The construction of confidence intervals for frequency analysis using resampling techniques

M.J. Hall[1], H.F.P. van den Boogaard[2], R.C. Fernando[3] and A.E. Mynett[1,2]

[1]UNESCO-IHE Institute for Water Education, PO Box 3015, 2601 DA, Delft, The Netherlands
[2]WL|Delft Hydraulics, PO Box 177, 2600 MH, Delft, The Netherlands
[3]Sri Lanka Land Reclamation and Development Corporation, Rajagiriya, Sri Lanka

Email for corresponding author: m.hall@unesco-ihe.org

Abstract

Resampling techniques such as the Bootstrap and the Jack-knife are generic methods for the estimation of uncertainties in statistics. When applied in frequency analysis, resampling techniques can provide estimates of the uncertainties in both distribution parameters and quantile estimates in circumstances in which confidence limits cannot be obtained theoretically. Test experiments using two different parameter estimation methods on two types of distributions with different initial sample sizes and numbers of resamples has confirmed the utility of such methods. However, care is necessary in evaluating the skewness of the resampled quantiles, especially with small initial sample sizes.

Keywords: Bootstrap, Jack-knife, frequency analysis, maximum likelihood method, maximum product of spacings method

Introduction

The increasing pressures of population growth are resulting in urban encroachment into what were previously acknowledged to be disaster-prone areas. The deltas of large river basins are becoming a particular focus for such development. The economic benefits to be gained from flood protection works for such areas are substantial but their computation depends upon the estimation of extreme floods, a statistical procedure that is subject to considerable uncertainties that deserve more than a cursory evaluation.

In general, the design event for a flood alleviation scheme is estimated as a predetermined quantile drawn from a frequency distribution that is chosen to represent the behaviour of observed floods at the site of interest. The standard error of such a quantile estimate may then be employed to develop confidence limits that provide a measure of uncertainty for the design flood (Kite, 1975), which can be taken into account in the decision making. The computation of the standard error of a quantile estimate depends upon both the form of the distribution and the method of parameter estimation. If a quantile can be expressed as a linear function of the distribution parameters,

then its standard error is expressed easily as a function of the variances and covariances of the estimated parameters. More specifically, if the method of moments (MOM) is applied to a two-parameter distribution, the variances of the sample mean and standard deviation are required, along with their covariance. For three parameter distributions, the sampling variance of the skewness and its covariances with the mean and standard deviation are also required. However, quantiles do not in general depend linearly on the third central moment. Approximations that ignore the variability of the skewness are available but such simplifications can lead to marked underestimation or overestimation of confidence intervals dependent upon return period, as exemplified by Bobée (1973) for the Pearson Type III (PE3) distribution. Another approximation technique that allows for the presence of the third parameter involves expanding the expression for the quantile as a function of the estimated parameters as a Taylor series about the population values, discarding all terms above the second order on squaring and taking expectations of the remaining terms. Jenkinson (1969) applied this approach to the General Extreme Value (GEV) distribution, with parameter estimation by the

Maximum Likelihood Method (MLH), but the resultant expression is not applicable for extreme quantiles and large positive values of the shape parameter, which exhibit a marked upper bound.

Further problems may be encountered with the application of different methods of parameter estimation. For example, MLH does not always provide satisfactory estimates of parameters for certain three-parameter distributions in which the density is positive only to the right hand side of a shifted origin, the latter being one of the unknown parameters. The Lognormal, Weibull and PE3 distributions all take this form. In determining their parameters, the shift parameter tends towards the smallest observation, the likelihood becomes unbounded and the estimates of the other parameters become inconsistent. Parameter estimation techniques that address these difficulties specifically, such as the Maximum Product of Spacings (MPS) method of Cheng and Amin (1983), can clearly be used to advantage for such distributions.

In general, therefore, the estimation of the standard error of a quantile from a predetermined frequency distribution is not altogether straightforward. A limited number of analytical expressions is available for both MOM and MLH parameter estimation methods applied to the Normal and Extreme Value Type I (EV1 or Gumbel) distributions (see NERC, 1975, Section 1.4). Analytical relationships for the PE3 distribution have been provided by Kaczmarek (1957) and Bobée (1973) and information on the standard errors of quantiles derived from a *peaks-over-a-threshold* model are summarised in NERC (1975; Section 2.7.5; see also Zucchini and Adamson, 1989). The possibilities for estimating the standard errors of quantiles using alternative approaches are therefore worthy of particular attention. One such possibility is the application of *resampling methods.*

In this paper, the construction of confidence limits to quantile estimates obtained from both two- and three-parameter frequency distributions is evaluated by means of Bootstrap resampling techniques. The EV1 (Gumbel) distribution was adopted as an example of a typical two-parameter distribution for which standard errors of quantile estimates can be derived analytically. Different sizes of samples were generated by Monte Carlo methods from an EVI population and the parameters of those samples were estimated using the MLH method. The analytical solution for the standard error of the quantile estimates is then compared with those obtained by application of the Bootstrap, with particular attention being paid to the number of resamples required for different initial sample sizes. The exercise was then repeated with a three-parameter Weibull distribution, for which, as noted above, the standard error of the quantile estimates is difficult to determine by analytical methods. These evaluations employed the MPS

method of parameter estimation instead of MLH because of the problems encountered with the latter for J-shaped distributions, as remarked earlier. The paper concludes with a discussion of the implications of the results for practical applications of resampling methods in hydrology and water resources engineering.

Resampling methods

Resampling methods are data-based simulation methods for statistical inference. The *Jack-knife* was introduced in the 1950s as a computer-based method for estimating biases and standard errors and the *Bootstrap* was originally proposed in the late 1970s as an approach for the estimation of standard errors (see Efron and Tibshirani, 1993, for further background). In effect, resampling creates an *ensemble* of data sets, each of which is replicated from the original sample. *Jack-knife algorithms* generate the new samples by deleting one (or more) specific data points. In contrast, *Bootstrap algorithms* create the new data sets by sampling with replacement, i.e. one or more data points may be absent, and one or more may be repeated more than once in any resampled data set. Therefore, given an original sample of size N, the total number of possible Bootstrap resamplings is $\binom{2N-1}{N}$. This quantity increases rapidly with N, and soon becomes infeasible. However, in practice, the number of resamples, B, is generally selected according to an accuracy or convergence criterion and the B resamples are a randomly chosen sub-set of the total number of possible resamples.

Despite their considerable attraction in both avoiding theoretical calculations and providing estimates of standard errors however mathematically complicated the form of the estimator, resampling techniques have attracted only limited attention in hydrology and water resources engineering. However, the Bootstrap has been applied by Zucchini and Adamson (1988) and Tasker and Dunne (1997) to assess the risks associated with reservoir operation. In addition, Tung and Mays (1981), Potter and Lettenmaier (1990) and Burn (2003) have applied Bootstrapping to regional flood frequency analysis and Moss and Tasker (1991) have employed the same technique for the design of hydrometric networks. The Jack-knife has been used by Benjamini and Harpaz (1986) to determine the statistical significance of changes in runoff from areas subjected to cloud seeding experiments. Applications of resampling in the analysis of hydrological time series have been reported by several authors, including Pereira *et al.* (1984), Oliveira *et al.* (1988), Vogel and Shallcross (1996), Lall and Sharma (1996) and Sharma *et al.* (1997).

As an illustration, consider the procedure for forming a Bootstrap estimate of the standard error of a parameter

estimate, $\hat{\theta}$. Given a sample vector $X = [x_1, x_2, \ldots, x_N]^T$, B independent samples, $X_1^*, X_2^*, \ldots, X_B^*$, each consisting of N values, can be drawn from the original sample with replacement. For each resample, an estimate of the required parameter, $\hat{\theta}_b^*$, $b = 1, 2, \ldots, B$, is computed. The standard error of the parameter, $SE(\hat{\theta})$, is then estimated by the standard deviation of the B replications:

$$SE(\hat{\theta}) = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} (\hat{\theta}_b^* - \hat{\theta}^*)^2} \; ; \quad \hat{\theta}^* = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_b^* \qquad (1)$$

Given the standard error of the parameter, the a-level confidence limits may be constructed by assuming that each quantile is normally distributed and computing $\hat{\theta}^* \pm z_\alpha SE(\hat{\theta})$, where z_α is the standardised normal deviate corresponding to the confidence level α (the *Gaussian method*). Alternatively, if a Gaussian approximation is not supportable and the value of B is large enough, the appropriate limiting values may be read from the list of the B estimates of the quantile ranked in ascending order of magnitude (the *percentile method*). A typical application in which the Bootstrap percentile method was applied to construct the confidence limits for quantiles of rainfall depth has been described by Dunn (2001).

The Bootstrap percentile method as described above represents a basic form of resampling. More refined procedures for defining confidence limits, such as the *Bootstrap-t* and the *bias-corrected and accelerated* (BC$_a$) methods, are available, as exemplified by the discussions of Efron and Tibshirani (1993; chapters 12–14) and Davison and Hinkley (1997; chapter 5). However, each of these more refined approaches may have its disadvantages, such as the estimation of the acceleration parameter in the BC$_a$ method, which were considered to add an extra level of complexity in this preliminary investigation. The majority of results are, therefore, presented only for the Bootstrap percentile method.

In addition, the Bootstrap may be applied in two different forms, involving either the original observations (often referred to as the *Bootstrap standard method*) or the residuals that remain after a model is fitted to the data. In terms of frequency analysis, the model is the distribution that must be fitted to the sample and the residuals are the differences between the observations and the points on the fitted distribution corresponding to the same standardised variates. However, in resampling residuals, particular attention must be paid to both their distributional properties and their independence. For example, in Bootstrapping regression models, the residuals are often assumed to be mutually independent and normally distributed with zero expectation (see, e.g. Freedman, 1981). If the residuals are serially correlated, then the variance of the resampled distribution

parameters will be suppressed and the derived confidence intervals will be too narrow. More detailed discussions of this problem have been provided by Zwiers (1990) and Wilks (1997). Preliminary investigation of the residuals from fitted distributions showed that the larger values tend to occur in the tails and that the residuals themselves were invariably quite strongly dependent. Therefore, only the results from the Bootstrap standard approach are reported herein.

During the preliminary phase of this study, some attention was also given to the Jack-knife. In Jack-knife resampling, a statistic is recomputed over a set of sub-samples of a data set $X := \{x_1, x_2, \cdots, x_N\}$. In the 'standard' *Jack-knife delete-1* scheme, these sub-samples of X are of the form $X^{(n)} = \{x_1, x_2, \cdots, x_{n-1}, x_{n+1}, \cdots, x_N\}$. In this way, N resamples of X of size $N-1$ are created by leaving out one observation at a time. Alternatively each resample $X^{(n)}$ can be seen as a selection of $N-1$ samples from X without replacement.

For a statistic $\theta = \Theta_N(x_1, x_2, \cdots, x_N)$ of the sample X (with the N samples x_n mutually independent and identically distributed), the Jack-knife estimate $\hat{\theta}_{Jack}$ and its standard error $SE(\hat{\theta}_{Jack})$ are defined by

$$\hat{\theta}_{Jack} = \frac{1}{N} \cdot \sum_{n=1}^{N} \Theta_{N-1}(x_1, x_2, \cdots, x_{n-1}, x_{n+1}, \cdots, x_N) \qquad (2)$$

$$SE^2(\hat{\theta}_{Jack}) = \frac{N-1}{N} \cdot \sum_{n=1}^{N}$$

$$\left(\hat{\theta}_{Jack} - \Theta_{N-1}(x_1, x_2, \cdots, x_{n-1}, x_{n+1}, \cdots, x_N) \right)^2 \qquad (3)$$

The literature on the Jack-knife (e.g. Efron and Tibshirani, 1993; Shao and Tu, 1995) shows that the consistency of estimates for the variance or standard error requires a sufficient amount of *smoothness* of the statistic θ as function of the N samples x_n. In frequency analysis, such smoothness will usually hold when the statistic(s) θ is (are) uncertain parameters in a probability density function (PDF) $f(\cdot \mid \theta)$. It 'merely' has to hold that $f(\cdot \mid \theta)$, and/or the associated cumulative distribution function (CDF) $F(\cdot \mid \theta)$, depend smoothly on θ. At the same time, the fitting method (such as MLH, MPS, MOM, etc.) must lead to an estimate $\theta = \Theta_N(x_1, x_2, \cdots, x_N)$ that is smooth with respect to all the x_n.

In preliminary simulations (not reported here), the Jack-knife delete-1 approach was indeed found to produce satisfactory estimates for the mean and standard errors of the parameters θ of a PDF $f(\cdot \mid \theta)$. This result was verified by comparing these Jack-knife estimates with either analytical expressions (for those special cases where such expressions are available), or the corresponding results using a Bootstrap form of resampling. In fact, in many cases the

Jack-knife estimates turned out to be quite accurate, even for relatively small sample sizes as low as $N = 30$.

In frequency analysis, estimates and uncertainties of the distribution's parameters are not necessarily the main issue. More important are the estimates and uncertainties of quantiles, i.e. the levels x_P associated to one or more prescribed probabilities of non-exceedence $P \in (0,1)$. For a given CDF, $F(\cdot \mid \theta)$, and an estimate $\hat{\theta}$ of the distribution's parameters θ, an estimate for a quantile x_P can be computed as $\hat{x}_P = F^{inv}(P \mid \hat{\theta}) = F^{inv}\left(P \mid \Theta_N\left(x_1, x_2, \cdots, x_N\right)\right)$. In this way, the x_P are estimated in a para-meterised form using an analytical distribution fitted to the sample $X = \{x_1, x_2, \cdots, x_N\}$, rather than being derived directly from the empirical distribution of the sample. For the latter case, the quantile estimate is *not* a smooth statistic of the sample and, as a result, the standard Jack-knife may produce highly inconsistent estimates for the standard error and/or other measures describing the uncertainty of the quantile estimate (Efron and Tibshirani, 1993; Shao and Tu, 1995). Since the estimate $\hat{x}_P = F^{inv}(P \mid \hat{\theta})$ for a quantile will be more smooth than that based on the empirical distribution of sample X, its Jack-knife estimate for the standard error may perform better. This possibility was confirmed by simulations that showed a good agreement between Jack-knife and Bootstrap estimates of the quantiles x_P and their standard errors. This agreement was obtained for large samples sizes, $N \geq 100$ say and, even for large probabilities of non-exceedence (i.e. P close to 1), only minor differences were found. For small sample sizes, $N < 50$ say, the situation was different, however, with Jack-knife and Bootstrap estimates of the quantiles x_P and their standard error agreeing well only if P was not larger than some value between $\frac{3}{4}$ and $\frac{N}{N+1}$. For larger P, and especially when $P > \frac{N}{N+1}$ (where $\frac{N}{N+1}$ can be seen as the empirical probability of non-exceedence associated with the largest data point of the sample X), large differences may occur. These differences are typical with forms of the Bootstrap and Jack-knife histograms of the ensemble of resampled x_P that are skewed. In this situation, the confidence interval of a quantile x_P will be also be skewed. Such skewed confidence intervals cannot be derived from merely the standard error of the quantile and should therefore be based on a percentile method. For the Bootstrap, this approach is feasible but not for the Jack-knife delete-1 because of the inconsistency of its histogram (see e.g. Shao and Tu, 1995).

The inconsistency of Jack-knife estimates of non-smooth statistics and/or histograms can be overcome using a more generalised form of sub-sampling by leaving out d ($d > 1$) different observations at a time instead of 1, as considered so far. The details of *Jack-knife delete-d* resampling can be found in Efron and Tibshirani (1993) and Shao and Tu (1995). Those authors show that in this case consistent standard errors and sampling histograms are obtained if and only if d and $r = N - d$ diverge to infinity. This condition, therefore, applies to large sample sizes. For limited sample sizes, the suitability of the Jack-knife delete-d is, therefore, doubtful for the construction of accurate skew confidence intervals of quantiles.

The above considerations lead to the following conclusions about the practical significance of the Jack-knife method for the assessment of uncertainties in frequency analysis. For large sample sizes ($N \geq 100$, say), a Jack-knife form of resampling can be used and may serve as an alternative to the Bootstrap. The estimates and confidence intervals of the parameters θ and/or quantiles $\hat{x}_P = F^{inv}(P \mid \hat{\theta})$ of a distribution $f(\cdot \mid \theta)$ produced by these methods appear to compare well. For $N \geq 100$, confidence intervals are virtually symmetrical so that these intervals can be based on the standard error and there is no real necessity to rely on a Jack-knife histogram and apply the percentile method. For small sample sizes ($N \leq 50$, say), a Jack-knife may still produce accurate estimates of the parameters θ and their standard errors. For the quantiles x_P, this also holds as long as the probability of non-exceedence is not too large. However, with small sample sizes, confidence intervals for the estimates of the parameters and quantiles are often quite skewed and cannot, therefore, be derived solely from the standard error. A percentile method based on the Jack-knife histogram should then be applied. Moreover, because of the small sample size N, no large d and $r = N - d$ can be used to guarantee the accuracy of that histogram. In practical applications of frequency analysis, large sample sizes are exceptional and often the number of samples is less than 50. In practice, therefore, a Jack-knife method tends to be less suited for the assessment of uncertainties and in the remainder of this paper this method will not be considered in further detail.

The distributions

The cumulative distribution function (CDF) of the EV1 (or Gumbel) distribution may be written in the form

$$F(x \mid \Theta) = P = \exp\left[-\exp\left(-\frac{x - \mu}{\sigma}\right)\right]; \quad \Theta = [\mu, \sigma]^T; \quad \sigma > 0 \tag{4}$$

where Q are the scale and location parameters, s and m, of the distribution of x. The quantiles, Q, of l are then a linear function of the standardised variate

$$H = -\ln[-\ln(P)]; \quad P \in (0,1) \tag{5}$$

given by

$$Q = \mu + H.\sigma \qquad (6)$$

Given the linear form of Eqn. (6), the variance of the estimated quantile, $\hat{Q}$, can be written in terms of the approximate variances and covariance of the estimated location and scale parameters, $\hat{\mu}$ and $\hat{\sigma}$, determined from the sample of N observations, according to

$$VAR(\hat{Q}) = VAR(\hat{\mu}) + 2H\,COV(\hat{\mu},\hat{\sigma}) + H^2\,VAR(\hat{\sigma}) \qquad (7)$$

Using MLH, the variances and covariance terms on the right-hand side of Eqn. (7) may be obtained from the variance-covariance matrix (NERC, 1975; Section 1.3), which is the inverse of the Hessian of the minus log likelihood function computed at $\mu = \hat{\mu}$ and $\sigma = \hat{\sigma}$ according to

$$\begin{pmatrix} VAR(\mu) & COV(\mu,\sigma) \\ COV(\mu,\sigma) & VAR(\sigma) \end{pmatrix} = \frac{\sigma^2}{N} \begin{pmatrix} 1 + \frac{6}{\pi^2}(1-\gamma)^2 & \frac{6}{\pi^2}(1-\gamma) \\ \frac{6}{\pi^2}(1-\gamma) & \frac{6}{\pi^2} \end{pmatrix}$$

$$= \frac{\sigma^2}{N} \begin{pmatrix} 1.11 & 0.26 \\ 0.26 & 0.61 \end{pmatrix} \qquad (8)$$

In this expression, γ is Euler's constant, 0.5772 …. Hence

$$se(\hat{Q}) = \sqrt{VAR(\hat{Q})} = \frac{\hat{\sigma}}{\sqrt{N}} \sqrt{1.11 + 0.52H + 0.61H^2} \qquad (9)$$

The three-parameter distribution chosen for the numerical experiments was the Weibull distribution. This distribution may be regarded as a 'reversed' Extreme Value type III (EV3) distribution. Its location parameter, σ_0, is equal to the sum of the EV3 location parameter and the ratio between the EV3 scale and shape parameters; the scale parameter, σ, is the ratio between the EV3 scale and shape parameters; and the shape parameter, a, is the inverse of the EV3 shape parameter. The CDF of the Weibull distribution may be written as

$$F(x|\Theta) = P = 1 - \exp\left[-\left(\frac{x - x_0}{\sigma}\right)^\alpha\right]; \quad x \geq x_0\ ;$$

$$\sigma, \alpha > 0\ ; \quad \Theta = [x_0, \sigma, \alpha]^T \qquad (10)$$

where Q are the location, scale and shape parameters, x_0, σ and α of the distribution of x. The quantiles, Q, of x are then a non-linear function of the Weibull probability plotting variate

$$H = \ln[-\ln(1-P)]; \quad P \in (0,1) \qquad (11)$$

given by

$$Q = x_0 + \sigma.\exp\left(\frac{H}{\alpha}\right) \qquad (12)$$

so that the standard error of the estimated quantile cannot be expressed explicitly as a function of the variances and covariances of the three estimated parameters.

The fitting methods

For the purposes of this study, attention was concentrated on the MLH and the MPS methods. MLH is based upon the maximisation of the log likelihood function. Given a set of N independent and identically distributed observations x_n described by a probability density function $f(x_n|q)$, the distribution parameters q are determined from the set of equations

$$\frac{\partial}{\partial\theta} \log\left[\sum_{n=1}^{N} f(x_n|\theta)\right] = 0 \qquad (13)$$

The MPS method, as proposed by Cheng and Amin (1983) and introduced independently by Ranneby (1984), is less well known than MLH but has better general robustness and identifiability features. The MPS method is related to MLH but, in MPS, a form of log likelihood function is optimised that is based on the *spacings* of the data points after these have been sorted in ascending order. In brief, given the ordered data points $\{x_{(1)}, x_{(2)}, \cdots, x_{(N)}\}$ of a sample $X = \{x_1, x_2, \cdots, x_N\}$, the following function must be maximised in MPS to find an estimate for the parameters Θ of some distribution with probability density function $f(\cdot|\Theta)$, or the equivalent cumulative distribution function $F(\cdot|\Theta)$:

$$H(\Theta) := \log\left(\prod_{n=1}^{N+1} D_n(\Theta)\right) = \sum_{n=1}^{N+1} \log\left(D_n(\Theta)\right) \qquad (14)$$

In this expression, the one-step spacings $D_n(\Theta)$ are defined by:

$$D_n(\Theta) := \int_{x_{(n-1)}}^{x_{(n)}} f(\xi|\Theta)\,d\xi = F(x_{(n)}|\Theta) - F(x_{(n-1)}|\Theta) \qquad (15)$$

The 'additional data points' $x_{(0)}$ and $x_{(N+1)}$ are prescribed by $x_{(0)} := -\infty$ and $x_{(N+1)} := \infty$ so that $F(x_{(0)}|\Theta) = 0$ and $F(x_{(N+1)}|\Theta) = 1$. If, for some *n*, the data points $x_{(n-1)}$ and $x_{(n)}$ coincide, the spacing $D_n(\Theta)$ is replaced by the appropriate limit $f(x_{(n)}|\Theta)$.

For an increasing size N of the sample $X = \{x_1, x_2, \cdots, x_N\}$, all, or virtually all, increments $x_{(n)} - x_{(n-1)}$ will converge gradually to zero so that $\log\left(D_n(\Theta)\right)$ will be of the form

239

$\log\big(f(x_{(n)}\,|\,\Theta)\big)+\log\big(x_{(n)}-x_{(n-1)}\big)$ + higher order terms in $x_{(n)}-x_{(n-1)}$. Apart from these higher order terms and the independence of $\log\big(x_{(n)}-x_{(n-1)}\big)$ on Θ, this shows that $\log\big(D_n(\Theta)\big)$ has the same asymptotic behaviour with respect to Θ as $\log\big(f(x_{(n)}\,|\,\Theta)\big)$. In this way, Cheng and Amin (1983) demonstrated that MPS and MLH estimation are asymptotically equal and have the same asymptotic sufficiency, consistency, and efficiency properties.

The MPS method can be applied to the estimation of parameters in any continuous univariate distribution. Compared to other estimation techniques, however, MPS is especially suited to cases in which one of the parameters is an unknown location parameter or shifted origin, x_0, in three parameter distributions such as the GEV, Generalised Pareto, Weibull, Lognormal, and Gamma or Pearson Type III distributions. For such distributions with the shifted origin as one of the unknown parameters, it is known that MLH estimation often fails because the likelihood is unbounded, and/or does not have a local maximum. In particular, this will be the case when the distribution is *J*-shaped (unbounded density near the shifted origin); this occurs for the Weibull and Gamma distributions when the shape parameter is less than unity. In this and other non-regular cases, MPS will provide consistent estimators where MLH or other techniques are bound to fail.

Apart from MPS, there are more general forms of estimation methods based on spacings. Ghosh and Rao Jammalamadaka (2001) consider a Generalised Spacings Estimator (GSE) in which the estimate of the parameters is defined as the argument Θ that *minimises* the function

$$T(\Theta) := \sum_{n=1}^{N+1} h\big(N \cdot D_n(\Theta)\big) \qquad (16)$$

The $h(\cdot)$ must be a strict convex function and some standard choices that have been used in the context of goodness of fit testing are $h(x)=-\log(x)$ (leading to the MPS method as described above), $h(x)=-x\cdot\log(x)$, $h(x)=x^2$, or $h(x)=|x-1|$. Ghosh and Rao Jammalamadaka (2001) discuss consistency and asymptotic normality of these more general spacings-based estimators and show that, within this class, MPS has the smallest asymptotic variance and is asymptotically equivalent to MLH when the latter exists. On the other hand, in small to moderate sample sizes, the simulation studies by those authors suggest that MPS may not always be the best among the spacings estimators.

The numerical experiments

For both Gumbel and Weibull distributions, with parameters fitted by both MLH and MPS methods, the Bootstrap resampling algorithm was applied to four different sizes of sample (30, 100, 200 and 400 data points). All calculations were carried out using a modified version of the EVA computer program developed by WL | Delft Hydraulics for the analysis of extreme values (van den Boogaard *et al.*, 1999). For both distributions, the data were generated by Monte Carlo methods, care being taken to vary the seeds for the sequences of random numbers to ensure independence of successive samples. Scale and location parameters of $\sigma = 41.42$ and $\mu = 70.35$, corresponding to the values for a site in the English Midlands, selected at random from the data in Volume IV of the *UK Flood Studies Report* (NERC, 1975), were employed for the Gumbel distribution. For the Weibull distributions, the location parameter was set to zero and the scale parameter *s* to unity but separate analyses were carried out for shape parameters *a* of 0.75 and two (a Rayleigh distribution). In each case, the confidence intervals obtained from the resampling were constructed for 101 probabilities of non-exceedance, *P*, ranging from 0.0001 and 0.9999 on a uniform grid of the standardised variate, *H* of Equation (5) for the Gumbel, and of the plotting position variate, *H* of Eqn. (11) for the Weibull distributions, using the percentile method to determine the limits of the 95 per cent confidence intervals.

The Gumbel distribution

For any resampling algorithm, the number of resamples that is drawn from the original set of observations is a particularly important variable. With the Bootstrap, the number of resamples should be sufficiently large for the values of the estimated parameters to stabilise but not so large as to require excessive computational resources. In these numerical experiments, preliminary runs were undertaken in which, for a given initial sample size, the number of Bootstrap resamples, *B*, was increased in increments of 100 from 100 to 4000. The results obtained from applying both MLH and MPS to estimate the location and scale parameters of the Gumbel distribution showed that, with an initial sample size of 30, the parameter estimates had stabilised after about 600 to 1000 resamples for both MLH and MPS. As the initial sample size increases, the parameter estimates stabilise more rapidly and MLH and MPS scale parameters were in closer agreement. However, based upon the recommendations of Davison and Hinkley (1997; chapter 2), unless otherwise stated, the results for all further Bootstrapping work were carried out using *B*=1000 resamples.

PARAMETER ESTIMATES

Table 1 presents estimates of the scale and location parameters of the Gumbel distribution and their standard

Table 1. Estimates of the scale and location parameters for the Gumbel distribution (Eqn. (4); population values μ=70.35 and σ=41.42 respectively) and their standard errors (in brackets) for 1000 Bootstrap resamples using the MLH method of parameter estimation with four different sample sizes N. The standard errors for the estimates based on the 'original sample' are derived via (the inverse of) the Hessian matrix of the minus log-likelihood function at the position where this function is minimal.

Location parameter μ	Sample size N			
	30	100	200	400
Original sample	61.47 (9.19)	66.34 (4.50)	68.13 (3.13)	67.15 (2.15)
Bootstrap	62.04 (8.96)	66.50 (4.43)	68.20 (3.06)	67.19 (2.12)
Scale parameter σ	Sample size N			
	30	100	200	400
Original sample	47.83 (7.11)	42.78 (3.43)	42.01 (2.34)	40.80 (1.60)
Bootstrap	47.11 (6.35)	42.56 (3.20)	41 88 (2.15)	40.73 (1.53)

Table 2. Estimates of the scale and location parameters for the Gumbel distribution (Eqn. (4); population values =70.35 and σ=41.42 respectively) and their standard errors (in brackets) for 1000 Bootstrap resamples using the MPS method of parameter estimation with four different sample sizes N. The standard errors for the estimates based on the 'original sample' are derived via (the inverse of) the Hessian matrix of the minus MPS-function at the position where this function is minimal.

Location parameter μ	Sample size N			
	30	100	200	400
Original sample	60.84 (9.99)	66.17 (4.66)	68.01 (3.20)	67.09 (2.17)
Bootstrap	61.45 (8.80)	66.32 (4.41)	68.09 (3.05)	67.13 (2.12)
Scale parameter σ	Sample size N			
	30	100	200	400
Original sample	52.59 (8.03)	44.48 (3.60)	43.04 (2.40)	41.35 (1.62)
Bootstrap	51.42 (6.70)	44.12 (3.27)	42 83 (2.17)	40.80 (1.54)

errors using MLH for all resamples, 1000 Bootstrap resamples and for four different initial sample sizes. The same information is shown in Table 2 but now for parameter estimation by MPS. In these tables, the entries for the standard errors of the parameters from the original sample were computed from the inverse of the Hessian of the minus log likelihood function for MLH or the function *H(.)* of Eqn. (14) for MPS. These tables demonstrate that there is good agreement between the parameter estimates obtained by Bootstrapping and those of the original sample, even for samples of size 30. The standard errors of the location parameters are always larger than those of the scale parameters and, as is to be expected, the values decrease with increasing sample size. Apart from the absolute values of the scale parameter estimates tending to be higher with MPS than with MLH, there were only minor differences associated with the method of parameter estimation.

CONFIDENCE INTERVALS

For the Gumbel distribution, the theoretical estimates of the standard errors of the quantile estimates using MLH are given by Eqn. (7) and the 95 percent confidence intervals may be set at ±1.96 standard errors around the values obtained from Eqn. (6). Using the Bootstrap resamples, means and standard deviations of each of the 101 selected quantile estimates were computed and confidence levels calculated using the percentile method, as explained above. The results for 1000 Bootstrap resamples of size 30 with MLH parameter estimation are compared with the theoretical quantiles in Fig. 1. In this case, the confidence intervals obtained by the percentile method closely matched those obtained using the Gaussian approximation. The frequency distributions of the location and scale parameters for the 1000 resamples could, therefore, be expected to be close to normal, as is confirmed visually by Fig. 2. Figure 1 shows that, comparted to the Gaussian apprimation, the width of the Bootstrap confidence band is slightly underestimated for higher quantiles corresponding to probabilities of exceedance greater than 0.998 and the lower confidence limit is marginally too high for probabilities of exceedance less than 0.1. The reduction in width of the confidence band may be associated with the smaller Bootstrap estimate of
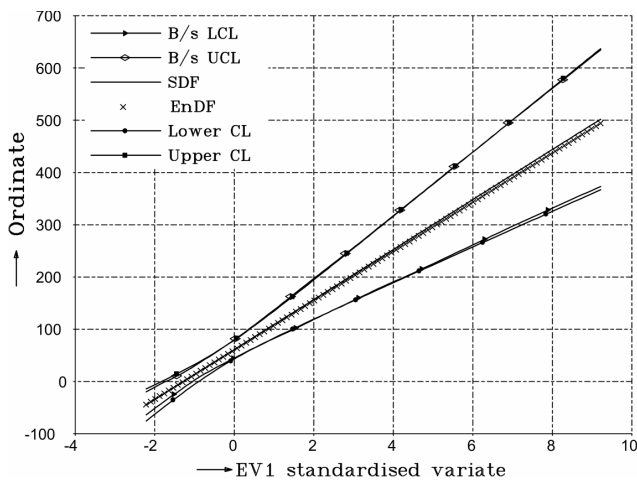
Fig. 1. *Ordinates of the Bootstrap ensemble distribution function (EnDF) and the upper and lower confidence limits computed from 1000 Bootstrap resamples from a sample of size 30 from a Gumbel distribution using MLH and the percentile approximation method compared with their sample distribution function (SDF). LCL denotes the lower, and UCL the upper confidence limit.*
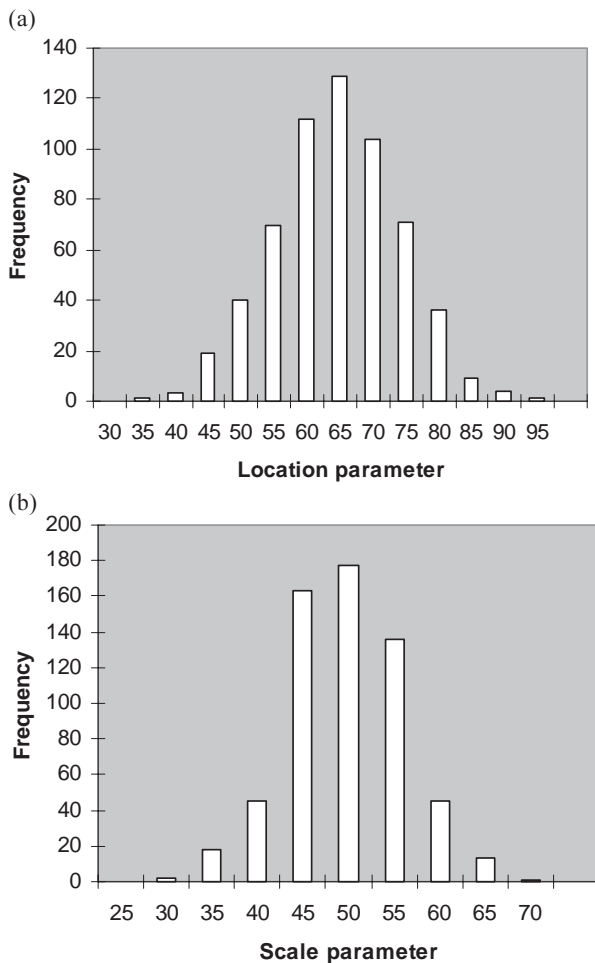
(a)

(b)

Fig. 2. *Frequency distributions of the (a) location and (b) scale parameters for 1000 resamples of size 30 from a Gumbel distribution using the MLH method of parameter estimation.*

the scale parameter of the EVI distribution evident in Table 1. This discrepancy reduces as the initial sample size increases, so that, for an initial sample size of 100, the computed and theoretical values were in close agreement.

The Weibull distribution

For the Weibull distribution, attention was confined to the cases with shape parameters of 0.75 and two. The MPS method of parameter estimation was employed and the Bootstrap applied for three different initial sampling sizes of 30, 100 and 400 using 1000 resamples. As noted above in the section on Fitting Methods, the MPS method has been shown by Cheng and Amin (1983) to provide consistent parameter estimates for three-parameter distributions with a shifted origin, such as the Weibull, in circumstances where the MLH method is prone to failure. However, initial results obtained from Bootstrapping samples from the Weibull distribution revealed that, under certain conditions, MPS may still fail to produce consistent estimates.

In their paper, Cheng and Amin (1983; example 2.2, p. 397) demonstrate that the MLH method may lead to inconsistent estimators when the true shape parameter a < 1. In this case, the MLH estimate for the location parameter x_0 converges to the smallest data point x_1 and an inconsistent estimate $\hat{\alpha} = 0$ is found for the shape parameter. For the MPS method, those authors argue that, for any α > 0, consistent estimates will be obtained for the parameters. However, further investigation has revealed that this condition holds only when the multiplicity of the smallest data point x_1 is unity. When the multiplicity of x_1 is $m > 1$, i.e. $x_1 = x_2 = \cdots = x_m$ while $x_{m+1} > x_1$, the true shape parameter α must be greater than $1 - \frac{1}{m}$ to obtain regular estimates. The arguments leading to this result are summarised separately in the Appendix.

For continuous distributions such as the Weibull distribution of Eqn. (10), the probability will be virtually zero that the multiplicity of the minimum x_1 of a random sample $X = \{x_1, x_2, \cdots, x_N\}$ is greater than 1. For an MPS estimate of Θ, the regularity condition derived in the Appendix then tends to be of more theoretical than practical relevance. This is not the case, however, when resampling techniques are applied to assess uncertainties for the estimates for the parameters Θ and the quantiles of the distribution $F(\cdot | \Theta)$. In the case of a Bootstrap resampling (a random selection of N samples from X with replacement), some fraction of the resamples $X^{(b)}$; $1 \leq b \leq B$ of X will have a smallest data point $X^{(b)}$ with a multiplicity larger than 1. For a part of this fraction, inconsistent MPS-estimates may then be found for Θ. The theoretical treatment of such resamples currently remains unclear. However, in the

experiments with Bootstrap resamplings applied to the Weibull distribution reported here, the strategy adopted was to ignore (or re-do) resamples that led to inconsistent estimates and the uncertainty measures (standard errors, confidence intervals) were based on the subset of resamples with regular estimates for the Θ-parameters.

The experiments with the Weibull distribution showed that the proportion of inconsistent estimates in a Bootstrap ensemble increases with (i) a decreasing population value of the shape parameter α (as can be expected from the regularity condition $\alpha > 1 - \frac{1}{m}$); and (ii) a decreasing sample size N. In fact, for a population value $\alpha = 0.75$, this proportion was found to be of the order of 23 per cent for N = 30 and about 11 per cent for $N = 100$. For the population value $\alpha = 2$, however, the proportions were only 0.5 and 0 per cent respectively.

Figures 3 and 4 present the results obtained for the shape parameters of 0.75 and 2 respectively. Each figure shows both the sample distribution function (SDF), which is based on MPS-estimates of the parameters for the original sample, and the Bootstrap ensemble distribution function (EnDF) that is obtained by averaging over all Bootstrap resamples, obtained using 1000 resamples, along with the 95 per cent confidence limits. For the purposes of discussion, the latter were computed using both the percentile and the Gaussian methods.

A particular feature of both plots is the marked spreading out of the confidence limits as the plotting position variate H increases from 1.5 (corresponding to a probability of non-exceedence $P=0.99$, see Eqn. 11) to 2.22 ($P=0.9999$). The limits computed using the Gaussian approximation show little agreement with those obtained using the percentile method. Indeed, for both shape factors, the Gaussian lower confidence limits are seen to *decrease* as the plotting position variate increases over 1.5. The reasons for this behaviour may be found in the distributional properties of both the quantiles for a given plotting position variate and the resampled parameters of the parent distribution. For the lower value of the shape parameter, Table 3 shows that, although the distributions of the Bootstrap location parameters are negatively skewed and those of the scale parameters are moderately positive, the distributions of the shape parameters are substantially positively skewed. Consequently, the distributions of the Bootstrap quantiles for probabilities of non-exceedance of 0.9, 0.99 and 0.999 are all positively skewed, with the skewness increasing as the probability increases. When the shape parameter α was increased to 2, Table 4 shows that, for an initial sample size of 30, the distribution of the Bootstrap location parameter exhibits moderate negative skewness but the scale and shape parameters have a moderate positive skewness. With the
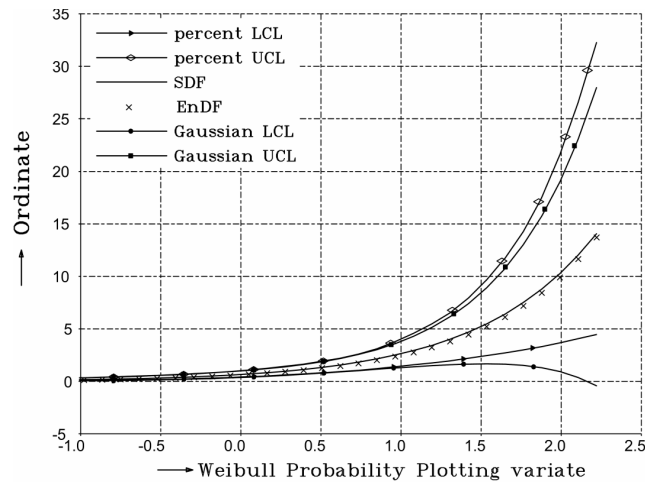

Fig. 3. *Ordinates of the Bootstrap ensemble distribution function (EnDF) and the 95 percent confidence limits for a Weibull distribution having a shape parameter of 0.75, based on 1000 Bootstrap resamples from a sample of size 30 using both Gaussian approximation and percentile methods, along with the sample distribution function (SDF). LCL denotes the lower, and UCL the upper confidence limit.*
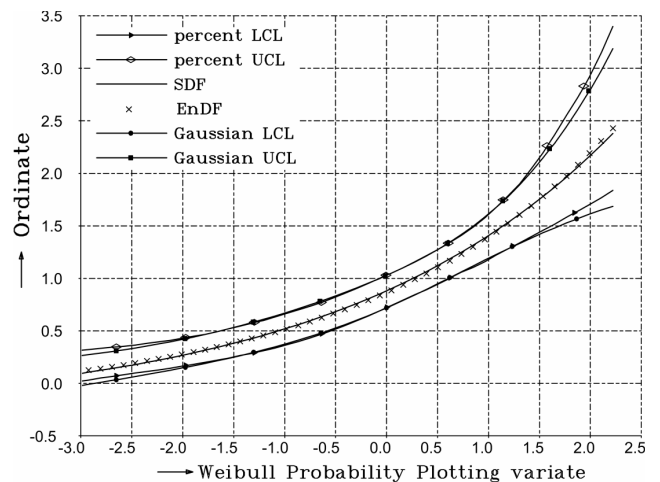

Fig. 4. *Ordinates of the Bootstrap ensemble distribution function (EnDF) and the 95 percent confidence limits for a Weibull distribution having a shape parameter of 2, based on 1000 Bootstrap resamples from a sample of size 30 using both Gaussian approximation and percentile methods, along with the sample distribution function (SDF). LCL denotes the lower, and UCL the upper confidence limit.*

exception of those for $P = 0.9$, the associated quantiles have a positive skewness that again increases rapidly with the probability of non-exceedance.

The dependence of results on the initial sample size is well illustrated by comparing Figs. 3 and 4 with Figs. 5 and 6, which apply to samples of size 100. For initial sample sizes of 100 and both values of the shape parameter, there is now good agreement between the Bootstrap Gaussian

Table 3. The skewness of the ensemble of resampled parameters, $\Theta = (\sigma, \alpha, x_0)$, and the skewness of the ensemble of quantiles, $x_P = F^{-1}(P \mid \Theta)$ for three different probabilities of non-exceedance for Weibull distributions with a shape parameter of $\alpha = 0.75$ using the Bootstrap resampling method on three different sample sizes and the MPS method of parameter estimation.

Sample size	Skewness of distribution of parameters			Skewness of distribution of quantiles		
	Scale	Shape	Location	P=0.9	P=0.99	P=0.999
30	0.487	4.638	-16.76	0.373	1.013	1.361
100	0.311	0.378	5.462	0.284	0.409	0.513
400	0.376	0.209	3.033	0.316	0.357	0.414

Table 4. The skewness of the ensemble of resampled parameters, $\Theta = (\sigma, \alpha, x_0)$, and the skewness of the ensemble of quantiles, $x_P = F^{-1}(P \mid \Theta)$ for three different probabilities of non-exceedance for Weibull distributions with a shape parameter of $\alpha = 2$ using the Bootstrap resampling method on three different sample sizes and the MPS method of parameter estimation.

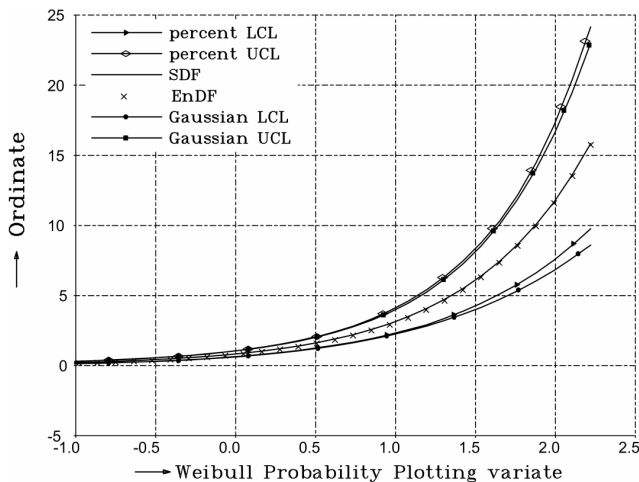| Sample size	Skewness of distribution of parameters			Skewness of distribution of quantiles		
	Scale	Shape	Location	P=0.9	P=0.99	P=0.999
30	0.448	0.518	-0.399	-0.106	0.485	0.980
100	0.051	0.906	0.230	0.062	0.055	0.110
400	0.029	0.322	0.260	0.166	0.127	0.139

Fig. 5. *Ordinates of the Bootstrap ensemble distribution function (EnDF) and the 95 percent confidence limits for a Weibull distribution having a shape parameter of 0.75, based on 1000 Bootstrap resamples from a sample of size 100 using both Gaussian approximation and percentile methods, along with the sample distribution function (SDF). LCL denotes the lower, and UCL the upper confidence limit.*
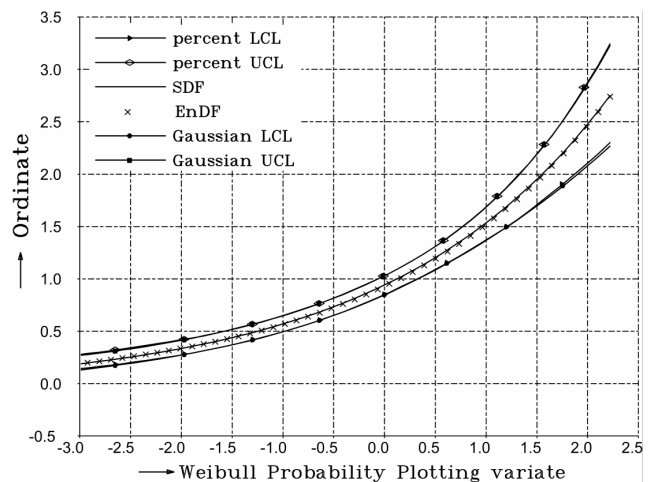
Fig. 6. *Ordinates of the Bootstrap ensemble distribution function (EnDF) and the 95 percent confidence limits for a Weibull distribution having a shape parameter of 2, based on 1000 Bootstrap resamples from a sample of size 100 using both Gaussian approximation and percentile methods, along with the sample distribution function (SDF). LCL denotes the lower, and UCL the upper confidence limit*

approximation and Bootstrap percentile confidence limits as well as between the SDF and the EnDF. Further scrutiny of Tables 3 and 4 reveals that, for both shape parameters, the distributions of the Bootstrap location, scale and shape parameters are close to symmetry for the value of 0.75 but have a more obvious positive skewness for the shape parameter at $\alpha = 2$. In the former case, the quantile distributions are close to symmetry but become more positively skewed as the probability increases. In the latter case, the distributions are even closer to symmetry.

A further increase in the initial sample size to 400 was found to improve the symmetry of the scale and shape

parameter distributions for Bootstrap resamples, although the location parameter had obvious positive skewness for a =0.75 (see Tables 3 and 4). The resultant quantile distributions are, therefore, reasonably symmetrical, especially for a = 2.

A further point of interest in Figs. 3 and 4 is the lack of agreement at higher standardised variates between the SDF and the EnDF. This discrepancy arises from the nature of the Bootstrap quantile estimates, which represent the averages of an ensemble of probability plots. Apart from the estimate of the mean, the ensemble can also be used to derive its uncertainty, either in the form of a spread or some (symmetrical or non-symmetrical) confidence interval. However, this EnDF can be quite different from the SDF, even if the Bootstrap parameter estimates agree closely with those of the original sample. Such differences can be found, in particular, where the quantiles are not a linear function of the distribution parameters, bearing in mind that the mean of a non-linear function is not merely the function value at the mean of the arguments. Equations (6) and (12) show that the Gumbel quantiles depend linearly on the parameters but the Weibull quantiles do not, most notably because of the shape parameter, α. Therefore, differences between the EnDF and the SDF may be expected for the Weibull distribution, especially if the sample size is small. However, when the sample size increases, these differences gradually disappear, since the distributions of the parameters will shrink to point distributions with increasing N.

Concluding remarks

When determining the confidence limits to a frequency distribution by Bootstrapping, the Gaussian approach is clearly vulnerable to the presence of skewness in the resampled quantiles, which may lead to pronounced under- or over-estimation depending upon the sign of the skewness. Clearly, for a Gaussian approximation to deliver good results, the skewness of the quantiles also needs to be close to zero. In turn, the skewness of the quantiles was found to depend upon the distributions of the resampled parameters of the parent frequency distribution. Compared to the Gumbel distribution, interactions for the Weibull distribution are more complex owing to the non-linear dependency of the quantiles on the parameters, thereby enhancing the skewness of the former.

With regard to the initial sample size, the results obtained for both the Gumbel and the Weibull distribution demonstrate that Bootstrapping samples of size 100 presents fewer problems in terms of resampled quantile distributions than samples of size 30. Unfortunately, the latter are generally more representative of the data sets generally

available for hydrological frequency analysis. Because of the poor behaviour of the Gaussian appriximation for the three-parameter distribution, the present analysis has been unable to validate the Bootstrap confidence intervals for sample sizes of 30: in other cases, the comparisons indicate that the Bootstrap confidence intervals will be relaible. In the exceptional case, the behaviour of the Bootstrap confidence limits arises from the distributional properties of the resampled quantiles and these have been shown to relate to the distributional properties of the resampled parameter values. These considerations suggest that the Bootstrap can provide reasonable results in practice, subject to critical evaluation. In such cases, the percentile method is to be preferred to the Gaussian approach.

Cheng and Amin (1983) have shown that the MPS method for estimating the parameters of a frequency distribution has particular advantages over the MLH method in cases involving an unknown location parameter. However, application of MPS to Bootstrap resamples has been found to be subject to a further condition relating to the multiplicity of the ordered resamples. This condition on the true shape parameter a is particularly important for small initial sample sizes, again of the order of those encountered in hydrological frequency analyses, and small values of the shape parameter; if ignored, it may lead to inconsistent parameter estimates and uncertainty measures.

References

Benjamini, Y. and Harpaz, Y., 1986. Observational rainfall-runoff analysis for estimating effects of cloud seeding on water resources in Northern Israel. *J. Hydrol.*, **83**, 299–306.

Bobée, B., 1973. Sample error of T-year events computed by fitting a Pearson Type 3 distribution. *Water Resour. Res.*, **9**, 1264–1270.

Boogaard, H.F.P. van den, Kramer, J. and Spaan, G.B.H., 1999. *Theoretical reference document, EVA Version 1.0*. WL | Delft Hydraulics, Delft, The Netherlands.

Burn, D.H., 2003. The use of resampling for estimating confidence intervals for single site and pooled frequency analysis. *Hydrolog. Sci. J.*, **48**, 25–38.

Cheng, R.C.H. and Amin, N.A.K., 1983. Estimating parameters in continuous univariate distributions with a shifted origin. *J. Roy. Statist. Soc.*, *Series B*, **45**, 394–403.

Davison, A.C. and Hinkley, D.V., 1997. *Bootstrap methods and their application.* Cambridge University Press, Cambridge, UK. 582pp.

Dunn, P.K., 2001. Bootstrap confidence intervals for predicted rainfall quantiles. *Int. J. Climatol.*, **21**, 89–94.

Efron, B. and Tibshirani, R.J., 1993. *An introduction to the Bootstrap*. Chapman and Hall, London, UK. 436pp.

Freedman, D.A., 1981. Bootstrapping regression models. *Ann. Statist.*, **9**, 1218–1228.

Ghosh, K. and Rao Jammalamadaka, S., 2001. A general estimation method using spacings. *J. Statist. Plan. Infer.*, **93**, 71–82.

Jenkinson, A.F., 1969. *Estimation of maximum floods*. Chapter 5 of Technical Note 98, World Meteorological Organisation, Geneva, Switzerland. WMO No 233, 183–257.

Kaczmarek, Z., 1957. Efficiency of the estimation of floods with a given return period. *IAHS Publication no. 45* (General Assembly, Toronto, Vol. 3), 144–159.

Kite, G.W., 1975. Confidence limits for design events. *Water Resour. Res.*, **11**, 48–53.

Lall, U. and Sharma, A., 1996. A nearest neighbour Bootstrap for resampling hydrologic time series. *Water Resour. Res.*, **32**, 679–693.

Moss, M.E. and Tasker, G.D., 1991. An intercomparison of hydrological network-design technologies. *Hydrolog. Sci. J.*, **36**, 209–221.

Natural Environment Research Council, 1975. *Flood Studies Report, Volume I: Hydrological Studies.* NERC, London, UK. 550pp.

Oliveira, G.C., Kelman, J., Periera, M.V.F. and Stedinger, J.R., 1988. A representation of spatial correlations in large stochastic seasonal streamflow models. *Water Resour. Res.*, **24**, 781–785.

Pereira, M.V.F., Oliveira, G.C., Costa, C.G. and Kelman, J., 1984. Stochastic streamflow models for hydroelectric schemes. *Water Resour. Res.*, **20**, 379–90.

Potter, K.W. and Lettenmaier, D.P., 1990. A comparison of regional flood frequency estimation methods using a resampling method. *Water Resour. Res.*, **26**, 415–424.

Ranneby, B., 1984. The maximum spacing method: an estimation method related to the maximum likelihood method. *Scand. J. Statist.*, **11**, 93–112.

Shao, J. and Tu, D., 1995. *The Jack-knife and Bootstrap.* Springer, New York, USA. 516pp.

Sharma, A., Tarboton, D.G. and Lall, U., 1997. Streamflow simulation: a non-parametric approach. *Water Resour. Res.*, **33**, 291–308.

Tasker, G.D. and Dunne, P., 1997. Bootstrap position analysis for forecasting low flow frequency. *J. Water Resour. Plan. Manage.*, **123**, 359–367.

Tung, Y.K. and Mays, L.W., 1981. Reducing hydrologic parameter uncertainty. *J. Water Resour. Plan. Manage.*, **107**, 245–262.

Vogel, R.M. and Shallcross, A.L., 1996. The moving blocks Bootstrap versus parametric time series models. *Water Resour. Res.*, **32**, 1875–1882.

Wilks, D.S., 1997. Resampling hypothesis tests for autocorrelated fields. *J. Climate*, **10**, 65–82.

Zucchini, W. and Adamson, P.T., 1988. On the application of the Bootstrap to assess the risk of deficient annual inflows to a reservoir. *Water Resour. Manage.*, **2**, 245–254.

Zucchini, W. and Adamson, P.T., 1989. Bootstrap confidence intervals for design storms from exceedance series. *Hydrolog. Sci. J.*, **34**, 41–48.

Zwiers, F.W., 1990. The effect of serial correlation on statistical inferences made with resampling procedures. *J. Climate*, **3**, 1452–1461.

Appendix

The occurrence of inconsistent MPS estimates for the parameters of a Weibull distribution

For an MPS estimate of the parameters Θ in the CDF $F(\cdot|\Theta)$ or corresponding PDF $f(\cdot|\Theta)$, the following function must be maximised:

$$H(\Theta) := \log\big(F(x_1|\Theta)\big) \; + $$
$$\sum_{n=2}^{N} \log\big(F(x_n|\Theta) - F(x_{n-1}|\Theta)\big) \; + $$
$$\log\big(1 - F(x_N|\Theta)\big) \tag{A.1}$$

If $x_n = x_{n-1}$ for some n, the term $\log\big(F(x_n|\Theta) - F(x_{n-1}|\Theta)\big)$ must be replaced by the appropriate limit $\log\big(f(x_n|\Theta)\big)$ which corresponds to the formulation used in the MLH criterion. In this way, the MPS function of Eqn.A.1 will have the following form when the multiplicity of the first data point x_1 is $m \geq 1$:

$$H(\Theta) := \log\big(F(x_1|\Theta)\big) \; + (m-1)\cdot\log\big(f(x_1|\Theta)\big) + H^{(m+1)}(\Theta)$$
$$\tag{A.2}$$

The part $H^{(m+1)}(\cdot)$ of the MPS function $H(\cdot)$ consists of a superposition of terms $\log\big(F(x_n|\Theta) - F(x_{n-1}|\Theta)\big)$ and/or $\log\big(f(x_n|\Theta)\big)$ for the sample points x_n with $n > m$. For the Weibull distribution of Eqn.(10), the MPS-function of Eqn.A.2 will then be of the form:

$$H(\Theta) := \log\big(1 - \exp(-\xi_1^\alpha)\big) \; + $$
$$(m-1)\cdot\Big\{\; \log\tfrac{\alpha}{\sigma} + (\alpha-1)\cdot\log(\xi_1^\alpha) - \xi_1^\alpha \;\Big\} \; + $$
$$H^{(m+1)}(\Theta) \tag{A.3}$$

The ξ_1 in this equation is defined by $\xi_1 := \frac{x_1 - x_0}{\sigma}$.

To obtain a regular estimate for the parameters Θ, the maximum of $H(\Theta)$ should be found for a location parameter x_0 that is strictly less than x_1, or equivalently $\xi_1 > 0$. For $x_0 \uparrow x_1$, and equivalently $\xi_1 \downarrow 0$, the combination $H^{(1)}(\Theta) := \log\big(1 - \exp(-\xi_1^\alpha)\big) + (m-1)\cdot(\alpha-1)\cdot\log(\xi_1^\alpha)$ will dominate all the other parts of $H(\Theta)$ because $\lim_{\xi_1 \downarrow 0}\big|H^{(1)}(\Theta)\big| = \infty$ while all other terms of $H(\Theta)$ (and in particular $H^{(m+1)}(\Theta)$) remain bounded. For $\xi_1 \downarrow 0$, $\log\big(1 - \exp(-\xi_1^\alpha)\big)$ equals $\alpha\cdot\log(\xi_1)$ plus a term that converges to zero. As a result $(m\cdot\alpha - m + 1)\cdot\log(\xi_1^\alpha)$ is the dominant part of the MPS function when $\xi_1 \downarrow 0$. For $m\cdot\alpha - m + 1 > 0$, it will hold that $\lim_{\xi_1 \downarrow 0} H(\Theta) = -\infty$, with the consequence that a maximum for th: $H(\Theta)$ is found for a $\xi_1 > 0$ or, equivalently, an estimate for the location parameter x_0 that is strictly less than the smallest sample point x_1. From the condition $m\cdot\alpha - m + 1 > 0$, it follows that $\alpha > 1 - \frac{1}{m}$.

This result shows that only for $m=1$, i.e. a single multiplicity of the smallest sample point x_1, will the MPS criterion always give consistent estimates for the parameters Θ. If this multiplicity is greater than 1, then the condition must be satisfied that the true shape parameter α is greater than $1 - \frac{1}{m}$. Note that this regularity condition for MPS is less strict than that for MLH, since in the latter case the shape parameter must always be greater than 1, irrespective the multiplicity of x_1.