Stochastic Hydrol. Hydraul. 2 (1988) 161–174

Stochastic Hydrology and Hydraulics

The generalized method of moments as applied to the generalized gamma distribution

F. Ashkar, B. Bobée, D. Leroux and **D. Morisette**

Institut National de la Researche Scientifique INRS-Eau, C.P. 7500, Sainte-Foy, Quebec G1V 4C7, Canada

Abstract: The generalized gamma (GG) distribution has a density function that can take on many possible forms commonly encountered in hydrologic applications. This fact has led many authors to study the properties of the distribution and to propose various estimation techniques (method of moments, mixed moments, maximum likelihood etc.). We discuss some of the most important properties of this flexible distribution and present a flexible method of parameter estimation, called the "generalized method of moments" (GMM) which combines any three moments of the GG distribution. The main advantage of this general method is that it has many of the previously proposed methods of estimation as special cases. We also give a general formula for the variance of the T-year event X_T obtained by the GMM along with a general formula for the variance of the parameter estimates and also for the covariances and correlation coefficients between any pair of such estimates. By applying the GMM and carefully choosing the order of the moments that are used in the estimation one can significantly reduce the variance of T-year events for the range of return periods that are of interest.

Key words: Floods, estimation, quantiles, generalized gamma, generalized moments, standard error.

1 Introduction

Flood frequency analysis is an important element in the design of hydraulic structures and in the planning and management of water resources projects. It is often concerned with the search of a statistical distribution which could be used to fit adequately a given sample of maximum annual flood flows. From the chosen distribution, one can make extrapolations in order to estimate an extreme flood event corresponding to a high return period i.e., to a low probability of exceedance (100-year flood, for example). This estimation can then be used to design flood control structures that are able to withstand this estimated extreme flood.

Associated with the choice of a statistical distribution is a certain amount of uncertainty called "model uncertainty". Consequently, one would want to choose a model for which this amount of uncertainty would be minimum.

The question of which distribution should be used to reach this objective has been discussed by several authors. On that issue, the U.S. Water Resources Council (Benson 1968) recommended the use of the log Pearson type III (LP) distribution for representing flood flows. Following that recommendation, several methods have been proposed for fitting this distribution.

Taking a close look at the LP distribution, one realizes that if the skew

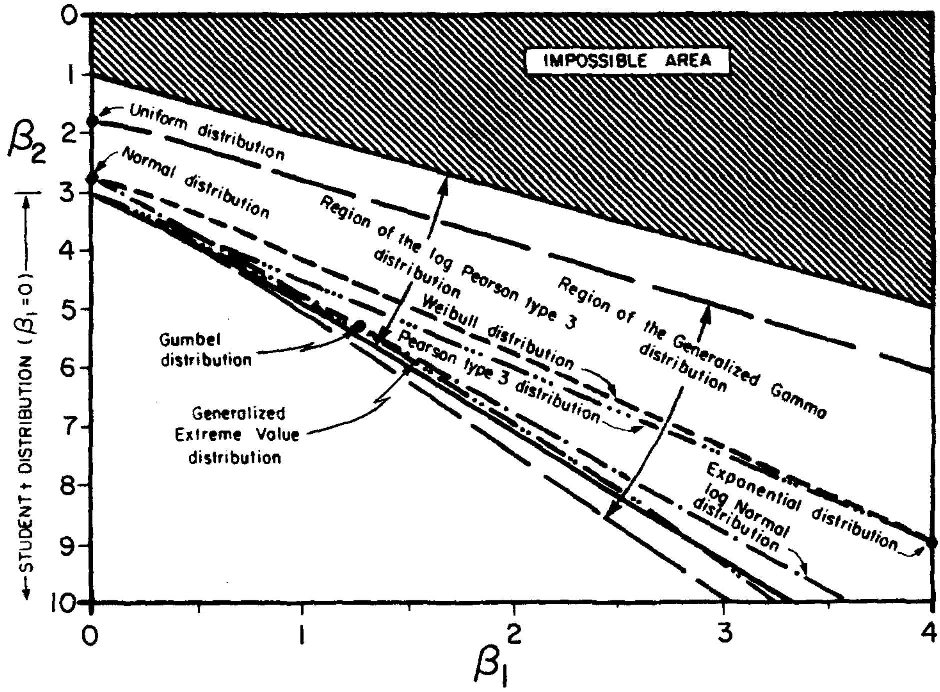

Figure 1. Regions in (β_1, β_2) plane for various distributions, where $\beta_1 = C_s^2$ and $\beta_2 = C_k$

coefficient in log space is negative, then the distribution has a finite upper bound. Some investigators, with or without valid justification, have considered this property to be unsatisfactory, and have looked for other distributions for fitting flood data. One of these, is the generalized gamma (GG) distribution which differs from the LP in that: (1) it is always unbounded above, and (2) it is always, bell-shaped (with a skew) which is one other property that is useful in practice.

The 3-parameter GG distribution which will be the subject of our investigation has two shape parameters and one scale parameter, but no location parameter. It has a fixed lower bound equal to zero (unlike other three-parameter distributions such as the Pearson type III, log Pearson type III and Weibull) but its lower tail can be flexible enough to be able to give a good fit to the lower values of flood flows even if the lower bound of these flows is different from zero. The four-parameter GG distribution can be obtained from its three-parameter counterpart by adding a location parameter m, i.e., if $X \sim GG(3)$ then $(X+m) \sim GG(4)$. In what follows, the abbreviation GG will be used to designate the generalized gamma distribution GG(3).

The GG distribution is very flexible in shape and can take on many possible forms commonly encountered in hydrologic applications. This fact can be verified by taking a look at the representation of this distribution in a moment ratio diagram (MRD) (Leroux et al. 1987) which shows the relationships between the coefficient of skewness (C_s) and the coefficient of kurtosis (C_k) of the distribution. In such a diagram (Fig. 1), GG occupies one of the largest regions, thus showing its large shape flexibility compared to other commonly used distributions. It is also interesting to note that the GG(4) distribution has many useful distribution as special cases like the Pearson type III (3-parameter), Weibull (3- and 2-parameter), gamma and exponential (2- and 1-parameter) as shown in Fig. 2. Obviously it also

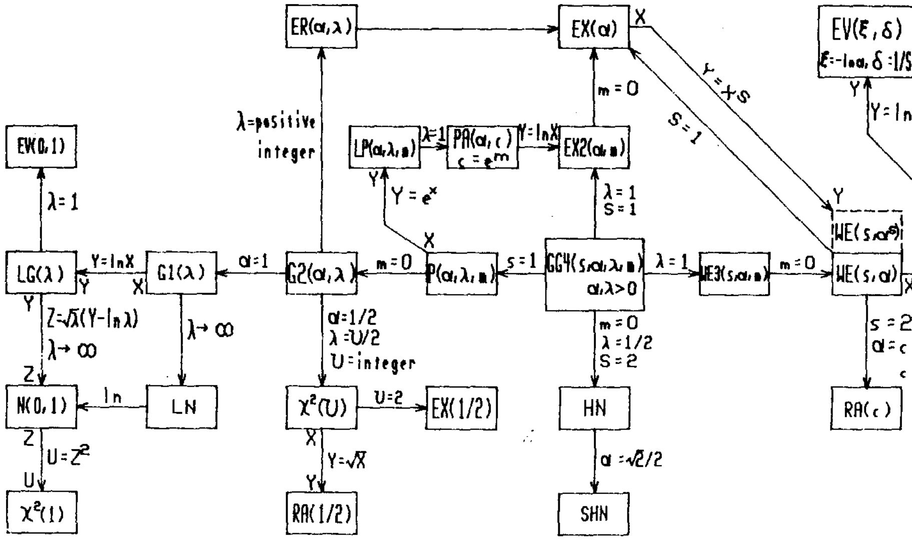

Figure 2. Relations between the GG(4) distribution and other well known distributions

$\chi^2(\nu)$	χ^2 with ν degrees of freedom		
ER	Erlang	LN	Log-normal
EV	Extreme value type 1 or Gumbel	LP	Log-Pearson type 3
EX	Exponential 1 - parameter	N(0,1)	Standard normal
EX2	Exponential 2 - parameter	P	Pearson type 3
G1	Gamma 1 - parameter	PA	Pareto type 1 (Pearson type 4)
G2	Gamma 2 - parameter	RA	Rayleigh
GG4	Generalized gamma 4 - parameter	SHN	Standard half-normal
HN	Half-normal	WE	Weibull 2 - parameter
LG	Log-gamma	WE3	Weibull 3 - parameter

has GG(3) and GG(2) (for which the scale parameter is set equal to unity) as special cases. Moreover, Fig. 2 shows that the 3-parameter log Normal (LN) and the Normal (N) distributions are limiting distributions of GG(4) (note that the 2-parameter LN is a limiting distribution of GG(3)). Also, the Gumbel distribution being a special case of LN(3) (Sangal and Biswas 1970) is also a limiting distribution of GG(4). Finally, studies have shown (Paradis and Bobée 1983) that the GG distribution can be a powerful tool for studying samples of various degrees of skewness.

Note that the GG distribution has first been used successfully in modelling pollution concentration (Nicholson 1975) and also in the study of life distributions (Parr and Webster 1967). Moreover, the GG is the most widely used distribution in the Soviet Union for flood frequency analysis (Kritsky and Menkel 1969). In hydrology, some extensive research (Hoshi and Yamaoka 1980, Paradis and Bobée 1983) has been done in order to derive mathematical and statistical properties of the GG distribution, to develop methods for estimating its parameters and to obtain the associated variance of the estimated T-year flood event (X_T).

Investigations concerning estimation of the parameters of the GG distribution led to several methods:

- indirect method of moments (Stacy and Mihram 1965) which uses the first three moments in log space,
- direct method of moments (Hoshi and Yamaoka 1980), which uses the first three moments in real space,
- method of maximum likelihood (Hager and Bain 1970), and
- four other methods (MM1, MM2, MM3, MM4) referred to as methods of "mixed moments" (Phien et al. 1987). They are all based upon the first two moments of X and the first two moments of $Y = \ln X$: they correspond to the four ways of choosing three moments among these four.

We shall propose for the GG distribution a general method of estimation which we shall call "generalized method of moments" (GMM) which combines any three moments of the distribution and we shall show how many of the previously proposed methods of moments are special cases of this more general method. We shall present a special case of the GMM which we shall call "sundry averages method" (SAM) which uses the harmonic, geometric and arithmetic means of the distribution. Our main aim will therefore be to focus on flexibility by considering:

(1) a very flexible distribution (the GG) which has the large majority of commonly used distributions as special cases; and

(2) a very flexible method of parameter estimation (GMM) which has many of the previously used methods of estimation as special cases.

In hydraulic design, one is concerned not only with the estimation of quantiles X_T (specified annual flood discharge corresponding to a specified return period of T-years) but also with the construction of confidence intervals for these quantiles. Note that it is always possible to derive asymptotic confidence intervals for X_T using the assumption that the estimation $\hat{X}_T$ is normally distributed with mean X_T and variance $\sigma^2(X_T)$ that can be calculated. These confidence intervals are only suitable for large N (where N is the sample size). However, for certain distributions, a method for constructing exact confidence intervals for X_T exists and is consequently also valid for small N. For the generalized gamma distribution no such method has yet been developed. This because, unlike many of the other distributions, the parameters of the generalized gamma distribution are not of the location-scale type. The same problem arose when dealing with the Pearson type III and log Pearson type III distributions but recently Ashkar and Bobée (1988) presented an approximate method for these distributions which performs well with small samples. So, until small-sample confidence intervals (Standard errors) for the GG distribution are developed, large-sample ones will have to be used as an approximation and in the present study we shall derive such large-sample standard errors.

2 Statistical properties of the generalized gamma distribution

The probability density function (p.d.f.) of the 3-parameter GG distribution is given by:

$$f(x;\, \beta,\lambda,s) = \frac{|\, s \,|\, x^{s\lambda-1} e^{-(x/\beta)^s}}{\beta^{s\lambda}\Gamma(\lambda)}; \qquad x > 0 \tag{1}$$

where β is a scale parameter and both λ and s are shape parameters. Both β and λ must be positive but s can be either positive or negative.

Non-central moments of the generalized gamma distribution are given by (Stacy 1962):

$$\mu_r'(x) = \frac{\beta^r\Gamma(\lambda+r/s)}{\Gamma(\lambda)}. \tag{2}$$

These moments are defined only if $(\lambda + r/s) > 0$, i.e., they exist for all r if $s > 0$ and only for $r > -s\lambda$ if $s < 0$. In the remainder of the present study, whenever we shall deal with $\mu_r'(x)$, we shall assume that $\mu_r'(x)$ exists, i.e., $(\lambda + r/s) > 0$.

Using $r = 1$ in Eq. (2), the mean $\mu_1'(x)$ of the GG distribution is obtained:

$$\mu_1'(x) = \frac{\beta\Gamma(\lambda+1/s)}{\Gamma(\lambda)}. \tag{3}$$

Central moments μ_r can be expressed in terms of non central moments μ_r using the following relationship (Kendall et al. 1987):

$$\mu_r(x) = \sum_{j=0}^{r} C_r^j \mu_{r-j}'(-\mu_1')^j. \tag{4}$$

Using $r = 2$ in the above equation, the variance of the GG distribution is obtained:

$$\mu_2(x) = \frac{\beta^2}{\Gamma^2(\lambda)}[\Gamma(\lambda)\Gamma(\lambda+2/s) - \Gamma^2(\lambda+1/s)]. \tag{5}$$

3 Description of the generalized method of moments and its application to the generalized gamma distribution

The generalized method of moments (GMM) applied to the GG distribution can be described as follows:

(1) For a given sample $x_1, \ldots, x_N$ drawn from a GG population, define the non central moment of order r as follows

$$m_r'(x) = \frac{1}{N}\sum_{i=1}^{r} x_i^r. \tag{6}$$

(2) Choose any three distinct real numbers t, u and v different from zero (the case where one of these numbers tends zero will be explicitly considered in the following section since it is an important special case of this method) and write the following three equations:

$$m_t'(x) = \mu_t'(x), \quad m_u'(x) = \mu_u'(x), \quad m_v'(x) = \mu_v'(x), \tag{7}$$

where terms of the left-hand side of Eq. (7) are moments of the sample (Eq. (6)) and those of the right-hand side are the corresponding moments of the population (Eq. (2));

(3) Solve the above equations for the parameters β, λ and s to give the desired estimates $\tilde{\beta}$, $\tilde{\lambda}$, $\tilde{s}$. This is done by isolating the parameter β in the first equation and substituting its value in the last two equations so that the resulting system of equations becomes:

$$g_1(\lambda,1/s) = \beta^u\Gamma(\lambda+u/s) - m_u'(x)\Gamma(\lambda) = 0$$

$$g_2(\lambda,1/s) = \beta^v\Gamma(\lambda+v/s) - m_v'(x)\Gamma(\lambda) = 0 \tag{8}$$

with $\beta = \dfrac{m_t'(x)\Gamma(\lambda)^{1/t}}{\Gamma(\lambda + t/s)}$.

This later system is then solved for λ and s using the Newton-Raphson method where partial derivatives needed for applying this method are:

$$\partial g_1/\partial\lambda = \beta^u\Gamma(\lambda+u/s)\{\psi(\lambda+u/s)+(u/t)[\psi(\lambda)-\psi(\lambda+t/s)]\}-m_u'(x)\Gamma(\lambda)\psi(\lambda) \quad (9)$$

$$\partial g_1/\partial(1/s) = u\beta^u\Gamma(\lambda+u/s)\{\psi(\lambda+u/s) - \psi(\lambda+t/s)\} \quad (10)$$

$$\partial g_2/\partial\lambda = \beta^v\Gamma(\lambda+v/s)\{\psi(\lambda+v/s)+(v/t)[\psi(\lambda)-\psi(\lambda+t/s)]\}-m_v'(x)\Gamma(\lambda)\psi(\lambda) \quad (11)$$

$$\partial g_2/\partial(1/s) = v\beta^v\Gamma(\lambda+v/s)\{\psi(\lambda+v/s) - \psi(\lambda+t/s)\} \quad (12)$$

where $\psi(\cdot)$ is the digamma function defined as $\Gamma'(\cdot)/\Gamma(\cdot)$ where $\Gamma'(\cdot)$ is the derivative of $\Gamma(\cdot)$.

The generalized method of moments which uses the system of Eq. (7) will be denoted by GMM (t,u,v).

4 A special case of the GMM: the sundry averages method (SAM)

It is obvious that putting $r = 0$ in Eqs. (2) and (6) give $\mu_0'(x) = 1$ and $m_0'(x) = 1$, respectively (independently of the parameters β, λ and s) since μ_0' corresponds to the area under the density curve. Therefore, the moment of order zero is not interesting because it cannot be used for estimating the parameters of a distribution. It is interesting to note, however, that when r tends to zero the equation:

$$m_r'(x) = \mu_r'(x) \quad (13)$$

becomes equivalent to (see Ashkar and Bobée 1986):

$$g_x = G_x \quad (14)$$

where g_x is the geometric mean of the sample and G_x is the geometric mean of the population. Let $\bar{y}$ be the arithmetic mean of the sample of logarithmic values $y_i = \ln x_i$ (which is the log of the geometric mean of the observed values x_i) and $\mu_1'(y)$ be the population mean of the variable $Y = \ln X$. Taking natural logs on both sides of Eq. (14) it can be shown that this equation is equivalent to:

$$\bar{y} = \mu_1'(y) \quad (15)$$

where $\mu_1'(y)$ csn be easily calculated using integration of Eq. (1) (see also Stacy and Mihram 1965) to give:

$$\mu_1'(y) = \ln\beta + \psi(\lambda)/s.$$

We shall therefore refer to this case of r tending to zero as the case of moment of order "quasi zero". In the remainder of the paper we shall use the simplified notation "$r = \mathbf{0}$" to mean "$r \to 0$" and shall define $m_{\mathbf{0}}'(x)$ and $\mu_{\mathbf{0}}'(x)$ as follows:

$$m_{\mathbf{0}}'(x) = \bar{y} = (\sum_i \ln x_i)/N \quad (16)$$

$$\mu_0'(x) = \mu_1'(y) = \ln\beta + \psi(\lambda)/s. \tag{17}$$

It is seen from Eq. (17) that $\mu_0'(x)$ does not have the same functional form as $\mu_r'(x)$ in Eq. (2). This remark will later be used for calculating the variance of X_T for certain interesting special cases of the GMM. Two of these interesting special cases are:

- the method MM1 introduced by Rao (1980) for the LP distribution and used by Phien et al. (1987) for the GG distribution. This method can be denoted by GMM(**0**,1,2) because it uses the moment of order **0** (geometric mean), 1 (arithmetic mean) and 2 (variance) of the distribution. It can be easily shown that, when the moment of order 1 is part of the system of equations, there is equivalence between using the non-central moment of order 2 (μ_2') or the central moment of order 2 (variance). Therefore in both cases it is correct to denote the method MM1 by GMM(**0**,1,2);
- a method which we shall call "sundry averages method" (SAM) which can be denoted by GMM(-1,**0**,1) because it uses moments of order -1 (harmonic mean), **0** (geometric mean) and 1 (arithmetic mean) of the distribution (and thus the name "Sundry Averages Method").

5 Estimation of X_T and calculation of its variance

The estimation $\tilde{X}_T$ of the flood magnitude with return period T under a GG distribution is obtained from:

$$\tilde{W}_T = \left(\frac{\tilde{X}_T}{\tilde{\beta}}\right)^s \quad \text{i.e.} \quad \tilde{X}_T = \tilde{\beta}\tilde{W}_T^{1/\tilde{s}} \tag{18}$$

where $\tilde{W}_T$ is the estimator of the event W_T of return period T under a one-parameter gamma distribution with p.d.f.:

$$f(w) = \frac{w^{\lambda-1}e^{-\lambda}}{\Gamma(\lambda)}; \qquad w > 0,\ \lambda > 0. \tag{19}$$

In practice, $\tilde{W}_T$ is calculated from:

$$\tilde{W}_T = \tilde{K}_T\sqrt{\tilde{\lambda}} + \tilde{\lambda} = 2\frac{\tilde{K}_T}{\tilde{C}_s} + \frac{4}{\tilde{C}_s^2} \qquad (\text{since } \tilde{C}_s = 2/\sqrt{\tilde{\lambda}}) \tag{20}$$

where $\tilde{K}_T$ is a frequency factor (standardized Pearson type 3 variate) which has been extensively tabulated as a function of the estimated coefficient of skewness $\tilde{C}_s$ of the random variable W (Harter 1969).

In subsequent calculations we shall drop the symbol "$\sim$", for simplicity. If X follows a GG distribution, then obviously X_T is a function of the three parameters of the distribution, and for large sample sizes we have:

$$\begin{aligned}\mathrm{Var}X_T = {} & \left(\frac{\partial X_T}{\partial\beta}\right)^2\mathrm{Var}\beta + \left(\frac{\partial X_T}{\partial\lambda}\right)^2\mathrm{Var}\lambda + \left(\frac{\partial X_T}{\partial s}\right)^2\mathrm{Var}s \\ & + 2\left(\frac{\partial X_T}{\partial\beta}\right)\left(\frac{\partial X_T}{\partial\lambda}\right)\mathrm{cov}(\beta,\lambda) + 2\left(\frac{\partial X_T}{\partial\beta}\right)\left(\frac{\partial X_T}{\partial s}\right)\mathrm{cov}(\beta,s)\end{aligned}$$

$$+ 2(\frac{\partial X_T}{\partial \lambda})(\frac{\partial X_T}{\partial s})\mathrm{cov}(\lambda,s). \tag{21}$$

Using $X_T = \beta W_T^{1/s}$ the partial derivatives in Eq. (21) are obtained as follows:

$$\frac{\partial X_T}{\partial \beta} = W_T^{1/s} \tag{22}$$

$$\frac{\partial X_T}{\partial \lambda} = (\beta/s)W_T^{(1/s-1)}(\partial W_T/\partial \lambda) \tag{23}$$

$$\frac{\partial X_T}{\partial s} = -(X_T/s^2)\ln W_T \tag{24}$$

and by Eq. (20), $(\partial W_T/\partial \lambda)$ is given by:

$$\partial W_T/\partial \lambda = 1 + \frac{K_T}{2\sqrt{\lambda}} - \frac{1}{\lambda}\frac{\partial K_T}{\partial C_s} \tag{25}$$

where $(\partial K_T/\partial C_s)$ is obtainable from the polynomial given by Bobée and Boucher (1981) or from the formula given by Hoshi and Burges (1981).

It is obvious that the parameter variances and covariances involved in Eq. (21) are functions of the variances and covariances of the three moments $\mu_t'(x)$, $\mu_u'(x)$ and $\mu_v'(x)$ that are employed in the estimation (Eq. (7)). Using matrix notation, these variances and covariances of the parameters are given by:

$$V_p = V^{-1}\cdot V_m \tag{26}$$

i.e.,

$$\begin{bmatrix} \mathrm{Var}\beta \\ \mathrm{Var}\lambda \\ \mathrm{Var}s \\ \mathrm{Cov}(\beta,\lambda) \\ \mathrm{Cov}(\beta,s) \\ \mathrm{Cov}(\lambda,s) \end{bmatrix} = \begin{bmatrix} V_{11} & \cdots & V_{16} \\ V_{21} & \cdots & V_{26} \\ V_{31} & \cdots & V_{36} \\ V_{41} & \cdots & V_{46} \\ V_{51} & \cdots & V_{56} \\ V_{61} & \cdots & V_{66} \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{Var}M_1 \\ \mathrm{Var}M_2 \\ \mathrm{Var}M_3 \\ \mathrm{Cov}(M_1, M_2) \\ \mathrm{Cov}(M_1, M_3) \\ \mathrm{Cov}(M_2, M_3) \end{bmatrix} \tag{27}$$

where $M_1 = \mu_t'(x)$; $M_2 = \mu_u'(x)$; $M_3 = \mu_v'(x)$.

The matrix V is given by:

$$\begin{bmatrix} A_{11}^2 & A_{12}^2 & A_{13}^2 & 2A_{11}A_{12} & 2A_{11}A_{13} & 2A_{12}A_{13} \\ A_{21}^2 & A_{22}^2 & A_{23}^2 & 2A_{21}A_{22} & 2A_{21}A_{23} & 2A_{22}A_{23} \\ A_{31}^2 & A_{32}^2 & A_{33}^2 & 2A_{31}A_{32} & 2A_{31}A_{33} & 2A_{32}A_{33} \\ A_{11}A_{21} & A_{12}A_{22} & A_{13}A_{23} & (A_{11}A_{22}+A_{12}A_{21}) & (A_{11}A_{23}+A_{13}A_{21}) & (A_{12}A_{23}+A_{13}A_{22}) \\ A_{11}A_{31} & A_{12}A_{32} & A_{13}A_{33} & (A_{11}A_{32}+A_{12}A_{31}) & (A_{11}A_{33}+A_{13}A_{31}) & (A_{12}A_{33}+A_{13}A_{32}) \\ A_{21}A_{31} & A_{22}A_{32} & A_{23}A_{33} & (A_{21}A_{32}+A_{22}A_{31}) & (A_{21}A_{33}+A_{23}A_{31}) & (A_{22}A_{33}+A_{23}A_{32}) \end{bmatrix} \tag{28}$$

where

$$A_{rj} = \frac{\partial M_r}{\partial \Theta_j} \qquad r,j = 1,2,3 \tag{29}$$

and $\Theta_1 = \beta$, $\Theta_2 = \lambda$, $\Theta_3 = s$.

The terms A_{rj} are therefore obtained from Eq. (2) by partial differentiation and are given in the following matrix:

$$A = (A_{rj}) = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{31}$$

where

$$\begin{aligned} A_{11} &= t\beta^{t-1}\frac{\Gamma(\lambda+t/s)}{\Gamma(\lambda)}, \quad A_{12} = \beta^t \frac{\Gamma(\lambda+t/s)}{\Gamma(\lambda)}[\psi(\lambda+t/s)-\psi(\lambda)] \\ A_{13} &= \frac{-t\beta^t}{s^2\Gamma(\lambda)}[\Gamma(\lambda+t/s)\psi(\lambda+t/s)] \\ A_{21} &= u\beta^{u-1}\frac{\Gamma(\lambda+u/s)}{\Gamma(\lambda)}, A_{22} = \beta^u \frac{\Gamma(\lambda+u/s)}{\Gamma(\lambda)}[\psi(\lambda+u/s)-\psi(\lambda)] \\ A_{23} &= \frac{-u\beta^u}{s^2\Gamma(\lambda)}[\Gamma(\lambda+u/s)\psi(\lambda+u/s)] \\ A_{31} &= v\beta^{v-1}\frac{\Gamma(\lambda+v/s)}{\Gamma(\lambda)}, A_{32} = \beta^v \frac{\Gamma(\lambda+v/s)}{\Gamma(\lambda)}[\psi(\lambda+v/s)-\psi(\lambda)] \\ A_{33} &= \frac{-v\beta^v}{s^2\Gamma(\lambda)}[\Gamma(\lambda+v/s)\psi(\lambda+v/s)]. \end{aligned} \tag{32}$$

The entries of the vector V_m in Eq. (26) are easily obtained if we recall from Kendall et al. (1987) that for arbitrary r and q and arbitrary sample size N, we have:

$$\mathrm{Var}(\mu_r') = (\mu_{2r}' - \mu_r'^2)/N \tag{33}$$

$$\mathrm{Cov}(\mu_r', \mu_q') = (\mu_{q+r}' - \mu_r'\mu_q')/N. \tag{34}$$

Since the elements of the vector V_m (Eq. (26)) and those of the matrix V (Eq. (28)) are now well determined, the vector V_p can be easily deduced using

$$V_p = V^{-1} \cdot V_m.$$

Hence by knowing the variances and covariances of the parameters, β, λ and s (i.e., the vector V_p) and the partial derivatives of X_T with respect to each of these parameters (Eqs. (22), (23), and (24)), the calculation of Var X_T becomes straightforward via Eq. (21) and can be easily programmed on a computer.

One can also be interested in computing correlation coefficients between moments or between parameters. This can be done by using the following relationship:

$$\rho(x,y) = \mathrm{cov}(x,y)/(\mathrm{var}x \cdot \mathrm{var}y)^{1/2} \tag{35}$$

and letting, for example, $x = M_r$ and $y = M_q$ in the moments case or, $x = \beta$ and $y = \lambda$ in the parameter case.

6 Special case of moment of order quasi zero

In the presentation of the sundry averages method (SAM) it was mentioned that the moment of order quasi zero, $\mu_0'(x)$, has a functional form different from that of $\mu_r'(x)$. This means that certain terms involved in the calculation of VarX_T would have to be modified whenever the moment $\mu_0'(x)$ is used in the estimation. Such is the case not only when the method SAM (GMM(**0**,-1,1)) is used, but also whenever the moment of order "quasi-zero" is involved (like in the method MM1 or GMM(**0**,1,2) for example). Terms to be modified are all those involving $t = \mathbf{0}$, or $M_1 = \mu_t'(x) = \mu_0'(x)$. These precisely are:

- the terms VarM_1, Cov(M_1, M_2) and Cov(M_1, M_3) of the vector V_m in Eq. (26);
- all terms in the first row of the matrix A in Eq. (31).

The new terms will be (see appendix):

$$\mathrm{Var}M_1 = \psi'(\lambda)/Ns^2 \tag{36}$$

where $\psi'(\cdot)$ is the trigamma function

$$\mathrm{Cov}(M_1, M_2) = \frac{\beta^u\Gamma(\lambda + u/s)}{Ns\Gamma(\lambda)}[\psi(\lambda + u/s) - \psi(\lambda)] \tag{37}$$

$$\mathrm{Cov}(M_1, M_3) = \frac{\beta^v\Gamma(\lambda + v/s)}{Ns\Gamma(\lambda)}[\psi(\lambda + v/s) - \psi(\lambda)] \tag{38}$$

$$A_{11} = 1/\beta \tag{39}$$

$$A_{12} = \psi'(\lambda)/s \tag{40}$$

$$A_{13} = -\psi'(\lambda)/s^2. \tag{41}$$

A comparison of VarX_T calculated by different versions of the GMM has been undertaken and the results are presented in the following section.

7 Comparison of VarX_T obtained by various versions of the GMM

The GMM is a method of estimation which is very broad in the sense that it allows one to combine orders of moments in such a way that the chosen combination performs better than others with respect to reducing VarX_T. Note however that one should not expect to find one method of estimation to be best for all return periods.

The example provided here consists of the comparison between the following methods:
GMM (-1, **0**, 1) or SAM, GMM (**0**, 1, 2) or MM1 and GMM (1, 2, 3)

The three different methods were applied to a set of observations made of annual maximum daily discharge recorded from 1916 to 1974 at the station located on St-Mary's river (at Stillwater) in Nova Scotia (Canada). This set of data was also used by Kite (1977).

Results are presented in Fig. 3 where the standard error of X_T is given for different values of the return period T. Each curve is associated with one of the three methods of estimation.

For this set of observations, Figure 3 indicates that lowest values for standard error of X_T are obtained by the sundry averages method or GMM(-1, **0**, 1) indicating that this method should probably be preferred over the two others in this particular case. Here, it appears that making use of moments of lower order (-1,

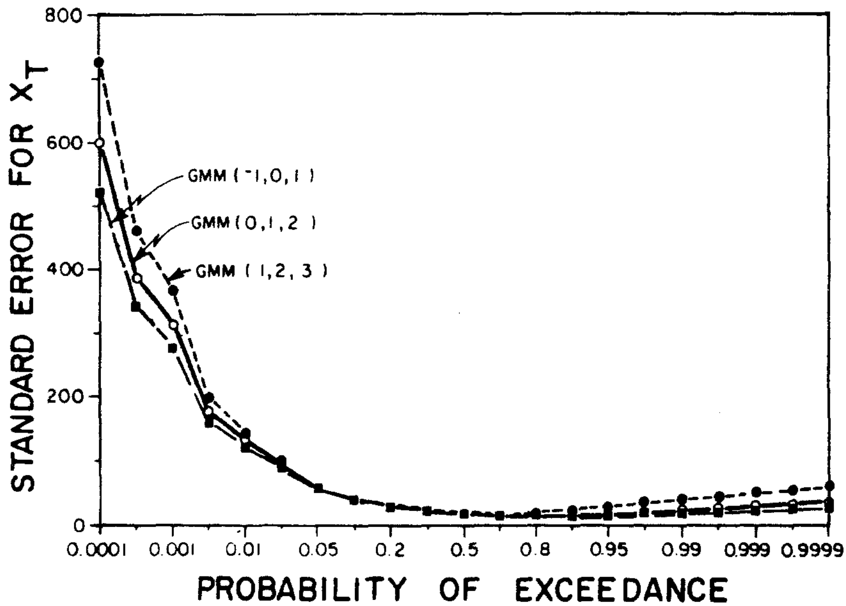

Figure 3. Standard error for X_T obtained by different versions of the GMM for various probabilities of exceedance

quasi-zero and 1 as compared to 1, 2 and 3 for example) helps in reducing the standard error of X_T for all return periods considered. Further investigations would be needed in order to confirm this preliminary result and to examine in more detail (by Monte Carlo simulation for example) how the different versions of the method of moments would compare to each other when applied to the GG distribution.

8 Summary and conclusion

The main purpose of the present study was to apply a very flexible method of parameter estimation (GMM) to a very flexible distribution (GG). This was done by:

(1) pointing out the interest of the generalized gamma distribution as a flexible distribution for hydrological applications and showing in a schematical diagram how many of the widely used distributions are either special cases or limiting cases of the GG distribution;

(2) showing how the generalized method of moments (GMM), can be considered as a generalization of the various methods of moments;

(3) deriving the general equations needed to apply the GMM to the GG distribution;

(4) presenting an interesting special case of the GMM which was called the sundry averages method (SAM). This method makes use of the arithmetic, geometric and harmonic means of the sample, which are respectively moments of order 1, quasi zero and -1. Reducing the order of the moments in this manner can lead to a lower value of Var(X_T) for certain values of T;

(5) determining the general formula for VarX_T for the GMM applied to the GG distribution including the special case of moment of order quasi zero;

(6) providing an example of the use of the GMM for annual flood series and pointing out the need for further comparison between the different versions of the GMM.

Acknowledgements

This study was made possible by grants from the Natural Sciences and Engineering Research Council of Canada (operating Grants A-8399 and A-5797 and Strategic Grant G-1617).

References

Ashkar, F.; Bobée, B. 1986: Comment on: Variance of the T-year event in the Log-Pearson type 3 distribution. J. of Hydrology 84, 181-187

Ashkar, F.; Bobée, B. 1988: Confidence intervals for extreme flood events under a Pearson type 3 or log Pearson type 3 distribution. Water Resour. Bulletin (accepted for publication)

Benson, M.A. 1968: Uniform flood frequency estimating methods for federal agencies. Water Resour. Res. 4, 891-908

Bobée, B.; Boucher, P. 1981: Calcul de la variance d'un évènement de période de retour T: cas des lois log-Pearson type 3 et log-gamma ajustés par la méthode des moments sur la série des valeurs observées. INRS-Eau, Rapp. Sci. 135, 17 pp

Hager, H.W.; Bain, L.J. 1970: Inferential procedures for the generalized gamma distribution. J. Amer. Statist. Assoc. 65, 1601-1609

Harter, H.L. 1969: A new table of percentage points of the Pearson type 3 distribution. Technometrics 11, 177-187

Hoshi, K.; Burges, S.J. 1981: Approximate estimation of the derivative of a standard gamma quantile for use in confidence interval estimates. J. of Hydrology 53, 317-325

Hoshi, K.; Yamaoka, I. 1980: The generalized gamma probability distribution and its application in hydrology. Hydraulics papers, 7. The research laboratory of civil and environmental engineering, Hokkaido University, Japan, 256 pp

Kendall, M.J.; Stuart, A.; Ord, J.K. 1987: Kendall's advanced theory of statistics. Vol. 1. 5th ed., Oxford University Press, New York, 604 pp

Kite, G.W. 1977: Frequency and risk analysis in hydrology. Water Resour. Publications, Colorado: Fort Collins, 224 pp

Kritsky, S.N.; Menkel, M.F. 1969: On principles of estimation methods of maximum discharge. Floods and their computation. A.I.H.S. 84

Leroux, D; Bobée, B.; Ashkar, F. 1987: Diagrammes des relations entre certaines fonctions de moments pour quelques lois fréquemment utilisées en hydrologie. INRS-Eau, Rapp. Sci. 237

Nicholson, D.L. 1975: Application of parameter estimation and hypothesis test for a generalized gamma distribution. Proc. IEEE conf. decis. control incl. symp. adapt. processes. Houston, Tex., Dec. 10-12, 321-326

Paradis, M.; Bobée, B. 1983: La distribution Gamma Généralisée et son application en hydrologie. INRS-Eau, Rapp. Sci. 156, 52 pp

Parr, V.B.; Webster, J.T. 1965: A method for discriminating between failure density functions used in reliability predictions. Technometrics 7, 1-10

Phien, H.N.; Nguyen, V.T.V.; Jung-Hua, K. 1987: Estimating the parameters of the generalized gamma distribution by mixed moments. In: V.P. Singh (ed.), Hydrologic Frequency Modeling. D. Reidel Publ. Co., Dordrecht, Holland

Rao, D.V. 1980: Log Pearson type 3 distribution: Method of mixed moments. J. Hydraul. Div. Am. Soc. Civ. Eng. 106, 999-1019

Sangal, B.P.; Biswas, A.K. 1970: The 3-parameter log normal distribution and its applications in hydrology. Water Resour. Res. 2, 505-515

Stacy, E.W. 1962: A generalization of the gamma distribution. Ann. of Math. Stat. 33, 1187-1192

Stacy, E.W.; Mihram, G.A. 1965: Parameter estimation for a generalized gamma distribution. Technometrics 7, 349-358

Appendix A Derivation of equations (36) through (41)

In Eqs. (36) through (41) we have: $M_1 = \bar{Y}$ estimate of $\mu_0'(x) = \mu_1'(y) = \ln\beta + \psi(\lambda)/s$
$M_2 = m_u'(x)$ estimate of $\mu_u'(x) = E(X^u) = \beta^u\Gamma(\lambda + u/s)/\Gamma(\lambda)$
$M_3 = m_v'(x)$ estimate of $\mu_v'(x) = E(X^v) = \beta^v\Gamma(\lambda + v/s)\Gamma(\lambda)$

From the expression for M_1, Eqs. (39), (40) and (41) are easily derived; in fact we have:

$$A_{11} = \partial M_1/\partial\beta = \partial(\ln\beta + \psi(\lambda)/s)/\partial\beta = 1/\beta,$$

$$A_{12} = \partial M_1/\partial\lambda = \partial(\ln\beta + \psi(\lambda)/s)/\partial\lambda = \psi'(\lambda)/s,$$

$$A_{13} = \partial M_1/\partial s = \partial(\ln\beta + \psi(\lambda)/s)/\partial s = -\psi'(\lambda)/s.$$

Equation (36) is obtained as follows:
Since $X \sim \mathrm{GG}\,(\beta, \lambda, s)$, let $Y = \ln X$ and $Z = \ln(X/\beta)^s$ i.e., $\ln X = (Z/s) + \ln\beta$.
Then:

$$\mathrm{Var}M_1 = \mathrm{Var}\bar{Y} = (1/N)\mathrm{Var}Y$$

where

$$\mathrm{Var}Y = \mathrm{Var}(\ln X) = \mathrm{Var}[(Z/s) + \ln\beta] = (1/s^2)\mathrm{Var}Z$$
$$= (1/s^2)\{E(Z^2) - [E(Z)]^2\}.$$

From Stacy and Mihram (1965),

$$E(Z^k) = \frac{\Gamma^{(k)}(\lambda)}{\Gamma(\lambda)}$$

where $\Gamma^{(k)}(\lambda)$ is the kth derivative of $\Gamma(\lambda)$ with respect to the parameter λ. Then after some manipulations we obtain:

$$\mathrm{Var}Y = (1/s^2)\{\frac{\Gamma^{(2)}(\lambda)}{\Gamma(\lambda)} - [\frac{\Gamma^{(1)}(\lambda)}{\Gamma(\lambda)}]^2\} = (1/s^2)\psi'(\lambda).$$

Finally,

$$\mathrm{Var}M_1 = \psi'(\lambda)/Ns^2.$$

Equation (37) is derived using M_1 and M_2 as follows:

$$\begin{aligned}\mathrm{Cov}(M_1, M_2) &= E[\bar{Y}\cdot m_u'(x)] - E(\bar{Y})\cdot E(m_u'(x)) \\ &= E[(1/N)\textstyle\sum Y_i)\cdot((1/N)\sum X_i^u)] - E(Y)\cdot E(X^u) \\ &= (1/N^2)[\textstyle\sum E(Y_i X_i^u) + (N^2 - N)E(Y)\cdot E(X^u)] - E(Y)\cdot E(X^u) \\ &= (1/N^2)[NE(YX^u) + (N^2 - N)E(Y)\cdot E(X^u)] - E(Y)\cdot E(X^u) \\ &= (1/N)E(YX^u) + (1 - 1/N)E(Y)\cdot E(X^u) - E(Y)\cdot E(X^u) \\ &= (1/N)[E(YX^u) - E(Y)\cdot E(X^u)] \qquad \text{(A1)}\end{aligned}$$

where:

$$E(Y) = \mu_1'(y) \qquad \text{(A2)}$$

$$E(X^u) = \mu_u'(x) \qquad \text{(A3)}$$

and

$$E(YX^u) = E(\ln X \cdot X^u) = \int x^u \ln x f(x)\, dx \tag{A4}$$

where $f(x)$ is the probability density function of the generalized gamma distribution (Eq. (1)) and the integration is over the whole domain of variation of this distribution.

Making the appropriate change of variable and after integration and some mathematical manipulations Eq. (A4) can be shown to be equal to:

$$E(X^u Y) = \frac{\beta^u \Gamma(\lambda + u/s)}{\Gamma(\lambda)} (\ln\beta + \psi(\lambda + u/s)/s). \tag{A5}$$

Substituting Eqs. (A5), (A3) and (A2) into Eq. (A1) gives the desired Eq. (37).

Equation (38) is derived in an exactly similar manner.

Accepted May 20, 1988.