Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

Nonparametric bootstrap confidence intervals for the Log-Pearson type III distribution

V. Fortin & B. Bobée

INRS-Eau, University of Québec, NSERC - Hydro-Québec Chair on Statistical Hydrology, 2800 rue Einstein, Sainte-Foy, Québec, Canada, G1V 4C7

Abstract

Estimating the frequency of floods is an important problem in hydrology, commonly solved by fitting a probability distribution to observed maximum annual floods. An essential step which must follow the estimation of a quantile is a quantification of its precision. First-order parametric approximations are commonly used to obtain confidence intervals (CIs) for flood flow quantiles. Nonparametric computer-intensive Bootstrap CIs are compared with parametric CIs for simulated samples, drawn from a log-Pearson type III (LP) distribution. Using this methodology, biased in favour of parametric CIs since the parent distribution is known, Bootstrap CIs are shown to be more accurate for small to moderate confidence level (≤80%), when parameters are estimated by the indirect method of moment (WRC). However, the actual level of Bootstrap CIs is almost always lower than the target level. It is expected that, compared to parametric CIs, Bootstrap CIs perform even better when applied to actual series of maximum annual floods, since they need not come from a LP distribution.

1 Introduction

The objective of flood frequency analysis (FFA) is to estimate the flood x_T which is exceeded in average once every T years. At gaged locations, x_T may be estimated by fitting a probability distribution with cumulative distribution function (cdf) $F(x;\underline{\theta})$ to observed maximum annual floods. If the parameters $\underline{\theta}$ are estimated by $\hat{\underline{\theta}}$, x_T can be estimated by: [3]

$$\hat{x}_T = F^{-1}(1 - 1/T;\hat{\underline{\theta}}) \qquad (1)$$

$\hat{x}_T$ being a random variable, it is important to quantify the precision with which it is known, either by estimating its standard error, or preferably by computing a confidence interval (CI) of confidence level $100(1-\alpha)\%$ for x_T, which is a range $[x_l, x_u]$ of values between which x_T has a $(1-\alpha)$ probability of lying. If different CIs may be derived using various methods and if they are all exact, that is they all achieve the desired confidence level, the CI giving the smaller range $[x_l, x_u]$ should be chosen.

The statistical distribution of $\hat{x}_T$ being usually unknown, it is not possible to derive an exact CI for x_T. However, first-order approximations, acceptable for large sample sizes, may be used [1,5]. Since hydrologic samples are typically of small size, these approximate CIs may lack accuracy. Using resampling methods, such as the Bootstrap [7], it is possible to estimate a CI for x_T without making assumptions on the sample size or the statistical distribution of $\hat{x}_T$. However, this nonparametric approach requires a great amount of computation.

We will compare parametric and Bootstrap CIs for the log-Pearson type III (LP) distribution [2], recommended to model flood flows in the United States by the Water Resource Council (WRC) [15]. After summarizing the main properties of this distribution, we will elaborate on the existing approaches used to obtain CIs for its quantiles, discuss the details of the study and present the results.

2 The Log-Pearson type III (LP) Distribution

The log-Pearson type III distribution has a scale parameter, m, and two shape parameters, β ($\neq$0) and λ (>0), making it very flexible. Its probability density function (pdf) is given by: [2]

$$f(x|\lambda,\beta,m) = |\beta| / \{x\Gamma(\lambda)\} \cdot \{\beta[\log(x) - m]\}^{\lambda-1} \exp\{-\beta[\log(x) - m]\} \qquad (2)$$

where log(·) denote Napierian logarithms. Notice that y=log(x) follows a Pearson type III (P) distribution. We shall later use the skewness of y, denoted γ, which can be computed as $\gamma = 2|\beta|/\beta \cdot \lambda^{1/2}$.

2.1 Estimation of parameters

The domain of x varying with m, maximum likelihood estimates are not optimal for small samples [2]. The parameters can be estimated using the method of moments applied to the logarithm of the observations, as recommended by the WRC [15], or using the direct method of moments (MM) [2].

2.2 Estimation of x_T

The cumulative distribution function (cdf), $F(x|\lambda,\beta,m)$, cannot be evaluated directly, making it difficult to estimate x_T through (1). $\hat{x}_T$ can however be computed from quantiles of the standardized P distribution, $K_T(\gamma)$, since: [2]

$$\hat{x}_T = \exp\left\{\hat{m} + \hat{\lambda}/\hat{\beta} + \sqrt{\hat{\lambda}} / |\hat{\beta}| \cdot K_T(\hat{\gamma})\right\} \qquad (3)$$

To approximate $K_T(\gamma)$, we shall use a fourth-order Chebyshev polynomial [12], shown [5] to be adequate in the usual range of γ.

3 Confidence Intervals for the LP Distribution

CIs based on asymptotic theory [1], along with CIs constructed using the noncentral t-distribution [5] are commonly used in practice. Contrasting with these parametric approaches, simple but computer-intensive methods, including the Jackknife [11] and the Bootstrap [7], may also be used. The former being almost a Taylor series approximation to the Bootstrap [7], only the latter will be discussed here. We will describe the different approaches in more details.

3.1 Asymptotic confidence intervals

An asymptotic 100(1-α)% CI for a statistic x given an estimator $\hat{x}$ is $\hat{x} \pm s(\hat{x}) z_{1-\alpha/2}$, where z_η is the η-quantile of the standardized normal distribution and $s(\hat{x})$ is an estimate of the asymptotic standard error of $\hat{x}$ [6]. Bobée [1] developped, for the Pearson type III (P) distribution, an estimate $s(\hat{y}_T)$ for the method of moments estimator $\hat{y}_T$ of a quantile y_T:

$$s^2(\hat{y}_T) = s_y^2/n \cdot \left\{1 + \hat{\gamma} K_T(\hat{\gamma}) + K_T^2(\hat{\gamma})\left(1/2 + 3\hat{\gamma}^2/8\right)\right.$$

$$\left. +3\hat{\gamma} K_T(\hat{\gamma})\left(1 + \hat{\gamma}^2/4\right) \frac{\partial K_T(\gamma)}{\partial \gamma}\bigg|_{\gamma=\hat{\gamma}} + \left(6 + 9\hat{\gamma}^2 + 15\hat{\gamma}^4/8\right)\left(\frac{\partial K_T(\gamma)}{\partial \gamma}\bigg|_{\gamma=\hat{\gamma}}\right)^2 \right\} \quad (4)$$

where s_y is the standard deviation of the Napierian logarithm of the observations, and n is the sample size. Using (4), two different asymptotic CIs may be derived. First, since $x_T = \exp\{y_T\}$, an asymptotic 100(1-α)% CI may be obtained from the asymptotic CI of the P distribution. We shall refer to this CI by the symbol [LN] since it assumes $\hat{x}_T$ to be log-normally distributed:

$$[LN] = \exp\left\{\hat{y}_T \pm s(\hat{y}_T) z_{1-\alpha/2}\right\} = \hat{x}_T \exp\left\{\pm s(\hat{y}_T) z_{1-\alpha/2}\right\} \quad (5)$$

Another asymptotic CI, which we shall name simply [N], may also be obtained by assuming $\hat{x}_T$ to be normally distributed since $s(\hat{x}_T) = \hat{x}_T\, s(\hat{y}_T)$ [2]:

$$[N] = \hat{x}_T \pm s(\hat{x}_T) z_{1-\alpha/2} = \hat{x}_T \left\{1 \pm s(\hat{y}_T) z_{1-\alpha/2}\right\} \quad (6)$$

(6) has the advantage of being systematically smaller than (5), and is therefore preferable for large samples. This can be easily shown from the Taylor series expansion for (5). However, for small samples, [LN] CIs may achieve the desired confidence level more accuratly and thus be preferable to [N].

3.2 A confidence interval based on the noncentral *t*-distribution

Chowdhury and Stedinger [5] propose a first-order correct CI for the LP distribution based on the non-central t-distribution, which we will name [t]

$$[t] = \left[\hat{x}_T \exp\left\{\kappa\left(\zeta_{\alpha/2,p} - z_p\right)\right\}, \hat{x}_T \exp\left\{\kappa\left(\zeta_{1-\alpha/2,p} - z_p\right)\right\}\right], \kappa = s(\hat{y}_T)\sqrt{\frac{2n}{2 + z_p^2}} \quad (7)$$

where $p = 1 - 1/T$ and $\zeta_{\eta,p}$ is the non-central t-distribution, which can be adequately approximated by (8) for $n \geq 15$: [5]

$$\zeta_{\eta,p} \approx \left\{z_p + \frac{z_\eta}{\sqrt{n}}\sqrt{1 + \left(n z_p^2 - z_\eta^2\right)/(2n-2)}\right\} \Big/ \left\{1 - z_\eta^2/(2n-2)\right\} \quad (8)$$

3.3 Bootstrap confidence intervals

The Bootstrap technique allows to derive a CI for x_T when only a random sample $X = \{x_1, x_2, \ldots, x_n\}$ is available, with no *a priori* information on the parent distribution [7]. The technique is simple: a large number, say $b \times n$, of observations are randomly drawn with replacement from X, and grouped in b samples of size n. x_T is then estimated for each sample and the values obtained are ordered, giving a Bootstrap ordered sample $X_b = \left\{\hat{x}_{T(1)} < \hat{x}_{T(2)} < \ldots < \hat{x}_{T(b)}\right\}$, from which may be estimated the distribution of $\hat{x}_T$, $G(x) \approx \Pr[\hat{x}_T \leq x]$.

$G^{-1}(x)$ may be computed using Hazen's [9] plotting position formula, which gives $G^{-1}(\hat{x}_{T(k)})=(k-0.5)/b$. For values of x not included in the sample X_b, linear interpolation may be used, given that $\hat{x}_{T(1)}<x<\hat{x}_{T(b)}$. If $\hat{x}_T$ is not biased, (9) is a straightforward CI for x_T, named the percentile method and noted [B]: [6]

$$[B] = [G^{-1}(\alpha/2), G^{-1}(1-\alpha/2)] \tag{9}$$

Otherwise, a bias-corrected CI, the [BC] method, may be used: [6]

$$[BC] = \left[G^{-1}\left(\Phi(2z' + z_{\alpha/2})\right), G^{-1}\left(\Phi(2z' - z_{\alpha/2})\right)\right] \tag{10}$$

where Φ is the cdf of the standardized normal distribution, and $z'=\Phi^{-1}(G^{-1}(\hat{x}_T))$. Notice that if $G(0.5)=\hat{x}_T$, then $z'=0$ and (10) is equivalent to (9). Another Bootstrap CI for x_T, shown to be second-order correct in a large number problems, is the accelerated [BC], or [BC$_a$]: [7]

$$[BC_a] = \left[G^{-1}\left(\Phi(z[\alpha])\right), G^{-1}\left(\Phi(z[1-\alpha])\right)\right], \quad z[\alpha] = z' + \frac{z' + z_{\alpha/2}}{1 - a(z' + z_{\alpha/2})} \tag{11}$$

where a is the acceleration constant. Notice that if a=0, (11) is equivalent to (10). a may be estimated [6] from the empirical influence function (IF) of $\hat{x}_T$:

$$a \approx \frac{1}{6}\left(\sum_{i=1}^{n} \mathrm{IF}^3(x_i)\right) \Big/ \left(\sum_{i=1}^{n} \mathrm{IF}^2(x_i)\right)^{3/2} \tag{12}$$

Considering $\hat{x}_T$ as a function of the sample X, $\hat{x}_T = h(\{x_1, x_2, ..., x_n\})$, IF($x$) may be estimated [8], for large n, by:

$$\mathrm{IF}(x) \approx (n+1)\left[h\left(\{x_1, x_2, \ldots, x_n, x\}\right) - h\left(\{x_1, x_2, \ldots, x_n\}\right)\right] \tag{13}$$

4 A Comparative Study of Confidence Intervals

To compare the 6 methods presented previously, we measured their effective confidence level, for samples drawn from a LP distribution. 3 sets of parameters were chosen, corresponding to typical values of the coefficients of variation (Cv=σ/μ) and skewness (Cs=μ_3/σ^3) for flood data from the provinces of Ontario and Québec, namely Cv=0.3 with Cs=0.1, 0.5 and 0.9. To simplify the analysis, emphasis was placed on Cs, and only one value of Cv was considered; other values could lead to different conclusions [14]. R=1000 samples of size n=25 and n=50 were simulated for each set of parameters, using Cheng's [4] algorithm coupled with the uniform random generator DRAND [10]. Each sample was then fitted both by the WRC method and the MM. For each CI, the number q of times that x_T fell between the bounds was measured for T=10,50 and 100 years, and for confidence levels 100(1-α)%=50,80 and 99%, allowing to compute the effective confidence level, $\hat{\alpha}=q/$R. As suggested by Efron and Tibshirani [7], the number of Bootstrap samples was set to b=1000. Figures 1-6 show the results of these simulations, displaying the difference between the effective confidence level and the target level ($\Delta\alpha=100[\hat{\alpha}-\alpha]$).

Figures 1-3 reveal that Bootstrap CIs ([B], [BC] and [BC$_a$]) have substantially higher values of $|\Delta\alpha|$, and thus perform poorly when the MM is used. This is especially true for large return periods and high confidence levels. [BC] and [BC$_a$] CIs are not much more accurate than the percentile method [B]. For T=10, Bootstrap CIs are acceptable, but their performance would not justify

Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

multiplying the computing time by 1000. Results differ for the WRC method, as displayed by figures 4-6: for confidence levels of 50% and 80%, Bootstrap CIs are more accurate. Even for a confidence level of 95%, Bootstrap CIs are much more acceptable for the WRC method than for the MM. However, in almost all cases the effective level of Bootstrap CIs is lower than the target level.

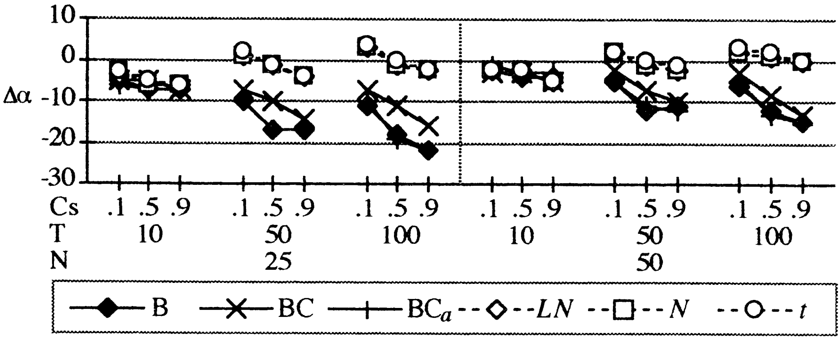

Figure 1. Δα obtained using the MM, for a confidence level of 50%

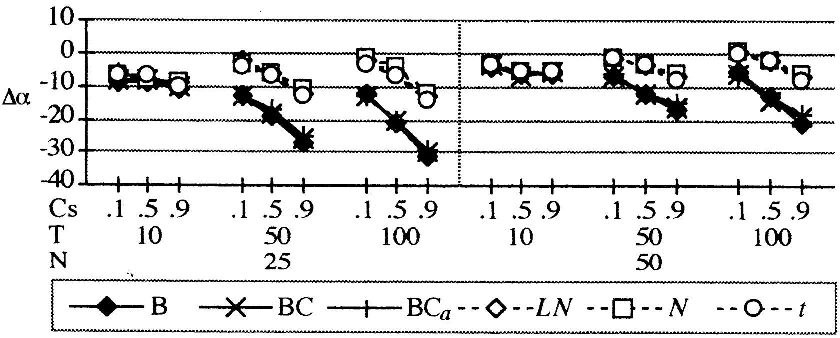

Figure 2. Δα obtained using the MM, for a confidence level of 80%

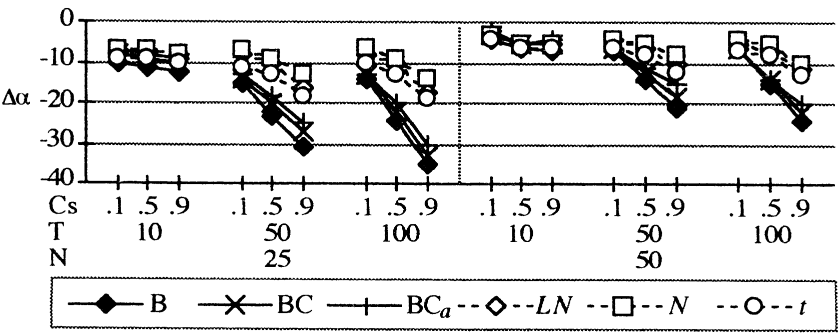

Figure 3. Δα obtained using the MM, for a confidence level of 95%

Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

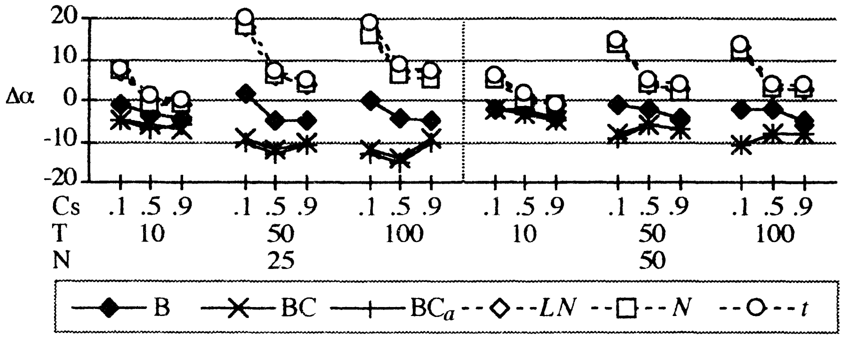

Figure 4. Δα obtained using the WRC method, for a confidence level of 50%

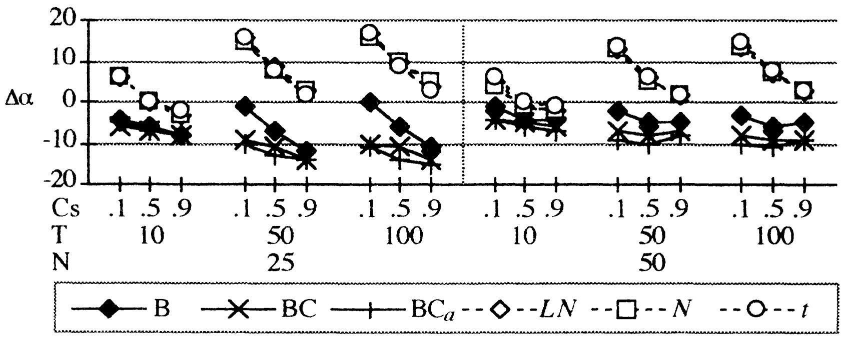

Figure 5. Δα obtained using the WRC method, for a confidence level of 80%

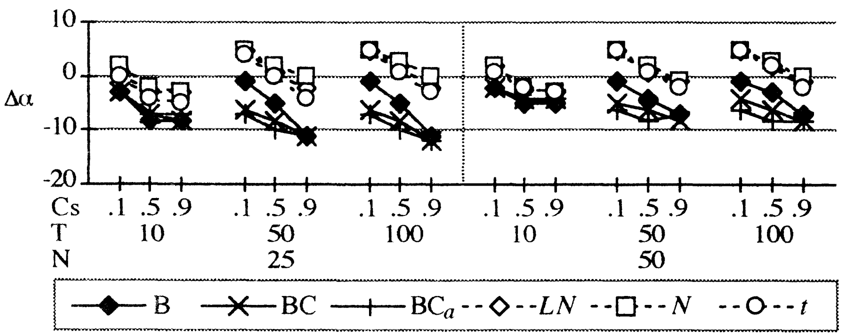

Figure 6. Δα obtained using the WRC method, for a confidence level of 95%

If one CI must be selected for given sample size, confidence level and return period T, results for each parameter sets (or value of Cs) must be combined. We considered selecting the CI for which the maximum absolute error, |Δα|, was the smallest, (minimax criterion), and we also considered

Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

selecting the CI for which the average absolute error was the smallest (lowest-average criterion). In all cases, both criteria lead to the selection of the same CI. Table 1 summarizes the results of these selection procedures.

Table 1. Minimax and lowest-average selection of the best CI

		MM			WRC		
n	T	50%	80%	95%	50%	80%	95%
	10	*t*				LN	LN
25	50	*t*	LN	LN	B	B	*t*
	100	LN				B	*t*
	10	BC_a	*t*				LN
50	50	LN	LN	LN	B	B	N
	100	LN	LN				LN

Table 1 makes it clear that Bootstrap CIs perform poorly for the MM, but much better for the WRC method. This may be explained by examining the bias of both methods, since the percentile method [B] needs unbiased estimators. Figure 7 shows the absolute standardized bias, $\sum_{i=1}^{b}|\hat{x}_{T_{(i)}} - x_T| / (x_T b)$, obtained for MM and WRC. The MM displays much more bias, especially for Cs=0.9 and large return periods. It is therefore not surprising that [B] CIs are more accurate for the WRC method. [BC] and [BC_a] CIs, which should take care of the bias, do not improve on the percentile method. As noted by Schenker [13], bias-corrected Bootstrap CIs are based on a number of assumptions which are in general not verified, and therefore often do not adequately correct the bias.

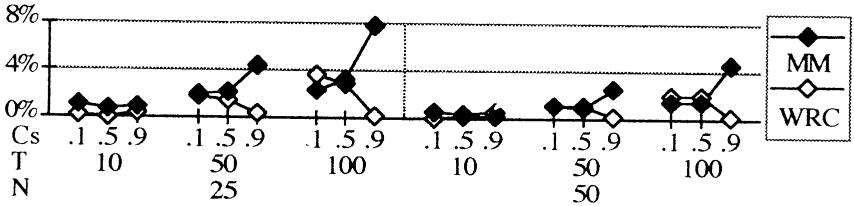

Figure 7. Absolute standardized bias of $\hat{x}_T$

It was expected that the accuracy of Bootstrap CIs would decrease with increasing confidence levels and also that the actual confidence level of Bootstrap CIs would be lower than the target level, since samples simulated using the Bootstrap technique never contain more information about outliers than the observed sample. Nothing being assumed about the tail of the distribution of x, $G(x)$ is a light-tail approximation of the distribution of $\hat{x}_T$: the probability of extreme events is underestimated, leading to smaller CIs.

Conclusion

For small to moderate confidence levels (≤80%), Bootstrap confidence intervals (CIs) for quantiles of the log-Pearson type III (LP) distribution revealed more

CIs when using the indirect method of moments (WRC), but performed poorly for large return periods (T≥50) when using the method of moments. The effective confidence level of Bootstrap CIs was almost always lower than the target level. Corrected Bootstrap CIs (BC and BC_a) did not give better results. Bootstrap CIs should not be used blindly, especially for high confidence levels, but may already be recommended for moderate confidence levels when using the WRC method.

Acknowledgments

We thank Luc Perreault for several useful discussions and suggestions.

References

1. Bobée, B. Sample Error of T-Year Events Computed by Fitting a Pearson Type 3 Distribution, *Water Resources Research*, 1973, **9**(5):1264-1270.
2. Bobée, B. and Ashkar F. *The Gamma Family and Derived Distributions Applied in Hydrology*, Water Resources Publication, Littleton, Colorado, USA, 203 p., 1991.
3. Bobée, B., Rasmussen, P., Perreault, L. and Ashkar, F. Risk Analysis of Hydrologic Data: Review and New Developments Concerning the Halphen Distributions, *NATO Advanced Study Institute on Engineering Risk and Reliability in a Changing Physical Environment*, Deauville, France, 23 May–5 June 1993.
4. Cheng, R.C.H. The Generation of Gamma Variables with Non-integral Shape Parameter, *Applied statistics*, 1977, **26**(1):71-75.
5. Chowdhury, J.U. and Stedinger, J.R. Confidence Interval for Design Floods with Estimated Skew Coefficient, *Journal of Hydraulic Engineering*, 1991, **117**(7):811-831.
6. Efron, B. Better Bootstrap Confidence Intervals, *Journal of the American Statistical Association, Theory and Methods*, 1987, **82**(397):171-185.
7. Efron, B. and Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, *Statistical Science*, 1986, **1**(1):54-77.
8. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel W.A. *Robust Statistics: The Approach Based on Influence Functions*, John Wiley & Sons, New York, USA, 502 p., 1986.
9. Hazen, A. Storage to be provided in impounding reservoirs for municipal water supply *Transactions of ASCE*, Paper 1308, 1914, **77**:1547-1550.
10. Law, A.M. and Kelton, W.D. *Simulation Modeling and Analysis*, McGraw-Hill, New York, USA, 400 p., 1982.
11. Miller, R.G. Jr. The jackknife–a review, *Biometrika*, 1974, **61**(1):1-15.
12. Perreault, L., Bobée, B. et Fortin V. *Approximation des quantiles de la loi Pearson type 3 standardisée par les polynômes de Tchebichef*, Scientific Report 346, INRS-Eau, Sainte-Foy, Québec, Canada, 36 p., 1992.
13. Schenker, N. Qualms About Bootstrap Confidence Intervals *Journal of the American Statistical Association, Theory and Methods*, 1985, **80**(390):360-361.
14. Stedinger, J.R. Confidence Intervals for Design Events, *Journal of Hydraulic Engineering*, 1983, **109**(1):13-27.
15. Water Resource Council, *Guidelines for Determining Flood Flow Frequency*, Bulletin 17B, Hydrology Committee, Washington D.C., 1981.