Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Hydrology 283 (2003) 1–18
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling

Eric Parent*, Jacques Bernier

GRESE (Laboratory for Risk Analysis in Water Sciences), ENGREF, 19, Avenue du Maine, F-75732 Paris Cedex 15, France

Received 13 September 2001; accepted 20 February 2003

Abstract

One of the main decisions to be made in operational hydrology is to estimate design floods for safety purposes. These floods are generally much rare events that have already been systematically recorded and consequently the results of any estimation process are subject to high levels of uncertainty. When adopting the frequentist framework of probability, the so called 'respect of scientific objectivity' shall forbid the hydrologists to introduce prior knowledge such as quantified hydrological expertise into the analysis. However, such an expertise can significantly improve the capability of a probabilistic model to extrapolate extreme value events. The Bayesian paradigm offers coherent tools to quantify the prior knowledge of experts. This paper develops an inference procedure for the peak over threshold (POT) model, using semi-conjugate informative priors. Such prior structures are convenient to encode a wide variety of prior expertise. They avoid recourse to Monte Carlo Markov Chain techniques which are presently the standard for Bayesian analyses, but such algorithms may be uneasy to implement. We show that prior expertise can significantly reduce uncertainty on design values.

Using the Garonne case study with a sample of systematic data spanning over the period 1913–1977, we point out that: (1) the elicitation approach for subjective prior information can be based on quantities with a definite practical hydrological meaning for the expert; (2) with respect to the usual Poisson–Generalized Pareto model, a semi-conjugate prior offers a flexible structure to assess expert knowledge about extreme behavior of the river flows. In addition, it leads to quasi-analytical formulations; (3) tractable algorithms can be implemented to approximate the prior uncertainty about POT parameters into these semi conjugate distribution forms via simple Monte Carlo simulations and normal approximations; (4) the design value and its credible interval are notably changed when incorporating prior knowledge into the risk analysis.

Keywords: Bayesian analysis; Monte Carlo methods; Peak over threshold model; Extreme value theory; Flood design; Elicitation of prior expertise

* Corresponding author.
E-mail addresses: parent@engref.fr (E. Parent); jacques.bernier2@wanadoo.fr (J. Bernier).

doi:10.1016/S0022-1694(03)00080-5

1. Prior knowledge is worth incorporating into extreme value analysis

For estimating probabilities of rare events, the usual way is to extrapolate distributions from common ordinary events to low probability events as qualitatively shown in Fig. 1. Such an extrapolation process is necessarily subject to many uncertainties: the target zone concerns the shape of the tail of the distribution model (i.e. probability with order of magnitude 10^{-3} to 10^{-4}) and can be reached only by extrapolation, encompassing therefore large model uncertainties as exemplified by Fig. 1: three distributions, not distinguishable by fitting them to ordinary events, may exhibit very different tail behavior. Particularly in assessing probability distribution functions of extremes in hydrology, many different models have been proposed without any clear and completely convincing justifications (see Bobée, 1999, for a review).

In addition to modeling uncertainties generated by the scientist's inability to assume correct or even realistic hypotheses and model, many other uncertainties are involved in the estimation of probabilities of extreme events and can interfere with decisional processes:

- *natural physical randomness* of phenomena under study (natural uncertainty),
- *measurement uncertainties* of data which often coexist with the sparsity of observations,
- *sampling uncertainties on parameters*, due the limited information available.

A large body of literature has been devoted to the study of these kinds of uncertainties in hydrology. Most of these studies were conducted mainly according the principles and criteria of classical statistics emphasizing unbiasedness and mean square errors (MSE) of percentiles or other parameters estimates. A list of classical arguments against the use of statistical extrapolation methods of extreme values distributions can be found in Coles and Powell (1996). Of course

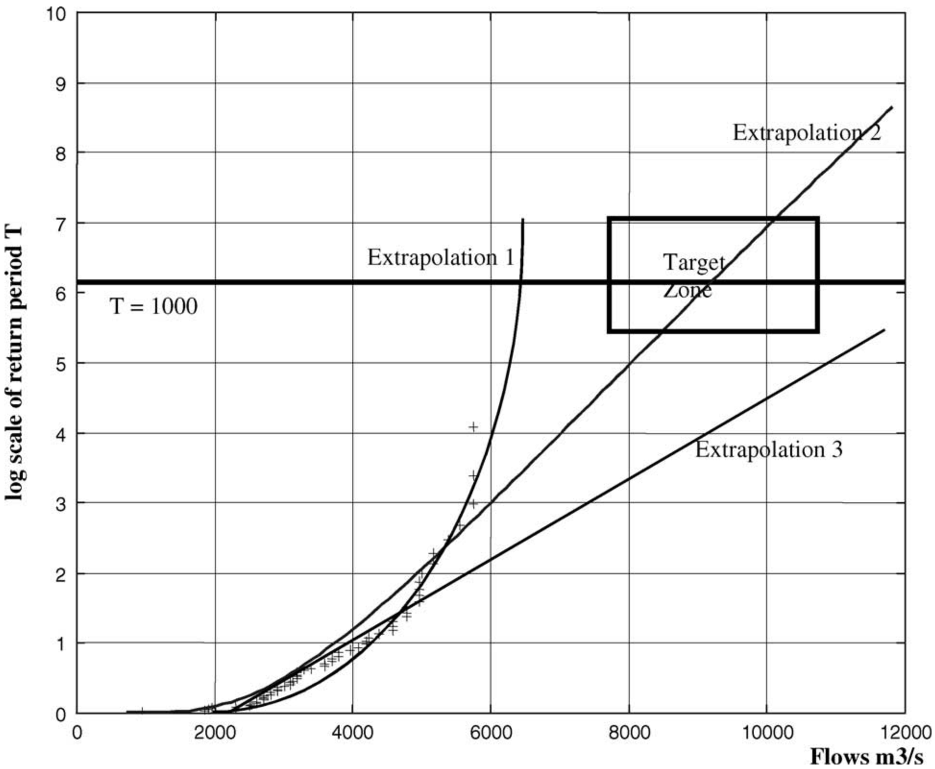

Fig. 1. Frequency curve extrapolation with a sample of 50 events.

the frequency concept supporting the interpretation of probability fails to be reasonably justified in the case of such rare events. The subjective concept of probability, as emphasized by the Bayesian paradigm (Bernier, 1967, 1987) does not suffer from this limitation because its background consists of decisional bets (Savage, 1972). The results of random draws from conceptual urns are shown to be equivalent to such decisional bets and the composition of the urn can be fitted to the conditions of occurrence of any rare events. However virtual this urn may be with regards to the phenomenon under study, its operational interpretation can be easily cast into the context of engineering decision making and needs no recourse to a stationary property on a large time scale. Conversely such a poor argument of stationarity is usually invoked for the frequentist interpretation of the probability concerning geophysical rare events.

1.1. Any source of information is worth bringing into the analysis to reduce uncertainty

To obtain sufficiently reliable estimates, the scientist must take all these uncertainties into account and try to control and reduce their effects on subsequent decisions. The rational way to achieve this reduction of uncertainty is to take advantage of available information of any kind. For extreme events, the systematic data series used are usually very short and do not allow reliable estimates of risks caused by rare events. Fortunately these series, generally considered site by site, are not the only useful information:

Further systematic information. The most common method for evaluating the risks associated with a random intensity of extremes in a given site, is to analyze the selected series of annual maximum values (AMS). But this selection, advocated for the respect of some independence assumptions, unduly reduces significant information. Peak over threshold (POT) methods on the contrary select all events above a threshold, and assumes independence and identical distribution of these high flows values. A thorough introduction to the Bayesian approach combined with the Poisson/exponential extreme value model can be found in Rasmussen and Rosbjerg (1991) and was generalized to the Poisson/Pareto extreme value model by Rosbjerg et al. (1992). The POT models have theoretical advantages (Pickands, 1975) such as asymptotic coherence insuring rational generalized Pareto parameters or percentile estimation whatever the values of thresholds may be. Furthermore this coherence usually reduces estimation errors. We shall use this kind of model in the following.

Regional information. No local geophysical event (hydometeorological or not) can be separated from its regional environment. The generating phenomena of these events (such as storms) generally act at a larger space scale. Furthermore similarities between sites can allow transposition of information from site to site. Describing such similarities is the purpose of popular regional models (GREHYS, 1996) such as the index flood method based on a frequentist approach. The so called ‘empirical Bayes methods’ were used by many authors following Madsen and Rosbjerg (1997). From a more complete Bayesian point of view, hierarchical models such as the ones given by Gelman et al. (1995) afford an interesting and reliable alternative and have been applied by Tawn (1993) to spatial modeling of extreme sea levels and by Madsen et al. (1995) to the regional analysis of extreme rainfalls.

Historical information. In countries inhabited for a long time, historical data are often available. Even if these data are sparse and often imprecise they can give valuable information about the behavior of distributions in the extreme domain. Using various specific models for representing these particular data brings a significant reduction of credible intervals of parameters of interest. This kind of approach has been developed in classical statistics terms by Stedinger and Cohn (1986). Interesting extensions concern the use of paleohydrologic data. Bayesian data augmentation algorithms were shown to solve in elegant way the incorporation of historical information into a POT analysis (Parent and Bernier, 2003).

Expert knowledge. Finally, it is worth taking advantage of expert knowledge as pointed out by O’Hagan (1998). Expert beliefs are perhaps the most common available source of information but also the most frequently neglected because frequentist studies consider their use as an act against scientific objectivity. Why should an engineer that has already built 20 dams deliberately wipe off his past experience and consider the hydraulic structure he is working on as the first one he has ever designed? In a Bayesian framework however, rational use of subjective prior expertise can be made and this paper shows how to incorporate such information into the statistical analysis by elicitation of priors bets on the parameters of a POT model.

1.2. Why is prior expertise most often neglected?

Statistical frequentist analysis is recommended usually as the only objective procedure allowing to assess probabilities of events. They are considered as true quantities belonging to the real objective world which should be the only object of study of a scientist. With respect to this philosophical attitude together with difficulties to assess prior probabilities, the Bayesian approach is considered useless and rejected as contrary to scientific objectivity. However in the Bayesian rationale, any conceptual and modeled object, and probability is such an object, has only a subjective meaning, that is a construct of the mind of the scientist. Concerning for instance uncertainty about a parameter θ, considered as a random quantity, the classical principle of objectivity is here replaced with the coherence principle between prior ideas about θ and posterior judgment, given the data, based on posterior probability. The 'information processor' between prior and posterior judgments is the Bayes theorem (Krzysztofowicz, 1983) using the likelihood of the phenomenological model as shown in Fig. 2.

The attention of Bayesian statisticians and practitioners has been recently renewed on the problem of elicitation of priors that is the translation in quantitative terms of opinion of experts, often qualitatively expressed, with due consideration of their own uncertainties. A deep discussion about this subject has been presented in 'The Statistician' (O'Hagan, 1998). Let us note that this discussion now has given up the philosophical issues (assuming the use of prior estimates as scientifically granted) to focus on the practical aspects of priors elicitation. This avenue of thought has been opened in meteorology in the 70s with the numerous pioneering works of Murphy and Winkler (1974a,b) on experimental (and practical) assessment of subjective distributions of meteorological variables. Probabilistic quantification of uncertainties is also a cornerstone of the concepts and methods that Krzysztofowicz (2001) developed to produce probabilistic quantitative prediction forecasts and probabilistic riverstage forecasts for the US National Weather Service. In the field of water resources planning, a recent landmark is the work

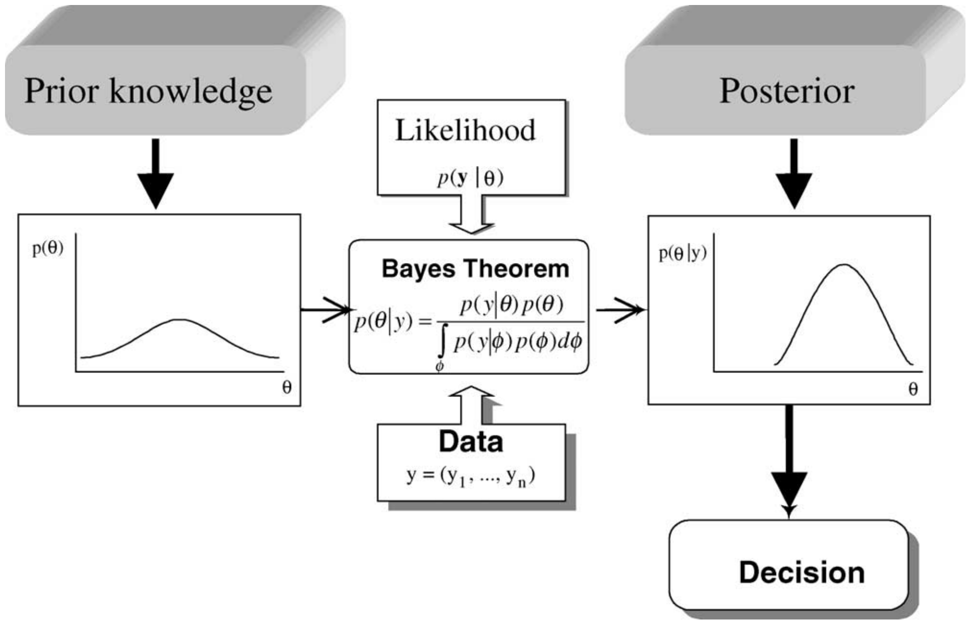

Fig. 2. Outline of the Bayesian paradigm.

by Coles and Tawn (1996) using prior expertise to assess percentiles of extreme daily precipitation distributions. Difficulties can appear in the elicitation process. Palmerini (1995), analyzing experiments in psychology, gives everyday examples of risky situations for which any individual, even a scientifically trained expert relies on psychological heuristics, very different from the usual probability rules and concepts. As a result, shortcuts appear in the line of reasoning, yielding cognitive biases about the prior judgments of experts and possible lack of coherence with the mathematical and probabilistic assessments. Such difficulties do not impede elicitation but must be given due consideration. At first the elicitation procedure must be clearly understood and accepted by the expert to yield reliable results. The direct elicitation of statistical parameters of models must be avoided when these do not have a direct understandable meaning. The parameters of POT models of extremes belong to this category. But easy understandable parameters like medians, means or quantiles are submitted to the expert for elicitation before deriving more complicated parameters. For a hydrology expert the concepts of quantiles, associated with not too large return periods (up to the 100 year event for instance) can be easily perceived as mentioned by O'Hagan (1998).

The paper exemplifies how such an elicitation method can be applied to prior beliefs involved in a POT model for the Garonne case study.

2. Risk analysis of extreme events via POT modeling

Consider a sequence of variables $X_1, X_2, \ldots, X_n$ independently distributed with the same distribution F. Pickands (1975) proved that, under general conditions, the limiting behavior for large u of the sequence $X_1, X_2, \ldots$ on the interval $[u, \infty[$ is a Poisson process with a generalized Pareto intensity function. The approach was developed in hydrology (Smith, 1984) as follows: we consider that u is fixed at a sufficiently high level so that this asymptotic approximation is realistic (Davison and Smith, 1990) and adopt it as a hydrological model for the peaks over the threshold u (POT). Guidelines for the choice of the threshold level can be found in Lang et al. (1999).

Introducing the parameters (μ, ρ, β), the POT model can be written as follows:

$$\Pr(X \le x | X \ge u) = G(x|\rho, \beta, u)$$

$$= \begin{cases} 1 - (1 - \beta(x-u))^{\rho/\beta} & \text{for } \beta \ne 0 \\ 1 - \exp(-\rho(x-u)) & \text{for } \beta = 0 \end{cases} \quad (1)$$

$$\Pr(\#\{X_i \ge u\} = k | \text{over } T \text{ years})$$

$$= \frac{(\mu T)^k \exp(-\mu T)}{k!} \quad (2)$$

β and ρ are the scaling parameters of the generalized Pareto distribution. ρ is strictly positive. $\#\{X_i \ge u\}$ is the random number of floods exceeding the threshold u. μ is the Poisson intensity parameter so that μT is the mean value of this random variable on T. Note that in theory, the model does not assume that the time period T is necessarily an integer nor greater than one year. As an immediate consequence,

$$\Pr(\max(X_1, X_2, \ldots) \le x | T \text{ years}, \ x \ge u)$$

$$= \begin{cases} \exp((-\mu T)(1 - \beta(x-u))^{\rho/\beta}) & \text{for } \beta \ne 0 \\ \exp((-\mu T)\exp(-\rho(x-u))) & \text{for } \beta = 0 \end{cases}$$

This distribution is truncated downwards at $x = u$. It appears to be directly expressed under one of the three asymptotic limiting forms (Gumbel, Frechet, Weibull) for Max values: a suitable change of parameters, $k = \beta/\rho$ and $\alpha = 1/\rho$ reveals the classical form of the well-known truncated GEV distribution:

$$\Pr(\max(X_1, X_2, \ldots) \le x)$$

$$= \begin{cases} \exp\left(-\left(1 - \frac{k}{\alpha}(x-u)\right)^{1/k}\right) & \text{for } k \ne 0 \\ \exp\left(-\exp\left(-\frac{1}{\alpha}(x-u)\right)\right) & \text{for } k = 0 \end{cases} \quad (3)$$

Although numerous hydrological applications of GEV model have taken k and α as working parameters, we will remain focused in the following development on μ, ρ, β: the 'natural' statistical parameters of POT model will be used here because there exists a more statistically tractable semi-conjugate prior model for μ, ρ, i.e. a conjugate prior given β known (Berger, 1985). The likelihood for $\boldsymbol{\theta} = (\mu, \rho, \beta)$ based on an observed series $\mathbf{x}$ of n data $x_1, x_2, \ldots, x_n$ over the threshold u during T years is given by:

$$L_u(\mathbf{x};\mu,\rho,\beta) = \left[\frac{(\mu T)^n \exp(-\mu T)}{n!}\right][\rho^n \exp((\rho-\beta)S_n(\mathbf{x},\beta))] \quad (4)$$

with

$$S_n(\mathbf{x},\beta) = \frac{1}{\beta}\sum_{i=1}^{n}\log((1-\beta(x_i-u))) \quad (5)$$

The special case $\beta = 0$ can be derived by continuity and gives $S_n(\mathbf{x},0) = -\sum_{i=1}^{n}(x_i-u)$).

Eq. (4) shows the lack of sufficient statistics, but, were β known, the POT model would merely belong to the exponential family.

2.1. Semi analytical prior model

A prior model must achieve a compromise between a specific structure for an easy computation of the posterior and a generic structure, large enough to encode a wide range of prior knowledge.

Assume for the moment that prior beliefs about (μ, ρ, β) can be represented by the following pdf:

$$p(\mu,\rho,\beta) = \frac{b_\mu^{a_\mu}}{\Gamma(a_\mu)}\mu^{a_\mu-1}\exp(-\mu/b_\mu)\frac{b^{-a}}{\Gamma(a)}\rho^{a-1} \exp(-\rho/b)\pi_0(\beta) \quad (6)$$

In other words, the marginal prior for β is $\pi_0(\beta)$. $\pi_0(\beta)$ is arbitrary and its functional shape is left for the expert to encode his prior knowledge. Conditionally to these prior beliefs for β, the pdf for ρ is approximated by a gamma distribution with hyper-parameters ($a(\beta)$, $b(\beta)$). If (a,b) are not functions of β, ρ and β are a priori independent. Throughout the paper, it will be assumed that the priors for μ and (β,ρ) are independent. Prior belief about μ is taken in the conjugate Poisson family, that is a gamma distribution with hyper-parameters (a_μ, b_μ) to be estimated when encoding prior expert's knowledge about the mean number of peaks over threshold by time unit.

Practical arguments for using such a prior model rely on the very large class of prior distributions covered with this structure: as functions $\pi_0(\beta)$, $a(\beta)$, $b(\beta)$ need not to be defined parametrically, their shapes will be freely chosen to match a great variety of prior beliefs.

On the theoretical side, however, advantage is kept from partial conjugacy; as the likelihood (4) belongs to a partly exponential family (conditioned upon β), prior and posterior pdf's for $\boldsymbol{\theta}$ will exhibit conjugate (given β) properties for ρ and μ.

This semi-conjugate prior can be used to incorporate further information such as historical (Parent and Bernier, 2003) and regional data. Furthermore, as shown in Section 2.2, a simple Monte Carlo procedure can be implemented to derive the posterior pdf.

2.2. Posterior evaluation of (μ, ρ, β)

The joint posterior pdf $p(\mu,\rho,\beta|\mathbf{x})$ is given via Bayes theorem by combining Eqs. (4) and (6). It has the same conditional structure as the prior pdf:

- μ remains a posteriori independent from (β, ρ). It follows a gamma distribution with (updated) parameters $a_\mu + n$ and $1/b_\mu + T$.
- Conditionally upon β, ρ is also gamma distributed with parameters $a(\beta)+n$ and $1/b(\beta) - S_n(\mathbf{x},\beta)$.
- The marginal posterior density of β is known up to a constant of normalization:

$$p(\beta|\mathbf{x}) \propto \left[\frac{\Gamma(a(\beta)+n)}{\Gamma(a(\beta))}b(\beta)^{-(a+n)}\right]\frac{\pi_0(\beta)\exp(-\beta S_n(\mathbf{x},\beta))}{(1/b(\beta)-S_n(\mathbf{x},\beta))^{\gamma+n}} \quad (7)$$

As β is a unidimensional quantity, this constant can be evaluated once for all by numerical integration by dividing the domain of variation of β into sub-intervals.

Note that, when (a,b) are not functions of β (i.e. prior independence between β and ρ), the first term in brackets of Eq. (7) vanishes. Even in this case, β and ρ are a posteriori dependent as shown by the posterior pdf of β conditioned on ρ :

$$p(\beta|\rho,\mathbf{x}) \propto \left[\frac{b(\beta)^{-a(\beta)}}{\Gamma(a(\beta))}\right]\rho^{a(\beta)-1}\exp\left(-\frac{\rho}{b(\beta)}\right) \times \exp((\rho-\beta)S_n(\mathbf{x},\beta))\pi_0(\beta)$$

Consequently, posterior samples for (μ, ρ, β) are easy to generate directly, avoiding recourse to Monte Carlo Markov Chain (MCMC) techniques (Kuczera and Parent, 1998). Although MCMC techniques are now the standard in Bayesian analysis, practitioners may encounter some unpleasant implementation troubles: very slow convergence, bad choice of tuning parameters for the Metropolis Hastings algorithm. Such trapping difficulties are avoided with our prior model (6): β can be drawn with a uniform number generator by inverting the univariate cumulative distribution function derived from Eq. (7), and μ as well as ρ given β can be obtained via gamma distributed random numbers. The main efforts of the modeler are now to be focused on how to derive functions $\pi_0(\beta)$, $a(\beta)$, $b(\beta)$ from prior expertise.

3. Prior elicitation of POT parameters (μ, ρ, β)

3.1. Hydrological interpretation of μ, q_2, r

The elicitation process should be easily understandable by the hydrologist (expert in his own field does not necessarily mean expert in statistics). The direct elicitation of natural parameters (μ, ρ, β) of POT model or even (μ, k, α) for the GEV one, does not make much sense for the expert in hydrology. A practically tractable method must distinguish between the phenomenological model (Poisson–Pareto POT) on one hand, and the simpler assumptions (elicitation) used to encode the expert opinion on the other hand. The following assumptions are of importance.

The directly assessed parameters are meaningful hydrological quantities such as quantiles or mean values. A first natural choice could be the 10, 100, 1000 year return floods Q_{10}, Q_{100}, Q_{1000}. In practice, these quantities are rather uneasy to assess directly: the hydrologist will give probabilistic statements for Q_{10} and Q_{100} but would feel rather reluctant to quantify his uncertainty about Q_{1000}. In addition, probabilistic judgments about possible values of the return quantiles (Q_{10}, Q_{100}, Q_{1000}) will not be made independently. In this paper we shall assume prior independence between the three following quantities: (1) μ, the annual expectation of the number of peaks over the threshold, (2) the difference of ordinary flood percentiles $\tilde{q}_2 = Q_{100} - Q_{10}$ and (3) the difference of high flood percentiles $\tilde{q}_3 = Q_{1000} - Q_{100}$. The independence hypothesis between μ and ($\tilde{q}_2$, $\tilde{q}_3$) is made because prior knowledge on the temporal process of flood generation is not connected with the intensity of the observed floods. Independence between the *increases* in quantiles $\tilde{q}_2$ and $\tilde{q}_3$ seems a much more reasonable assumption than independence between the return quantiles themselves. In addition the operational interpretation of μ and $\tilde{q}_2$ is straightforward: μ is linked with the hydrologist's past experience for the annual numbers of damaging floods on similar watersheds, $\tilde{q}_2 = Q_{100} - Q_{10}$ stems from hydraulical knowledge arising from extensive analysis of flood data. Q_{1000} (and hence $\tilde{q}_3 = Q_{1000} - Q_{100}$) lies well beyond gaged experience and the probabilistic judgments cannot be based on direct analogies. They have to include knowledge about drainage basin or meteorological characteristics as well. In Coles and Tawn (1996) the expert was able to make such judgments about difference of extreme quantile of precipitations. In our case, the expert felt better working with the ratio of quantiles differences $r = \tilde{q}_3/\tilde{q}_2$ than with $\tilde{q}_3$ because r appears as a dimensionless relative order of magnitude between ordinary and high flood situations.

It also bears a hydrological significance because it is functionally linked with parameter k inside the Poisson–Pareto POT or GEV family of distributions as shown in Fig. 3:

$$r = \frac{\tilde{q}_3}{\tilde{q}_2} = \frac{(-\log(0.99))^k - (-\log(0.999))^k}{(-\log(0.9))^k - (-\log(0.99))^k} \tag{8}$$

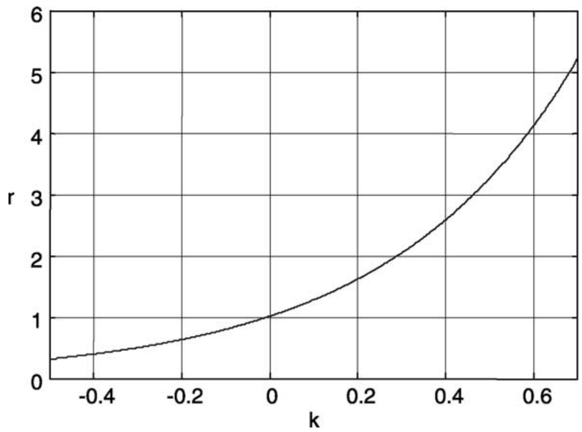

Fig. 3. Ratio r as a function of k.

The particular value $r = 1$ corresponds to $k = 0$, i.e. the Gumbel limiting distribution for the annual maxima. When $r > 1$ the distribution belongs to the Weibull domain and the relative flood magnitudes increase to a greater extend when passing from the 100 year return flood to the 1000 year return one than when passing from the 10 year to the one year return floods.

The expert will also be asked about his or her prior beliefs about r, and therefore this subjective information will in turn give credibilities about the domains to which the limiting distribution for the Max may belong to.

The elicitation model contains the complementary assumptions allowing to encode the knowledge of the expert into a pdf. It is rather easy to ask the expert's opinion about each parameter μ, $\tilde{q}_2$ and r separately, but it is more difficult to go beyond marginal distributions and make him assess a complete prior shape of distributions including prior dependence properties between these parameters. Therefore complementary assumptions have to be made concerning interrelations between parameters. For no reason but mathematical convenience, the gamma pdf is generally adopted as a model for the variation of a positive unimodal variable. Following Coles and Tawn (1996) we shall assume that μ, $\tilde{q}_2$ and $\tilde{q}_3$ are prior independent and gamma distributed. As there is a one to one mapping between (μ, $\tilde{q}_2$, $\tilde{q}_3 = \tilde{q}_2/r$) and ($\mu$, ρ, β), the elicitation model (μ, $\tilde{q}_2$, r) could be transferred in term of an elicitated prior model for (μ, ρ, β) using Jacobian transformations.

The family of modeled prior functions, i.e. the set of semi-conjugate priors $p(\mu, \rho, \beta)$ associated with the phenomenological model (POT) and chosen for theoretical and ease of computation reasons is given by the specific structure (6). In general, it will not contain the exact elicitation model that was previously described. But as many degrees of freedom are given by the wide choice left for nonparametric functions $\pi_0(\beta)$, $a(\beta)$, $b(\beta)$ in Eq. (6), there exist members from this family which are close enough to the elicitation model for practical purposes. In what follows, a simulation-based technique is described to select such a semi-conjugate prior from Eq. (6).

3.2. Encoding prior knowledge

The Garonne river near Agen (drainage basin area of 52,000 km^2) is used as a case study. No records of data is presented to the expert but the main rivers of the region and their tributaries were well documented in Pardé (1935, 1963). These works published in the beginning of the 1900s are famous among French hydrologists. The hydrological background of the second author was used as expert's knowledge; the practical assumptions of the elicitation process were chosen as simple as possible and the expert was asked to assess prior judgments as reasonably simple as it could be made in such a context. Using the elicitation assumptions, expert's knowledge is encoded into a semi-conjugate prior by a seven-step procedure. The units for all flows given in the subsequent graphs and tables are in m^3/s.

3.2.1. Decisional bets to encode uncertainty about meaningful hydrological quantities

In the Bayesian paradigm, the quantitative estimation of the expert belief relies on direct probabilistic judgments. There exist many procedures as described in the special issue of 'The Statistician' (Kadane et al., 1998) devoted to prior elicitation. These procedures generally begin with the calibration of expert's judgments by means of hydrological guesses compared by the expert with 'objective' guesses (lotteries) (see Berger, 1985; Bernier et al., 2000).

The interest of the Bayesian analyst does not lie in the psychological or logical connections in the mind of expert, nor in the means to make his conceptual short cuts acceptable for others hydrologists. The Bayesian conception of probability is subjective and personal and only under asymptotical circumstances (when the information stemming from the data dominates the prior knowledge), two experts with different priors will end up sharing the same posterior.

Therefore, the first step of the elicitation process is to ask the expert to give two degrees of belief θ_p, concerning each elicitated parameter, obtained by subjective bets with $p = 0.5$ (i.e. with odds 1 against 1 to be larger than the assessed value) and $p = 0.9$ (with odds 1 against 9), for each parameter $\theta = \mu$, $\tilde{q}_2$ and r. Table 1 shows expert's answers for our case study

Table 1
Expert's beliefs for meaningful hydrological quantities μ, $\tilde{q}_2$ and r

Hydrological meaning	Elicitated θ	Median ($p = 0.5$)	90% Quantile
Annual mean number of floods >2500 m^3/s	μ	1.7	2.1
Difference between 100 and 10 year return floods	$\tilde{q}_2 = Q_{100} - Q_{10}$	1000 m^3/s	1600 m^3/s
Ratio of quantile differences	$r = \dfrac{Q_{1000} - Q_{100}}{Q_{100} - Q_{10}}$	2	3.5

(annual maximum floods of the Garonne river near Agen with a threshold $u = 2500$ m^3/s).

During the assessment process, the expert carefully considered the differences between the 90% percentile and the median for the all quantities to be elicitated. He was not really feeling at ease with the values he gave for r, arguing that he was somewhat too overconfident when thinking indirectly. Indeed, he mentally considered a rather small collection of possible values for $\tilde{q}_2$ and $\tilde{q}_3$, and informally 'evaluate' their ratios in order to pick a 'median' and a 'high' percentile for r. However, consideration of Fig. 3 brought some relief since he got aware that his upper value $r = 3.5$ fairly corresponds to $k = 0.5$ that is traditionally assumed as a rare upper value for k in many floods studies. Finally all these assessed values were accepted for the present elicitation exercise.

3.2.2. From hydrological assessments to a tentative prior distribution for intermediate parameters μ, $\tilde{q}_2$, $\tilde{q}_3$

μ, $\tilde{q}_2$, r are considered as hydrological model parameters of which expert assesses a prior distribution described with 'estimated hyperparameters' based on his previous judgments. Hence assuming μ, $\tilde{q}_2$, $\tilde{q}_3$, independently distributed according to gamma distributions with hyperparameters (a_μ, b_μ), (a_2, b_2) and (a_3, b_3), the corresponding couples of parameters are determined to satisfy the previous constraints. With these assumptions, $z = 1/(1 + (b_2/b_3)r)$ is distributed according to a beta distribution with parameters (a_2, a_3) allowing to calculate a_3, b_3 from the prior beliefs on r knowing a_2, b_2. The numerical values for these hyperparameters are given in Table 2.

Table 2
Hyper-parameters for prior gamma distributions of μ, $\tilde{q}_2$ and $\tilde{q}_3$

a_μ	b_μ	a_2	b_2	a_3	b_3
34	0.05	6.50	162.08	100	20

With these parameters in hand, we proceed to classical Monte Carlo simulation of 20,000 independent sample values of μ, $\tilde{q}_2$, $\tilde{q}_3$ and r (according to the estimated a_2, b_2, a_3, b_3). Computations of the corresponding values of parameters μ, β, ρ are straightforward performed with the idea to fit a semi-conjugate structure as given by Eq. (6). μ has the desired form of the semi-conjugate model and his posterior inference can be dealt separately from (β, ρ). Unfortunately when plotting the marginal distribution of β as in Fig. 4, no candidate model for the marginal $\pi_0(\beta)$ comes to the analyst's mind: the high skewness would prevent from fitting a traditional normal pdf, the shape with the quick decreasing right tail does not agree with a possible gamma or inverse gamma behavior.

3.2.3. Monte Carlo simulations to by-pass Jacobian transformation from μ, $\tilde{q}_2$, $\tilde{q}_3$ into , ,

The traditional GEV parameters $k = \beta/\rho$ and $\alpha = 1/\rho$ can also be simulated from the previous 20,000 independent sample values of μ, $\tilde{q}_2$ and r. Luckily, their prior pdfs exhibit smooth normal-like behavior.

3.2.4. Normal approximation of intermediate parameters $k = \dfrac{\beta}{\rho}$ and $\lambda = \log(\rho)$

Fig. 5 shows the histogram of k and its normal fit $N(m_k, \sigma_k^2)$. It is also possible to take into account skewness and kurtosis with a normal mixture model (not used in this case). In the following, we will assume:

$$k = N(m_k, \sigma_k^2) \tag{9}$$

Values of m_k, σ_k^2 are, respectively, 0.298 and 0.032 when estimated by max-likelihood.

Fig. 6 shows that the conditional model of $\lambda = \log(\rho)$ given k is pretty well fitted by a linear regression with residuals distributed according to a normal mixture, i.e.

$$\begin{aligned} &\lambda = l_0 + l_1 k + \epsilon \\ &\epsilon = (\delta)N(0, \sigma_1^2) + (1 - \delta)N(0, \sigma_2^2) \end{aligned} \tag{10}$$

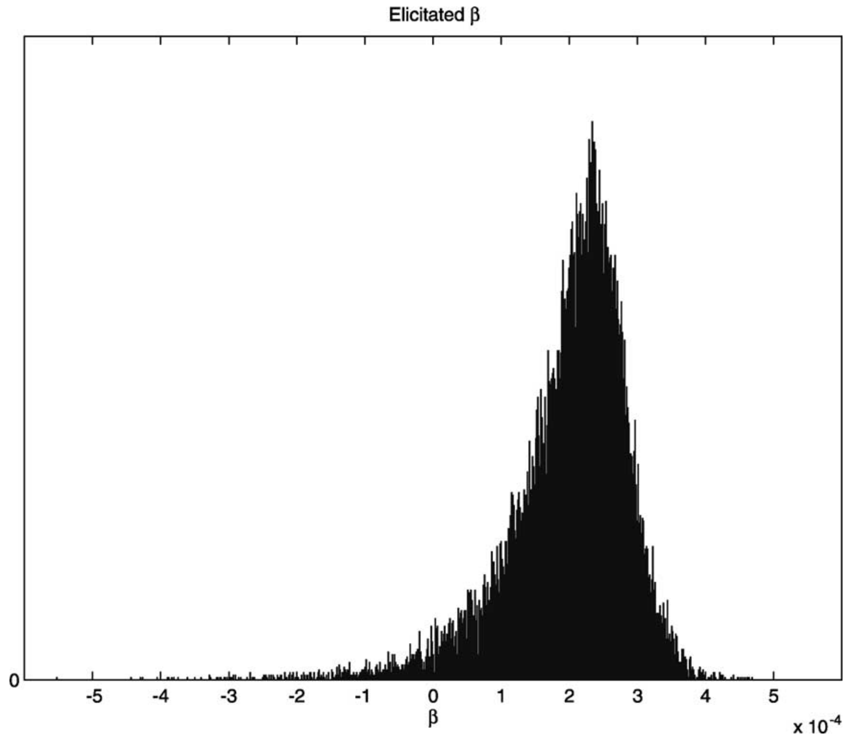

Fig. 4. The distribution of 20,000 simulated values of β is neither normal nor gamma.

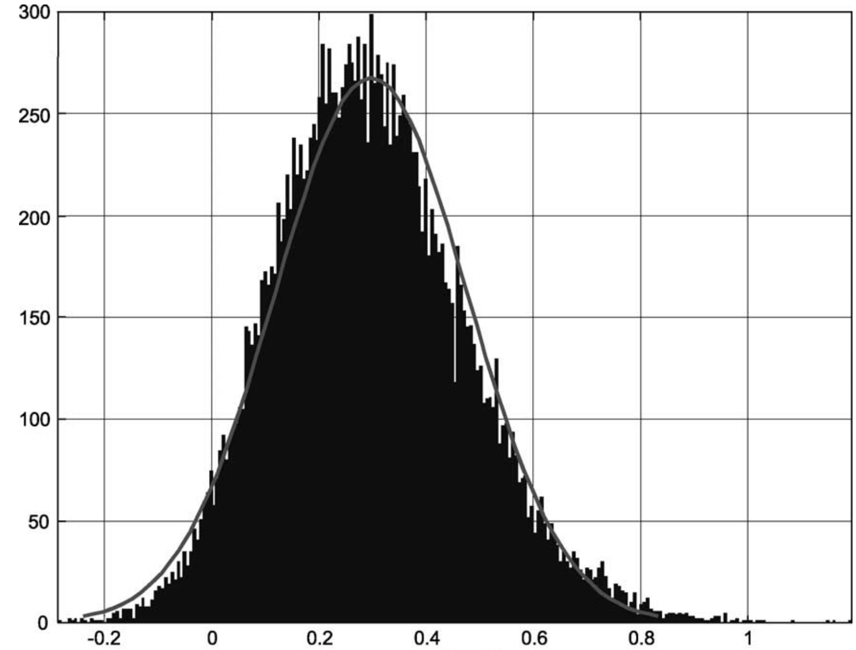

Fig. 5. Distribution of 20,000 simulated values of k.

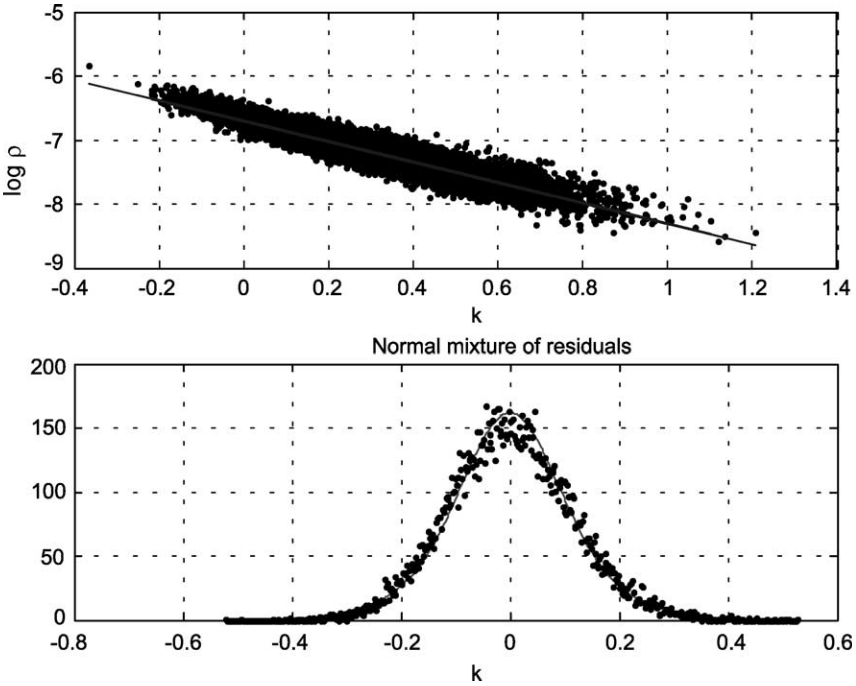

Fig. 6. The two parameters of the gamma pdf for $p(\rho|\beta)$ as a function of β.

These formulae are the more general expressions used in our study. For our case study, the simpler model with $\delta = 0.5$ works well. The parameter values are given in Table 3.

This approximation, namely 'k—normally distributed, $\lambda|k$, conditionally distributed according a normal mixture with linear dependence in mean with k, introduces seven additional hyper-parameters (m_k, σ_k, l_0, l_1, δ, σ_1, σ_2) allowing to write the density function of k, λ under an analytical form based on normal expressions. Moreover Eq. (10) shows that only Bernoulli and Normal random generators are used to sample (k, λ) from:

$$p(k,\lambda|m_k,\sigma_k,l_0,l_1,\delta,\sigma_1,\sigma_2)$$
$$= p(\lambda|k,l_0,l_1,\delta,\sigma_1,\sigma_2)p(k|m_k,\sigma_k)$$

The elicitation model for the joint distribution of β and ρ can be analytically derived via Jacobian transformations. Formally it reads as $p(k,\lambda|m_k,\sigma_k, l_0,l_1,\delta,\sigma_1,\sigma_2)\ |\partial(\beta,\rho)/\partial(k,\lambda)|$. As long as sampling issues are addressed, there is no need to evaluate this expression in closed form. Samples for (ρ, β) can be derived from the previous samples (k,λ) since $\rho = \exp(\lambda)$ and $\beta = k\exp(\lambda)$.

3.2.5. Picking a semi-conjugate prior

Formally the prior for β is a function of the seven hyper-parameters (m_k, σ_k, l_0, l_1, δ, σ_1, σ_2) which reads:

$$\pi_0(\beta) = \int_\rho p(k,\lambda|m_k,\sigma_k,l_0,l_1,\delta,\sigma_1,\sigma_2)\left|\frac{\partial(\beta,\rho)}{\partial(k,\lambda)}\right|\mathrm{d}\rho$$

The reader wishing to struggle with formal calculus of integrals may take some pleasure in searching for the analytic expression of $\pi_0(\beta)$, but as before, such a task is useless since prior samples of any size for β can be derived via the previous sampling scheme using Eqs. (9) and (10).

Table 3
Hyper-parameters for normal distributions describing (κ, $\log(\lambda)$)

m_k	σ_k^2	δ	l_0	l_1	σ_1^2	σ_2^2
0.298	0.032	0.5	− 6.7029	− 1.6029	0.01989	0.00663

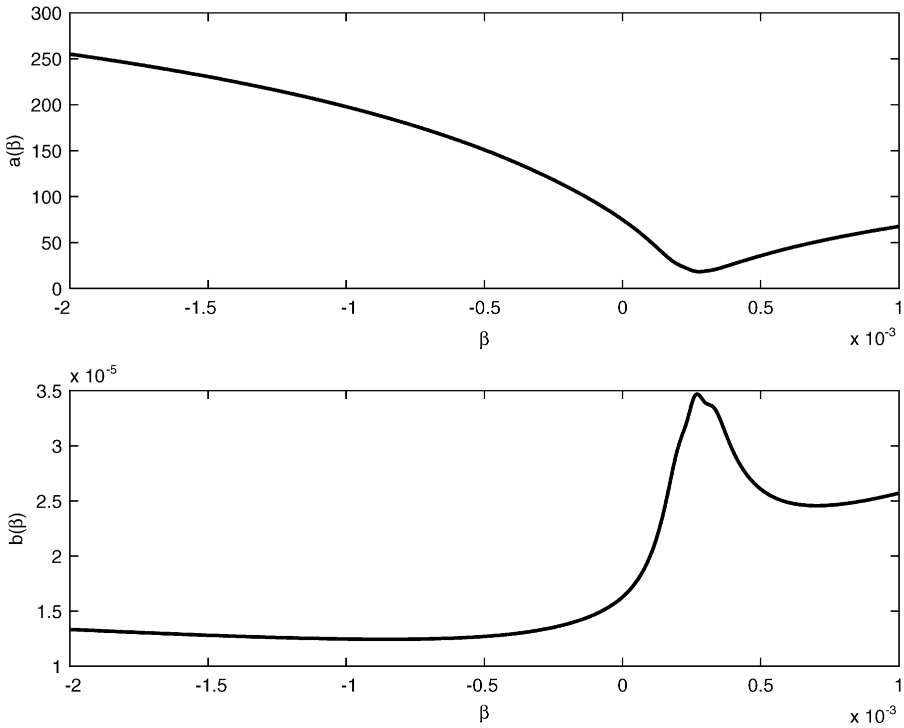

Fig. 7. Conditional mean of log p given k and residuals.

The following sub-step consists in the approximation of the conditional $p(\rho|\beta)$ with the modeled conditional prior following a gamma distribution with parameters $a(\beta)$, $b(\beta)$. The range of β is divided into sub-intervals and for each of these sub-intervals, estimates for conditional expectations $\tilde{E}(\rho|\beta)$ and variances $\tilde{V}(\rho|\beta)$ are computed from the (ρ, β) sample. Then the center of the sub-intervals is related to the corresponding values of the gamma hyperparameters a and b :

$$a(\beta) = \frac{(\tilde{E}(\rho|\beta))^2}{\tilde{V}(\rho|\beta)} \qquad b(\beta) = \frac{\tilde{V}(\rho|\beta)}{\tilde{E}(\rho|\beta)} \tag{11}$$

Fig. 7 is based on a file keeping triplets of (β, $a(\beta)$, $b(\beta)$) for further applications. A low order polynomial approximation could also have been fitted. The weird shape of the functions $a(\beta)$, $b(\beta)$ is not surprising: the semi conjugate prior Eq. (6) inherits from the functional structure of the likelihood (4). $a(\beta)$ corresponds to the sample size n in Eq. (4) and $b(\beta)$ corresponds to the function $S_n(x, \beta)$ given by Eq. (5).

At this step, the independent gamma distribution for μ (with parameters (a_μ, b_μ)), the marginal $\pi_0(\beta)$ and the conditional gamma distribution for ρ (with parameters $a(\beta)$, $b(\beta)$) are taken as a modeled joint prior belonging to the desired semi conjugate family of priors.

3.2.6. Checking that elicitated prior and modeled priors are close enough

Finally for testing purposes, 5000 values of the parameters are sampled in this modeled prior and the results compared with those of the previous values obtained with the elicitation model. Fig. 8 gives three histograms for the prior marginals of (k, β, ρ) and one scatter plot to represent covariations between ρ and β based on the 20,000 simulations of POT parameters obtained with the 'elicitation model' (functions of quantiles parameters independently distributed as gamma). This set must be compared with the same graphs given in Fig. 9 resulting from simulations obtained with the semi conjugate modeled prior.

Similarities between the two sets appear to be very strong both qualitatively and quantitatively.

Furthermore Table 4 gives results about medians of the sample values performed for the elicitated

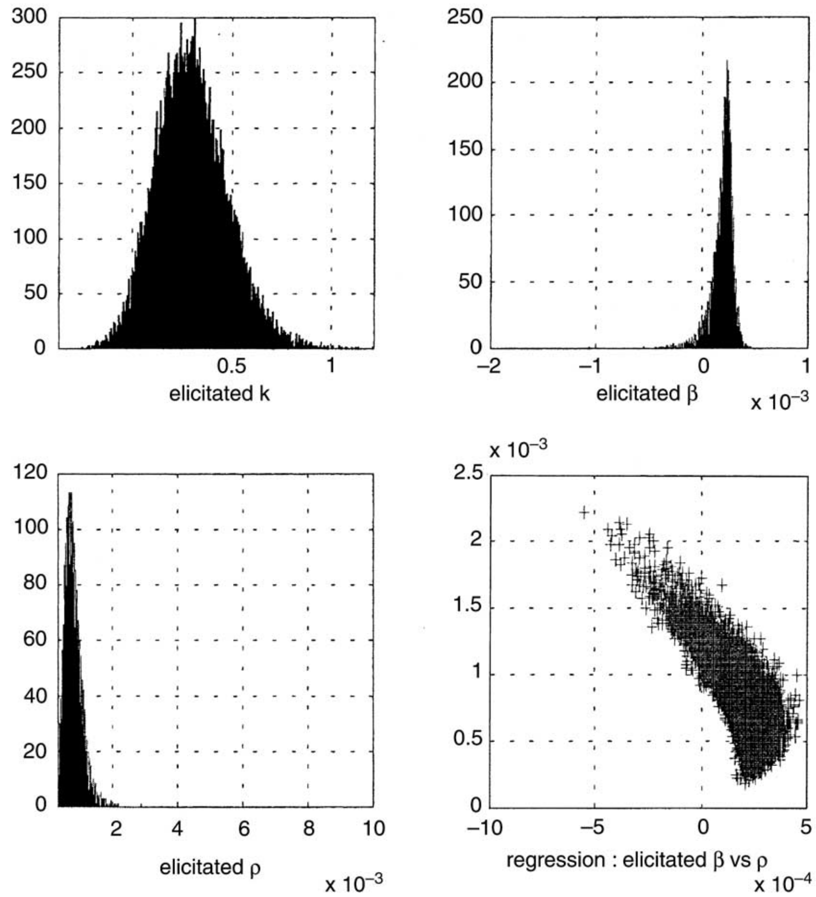

Fig. 8. Results on the elicitation model prior.

and modeled parameters. Table 3 shows the 90% empirical quantiles for μ, $\tilde{q}_2$, $\tilde{q}_3$ and r from the samples. Tables 5 and 6 exhibit similar information for 10, 100 and 1000 year return floods.

Tables 4–7 show elicitated versus modeled results which seem to be quite comparable (relative differences less than 10%) except for $\check{q}_3$ for which the differences are important. This result concerning $\check{q}_3$ illustrates some incompatibility between the analytical assumptions (prior gamma distribution and setting independence of quantiles variations on one hand, the Poisson–Pareto POT model on the other hand). However the differences are much weaker in terms of absolute quantiles themselves.

3.2.7. Overall checking of the method

A sensitivity analysis is performed on quantiles assessed from the expert. This semi-analytical, semi Monte Carlo method for determining prior distribution of β, ρ, marginal prior $\pi_0(\beta)$ and conditional $p(\rho|\beta)$ is checked to be stable and robust.

4. Posterior analysis of the Garonne case study with informative prior

The flood data of the Garonne river near Agen at the same site (Mas d'Agenais) have been recorded for the period 1913–1977. The peak flows above 2500 m^3/s were selected. 151 peaks were obtained in 65 years. All posterior estimates with informative prior and non-informative prior modeling prior ignorance have been evaluated via MATLAB scripts that call to ready-made random number generators (Normal, Bernoulli, Beta, Gamma). The posterior analysis is based on direct Monte Carlo sampling after

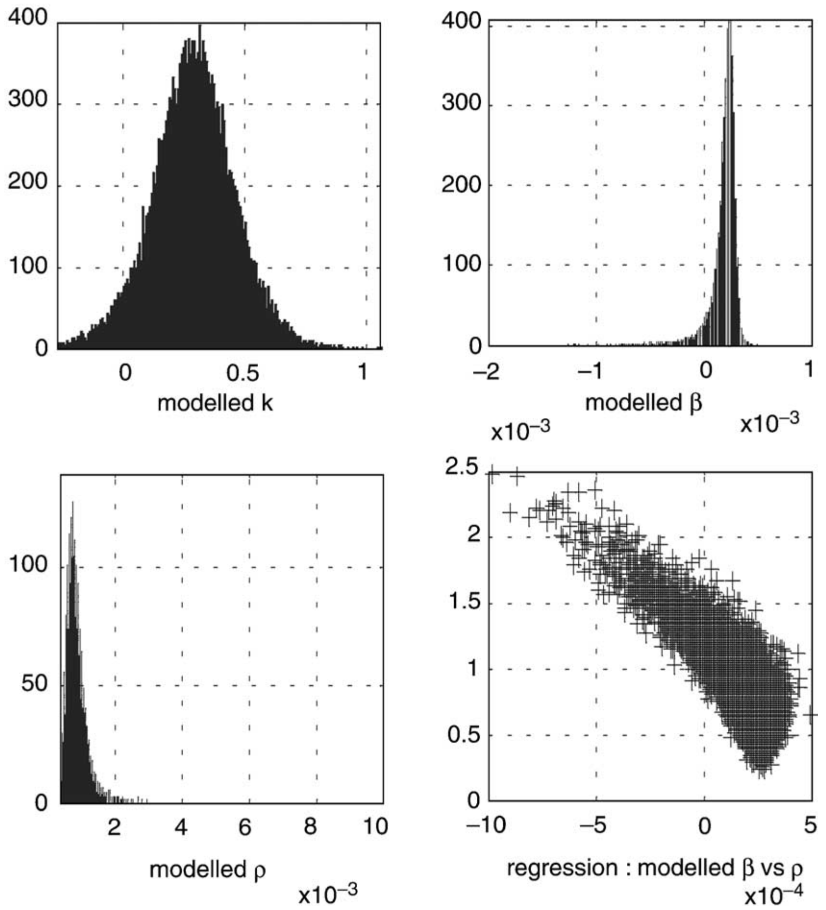

Fig. 9. Results on the modeled prior.

Eq. (7) has been evaluated at the centers of subintervals partitioning the range of β. A sample of size 200 was used for the statistical analysis.

Fig. 10 shows density estimates of 10, 100, 1000 year floods with observed data taking in account:

- non-informative prior,
- the previously informative elicitated prior,

Fig. 11 presents posterior mean of quantiles together with 5 and 95% credible limits as a function of return periods.

The POT model constrains quantiles to be linked together and the expert's opinion plays some role on the location of design quantiles, bringing them towards his or her prior expertise. More clearly it

Table 4
Comparison of sample prior medians for elicitated and modeled μ, $\tilde{q}_2$, $\tilde{q}_3$ and r

Medians for	μ	$\tilde{q}_2$	$\tilde{q}_3$	r
Elicitated	1.69	1001	2009	2.00
Modeled	1.71	987	490	2.02

Table 5
Comparison of prior 90% credible bounds for elicitated and modeled μ, $\tilde{q}_2$, $\tilde{q}_3$ and r

90% Credible bound for	μ	$\tilde{q}_2$	$\tilde{q}_3$	r
Elicitated	2.10	1609	2265	3.51
Modeled	2.11	1654	1349	3.30

Table 6
Comparison of prior sample medians for elicitated and modeled 10, 100 and 1000 year return floods Q_{10}, Q_{100}, Q_{1000}

Medians for	Q_{10}	Q_{100}	Q_{1000}
Elicitated	4987	6062	6596
Modeled	5001	6028	6515

Table 7
Comparison of prior 90% credible bounds for elicitated and modeled 10, 100 and 1000 year return floods Q_{10}, Q_{100}, Q_{1000}

90% Credible bound for	Q_{10}	Q_{100}	Q_{1000}
Elicitated	5550	6708	7842
Modeled	5473	6755	7962

appears, on all these curves, the uncertainty reduction effect of taking into account prior knowledge with reference to the usual non informative prior.

Table 8 considers the one hundred year return flood Q_{100}, with its 90% credible interval (Bayesian counterpart of classical confidence intervals). The range Δ of this 90% credible interval ($\Delta = 95\%$ credible bound $-$ 5% credible bound) is cut in two when incorporating prior expertise.

This is a large reduction due to the prior statements of the expert. In a frequentist-based hydrological study on the same data, Miquel (1984) showed similar results using non-informative prior but taking into account 12 historical records observed in the previous 150 years. The present reduction in Δ is almost equal to the one obtained with such historical information.

4.1. Discussion

Coles and Tawn (1996) presented a similar elicitation exercise (with Professor Duncan Reed as the expert) in the connected field of meteorology. They did not use the distinction between elicitation and modeled priors and directly introduced the elicitation model into the Bayes rule to compute the posteriors. Their method may appear much easier since it avoids approximation and transformation. But recourse is to be made to the Metropolis Hasting algorithm (Kuczera and Parent, 1998) the latest brute force development from the MCMC toolbox of statistical estimation techniques. The use of the semi-conjugate structure of the modeled prior has many advantages.

The first advantages are computational. The method presented in the elicitation process relies on direct simulation of posterior distributions, not MCMC ones for which caution must be taken. MCMC runs only asymptotically provide samples from the posterior pdf, and the initial non-ergodic period may take long. In addition, convergence tests for MCMC techniques still belong to the open fields of statistical research. Conditional posterior distributions of parameters are partly analytically known so that Rao Blackwell

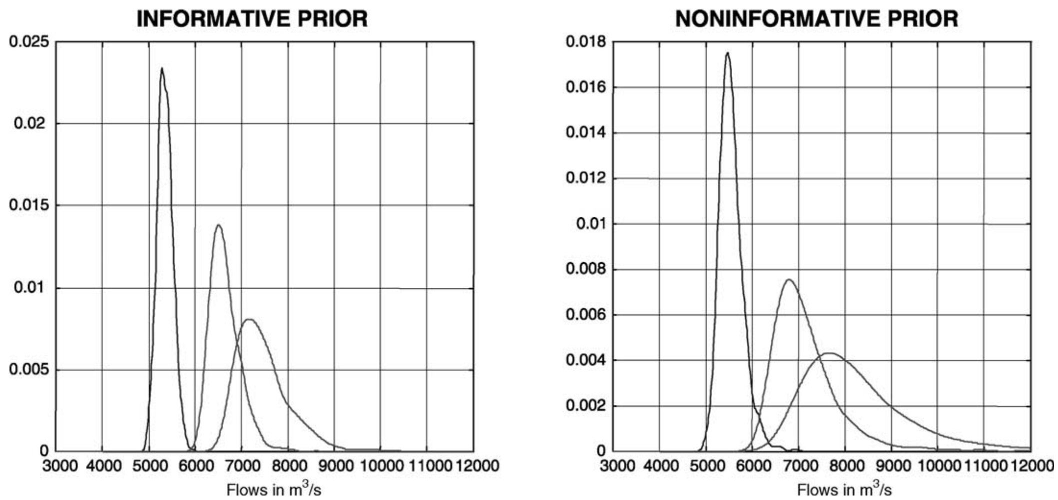

Fig. 10. Density estimates for Q_{10}, Q_{100}, Q_{1000}.

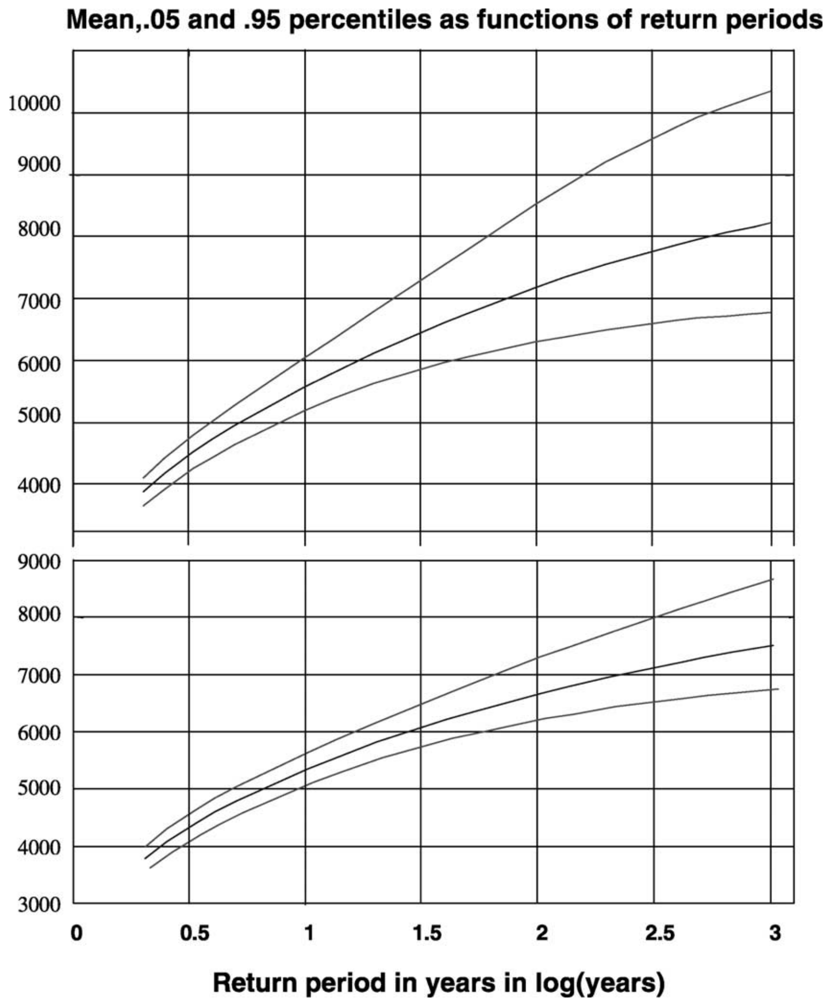

Fig. 11. Posterior mean and percentiles of Q_T as functions of return period T.

variance reduction techniques are used here to provide smooth estimates, a trick that could not be obtained when using Metropolis Hastings algorithms. Further important developments of the POT model (introduction of historical data, predictive analysis of decision, regional analysis of floods) will increase its complexity and may necessitate recourse to MCMC techniques. But as, even in this case, conditional posterior distributions will remain partly known, so that most posterior analysis would be performed via Gibbs sampling, a much more efficient MCMC technique than the Metropolis Hasting algorithm.

There are also conceptual advantages stemming from the semi-conjugate prior pdf structure. Of course

Table 8
Precision about the one hundred year return flood Q_{100}, has doubled when incorporating expertise

Design value Q_{100}	Post-median	5% Credible lim.	95% Credible lim.	Δ
Q_{100} (with non-informative prior)	7000	6290	8500	2210
Q_{100} (with informative prior)	6590	6190	7240	1050

picking a prior from the POT semi-conjugate family restricts the modeler's choice: for instance, the POT model structure imposes that, at last a posteriori, quantiles are rather strongly correlated. Nevertheless we believe that the semi-conjugate modeled prior is versatile enough to take into account most quantified opinions of experts.

An elicitation model is just a way of encoding prior expertise. Coles and Tawn (1996) assumed that Q_{10}, $\tilde{q}_2$ and $\tilde{q}_3$ are independently gamma distributed. We adopted and adapted these assumptions because they are convenient but other elicitation models can be made. What is the effect of these assumptions? An other elicitation model, or another expert's opinion could have been encoded following the same way detailed here into a semi-conjugate prior. Two different elicitation models are equivalent if they lead to similar posterior inference and operational results for flood design. The last word is left to the expert to say whether the distance between two priors or between an elicitated prior and its semi-conjugate approximation is acceptable or not: simulations are easy to perform, allowing to compute various meaningful hydrological quantities to be checked by the hydrologist.

5. Conclusions

The following conclusions can be made:

- The Garonne case study exemplifies how valuable information carried by prior expertise can reduce flood design uncertainty. On the Garonne example, the design value and its credible interval are notably changed when incorporating prior evidence into the study, whatever the way of encoding this information.
- A complete way of encoding prior expertise is proposed via the semi conjugate structure for POT prior. This prior structure is flexible enough to encode a wide variety of prior degrees of belief. It also contributes to a quick and easy direct posterior inference.
- Fruitful results are obtained by adopting the Bayesian perspective for risk analysis of extreme events: professional expertise, even qualitative and subjective can be associated with quantitative experimental data. From an engineer's point of view, new practical issues opened by the Bayesian approach for hydraulic structure design will overcome old philosophical controversies about the subjective nature of probabilities.

References

Berger, J.A., 1985. Statistical Decision Theory and Bayesian Analysis, Second ed, Springer, Berlin.

Bernier, J., 1967. Les méthodes Bayesiennes en hydrologie. International Hydrology Symposium, Fort Collins.

Bernier, J., 1987. Elements of Bayesian analysis of uncertainties in hydrology reliability and risk models. In: Duckstein, L., Plate, E.J. (Eds.), Engineering Reliability and Risk in Water Resources, NATO ASI series, Tucson, pp. 405–422.

Bernier, J., Parent, E., Boreux, J.J., 2000. Statistique pour l'environnement. Traitement bayésien de l'incertitude. Lavoisier, Tec § Doc

Bobée, B., 1999. Estimation des événements extrêmes de crue par l'analyse fréquentielle: une revue critique. La Houille Blanche (7–8), 100–105.

Coles, S.G., Powell, E.A., 1996. Bayesian methods in extreme value modelling. International Statistics Review 64 (1), 119–136.

Coles, S.G., Tawn, J.A., 1996. A Bayesian analysis of extreme rainfall data. Applied Statistics 45 (4), 463–478.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. Journal of the Royal Statistical Society 52 (3), 393–442.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis, Chapman & Hall, New York.

GREHYS, 1996. Research group in statistical hydrology: presentation and review of some methods for regional flood frequency analysis. Journal of Hydrology 186, 63–84.

Kadane, J.B., Wolson, L.J., O'Hagan, A., Craig, et al., 1998. Papers on elicitation and discussions. The Statistician 47 (1), 3–53.

Krzysztofowicz, R., 1983. Why should a forecaster and a decision maker use bayes theorem? Water Resources Research 19 (2), 327–336.

Krzysztofowicz, R., 2001. The case of probabilistic forecast in hydrology. Journal of Hydrology 249, 2–9.

Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. Journal of Hydrology 211 (4), 69–85.

Lang, M., Ouarda, T.B.M.J., Bobée, B., 1999. Towards operational guidelines for over-threshold modeling. Journal of Hydrology 225 (3–4), 103–117.

Madsen, H., Rosbjerg, D., 1997. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modelling. Water Resources Research 33 (4), 771–781.

Madsen, H., Rosbjerg, D., Harremoës, P., 1995. Application of the Bayesian approach in the regional analysis of extreme rainfalls. Stochastic Hydrology and Hydraulics 9 (1), 77–88.

Miquel, J., 1984. Guide Pratique d'Estimation des Probabilités de Crues, Eyrolles, Paris.

Murphy, A.H., Winkler, R.L., 1974a. Credible intervals temperature forecasting: some experimental results. Monthly Weather Review 102, 784–794.

Murphy, A.H., Winkler, R.L., 1974b. Probabilistic temperature forecasts: the Casefor an operational program. Bulletin of American Meteorological Society. 102, 12–19.

O'Hagan, 1998. Elicitation expert beliefs in substantial practical applications. The Statistician 47 (1), 21–35.

Palmerini, M.P., 1995. La Réforme du Jugement ou Comment ne plus se Tromper, O. Jacobs, Paris.

Pardé, M., 1935. Le régime de la Garonne. Revue Géographique des Pyrénées et Sud-Ouest.

Pardé, M., 1963. Etude de quelques grandes crues caractéristiques en France. Revue Générale.

Parent, E., Bernier, J., 2003. Bayesian POT modeling for historical data. Journal of Hydrology, 274, 95–108.

Pickands, J., 1975. Statistical inference using extreme order statistics. Annals of Statistics 3 (1), 119–131.

Rasmussen, P.F., Rosbjerg, D., 1991. Evaluation of risk concepts in partial duration series. Stochastic Hydrology and Hydraulics 5 (1), 1–16.

Rosbjerg, D., Madsen, H., Rasmussen, P.F., 1992. Prediction in partial duration series with generalized Pareto-distributed exceedances. Water Resources Research 28 (11), 3001–3010.

Savage, L.J., 1972. The Foundations of Statistics, Dover, New York.

Smith, R.L., 1984. Threshold model for sample extreme. In: de Oliveira, T., (Ed.), Statistical Extremes and Applications, Reidel, Dordrecht, pp. 621–638.

Stedinger, J.R., Cohn, T.A., 1986. Flood frequency analysis with historical and paleoflood information. Water Resources Research 22 (5), 785–793.

Tawn, J.A., 1993. Extreme sea levels. In: Barnet, V., Turkman, K.F. (Eds.), Statistics for the Environment, Wiley, New York, pp. 243–263.