoods for deriving matching priors, with particular focus on the AC likelihood. Derivation of  $\pi^*(\cdot)$  along with the estimation of the coverage errors of posterior credible intervals (C.I.) are contained in Section 3. An application to hydrology data is illustrated in Section 4, while Section 5 contains further remarks.

## 2 Matching priors from pseudo-likelihoods under orthogonal parametrization

In this section we resume, under a more general setup, the procedure in [36] that leads to the formal rule here employed in order to obtain  $\pi^*(\cdot)$ . We consider a partition of the parameter  $\theta \in \Theta \subseteq \mathbb{R}^3$ , as  $\theta = (\psi, \lambda)$ , where  $\psi$  represents the parameter of interest and  $\lambda$  is the nuisance parameter. The elimination of  $\lambda$  can be achieved using an appropriate pseudo-likelihood function,  $L^*(\psi)$ , which is a function of  $\psi$  and  $\mathbf{y}$ , with properties similar to those of a genuine likelihood function. Commonly used pseudo-likelihood functions include: marginal, usually employed in Bayesian analysis, conditional, profile and modified profile likelihoods; see, for instance, [2], [28], [31]. The use of  $L^*(\psi)$ , as if it were a true likelihood, leads to the posterior distribution for  $\psi$ ,

$$\pi^*(\psi|\mathbf{y}) \propto \pi(\psi)L^*(\psi),$$

where  $\pi(\psi)$  is a suitable prior distribution on  $\psi$  only. Under a non-orthodox Bayesian point of view, there are two main advantages in computing  $\pi^*(\psi|\mathbf{y})$  in-

stead of the usual marginal posterior

$$\pi(\psi|\mathbf{y}) \propto \int_{\lambda \in \Lambda} L(\theta) \pi(\lambda|\psi) \pi(\psi) d\lambda,$$

based on the full likelihood  $L(\theta)$ . In particular,

- we avoid elicitation of  $\pi(\lambda|\psi)$
- numerical integration on the full parameter space  $\Theta$  is no longer needed.

Bayesian inference based on  $L^*(\psi)$  has been considered, for instance, in [19], [4], [5], [20], [6], [26], [10], [22], [36], [37] [30].

We define a matching prior  $\pi^*(\psi)$  as the prior that satisfies the following first order partial differential equation ([14], Chp. 2)

$$P_\theta(\psi \leq \psi_\alpha) = P_\pi(\psi \leq \psi_\alpha) + O_p(n^{-1}),$$

where  $P_\theta(\cdot)$  denotes the frequentist probability under  $\theta$ ,  $P_\pi(\cdot)$  is the posterior probability and  $\psi_\alpha$  is the  $\alpha$ -quantile of  $\pi^*(\psi|\mathbf{y})$ . If  $\theta$  is a vector of orthogonal parameters (i. e. the Fisher information matrix is strictly diagonal) then  $\pi^*(\psi)$  is available via the approximate conditional likelihood in [13]

$$L_{AC}(\psi) = L_p(\psi) |j_{\lambda\lambda}(\psi, \hat{\lambda}_\psi)|^{-1/2},$$

where  $\hat{\lambda}_\psi$  is the restricted MLE of  $\lambda$  for fixed  $\psi$ ,  $L_p(\psi) = L(\psi, \hat{\lambda}_\psi)$  is the profile likelihood and  $j_{\lambda\lambda}(\psi, \lambda) = -\partial^2 \ell(\psi, \lambda) / \partial \lambda \partial \lambda^\top$  is the observed Fisher information for  $\lambda$ , with  $\ell(\psi, \lambda) = \log L(\psi, \lambda)$ . According to Proposition 2.1 in [36], the matching prior for  $\psi$  associated to  $L_{AC}(\psi)$  is given by

$$\pi^*(\psi) \propto i_{\psi\psi}(\psi, \hat{\lambda}_\psi)^{1/2}, \quad (2)$$

where  $i_{\psi\psi}(\psi, \lambda)$  is the  $(\psi, \psi)$ -component of the Fisher information matrix  $i(\psi, \lambda)$  based on  $L(\psi, \lambda)$ . The corresponding pseudo-posterior is

$$\pi^*(\psi|\mathbf{y}) \propto L_{AC}(\psi) i_{\psi\psi}(\psi, \hat{\lambda}_\psi)^{1/2};$$

see also [35], [34] and [10].

### 3 Matching priors for parameters in the Poisson-GPD model

Let us consider a reparametrization of the GPD density in terms of orthogonal parameters  $\xi$  and  $\nu = \sigma(1 + \xi)$  according to [11]. The GPD density becomes

$$f(y|\xi, \nu) = \begin{cases} \nu^{-1}(1 + \xi)(1 + y\xi(1 + \xi)/\nu)^{-(1+\xi)/\xi}, & (1 + y\xi(1 + \xi)/\nu) > 0 \\ \nu^{-1} \exp(-y/\nu), & \xi = 0 \end{cases}. \quad (3)$$

The corresponding log-likelihood for the Poisson-GPD model, for  $\xi \neq 0$ , is

$$\ell(\theta) = n \log(\gamma) - \gamma T - n \log(v) + n \log(1 + \xi) - \frac{1 + \xi}{\xi} \sum_{i=1}^n \log \left( 1 + \frac{\xi(1 + \xi)}{v} y_i \right),$$

where, with an abuse of notation,  $\theta = (\gamma, \xi, v)$ , and the expected Fisher information matrix, for one observation, is

$$i_{\theta}(\theta) = \begin{pmatrix} \gamma^{-1} & 0 & 0 \\ 0 & (1 + \xi)^{-2} & 0 \\ 0 & 0 & v^{-2}(1 + 2\xi)^{-1} \end{pmatrix}.$$

We consider four relevant different partitions of  $\theta$ : (i)  $\psi = \xi$ , (ii)  $\psi = v$ , (iii)  $\psi = \gamma$  and (iv)  $\psi = (\xi, v)$ . It is worth nothing that the conditional MLE of the Poisson rate,  $\hat{\gamma}_{\psi} = n/T$ , does not depend on  $\xi$  and  $v$ . The corresponding priors (2), for each partition, are all improper and illustrated in Table 1.

**Table 1** Partitions of  $\theta$  along with corresponding priors and posteriors.

| Case  | Partition                           | Prior $\pi^*(\psi)$                                                   | Posterior $\pi^*(\psi \mathbf{y})$ |
|-------|-------------------------------------|-----------------------------------------------------------------------|------------------------------------|
| (i)   | $\psi = \xi, \lambda = (\gamma, v)$ | $\propto (1 + \xi)^{-1}$                                              | numerically                        |
| (ii)  | $\psi = v, \lambda = (\gamma, \xi)$ | $\propto v^{-1}(1 + 2\hat{\xi}_v)^{-1}$                               | numerically                        |
| (iii) | $\psi = \gamma, \lambda = (\xi, v)$ | $\propto \gamma^{-1/2}$                                               | $\text{Gamma}(n + 1/2, 1)$         |
| (iv)  | $\psi = (\xi, v), \lambda = \gamma$ | $\propto \pi^J(\xi, v) \propto v^{-1}(1 + \xi)^{-1}(1 + 2\xi)^{-1/2}$ | see [8]                            |

Only for case (ii) we have a data dependent prior because the conditional MLE  $\hat{\xi}_v$  enters into the prior.

Prior for case (iii) is the Jeffreys/Reference and the first order matching prior for the Poisson rate. Moreover, because of factorization  $L(\theta) = L(\gamma) \times L(\xi, v)$ ,  $\pi^*(\gamma)$  is also the unique first order matching prior for the Poisson-GPD model according to Theorem 2.5.1 in [14].

For case (iv), we have a vector of parameters of interest and the resulting prior  $\pi^*(\psi)$  equals the Jeffreys' prior  $\pi^J(\xi, v)$ . Such prior is the same in [8],  $\pi^J(\xi, \sigma)$ , under the non-orthogonal parametrization. This can be easily seen, by recalling that the Jeffreys' prior is invariant under reparametrization. However, although Jeffreys' and [36] procedures both lead to the same prior, in this paper we only have numerical evidence that  $\pi^*(\xi, v|\mathbf{y})$  is a joint matching prior according to the definition in [14] that is,

$$P_{\theta}(\xi \leq \xi_{\alpha}, v \leq v_{\alpha}) = P_{\pi}(\xi \leq \xi_{\alpha}, v \leq v_{\alpha}) + O_p(n^{-1}).$$

Such evidence is substantial for the regular case of  $\xi > 0$ , while coverages of posterior C.I., for  $\xi < 0$ , are less than the nominal value.

We use Monte Carlo simulations to evaluate frequentist coverages of 95% posterior C.I. for the corresponding priors in Table 1 for cases (i), (ii) and (iv). We use,

for each combination of  $\xi$ ,  $v$  and  $n$ , 5000 simulated C.I.. Cases (i) and (ii) are new to the literature and coverages are shown in Table 2 with significant digits, while coverages for case (iv) are showed in Figure 1.

**Table 2** Empirical coverages for case (i) with  $v = 1$  and (ii) with  $\xi = 1$ .

| $\xi$       | $n$  |      |      |
|-------------|------|------|------|
|             | 10   | 20   | 30   |
| <b>-0.4</b> | 0.94 | 0.96 | 0.95 |
| <b>0.0</b>  | 0.98 | 0.97 | 0.96 |
| <b>1.0</b>  | 0.97 | 0.96 | 0.95 |
| <b>2.0</b>  | 0.97 | 0.97 | 0.96 |
| $v$         |      |      |      |
| <b>0.5</b>  | 0.95 | 0.95 | 0.95 |
| <b>1.0</b>  | 0.94 | 0.95 | 0.94 |
| <b>2.0</b>  | 0.95 | 0.96 | 0.95 |
| <b>5.0</b>  | 0.95 | 0.96 | 0.95 |

We can see from Table 2 and Figure 1 that with  $n = 30$  coverages are all compatible with the nominal value of 95%.

For the mean return level, we propose to use cases (iii) and (iv) to derive the prior for  $R_k(\theta)$  by means of

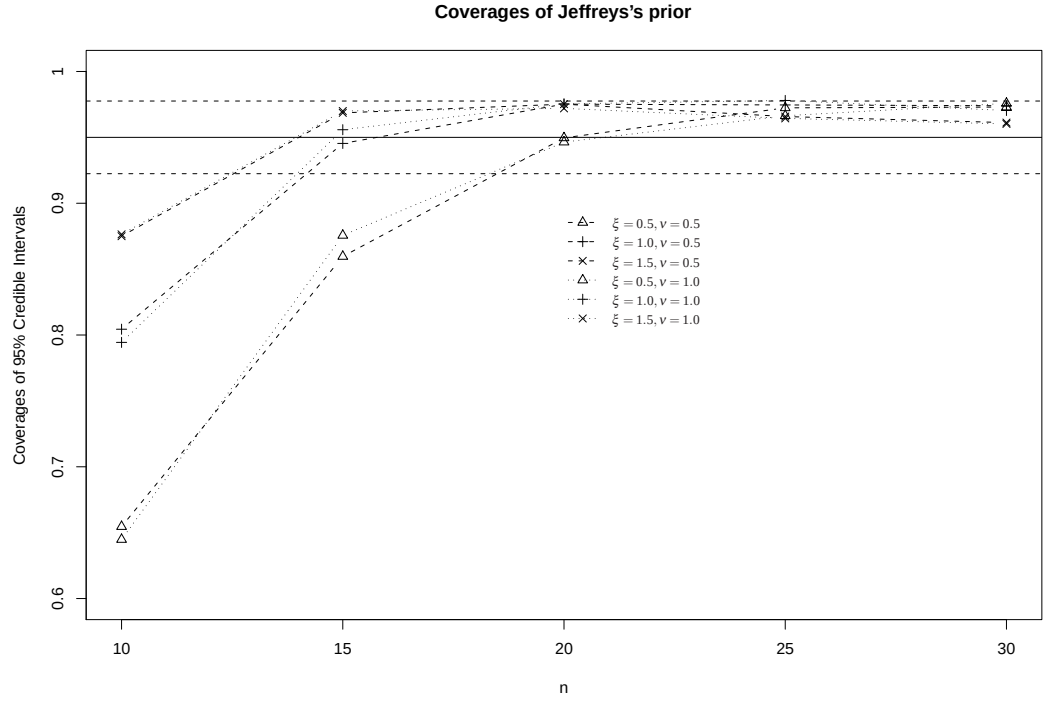
$$\pi^J(\theta) \propto \pi^*(\gamma)\pi^*(\xi, v).$$

Such prior is the Jeffreys' prior for the Poisson-GPD model. The corresponding posterior,  $\pi^J(R_k(\theta)|\mathbf{y})$ , has been also used in [8] and here we study coverages of the corresponding 95% posterior C.I. that are showed in Figure 2 for  $\xi > 0$  and  $k = 1$  year. From  $n = 40$  coverages are compatible with the nominal 95%.

Finally, it is possible to note, from Figures 1 and 2, that model irregularities, occurring when  $\xi \rightarrow -1/2$ , have effect on the coverages of posterior C.I.. In fact, coverages tends to be smaller when  $\xi = 0.5$  with respect to the other values of  $\xi$  considered in the simulation study. However, in real applications, negative values of  $\xi$  are not very common, while large positive values of  $\xi$  denote the presence of heavy tails in the distribution of  $Y$ .

## 4 An application to hydrology

We apply the proposed inference procedure to a sample of 154 exceedances over the level  $65 \text{ m}^3/\text{sec.}$  by the River Nidd at Hunsingore Weir from 1934 to 1969 ([27]). This data set (see Figure 3) has been analyzed by several authors: [16], [25], [8]. These data involve problems of trend, seasonality and threshold selection. Here we use them to demonstrate our procedure. One of the main feature of this sample is



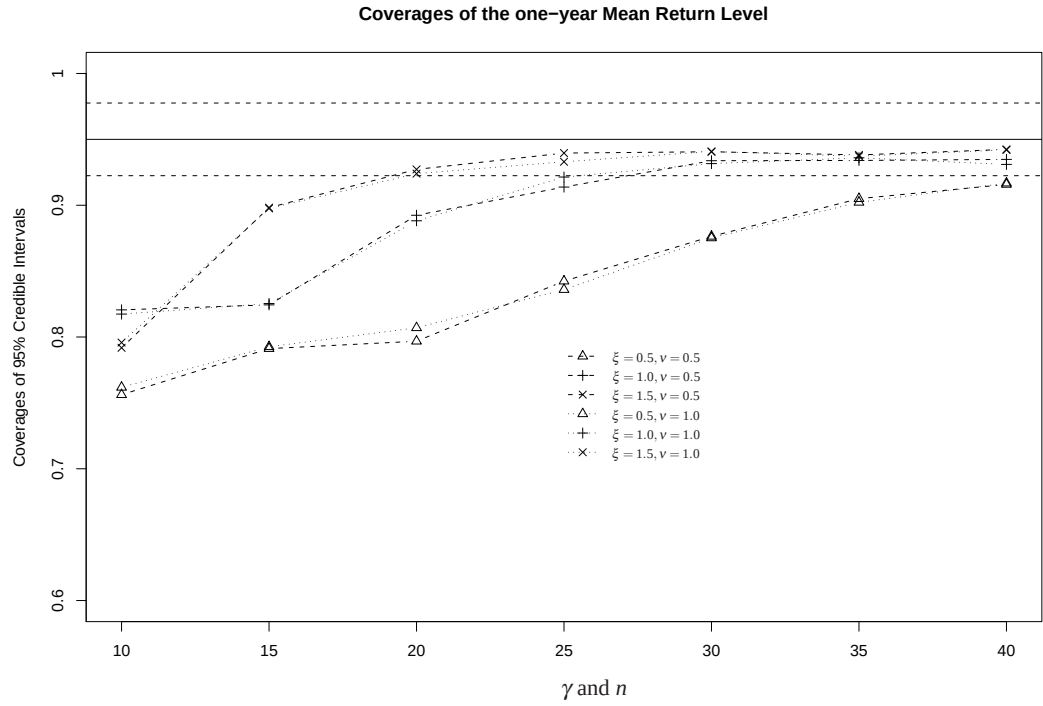
**Fig. 1** Empirical coverages of  $\pi^*(\xi, \sigma|\mathbf{y})$ , for the regular case of  $\xi > 0$ . Horizontal lines indicate coverages compatible with the nominal 95%.

the tail bumpiness for thresholds between 100 and 120  $m^3/sec.$ , because the number of observations over the threshold  $u$  does not smoothly decreases as  $u$  increases. This feature complicates the threshold selection problem, which is not explicitly addressed here. For different thresholds, the C.I. for  $\xi$  and  $\nu$  are showed in Table 3, while C.I. for  $R_k$  are the same intervals in [8]. The length of  $\xi$  and  $\nu$  C.I. increases with  $u$  because of the sample size. Inference is sensible to  $u$  because values of  $u$  up to 100  $m^3/sec.$  provide similar inference (i.e. all C.I. contains  $\xi = 0$ ) and this differs from that obtained with  $u > 100$ . For this data, it could be of interest whether or not the exceedances of the Nidd River are bounded. This amounts to calculate

$$BF^* = \pi^*(\xi > 0|\mathbf{y})/\pi^*(\xi < 0|\mathbf{y}),$$

that is the pseudo-Bayes Factor on the sign of  $\xi$ . Differently from the usual  $BF$ , based on  $L(\theta)$ , as that in [8],  $BF^*$  is not based on the posterior marginal distribution of  $\xi$ , but instead on pseudo-likelihood  $L^*(\xi)$ . The general use of  $L^*(\psi)$  to obtain  $BF^*$  is further discussed in [7].

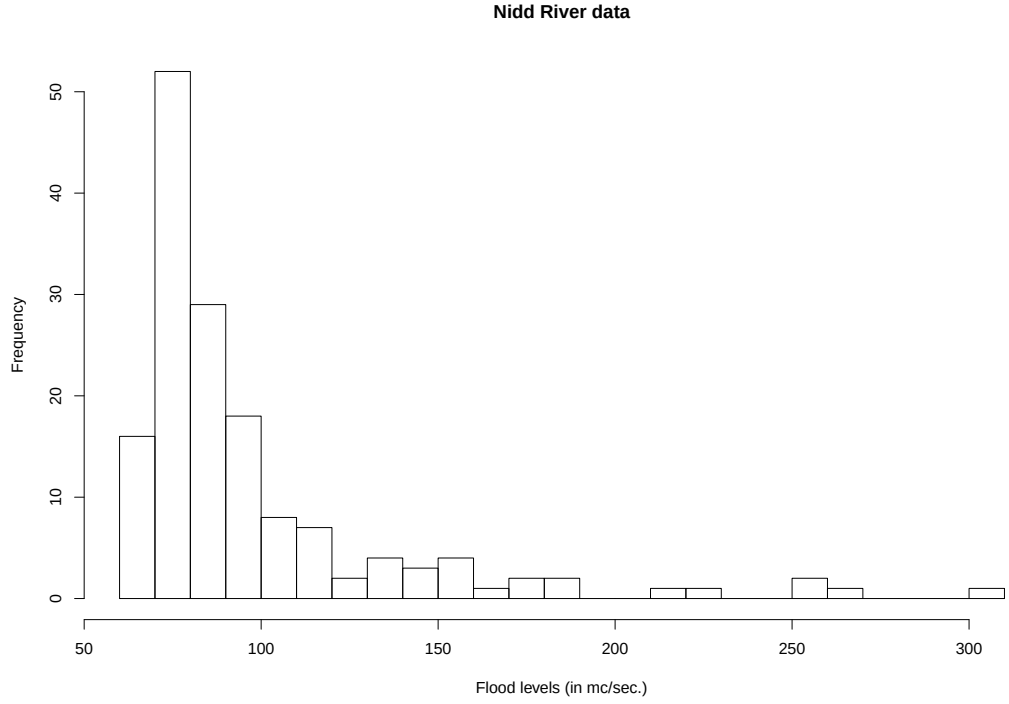
The weight of evidence,  $\log BF^*$ , is showed in Table 3. For  $u < 110$ , the evidence in favour of  $\xi > 0$  is substantial and it is larger than the one obtained in [8].



**Fig. 2** Empirical coverages of  $\pi^J(R_k(\theta)|\mathbf{y})$ , for the regular case of  $\xi > 0$ . Horizontal lines indicate coverages compatible with the nominal 95%.

**Table 3** 95% credible intervals for  $\xi$  and  $\nu$  at different thresholds for Nidd data.

| $u$ | $n$ | 95% C.I.     | $\xi$  | $\log BF^*$ | 95% C.I. | $\nu$  |
|-----|-----|--------------|--------|-------------|----------|--------|
|     |     |              | Length |             |          | Length |
| 70  | 138 | (0.14,0.60)  | 0.46   | 9           | (23,36)  | 13     |
| 80  | 86  | (0.09,0.76)  | 0.68   | 6           | (26,45)  | 20     |
| 90  | 57  | (-0.07,0.79) | 0.86   | 3           | (30,58)  | 29     |
| 100 | 39  | (-0.30,0.65) | 0.95   | 1           | (37,75)  | 38     |
| 110 | 31  | (-0.37,0.71) | 1.08   | 0           | (37,81)  | 44     |
| 120 | 24  | (-0.44,0.58) | 1.02   | -1          | (40,89)  | 49     |
| 130 | 22  | (-0.42,0.87) | 1.29   | 0           | (35,88)  | 54     |
| 140 | 18  | (-0.44,0.97) | 1.41   | 0           | (34,95)  | 61     |



**Fig. 3** Flood levels above  $65 \text{ m}^3/\text{sec.}$  of River Nidd at Hunsingore Weir from 1934 to 1969.

## 5 Concluding Remarks

For the Poisson-GPD model (1) and a reduced set of parameters of interest, we derived priors for inference based on the AC likelihood. Some of these priors are also first order matching priors and some of them coincide with Jeffreys' priors. The main objective of this work is to show the potential of the approach based on pseudo-likelihoods to derive matching priors [36]. In fact, for the Poisson-GPD model, matching priors are derived quite easily.

Finally, one may employ priors  $\pi^*(\psi)$  here developed and obtain the posterior  $\pi(\psi|\mathbf{y})$  with the usual integration of the marginal likelihood

$$L^m(\psi) \propto \int_{\lambda \in \Lambda} L(\theta) \pi(\lambda|\psi) d\lambda,$$

such that the posterior is

$$\pi(\psi|\mathbf{y}) \propto L^m(\psi) \pi^*(\psi).$$

However, the procedure here proposed avoids elicitation  $\pi(\lambda|\psi)$  that can be particularly cumbersome for the parameters of the Poisson-GPD model, moreover in-

tegration over  $\lambda$  is avoid in our analysis. Finally, the matching argument holds for  $\pi^*(\psi|\mathbf{y})$ , while the marginal posterior  $\pi(\psi|\mathbf{y})$  it is no longer guaranteed to be matching.

**Acknowledgements** This work has been partially supported by M.I.U.R. of Italy.

## References

1. Arnold, B., Press, S.J.: Bayesian estimation and prediction for Pareto data. *J. Amer. Statist. Assoc.* **84**, 1079-1084 (1989).
2. Barndorff-Nielsen O., Cox D.: *Inference and Asymptotics*. Chapman and Hall/CRC Press, London (1994)
3. Behrens C.N., Lopes H.F., Gamerman D.: Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling* **4**, 227-244 (2004)
4. Bertolino F., Racugno W.: Analysis of the linear correlation coefficient using pseudo-likelihoods. *Journal of the Italian Statistical Society* **1**, 33-50 (1992)
5. Bertolino F., Racugno W.: Robust bayesian analysis of variance and the  $\chi^2$ -test by using marginal likelihoods. *The Statistician* **43**, 191-201 (1994)
6. Cabras S., Racugno W., Ventura L.: Bayesian inference on the scalar skew-normal distribution. In, 17th Compstat Symposium of the IASC (Roma-Italy, 28/8-1/9) (ISBN, 3-7908-1708-2), 1373-1381 (2006)
7. Cabras S., Racugno W., Ventura V.: Pseudo-bayes factors. In, 19th Compstat Symposium of the IASC (Paris-France, 22/8-27/8) (to appear) (2010)
8. Castellanos M., Cabras S.: A default bayesian procedure for the generalized pareto distribution. *Journal of Statistical Planning and Inference*, **137**, 473-483 (2007)
9. Castillo E., Hadi A.: Fitting the generalized pareto distribution to data. *Journal of the American Statistical Association* **92**, 1609-1620 (1997)
10. Chang H., Mukerjee R.: Probability matching property of adjusted likelihoods. *Statistics and Probability Letters* **76**, 838-842 (2006)
11. Chavez-Demoulin V., Davison A.: Generalized additive modelling of sample extremes. *Applied Statistics* **54**, 207-222 (2005)
12. Coles S.: *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, NY (2001)
13. Cox D., Reid N.: Parameter orthogonality and approximate conditional inference (with discussion). *J. Roy. Stat. Soc. B*, **49**, 1-39 (1987)
14. Datta G.S., Mukerjee R.: *Probability Matching Priors, Higher Order Asymptotics*, Lecture Notes in Statistics, **178**. Spriner-Verlag (2004)
15. Davison A.: *Statistical extremes and Applications*, Reidel, Dordrecht, Chp. Modelling excesses over high thresholds, with an application, 461-482 (1984)
16. Davison A.C., Smith R.L.: Models for exceedance over high threshold. *J. Roy. Stat. Soc. B* **52**, 393-442 (1990)
17. de Zea Bermudez P., Turkman M.A.: Bayesian approach to parameter estimation of the generalized pareto distribution. *Test* **12**, 259-277 (2003)
18. DuMouchel W.: Estimating the stable index  $\alpha$  in order to measure the thickness, a critique. *The Annals of Statistics* **11**, 1019-1036 (1983)
19. Efron B.: Bayes and likelihood calculations from confidence intervals. *Biometrika* **80**, 3-26 (1993)
20. Fraser D.A.S., Reid N., Wong A., Yun Yi G.: Direct bayes for interest parameters. In Valencia **7**, 529-533 (2003)
21. Gnedenko B.: Sur la distribution limite du terme maximum d'une serie aleatoire. *Ann Mathematics* **44**, 423-453 (1943)

22. Greco L., Racugno W., Ventura L.: Robust likelihood functions in bayesian inference. *Journal of Statistical Planning and Inference* **138**, 1258-1270 (2008)
23. Grimshaw S. D.: Computing maximum likelihood estimates for the generalized pareto distribution. *Technometrics* **35**, 185-191 (1993)
24. Ho K.W.: A matching prior for extreme quantile estimation of the generalized pareto distribution. *Journal of Statistical Planning and Inference* **140**, 1513-1518 (2010)
25. Hosking J., Wallis J.: Parameter and quantile estimation for the generalized pareto distribution. *Technometrics* **29**, 339-349 (1987)
26. Lin L.: Quasi bayesian likelihood. *Statistical Methodology* **3**, 444-455 (2006)
27. Natural Environment Research Council: Flood Studies Report. Natural Environment Research Council (1975)
28. Pace L., Salvan A.: *Principles of Statistical Inference*. World Scientific, Singapore (1997)
29. Pickands J.: Statistical inference using extreme order statistics. *Annals of Statistics*, **3**, 119-131 (1975)
30. Racugno, W., Salvan, A., Ventura, L.: Bayesian analysis in regression models using pseudo-likelihoods, *Communications in Statistics - Theory and Methods*, to appear (2010)
31. Severini T.A.: *Likelihood Methods in Statistics*. Oxford University Press (2000)
32. Smith R.: *Extreme values in finance, telecommunications and the environment*, Chapman and Hall/CRC Press (2003)
33. Smith R.L.: *Statistical extremes and Applications*, Reidel, Dordrecht (1984)
34. Sweeting, T.J.: On the implementation of local probability matching priors for interest parameters, *Biometrika*, **92**, 47-57 (2005)
35. Tibshirani, R.: Noninformative priors for one parameter of many, *Biometrika*, **76**, 604-608 (1989)
36. Ventura L., Cabras S., Racugno W.: Prior distributions from pseudo-likelihoods in the presence of nuisance parameters. *Journal of the American Statistical Association* **104**, 768-774 (2009)
37. Ventura L., Cabras S., Racugno W.: Default prior distributions from quasi- and quasi-profile likelihoods. *Journal of Statistical Planning and Inference* doi:10.1016/j.jspi.2010.04.003 (2010)