Journal of Hydrology (2006) 328, 694–703

available at www.sciencedirect.com

journal homepage: www.elsevier.com/locate/jhydrol

Fitting the log-logistic distribution by generalized moments

Fahim Ashkar [a,*], Smail Mahdi [b,1]

[a] Department of Mathematics and Statistics, Université de Moncton, Moncton, NB, Canada E1A 3E9
[b] Department of Computer Science, Mathematics and Physics, University of the West Indies, Cave Hill campus, Barbados

Received 25 May 2005; received in revised form 22 January 2006; accepted 23 January 2006

KEYWORDS
Frequency analysis;
Low stream flow;
Log-logistic distribution;
Generalized moments;
Generalized probability weighted moments;
Maximum likelihood

Summary The method of generalized moments (GM) is investigated for parameter and quantile estimation in the 2-parameter log-logistic (LL2) model. Point estimators for the shape and scale parameters and quantiles are derived. Asymptotic variances and covariances for these estimators are presented, along with simulation results on the performance of the GM method versus the methods of generalized probability weighted moments (GPWM), of maximum likelihood (ML), and of classical moments applied to $Y = \ln X$. The GPWM and ML methods have already been investigated by the authors. Some mathematical properties of the LL2 model and some relationships between GM and GPWM are highlighted. The simulation results show the GM method to outperform the other competitive methods in the LL2 case, when moment orders are appropriately chosen. It is also shown that a mixture of moments of positive and negative orders is needed for optimal estimation under an LL2 model, and how this mixture can be implemented using the GM method. However, further research into the area of optimal choice of moment orders is still needed. Mixing positive and negative moments in the estimation is demonstrated by a hydrological example involving low stream flow.

Introduction

Statistical frequency analysis using probability distributions is widely employed in hydrology for estimating the relation between the magnitude and occurrence frequency of various hydrological events. A procedure commonly used involves: (i) selecting a sample of values of the hydrological variable that satisfies the criteria of randomness, independence, homogeneity and stationarity; (ii) fitting a probability distribution to this sample by an appropriate fitting method, and (iii) using this fitted distribution to make statistical inferences about the underlying population.

Following the publication of the Flood Estimation Handbook (FEH) in the UK (IH, 1999) and the recommendation of the generalized logistic (GL) distribution as the standard

* Corresponding author. Tel.: +1 506 858 4312; fax: +1 506 858 4396.
E-mail addresses: ashkarf@umoncton.ca (F. Ashkar), smahdi@uwichill.edu.bb (S. Mahdi).
[1] Tel.: +1 246 417 4367; fax: +1 246 425 1325.

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2006.01.014

for flood frequency analysis in that country, the use of this distribution has increased in popularity in hydrology. The GL distribution, as used in the FEH, is in fact a generalization of the 3-parameter log-logistic (LL3) distribution, which had earlier been examined by Ahmad et al., 1988) in an at-site and regional study involving Scottish flood data. Ahmad et al., 1988) compared the LL3 distribution to the GEV, the 3-parameter lognormal, and the Pearson type 3 models, and found that the LL3 model performed ''extremely well'' compared to these other models, according to a set of criteria chosen by the authors. For the exact relationship between the GL, LL3 and logistic distributions, the reader may refer to the two studies: Kjeldsen and Jones (2004), and Ahmad et al., 1988). From these two studies, it can be deduced that the LL3 distribution of Ahmad et al., 1988) is a special case of the GL distribution (corresponding to the case $k < 0$, according to the notation by Kjeldsen and Jones (2004)).

On the other hand, Shoukri et al., 1988) showed a good fit of the 2-parameter LL model (LL2) to precipitation data from various Canadian regions. More recently, Ashkar and Mahdi (2003) also compared the LL2 model to the 2-parameter lognormal, the 2-parameter Weibull, and the extreme value type 1 distributions for fitting maximum annual stream flow data. This comparison showed the good fitting potential of the LL2 distribution to a large data set (114 hydrometric series). In a separate study, various 2-parameter distributions were also considered by Ashkar et al., 2004) for fitting low stream-flow data by the deficit-below-threshold (DBT) approach, as will be described in ''Hydrological example'' of the present study. Following that study, in which seven types of distributions were considered, the LL2 distribution was one of the models recommended for fitting low-flow volume, intensity and duration.

Among the methods used to fit statistical distributions to hydrological data, the maximum likelihood (ML) method has long been considered important, due to its asymptotic efficiency. On the other hand, the method of moments (MM) has been popular due to its ease of application, and the method of probability-weighted moments (PWM) (Greenwood et al., 1979; Hosking et al., 1985; Hosking, 1986) has been quite widely applied as an alternative to the MM and ML methods.

In 2001, Rasmussen proposed a ''generalization'' of the PWM method, which he called method of generalized probability weighted moments (GPWM), and applied this method to the 2-parameter Pareto distribution (although according to a referee, this ''generalized'' PWM method was the one originally proposed by Greenwood et al., 1979). Ashkar and Mahdi (2003) developed the GPWM for the LL2 distribution, and compared it to the ML method.

In the present study, we will develop the method of generalized moments (GM) for parameter and quantile estimation under an LL2 model. Point estimators will be derived and asymptotic variances and covariances of these estimators will be presented. Simulations will also be performed to compare the GM, GPWM and ML methods, along with the method of ''log-moments'' (LM), which is the classical method of moments applied to the logistic random variable $Y = \ln X$. Special attention will be paid to the flexibility of moment orders that can (and should) be used to fit statistical distributions to observed data. We will also make some comparisons between the GM and GPWM approaches to parameter estimation and, along the way, point out some interesting results concerning the LL2 distribution.

We will organize the paper as follows. In ''GM versus GPWM'', we will make some comparisons between the GM and GPWM methods for parameter estimation and in ''Generalities concerning LL2 and estimation using GM'' will give some generalities concerning the LL2 distribution and develop the necessary theory for GM estimation under this model. ''Comparison of the GM, GPWM and ML and LM methods'' will present some simulation results that compare the GM, GPWM, ML and LM methods, and ''Hydrological example'' will contain a hydrological example. Finally, ''Conclusion'' will be devoted to some concluding remarks. We will present the asymptotic variances and covariances of parameter and quantile estimators obtained by the GM method under an LL2 model in Appendix.

GM versus GPWM

In the next section, we will mathematically develop the GM method specifically for the LL2 distribution. However, we will start in the present section by making some basic reflections on the GM and GPWM methods, as applied to *any* distribution.

To estimate the parameters of a distribution, the studies by Rasmussen (2001), and by Ashkar and Mahdi (2003), used PWMs of the form

$$M_{l,r} = E[X^l F^r] = \int_{-\infty}^{\infty} x^l F^r(x) f(x) \mathrm{d}x, \tag{1}$$

where X is the hydrological continuous variable whose distribution is being estimated and F is the cumulative distribution function (cdf) of X, (Rasmussen's study also considered PWMs of the form $E[X^l(1-F)^r]$). The PWM and GPWM methods are computationally practical when the inverse of F can be calculated analytically, and the integral in (1) can be evaluated to yield an analytical expression of PWMs, as functions of distribution parameters. However, it is worth noting that, once $M_{l,r}$ in (1) is calculated, it serves as a direct basis for applying not only the GPWM method, but also the GM method, as it will clearly be seen in the following section. The GM method has been earlier used in hydrology, for instance, by Ashkar and Bobée (1987).

The idea behind the PWM and GPWM methods is to obtain parameter estimates by equating PWMs ($M_{l,r}$), to sample PWM estimates $(\hat{M}_{l,r})$, and solving the resulting system of equations for the distribution parameters. In the ''traditional'' PWM method, r values are chosen to be nonnegative integers that are as small as possible, so for a 2-parameter distribution, this method involves consideration of $r = 0$ and $r = 1$ in (1) (with $l = 1$). In the ''GPWM'' method on the other hand (as called by Rasmussen, 2001), r has neither to be small, nor a nonnegative integer.

PWM and GPWM applications have restricted attention to the case where $l = 1$ in Eq. (1), i.e., they have used only PWMs of the form $M_{l,r;l=1}$ in the estimation. The justification often provided for this restriction, is that using moments of X, that are of order greater than 1, ''should be avoided''. On the other hand, in the classical method of (product) moments, as well as in the method of generalized (product) moments (GM), which is the subject of the present paper, moments of order different from 1 are involved. Thus, whereas the GM method considers moments of the form $M_{l,r;r=0} = E[X^l F^0] = E[X^l]$, the GPWM method considers moments of the form $M_{l,r;l=1} = E[XF^r]$. In the GM method, the power l, of X, acts as a weight applied to the sample values, whereas in the GPWM method F^r plays this same role.

Generalities concerning LL2 and estimation using GM

Hereafter, we shall denote the LL2 distribution as ''LL'', for simplicity. In the present section, a few equations that have already been presented in Ashkar and Mahdi (2003) will be repeated, because they serve as a basis for applying both the GPWM and GM methods.

The relationship of LL to the logistic distribution is the same as that of the lognormal (LN2) to the normal distribution. The logistic density curve, like that of the normal, is bell shaped and symmetrical, but has thicker tails, as compared to the normal. Adding a location parameter to LL, to get LL3, is the same as adding a location parameter to LN2 to obtain LN3. Probably, the most popular use of the logistic distribution is in a certain type of regression called *logistic regression*. This type of regression is used in situations where one is concerned with a response variable Y which is dichotomous (i.e., taking one of two values: ''1'', representing ''success'', or ''0'', representing ''failure'') rather than continuous (whereas in ordinary regression, Y is assumed to be continuous). If one would like to estimate the probability p of success, with the help of some explanatory variables, a popular way would be to use logistic regression. We believe that such type of regression (see, e.g., Kutner et al., 2004) needs to be investigated more closely for possible applications in hydrology, because it has already received numerous useful applications in other fields. However, logistic regression models will be outside the scope of the present study.

The cumulative distribution function (cdf) of the LL model is given by

$$F(x) = \frac{(x/\alpha)^\beta}{1+(x/\alpha)^\beta}, \quad x > 0; \ \alpha > 0, \ \beta \geqslant 1, \tag{2}$$

where α is a scale parameter and β a shape parameter. In fact, the shape parameter β in Eq. (2) could be chosen such that $\beta > 0$, but for $0 < \beta < 1$, the mean of the distribution does not exist, so we have chosen to consider only the case $\beta \geqslant 1$, as has been done by Shoukri et al., 1988). The corresponding probability density function (pdf) of the LL distribution is given by

$$f(x) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{[1+(x/\alpha)^\beta]^2}. \tag{3}$$

The quantile x_T, corresponding to return period T, is obtained by solving the equation

$$T = \frac{1}{1-F(x_T)} \tag{4}$$

for x_T, in terms of T. This solution yields:

$$x_T = \alpha(T-1)^{1/\beta}. \tag{5}$$

It is therefore readily noted that the LL model has a cdf and inverse cdf that are quite simple mathematically.

It is also rather easy to calculate $M_{l,r}$ for the LL distribution. The result, from Ashkar and Mahdi (2003), is given by

$$M_{l,r} = \int_0^1 \alpha^l \left(\frac{F}{1-F}\right)^{l/\beta} F^r \,dF = \alpha^l B(r+1+l/\beta, 1-l/\beta), \tag{6}$$

where $B(\cdot,\cdot)$ stands for the Beta function. As noted before, this calculation of $M_{l,r}$ serves as a basis for applying both the GPWM and GM methods. Note also, that for $M_{l,r}$ in (6) to exist, we need to have $r > -1 - l/\beta$, and $l/\beta < 1$.

Ashkar and Mahdi (2003) discussed the choice of r (power of F) in Eq. (6). However, in the GM method, r is taken equal to 0, and l is taken as arbitrary, as previously discussed. Therefore, in order for (6) to be defined, we need to have $-\beta < l < \beta$ in the GM method.

Another interesting property of the LL model is now readily apparent: positive and negative order moments exist for this model in a symmetrical fashion, over the interval $(-\beta, \beta)$. In other words, if the moment of order $+l$ exists, then the moment of order $-l$ also exists. Therefore, one may argue that positive moments need not have preference over negative moments under an LL model.

Now taking $r = 0$ in (6), yields

$$M_{l,0} = \alpha^l B(1+l/\beta, 1-l/\beta) \tag{7}$$

$$= \frac{\alpha^l l\pi}{\beta \sin(\pi l/\beta)}, \quad -\beta < l < \beta, \tag{8}$$

where (8) is obtained from (7) by use of Eq. 6.1.17 of Abramowitz and Stegun (1970).

To apply the GM method, we use Eq. (8) and choose two values of l: $l = l_1$, l_2, not necessarily integers, nor positive, to form two equations from which the estimators of α and β are calculated. Note that in the classical method of moments (MM), we simply take $l_1 = 1$ and $l_2 = 2$. However, in the GM method we allow l_1 and l_2 to be positive or negative, which provides more flexibility in assigning weights to the two tails of the sample distribution.

Taking $l = l_1$, l_2 ($l_1 \neq l_2$), in Eq. (8), we obtain the system of equations:

$$M_{l_1,0} = \frac{\alpha^{l_1}\pi l_1}{\beta \sin(\pi l_1/\beta)} \tag{9}$$

and

$$M_{l_2,0} = \frac{\alpha^{l_2}\pi l_2}{\beta \sin(\pi l_2/\beta)}. \tag{10}$$

Solution of these two equations is equivalent to first solving the following equation numerically for β:

$$M_{l_2,0} = l_2 l_1^{-l_2/l_1} (\pi/\beta)^{(l_1-l_2)/l_1} \times [M_{l_1,0}\sin(\pi l_1/\beta)]^{l_2/l_1} [\sin(\pi l_2/\beta)]^{-1} \tag{11}$$

and then calculating α explicitly from the following equation:

$$\alpha = [\beta(\pi l_1)^{-1} M_{l_1,0} \sin(\pi l_1/\beta)]^{1/l_1}. \tag{12}$$

The quantile estimator $\hat{X}_T$ is then obtained from Eq. (5) by replacing α and β by their respective estimators.

Another moment that is often of interest in estimating the parameters of a statistical distribution is the geometric mean, g, or its natural logarithm, $\ln g$. According to Ashkar and Bobée (1987), $\ln g$ is the moment of order ''quasi-zero'', because it corresponds to the moment $M_{l,0}$, as l tends to zero. This moment is usually denoted by $M_{\underline{0},0}$ ($M_{l,0}$, ''$l = \underline{0}$''). But since the natural logarithm of the geometric mean (g) of a random variable X, is the arithmetic mean of $Y = \ln X$, we may write:

$$l = \underline{0} : M_{\underline{0},0}(X) = \ln g = M_{1,0}(Y); Y = \ln X. \tag{13}$$

In the case of the LL distribution, the moments of order quasi-zero for the sample, and for the population, are respectively given by:

$$\overline{\ln x} = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} \ln(x_i), \tag{14}$$

$$M_{\underline{0},0}(X) = M_{1,0}(Y) = \ln \alpha. \tag{15}$$

The parameters of the LL distribution may be estimated by using the moments of order $\underline{0}$ and 1 of X (geometric mean and arithmetic mean of X), which is equivalent to using the arithmetic mean of X, and the arithmetic mean of $Y = \ln X$. This method will be denoted by $GM(\underline{0}, 1)$. This ''mixing'' of moments of X and $\ln X$ in the estimation, has already been done by Rao (1980), and by Ashkar and Bobée (1987), among others. Application of the method GM($\underline{0}$, 1) to the LL distribution is done according to the following two steps:

1. estimate α explicitly using $\alpha = \exp(\overline{\ln x})$;
2. calculate β numerically by the equation $\bar{x} = (\alpha\pi/\beta)/\sin(\pi/\beta)$ that derives from Eq. (8) for $l = 1$.

Another estimation method that needs to be considered is the LM method, which is the classical method of moments applied to the logistic random variable $Y = \ln X$. In this method, α and β are estimated respectively using $\alpha = \exp(\overline{\ln x})$ and $\beta = \pi/\sqrt{3[\overline{\ln^2 x} - (\overline{\ln x})^2]}$.

In Appendix, we derive the asymptotic covariance matrix of $(\hat{\alpha}, \hat{\beta})$ and the asymptotic variance of $\hat{X}_T$, for the GM method. The corresponding derivations for the GPWM and ML methods were given in Ashkar and Mahdi (2003).

Comparison of the GM, GPWM and ML and LM methods

The results derived in Appendix are useful for studying the variances and covariances among parameter and quantile estimators, when the sample is large. However, as many studies have shown, such asymptotic results are seldom quite accurate with small samples typically encountered in hydrology. With such small samples, numerical studies need to be performed, which can nonetheless be also designed to include large samples.

A simulation study was carried out, with several sample sizes n, shape parameters β of the LL population, and return periods T. With no loss of generality, the scale parameter α was set equal to 1. The values n = 10, 15, 30, 60, 120, 240, 480, 960 were chosen, along with the values: T = 10, 50, 100, 200 and β = 1.5, 3, 6. These are the same simulation parameters chosen by Ashkar and Mahdi (2003) to compare the GPWM and ML methods. The chosen β values account for different shapes of the LL distribution, as shown in Fig. 1. We shall only report results for $n \leqslant 120$, since larger samples are seldom encountered in hydrology.

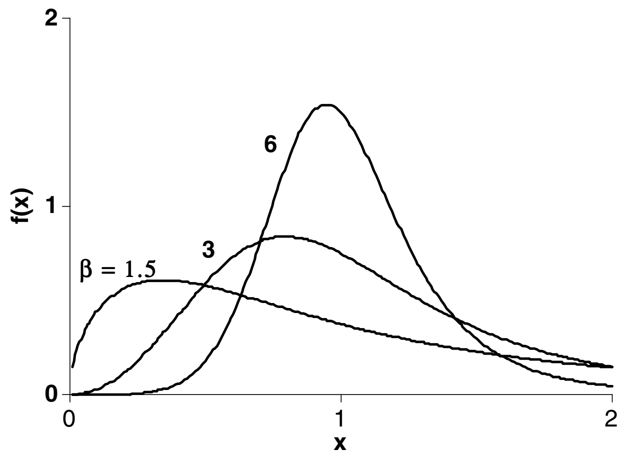

Figure 1 Log-logistic probability density function for shape parameter β = 1.5, 3 and 6.

The simulation was designed to consider moment orders l_1 and l_2 between -0.99 and 0.99, in steps of 0.05. However, the study readily showed that it was sufficient to consider the closed interval $[-0.95, 0.95]$. No orders outside the open interval $(-1, 1)$ were considered because from Eq. (8), the value of l should be within the interval $(-\beta, \beta)$, and since $\beta \geqslant 1$ (Eq. (2)), the interval $(-1, 1)$ is the only one suiting all allowable β values. One objective was to identify moment orders (l_1, l_2) in the range $(-1, +1)$ that provide optimal parameter and quantile estimates by the GM method.

For each combination of n, β, and order pair (l_1, l_2), 1000 LL samples were randomly generated, but for smaller sample sizes (n = 10, 15) 5000 samples were generated. The root mean square errors (RMSEs) for the α, β and quantile estimates were then calculated. Note that the conditions given in equations (A.4) through (A.6) of Appendix would restrict the values of l_1 and l_2 to the global interval $(-0.5, 0.5)$. However, it is always possible to compute empirical RMSE, provided only the existence of parameter or quantile estimates. The bisection method, (Zachary, 1998) which is computationally very simple, was successful for calculating GM estimates.

Focus was on estimating quantiles and the shape parameter β of the LL distribution, since they constitute the main interest in hydrology. Optimum RMSE results with respect to order pairs (l_1, l_2) using the GM method were explored, and attention was given to the methods GM(1, 2) (classical method of moments), GM($\underline{0}$, 1), and LM, as described earlier. The simulations led to the following main conclusions:

1. For estimating β, Ashkar and Mahdi (2003) had found that the GPWM method almost always provided better results than the ML method, when optimum powers of F were used in the GPWM method. However, the difference between the two methods became negligible as the sample size increased. Hereafter, when optimum order pairs (l_1, l_2) are used with the GM method, the resulting

method will be denoted by ''GM-opt''. Similarly, when optimum powers of F are used with the GPWM method, this will be denoted by ''GPWM-opt''. The simulations of the present study showed that the GPWM-opt method outperforms the GM-opt method and the ML method (in that order), only when the sample size is relatively small ($n < 50$). For n greater than about 50, the GM-opt, GPWM-opt and ML methods show similar performance (with a slight out performance by the GM-opt method). These results can be seen from Fig. 2(a)–(c), which presents relative RMSE (i.e., RMSE(β)/β). This figure also shows that the GM-opt, GPWM-opt, and ML methods, outperform the LM, GM($\underline{0}$, 1) and GM(1, 2) methods. Note also that GM(1, 2) is not represented in Fig. 2(a), because for $\beta = 1.5$, the GM(1, 2) method is not applicable [moment of order $l = 2$ does not exist, by Eq. (8)]. From Fig. 2(a)–(c) it can also be seen that relative RMSE for β decreases as β increases.

2. For estimating quantiles X_T, Ashkar and Mahdi (2003) had found that the GPWM-opt method provided better results than the ML method only for very small sample sizes, and small β, i.e., when the distribution was quite asymmetrical. The simulations of the present study have shown that the GM-opt method provides better results than both the GPWM-opt and ML methods. This conclusion can be seen from Fig. 3(a)–(c), which displays RMSE(X_T)/X_T for the case $T = 100$. For other T values, similar results were obtained, but are not reported. Fig. 3(a)–(c) also shows that the GM-opt, GPWM-opt, and ML methods, generally outperform the LM, GM($\underline{0}$, 1) and GM(1, 2) methods (note again that GM(1, 2) is not represented in Fig. 3(a), for the same reason specified earlier). From Fig. 3(a)–(c) it is also seen that relative RMSE for X_T decreases as β increases.

Optimal order pairs (l_1, l_2) for GM are reported in Table 1 for estimating β, and in Table 2 for estimating X_T, $T = 100$. The GM-opt method seems to perform well for estimating β (Fig. 2(a)–(c)), especially when the sample size is quite large ($n > 50$). It is also interesting to note that this method performs well for estimating quantiles X_T, for all sample sizes, all β values, and all T values considered. This can be seen from Fig. 3(a)–(c), for $T = 100$. From Tables 1 and 2, it is also readily noted that a mixture of *both positive and negative* orders (l_1, l_2) are needed for an optimal estimation of both β and X_T, by the GM method.

We have also attempted, as was done by Rasmussen (2001), to determine some kind of analytical rule for finding the optimal orders (l_1, l_2), to use with the GM method. Unfortunately, we did not succeed in this, since the optimal order values were found to change rapidly as the sample size and/or the β values were changed. Further research focusing into this specific area of optimal choice of orders (l_1, l_2) is still needed.

Hydrological example

In managing water resources, scientists and engineers often have to handle conflicts between water demand and water availability during periods of low stream flow. Hydrometric data serve to estimate frequencies of various characteristics of low flows in rivers. A method used to estimate low-flow frequency, bases the analysis on extreme low flows, by considering only the most severe event within a time interval, such as the year. However, in order to better characterize low flows in terms of their duration (D), volume (V), and/or intensity (I), the deficit-below-threshold (DBT) approach seems to be more appropriate (e.g., Sen, 1980; Zelenhasic and Salvai, 1987; Ashkar et al., 2004). This method analyses all flows below a chosen threshold; and 2-parameter distributions such as LL, are the ones typically needed in applying this method.

Ashkar et al., 2004) used statistical distributions to study low stream flows in Atlantic Canada by the DBT approach. In this region of Canada, the occurrence of low stream flow has important environmental and economic implications, particularly in relation to the protection of aquatic habitat (Caissie and El-Jabi, 1995). Thirty-one hydrometric stations were analyzed by the DBT approach, and the results have been reported by Ashkar et al., 2004). The threshold value chosen for the analysis was the median flow for the month of August. Various 2-parameter distributions were considered for fitting low-flow data. Following that analysis, in which a total of seven distributions were considered, the LL distribution was one of the distributions recommended for fitting low-flow volume (V), intensity (I) and duration (D).

The low-flow variable that we will consider here, as an example, is the volume (V) of low-flow events, expressed in m^3/s days, where V was defined as the cumulative stream flow deficit for the duration of the low-flow event.

We will consider hydrometric Station 01AE001 on the Fish River, near Fort Kent, New Brunswick, Canada. This station drains an area of 2260 km^2. During the period 1981–1999, there were $n = 42$ low-flow events below the threshold discharge of 13.237 m^3/s, which represents the median flow for the month of August. Some sample statistics are: sample size = 42; mean = 166.3 m^3/s days; median = 123.5 m^3/s days; standard deviation = 181.5 m^3/s days.

Application of the ML method to the data gave the parameter estimates $\hat{\alpha} = 100.36$, and $\hat{\beta} = 1.45$. Fig. 4 shows the LL fit to the data by the ML method (along with some other methods). Different powers (r_1, r_2) of F were also used with the GPWM method in order to identify pairs (r_1, r_2) that provide adequate fit to the data. Ashkar and Mahdi (2003) observed that negative powers of F are often needed to reduce quantile estimation error by the GPWM method. Fitting the data using the order pairs $(r_1, r_2) = (0.5, 1.5)$; $(0, 1)$; $(-0.5, 0.5)$; $(-1, 0)$; and $(-1.5, -0.5)$, for example, showed that the last two order pairs provided the best fit to the data, especially in the distribution tails. In fact, the order pair $(-1.5, -0.5)$ gave the best fit in the distribution tails, but the pair $(-1, 0)$ provided a slightly better fit in the central part of the data. As expected, different order pairs gave significantly different fits to the data, because each method assigns different weights to the two tails of the sample distribution. Fig. 4 presents some fits to the data by the GPWM method, and Table 3 displays some of the $\hat{\alpha}$ and $\hat{\beta}$ values obtained.

It is observed from Table 3 that the majority of $\hat{\beta}$ values obtained by the various methods are smaller than 2, which could be taken as an indication that the β value of the hypothesized LL population is smaller than 2. From Eq. (8), this implies that the second moment of the population

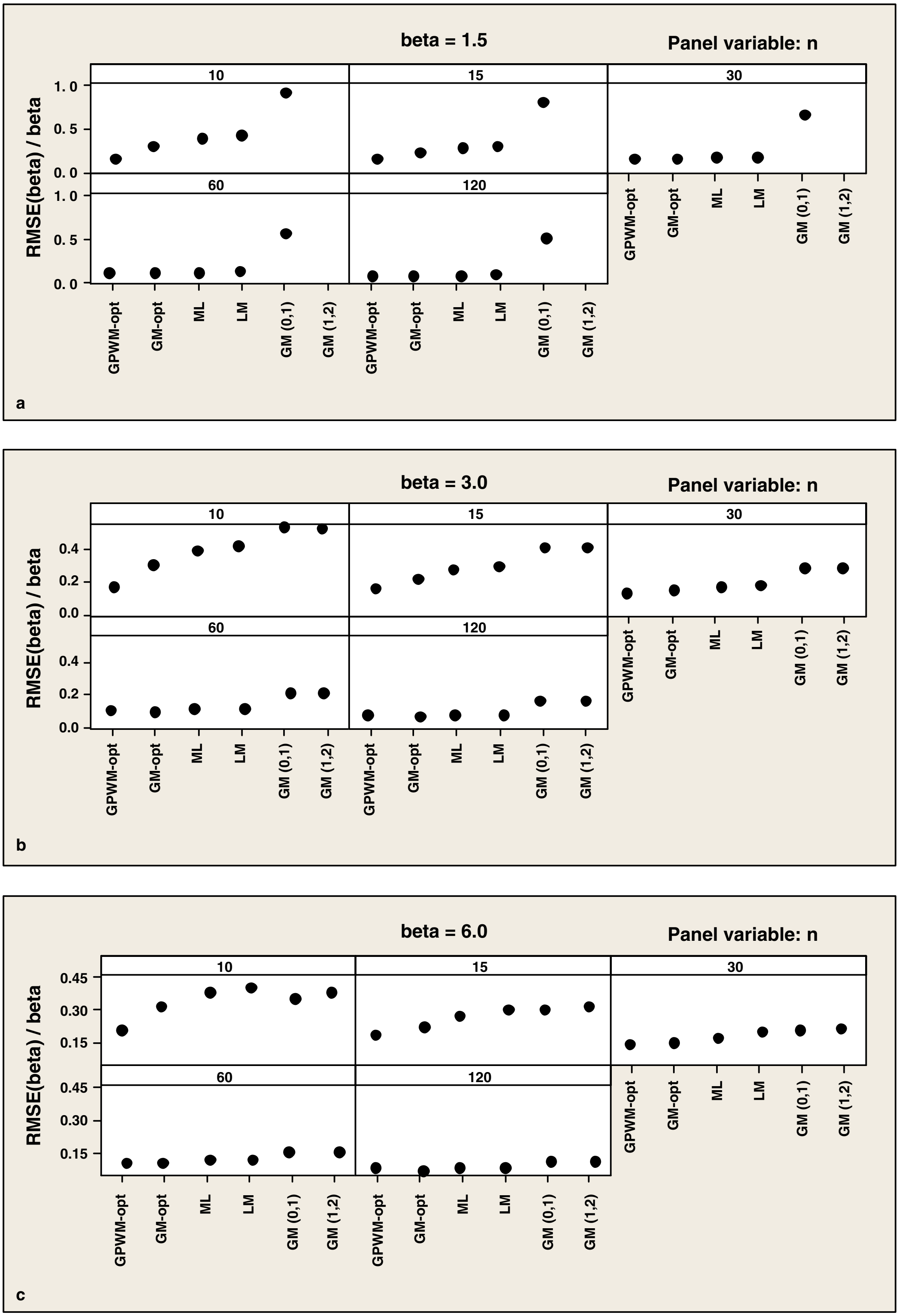

Figure 2 (a)—(c) Relative RMSE for estimates of β. (a) $\beta = 1.5$; (b) $\beta = 3$; (c) $\beta = 6$.

does not exist, which in turn implies that estimation by the classical MM method [i.e. GM(1,2)], cannot be used. However, other versions of the GM method are applicable, including ones that mix moments of positive and negative orders, as recommended in the previous section, as well as the methods GM($\underline{0}$,1) and LM. A visual inspection of a

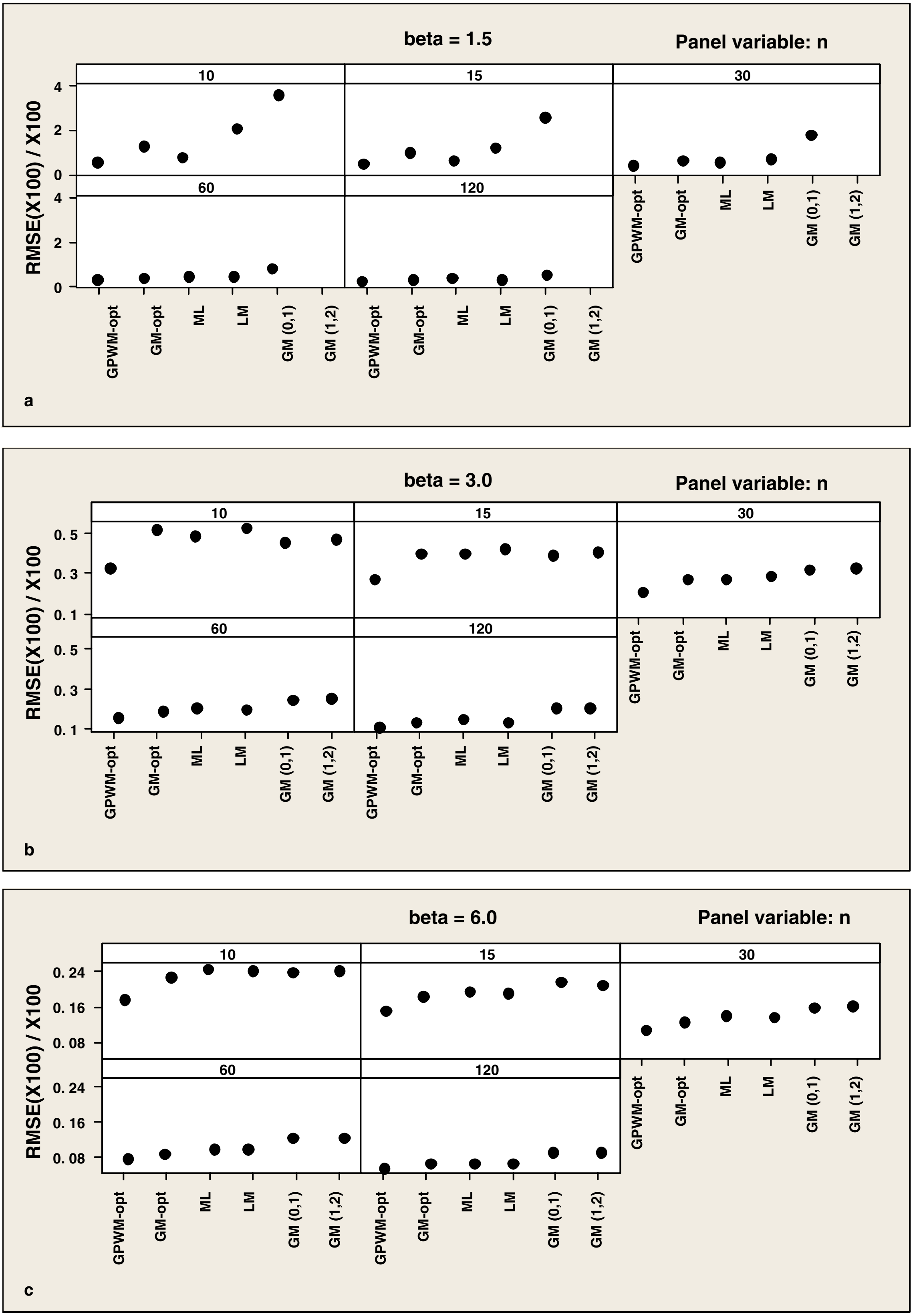

Figure 3 (a)—(c) Relative RMSE for estimates of X_T, $T = 100$. (a) $\beta = 1.5$; (b) $\beta = 3$; (c) $\beta = 6$.

set of order pairs (l_1, l_2) showed that pairs such as $(-0.5, 0.5)$ and $(-0.65, 0.15)$ provided a substantially better fit to the data than pairs such as $(-1, 0.5)$ or $(-1, 1)$ or $(\underline{0}, 1)$. As with the GPWM method, use of different order pairs (l_1, l_2) in the GM method gave significantly different fits to the data because each method assigns different weights to

Table 1 Orders l_1 and l_2 that provide the best performance of the GM estimator for β in the case n = 10, 15, 30, 60, 120; β = 1.5, 3, 6

n	β	l_1	l_2
10	1.5	0.65	−0.40
	3	0.65	−0.40
	6	0.15	−0.90
15	1.5	0.35	−0.20
	3	0.45	−0.50
	6	0.05	−0.90
30	1.5	0.35	−0.20
	3	0.55	−0.55
	6	0.35	−0.30
60	1.5	0.20	−0.20
	3	0.45	−0.40
	6	0.80	−0.75
120	1.5	0.10	−0.25
	3	0.35	−0.45
	6	0.05	−0.90

Table 2 Orders l_1 and l_2 that provide the best performance of the GM estimator for X_T, T = 100, in the case n = 10, 15, 30, 60, 120; β = 1.5, 3, 6

n	β	l_1	l_2
10	1.5	0.25	−0.65
	3	0.35	−0.95
	6	0.50	−0.70
15	1.5	0.30	−0.80
	3	0.60	−0.60
	6	0.10	−0.25
30	1.5	0.05	−0.90
	3	0.90	−0.95
	6	0.35	−0.40
60	1.5	0.15	−0.65
	3	0.15	−0.55
	6	0.75	−0.45
120	1.5	0.40	−0.25
	3	0.20	−0.35
	6	0.65	−0.70

the distribution tails. Fig. 4 shows the fit to the data by the GM(−0.5, 0.5) method and Table 3 displays some $\hat{\alpha}$ and $\hat{\beta}$ values obtained using GM.

Suppose that the user decides to estimate the event with return period T = 100, by the GM(−0.5, 0.5) method. The values $\hat{\alpha} = 90.564$, $\hat{\beta} = 1.56$ (Table 3), and T = 100, are substituted into Eq. (5), to give the estimate $\hat{x}_T =$

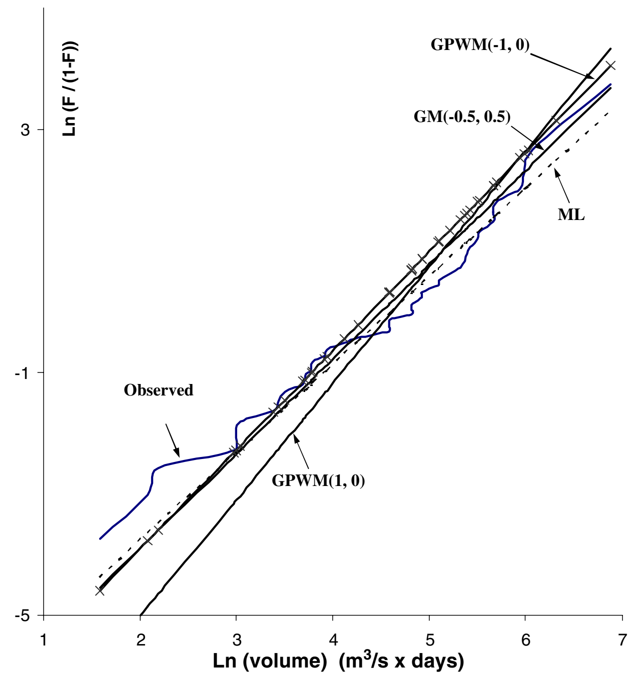

Figure 4 Log-logistic fit to low flow volume data by different fitting methods.

Table 3 Some estimates ($\hat{\alpha}$, $\hat{\beta}$) obtained

Method	$\hat{\alpha}$	$\hat{\beta}$
ML	100.4	1.449
GPWM(−1.5, −0.5)	60.81	1.386
GPWM(−1, 0)	81.13	1.633
GPWM(−0.5, 0.5)	92.25	1.789
GPWM(0, 1)	101.3	1.918
GPWM(0.5, 1.5)	109.2	2.03
LM	93.47	1.522
GM(0, 1)	93.47	1.795
GM(−1, 1)	84.15	1.668
GM(−1, 0.5)	91.84	1.59
GM(−0.5, 0.5)	90.564	1.56
GM(−0.65, 0.15)	93.01	1.512

1722.7 m^3/s days. Eq. (A.1) of Appendix is then applied to give $\mathrm{Var}(\hat{\alpha}) = 338.2$; $\mathrm{Var}(\hat{\beta}) = 0.0604$; and $\mathrm{Cov}(\hat{\alpha},\ \hat{\beta}) \approx 0$. Substituting these into Eq. (A.9) of Appendix, gives a standard error of $\hat{x}_T$ equal to 872 m^3/s days. It should be stressed, however, that this standard error is only approximate, because it is based on asymptotic theory. Research is still needed for the development of small-sample methods that are capable of better quantifying the error of estimation of distribution quantiles, by the GM and other estimation methods.

Conclusion

After analytically deriving GM parameter and quantile estimators for the LL distribution, a numerical analysis of the performance of this method was carried out. A comparison with competitive estimation procedures, namely, the GPWM, ML and LM methods, was done. Relative root mean square error was used as performance index. The study revealed that when optimal order pairs (l_1, l_2) are used, the GM method provides competitive results for estimating the shape parameter β, especially when the sample size is large enough, say, $n > 50$. Moreover, GM results for estimating quantiles, were competitive throughout. It was also noted that a mixture of positive and negative orders (l_1, l_2) were needed for an optimal estimation of β, and of quantiles.

The hydrological example showed that the fit of a model such as LL by GM, or by GPWM, depends significantly on the choice of moment orders (or powers of F) that are used. Use of moments of negative order in the estimation can improve the estimation significantly. However, a concrete analytical rule for finding optimal orders (l_1, l_2) is not yet available, which points to a challenging area of research that deserves further attention.

In certain hydrological applications, 2-parameter distributions are the most appropriate in the modeling, as discussed in the applications section of this study. In other types of applications, where more parameters are needed in the modeling, the 3-parameter log-logistic (LL3) distribution (the relationship between LL3 and LL has been described earlier), or the generalized logistic (GL) distribution (see, e.g., Hosking, 1994), may be useful candidates. The GL distribution (3 parameter) is a special case of the 4-parameter Kappa distribution (Hosking, 1994), and the logistic distribution (2 parameter) is a special case of GL.

For certain probability distributions, especially those that take both positive and negative values, negative-order moments do not exist, and therefore they cannot be used in GM-based estimation. One interesting feature of the LL distribution is that positive and negative order moments exist *in a symmetrical fashion*, over the interval $(-\beta, \beta)$, as we pointed out earlier.

It is finally worth mentioning that the simple bisection method was successful in numerically obtaining parameter estimates for the LL distribution by the GM method. Our numerical studies were carried out with Mathematica, Gauss (Unix Version 3.2.42) and SPSS Release 11. Figs. 2(a)–(c) and Fig. 3(a)–(c) were produced using Minitab.

Acknowledgement

The financial support of UWI and of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. The simulation studies and other computer work were carried out with UWI and UM computer facilities. The authors gratefully acknowledge the comments of the two reviewers, and Miss Lina Ashkar for her assistance.

Appendix

We derive hereafter the asymptotic covariance matrix of $(\hat{\alpha}, \hat{\beta})$, and the asymptotic variance of $\hat{X}_T$.

The asymptotic variances and covariance of $\hat{\alpha}$ and $\hat{\beta}$ are calculated from the variances and covariance of the sample moments $\hat{M}_{l_1,0}$ and $\hat{M}_{l_2,0}$ as follows:

$$\begin{bmatrix} \mathrm{Var}(\hat{\alpha}) \\ \mathrm{Var}(\hat{\beta}) \\ \mathrm{Cov}(\hat{\alpha}, \hat{\beta}) \end{bmatrix} = \begin{bmatrix} M_{11}^2 & M_{12}^2 & 2M_{11}M_{12} \\ M_{21}^2 & M_{22}^2 & 2M_{21}M_{22} \\ M_{11}M_{21} & M_{12}M_{22} & M_{11}M_{22} + M_{21}M_{12} \end{bmatrix}^{-1} \times \begin{bmatrix} \mathrm{Var}(\hat{M}_{l_1,0}) \\ \mathrm{Var}(\hat{M}_{l_2,0}) \\ \mathrm{Cov}(\hat{M}_{l_1,0}, \hat{M}_{l_2,0}) \end{bmatrix}, \tag{A.1}$$

where:

$$M_{11} = \frac{\partial M_{l_1,0}}{\partial \alpha} = \frac{l_1^2 \alpha^{l_1-1}\pi}{\beta}[\sin(\pi l_1/\beta)]^{-1}, \tag{A.2}$$

$$M_{12} = \frac{\partial M_{l_1,0}}{\partial \beta} = \frac{\alpha^{l_1}\pi l_1}{\beta^2}\left\{\frac{\pi l_1}{\beta}[\sin(\pi l_1/\beta)]^{-2}\cos(\pi l_1/\beta) - [\sin(\pi l_1/\beta)]^{-1}\right\}. \tag{A.3}$$

M_{21} and M_{22} are, respectively, obtained from M_{11} and M_{12} by simply replacing l_1 by l_2.

We also have:

$$\mathrm{Var}(\hat{M}_{l_{1,0}}) = \frac{1}{n}[M_{2l_1,0} - M_{l_1,0}^2] = \frac{\alpha^{2l_1}\pi l_1}{n\beta}\left\{2[\sin(2l_1\pi/\beta)]^{-1} - \frac{\pi l_1}{\beta}[\sin(\pi l_1/\beta)]^{-2}\right\}, \quad -\beta/2 < l_1 < \beta/2, \tag{A.4}$$

$$\mathrm{Var}(\hat{M}_{l_{2,0}}) = \frac{1}{n}[M_{2l_2,0} - M_{l_2,0}^2], \quad -\beta/2 < l_2 < \beta/2, \tag{A.5}$$

where (A.5) is obtained from (A.4) by simply replacing l_1 by l_2;

$$\mathrm{Cov}(\hat{M}_{l_1,0}, \hat{M}_{l_2,0}) = \frac{1}{n}[M_{l_1+l_2,0} - M_{l_1,0}M_{l_2,0}] = \frac{\alpha^{l_1+l_2}\pi}{n\beta}\left\{(l_1+l_2)\left[\sin\frac{\pi(l_1+l_2)}{\beta}\right]^{-1} - \frac{\pi l_1 l_2}{\beta}\left[\sin\frac{\pi l_1}{\beta}\right]^{-1}\left[\sin\frac{\pi l_2}{\beta}\right]^{-1}\right\}, \quad -\beta < l_1 + l_2 < \beta. \tag{A.6}$$

By substituting (A.2) through (A.6), into (A.1), the desired variances and covariance of $(\hat{\alpha}, \hat{\beta})$ are obtained.

In the case of estimation by the GM($\underline{0}$, 1) method, Eqs. (A.1)–(A.3) remain the same, but with $l_1 = 1$ and $l_2 = \underline{0}$. However, we now have: $M_{21} = 1/\alpha$ and $M_{22} = 0$. Eq. (A.4) remains also the same, but with $l_1 = 1$. Eq. (A.5) becomes:

$$\mathrm{Var}(\hat{M}_{0,0}) = \pi^2/3n\beta^2. \tag{A.7}$$

In addition, Eq. (A.6) becomes:

$$\mathrm{Cov}(\hat{M}_{0,0}, \hat{M}_{1,0}) = \frac{\pi\alpha[1 - \beta^{-1} ctg(\pi/\beta)]}{n\beta\sin(\pi/\beta)}. \tag{A.8}$$

The asymptotic variance of the quantile estimator $\hat{X}_T$ is calculated by the equation

$$\mathrm{Var}(\hat{X}_T) = \left(\frac{\partial X_T}{\partial \alpha}\right)^2 \mathrm{Var}(\hat{\alpha}) + \left(\frac{\partial X_T}{\partial \beta}\right)^2 \mathrm{Var}(\hat{\beta}) + 2\left(\frac{\partial X_T}{\partial \alpha}\right)\left(\frac{\partial X_T}{\partial \beta}\right)\mathrm{Cov}(\hat{\alpha}, \hat{\beta}), \tag{A.9}$$

where

$$\frac{\partial X_T}{\partial \alpha} = (T-1)^{1/\beta}, \tag{A.10}$$

$$\frac{\partial X_T}{\partial \beta} = -\frac{\alpha}{\beta^2}(T-1)^{1/\beta}\ln(T-1). \tag{A.11}$$

References

Abramowitz, M., Stegun, I.A. (Eds.), 1970. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications, New York.

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

Ashkar, F., Bobée, B., 1987. The generalized method of moments as applied to problems of flood frequency analysis: some practical results for the log-Pearson type 3 distribution. J. Hydrol. 90, 199–217.

Ashkar, F., Mahdi, S., 2003. Comparison of two fitting methods for the log-logistic distribution. Water Resour. Res. 39 (8). Art. No. 1217.

Ashkar, F., Bayentin, J., Savoie, N., El-Jabi, N., 2004. Frequency analysis of low stream-flow characteristics using statistical distributions. In: Proceedings of the World Water and Environmental Resources Congress, June 27–July 1, Salt Lake City, Utah, USA.

Caissie, D., El-Jabi, N., 1995. Comparison and regionalization of hydrologically based instream flow techniques in Atlantic Canada. Can. J. Civil Eng. 22, 235–246.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C., Wallis, J.R., 1979. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. Water Resour. Res. 15 (5), 1049–1054.

Hosking, J.R.M., 1986. The Theory of Probability Weighted Moments. Research Report RC12210, IBM Thomas J. Watson Research Center, New York.

Hosking, J.R.M., 1994. The four-parameter kappa distribution. IBM J. Res. Dev. 38 (3).

Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. Estimation of the generalized extreme-value distribution by the method of probability weighted moments. Technometrics 27, 251–261.

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

Kjeldsen, T.R., Jones, D.A., 2004. Sampling variance of flood quantiles from the generalized logistic distribution estimated using the method of L-moments. Hydrol. Earth Syst. Sci. 8 (2), 183–190.

Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., 2004. Applied Linear Statistical Models, fifth ed. McGraw-Hill/Irwin.

Rao, D.V., 1980. Log Pearson type 3 distribution: method of mixed moments. J. Hydr. Div. ASCE 106 (6), 999–1019.

Rasmussen, P., 2001. Generalized probability weighted moments: application to the generalized Pareto distribution. Water Resour. Res. 37 (6), 1745–1751.

Sen, Z., 1980. Statistical analysis of hydrologic critical droughts. J. Hydr. Div. ASCE 106 (HY1), 99–115.

Shoukri, M.M., Mian, I.U.H., Tracy, D.S., 1988. Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data. Can. J. Stat. 16 (3), 223–236.

Zachary, J.L., 1998. Introduction to Scientific Programming Computational Problem Solving Using Mathematica and C. TELOS/Springer.

Zelenhasic, E., Salvai, A., 1987. A method of streamflow drought analysis. Water Resour. Res. 23, 156–168.