The effect of scale in daily precipitation hazard assessment

J. J. Egozcue, V. Pawlowsky-Glahn, M. I. Ortego, R. Tolosana-Delgado

Nat. Hazards Earth Syst. Sci., 6, 459–470, 2006
www.nat-hazards-earth-syst-sci.net/6/459/2006/

Natural Hazards and Earth System Sciences

The effect of scale in daily precipitation hazard assessment

J. J. Egozcue[1], **V. Pawlowsky-Glahn**[2], **M. I. Ortego**[1], **and R. Tolosana-Delgado**[2]

[1]Dept. de Matemàtica Aplicada III, Universitat Politècnica de Catalunya, Barcelona, Spain
[2]Dept. d'Informàtica i Matemàtica Aplicada, Universitat de Girona, Girona, Spain

Received: 23 September 2005 – Revised: 10 February 2006 – Accepted: 6 March 2006 – Published: 6 June 2006

Abstract. Daily precipitation is recorded as the total amount of water collected by a rain-gauge in 24 h. Events are modelled as a Poisson process and the 24 h precipitation by a Generalised Pareto Distribution (GPD) of excesses. Hazard assessment is complete when estimates of the Poisson rate and the distribution parameters, together with a measure of their uncertainty, are obtained. The shape parameter of the GPD determines the support of the variable: Weibull domain of attraction (DA) corresponds to finite support variables as should be for natural phenomena. However, Fréchet DA has been reported for daily precipitation, which implies an infinite support and a heavy-tailed distribution. Bayesian techniques are used to estimate the parameters. The approach is illustrated with precipitation data from the Eastern coast of the Iberian Peninsula affected by severe convective precipitation. The estimated GPD is mainly in the Fréchet DA, something incompatible with the common sense assumption of that precipitation is a bounded phenomenon. The bounded character of precipitation is then taken as a priori hypothesis. Consistency of this hypothesis with the data is checked in two cases: using the raw-data (in mm) and using log-transformed data. As expected, a Bayesian model checking clearly rejects the model in the raw-data case. However, log-transformed data seem to be consistent with the model. This fact may be due to the adequacy of the log-scale to represent positive measurements for which differences are better relative than absolute.

1 Introduction

The goal of hazard assessment is to estimate the probability of occurrence of large events in a given lifetime. Hazardous events due to natural or anthropogenic phenomena (precipitation, earthquakes, wind, eruptions, floods, fires, etc.) are often modelled by marked Poisson processes: events are assumed to occur as a point Poisson process in time, and magnitude of events is assumed to be random, independent from the time-occurrence process and from event to event.

This simple model may be useful in situations when one is interested in rare but dangerous events. However, the scarcity of data leads to highly uncertain parameter estimates, a problem which can be overcome using Bayesian estimation (Gelman et al., 1995; Coles, 2001) to account for uncertainty. A standard model for large magnitude events is the Generalised Pareto Distribution (GPD) (Embrechts et al., 1997), leading to a global model with four parameters: the rate of the Poisson process; the scale and shape for the GPD; and a reference threshold. The reference threshold is assessed empirically and afterwards validated. This assessment is a key point of the analysis because a trade-off must be made between a high threshold, guaranteeing a better model fit, and the number of available data with magnitude over it. The other three parameters are considered jointly distributed, and are estimated using Bayesian techniques. Prior information is obtained from expert opinions or physical knowledge.

This approach was used by Egozcue and Ramis (2001) to analyse precipitation in Eastern Spain using a database covering 30 years (Romero et al., 1998). Heavy precipitation is a serious weather hazard in the Valencia region, especially in autumn. Every year several events exceeding 100 mm daily precipitation occur. Strong convective systems are responsible for it, and precipitation tends to discharge over short periods. For example, in Gandía, on 3 November 1987, more than 800 mm were recorded in 24 h. Some of these events produce floods and severe damage to properties, infrastructure and agriculture, like the one that destroyed the Tous dam (Valencia) on 20 October 1982. The main problem in the study performed by Egozcue and Ramis (2001) appeared to be that excesses exhibit a heavy, unbounded, upper tail, something contradictory with the naturally bounded

Correspondence to: J. J. Egozcue
(juan.jose.egozcue@upc.edu)

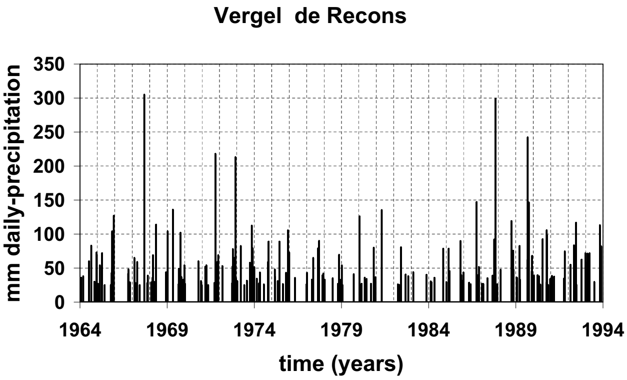

Fig. 1. Observed daily precipitation at Vergel de Racons (Alicante, Spain), 1964–1993.

character of precipitation. Precipitation in 24 h must be limited due to several physical reasons: water content in the atmosphere is limited, the movement of convective cells is limited, and also the time of precipitation. These facts have been neglected by many authors for the sake of simplicity, because actual physical upper limit is not known (e.g. Coles and Tawn, 1996; Egozcue and Ramis, 2001; DeMichele and Salvadori, 2003).

While analysing the underlaying reasons for heavy tail behaviour of precipitation, one can realise that 10 mm and 20 mm precipitations are one double of the other, whereas 500 mm and 510 mm are both large rainfalls and may be considered similar. Only few observers would characterise these differences using the ordinary differences of 10 mm. Most people would prefer a relative measure of difference, i.e. the first case corresponds $100{\cdot}(20-10)/10{=}100\%$ increment of the first rainfall, while in the second case, it is only $100{\cdot}(510-500)/500{=}2\%$. The comparison of 0 mm precipitation and 10 mm is more dramatic: the first one corresponds to “no rainfall” or “no event”; the second one is definitively “rainfall” and, therefore, the difference is no longer quantitative but qualitative. These appreciations are mainly subjective and may depend on the particular application but, at least, put a question mark on the use of the ordinary distance to quantify differences between rainfall events. A logarithmic transformation of rainfall data gives a way to handle these problems of scale: the relative scale is transformed into the ordinary one and positive data become just real (positive or negative). This kind of discussion on the natural scale of positive data is general (Pawlowsky-Glahn and Egozcue, 2001) and logarithmic transformation of positive data may be recommended in many applications.

The present goal is to compare the behavior of a precipitation sample in its original scale (raw-data) and log-transformed data ($\log_{10}$-data) using the above described hazard model and Bayesian estimation of parameters. The considered model is based on the following hypotheses: a) daily precipitation is bounded in the upper tail and the uppermost value is assumed unknown; b) large events occur in time as a Poisson process; c) precipitation in 24 h is random and distributed GPD over a suitable threshold. Moreover, it is independent from event to event and the time occurrence.

Hypothesis a), being a common sense one, appears to be quite inconsistent with the data when the analysis is performed using raw-data. However, $\log_{10}$-data do not show this inconsistency. The whole model is checked using Bayesian p-values for both data sets.

2 Precipitation data

To study the effect of scale transformation, a single rain-gauge is selected from the data base by Romero et al. (1998). This data base contains records of daily precipitation along 30 years (1964–1993) for 72 rain-gauges in Valencia region (Eastern Iberian Peninsula). The selected rain-gauge is located at Vergel de Racons (Alicante; 38.842° N, 0.012° E). Only daily precipitation over 25 mm is considered, and the maximum daily precipitation in a moving window of 7 days is extracted. In order to obtain independent events, consecutive maxima are forced to be separated more than three days; whenever two or more are closer than three days, only the largest daily precipitation is considered. The obtained set of 192 events is represented in Fig. 1. They range from 25 mm up to the maximum observed, namely 305 mm daily precipitation. Along this study the original precipitation values are used (raw-data) as well as the $\log_{10}$-transformed data ($\log_{10}$-data).

3 Hazard model

The Poisson point-process model used for natural event occurrences is assumed stationary, something contradictory with the seasonal character of precipitation, or even with some long term trend. However, if attention is restricted to high magnitude events, the yearly periodic component can be neglected because these events occur with return periods much longer than one year. Moreover, long term trends like those possibly induced by a climate change are hardly detected along a 30 years record. Hence, results of the analysis using stationary models should be considered carefully for predictive uses. Precipitation events exceeding a threshold u are modelled as points in time. For each event, a magnitude or size X (daily precipitation in this case) is usually taken as a random variable. The assumptions used are: (a) magnitudes are independent from the point process itself; (b) they are independent from event to event; (c) all magnitudes have the same distribution F_X; and (d) the occurrence of the events in time follows a Poisson process. Thus, the number of events, $N(u)$, occurring in a given arbitrary time interval, t, is governed by the Poisson probability ($n{=}0, 1, 2, \ldots$)

$$\mathrm{P}[N(u) = n|\lambda(u), t] = \frac{1}{n!}(\lambda(u)t)^n e^{-\lambda(u)t}, \quad (1)$$

where $\lambda(u)$ stands for the rate of the Poisson process given the threshold u. This model is sometimes known as the Cramér-Lundberg model.

The event magnitude, X, is modelled only in the upper tail of its distribution using the peak-over-threshold method (Embrechts et al., 1997). The excess over the threshold u is defined as $Y=\{X-u|X>u\}$ and uses the relationship $(y=x-u>0)$

$$1 - F_Y(y) = \mathrm{P}[Y > y|X > u] = \frac{1 - F_X(x)}{1 - F_X(u)} , \tag{2}$$

where F_Y is the excess distribution. Equation (2) links the distributions of X and Y.

Hazard parameters can be obtained from the model. For instance, the return period of events whose magnitude X exceeds x is

$$\tau(x) = \frac{\tau(u)}{1 - F_Y(x-u)} , \ u \le x , \tag{3}$$

where $\tau(u)=1/\lambda(u)$. Also, non-exceedance probabilities of the threshold u in a lifetime L are

$$\mathrm{P}[N(x) = 0|L] = \exp[-\lambda(x)L] , \tag{4}$$

where $\lambda(x)=1/\tau(x)$.

When F_Y is specified, the hazard model is complete for excess values $y>u$. The option here selected is the Generalised Pareto Distribution (GPD). GPD is a simple and parsimonious model for excesses, as it is the limiting distribution for excesses whenever u is high enough (Pickands, 1975) and it is general enough as it includes both finite and infinite support distributions. GPD is given by $(\beta>0,\ y>0)$

$$F_Y(y|\xi,\beta) = 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-1/\xi} , \tag{5}$$

with ξ and β the shape and scale parameters. The support of Y is the positive real line $\mathbb{R}^+$ for $\xi\ge 0$, while it is bounded in the interval $[0, -\beta/\xi]$ for $\xi<0$. For $\xi=0$, Eq. (5) takes the exponential form

$$F_Y(y|\xi = 0,\beta) = 1 - \exp\left(-\frac{y}{\beta}\right) . \tag{6}$$

Asymptotically, GPD approaches the upper tail of continuous distributions. According to the type of the upper tail, maxima extracted from such distributions correspond to different extreme value distributions, namely Gumbel, Weibull and Fréchet, frequently merged in the so-called generalised extreme value distribution (Embrechts et al., 1997). GPD's are classified into three domains of attraction (DA) referred to the kind of maxima they generate. Distributions with exponentially decaying upper tails belong to the Gumbel DA and are approached by a GPD with $\xi=0$ (exponential). Bounded upper tail distributions correspond to the Weibull DA and are approached by a GPD with $\xi<0$. The Fréchet DA contains distributions with heavy upper tails corresponding to GPD's with $\xi>0$. GPD's in the Weibull and Gumbel domain have both mean and variance, but this does not apply to Fréchet DA. If $\xi\ge 1$, the mean does not exist, and variance is not defined for $\xi\ge 0.5$, thus remarking the heavy tail behavior of GPD's in the Fréchet DA.

Natural phenomena are physically bounded; thus, their magnitude should be in the Weibull DA. However, heavy tail distributions have been reported, in particular for intense precipitation in different climates (Coles and Tawn, 1996; Egozcue and Ramis, 2001). Data in Fig. 1, reported in raw scale, behave in this way, as is shown later. This fact is usually considered to be due to lack of data, but the reason might be simply an inappropriate scale. In the present case study, a logarithmic scale reveals a clear Weibull DA for precipitation events.

This is not unimportant, as one should be aware that, from a theoretical point of view, a log-transformation of a GPD random variable cannot transform an unbounded support into a limited one. In fact, log-transformation of a distribution in the Gumbel and Fréchet DA's are transformed into distributions in the Gumbel DA.

4 Bayesian estimation of parameters

The selected hazard model has four parameters: the absolute threshold u; the rate of the Poisson process of events which magnitude is larger than u, $\lambda(u)$; and the shape and scale parameters of the GPD, ξ and β, for excesses over u. Due to scale reasons, we prefer to re-define the Poisson parameter as $z=z(u)=\log_{10}\lambda(u)=-\log_{10}\tau(u)$. The four parameters must be estimated from the data set to evaluate hazard parameters.

The absolute threshold u is selected according to criteria discussed in Sect. 5 and, here, we assume that it has been already chosen. The remaining three parameters, z, ξ and β, are estimated using Bayesian methods developed in Egozcue and Ramis (2001) and Egozcue and Tolosana-Delgado (2002). According to the Bayesian paradigm, these parameters are assumed to be random variables. Their joint probability densities, $f_{z\xi\beta}(z,\xi,\beta)$ and $f_{z\xi\beta}(z,\xi,\beta|D)$, account for their uncertainty before (prior) and after (posterior) using the data sample, symbolised by D.

The prior density represents our knowledge about parameters previous to D. The key assumption is that $z(u)$ is independent from (ξ,β), i.e. $f_{z\xi\beta}(z,\xi,\beta)=f_z(z)\cdot f_{\xi\beta}(\xi,\beta)$. Section 6 describes the assessment of the prior in the present case study.

The core of Bayesian estimation is to use Bayes' theorem to obtain the posterior from the prior and the data. Bayes' theorem, under the stated assumption about the prior, states that

$$f_{z\xi\beta}(z,\xi,\beta|D) = C\ L(z,\xi,\beta|D) \cdot f_z(z) \cdot f_{\xi\beta}(\xi,\beta) , \tag{7}$$

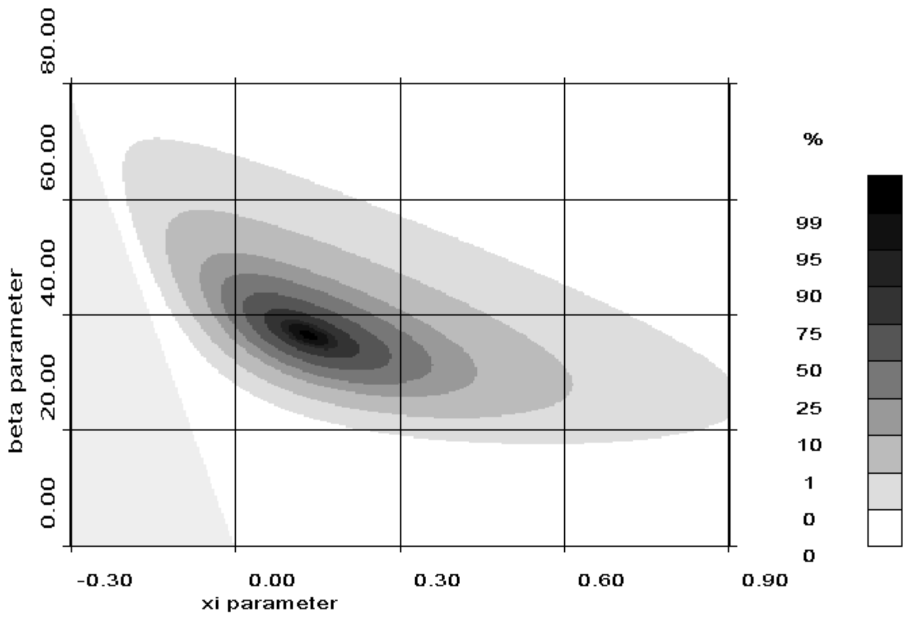

Fig. 2. Posterior density of ξ, β for raw-data. u=45 mm. Flat prior. High probability of Fréchet DA. Lower-Left grey triangle is a null-likelihood zone due to observations.

where C is a normalising constant and $L(z, \xi, \beta|D)$ is the likelihood of the data. The likelihood can also be factorised,

$$L(z, \xi, \beta|D) = \mathrm{P}[N(u) = n|z, t_0] \cdot \prod_{j=1}^{m} f_Y(y_j|\xi, \beta) , \quad (8)$$

where D has been made explicit as the number $N(u)=n$ of excesses $y_j=x_j-u$ over u, the excesses themselves, and the observation time t_0.

The posterior density in Eq. (7) is itself the result of the Bayesian estimation, but it is also the basis to obtain the distribution of other hazard parameters, like return periods, Eq. (3), or non-exceedance probabilities, Eq. (4). When it is assumed that the estimated GPD is in the Weibull DA with probability 1, i.e. the magnitude X is surely limited by $y_{\sup}=-\beta/\xi$, two additional hazard parameters can be considered: the maximum attainable magnitude, $x_{\sup}=u+y_{\sup}$, and the probability of a magnitude x to be attainable, i.e. $\mathrm{P}[x<u+y_{\sup}|D]$.

These hazard parameters are easily approached from the posterior density. Since z, ξ, and β, are random and jointly distributed as $f_{z\xi\beta}(z, \xi, \beta|D)$, hazard parameters are also random and can be described by their probability densities. A simulated sample of z, ξ, and β, generates a sample of each desired hazard parameter. From this derived sample, central tendency parameters, e.g. the median, provide point estimates, and the sample quantiles determine credible intervals. These type of estimates of hazard parameters (return periods, exceedance probabilities, $y_{\sup}$) will be used in Sect. 7.

As an example of Bayesian estimation, Fig. 2 shows the estimation of ξ and β for raw-data using an absolute threshold u=45 mm. The prior has been set to a uniform distribution over a large enough square of (ξ, β)-values. Therefore, the posterior is, up to a constant, the likelihood of the excesses. Figure 2 represents the posterior density in the plane ξ, β. Each point in the plane (ξ, β) corresponds to a GPD, and the value of $f_{\xi\beta}(\xi, \beta|D)$ to the relative likelihood of the parameters. An important feature is that $\xi\geq 0$ (Fréchet DA) contains the most likely points for the parameters and, consequently, Fréchet DA is clearly more probable than Weibull DA, thus confirming the Fréchet DA behaviour of the raw-data. In contrast, Fig. 3 shows the posterior for $\log_{10}$-data. Again the prior was uniform in a large enough domain, so that the posterior is essentially the normalised likelihood. In this case, Weibull DA is largely more probable than Fréchet DA, thus remarking the opposite behaviour of raw-data and $\log_{10}$-data.

The mode of (ξ, β) is easily located in Figs. 2 and 3, they correspond to the maximum likelihood estimates of the parameters. For raw-data, they are $\hat{\xi}_{ML}$=0.14 (Fréchet DA), $\hat{\beta}_{ML}$=37; for $\log_{10}$-data, the results are $\hat{\xi}_{ML}=-0.24$ (Weibull DA), $\hat{\beta}_{ML}$=0.29. The sample distribution and the estimated GPD can be represented in a QQ-plot, Fig. 4 for raw-data and Fig. 5 for $\log_{10}$-data. In both figures the reference is an exponential distribution (GPD-Gumbel DA) fitted to the data (diagonal line). In Fig. 5 quantile, given in $\log_{10}$ scale have been translated back into raw scale in mm to make the comparison easier. Sample distributions are represented only by markers (filled for upper corners; hollow for lower corners). In both figures the maximum likelihood estimated GPD's are also shown (full lines) together with a band

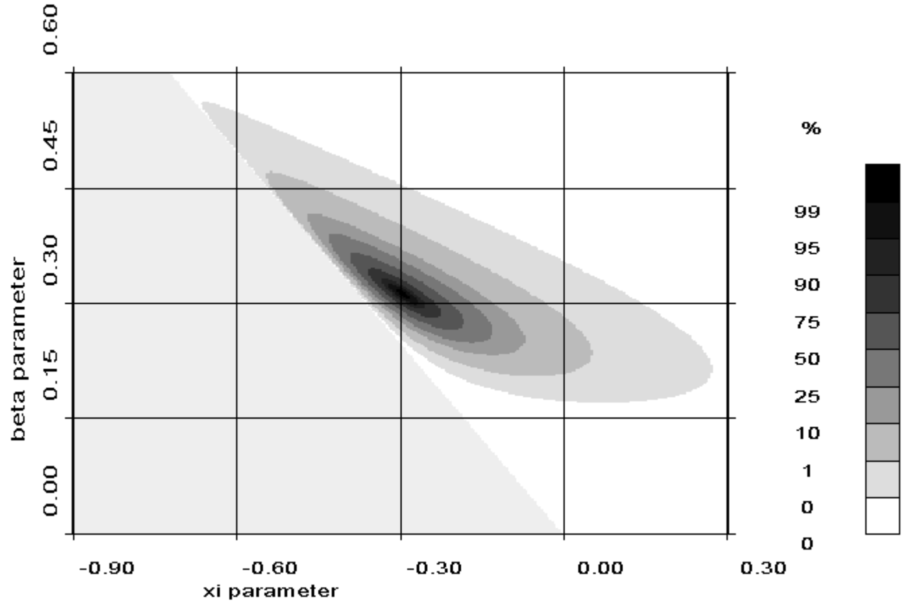

Fig. 3. Posterior density of ξ, β for $\log_{10}$-data. $u=1.653=\log_{10}(45)$. Flat prior. High probability of Weibull DA. Lower-left grey triangle is a null-likelihood zone due to observations.

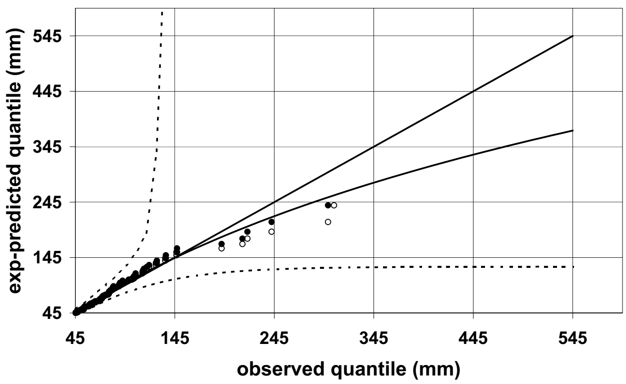

Fig. 4. QQ-plot with GPD($\xi=0$) as reference (diagonal line). Sample distribution of raw-data: filled and hollow markers correspond to upper and lower corners of the stepwise function; maximum likelihood estimated GPD, full line; dotted lines: 5% acceptance region for Kolmogorov-Smirnov test.

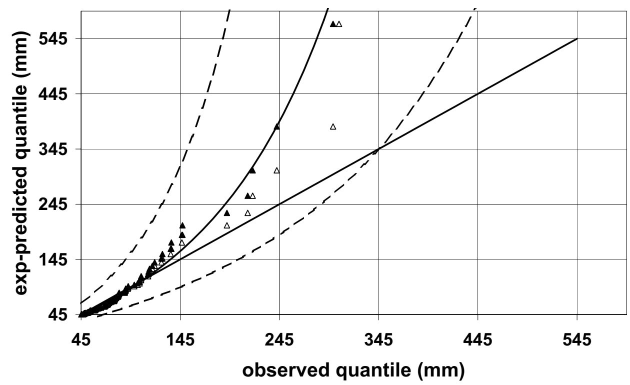

Fig. 5. QQ-plot with GPD($\xi=0$) as reference (diagonal line). Sample distribution of $\log_{10}$-data: filled and hollow markers correspond to upper and lower corners of the stepwise function; maximum likelihood estimated GPD, full line; dotted lines: 5% acceptance region for Kolmogorov-Smirnov test.

(dotted or dashed lines) that represents the 5% acceptance region for the Kolmogorov-Smirnov goodness-of-fit test. The different behavior of the data and fitted GPD with respect to the exponential confirms the Fréchet DA behavior of raw-data (Fig. 4) and the Weibull DA in the case of $\log_{10}$-data (Fig. 5). In both cases, the estimated GPD could not be rejected at 5% significance level.

5 Absolute threshold selection

The first step in the estimation is the selection of an appropriate reference threshold u. A graphical technique (Embrechts et al., 1997) can be applied attending to the fact that, for GPD distributed excesses and for $u' \geq u$, the mean excess is linear with respect to u' ($\xi < 1$),

$$\mathrm{E}[X-u'|X>u'\geq u]=\frac{\beta+\xi u'}{1-\xi}\,. \tag{9}$$

Inspection of the mean excess function was performed for the raw data and $\log_{10}$-data. Figures 6 and 7 show sample

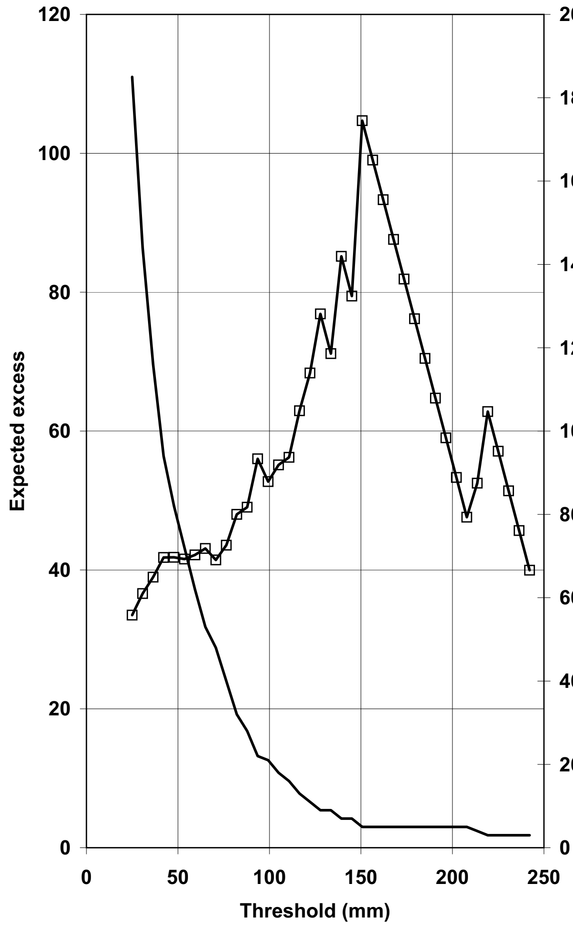

Fig. 6. Mean excess function for raw-data, squares. Number of excesses, no marker, secondary axis.

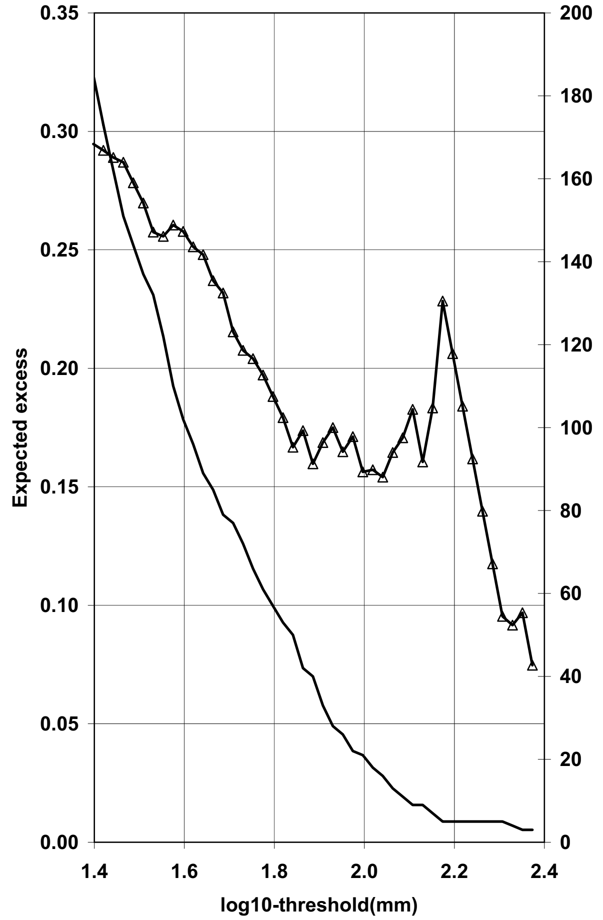

Fig. 7. Mean excess function log-data, triangles. Number of excesses, no marker, secondary axis.

estimates of the mean excess function for both data samples. The lines with markers correspond to the sample average excess over each threshold. The number of excesses used to compute the mean excess is superimposed to indicate the accuracy of the estimated mean. From Eq. (9), positive slopes of the mean excess function indicate that the data set corresponds to a Fréchet DA distribution (heavy and unlimited tails), whereas negative slopes suggest data from GPD's of the Weibull DA (limited support). In order to guess an absolute threshold, we look for the smallest threshold for which the mean excess function can be assumed to be linear from this point on, according to Eq. (9). In general, a tradeoff between a good fit and the number of excesses should be done to retain a good deal of data in the subsequent estimation. From the expected excess functions we have guessed u=45 mm for raw-data, and u=1.85 for $\log_{10}$-data, which corresponds to 63 mm. However, lower absolute thresholds seem to be also acceptable for $\log_{10}$-data.

This kind of selection of absolute threshold is mainly subjective, and a statistical validation is convenient. A way to do this is to check the goodness-of-fit of excesses to the GPD, but this requires the previous estimation of the GPD. The procedure is the following: Fix an absolute threshold as a candidate and obtain the marginal posterior of ξ and β, $f_{\xi\beta}(\xi,\beta|D)$; for each (ξ,β) on a grid, compute the p-value of the Kolmogorov-Smirnov (KS) goodness-of-fit test $\alpha(\xi,\beta)$; then, approximate the predictive KS-p-value,

$$\alpha_{\text{pred}} = \int \alpha(\xi,\beta)\ f_{\xi\beta}(\xi,\beta|D)\ d\xi\ d\beta\ , \qquad (10)$$

by a weighted average. Figure 8 shows the values of $\alpha(\xi,\beta)$ for $\log_{10}$-data assuming an absolute threshold u=1.653, equivalent to 45 mm. The p-value has been computed only for points in which the posterior is not negligible. Figure 9 shows the results of α_{pred} for both raw-data and $\log_{10}$-data using different absolute thresholds. These results suggest that, with the exception of thresholds near 25 mm for raw-data, there are no evidences for rejection of the GPD model. For further comparisons between raw-data and $\log_{10}$-data, thresholds of u=45 mm for raw-data and

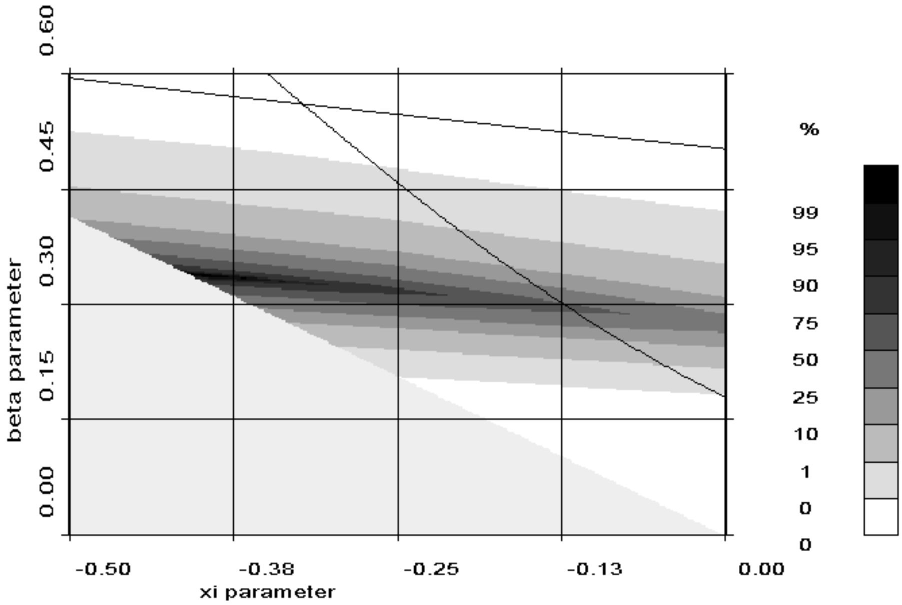

Fig. 8. Significance of KS goodness-of-fit test, $\alpha(\xi, \beta)$, for $\log_{10}$-data. u=1.653 mm.

$u=\log_{10}(45)=1.653$ have been finally chosen. This decision is supported by a KS-predictive p-value of about 0.4.

6 Prior knowledge implementation

Prior density for z has been assumed uniform for a very wide range of values corresponding to a non-informative prior. We have renounced to an informative prior because we assume there are enough data (88 excesses over u=45 mm) to estimate z quite accurately, and any soft information would be superseded by the likelihood of the data. Prior density for GPD parameters has been assessed following the methods developed in Egozcue and Ramis (2001) and Egozcue and Tolosana-Delgado (2002).

Prior information is mainly used to give bounds of the admissible domain for the GPD-parameters ξ and β. Our prior hypothesis are referred to the Mediterranean area. They are ordered by decreasing influence in the estimation:

A. Daily precipitation is a bounded phenomenon.

B. A daily precipitation of 2000 mm is almost impossible. Precipitation events larger than 45 mm attain 2000 mm with probability less than 10^{-4} (almost impossible events).

C. Precipitation events larger than 45 mm exceed 100 mm with probability less than 0.5 (upper probability of characteristic events).

D. A precipitation of 400 mm must be possible, although it may not be observed (surely attainable event).

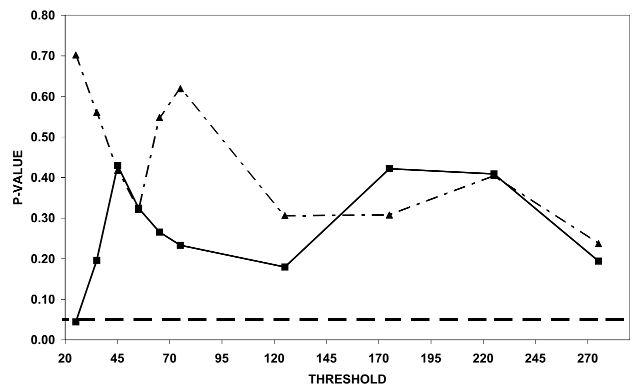

Fig. 9. Predictive p-values of the KS-goodness-of-fit test. Raw-data, full line and squares; $\log_{10}$-data, dotted-dashed line and triangles. Critical value of 0.05, dashed line.

E. Frequencies of events are decreasing with precipitation; the decreasing is, at least, linear.

F. A typical event with precipitation larger than 45 mm is 100 mm, which occurs approximately with probability 0.1 (characteristic event).

The most important assumption is A. It essentially implies that we only consider GPD's in the Weibull DA ($\xi<0$) for both raw-data and $\log_{10}$-data. Figures 10 and 11 show the corresponding prior densities for ξ and β. Assumption A is reflected in the upper bound of ξ. Assumption B (almost impossible events) is physically based. In Fig. 11 it appears as a border cutting the upper-right corner of the represented domain, so excluding all GPD's not fulfilling B.

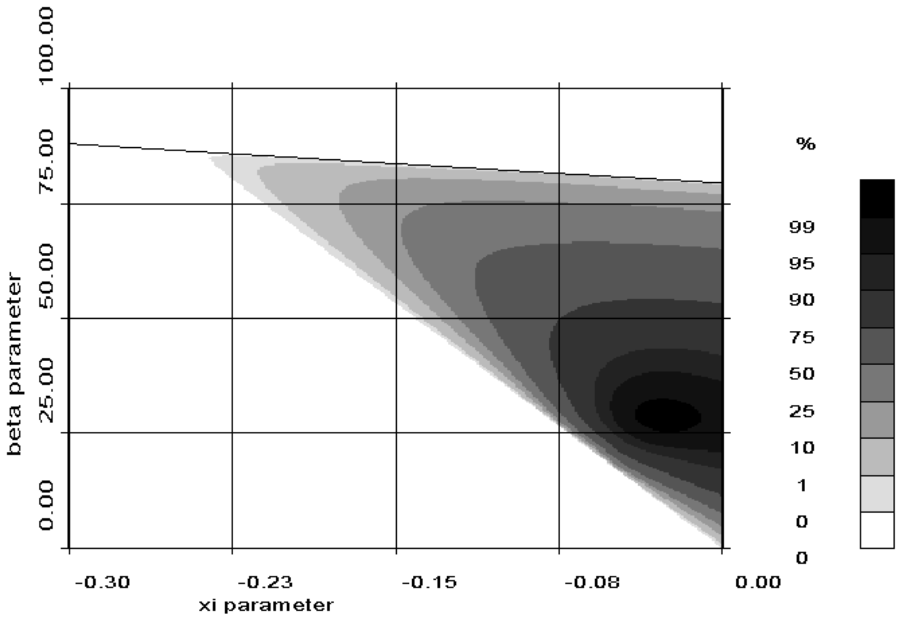

Fig. 10. Prior density for GPD parameters, ξ, β, for raw-data. u=45 mm.

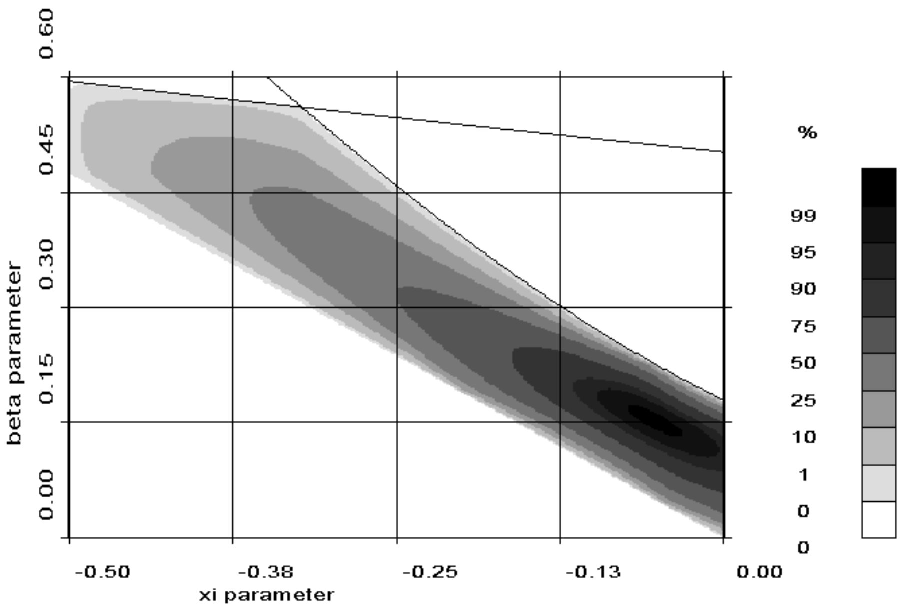

Fig. 11. Prior density for GPD parameters, ξ, β, for $\log_{10}$-data. u=1.653.

This condition does not appear in the represented domain in Fig. 10. It could be verified at whichever rain-gauge in the Mediterranean zone. Assumption C is a weak condition based on the behaviour of convective precipitation in the Eastern Mediterranean sea; it can be observed regularly in this region. It produces a similar effect to condition B; the corresponding border appears in both figures as an upper limit of the prior support, thus excluding too large β values. Assumption D is quite subjective, although, in general, experience in the Mediterranean area clearly supports it and there is no physical reason to assume that 400 mm is not attainable. It excludes GPD's in the lower-left triangle with a vertex in (ξ=0, β=0); for these distributions precipitation of 400 is not attainable. For raw-data, assumption E means that $-0.5<\xi$, because $\xi=-0.5$ corresponds to a triangular GPD. For $\log_{10}$-data we apply the same rule. However, the statement E does not match exactly $-0.5<\xi$ in the log-scale. Figure 10 does not show the value $\xi=-0.5$ because conditions C

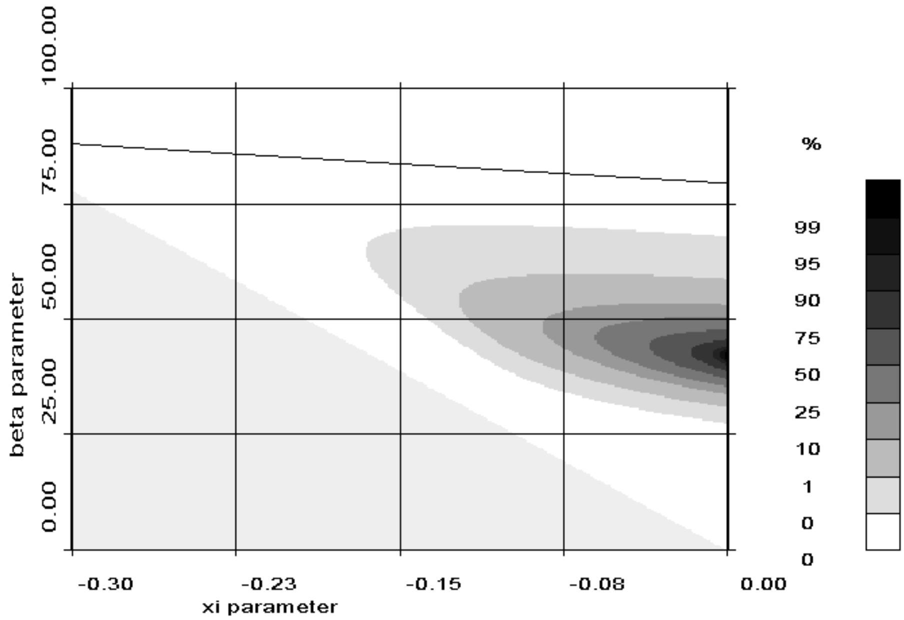

Fig. 12. Posterior density for GPD parameters, ξ, β, for raw-data. u=45 mm.

and D imply E. Both priors have been forced to be flat inside the domain defined by conditions A through E, but decreasing to zero at the borders. Finally, assumption F roughly defines a characteristic event that is used to place the mode of both priors; its influence is weak. Figures 10 and 11 show contours of the resulting priors for raw-data and $\log_{10}$-data respectively.

Although the stated assumptions are weak and reasonable, they collate with the information from raw-data. Figure 2 shows that most of the posterior probability is placed in the Fréchet DA ($\xi>0$) when the prior does not constrain it to the Weibull DA ($\xi<0$). However, according to our main hypothesis, i.e. daily precipitation is a bounded phenomenon, the prior for raw-data (Fig. 10) is restricted to the Weibull DA. This contradiction between observed raw-data and prior assumptions does not hold for $\log_{10}$-data because their likelihood is mainly located in the Weibull DA, as shown in Fig. 3.

7 Daily precipitation hazard estimation

For both raw-data and $\log_{10}$-data, hazard parameters are estimated using the priors in Figs. 10 and 11, in agreement with the assumption that daily precipitation is a bounded phenomenon. The obtained posteriors are shown in Figs. 12 and 13, respectively. Simulated samples of (z, ξ, β) are obtained according to their joint posterior distribution, and characteristics of the posterior distribution of hazard parameters are then estimated.

A typical hazard parameter, introduced in Eq. (3), is the return period of events which magnitude exceed a given threshold. Figure 14 shows the results. At a first sight, one confirms the large uncertainty of the estimates, visualised by the $(0.05, 0.95)$ credible interval for both raw-data and $\log_{10}$-data (thin lines-hollow markers). Other important differences between the two samples are also revealed. Median return periods (thick lines-filled markers) for raw-data are clearly larger than for $\log_{10}$-data. For instance, for the 300 mm daily precipitation, 190 years using raw-data are obtained, and 29 years using $\log_{10}$-data, a difference of one order magnitude. This is not negligible, even taking into account the large uncertainty of the results. To understand the importance intuitively, one should take into account that one event of more than 300 mm has been recorded in the available sample running along 30 years. This fact predisposes us to accept the results based on $\log_{10}$-data more easily than those based on raw data.

Differences between the two samples appear as well when comparing exceedance probabilities of a given level of precipitation in 50 years. Exceedance probabilities are complementary to non-exceedance probabilities, introduced in Eq. (4). Results for both samples are shown in Fig. 15. For instance, for the level of 400 mm, raw-data gives a median exceedance probability of 0.015 in 50 years, whereas for $\log_{10}$-data this probability is 0.37.

In a particular case study, it may be difficult to believe that certain precipitation levels may be attainable. Therefore, raising the question whether they are actually attainable or not makes sense. The estimation of the probability of one particular value being attainable is possible because we are assuming bounded distributions for precipitation (Weibull DA). Figure 15 shows these probabilities. They are very important from the engineering point of view, because civil

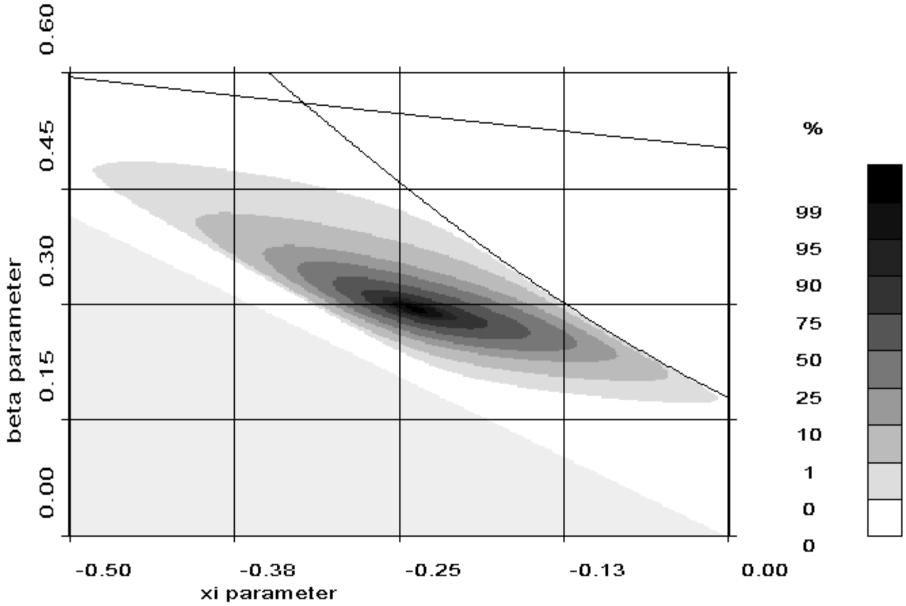

Fig. 13. Posterior density for GPD parameters, ξ, β, for $\log_{10}$-data. u=1.653.

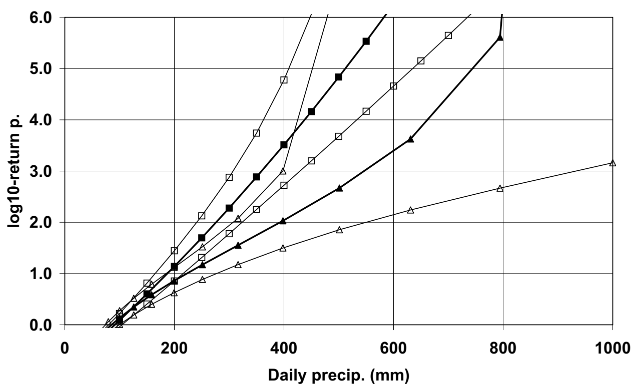

Fig. 14. Estimated $\log_{10}$-return periods, $\log_{10}\tau(x)$. Squares, raw-data; triangles, $\log_{10}$-data. Thick lines-filled markers, median of posterior distribution; thin lines-hollow markers, 0.05 and 0.95 quantiles.

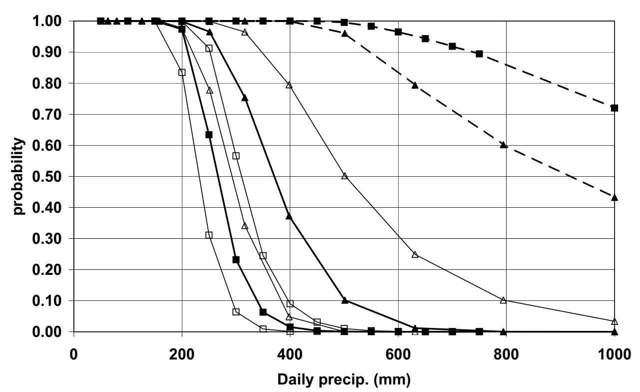

Fig. 15. Exceedance probabilities in 50 years. Squares, raw-data; triangles, $\log_{10}$-data. Thick lines-filled markers, median of posterior distribution; thin lines-hollow markers, 0.05 and 0.95 quantiles. Probability of attainable precipitation, dashed lines.

works do not need to be designed for non attainable actions. Consider for example precipitation of 700 mm, which has a probability of 0.92 of being attainable using raw-data, and a lower one, 0.71, using $\log_{10}$-data. Looking at this and higher values for precipitation, one recognizes the conservative character of conclusions drawn from raw-data, compared to those derived from $\log_{10}$-data. In the same Fig. 15, it is easy to see that raw-data give overall lower probabilities of exceedance in 50 years compared to probabilities given by $\log_{10}$-data. Thus, while raw-data are conservative concerning the attainable character of a precipitation level, $\log_{10}$-data are so for probabilities of exceedance in 50 years.

An estimation of the upper-limit of precipitation is also possible, although it is very uncertain and the corresponding distribution is very asymmetric. The mode of the posterior distribution of this upper limit, $-\beta/\xi$, was 788 mm for raw-data and 759 mm for $\log_{10}$-data. However, these values correspond to low quantiles: probability of the upper-limit being larger than 788 mm is about 0.87 for raw-data; for $\log_{10}$-data, the probability of an upper-limit greater than 759 mm is approximately 0.65 thus showing the asymmetry of these estimates. In this case the difference between the two samples is not so large.

Two main points derive from these results: uncertainty of the estimates is large and the results from raw-data and

$\log_{10}$-data differ substantially. The first one is a consequence of the short record (30 years) we are using to estimate occurrence of long-term large events. The second one confronts the analyst with the choice of the appropriate model for the data: which scale gives more sensible results? Reasons supporting the $\log_{10}$ scale option have been exposed in the previous sections. A quantitative study on adequacy of the model for both samples follows.

8 Bayesian Model checking

In order to quantify consistency of a model in a Bayesian framework, model checking techniques have been developed recently (Meng, 1994; Gelman et al., 1996). They are based on the comparison of the observed data with replicated data obtained using the estimated model. We centre our attention on the GPD-model for excesses. It includes the (ξ, β)-prior, the corresponding posterior, and the GPD itself.

To replicate data from the estimated model the steps are: (a) Simulate GPD parameters, $\xi^{(j)}, \beta^{(j)}$, according to the obtained posterior distribution; (b) simulate a n-sample, $\boldsymbol{y}_{\mathrm{rep}}^{(j)}$, from a GPD with parameters $\xi^{(j)}, \beta^{(j)}$. This can be repeated a large enough number of times. Each replicated n-sample, $\boldsymbol{y}_{\mathrm{rep}}^{(j)}$, is compared to the original n-sample of excesses, $\boldsymbol{y}_{\mathrm{obs}}$, using a function, $T{=}T(\boldsymbol{y}, \xi, \beta)$, known as discrepancy function. T is intended to express the features one wants to check in the model and can be defined ad hoc. The actual comparison is performed using a Bayesian p-value,

$$\alpha_p = \mathrm{P}[T(\boldsymbol{y}_{\mathrm{rep}}, \xi, \beta) \leq T(\boldsymbol{y}_{\mathrm{obs}}, \xi, \beta)] \,, \qquad (11)$$

where ξ and β are distributed according to their joint posterior, and the sample $\boldsymbol{y}_{\mathrm{rep}}$ as a GPD conditional to the parameters ξ and β. In a consistent model, the observed data should appear (through the discrepancy function) mixed with replicated samples and not separated from them. From this point of view, p-values close to 0 or to 1 indicate some inconsistency in the model. Clearly, the kind of inconsistency checked depends on the discrepancy function. Once the replicated samples have been obtained, α_p is easily estimated as the proportion of times in which $T(\boldsymbol{y}_{\mathrm{rep}}^{(j)}, \xi^{(j)}, \beta^{(j)}){\leq}T(\boldsymbol{y}_{\mathrm{obs}}, \xi^{(j)}, \beta^{(j)})$.

We are mainly interested in checking the compatibility of observed data with the GPD model of excesses, and our prior assumptions concerning the domain of attraction of the data-sample. Therefore, the value of ξ corresponding to a particular data-set is a suitable explanatory function. An estimator of ξ given the sample of excesses is then a reasonable choice for a discrepancy function. Rough, but intuitively effective estimators of ξ, $(\xi{<}1)$, and β are $\hat{\xi}$ and $\hat{\beta}$ minimising

$$\sum_{k=1}^{m} n(u_k)\,(e_k - (a + bu_k))^2 \ , \ a = \frac{\beta}{1-\xi} \ , \ b = \frac{\xi}{1-\xi} \ , (12)$$

where u_i, $u_i \geq u$, are pre-defined thresholds for excesses, $n(u_k)$ is the number of excesses in the sample exceeding u_k,

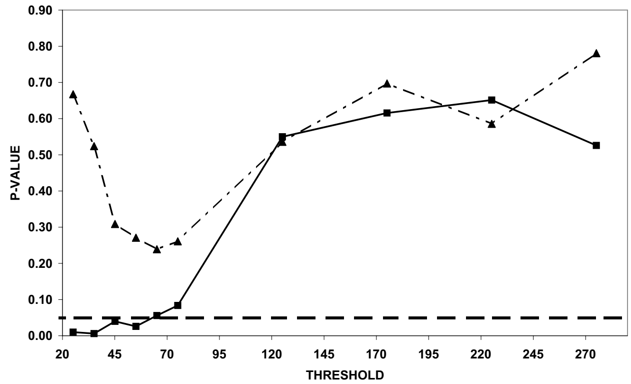

Fig. 16. Bayesian p-values. Raw-data: full line, squares. $\log_{10}$-data: dashed-dotted line, triangles. 0.05 reference level: dashed line.

and e_k is the average excess over u_k. This is equivalent to fitting a regression line to the estimated mean excesses as a function of the threshold, as suggested by Eq. (9). The regression is actually weighted by the number of excesses used to estimate e_k.

For each excess sample, observed or replicated, we define $T{=}\hat{\xi}$. We realise that ξ and β do not appear explicitly in T but, when calculated for a replicated sample, $\xi^{(j)}$ and $\beta^{(j)}$ play a role through the GPD distribution. We expect this discrepancy function to be sensible to situations in which prior and likelihood disagree.

Figure 16 shows the p-values of Eq. (11) for a number of different absolute thresholds and for both raw-data and $\log_{10}$-data. For raw-data and thresholds under 80 mm, the model is suspicious of inconsistencies. Recall that the chosen threshold was 45 mm both for raw-data and for $\log_{10}$-data. For higher thresholds, the model becomes acceptable, although the number of available data decreases to less than 30 excesses (Fig. 6) and uncertainty increases accordingly. Contrarily, for $\log_{10}$-data, the model checking is satisfactory for all used thresholds. These results quantitatively confirm that raw-data are hardly compatible with the assumption that excesses have a bounded distribution.

9 Conclusions

Estimation of hazard parameters is highly uncertain mainly due to lack of data. Using a Bayesian approach, this uncertainty can be monitored as shown in the study of a 30-year precipitation series obtained from a rain-gauge at Vergel de Racons (Alicante). Events can be modelled as a Poisson process, and the upper tail of daily precipitation as Generalised Pareto distributed. The final model is validated using a Bayesian model checking. Two considerations about precipitation lead the approach: (a) differences in precipitation are relative, calling for a log-scale, and (b) precipitation is finite.

The present case study shows that a log-transformation, which satisfies the first condition, naturally leads to a model which also satisfies the second one. In fact, excesses of the $\log_{10}$ transformed data can be assumed to follow a Generalised Pareto distribution in the Weibull domain of attraction (bounded upper tail), in agreement with the finite character of precipitation. Results for untransformed data showed that a Generalised Pareto distribution in the Weibull domain of attraction is not compatible with the upper tail of the observations. This result stands in agreement with a preliminary study which leads to some distribution in the Fréchet domain of attraction (heavy, unbounded upper tail), a behaviour previously reported for other precipitation data sets.

These results support the idea that, for strictly positive data, not only for precipitation, for which a relative scale appears to be natural, taking log's might be not unimportant. When using a natural scale, some modelling requirements, such as goodness-of-fit or compatibility with physical assumptions, appear to be fulfilled in an easier way than when using the usual raw scale in real space.

Acknowledgements. The authors thank O. Serrano (Puertos del Estado, Spain) for his encouragement and advice. Data come from the Instituto Nacional de Meteorología of Spain. BGPE program and this study were supported under agreement of the Universidad Politécnica de Catalunya and Puertos del Estado (Spain). This research has also received financial support from the Dirección General de Investigación of the Spanish Ministry for Science and Technology through the project BFM2003-05640/MATE.

Edited by: L. Ferraris
Reviewed by: two referees

References

Coles, S. G.: An Introduction to Statistical Modeling of Extreme Values, Springer, London, GB, 208 p., 2001.

Coles, S. G. and Tawn, J. A.: A Bayesian analysis of extreme rainfall data, Applied Statistics, 45, 463–478, 1996.

DeMichele, C. and Salvadori, G.: A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas, J. Geophys. Res., 108(D2), 4067, doi:10.1029/2002JD002534, 2003.

Egozcue, J. J. and Ramis, C.: Bayesian Hazard Analysis of Heavy Precipitation in Eastern Spain, Int. J. Climatol., 21, 1263–1279, 2001.

Egozcue, J. J. and Tolosana-Delgado, R.: Program BGPE: Bayesian Generalized Pareto Estimation, CD-ROM, ISBN 84-69999125, Barcelona, Spain, 2002.

Embrechts, P., Klüppelberg, C., and Mikosch, T.: Modelling extremal values, Springer Verlag, Berlin, 1997.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B.: Bayesian data analysis, Wiley, New York, NY,USA, 1995.

Gelman, A., Meng, X. L., and Stern, H.: Posterior predictive assessment of model fitness via realized discrepancies (with discussion), Statistica Sinica, 6, 733–807, 1996.

Meng, X. L.: Posterior predictive p-values, Ann. Stat., 22, 1142–1160, 1994.

Pawlowsky-Glahn, V. and Egozcue, J. J.: Geometric approach to statistical analysis on the simplex, Stochastic Environmental Research and Risk Assessment, 15, 384–398, 2001.

Pickands, J.: Statistical inference using extreme order statistics, Annals of Statistics, 3, 119–131, 1975.

Romero, R., Guijarro, J. A., Ramis, C., and Alonso, S.: A 30-years (1964-93) daily rainfall data base for the Spanish Mediterranean regions: First exploratory study, Int. J. Climatol., 18, 541–560, 1998.