Hydrology and Earth System Sciences, 8(2), *183–190* (2004) © EGU

Hydrology & Earth System Sciences

Sampling variance of flood quantiles from the generalised logistic distribution estimated using the method of L-moments

Thomas R. Kjeldsen and David A. Jones

Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

E-mail for corresponding author: trkj@ceh.ac.uk

Abstract

The method of L-moments is the recommended method for fitting the three parameters (location, scale and shape) of a Generalised Logistic (GLO) distribution when conducting flood frequency analyses in the UK. This paper examines the sampling uncertainty of quantile estimates obtained using the GLO distribution for single site analysis using the median to estimate the location parameter. Analytical expressions for the variance of the quantile estimates were derived, based on asymptotic theory. This has involved deriving expressions for the covariance between the sampling median (location parameter) and the quantiles of the estimated unit-median GLO distribution (growth curve). The accuracy of the asymptotic approximations for many of these intermediate results and for the quantile estimates was investigated by comparing the approximations to the outcome of a series of Monte Carlo experiments. The approximations were found to be adequate for GLO shape parameter values between –0.35 and 0.25, which is an interval that includes the shape parameter estimates for most British catchments. An investigation into the contribution of different components to the total uncertainty showed that for large returns periods, the variance of the growth curve is larger than the contribution of the median. Therefore, statistical methods using regional information to estimate the growth curve should be considered when estimating design events at large return periods.

Keywords: flood frequency analysis, Flood Estimation Handbook, single site, annual maximum series, generalised logistic distribution, uncertainty

Introduction

Following the publication of the Flood Estimation Handbook (FEH), the generalised logistic (GLO) distribution is recommended as the standard for flood frequency analysis in the UK (IH, 1999) with a variant of the Method of L-moments (Hosking and Wallis, 1997) for estimating model parameters. The GLO distribution has not been applied in hydrology to the same extent as the Generalised Extreme Value (GEV) distribution, but Ahmad *et al.* (1988) found the three parameter Log-logistic distribution to perform better than other, more commonly encountered, distributions for modelling floods in Scotland. The GLO distribution adopted by IH (1999) is a re-parameterised version of the Log-logistic distribution used by Ahmad *et al.* (1988).

The sampling variance of the T-year events can be estimated through a number of different methods, such as the computer oriented methods Jack-knife and Bootstrap (Efron and Tibshirani 1993) or, alternatively, analytically through Taylor approximations. Analytical expressions of sampling variance have been derived for a number of distributions with different parameter estimation techniques used in hydrology, for example, the generalised extreme value distribution (Hosking, 1985) and the Log-normal distribution (Hoshi *et al.* 1984). This paper presents an analytical approach for the GLO distribution for the particular parameter estimation method adopted in the FEH. The method considers only uncertainty arising as a result of limited sample records and not other sources of uncertainty encountered in practical flood frequency analysis, such as mis-specification of statistical distributions and errors in the measured data.

Definitions of L-moments

Similar to ordinary product moments, L-moments and probability weighted moments (PWM) can be used to summarise probability distributions and observed samples. L-moments, as defined by Hosking (1990), are linear combinations of PWMs. Greenwood *et al.* (1979) summarises the theory of PWM and define them as

$$\beta_r = E\{X[F_X(x)]^r\} \quad (1)$$

where β_r is the r'th order PWM and $F_X(x)$ is the cumulative distribution function (cdf) of X. Unbiased estimators (b_i) of the first three PWMs are given by Hosking and Wallis (1997) as

$$b_r = n^{-1}\sum_{j=r+1}^{n}\frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)}x_{j:n} \quad (2)$$

where n is the sample size and $x_{j:n}$ represents an ordered sample $x_{1:n} \le x_{2:n} \le \ldots \le x_{n:n}$. The sample L-moments ($l_r$) are linear combinations of sample PWM calculated as

$$\begin{aligned} l_1 &= b_0 \\ l_2 &= 2b_1 - b_0 \\ l_3 &= 6b_2 - 6b_1 + b_0 \end{aligned} \quad (3)$$

The sample L-moment ratios, t_r, are based on the sample L-moments and defined as

$$\begin{aligned} t_2 &= \frac{l_2}{l_1} \\ t_3 &= \frac{l_3}{l_2} \end{aligned} \quad (4)$$

where t_2 is the L-CV and t_3 the L-SKEW (Hosking and Wallis, 1997).

The GLO distribution

The probability density function (pdf) of the three parameter GLO distribution is defined by Hosking and Wallis (1997) as

$$f(x) = \frac{\alpha^{-1}\exp(-(1-\kappa)y)}{(1+\exp(-y))^2},$$

$$y = \begin{cases} -\kappa^{-1}\ln(1-\kappa(x-\xi)/\alpha) & \kappa \ne 0 \\ (x-\xi)/\alpha & \kappa = 0 \end{cases} \quad (5)$$

where (ξ,α,κ) are the location, scale and shape parameters, respectively. The range of x is defined as: $-\infty < x \le \xi + \alpha/\kappa$ if $k > 0$; $-\infty < x < \infty$ if $k = 0$; $\xi + \alpha/\kappa \le x < \infty$ if $k < 0$. The T-year event x_T is defined as

$$x_T = \xi + \frac{\alpha}{\kappa}\left(1-(T-1)^{-\kappa}\right) = \xi\left[1+\frac{\beta}{\kappa}\left(1-(T-1)^{-\kappa}\right)\right] = \xi\, z_T \quad (6)$$

where $\beta = \alpha/\xi$, T is the return period and z_T is the growth curve at T, defined by the square brackets in Eqn. (6). In the special case where the shape parameter $\kappa = 0$, the GLO distribution reduces to a 2 parameter Logistic distribution.

The parameter estimation method adopted in this study is the method adopted by IH (1999) and is a variant of the method of L-moments described by Hosking and Wallis (1997). The location parameter (ξ) is estimated by first equating the distribution median to the sample median (m), i.e.

$$\hat{\xi} = m \quad (7)$$

Next, the shape parameter (κ) and the scale parameter (β) are estimated as

$$\begin{aligned} \hat{\kappa} &= -t_3 \\ \hat{\beta} &= \frac{t_2\hat{\kappa}\sin(\pi\hat{\kappa})}{\hat{\kappa}\pi\sin(\hat{\kappa}+t_2) - t_2\sin(\pi\hat{\kappa})} \end{aligned} \quad (8)$$

By using the estimation method outlined above, the median of the GLO distribution is fitted to the sample median rather than the distribution mean value being fitted to the sample mean, as is most commonly the case (see, e.g. Hosking and Wallis, 1997). The median was adopted in FEH to minimise the potential impact of outliers in the observed series (IH, 1999).

Variance of T-year event

The method for approximating variances used in this paper is based on Taylor series expansions. Consider an estimator $\hat{x}$ of x, which is derived from a vector of estimated parameters $\hat{\alpha}$, whose true value is α. Suppose that $\hat{x} = g(\hat{\alpha})$, and that the covariance matrix, V, of the estimated parameters is small. Then $\hat{x} \approx g(\alpha) + d^T(\hat{\alpha}-\alpha)$, where the elements d_i in the vector d are given as $d_i = \partial g/\partial\alpha_i$. It then follows that $\mathrm{var}\{\hat{x}\} \approx d^T V d$. In this paper, the quantity being estimated, , is the T-year flood event. This estimate is derived from estimates of the GLO distribution's parameters, which themselves are based on estimated L-moment ratios, which again, are based on L-moments for which statistical properties are known, based on the properties of the PWMs.

To get an expression of the variance of the T-year estimate, a first order Taylor expression of Eqn. (6) at the population values of the median and the growth curve (ξ and z_T) is applied as described by Hosking and Wallis (1997)

$$\mathrm{var}\{\hat{x}_T\} = (z_T)^2\,\mathrm{var}\{m\} + \xi^2\,\mathrm{var}\{\hat{z}_T\} + 2z_T\xi\,\mathrm{cov}\{m,\hat{z}_T\} \quad (9)$$

The contribution from each of the three terms in the equation above will be discussed later. However, the components of the equation can be derived as follows. The approximate variance of the sampling median, var$\{m\}$, is given by Kendall and Stuart (1963) as

$$\mathrm{var}\{m\} = \frac{1}{4nf^2(\xi)} = \frac{4\alpha^2}{n} \tag{10}$$

where f is the pdf of the GLO distribution defined in Eqn. (5). The variance of the growth curve, $\mathrm{var}\{\hat{z}_T\}$, is derived using a Taylor expansion

$$\mathrm{var}\{\hat{z}_T\} \approx \left(\frac{\partial z_T}{\partial \beta}\right)_p^2 \mathrm{var}\{\hat{\beta}_i\} + \left(\frac{\partial z_T}{\partial \kappa}\right)_p^2 \mathrm{var}\{\hat{\kappa}_i\} + 2\left(\frac{\partial z_T}{\partial \beta}\right)_p \left(\frac{\partial z_T}{\partial \kappa}\right)_p \mathrm{cov}\{\hat{\beta},\hat{\kappa}\} \tag{11}$$

where the index p indicates that the derivatives are evaluated at their population values. The derivatives are not displayed but are easily obtained. Similarly, an analytical expression of the covariance between the index flood and the growth curve is given as

$$\mathrm{cov}\{m,\hat{z}_T\} \approx \mathrm{cov}\left\{m, \left(\frac{\partial z_T}{\partial \beta}\right)_p \hat{\beta} + \left(\frac{\partial z_T}{\partial \kappa}\right)_p \hat{\kappa}\right\} = \left(\frac{\partial z_T}{\partial \beta}\right)_p \mathrm{cov}\{m,\hat{\beta}\} + \left(\frac{\partial z_T}{\partial \kappa}\right)_p \mathrm{cov}\{m,\hat{\kappa}\} \tag{12}$$

In the following sections, the necessary expressions of variance-covariance of the parameters will be deduced as functions of the variance-covariance of PWMs and the median using asymptotic results from Hosking (1986) and a set of new general expressions derived for this study.

VARIANCE OF PARAMETERS

The variance-covariance of the parameters is needed to evaluate the variance of the growth curve in Eqn. (11) and is derived approximately by a Taylor expansion around the population values of the L-moment ratios and the median. First, the scale parameter, β, is considered

$$\hat{\beta} = \frac{t_2 \hat{\kappa} \sin(\pi\hat{\kappa})}{\hat{\kappa}\pi(\kappa + t_2) - t_2 \sin(\pi\hat{\kappa})} = g(t_2, t_3) \tag{13}$$

$$\mathrm{var}\{\hat{\beta}\} \approx \left(\frac{\partial g}{\partial t_2}\right)_p^2 \mathrm{var}\{t_2\} + \left(\frac{\partial g}{\partial t_3}\right)_p^2 \mathrm{var}\{t_3\} + 2\left(\frac{\partial g}{\partial t_2}\right)_p \left(\frac{\partial g}{\partial t_3}\right)_p \mathrm{cov}\{t_2, t_3\} \tag{14}$$

where p indicates that the Taylor expansion is made at the population values of the L-moment ratios. Again, the derivatives are easily obtained but not displayed. Next, the shape parameter, κ, is considered, where

$$\hat{\kappa} = -t_3 \qquad \mathrm{var}\{\hat{\kappa}\} = \mathrm{var}\{t_3\} \tag{15}$$

The covariance between scale and shape parameters is derived as

$$\mathrm{cov}\{\hat{\beta},\hat{\kappa}\} \approx \mathrm{cov}\left\{\left(\frac{\partial g}{\partial t_2}\right) t_2 + \left(\frac{\partial g}{\partial t_3}\right) t_3, -t_3\right\} = -\left(\frac{\partial g}{\partial t_2}\right) \mathrm{cov}\{t_2, t_3\} - \left(\frac{\partial g}{\partial t_3}\right) \mathrm{var}\{t_3\} \tag{16}$$

The sampling variance-covariance of the L-moment ratios, used in Eqns. (14), (15) and (16), is estimated using Taylor approximations of Eqn. (4) around the population values of the L-moments, thus the variance-covariance is derived straightforwardly as:

$$\mathrm{cov}\{t_r, t_k\} \approx \sum \left(\frac{\partial t_r}{\partial l_i}\right)_p \left(\frac{\partial t_k}{\partial l_j}\right)_p \mathrm{cov}\{l_i, l_j\}, \quad i,j = 1,2,3 \tag{17}$$

where the functions t_r and t_k are the relevant L-moment ratio in Eqn. (4). The variance-covariance of the L-moments, cov$\{l_i, l_j\}$, is derived straight forwardly by using the covariance operator in combination with Eqn. (3) pending the variance-covariance of the PWMs derived in the following section.

Next, the covariances between the median and the scale and shape parameters are derived, as they are needed to calculate the covariance between the location parameter and the growth curve specified in Eqn. (12). The results have been obtained by approximating the parameter estimators $\hat{\beta}$ and $\hat{\kappa}$ in Eqns. (13) and (15) by linear functions of the L-moment ratios, which themselves are approximated by linear functions of the L-moments. Thus, the covariance between the median and the parameters will be reduced to expressions of covariances between the median and the L-moments.

First, the covariance between location and scale parameter is derived as a function of the covariance between the median and the L-moments as

$$\mathrm{cov}\{m,\hat{\beta}\} \approx \mathrm{cov}\left\{m, \left(\frac{\partial g}{\partial t_2}\right)_p t_2 + \left(\frac{\partial g}{\partial t_3}\right)_p t_3\right\} = \left(\frac{\partial g}{\partial t_2}\right)_p \mathrm{cov}\{m, t_2\} + \left(\frac{\partial g}{\partial t_3}\right)_p \mathrm{cov}\{m, t_3\}$$

$$= \left(\frac{\partial g}{\partial t_2}\right)_p \mathrm{cov}\left\{m, \left(\frac{\partial t_2}{\partial l_1}\right)_p l_1 + \left(\frac{\partial t_2}{\partial l_2}\right)_p l_2\right\} +$$

$$\left(\frac{\partial g}{\partial t_3}\right)_p \operatorname{cov}\left\{m, \left(\frac{\partial t_3}{\partial l_2}\right)_p l_2 + \left(\frac{\partial t_3}{\partial l_3}\right)_p l_3\right\}$$

$$= \left(\frac{\partial g}{\partial t_2}\right)_p \left[\left(\frac{\partial t_2}{\partial l_1}\right)_p \operatorname{cov}\{m, l_1\} + \left(\frac{\partial t_2}{\partial l_2}\right)_p \operatorname{cov}\{m, l_2\}\right] + \left(\frac{\partial g}{\partial t_3}\right)_p \left[\left(\frac{\partial t_3}{\partial l_2}\right)_p \operatorname{cov}\{m, l_2\} + \left(\frac{\partial t_3}{\partial l_3}\right)_p \operatorname{cov}\{m, l_3\}\right] \quad (18)$$

where Taylor expansions of both the scale parameter (β) and the L-CV (t_2) and L-SKEW (t_3) have been applied. Next, the covariance between location and shape parameter is derived

$$\operatorname{cov}\{m, \hat{\kappa}\} = -\operatorname{cov}\{m, t_3\} \approx -\operatorname{cov}\left\{m, \left(\frac{\partial t_3}{\partial l_2}\right) l_2 + \left(\frac{\partial t_3}{\partial l_3}\right) l_3\right\} = -\left(\frac{\partial t_3}{\partial l_2}\right)\operatorname{cov}\{m, l_2\} - \left(\frac{\partial t_3}{\partial l_3}\right)\operatorname{cov}\{m, l_3\} \quad (19)$$

The covariance between the median and the L-moments will be derived in the following section.

COVARIANCE OF PWM AND THE SAMPLE MEDIAN

The asymptotic variance-covariance of the sample probability weighted moments (PWM) are given by Hosking (1986) as

$$n\operatorname{cov}\{b_r, b_s\} = B_{rs} = J_{rs} + J_{sr} \quad (20)$$

where n is the sample size and J_{rs} is calculated as

$$J_{rs} = \frac{\alpha^2}{1+\kappa} \frac{\Gamma(1+2\kappa)\Gamma(r+s+1-2\kappa)}{\Gamma(r+s+2)} \; {}_3F_2\left[\begin{matrix} 1, s+1, 1+2\kappa \\ r+s+2, 2+\kappa \end{matrix}\right], \quad |\kappa| < \frac{1}{2} \quad (21)$$

where ${}_3F_2$ is a generalised hypergeometric function of unit argument.

As the L-moments are linear combinations of the PWMs, as expressed in Eqn. (3), the estimation of the variance-covariance of the L-moments is straightforward.

The sampling covariance between the median and the PWM is derived in Appendix A, and the final expression stated here

$$n\operatorname{cov}\{m, b_r\} = \frac{1}{f(\xi)}\left[\frac{1}{2}\{(r+1)\beta_r - r\beta_{r-1}\} + \frac{1}{2} r\left\{\int_0^{1/2} t^{r-1} x(t)dt\right\} - (r+1)\left\{\int_0^{1/2} t^r x(t)dt\right\}\right] \quad (22)$$

where $f(\xi)$ is the pdf of the GLO distribution evaluated at the median and $x(t)$ is the inverse of the cdf of the considered distribution. In the case of the GLO distribution, the integral in the expression above can be solved as

$$\int_0^{1/2} t^r x(t)dt = \left(\xi + \frac{\alpha}{\kappa}\right)\frac{1}{r+1}\left(\frac{1}{2}\right)^{r+1} - \frac{\alpha}{\kappa} G(r, \kappa) \quad (23)$$

where G is the incomplete beta function defined by Gradshteyn and Ryzhik (1980) as

$$G(r, \kappa) = \int_0^{1/2} t^r \left(\frac{1}{t} - 1\right)^\kappa dt = \left(\frac{1}{2}\right)^{r+1} \frac{1}{r+1} \, {}_2F_1\left[r+1, -\kappa, r+2, \frac{1}{2}\right] \quad (24)$$

where ${}_2F_1$ is the hypergeometric function. For the case where $\kappa = 0$, i.e. the 2 parameter logistic distribution (IH, 1999), the relevant covariances reduce to

$$\begin{aligned} n\operatorname{cov}\{m, b_0\} &= \alpha^2 4\ln(2) \\ n\operatorname{cov}\{m, b_1\} &= \alpha^2 2\ln(2) \end{aligned} \quad (25)$$

The accuracy of these expressions will be evaluated through Monte Carlo simulations. No similar expressions were identified in the literature and therefore this is believed to be the first derivation of this covariance. Finally, the covariance between the median and the L-moments are easily derived by combining Eqn. (22) and Eqn. (3). The results from this section are fed back into the expressions for variance-covariance of the parameters, which, again, are needed for evaluation of the variance of the quantile estimator in Eqn. (9).

Monte Carlo simulations

To assess the performance of the derived estimator of sampling variance of the T-year event in Eqn. (9) and the covariance between the median and the PWM, a series of Monte Carlo simulations was carried out. In each experiment, the Monte Carlo simulations are carried out by generating 10 000 random GLO samples of varying sample size (n=10, 100, 5000). Through a Monte Carlo study it is possible to investigate each of the analytical expressions derived as part of this study. However, for practical reasons, only two properties will be investigated. Firstly, the newly derived expression for covariance between the median and the PWM is the fundamental building block of many of the preceding expressions, hence, verification of this result is necessary. Secondly, the variance of the T-year event will be investigated as it is of direct interest for practical flood risk assessment.

The GLO distribution has an upper bound of $\xi + \alpha/\kappa$ for $\kappa > 0$ but no lower bound, i.e. negative values of annual maximum peak flow can occur in the simulation study. To

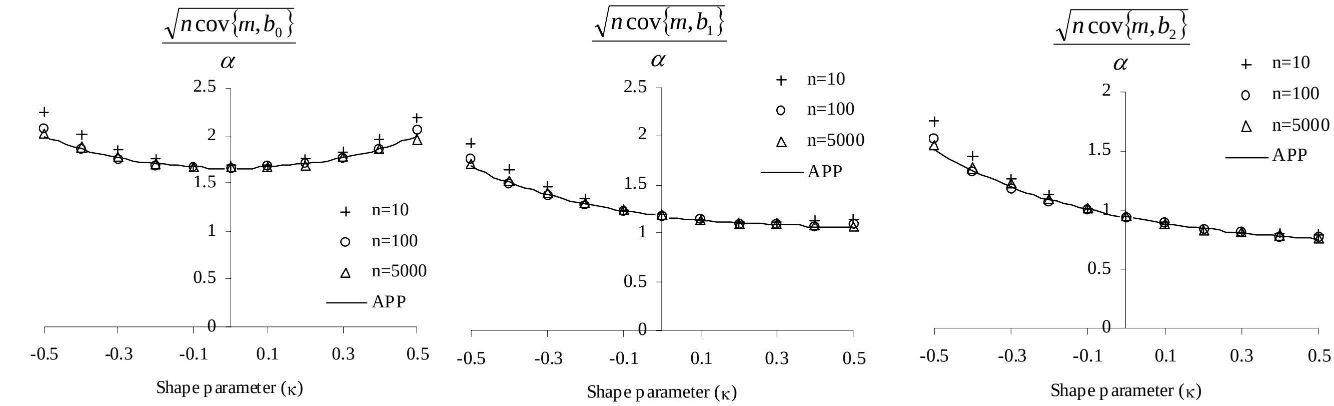

Fig. 1. *Comparison of analytical solution (APP) and results from Monte Carlo simulations for (a)* $\mathrm{cov}\{\hat{m}, b_0\}$ *(b)* $\mathrm{cov}\{\hat{m}, b_1\}$ *and (c)* $\mathrm{cov}\{\hat{m}, b_2\}$.

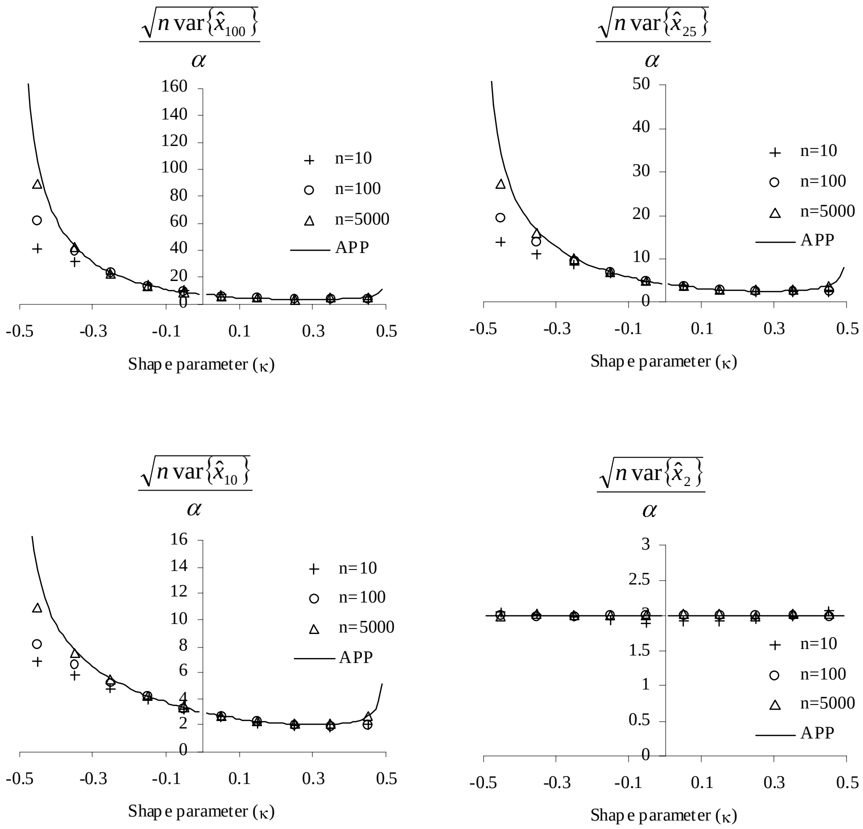

Fig. 2. *Comparison of analytical solution (APP) and results from Monte Carlo simulations for (a)* $\mathrm{var}\{\hat{x}_{10}\}$ *(b)* $\mathrm{var}\{\hat{x}_{25}\}$ *and (c)* $\mathrm{var}\{\hat{x}_{100}\}$.

minimise the problem, a set of parameter values was estimated from an observed annual maximum series. Subsequent comparison of the Monte Carlo results with the results from a censored simulation showed that the effect of the negative values was insignificant.

Figure 1a–c shows comparisons of the analytical expression of covariance between the median and the PWM, as outlined in Eqn. (22), and the results obtained from a Monte Carlo simulation for different sample sizes. Generally, the analytical approach corresponds well to the Monte Carlo results, even for small samples. However, some deviation is observed for $\kappa < -0.25$, where the analytical solution slightly underestimates the covariance, especially for small samples.

Comparisons of the analytical variance of the T-year event with the outcome of the Monte Carlo simulations are shown in Fig. 2a–d for different sample sizes. The goodness of fit

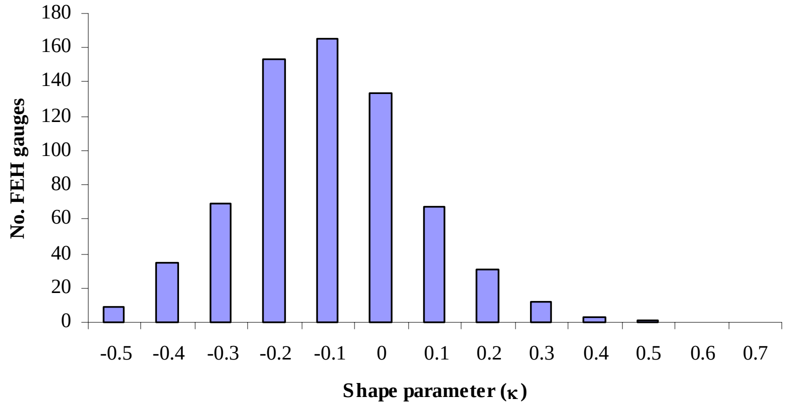

Fig. 3. *Histogram of estimated shape parameter κ for 955 UK gauges.*

of the analytical approximations depends on the shape parameter, the sample size and the considered return period. The approximations can only be evaluated for $-0.5 < \kappa < 0.5$ as determined by the covariance of the PWMs as expressed in Eqn. (21). For values of κ close to the edges of the interval, the analytical expressions deteriorate regardless of sample size. However, for $-0.35 \le \kappa \le 0.35$ the approximations perform relatively well, even for small samples. Comparing this interval with the distribution of estimates of shape parameters obtained from 955 UK gauges, as shown in Fig. 3, it is noted that a large fraction of the gauges falls within this interval. Finally, the considered return period has little effect on the performance of the analytical expressions.

Application example

The derived method was applied to two relatively long annual maximum series of peak flow from the UK. A GLO distribution was fitted to each of the two series using the method of L-moments described earlier. The resulting flood frequency curves are plotted in Fig.4, together with the 95% confidence intervals (Stedinger *et al.*, 1993) given as

$$\left(\hat{x}_T - z_{1-\alpha/2}\sqrt{\mathrm{var}\{\hat{x}_T\}},\ \hat{x}_T + z_{1-\alpha/2}\sqrt{\mathrm{var}\{\hat{x}_T\}}\right) \qquad (26)$$

where $z_{1-\alpha/2}$ is the upper 100(α/2)% percentile of the normal distribution.

While such approximate confidence intervals do not account for the skewness of the actual sampling distribution of the estimate, they are useful here in allowing a simple graphical assessment of the relative size of uncertainty at different return periods and of the importance of different sources of uncertainty. As the population values of the L-moment ratios and parameters are unknown the expression

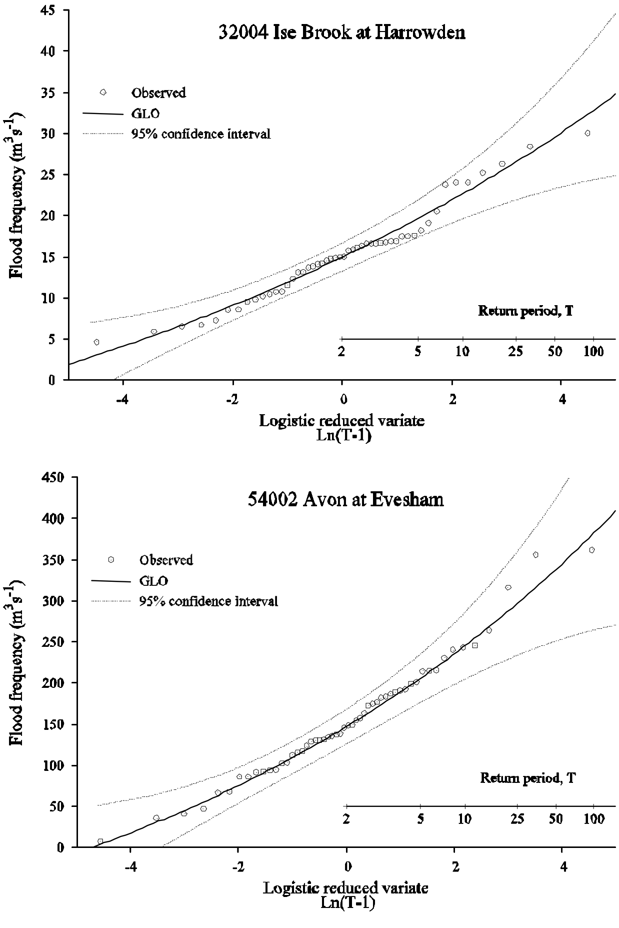

Fig. 4. *Flood frequency curves with 95% confidence intervals for gauging stations 32004 and 54002.*

of $\mathrm{var}\{\hat{x}_T\}$ in Eqn. (26) is evaluated using the sample values of these from the annual maximum series rather than population values. For both series the derived confidence intervals appear plausible from a visual inspection. Note

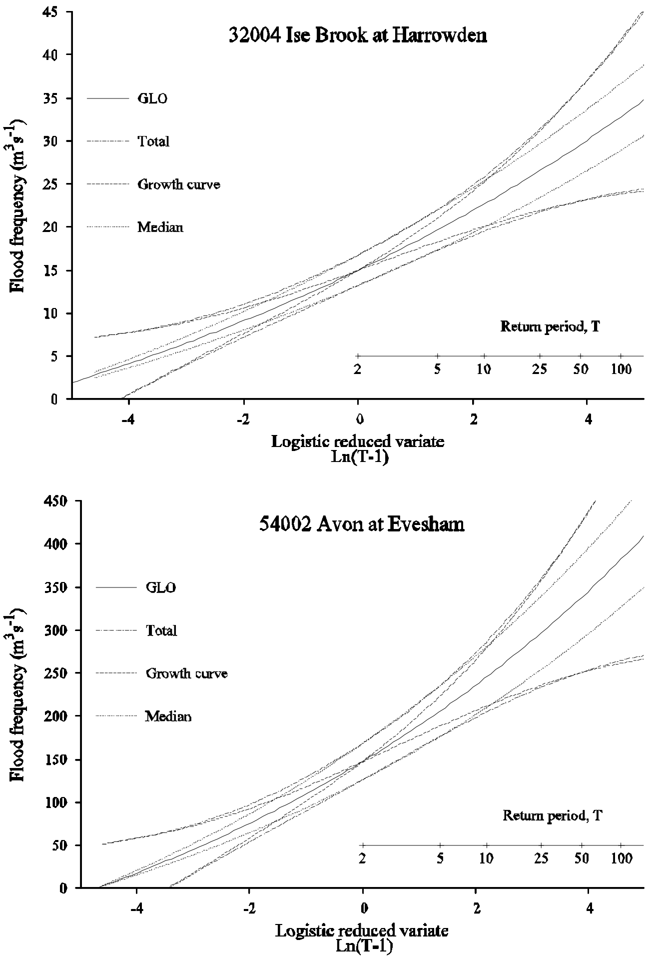

Fig. 5. *Comparison of 95% confidence intervals calculated through Eqn. (26) for (1) total sample variance (2) known median (3) known growth curve.*

that for very low return periods, the lower limit of the 95% confidence interval can fall below zero.

Uncertainty contributions

The total prediction uncertainty of the T-year event, as shown in Eqn. (9), can be divided into three contributing sources: (1) variability of sample median, (2) variability of growth curve and (3) covariance between sample median and growth curve. Consider a situation where the population values of either the median or the growth curves are known. In both cases, the variance of the known entity and the covariance between the sample median and the growth curve would equal zero. To assess the importance of each of these three sources of uncertainty, each of the three terms in Eqn. (9) was evaluated using the Taylor approximations for the two test catchments. The 95% confidence intervals for each of the two cases were calculated and compared to the confidence interval derived using the total variance, as shown in Fig. 5. At high return periods the sample uncertainty of the growth curve provides a good approximation to the total sample uncertainty and at return periods around $T = 2$, the sampling uncertainty of the median dominates, as the sample uncertainty of the growth curve equals zero at $T = 2$. For both catchments, the sample uncertainty of the median becomes of secondary importance at return periods around $T = 25$.

Conclusions

Asymptotic expressions of the sampling variance of quantiles from a GLO distribution, with parameters estimated using a parameter estimation scheme used in the Flood Estimation Handbook (IH, 1999), were derived. These expressions included the covariance between the sample median and the PWMs which, to the best of the authors' knowledge, has not been derived previously. The asymptotic expressions were evaluated through a Monte Carlo simulation study. In general, the expressions were found to provide an accurate estimate of the sampling variance when the shape parameter takes values of $-0.35 < \kappa < 0.35$, which is where the sample estimates from a large proportion of annual maximum series from UK generally fall. However, outside this range the model performs rather poorly for the moderate sample sizes most often encountered in practical flood frequency analysis. Furthermore, the practical use of the method is limited by the fact that sampling variance can only be evaluated for shape parameters in the interval $0.5 < \kappa < 0.5$. However, sample estimates of the shape parameter obtained from real data can fall outside this range. Another possible limitation of the practical use is the complexity of the equation involved, often requiring numerical method for evaluation.

An investigation into the contribution of different components to the total uncertainty showed that, for large return periods, the variance of the growth curve is larger than the contribution of the median. Therefore, statistical methods using regional information to estimate the growth curve should be considered when estimating design events at large return periods.

References

Ahmad, M.I., Sinclair, C.D. and Werritty, A., 1988. Log-logistic flood frequency analysis. *J. Hydrol.*, **98**, 205–224.

Efron, B. and Tibshirani, R.J., 1993. *An Introduction to the Bootstrap.* Monographs on Statistics and Applied Probability 57, Chapman & Hall, USA.

Ferguson T.S., 1998. Asymptotic Joint distribution of Sample Mean and a Sample Quantile. Unpublished, but available at http://www.math.ucla.edu/~tom/papers/unpublished/meanmed.pdf

Gradshteyn, I.S. and Ryzhik, I.M., 1980. *Table of Integrals, Series and Products.* Academic Press, USA.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis J.R., 1979. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resour. Res.*, **15**, 1049–1054.

Hoshi, K., Stedinger J.R. and Burges S.J., 1984. Estimation of log-normal quantiles: Monte Carlo results and first order approximations. *J. Hydrol.*, **71**, 1–30.

Hosking, J.R.M., 1986. The Theory of Probability Weighted Moments. *Research Report RC12210*, IBM Research Division, Yorktown Heights, New York.

Hosking J.R.M. and Wallis J.R., 1997. *Regional Flood Frequency Analysis: an approach based on L-moments.* Cambridge University Press, USA.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F., 1985. Estimation of the generalised extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251–61.

IH, 1999. *The Flood Estimation Handbook.* Institute of Hydrology, Wallingford, UK.

Kendall, M.G. and Stuart, M.A., 1977. *The Advanced Theory of Statistics, Vol. 1. Distribution Theory.* 4th edn. Griffin, London, UK.

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

Appendix A: Covariance of PWMs and median

This appendix outlines the derivation of asymptotic formulae for the covariances of the PWMs and the median. For notational convenience, the derivation is slightly more general than this and the formulae actually relate to the simple sample estimator of any population quantile. Let p be the selected probability point and define the sample quantile estimator (of $m_p = x(p)$) to be

$$\hat{m}_p = x_{k:n}$$

where k is the integer nearest to pn, and where $x_{j:n}$ represents the ordered sample $x_{1:n} \le x_{2:n} \le \ldots \le x_{n:n}$. The median corresponds to p = ½.

The formulae can be derived by a minor modification to the theory provided by Ferguson (1998) in an unpublished paper. Firstly, define the ordered uniforms $u_{j:n}$ equivalent to $x_{j:n}$, so that $x_{j:n} = x(u_{j:n})$ and define also $p_{j:n} = j/(n+1)$. The first step in the derivation is to note that, for the purposes of asymptotic arguments, the sample (unbiased) PWMs can be approximated by the 'plotting-position' estimators

$$\tilde{b}_r = n^{-1}\sum_{j=1}^{n}(p_{j:n})^r x_{j:n} = n^{-1}\sum_{j=1}^{n}(p_{j:n})^r x(u_{j:n})$$

For formal purposes, one can define

$$\tilde{\beta}_r = n^{-1}\sum_{j=1}^{n}(p_{j:n})^r x(p_{j:n})$$

and note that $\tilde{\beta}_r \to \beta_r$ as $n \to \infty$.

The arguments required to apply the approach of Ferguson (1998) are based on the simple expansion

$$\tilde{b}_r = n^{-1}\sum_{j=1}^{n}(p_{j:n})^r \{x(p_{j:n}) + x'(p_{j:n})(u_{j:n} - p_{j:n}) + \ldots\}$$

and this then leads to the following joint asymptotic description for the sample quantile and the PWM estimators

$$\sqrt{n}(\tilde{b}_r - \tilde{\beta}_r) \xrightarrow{L} \int_0^1 t^r x'(t)\{W(t) - tW(1)\}\mathrm{d}t$$
$$\sqrt{n}(\hat{m}_p - m_p) \xrightarrow{L0} x'(p)\{W(p) - W(1)\}$$

where $W(t)$ is Brownian motion and where the symbol "$\xrightarrow{L}$" denotes convergence in law (convergence in distribution). For the second of the above equations, the finite summation and the integral can be related by noting the following correspondences:

$$(p_{j:n})^r \approx t^r, \quad x'(p_{j:n}) \approx x'(t), \quad u_{j:n} - p_{j:n} \approx W(t) - tW(1)$$

.

Note that $W(t) - tW(1)$ is a Brownian Bridge, *i.e.* Brownian motion conditioned on starting and ending at zero on the interval (0,1).

. It then follows that $n\,\mathrm{cov}(\hat{m}_p - m_p, \tilde{b}_r - \tilde{\beta}_r) \approx$

$$\approx x'(p)\int_0^1 t^r x'(t) E[\{W(t) - tW(1)\}\{W(p) - W(1)\}]\mathrm{d}t$$

$$= x'(p)\int_0^1 t^r x'(t)\{\min(p,t) - pt\}\mathrm{d}t$$

$$= x'(p)\left\{\int_0^p t^r x'(t)\{(1-p)t\}\mathrm{d}t + \int_p^1 t^r x'(t)\{p(1-t)\}\mathrm{d}t\right\}$$

$$= x'(p)\left\{(1-p)\int_0^p t^{r+1}x'(t)\,\mathrm{d}t + p\int_p^1 (t^r - t^{r+1})x'(t)\,\mathrm{d}t\right\}$$

$$= x'(p)\left\{\int_0^p t^{r+1}x'(t)\,\mathrm{d}t + p\int_p^1 t^r x'(t)\,\mathrm{d}t - p\int_0^1 t^{r+1}x'(t)\,\mathrm{d}t\right\}$$

The term in braces here can be evaluated by integrating by parts, giving

$$(1-p)\int_0^p t^{r+1}x'(t)\,\mathrm{d}t + p\int_p^1 (t^r - t^{r+1})x'(t)\,\mathrm{d}t$$

$$= (1-p)\left\{p^{r+1}x(p) - \int_0^p (r+1)t^r x(t)\,\mathrm{d}t\right\} + p\left\{-(p^r - p^{r+1})\, x(p) - \int_p^1 (rt^{r-1} - (r+1)t^r)x(t)\,\mathrm{d}t\right\}$$

$$= (1-p)\left\{-\int_0^p (r+1)t^r x(t)\,\mathrm{d}t\right\} + p\left\{\int_p^1 ((r+1)t^r - rt^{r-1})x(t)\,\mathrm{d}t\right\}$$

$$= (1-p)\left\{-\int_0^p (r+1)t^r x(t)\,\mathrm{d}t\right\} + p\left\{(r+1)\beta_r - r\beta_{r-1} - \int_0^p ((r+1)t^r - rt^{r-1})x(t)\,\mathrm{d}t\right\}$$

$$= p\{(r+1)\beta_r - r\beta_{r-1}\} + pr\int_0^p t^{r-1}x(t)\,\mathrm{d}t - (r+1)\int_0^p t^r x(t)\,\mathrm{d}t$$