Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Hydrology 313 (2005) 97–116

Journal of Hydrology
www.elsevier.com/locate/jhydrol

# Bayesian MCMC flood frequency analysis with historical information

Dirceu S. Reis Jr., Jery R. Stedinger*

*School of Civil and Environmental Engineering, Hollister Hall, Cornell University, Ithaca, NY 14853-3501, USA*

**Abstract**

This paper explores Bayesian Markov Chain Monte Carlo (MCMC) methods for evaluation of the posterior distributions of flood quantiles, flood risk, and parameters of both the log-normal and Log-Pearson Type 3 distributions. Bayesian methods allow a richer and more complete representation of large flood records and historical flood information and their uncertainty (particularly measurement and discharge errors) than is computationally convenient with maximum likelihood and moment estimators. Bayesian MCMC provides a computationally attractive and straightforward method to develop a full and complete description of the uncertainty in parameters, quantiles and performance metrics. Examples illustrate limitations of traditional first-order second-moment analyses based upon the Fisher Information matrix.


*Keywords:* Bayesian MCMC; Log-Pearson Type 3 distribution; Log-normal distribution

## 1. Introduction

Estimation of flood quantiles and other extremes is an important issue. Determination of the magnitude of design floods with a specified exceedance probability is required for many engineering works, such as the design of bridges, dams, canals, and water intakes, and for the development of flood risk management projects. But determining the magnitude of the 100-year flood is not enough. Decision makers should be provided estimates of the precision of the estimated quantiles so they can appreciate the uncertainties. Uncertainties in quantile estimates are often based on asymptotic approximations. Papers and books describe approximate confidence intervals for quantiles of distributions used in hydrology (USWRC, 1982; Stedinger, 1983a; Kite, 1988; Chowdhury and Stedinger, 1991; Stedinger et al., 1993; Bobee and Ashkar, 1991; Ashkar and Ouarda, 1998; Whitley and Hromadka, 1999; Cohn et al., 2001; Frances, 2001; Coles and Pericchi, 2003) illustrating the importance of the provision of measures of precision. However, uncertainties in anticipated expected benefits and other project performance indices are also desirable, particularly when decision-making is cast in a more comprehensive risk and uncertainty framework (NRC, 1995; USACE, 1996; Al-Futaisi and Stedinger, 1999; National Research Council, 2000).

The data available for a single site is usually insufficient to obtain estimates of flood quantiles with

* Corresponding author. Tel.: +1 607 255 2351; fax: +1 607 255 9004.

*E-mail address:* jrs5@cornell.edu (J.R. Stedinger).

0022-1694/$ - see front matter 
doi:10.1016/j.jhydrol.2005.02.028

great precision. Regional and historical information can be incorporated into flood frequency analyses to increase the precision of the estimators. The former uses data from different sites in an attempt to substitute space for time (National Research Council, 1998; Stedinger and Lu, 1995; Hosking and Wallis, 1997). The latter incorporates into the analysis information about floods that occurred before the beginning of a systematic gauged record (Jarrett and Tomlinson, 2000; House et al., 2002; Benito and Thorndycraft, 2004). Several studies have illustrated the use of historical information in flood frequency analysis (Stedinger and Cohn, 1986; Jin and Stedinger, 1989; Pilon and Adamowski, 1993; Salas et al., 1994; Stevens, 1994; Frances et al., 1994; Cohn et al., 1997; Kuczera, 1999; Martins and Stedinger, 2001; Ostenaa et al., 2002; O'Connell et al., 2002; Blainey et al., 2002; Parent and Bernier, 2003; Thorndycraft et al., 2002; Benito et al., 2004; Frances, 2004; Ouarda et al., 2004). These studies generally show that the use of historical information can be of great value in the reduction of the uncertainty in flood quantiles estimators, though many have raised concerns with measurement and recording errors (Hosking and Wallis, 1986, 1997; Sho et al., 2000; Blainey et al., 2002; Baker, 2003).

The most common estimation methods in flood frequency analysis are based on either the method of moments or the maximum likelihood. Other authors have proposed the use of the Bayesian approach with systematic data for special cases when it is possible to use conjugate priors that yeild a closed form posterior distribution for the parameters (Vicens et al., 1975; Wood and Rodriguez-Iturbe, 1975; and Stedinger, 1983b). More recently, Parent and Bernier (2003) developed a Bayesian procedure for the peak over threshold (POT) model that uses semi-conjugate informative priors leading to a quasi-analytical formulation. Coles and Pericchi (2003) and Coles et al. (2003) argue for use of a Bayesian approach and employed a MCMC algorithm with the three-parameter GEV distribution for modeling systematic extreme rainfall data in Venezuela.

Bulletin 17B recommends weighted moments estimators for use with the LP3 distribution when historical information is available. Cohn et al. (1997) propose the Expected Moment algorithm (EMA) as an alternative to the weighted moment method. EMA was extended to include regional skew information in Griffis et al. (2004).

Kuczera (1996) discusses problems posed by measurement errors and uses a maximum likelihood procedure to incorporate the distribution of possible measurement errors into an analysis for the three-parameter GEV distribution. Subsequently, Kuczera (1999) used importance sampling with Monte Carlo simulation to conduct a full Bayesian analysis of a record with systematic and historical information. (His algorithm is described below.) O'Connell et al. (2002) also took a full Bayesian approach to incorporate historical information and measurement errors into their flood frequency analysis. O'Connell et al. (2002) searched for the primary modes of the parameters using simulated annealing and the simplex method. The credible intervals of the parameters are calculated numerically through a sophisticated adaptive grid algorithm to support numerical integration of the posterior distribution of the three parameters over their infinite range of possible values.

This paper employs a fully Bayesian approach for flood frequency analysis with historical information and measurement error. Bayesian Markov Chain Monte Carlo (MCMC) methods are employed to derive an empirical approximation of the posterior distribution of parameters, flood quantiles and other functions of the parameters of both the log-normal and Log-Pearson Type 3 distributions. Relative to MLE-based approximations, a full Bayesian analysis provides a more accurate description of flood risk and parameters uncertainties, and more realistic credible intervals for both flood quantiles and flood damages for data sets with historical information reflecting the joint distribution of possible measurement errors. The use of Bayesian MCMC to fit a two-parameter log-normal distribution or a three-parameter Log-Pearson Type 3 distribution with systematic and historical information is attractive because it: (i) allows a full Bayesian analysis of the data including appropriate descriptions of the uncertainty in any function of the parameters, (ii) it allows a wide range of descriptions of the joint distribution of uncertainties in the data including the magnitude of historical peaks and exceedance thresholds, and (iii) the numerical procedures are relative straightforward without significant increase in complexity or computational effort with additional descriptions of data uncertainty, nor does it require that the parameter space be represented by a grid for numerical integration.

## 2. Parameter estimation with historical information

This section describes four methods that have been used to estimate the parameters of a distribution with flood records that may include historical information. The Bayesian approach is introduced last and its computational implementation with Monte Carlo algorithms is described. The focus of this paper is the performance of the Bayesian methods and their ability to represent hydrologic data and its precision, and to improve upon the characterization of uncertainty provided by maximum likelihood estimators.

### *2.1. Weighted moments*

Bulletin 17B (USWRC, 1982) describes procedures employed by US federal agencies for flood frequency analysis. It recommends the use of the adjusted-moment estimator to fit the Log-Pearson Type 3 distribution when historical information is available. It requires that all historical floods that exceed a threshold of perception have a numerical value to represent their magnitude. The values of unobserved historical floods below the perception threshold are represented by the sample average and variability of below-threshold floods during the system record period.

The performance of this method has been explored by Stedinger and Cohn (1986) and Cohn et al. (1997). Both studies show that the adjusted moments method is not as good as competing methods, such as maximum likelihood analysis. Thus attention here will focus on Bayesian methods as an extension of the likelihood principle upon which MLEs are based.

### *2.2. Expected moment*

Cohn et al. (1997) present the expected moments algorithm (EMA), which is another moment-based quantile estimator developed to use systematic and historical information. EMA uses an iterative procedure to compute parameter estimates that reflect both the systematic and historical flood records. They report Monte Carlo results showing that EMA performs better than the adjusted-moment estimator when fitting a Log-Pearson Type 3 distribution, and as well as MLEs in a case where it regularly converged. Cohn et al. (2001) provides approximate variance estimates that can be used to compute confidence intervals for quantiles correcting for the correlation between the quantile and standard deviation estimators. Monte Carlo experiments demonstrate that the confidence intervals did cover the actual quantiles with nearly the target frequency.

EMA is an attractive alternative to the weighted moment estimators currently recommended by Bulletin 17B, particularly for the LP3 distribution for which MLEs have trouble. Interest here is to see if Bayesian methods can also overcome the problems with MLEs. Bayesian methods naturally incorporate a wide range of data uncertainties and provide the posterior distribution of any function of the parameters without use of the specialized corrections for EMA provided by Cohn et al. (2001).

### *2.3. MLE approach*

Maximum likelihood estimators are one of the standards in statistical inference because of their theoretical motivation and asymptotic efficiency for distribution satisfying several regularity conditions (Bickel and Doksum, 1977). Stedinger and Cohn (1986) develop maximum likelihood estimators that incorporate historical and paleoflood information into a flood frequency analysis. Their Monte Carlo results for the log-normal distribution demonstrate that the MLE procedure is flexible and robust, and more efficient than the adjust-moment estimator suggested by Bulletin 17. The maximum likelihood representation of the problem proceeds as follows.

Suppose one has $s$ years of systematic observations denoted by $\{x_1, \ldots, x_s\}$, and knows both the number of floods $k$ that exceeded a perception threshold $X_0$ over an $h$ year period and the magnitudes of those historical floods, which are denoted $\{y_1, \ldots, y_k\}$. When the annual floods are independent events, the likelihood function for this data is simply the product of three terms, the probability of seeing the systematic data, the probability of observing $k$ floods above $X_0$ in $h$

years, and the likelihood of seeing $\{y_1, \ldots, y_k\}$ given they were greater than $X_0$

$$\ell(D|\theta) = \prod_{i=1}^{s} f_X(x_i) \left\{ \begin{bmatrix} h \\ k \end{bmatrix} F_X(X_0)^{(h-k)} [1 - F_X(X_0)]^k \right\} \times \prod_{j=1}^{k} f_X(y_j) \quad (1)$$

where $D$ represent the data and $\theta$ represents the parameters, $f_X(\ )$ and $F_X(\ )$ are the probability density function (pdf) and the cumulative density function for $X$, and $f_Y(\ )$ is the pdf for the historical floods. Because we know $Y$ is greater than $X_0$:

$$f_Y(y) = \frac{f_X(y)}{[1 - F_X(X_0)]}$$

With this substitution, the likelihood function becomes

$$\ell(D|\theta) = \begin{bmatrix} h \\ k \end{bmatrix} F_X(X_0)^{(h-k)} \prod_{i=1}^{s} f_X(x_i) \prod_{j=1}^{k} f_X(y_j) \quad (2)$$

In some cases, due to the recording limitations on the estimation of the magnitudes of the historical floods, one may prefer to rely only on the observation that $k$ floods exceeded the threshold $X_0$ in $h$ years. Stedinger and Cohn (1986) called this type of representation of information binomial-censored data.

Omitting the third term in (1) yields the likelihood function for binomial-censored data:

$$\ell(D|\theta) = \prod_{i=1}^{s} f_X(x_i) \left\{ \begin{bmatrix} h \\ k \end{bmatrix} F_X(X_0)^{(h-k)} [1 - F_X(X_0)]^k \right\} \quad (3)$$

The MLEs are the values of the parameters that maximize the likelihood function. However, finding the best point estimate of the parameters is not enough. Quantifying parameter uncertainty is also important.

The uncertainty in the MLE estimators is usually estimated through a quadratic approximation of the likelihood function which represents the inverse of the Fisher Information matrix (Bickel and Doksum, 1977). The quadratic approximation of the log likelihood function for LN2 with censored data is reported in Appendix. The performance of this approximation is evaluated in Section 3.

A drawback of the MLE approach is that it does not always work well, such as when fitting a Log-Pearson Type 3 distribution (Cohn et al., 2001). Bobee and Ashkar (1991) show that with the first-order conditions for a local maximum of the likelihood function for the LP3 distribution, $\alpha$ is never less than one. However, the LP3 distribution can always be made to go to infinity for any $0 < \alpha < 1$ by letting the lower bound approach the smallest observation (largest for $\beta < 0$). Some authors have reported failures on finding a local and reasonable maximum of the likelihood, especially for small sample sizes. Hirose (1995) showed there may be more than one local maximum for a given sample size. Kuczera (1996) and O'Connell et al. (2002) show that when measurement error is present, the likelihood function may be multi-modal. It is hoped that use of the likelihood function in a Bayesian framework may result in good estimators because a Bayesian analysis is based upon the whole likelihood function, and not just the maximum which may not be representative of the character of the entire likelihood function.

### 2.4. Bayesian inference and Monte Carlo simulation

Bayesian inference is an alternative to the classical statistical inference. With a Bayesian approach, our understanding of the likelihood the parameters have different values is described by a probability density function. A Bayesian analysis combines the information in the data represented by the entire likelihood function with prior knowledge about the parameters, which may come from other data sets or a modeler's experience and physical intuition. Parameter estimation is made through the posterior distribution which is computed using Bayes' Theorem

$$p(\theta|D) = \frac{\ell(D|\theta)\xi(\theta)}{\int \ell(D|\theta)\xi(\theta)\mathrm{d}\theta} \quad (4)$$

where $p(\theta|D)$ is the posterior distribution of the parameters $\theta$, $\ell(D|\theta)$ is the likelihood function, and $\xi(\theta)$ is the prior distribution of $\theta$. The denominator is a normalizing constant that scales the posterior so that the area under the posterior pdf equals one.

Providing a full posterior distribution of the parameters is an advantage of the Bayesian approach over classical methods, which usually provide only a point estimate of the parameters represented by the mode of the likelihood function, and make use of asymptotic normality assumptions and a quadratic approximation of the log-likelihood function to describe uncertainties. With the Bayesian framework, one does not have to use any approximation to evaluate the uncertainties because the full posterior distribution of the parameters is available. Moreover, a Bayesian analysis can provide credible intervals for parameters or any function of the parameters which are more easily interpreted than the concept of confidence interval in classical statistics (Congdon, 2001).

Kuczera (1996) shows that correlated measurement errors from extrapolation of the rating curve beyond the range of flow measurements decrease the precision of estimated flood quantiles. Potter and Walker (1985), Kuczera (1996), O'Connell et al. (2002) and Carling et al. (2003) provide estimates of the uncertainty in estimated flood peaks. For instance, interpolation within the range of flow measurements yields errors of about 1–5% (error coefficient of variation), while extrapolating increases the uncertainty up to 30% in poor situations. Estimating peak discharge of historical floods by hydraulic calculations have errors of about 15–20% and may reach 25% for paleostage measurements. O'Connell et al. (2002) use discrete three-point error distributions to define a likelihood function to include such uncertainties in their Bayesian analysis.

Sometimes the errors in different flood measurements are independent. However, a common and important case is when rating-curve errors affect the computed value of several flood peaks and flood exceedance thresholds (Kuczera, 1999). This case requires integrating the entire likelihood function over the distribution of the possible errors in the rating curve or flood routing model to yield the appropriate likelihood function

$$\ell(D|\theta) = \int \ell(\{x_i(\omega)\}, \{y_i(\omega)\}, X_0(\omega)|\theta) g(\omega) \mathrm{d}\omega$$

where the vector $\omega$ describes the possible errors in individual flood measurements and the uncertain in parameters of the rating curve (or routing model) that affect the computed flow values $[\{x_i(\omega)\}, \{y_i(\omega)\}, X_0(\omega)]$, associated with stage records for the systematic floods, historical floods and the perception threshold; here $g(\omega)$ is the pdf for the parameter vector $\omega$ describing possible errors in the stage-discharge function and in individual measurements. Such a computation can easily be incorporated with a MCMC analysis with relatively little impact on the computational effort. One just adds $\omega$ as an additional vector of unknown parameters with prior $g(\omega)$, and the MCMC analysis will provide the required computation.

Computing the denominator of Eq. (4) may not be easy. There are simple cases for which one can choose a conjugate prior to obtain a posterior that is analytically tractable (Zellner, 1971). In these cases, the normalizing constant can be easily computed. When choosing a conjugate prior is not possible, the normalizing constant can be computed numerically in low-dimensional cases, as illustrated by O'Connell et al. (2002). However, in more complicated cases where the dimension of $\omega$ and the parameter vector $\theta$ is large, computing the normalizing constant can be computationally infeasible. Such problems had limited the use of the Bayesian approach for some time (Gelman et al., 1995).

The rapid development of computers in the last two decades provided a tool for numerically intense methods in statistical inference, including Markov Chain Monte Carlo (MCMC) methods. MCMC samples values of the parameters from the posterior distribution without computing the normalizing constant. The two most popular MCMC algorithms are the Gibbs sampler, named by Geman and Geman (1984) and examined by Casella and George (1992), and the Metropolis-Hastings algorithm based on the papers by Metropolis et al. (1953) and Hastings (1970).

Standard and efficient MCMC routines for uncensored and censored normal samples are used in Section 3.2 to simulate the posterior distribution of the parameters of the two-parameter log-normal distribution. The Metropolis–Hastings algorithm is used in Section 4.2 to generate the posterior distribution of the parameters of the LP3 distribution. With the LP3 care is taken to employ proposal distributions that provide an efficient simulation with a description of the likelihood function valid with log-space skew coefficients near zero to avoid the numerical problems that would result had the natural parameterization of the gamma density function been employed.

Tierney (1994), Chib and Greenberg (1995) and Gelman et al. (1995) provide a theoretical description of the MCMC algorithm. The Metropolis–Hastings algorithm simulates a Markov Chain in a Monte Carlo study to generate a set of points whose distribution converges to the posterior distribution. Let $\theta_j$ be the vector of parameters at the $j$th iteration. Then one employs a proposal transition distribution $q(\theta_j,\theta_{j+1})$ that satisfies the reversibility condition

$$\pi(\theta_j)q(\theta_j,\theta_{j+1}) = \pi(\theta_{j+1})q(\theta_{j+1},\theta_j) \tag{5}$$

where $\pi(\theta)$ is the desired posterior distribution for $\theta$. The reversibility condition is a sufficient condition for $\pi(\theta)$ be the equilibrium distribution of the chain. This condition basically requires that the unconditional probability of moving from one set of parameters $\theta_j$ to a new set $\theta_{j+1}$ must be equal to the unconditional probability of moving from $\theta_{j+1}$ to $\theta_j$. For any trial proposed distribution, a transition distribution that satisfies (5) is obtained by assigning probabilities $\alpha$ of moving from $\theta_j$ to $\theta_{j+1}$ and vice-versa so that the reversibility condition is satisfied

$$\pi(\theta_j)q(\theta_j,\theta_{j+1})\alpha(\theta_j,\theta_{j+1})$$
$$= \pi(\theta_{j+1})q(\theta_{j+1},\theta_j)\alpha(\theta_{j+1},\theta_j), \quad \theta_j \neq \theta_{j+1}$$

Chib and Greenberg (1995) show that the needed probability $\alpha$ is

$$\alpha(\theta_j,\theta_{j+1}) = \min\left[\frac{\pi(\theta_{j+1})q(\theta_{j+1},\theta_j)}{\pi(\theta_j)q(\theta_j,\theta_{j+1})}, 1\right] \tag{6}$$

The probability of a move is also called the acceptance probability. In Eq. (6), it depends on the ratio between the values of the posterior distribution at the two points. The trial proposal distribution $q$ generates only a possible set of new parameters, which are accepted or rejected depending on the ratio in Eq. (6). If the unconditional probability of moving from $\theta_j$ to $\theta_{j+1}$ is larger than the unconditional probability of moving in the other direction, then the move is accepted, otherwise the proposal is accepted with probability $\alpha$. If the move is not accepted, the parameters retain their values from the last iteration.

The choice of an adequate proposal distribution $q$ is key to an efficient implementation of the Metropolis–Hastings algorithm. One wants a proposal distribution that gives a reasonable acceptance rate, which is the average number of acceptances in the simulation, and generates a series that covers the entire parameters space so that the final sample accurately describes the posterior distribution. If the proposal distribution nominates values of the parameters that are very far from the current position, it is likely it will result in a low acceptance rate and the simulation procedure will be inefficient. On the other hand, if it only proposes values very close to the current position, the acceptance rate may be high, but it will take a long time to cover the parameter space thoroughly and again the simulation procedure will be inefficient. In both extreme cases, one sees a highly correlated series of values which results in an uncertain estimate of the distribution of the parameters for a fixed simulation run length (Gamerman, 1997).

After sampling the parameters one can estimate the marginal density distributions, compute means and standard errors, and estimate credible intervals not only of the parameters but also of any function of them, such as desired quantiles and flood damages.

### 2.5. Importance sampling and measurement error

Kuczera (1999) presents a conceptually similar approach for evaluating the posterior distribution of the parameters and other quantities; instead of using a general Markov Chain Monte Carlo simulation to evaluate the posterior distribution of such functions, he uses a specialized and potentially more efficient importance sampling algorithm (Gamerman, 1997). For this purpose he develops a quadratic approximation of the posterior probability density function which is used as the basis of a multivariate normal approximation. To evaluate the expected value of any function $h(q|\theta)$ of the flow $q$ for parameter values $\theta$, where $h$ may be the exceedance probability of $q$ or the pdf at $q$, one needs to be able to compute:

$$E\{h(q|\theta)\} = \int h(q|\theta)p(\theta|D)\mathrm{d}\theta \tag{7}$$

The trick to facilitate a Monte Carlo evaluation of (7) is to write the required integration as

$$\int h(q|\theta)p(\theta|D)\mathrm{d}\theta = C\int h(q|\theta)l(D|\theta)\xi(\theta)\mathrm{d}\theta = C\int h(q|\theta)\frac{l(D|\theta)\xi(\theta)}{f(\theta)}f(\theta)\mathrm{d}\theta \quad (8)$$

where $C$ is a normalizing constant corresponding to the denominator in (4), and $f(\theta)$ is the pdf of an 'importance' distribution chosen to approximate $p(\theta|D)$. Kuczera uses a multivariate normal distribution for $f(\theta)$ with an inflated variance to ensure good coverage. His Monte Carlo algorithm draws a large number of points $\theta_i$ from the distribution $f(\theta)$, and assigns to each the probability

$$p_i = w_i/S \text{ where } w_i = \frac{\ell(D|\theta_i)\xi(\theta_i)}{f(\theta_i)} \text{ and } S = \sum w_i \quad (9)$$

Thus $(\theta_i, p_i)$ jointly represent the posterior distribution of $\theta$, and the computation of the mean value of $h(q|\theta)$ at these points with the assigned probabilities yields a convenient numerical approximation of the value of the integral $E\{h(q|\theta)\}$. The computation of $S$ provides the required normalization of the likelihood function and corresponds to the computation of $C$; this step introduces some bias in small samples because the estimator of the integral in (7) using (8) will be the ratio of two random variables $[\sum w_i h(q|\theta_i)]$ and $S$ (Gamerman, 1997).

The efficiency and accuracy of this importance sampling approach depends upon how well $f(\theta)$ approximates the posterior distribution of the parameters; the most critical issue is that $f(\theta)$ covers all of the likely values of $\theta$, otherwise a few of the probabilities $p_i$ can be very large (Gelman et al., 1995). As shown by examples in Kuczera (1996, 1999), and the examples below, the likelihood function need not have the shape of a multivariate normal distribution, and thus the appeal of the more general and flexible MCMC simulation methods employed here, and by Parent and Bernier (2003).

If the likelihood function is defined by an integral over the measurement-error vector $\omega$, then this importance sampling approach would still work if one simulates jointly the values of $\theta$ and $\omega$ using a importance distribution $f(\theta,\omega)$ that well approximates $\ell(D|\theta,\omega)\xi(\theta)g(\omega)$ wherein $g(\omega)$ is the pdf of the measurement-error distribution. For example, MCMC analysis with $\theta$ and $\omega$ treated as parameters would for any function $h(q|\theta)$ evaluate the integral

$$\begin{aligned} E_{\theta,\omega}\{h(q|\theta)\} &= \int h(q|\theta)p(\theta,\omega|D)\mathrm{d}\theta\,\mathrm{d}\omega \\ &= \iint h(q|\theta)(C)\ell(\{x_i(\omega)\},\{y_i(\omega)\}, X_0(\omega)|\theta)\xi(\theta)g(\omega)d\theta\,d\omega \\ &= \int h(q|\theta)(C)\left[\int \ell(\{x_i(\omega)\},\{y_i(\omega)\}, X_0(\omega)|\theta)g(\omega)\mathrm{d}\omega\right]\xi(\theta)\mathrm{d}\theta \end{aligned} \quad (10)$$

where this last expression clearly is the quantity that is required in that it includes the average over the distribution of $\omega$ of the likelihood function for the data, and where $C$ is the required normalization constant from Eq. (4):

$$\begin{aligned} 1/C &= \int p(\theta,\omega|D)\mathrm{d}\theta\,\mathrm{d}\omega \\ &= \int\left[\int \ell(\{x_i(\omega)\},\{y_i(\omega)\}, X_0(\omega)|\theta)g(\omega)\mathrm{d}\omega\right]\xi(\theta)\mathrm{d}\theta \end{aligned} \quad (11)$$

## 3. LN2 example

This first and simple example illustrates the differences between the Bayesian approach and the classical statistical method, which finds the point estimate through MLE and then assumes that the uncertainties can be approximated by a quadratic function. The analysis considers the Bayesian and asymptotic-MLE descriptions of the uncertainty in the parameters and quantiles of a LN2 distribution with historical information.

### *3.1. Description of the data*

A data set with 120 years of observations was generated using a log-normal distribution with a log-space (natural logarithms) mean $\mu=6.9$ and log-space standard deviation $\sigma=0.83$. Thus the coefficient of variation CV equaled 1 and the real-space mean is equal to 1000 (Stedinger, 1980). Only the last 20 years were considered to be systematic data, and the censoring threshold of the earlier historical period was chosen to be equal to the 99th quantile ($X_0=6842$). In the sample, two floods exceeded

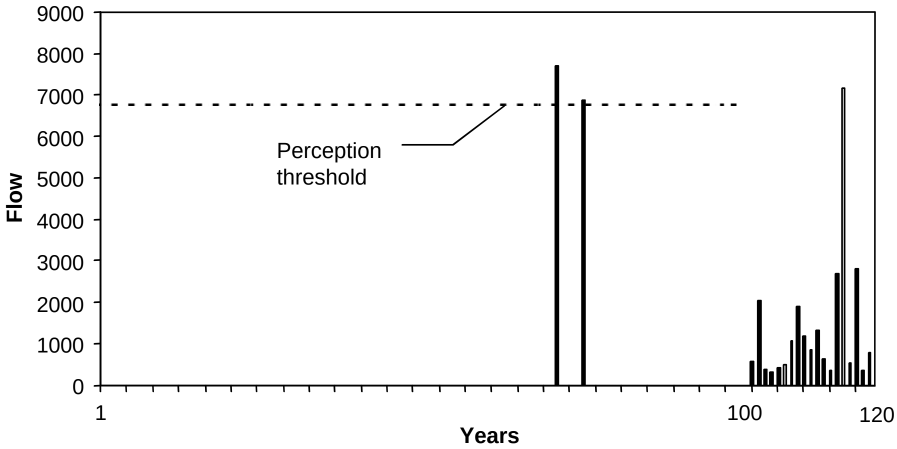


Fig. 1. Flood record in real space with systematic and historical data. $N=20$, $h=100$, $k=2$, and the perception threshold is equal to the 99th quantile.

the 99th quantile in the 100 year historical period, and a third flood exceeded the threshold in the systematic-record period, as can be seen in Fig. 1. The magnitudes of those three floods were 6890, 7720, and 7180, respectively.

In order to illustrate the consequences of using the asymptotic MLE and Bayesian approaches in flood frequency and uncertainty analyses, we examined the estimated uncertainties in the estimator of the flood quantiles and of the expected value of flood damages employing the damage function $D(q)$ developed in Al-Futaisi and Stedinger (1999). For a flood of magnitude $q$ in a hypothetical basin

$$D(q) = \alpha_D(q^{3/5} - q_0^{3/5})\lfloor H + (q^{3/5} - q_0^{3/5})\rfloor \quad (12)$$

where $\alpha_D=24$, $q_0$ is the discharge threshold above which damage begins and is set equal to the 80th quantile, and $H=300$ (see Al-Futaisi and Stedinger (1999) for details). The expected value of flood damage is estimated by numerical evaluation of the integral using the pdf for $q$:

$$E[D(\theta)] = \int_{q_0}^{\infty} D(\tilde{q}) f(\tilde{q}|\theta)\mathrm{d}\tilde{q} \quad (13)$$

### 3.2. MLE and asymptotic normality

Fig. 2 shows the contour lines of the likelihood function for two different situations, one with only the systematic data and the other in which the number and magnitudes of two historical floods ($k=2$) above the perception threshold (99th quantile) during the historical period ($h=100$ years) is included to the likelihood function (censored data as in Eq. (2)). The values of the likelihood function were scales so that the maximum value is equal to the unity. The first contour line is equal to 0.98 and gives a good idea of the location of the maximum. Subsequent contour lines correspond to decrease of a factor of 3.16 in the likelihood function (every two contour lines is a factor of 10). While better behaved than the LP3

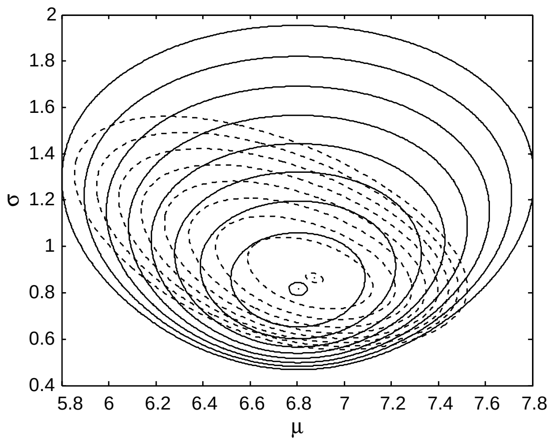


Fig. 2. Contour lines of the likelihood function of the LN2 distribution. Solid lines represent the likelihood with only 20 years of systematic data, while the dotted lines represent the likelihood with censoring data with $h=100$ and $k=2$.

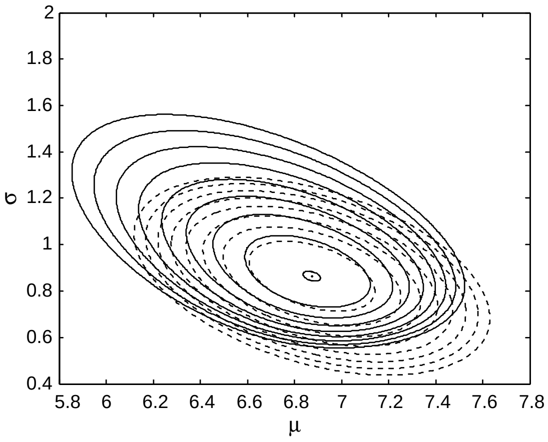


Fig. 3. Solid lines represent the likelihood function of LN2 distribution with censored data. Dotted lines represent the quadratic approximation of the likelihood function.

contour lines displayed in Kuczera (1999), one can still see that in both cases the likelihood function is far from being concentric ellipses as assumed by an asymptotical second-order analysis. This occurs in part because the population variance $\sigma^2$ is bounded below by zero, resulting in a positively skewed likelihood function. However, the uncertainty in $\mu$ also increases with the value of $\sigma^2$.

This graph also shows that the censored data has a dramatic impact on the likelihood function, which becomes even less quadratic. Fig. 3 illustrates how poor the quadratic approximation is in this specific case. The contour lines of the quadratic approximation of the likelihood function are plotted over the contours of the likelihood function for censored data. The asymptotic quadratic approximation misrepresents the real likelihood, and thus may provide a poor description of parameter uncertainty.

### 3.3. Bayesian MCMC

The Bayesian parameter estimation of the LN2 with historical information was done using the software Winbugs (Spiegelhalter et al., 2000). This requires a prior distribution for the parameters and the likelihood function. Winbugs has a built-in likelihood function for uncensored and censored normal data that was used in these simulations.

To check on the role the prior distribution in the estimation process, two different prior distributions were considered. One with an almost non-informative prior for both $\mu$ and $\sigma$, where $\mu$ was normal with mean zero and variance 1000, and $\sigma$ had a gamma distribution with mean 1 and variance 100. With the other prior, $\sigma$ had a more informative prior described by a gamma distribution with mean 0.7 and standard deviation 0.3. This more informative prior for $\sigma$ is consistent with the values of the coefficient of variation of annual maximum discharges in American rivers that varies from 0.20 to 1.50 (Landwehr et al., 1978). This prior for $\sigma$ assigns only a 10% probability the CV is outside of (0.30,1.96).

For both priors, Winbugs was used to generate 50,000 realizations of $\mu$ and $\sigma$ from the posterior distribution. The last 40,000 were used in the estimation of the mean, standard deviation, and the marginal density distribution of the parameters and quantiles. In order to facilitate visualization in Fig. 4, that sample was thinned to 5000 values by taking every eighth value.

Fig. 4 presents the contour lines of the posterior distribution with censored data and the MCMC sample when one uses a non-informative prior for both $\mu$ and $\sigma$. A picture that corresponded to use of the informative prior was very similar—the most significant difference being the decreased likelihood

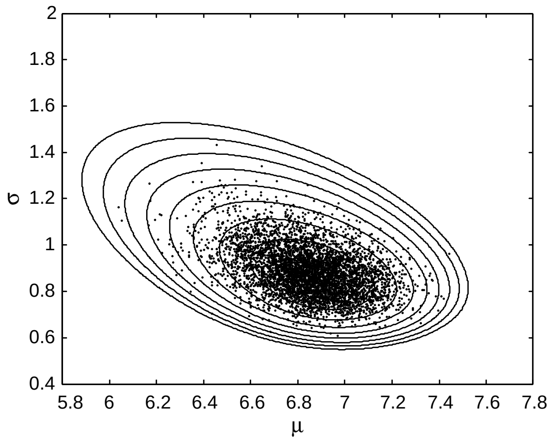


Fig. 4. Posterior distribution of the LN2 with censored data and MCMC sample of parameters with non-informative prior on both $\mu$ and $\sigma$.

Table 1
Results of parameter estimation using MLEs and Bayesian MCMC

| Method | $\mu$ | | | $\sigma$ | | | $Q_{99}$/1000 | | | $E$(damage)/1000 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2.5% | Mode/mean[a] | 97.5% | 2.5% | Mode/mean[a] | 97.5% | 2.5% | Mode/mean[a] | 97.5% | 2.5% | Mode/mean[a] | 97.5% |
| MLE quadratic approx. | 6.53 | 6.87 | 7.22 | 0.67 | 0.86 | 1.06 | 4.7 | 7.2 | 11.2 | 11.8 | 65.8 | 119.8 |
| MCMC non-informative prior | 6.46 | 6.84 | 7.17 | 0.71 | 0.89 | 1.13 | 4.8 | 7.6 | 12.2 | 26.6 | 72.7 | 149.0 |
| MCMC informative prior on $\sigma$ | 6.48 | 6.86 | 7.18 | 0.70 | 0.87 | 1.09 | 4.7 | 7.4 | 11.6 | 25.9 | 69.9 | 142.2 |

[a] Mode for MLE and mean for MCMC.

of very small $\sigma$ values. One can see that the sample generated by Winbugs represents the posterior distribution very well, recognizing that the probability of seeing a point within a ring is the product of the posterior probability, described by the contour lines, and the width of the ring.

Table 1 presents the results of the parameter, quantile, and expected value of flood damage estimation using both MLE along with quadratic-likelihood approximation to determine uncertainties (following Cohn (1984)), and the Bayesian approach with the two different priors. One can see that uncertainties in both the parameters and flood quantiles given by the Bayesian analysis with a non-informative prior are larger than those obtained using a quadratic approximation of the likelihood function. The differences are most apparent in the upper bound on the standard deviation $\sigma$ and the 100-year flood $Q_{0.99}$, except for $\sigma$ where the Bayesian result with a more informative prior on $\sigma$ provides a credible interval about the same size as the confidence interval.

Numerical estimates of the partial derivatives of the expected damage function in (13) with respect to the two parameters were used with a first-order/second moment approximation of the variance of the estimated damages to compute an approximate 95% confidence interval of the expected flood damages based upon a normal error distribution (Benjamin and Cornell, 1970; Cohn, 1984). The linear approximation is not an accurate description of the nonlinear expected damage function. Therefore, the Bayesian MCMC credible interval of the expected damage, which is estimated using the 40,000 values of $\mu$ and $\sigma$, provides a better representation of the uncertainty distribution of the expected flood damages. Uncertainty intervals for various quantiles are presented in Fig. 5, which shows the 95% credible interval generated by MCMC with non-informative prior, and the 95% confidence interval given by the quadratic approximation for flood quantiles with different exceedance probabilities.

The use of a more informative but modest prior on $\sigma$ lead to slightly tighter credible intervals on all parameters, flood quantiles, and flood damages in Table 1. The largest differences occurred for expected damages, which mostly depends upon rare floods with exceedance probabilities less than 5%. Moreover the Bayesian credible intervals for expected damages are very asymmetric, being almost twice as long above the mean as below the mean, whereas the use of a normal approximation based upon the asymptotic variance of the MLE results in a less-appropriate symmetric confidence interval.

## 4. Log-Pearson Type 3 distribution

This section provides a Bayesian analysis for the LP3 distribution using a Metropolis–Hastings algorithm. First, the LP3 distribution is described. Then, the prior distributions for the parameters, the likelihood function of the data, and the choice of an adequate proposal distribution are presented. Finally, the method is applied to several samples. Because MLEs can have trouble when fitting the three-parameter LP3 distribution, it would be convenient if Bayesian procedures were able to deal with this distribution while still using the likelihood function for the data. A comparison of credible regions generated by the Bayesian analysis with confidence intervals generated using asymptotic results shows the problems with asymptotic approximations, as was to be expected because of the nonlinear relationship

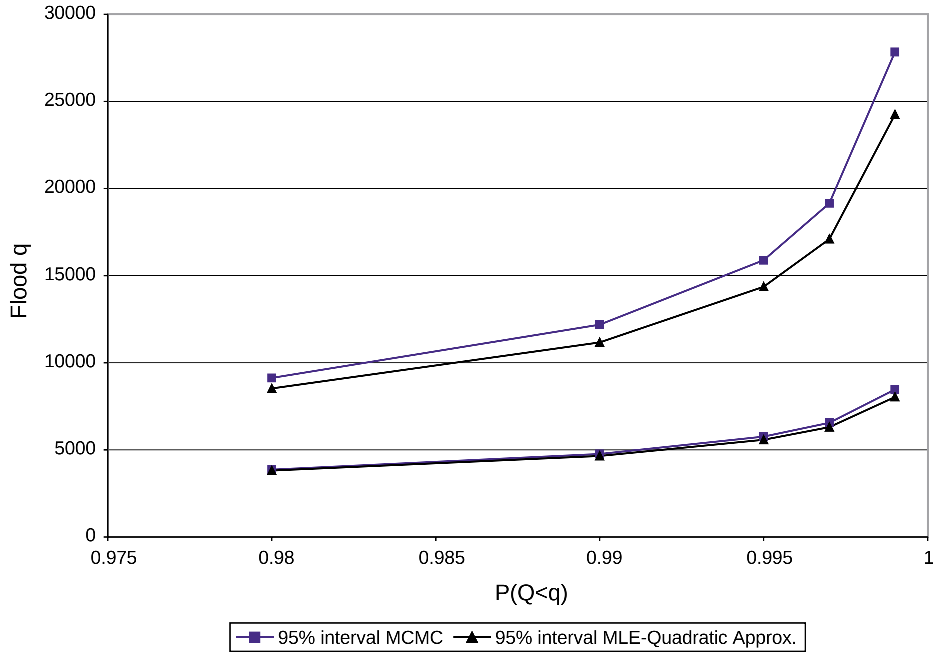


Fig. 5. Ninety-five percent confidence interval of flood quantiles given by the quadratic approximation of the likelihood function and 95% credible interval of flood quantiles given by Bayesian MCMC.

between the highly uncertain skew parameter and flood quantiles and expected flood damages.

### 4.1. The model

If floods $\{Q_1, \ldots, Q_N\}$ are distributed as a Log-Pearson Type 3 (LP-3) distribution, then $X=\log(Q)$ has a Pearson Type 3 variate with probability density function

$$f_X(x) = \frac{|\beta|}{\Gamma(\alpha)}[\beta(x-\tau)]^{\alpha-1}\mathrm{e}^{-\beta(x-\tau)} \tag{14}$$

where $\alpha$, $\beta$, and $\tau$ are the shape, scale and location parameters, respectively, and $\Gamma(\alpha)$ is the gamma function. If $\beta>0$, the distribution is positively skewed and $\tau$ is a lower bound; if $\beta<0$, the distribution is negatively skewed with an upper bound $\tau$. As the shape parameter $\alpha$ goes to infinity, the skew coefficient goes to zero, and the Log-Pearson Type 3 reduces to the two-parameter log-normal distribution. Because interest is often in distributions with small skews, numerical problems can result when fitting an LP3 distribution with MLEs.

Another parameterization based on the mean, standard deviation, and skewness is often used to calculate the $p$th quantile

$$x_p = \mu + \sigma K_p(\gamma) \tag{15}$$

where $K_p(\gamma)$ is the frequency factor which is the $p$th quantile of the P3 distribution with mean zero and standard deviation 1, and skewness $\gamma$. The frequency factor $K_p$ can be well approximated by the Wilson–Hilferty transformation (Kirby, 1972) for $|\gamma|<2$ and $0.01\le p\le 0.99$

$$K_p(\gamma) = \frac{2}{\gamma}\left(1+\frac{\gamma z_p}{6}-\frac{\gamma^2}{36}\right)^3 - \frac{2}{\gamma} \tag{16}$$

where $z_p$ is the $p$th quantile of the standard normal distribution.

### 4.2. A Bayesian framework with a MCMC procedure

For historical and other censored data, the Bayesian approach is an alternative to the adjusted-moment procedure recommended by Bulletin 17B and to EMA. Bayesian inference uses the likelihood function, which is an effective and flexible way to employ a wide range of data types as well as including measurement errors in the analysis. Here we use the Metropolis–Hastings algorithm to generate a sample that represents the posterior distribution of the parameters $\mu$, $\sigma$, $\gamma$, and of the quantiles $x_p$ and damages using Eqs. (12) and (13) This requires a prior and proposal distributions for the parameters, and the likelihood function.

*4.2.1. Prior distributions*

For illustration we use relatively non-informative priors for $\mu$, $\sigma$, and a more informative prior for $\gamma$. Because $\mu$ can be negative, the prior distribution is represented by a normal distribution with mean zero and variance 1000. The standard deviation $\sigma$ must be strictly positive. Zellner (1971) discusses the use of non-informative prior for $\sigma$ and suggests it should be proportional to its reciprocal. Previous studies and physical intuition indicate that the skewness coefficient should be well within $\pm 1.4$, which are plausible bounds on the skewness coefficient. Were population skews uniformily distributed on $\pm 1.0$, their variance would be 0.33. The adopted prior distribution for $\gamma$ was a normal distribution with mean zero and variance 0.3, though this number is certainly too large a variance (Tasker and Stedinger, 1986; Reis et al., 2003)

$$\xi(\mu) \sim N(0, 10000) \tag{17}$$

$$\xi(\sigma) \propto \frac{1}{\sigma}$$

$$\xi(\gamma) \sim N(0, 0.3)$$

For a non-informative $\gamma$-prior we used $\xi(\gamma) \sim N(0, 10{,}000)$.

*4.2.2. Likelihood function*

Section 2.3 describes a likelihood function that incorporates historical information. Eqs. (2) and (3) represent the likelihood function for censored and binomial-censored data, respectively. It requires calculation of $f_X(x)$ of all observations and the cdf evaluated at the perception threshold $F_X(x_0)$. Using the Wilson–Hilferty transformation to relate the Pearson variate $X$ to the standard normal variable $z$ for modest skews $\gamma$ yields

$$z = \frac{\gamma}{6} + \frac{6}{\gamma}\left[\frac{\gamma}{2}\left(\frac{X-\mu}{\sigma}\right) + 1\right]^{1/3} - \frac{6}{\gamma} \tag{18}$$

Thus, the LP3 probability density function for $X$ can be computed based on an adjustment to the standard normal probability density function for $z$

$$f_X(x) = \phi(z)\frac{\mathrm{d}z}{\mathrm{d}X} = \frac{\phi(z)}{\sigma\left[\frac{\gamma}{2}\left(\frac{x-\mu}{\sigma}\right) + 1\right]^{2/3}} \tag{19}$$

Substituting Eqs. (18) and (19) into either Eq. (2) or (3) gives the likelihood function to be used in the Metropolis–Hastings algorithm. This approximation is relatively well behaved near $\gamma = 0$ so that the entire algorithm is computationally stable.

Other mathematical approximations are available for the Pearson Type 3 or the three-parameter Gamma distribution. But the mathematical definition of its pdf involves the gamma function evaluated at $\alpha = (2/\gamma)^2$, which is very large for $\gamma$ approaching zero. Thus, use of the Wilson–Hilferty transformation avoids the numerical problems that use of the mathematical definition of the P3 pdf and cdf would have introduced. The sampling uncertainty associated with the parameters is much more important than any error introduced by this numerical approximation, which could be improved with approximations discussed by Kirby (1972) and Chowdhury and Stedinger (1991).

*4.2.3. Proposal distribution*

The proposed values of the three parameters $\mu$, $\sigma$, and $\gamma$ are generated independently depending on their values at the previous iteration. The proposal distribution for the mean $\mu$ is a normal with mean equal to $\mu_{t-1}$ and variance equal to $\sigma^2_{t-1}$ divided by the number of observations $N$. The number of observations $N$ in this case is equal to either $(s+k)$ if censored data is used or to $s$ when one uses binomial-censored data. Thus

$$\mu_t \sim N\left(\mu_{t-1}, \sigma^2_{t-1}/N\right) \tag{20}$$

The proposal distribution for $\sigma$ is a gamma distribution with mean equal to $\sigma_{t-1}$ and variance modeled as a function of both $\sigma_{t-1}$ and $\gamma_{t-1}$, as described in Stedinger and Tasker (1986)

$\sigma_t \sim \gamma(a, b)$

$$a = \frac{\sigma^2_{t-1}}{\mathrm{Var}(\sigma_{t-1})}; \quad b = \frac{\mathrm{Var}(\sigma_{t-1})}{\sigma_{t-1}} \tag{21}$$

$$\mathrm{Var}(\sigma_{t-1}) = \frac{\sigma^2_{t-1}(1 + 0.75\gamma^2_{t-1})}{2N}$$

The proposal distribution for $\gamma$ is a normal distribution with mean equal to $\gamma_{t-1}$ and variance modeled as a function of $\gamma_{t-1}$ and $N$, as described in Bulletin 17B and Tasker and Stedinger (1986).

$\gamma_t \sim N[\gamma_{t-1}, \mathrm{Var}(\gamma)]$

$$\mathrm{Var}(\gamma) = \left[1 + \frac{6}{N}\right]^2 10^{[a-b\ \log(N/10)]}$$

$$a = \begin{cases} -0.33 + 0.08|\gamma_{t-1}| & \text{if } |\gamma_{t-1}| < 0.90 \\ -0.52 + 0.30|\gamma_{t-1}| & \text{if } |\gamma_{t-1}| > 0.90 \end{cases} \tag{22}$$

$$b = \begin{cases} 0.94 - 0.26|\gamma_{t-1}| & \text{if } |\gamma_{t-1}| < 1.50 \\ 0.55 & \text{if } |\gamma_{t-1}| > 1.50 \end{cases}$$

As explained in Section 4.1, $\tau$ is either a lower or upper bound depending on the sign of $\beta$. Every time the proposed $\tau$ is either greater than the smallest observation when $\beta$ is positive, or is less than the largest observation when $\beta$ is negative, the proposed parameters are rejected because the likelihood function would be zero for such a set of parameters. Given the uncertainty in $\mu_{t-1}$, $\sigma_{t-1}$, and $\gamma_{t-1}$, these proposal distributions will have larger variance than the posterior distribution of the parameters and thus ensure that the algorithm explores the entire parameter space.

### 4.3. Results

The first example illustrates how poor the first-order asymptotic approximation of the variance of flood quantiles can be in the case of the LP3 distribution. A systematic sample with 25 observations was generated using a LP3 distribution. The sample mean, standard deviation, and skew are equal to 6.81, 1.17, and −0.66, respectively. No historical information was considered. Bayesian MCMC and method-of-moments along with first-order asymptotic approximations of the variance of quantiles were used to fit the data and estimate the 95% credible/confidence intervals. The asymptotic variances of flood quantiles were computed using a normal distribution with the variance formula provided by Chowdhury and Stedinger (1991) to be used when the skew is estimated by the at-site sample.

Fig. 6 displays the 95% credible/confidence intervals in the frequency plot of the data fitted by the Bayesian MCMC and method-of-moments. One can see the lower bound of the asymptotic confidence interval starts to decrease beyond an exceedance probability equal to 2%. Coles and Pericchi (2003, Fig. 2) illustrate a similar problem with the analysis of extreme rainfall in Venezuela. This means that our lower bound of the asymptotic confidence interval of the 100-year flood is smaller than the one of the 50-year flood. Clearly, the first-order asymptotic approximation of the variance of flood quantiles is not adequate. One can also see the Bayesian 95% credible interval provides a consistent result here, as it did for Coles and Pericchi (2003). The Bayesian

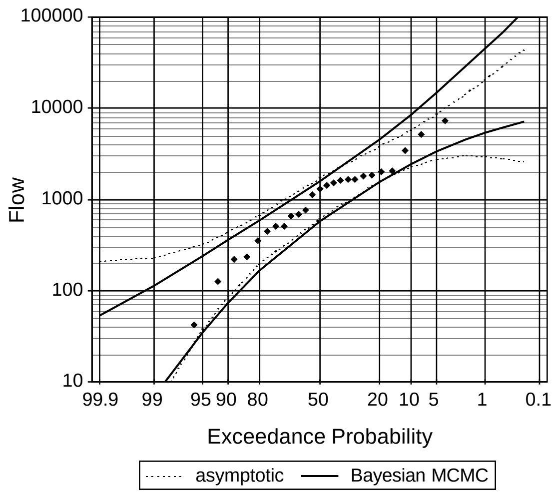


Fig. 6. Flood frequency plot of a LP3 distribution fitted by Bayesian MCMC and method-of-moments. Only systematic data was used. The solid lines represent the 95% credible interval on flood quantiles obtained by Bayesian MCMC while the dotted lines represent the 95% confidence interval of flood quantiles computed using first-order asymptotic approximation of the variance. The sample, generated by a LP3 distribution, consists of 25 years of systematic data with sample mean, standard deviation, and skew equal to 6.81, 1.17, and −0.66, respectively.

Table 2
Results of parameter estimation of LP3 with historical information using MLE, Bayesian MCMC and a modified MLE with a prior on $\gamma$

| Method | $\mu$ | | | $\sigma$ | | | $\gamma$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2.5% | Mode/mean[a] | 97.5% | 2.5% | Mode/mean[a] | 97.5% | 2.5% | Mode/mean[a] | 97.5% |
| MLE | | 6.67 | | | 0.82 | | | 1.85 | |
| Bayesian MCMC: non-informative prior on $\gamma$ | 6.46 | 6.76 | 7.08 | 0.63 | 0.82 | 1.08 | −0.28 | 1.02 | 2.10 |
| MLE: informative prior on $\gamma$ | 6.55 | 6.85 | 7.15 | 0.59 | 0.79 | 0.99 | −0.45 | 0.49 | 1.42 |
| Bayesian MCMC | 6.48 | 6.82 | 7.13 | 0.65 | 0.84 | 1.12 | −0.45 | 0.42 | 1.28 |

[a] Mode for MLE and mean for MCMC.

MCMC is not based on asymptotic approximations, but on the true posterior distribution of the flood quantiles thereby capturing the skewness of their posterior distribution.

The second example uses the same sample generated in Section 3 (Fig. 1) to illustrate the application of a full and accurate Bayesian approach for the LP3 distribution with historical information. Table 2 presents the results of the parameter estimation given by Bayesian MCMC and MLE. In the Bayesian analysis, 70,000 values of the parameters were generated though only 65,000 were used in the computation of the mean, standard error, and credible interval of the parameters and quantiles. The acceptance rate was about 29%. In the first case already presented, the acceptance rate was 33%. In a verification run with a fixed skew of zero (log-normal distribution) and only systematic data, the acceptance rate was 50%. Gamerman (1997) indicates that optimization of the trade-off between acceptance rate and coverage has generally resulted in acceptance rates in 20–50% range.

The maximum likelihood procedure had trouble finding a result that made sense: a local maximum of the likelihood function occurs where $\gamma$ has a value near 1.85 and a lower bound $\tau$ almost equal to the smallest observation. This is very near the boundary of the feasible parameter space and thus this maximum is not a good representation of the entire likelihood function. To compare the results of MLE and the Bayesian approach, a prior distribution for $\gamma$ was introduced into the MLE computation. This prior distribution is normal with zero mean and variance equal to 0.3, which gives less than a 1% chance that $|\gamma|$ is greater than 1.4. A second Bayesian simulation was also performed with a non-informative prior for $\gamma$ to show that even in cases where MLE finds a maximum that does not represent the likelihood function, Bayesian MCMC provides a more reasonable result.

Table 2 presents the results of parameter estimation using MLE with and without a prior distribution on $\gamma$, and Bayesian MCMC for those two cases. Table 3 presents the results of flood quantiles and expected flood damages. The expected flood damages were calculated according to (13). The uncertainties in both quantiles and expected flood damages in the MLE case were computed as in Section 3. (Pilon and Adamowski (1993), illustrate the asymptotic analysis.)

Comparing MLE with a prior distribution on $\gamma$ and Bayesian MCMC with a prior on $\gamma$, one sees big differences in $\sigma$ and $\gamma$, the 100-year flood,

Table 3
Credible intervals and confidence intervals for quantiles and flood damages

| Method | $Q_{99}$/1000 | | | $E$(damage)/1000 | | |
|---|---|---|---|---|---|---|
| | 2.5% | Mode/mean[a] | 97.5% | 2.5% | Mode/mean[a] | 97.5% |
| MLE | | 14.0 | | | 122.7 | |
| Bayesian MCMC: non-informative prior on $\gamma$ | 5.1 | 10.1 | 29.9 | 27.8 | 98.5 | 246.0 |
| MLE: informative prior on $\gamma$ | 4.4 | 7.8 | 14.0 | 2.2 | 65.1 | 128.0 |
| Bayesian MCMC | 4.8 | 8.2 | 16.4 | 25.6 | 78.5 | 179.9 |

[a] Mode for MLE and mean for MCMC.

and the expected flood damages. Bayesian MCMC provides an accurate and skewed credible interval for quantiles and flood damages. It is not based on asymptotic assumptions but on the actual posterior distributions.

Bayesian MCMC with a non-informative prior on $\gamma$ provides a larger mean for $\gamma$ than with an informative prior and higly skewed credible region describing $\gamma$ uncertainty. That increases the values of the mean of quantiles and expected flood damages as well as increasing the uncertainties in both. Bayesian MCMC is able to provide a reasonable description of the parameters and quantiles even in cases where a maximum of the likelihood function provides a poor summary of the information in the data. O'Connell et al. (2002, pp. 16–9) observe that when measurement error is introduced, the likelihood function has multiple optima demonstrating 'how meaningless single-mode flood frequency estimates can be in the context of probabilistic risk assessment'.

Fig. 7 shows the flood frequency plot of the LP3 distribution fitted by the Bayesian approach using historical information along with the 95% credible interval for the flood quantiles. The largest flood that occurred in the 20 years of the systematic data actually has only a 1% chance of being exceeded. The availability of historical information has reduced the uncertainty in the more extreme flows relative to the smallest values (consider $p=0.999$ versus 0.001 in Fig. 7).

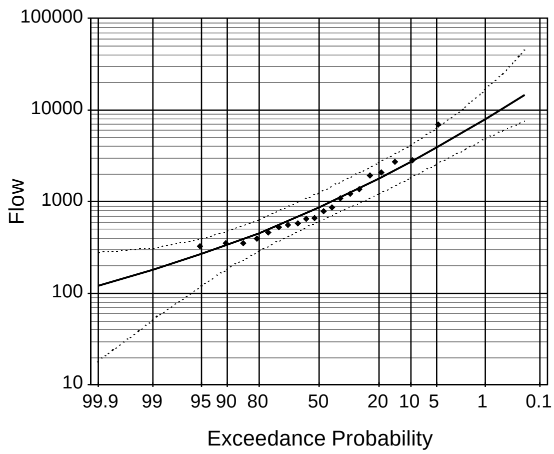


Fig. 7. Flood frequency plot of a LP3 distribution fitted by Bayesian MCMC with historical information. The dotted lines represent the 95% credible interval for flood quantiles. The sample, generated by a log-normal distribution, consists of 20 years of systematic data and 100 years of historical information, in which two floods exceeded the perception threshold (99th quantile).

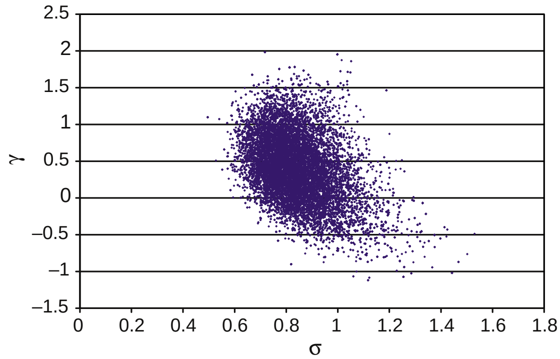


Fig. 8. Figure shows the joint distribution of the standard deviation σ and skewness coefficient γ of a LP3 distribution generated by the MCMC Bayesian simulation.

It is worth mentioning that the sample was generated by a log-normal distribution, a special case of the LP3 distribution with parameter $\gamma=0$. Fitting a distribution with 3 parameters increases the uncertainty in the estimates of both the parameters and flood quantiles as can be seen by the increased uncertainty in the quantiles. One can also see the large uncertainty in $\gamma$, with 95% credible interval ranging from $-0.45$ to $+1.28$. Fig. 8 shows the very unusual relationship between the posterior distribution of the standard deviation and skewness coefficient which is revealed by the Bayesian analysis (Kuczera (1999) provides a similar figure).

### 4.4. Impact of measurement error

Consider now the impact of including measurement errors, represented by $\omega$, into the flood frequency analysis as suggested by Kuczera (1996). In order to illustrate how these errors affect the results, suppose that all the unusually large floods that exceeded the perception threshold in both the systematic and historical data, and the discharge associated with the threshold, are subject to potentially large measurement errors that result from the hydrologic routing computation used to assign discharge values to the events. Consider three cases where that potential error results in a common multiplicative flood-flow measurement

Table 4
Results of Bayesian MCMC with and without historical information including measurement errors (ME) $\omega$, which is log-normally distributed with mean one and CV=0.10, 0.20, 0.30

| | $\gamma$ | | | $Q_{99}$/1000 | | | $E$(damage)/1000 | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2.5% | Mean | 97.5% | 2.5% | Mean | 97.5% | 2.5% | Mean | 97.5% |
| Systematic no ME | −0.30 | 0.57 | 1.48 | 3.8 | 9.1 | 39.9 | 14.0 | 104.4 | 407.5 |
| Historical no ME | −0.45 | 0.42 | 1.28 | 4.8 | 8.2 | 16.4 | 25.6 | 78.5 | 179.9 |
| Historical CV=0.10 | −0.49 | 0.41 | 1.28 | 4.7 | 7.9 | 15.9 | 23.7 | 75.0 | 173.0 |
| Historical CV=0.20 | −0.55 | 0.36 | 1.23 | 4.2 | 7.2 | 15.5 | 18.2 | 66.2 | 168.8 |
| Historical CV=0.30 | −0.65 | 0.30 | 1.21 | 3.5 | 6.5 | 14.8 | 11.7 | 57.3 | 161.2 |
| Systematic CV=0.30 | −0.43 | 0.50 | 1.44 | 3.6 | 8.5 | 39.3 | 12.0 | 97.4 | 392.5 |

error factor that is log-normally distributed with mean one and coefficient of variation equal to 0.10, 0.20, or 0.30. Table 4 presents results of the subsequent analysis of the flood data set. The first two rows present a statistical analysis for those cases in which measurement errors are not considered. Rows 3–5 show the results when measurement errors are considered in those floods that exceeded the perception threshold and in the discharge associated with the threshold. The last row presents the result for the case in which only systematic data is used and a measurement error factor with coefficient of variation equal to 0.30 is considered in the single flood that exceeded the threshold.

The effect of a common measurement error in this case is to place less weight on the observed values of the largest floods, and the historical information represented by the historical flood threshold which would have experienced the same error. When less weight is placed on those values, as suggested by in Fig. 2, the mean, standard deviation and skew of the fitted flood model decrease, resulting in a smaller 100-year flood and smaller flood damage estimates. This is reasonable for this sample because the historical peaks were unusually large.

Fig. 9 presents the width of the 95% credible interval for skew, and the relative width of the 95% credible interval for the 100-year flood, and the expected

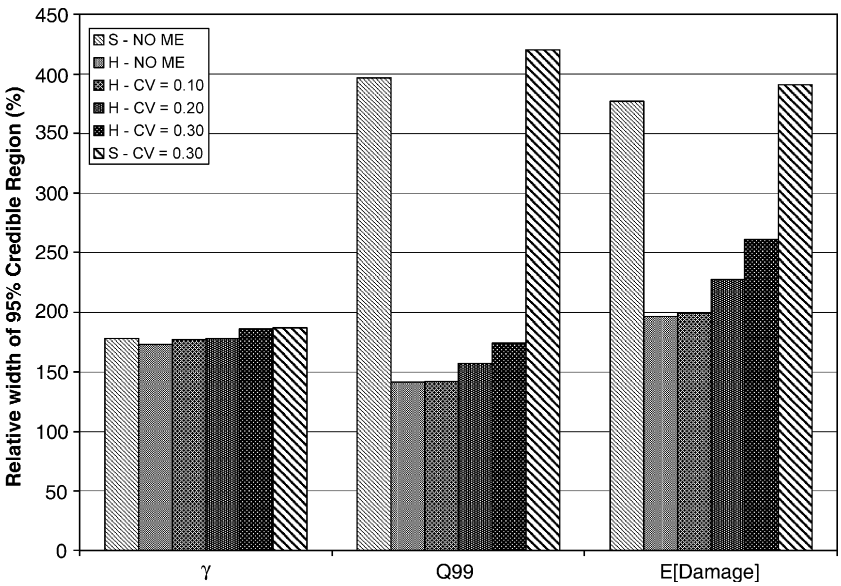


Fig. 9. Impact of measurement error on LP3 precision: Relative width of 95% credible interval. $s=20$, $k=2$, $h=100$, threshold=99th quantile.

damages. It shows that the use of historical information in this case does not improve significantly the estimation of the skew. This is explained by the fact that historical data provides information about the right tail of the distribution, but gives no information about the left tail. Fig. 9 does show that the historical information decreases drastically the uncertainty in the 100-year flood and expected flood damage even when relatively large measurement errors are considered. One can also see that the relative width of the 95% credible intervals increase as the coefficient of variation of the measurement errors increases. A Bayesian MCMC analysis allows the joint distribution of measurement errors to be integrated into the analysis, as this example illustrates.

## 5. Conclusions

This paper focuses on the use of a Bayesian framework for flood frequency analysis with historical information. The required computations are performed using Markov Chain Monte Carlo (MCMC) simulation. The Bayesian approach is an alternative to classical statistical estimators, such as adjusted-moment, maximum likelihood, and EMA. A Bayesian analysis provides the full posterior distribution of the parameters, or of any function of the parameters, such as desired quantiles and flood damages, by numerically sampling from the posterior distribution. It was shown through an example using log-normal and LP3 distributions that an uncertainty analysis using the Fisher Information matrix does not provide an accurate description of the actual uncertainty in some quantiles and other functions.

For the Log-Pearson Type 3 distribution with historical information, Bayesian MCMC has an advantage over the method of moments (either adjusted-moments and EMA), because it uses the full likelihood function, which is an effective and flexible way to represent information for a site, whether it be counts, intervals or particular magnitudes. Alternatives such as direct numerical integration (O'Connell et al., 2002) pose challenges because of the unbounded range of the three parameters, and experience additional difficulties as the dimension of the problem is increased as would occur if additional variables were added to represent uncertainty in the rating curve or individual flood measurements. The use of the Metropolis–Hastings algorithm avoids the numerical problems MLE faces when fitting the LP3 distribution, such as the non-existence of a maximum of the likelihood function, multiple maximum or a maximum value that is not a reasonable description of the information in a sample. Importance sampling, described by Kuczera (1999) is another alternative.

Clearly measurement error can be important when dealing with historical and paleoflood information, and it seems critical that the joint distribution of such errors be appropriately represented. A Bayesian MCMC approach provides a computationally and conceptually simple way of appropriately incorporating into flood frequency analysis the joint distribution of possible errors in rating curves and individual observations.

## Acknowledgements

We thank the Brazilian Government agency CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) who provided support for much of this research. We also thank George Kuzcera for leading the way and his comments on the manuscript.

## Appendix

This appendix presents the quadratic approximation of the log likelihood function for a two-parameter log-normal distribution when censored data is available (Cohn, 1984). This approximation was used in Section 3 to estimate the uncertainties in the MLEs.

Let $L'$ be the likelihood function with censored data for a log-normal distribution, The quadratic approximation of $L'$ based on the Fisher Information matrix can be written

$$L'_{\text{approx}}(\mu,\sigma) = L'(\mu_{\text{mle}},\sigma_{\text{mle}}) + 0.5\big[A(\mu-\mu_{\text{mle}})^2 + B(\sigma-\sigma_{\text{mle}})^2 + 2C(\mu-\mu_{\text{mle}})(\sigma-\sigma_{\text{mle}})\big]$$

where

$$A = E\left[\frac{\partial^2 L'}{\partial\mu^2}\right] = -\frac{1}{\sigma^2_{\text{mle}}}\left\{s + h[1-\Phi(x_0)] + \frac{h\phi(x_0)}{\Phi(x_0)}[\xi\Phi(x_0)+\phi(x_0)]\right\}$$

$$B = E\left[\frac{\partial^2 L'}{\partial \sigma^2}\right] = -\frac{1}{\sigma_{\text{mle}}^2}\left\{2[s+(1-\Phi(x_0))h] + 3h\phi(x_0)\xi - \frac{h\phi(x_0)}{\Phi(x_0)}[\Phi(x_0)(2-\xi^2) - \xi\phi(x_0)]\right\}$$

$$C = E\left[\frac{\partial^2 L'}{\partial \mu\,\partial \sigma}\right] = -\frac{1}{\sigma_{\text{mle}}^2}\left\{2h\phi(x_0) - \frac{h\phi(x_0)}{\Phi(x_0)}\left[\Phi(x_0)(1-\xi^2) - \xi\phi(x_0)\right]\right\}$$

$$\xi = \frac{(x_0 - \mu_{\text{mle}})}{\sigma_{\text{mle}}}$$

## References


Al-Futaisi, A., Stedinger, J.R., 1999. Hydrologic and economic uncertainties and flood-risk project design. Journal of Water Resources Planning and Management 125 (6), 314–324.

Ashkar, F., Ouarda, T.B.M.J., 1998. Approximate confidence intervals for quantiles of gamma and generalized gamma distributions. Journal of Hydrologic Engineering 3 (1), 43–51.

Baker, V.R., 2003. A bright future for old flows: origins, status and future of paleoflood hydrology. In: Thorndycraft, V.R., Benito, G., Barriendos, M., Llasat, M.C. (Eds.), Paleofloods, Historical Floods and Climatic Variability: Applications in Flood Risk Assessment Proceedings of the PHERA Workshop, Barcelona, 16–19th October, 2002. CSIC, Madrid, pp. 13–18.

Benito, G., Thorndycraft, V.R. (Eds.), 2004. Systematic, Paleoflood and Historical Data for the Improvement of Flood Risk Estimation: Methodological Guidelines. CSIC, Madrid.

Benito, G., Lang, M., Barriendos, M., Llasat, M., Frances, F., Ouarda, T., Thorndycraft, V., Enzel, Y., Bardossy, A., Couer, D., Bobbe, B., 2004. Use of systematic, paleoflood and historical data for the improvement of flood risk estimation: review of scientific methods. Natural Hazards 31, 623–643.

Benjamin, J., Cornell, A., 1970. Probability, Statistics and Decisions for Civil Engineering. McGraw-Hill, New York.

Bickel, P.J., Doksum, K.A., 1977. Mathematical Statistics: Basic Ideas and Selected Topics. Holden-Day, San Francisco.

Blainey, J.B., Webb, R.H., Moss, M.E., Baker, V.R., 2002. Bias and information content of paleoflood data in flood frequency analysis. In: House, P.K., Webb, R.H., Baker, V.R., Levish, D.R. (Eds.), Ancient Floods, Modern Hazards: Principles and Applications of Paleoflood Hydrology. American Geophysical Union, Washington, DC.

Bobee, B.B., Ashkar, F., 1991. The Gamma Family and Derived Distributions Applied in Hydrology. Water Resources Publications, Littleton, Colorado.

Carling, P.A., Kidson, R., Cao, Z., Herget, J., 2003. Palaeohydraulics of exterme flood events. In: Gregory, K.J., Benito, G. (Eds.), Paleohydrology: Understanding Global Change. Wiley, New York (Chapter 19).

Casella, G., George, E.I., 1992. Explaining the Gibbs Sampler. American Statistical Association 46 (3), 167–174.

Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. The American Statistician 49 (4), 327–335.

Chowdhury, J.U., Stedinger, J.R., 1991. Confidence interval for design flood with estimated skew coefficient. Journal of Hydraulic Engineering 117 (7), 811–831.

Cohn, T.A., 1984. The Incorporation of Historical Information in Flood Frequency Analysis. MS Thesis, School of Civil and Environmental Engineering, Cornell University, Ithaca, New York.

Cohn, T.A., Lane, W.L., Baier, W.G., 1997. An algorithm for computing moments-based flood quantile estimates when historical flood information is available. Water Resources Research 33 (9), 2089–2096.

Cohn, T.A., Lane, W.L., Stedinger, J.R., 2001. Confidence intervals for expected moments algorithm flood quantile estimates. Water Resources Research 37 (6), 1695–1706.

Coles, S., Pericchi, L., 2003. Anticipating catastrophes through extreme value modelling. Applied Statistics 52 (Part A), 405–416.

Coles, S., Pericchi, L.R., Sisson, S., 2003. A fully probabilistic approach to extreme rainfall modeling. Journal of Hydrology 273, 35–50.

Congdon, P., 2001. Bayesian Statistical Modelling. Wiley, West Sussex.

Frances, F., 2001. Incorporating non-systematic information to flood frequency analysis using the maximum likelihood estimation method. In: Glade, T., Alini, P., Frances, F. (Eds.), The Use of Historical Data in Natural Hazard Assessments. Kluwer Academic Publishers, Netherlands, pp. 89–99.

Frances, F., 2004. Flood frequency analysis using systematic and non-systematic information. In: Benito, G., Thorndycraft, V.R. (Eds.), Systematic, Paleoflood and Historical Data for the Improvement of Flood Risk Estimation: Methodological Guidelines. CSIC, Madrid, pp. 55–70.

Frances, F., Salas, J.D., Boes, D.C., 1994. Flood frequency analysis with systematic and historical or paleoflod data based on the two-parameter general extreme vale models. Water Resources Research 29 (6), 1653–1664.

Gamerman, D., 1997. Markov Chain Monte Carlo—Stochastic Simulation for Bayesian Inference. Chapman & Hall, London, UK.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman & Hall/CRC, London/Boca Raton.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721–741.

Griffis, V.W., Stedinger, J.R., Cohn, T.A., 2004. LP3 quantile estimators with reginal skew information and low outlier adjustments. Water Resources Research 40, W07503.

Hastings, W.K., 1970. Monte Carlo sampling methods using markov chains and their applications. Biometrika 57, 97–109.

Hirose, H., 1995. Maximum likelihood parameter estimation in the three-parameter gamma distribution. Computational Statistics and Data Analysis 20, 343–354.

Hosking, J.R.M., Wallis, J.R., 1986. Paleoflood hydrology and flood frequency analysis. Water Resources Research 22 (4), 543–550.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-moments. Cambridge University Press, Cambridge, UK.

House, P.K., Webb, R.H., Baker, V.R., Levish, D.R. (Eds.), 2002. Ancient Floods, Modern Hazards: Principles, and Applications of Paleoflood Hydrology Water Science and Application Series, vol. 5. AGU, Washington, p. 386.

Jarrett, R.D., Tomlinson, E.M., 2000. Regional interdisciplinary paleoflood approach to assess extreme flood potential. Water Resources Research 36 (10), 2957–2984.

Jin, M., Stedinger, J.R., 1989. Flood frequency analysis with regional and historical information. Water Resources Research 25 (5), 925–936.

Kirby, W., 1972. Computer oriented Wilson–Hilferty transformation that preserves the first 3 moments and lower bound of the Pearson type 3 distribution. Water Resources Research 10 (2), 220–222.

Kite, G.W., 1988. Frequency and Risk Analysis in Hydrology. Water Resources Publication, Littleton, CO.

Kuczera, G., 1996. Correlated rating curve error in flood frequency inference. Water Resources Research 32 (7), 2119–2127.

Kuczera, G., 1999. Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. Water Resources Research 35 (5), 1551–1557.

Landwehr, J.M., Matalas, N.C., Wallis, J.R., 1978. Some comparisons of flood statistics in real and log space. Water Resources Research 14 (5), 902–920.

Martins, E.S., Stedinger, J.R., 2001. Historical information in a GMLE-GEV framework with partial duration and annual maximum series. Water Resources Research 37 (10). 2551–2557.

Metropolis, N., Rosenbluth, A.W., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast computing machines. Journal of Chemical Physics 21, 1087–1092.

National Research Council, 1995. Flood Risk Management and the American River Basin: An Evaluation. National Academy Press, Washington, DC.

National Research Council, 1998. Estimating Probabilities of Extreme Floods: Methods and Recommended Research. National Academy Press, Washington, DC. 141 pp.

National Research Council, 2000. Risk Analysis and Uncertainty in Flood Damage Reduction Studies. National Academy Press, Washington, DC.

O'Connell, D.R.H., Ostenaa, D.A., Levish, D.R., Klinger, R.E., 2002. Bayesian flood frequency analysis with paleohydrologic bound data. Water Resources Research 38 (5).

Ostenaa, D.A., O'Connell, D.R.H., Walters, R.A., Creed, R.J., 2002. Holocene paleoflood hydrology of the big lost river, western INEEL, Idaho. In: Link, P.K., Mink, L.L. (Eds.), Geology, Hydrogeology and Environmental Remediation Special Paper of Geological Society of America 353. Idaho National Engineering and Environmental Laboratory, Eastern Snake River Plain, Idaho, pp. 91–110.

Ouarda, T., Hamdi, Y., Bobee, B., 2004. A general system for frequency estimation in hydrology (FRESH) with historical data. In: Benito, G., Thorndycraft, V.R. (Eds.), Systematic, Paleoflood and Historical Data for the Improvement of Flood Risk Estimation: Methodological Guidelines. CSIC, Madrid, pp. 55–70.

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

Pilon, P.J., Adamowski, K., 1993. Asymptotic variance of flood quantile in log Pearson type III distribution with historical information. Journal of Hydrology 143, 481–503.

Potter, K.W., Walker, J.F., 1985. An empirical study of flood measurement error. Water Resources Research 21 (3), 403–406.

Reis Jr., D.S., Stedinger, J.R., Martins, E.S., 2003. Bayesian GLS regression with application to LP3 regional skew estimation. In: Bizier, P., DeBarry, P. (Eds.), Proceedings World Water and Environmental Resources Congress 2003, June 23–26, 2003, Paper #811. American Society of Civil Engineers, Reston, Virginia, Philadelphia, PA.

Salas, J.D., Wold, E.E., Jarrett, R.D., 1994. Determination of flood characteristics using systematic, historical and paleoflood data. In: Rossi, G., Harmancioglu, N., Yevjevich, V. (Eds.), Coping with Floods. Kluwer Academic Publishers, Netherlands, pp. 111–134.

Sho, K., Iwasaki, S., Tominaga, A., 2000. Effect of introducing uncertain historical hydrologic data on quantiles estimation accuracy. Journal of Hydroscience and Hydraulic Engineering 18 (1), 45–52.

Spiegelhalter, D., Thoms, A., Best, N., 2000. WinBugs Version 1.3 User Manual. Medical Research Council Biostatistics Unit, Cambridge, UK.

Stedinger, J.R., 1980. Fitting log normal distributions to hydrologic data. Water Resources Research 16 (3), 481–490.

Stedinger, J.R., 1983a. Confidence intervals for design events. Journal of Hydraulic Engineering 109 (1), 13–27.

Stedinger, J.R., 1983b. Design events with specified flood risk. Water Resources Research 19 (2), 511–522.

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

Stedinger, J.R., Lu, L., 1995. Appraisal of regional and index flood quantile estimators. Stochastic Hydrology and Hydraulics 9 (1), 49–75.

Stedinger, J.R., Tasker, G.D., 1986. Regional hydrologic analysis, 2, model-error estimators, estimations of sigma and log-Pearson type 3 distributions. Water Resources Research 22 (10), 1487–1499.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidment, D. (Ed.), Handbook of Hydrology. McGraw-Hill, New York (Chapter 18).

Stevens, E.W., 1994. Multilevel model for gage and paleoflood data. Journal of Water Resources Planning and Management 120 (4), 444–457.

Tasker, G.D., Stedinger, J.R., 1986. Regional skew with weighted LS regression. Journal of Water Resources Planning and Management 112 (2), 225–237.

Thorndycraft, V.R., Benito, G., Barriendos, M., Llasat, M.C. (Eds.), 2002. Paleofloods, Historical Floods and Climatic Variability: Applications in Flood Risk Assesment Proceedings of the PHERA Workshop, Barcelona, 16–19th October. CSIC, Madrid.

Tierney, L., 1994. Markov Chains for exploring posterior distributions. Annals of Statistics 22 (4), 1701–1728.

United States Army Corps of Engineers (USACE), 1996. Risk-Based Analysis for Evaluation for Hydrology/Hydraulics, Geotechnical Stability, and Economics in Flood Damage Reduction Studies. ER 1105-2-101, Washington, DC.

United States Water Resources Council Hydrology Committee (USWRC), 1982. Guidelines for Determining Flood Frequency, Bull. 17B (revised). US Government Printing Office, Washington, DC.

Vicens, G.J., Rodríguez-Iturbe, I., Schaake Jr., C., 1975. A bayesian framework for the use of regional information in hydrology. Water Resources Research 11 (3), 405–414.

Whitley, R.J., Hromadka II., T.V., 1999. Approximate confidence intervals for design floods for a single site using a neural network. Water Resources Research 35 (1), 203–209.

Wood, E.F., Rodriguez-Iturbe, I., 1975. Bayesian inference and decision making for extreme hydrologic events. Water Resources Research 11 (4), 533–542.

Zellner, A., 1971. An Introduction to Bayesian Inference in Econometrics. Wiley, New York.