ELSEVIER

Journal of Hydrology 258 (2002) 122–148

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Asymptotic bias of estimation methods caused by the assumption of false probability distribution

Witold G. Strupczewski[a,*], Vijay P. Singh[b,1], Stanislaw Weglarczyk[c,2]

[a]Department of Water Resources, Institute of Geophysics, Polish Academy of Sciences, ul. Ksiecia Janusza 64, 01-452 Warsaw, Poland
[b]Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA 70803-6405, USA
[c]Institute of Water Engineering and Water Management, Cracow University of Technology, Warszawska 24, 31-155 Cracow, Poland

Received 5 January 2001; revised 12 September 2001; accepted 24 October 2001

Abstract

Asymptotic bias in large quantiles and moments for four parameter estimation methods, including the maximum likelihood method (MLM), method of moments (MOM), method of L-moments (LMM), and least squares method (LSM), is derived when a probability distribution function (PDF) is falsely assumed. The first three estimation methods are illustrated using the lognormal and gamma distributions forming an alternative set of PDFs. It is shown that for every method when either the gamma or lognormal distribution serves as the true distribution, the relative asymptotic bias (RB) of moments and quantiles corresponding to the upper tail is an increasing function of the true value of the coefficient of variation (c_v), except that RB of moments for MOM is zero. The value of RB is the smallest for MOM and largest for MLM. The bias of LMM occupies an intermediate position. The value of RB from MLM is larger for the lognormal distribution as a hypothetical distribution with the gamma distribution being assumed to be the true distribution than it would be in the opposite case. For $c_v = 1$ and MLM, it equals 30, 600, 320% for mean, variance and 0.1% quantile, respectively, while for MOM, the moments are asymptotically unbiased and the bias for 0.1% quantile amounts to 35%. An analysis of 39 70-year long annual peak flow series of Polish rivers provides an empirical evidence for the necessity to include bias in evaluation of the efficiency of PDF estimation methods.

Keywords: Flood frequency; Asymptotical bias; Moment estimation; Quantile estimation; Maximum likelihood; Method of moments

1. Introduction

Flood frequency analysis (FFA) entails estimation of the upper tail of a probability distribution function (PDF) of peak flows obtained from either the annual duration series or partial duration series, although the upper part of the PDF may usually be out of the range of observations. The usual empirical approach is to fit an a priori assumed PDF to the peak flow data, where the fitting involves estimating the parameters of the PDF, which, in turn, requires the knowledge of the PDF. Thus, one tries to find and use the most robust method of parameter estimation for a given sample size. Unfortunately, the true PDF is not known and even if it were known it might, in all probability, contain too many parameters. These parameters cannot possibly be estimated reliably and efficiently from a hydrological sample, which is of relatively

* Corresponding author. Fax: +48-22-6915915.
E-mail addresses: wgs@igf.edu.pl (W.G. Strupczewski), cesing@unix1.sncc.lsu.edu (V.P. Singh), sweglar@lajkonik.wis.pk.edu.pl (S. Weglarczyk).
[1] Fax: +1-225-578-6588.
[2] Fax: +48-12-6282041.

PII: S0022-1694(01)00563-7

small size, meaning that strictly speaking such a PDF cannot be applied. Therefore, the task of FFA reduces to (1) choosing the PDF which can be derived either by 'at site' or 'regional' analysis; and (2) finding and using the most robust method of parameter estimation which produces the smallest mean square error (MSE) and bias in moments as well as in quantiles of interest for a given sample size and the chosen distribution.

A number of two- and three-parameter PDFs have been discussed in the literature for hydrologic FFA (Hosking and Wallis, 1997; Singh, 1998; Rao and Hamed, 2000). Although three-parameter probability models are often recommended for FFA, two-parameter distributions were chosen in this study for two reasons. First, the constraints with respect to the parameters representing location, scale and shape are very rigid for hydrologic problems and even three parameters may be too many in case of normal hydrological sample sizes when the regional flood information is not exploited (e.g. Landwehr et al., 1980; Kuczera, 1982; Strupczewski et al., 2001). Second, the objective is to show the significance of bias using simpler models which are still in use in many parts of the world. Landwehr et al. (1980) showed that the two-parameter Gumbel and three-parameter lognormal distribution produced lower MSE than the five-parameter Wakeby distribution. For a sample size of 31, the share of bias in MSE was found to be as high as 93% for 0.1% probability of exceedance for Gumbel and up to 33% for the lognormal. In a simulation study, Kuczera (1982) found the two-parameter lognormal distribution as the best model. The Gumbel distribution with its parameters estimated by either maximum likelihood method (MLM) or probability weighted moments (PWM) also displayed comparable performance. This shows that a uniform consensus on the choice of a model for FFA is lacking and the choice is critically important.

Several parameter estimation methods for these PDFs have been developed (Singh, 1988; Rao and Hamed, 2000). Since the statistics used in every estimation method differ from each other, a method of fitting a theoretical distribution to an empirical one depends on the estimation method itself and in case of MLE on the distribution function as well. The differences in fitting may become crucial if an assumed PDF differs from the true one while for practical reasons the interest is in high accuracy of estimation in a certain range of variability, i.e. in the upper tail of the distribution. The MLM is considered as the most theoretically correct method in the sense that it produces the most efficient parameter estimates. The secret of the high efficiency of the MLM lies in its ability to extract greater amount of information from the assumed distribution function, which is required for the use of MLM. The question then arises: do we possess sufficient knowledge to assume a distributional form for statistical parameter estimation? The assumption of a PDF results in biased estimates of the distribution moments as well as quantiles.

The objective of the present study is to analytically derive asymptotic bias in moments as well as in large quantiles and illustrate it with examples. A study of asymptotic bias can serve as a basis to assess its magnitude and give an idea about the bias for any sample size, and whether the difference in the bias due to various parameter estimation methods can be counterbalanced by their efficiency of estimation. It would also be useful to verify the correctness of Monte Carlo experiments. The paper is organized as follows. Providing a short review of flood frequency modelling in Section 2, the problem of estimation of bias and relative bias in flood statistical characteristics is introduced in Section 3. Section 4 introduces PDFs considered in the study. Section 5 discusses expressing distribution parameters in terms of moments. Section 6 is the largest one, dealing with the MLM as an approximation method illustrated by an example where the gamma PDF is taken as the true distribution function while the lognormal distribution as the hypothetical one. The opposite case is presented in Appendix A. The method of moments (MOM) is dealt with in Section 7 and the method of L-moments (LMM) in Section 8. The least squares approximation is discussed in Section 9. Section 10 discusses empirical testing and Section 11 concludes the paper.

2. Review of literature

There is vast literature on FFA and no effort is made to review it here. However, a short discussion of four aspects considered relevant to the objective of this study is presented.

Table 1
Statistical characteristics of six Wakeby specific distributions with zero lower bound used by Landwehr et al. (1980)

PDF	Statistical characteristics				
	μ	σ	c_v	c_s	λ
Wakeby-1	1.94	1.34	0.69	4.14	63.74
Wakeby-2	1.56	0.90	0.58	2.01	14.08
Wakeby-3	1.18	1.03	0.87	1.91	10.73
Wakeby-4	1.36	0.51	0.38	1.10	7.69
Wakeby-5	0.92	0.70	0.76	1.11	4.73
Wakeby-6	0.92	0.46	0.50	0.00	2.65

2.1. Flood frequency models

There is a wide range of flood frequency models developed in hydrology (Greis, 1983). These can be grouped into four types (Singh and Adrian, 2000): (a) empirical, (b) phenomenological, (c) information-based, and (d) physically based. A thought-provoking critique of FFA models has been presented by Klemes (2000a,b). Bras et al. (1985) compared three physically based flood frequency models, which generate flood frequency distributions without the use of streamflow records, and tested them on five different river basins in the US. None of the methods compared well with data-based methods. The most popular types are empirical models which Cunnane (1985) classified into annual maximum series, partial duration series, and time series models. These models are based on fitting a probability distribution to empirical data. A number of probability distributions have been used for FFA, and a discussion of these distributions is given in Rao and Hamed (2000).

2.2. Choosing a flood frequency model

There are a number of methods by which to choose a flood probability distribution, but the treatment considering the resistance of methods with respect to the distribution choice has not been investigated as fully (Landwehr et al., 1980; Kuczera, 1982). Mockus (1960) presented some of the elements in selecting a method for frequency analyses of hydrologic data. Cunnane (1985) discussed factors affecting the choice of a distribution for FFA, including the method of parameter estimation, treatment of outliers, inclusion of large historical flood values, data transformations, and causal composition of flood population. He concluded that distribution choice could not be based on theoretical arguments alone or one criterion. Gupta (1970) presented a method for selecting among 10 commonly used FFA methods to fit frequency distributions to hydrologic data. Spence (1973) used the correlation coefficients for selecting the best of four flood frequency distributions for fitting annual flood flow data from 161 drainage basins in Canadian plains.

Turkman (1985) proposed the Akaike's information criterion (AIC) for the choice of extremal models and analysed its effectiveness in choosing the most likely among the Gumbel, Frechet and Weibull models. Mutua (1994) used AIC in the identification of an optimum flood frequency model in Kenya from the class comprised of seven three-parameter and two five-parameter flood frequency models and for testing the existence of outliers. Chong and Moore (1983) used the residual sum of squares (RSS) to compare two-parameter lognormal (LN2), three-parameter lognormal (LN3), Pearson type 3 (PT3), and log-Pearson type 3 (LPT3) distributions and selected the distribution that produced the smallest value of RSS for developing a regional curve. However, it was difficult to say which method was the best for regional FFA.

2.3. Methods of parameter estimation

Popular methods of parameter estimation are the MOM (Nash, 1959), PWM (Greewood et al., 1979; Landwehr et al., 1979), LMM (Hosking, 1990; Hosking and Wallis, 1997), MLM (Douglas et al., 1976), maximum entropy method (MEM; Singh and Rajagopal, 1986; Singh, 1998), and least squares method (LSM; Snyder, 1972; Stedinger and Tasker, 1985). Several studies have compared methods of parameter estimation using the standard error of estimate as a criterion. Landwehr et al. (1980) and Kuczera (1982) presented an interesting evidence of the statistical overparameterization. They showed that the knowledge of the 'true' model is not sufficient to accept such a model as data may be too short to calibrate it. In their simulation experiments, Landwehr et al. (1980) considered six specific Wakeby distributions lower-bounded at zero as the parent (i.e. true) distributions and a sample size of 31. Values of the

Table 2
Share of bias (in %) in MSE computed by Landwehr et al. (1980) for six variants of the Wakeby distribution shown in Table 1. Sample size $N = 31$

True PDF	Fitted PDF	Estimation method	Probability of exceedance	
			1%	0.1%
Wakeby-1	Wakeby	LMM	0	1
	Gumbel	LMM	64	92
		MLM	86	96
		MOM	34	78
	Lognormal	MOM	14	35
Wakeby-2	Wakeby	LMM	0	4
	Gumbel	LMM	34	76
		MLM	48	86
		MOM	22	64
	Lognormal	MOM	16	33
Wakeby-3	Wakeby	LMM	0	0
	Gumbel	LMM	26	65
		MLM	64	85
		MOM	18	54
	Lognormal	MOM	10	20
Wakeby-4	Wakeby	LMM	3	11
	Gumbel	LMM	12	38
		MLM	11	0
		MOM	8	33
	Lognormal	MOM	21	31
Wakeby-5	Wakeby	LMM	0	0
	Gumbel	LMM	0	10
		MLM	3	1
		MOM	0	0
	Lognormal	MOM	0	0
Wakeby-6	Wakeby	LMM	0	7
	Gumbel	LMM	85	89
		MLM	92	93
		MOM	81	87
	Lognormal	MOM	0	20

statistical characteristics, mean, standard deviation, coefficient of variation, skewness and kurtosis (μ, σ, c_v, c_s, λ) are given in Table 1. They found that the five-parameter Wakeby as hypothetical distribution happened to be worse with respect to MSE of upper quantiles than the three-parameter log-normal distribution and the (two-parameter) Gumbel model. The last one combined with MOM usually produced the lowest value of MSE. Therefore, the model can be a best one but the hydrologic data are too short to acknowledge it. As shown in Table 2, the share of bias in MSE for 0.1% probability of exceedance was found to be negligibly small for Wakeby but as high as 96% for Gumbel and MLM and up to 33% for the lognormal. It shows that the bias share of MSE can be significant even for small samples for long return periods and that the study of asymptotic bias has relevance for the sample sizes encountered in practice. In a similar simulation study, Kuczera (1982) found the two-parameter lognormal as the best model and the Gumbel distribution with its parameters estimated by either MLM or PWM also displaying a comparable performance. These studies indicate that the constraints with respect to the number of parameters are rigid for normal hydrological sample sizes and that the doctrine of parameter parsimony should be observed in the distribution choice.

There are numerous hydrologic studies dealing with comparison of the accuracy of various methods of parameter estimation for various distributions and Monte Carlo simulated sample sizes. For choosing an estimation method, the approach used in FFA follows the findings based on the case of a known distribution form, where the robustness of the methods is considered. A robust method performs well over a range of situations and is able to withstand a certain amount of abuse without breaking down. It is not necessarily the best estimation method for any one model, and is characterized in terms of stability and consistency of parameter estimates. Stable estimates are characterized by small estimator dispersion or variance, while consistency implies estimates converge in probability to the 'true' value of the parameters as the number of observation becomes large.

2.4. Model and sampling errors

Model errors are caused by the wrong choice of the model, the wrong estimation of model parameters and the inadequate sample size for parameter estimation. Landwehr et al. (1980) showed that the two-parameter Gumbel and three-parameter lognormal distribution produced lower MSE than the five-parameter Wakeby distribution. For a sample size of 31, the share of bias in MSE was found to be as high as 93% for 0.1% probability of exceedance for Gumbel and up to 33% for the lognormal. In a simulation study, Kuczera (1982) found the two-parameter lognormal distribution as the best model. The Gumbel distribution with its parameters estimated by either MLM or PWM also displayed comparable performance. This shows that a uniform consensus on the choice of a model for FFA is lacking and the choice is critically important.

Bobee (1973) derived the sample error of *T*-year events computed by fitting a Pearson type 3 distribution. Condie (1977, 1986) expressed a *T*-year event derived from LPT3 or LN3 distribution by MLM as a function of parameters, which are subject to sampling variance and covariance. By comparison with MOM, MLM exhibited less bias for LPT3 distribution. Condie (1986) derived asymptotic standard error of estimate of the *T*-year flood event. A combined analysis of systematic record and historical flood would give upwardly biased estimate of the *T*-year flood. Censoring theory helped reduce the bias. Hoshi and Burges (1981) derived for LP3 populations the variance of the *T*-year event, when MOM was used. Phien and Hsu (1985) used the asymptotic variance of the *T*-year event for evaluating the performance of parameter estimation methods for fitting LP3 distribution to a set of observed data.

Wang and Singh (1994) derived the sampling variance of a *T*-year flood estimated by curve fitting using plotting positions and showed that the error due to the plotting position contributed more to the sampling variance than the error in hydrologic observations and model fitting. Stevens (1992) showed for three hypothetical populations that there was a considerable reduction in bias and variance of the extreme flood (500-year to 2000-year) if historical data was used, as opposed to basing the estimates on gauge record alone.

Buishand (1990) discussed approximations to the bias of a *T*-year flood and showed that small departures from the assumed model could have a large impact on the variance of the flood. The ML estimate of the *T*-year flood was highly biased. By applying MLM and censored sample theory to LP3 distribution for FFA, Pilon and Adamowski (1993) derived the asymptotic standard error of estimate of the *T*-year flood.

3. Asymptotic bias

The MSE of any statistical characteristic, Z, can be expressed as

$$\mathrm{MSE}(Z) = \mathrm{var}(Z) + [\mathrm{Bias}(Z)]^2 \tag{1}$$

where var(Z) is the variance of Z and Bias(Z) is the bias of Z. For a given sample size, the ratio of the two terms in Eq. (4) depends on both the PDF model and the parameter estimation method. An increase in the number of model parameters (degrees of freedom) increases the first term and decreases the second one. For large samples, the standard deviation of the Z estimate becomes small in comparison to the bias caused by the wrong distribution choice (i.e. by the model error) and therefore MSE approaches the square of the asymptotic bias.

$$B(Z) = \lim_{N\to\infty} \mathrm{Bias}(Z) \tag{2}$$

It should be remarked that the statistical theory is based on asymptotic properties and only a limited number of results are available for finite samples. As a result, asymptotic formulae are used for small samples if the theoretical approach fails and the Monte Carlo assessment is not available. For this reason, one can be dubious when dealing with the asymptotic bias in hydrology for limited sample sizes available. Evaluating the asymptotic bias caused by the false distributional assumption, one should realise that for a finite sample, the bias would likely be a little greater. It is because even for the correct distributional assumption, any estimation method is not bias free for small samples. Applying the five-parameter Wakeby distribution to six specific variants of the Wakeby distribution with lower-bound at zero, Landwehr et al. (1980) found for a sample size of 31 that the share of the bias in MSE was 1, 4, 0, 11, 0 and 7%, depending on the parameter values of the parent distribution (Tables 1 and 2). This shows that the asymptotic bias can give an idea about the magnitude of bias for any sample size. When dealing with small samples, the bias caused by the wrong distributional assumption with respect to any statistical characteristic should be compared for a given estimation method with one for the proper choice of the distribution. This can be accomplished by simulation techniques the discussion of which is beyond the scope of this study. An advantage of the analytical approach applied here over the Monte Carlo simulation experiments is, except for accuracy, the possibility of getting functional relationships between the asymptotic bias and population parameters, which in the simulation approach has to be reproduced point by point. Therefore, an analytically derived asymptotic bias can serve to verify the correctness of Monte Carlo simulation experiments for samples of various sizes. Furthermore, a comparison of the values of the asymptotic bias got by various approximation methods may give an idea whether the differences between them can be counterbalanced by the difference in the efficiency of estimation methods.

In order to derive the asymptotic bias (B) of Z caused by the false (F) choice of the distributional hypothesis (H), the knowledge of the true distribution (T) together with the value of its parameters is necessary. Then, the problem is defined as an approximation of the T-function by the F-function and it, therefore, remains no longer a statistical estimation problem. Having approximated T by F, one can find for any characteristic Z both the value of z of the approximated function, i.e. $z(H = T)$ and the corresponding value of z of the approximating function, i.e. $z(H = F|T)$. Thus, the asymptotic bias of any statistical characteristic Z, $B(Z)$, is defined as

$$B(Z) = z(H = F|T) - z(H = T) \quad (3)$$

and the relative asymptotic bias, $\mathrm{RB}(Z)$, as

$$\mathrm{RB}(Z) = \frac{z(H = F|T) - z(H = T)}{z(H = T)} \quad (4)$$

where H, F and T stand for hypothetical, false and true distributions, respectively.

It is clear that the bias depends not only on the chosen PDF but also on the estimation method, and this constitutes the subject of this study. Various estimation methods can be analysed and compared with respect to resistance to the false distribution assumption. The MOM, the method of maximum likelihood (MLM), the method of linear moments (LMM) and the LSM are considered in this study. For the sake of brevity, only the first three methods are illustrated by a numerical example. However, any other method, such as (weighted) least squares or entropy, can be easily included. One can also use the estimation methods as the approximation methods of one distribution function by another.

Since statistical moments of the PDF are of interest in both regional and non-stationary approaches to FFA, the analysis covers both the quantiles and the first two moments, i.e. $z = x_p$, α_1, μ_2, where p denotes the probability of exceedance. Obviously our interest is in quantiles of small probability of exceedance, i.e. $p < 10\%$.

To derive the bias, one distribution is considered here as the true one (T) while another hypothetical distribution as the false one (F). Among various criteria applied for the assessment of a distribution fit to the data, the L-ratio and its extension, AIC, are frequently used. Therefore, the difference in the values of the maximum of $\ln L$ function per one element of an infinite sample, denoted hereafter M_∞, between the case of the correct distribution choice ($H = T$) and that of the wrong distribution choice ($H = F$), i.e.

$$\Delta M_\infty(F|T) = M_\infty(T|T) - M_\infty(F|T) \quad (5)$$

where

$$M_\infty(T|T) = \lim_{N\to\infty} \frac{1}{N} \max \ln L(T|T) \quad (6)$$

and

$$M_\infty(F|T) = \lim_{N\to\infty} \frac{1}{N} \max \ln L(F|T) \quad (7)$$

is also of interest.

4. Set of PDFs

Only two-parameter PDFs are selected for analysis in this study. All estimation methods are used as approximation methods, except for MLM and entropy. Thus, there is no mathematical restriction to the selection of any PDF from a set of alternative PDFs (APDFS) of annual peak flows, except the overlapping domains of the selected distributions. To use MLM, the domains of T and F distribution functions must be either the same or the domain of F must cover the domain of T, e.g. $F =$ Gumbel I type and $T =$ Gamma.

It is convenient if APDFS consist of PDFs of the same range, which is assumed here as $(0, +\infty)$, i.e. the range of the distribution of annual peak flows is regarded here as a priori information. The choice is limited to PDFs with existing moments of all orders for the whole range of parameters. We acknowledge the existence of some pathological exceptions, which have recently appeared in the hydrological literature, such as the log-Gumbel or the log-logistic distribution; these distributions are, however, of more mathematical interest than of practical importance.

It is assumed that the parameters of any distribution of APDFS can be explicitly expressed by the distribution moments. For convenience, let APDFS consist of two-parameter distributions

$$\varphi_j(x;\theta^{(j)}), \qquad \theta^{(j)} = \left(\theta_1^{(j)}, \theta_2^{(j)}\right);\ j \in \text{APDFS} \tag{8}$$

where x is a random variable, $\varphi_j(x)$ is the jth distribution function of x with parameter set $\theta^{(j)}$. One can argue that the unknown true PDF contains too many parameters. To comply with it makes a bias dependent on the number of parameters and therefore results less transparent.

Let the set of APDFS consist, for simplicity, of the two two-parameter distributions defined in a semi-infinite domain $(0,\infty)$: the lognormal (LN)

$$\varphi(x;\mu,\sigma) = \frac{1}{\sigma x\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}(\ln x - \mu)^2\right] \tag{9}$$

and the gamma (Γ)

$$\varphi(x;\alpha,\lambda) = \frac{\alpha^\lambda}{\Gamma(\lambda)}x^{\lambda-1}\exp(-\alpha x) \tag{10}$$

where μ, σ, λ and α are parameters, and $\Gamma(\lambda)$ is the gamma function. The case ($H = \text{LN}$, $T = \Gamma$) and its application are presented in what follows, while the opposite case, i.e. ($H = \Gamma$, $T = \text{LN}$), constitutes the subject of Appendix A. It may be noted that to extend the results obtained in this study to three-parameter lognormal and gamma distributions, it would suffice to start with the matching of the lower bounds (or scale parameters) of both distributions.

5. Parameters replaced by moments

In order to unify the distributions with respect to parameters, the original set of parameters can be replaced by moments using the relationships available in standard statistical handbooks. Then, the density function is denoted as $\varphi_j(x;\alpha_1,\mu_2)$ where $\alpha_1 = f_1^{(j)}(\theta_1^{(j)},\theta_2^{(j)})$ and $\mu_2 = f_2^{(j)}(\theta_1^{(j)},\theta_2^{(j)})$ are, respectively, the first moment about the origin and the second moment about the centroid.

The first moments of the LN distribution are

$$\alpha_1 = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{11}$$

$$\begin{aligned}\mu_2 &= [\exp(\sigma^2) - 1]\exp(2\mu + \sigma^2)\\ &= [\exp(\sigma^2) - 1]\alpha_1^2\end{aligned} \tag{12}$$

$$c_v = \sqrt{\exp(\sigma^2) - 1} \tag{13}$$

$$c_s = 3c_v + c_v^3 \tag{14}$$

where c_v is the coefficient of variation, and c_s is the coefficient of skewness. Hence,

$$\sigma^2 = \ln\left[1 + c_v^2\right] \tag{15}$$

and

$$\mu = \ln\alpha_1 - 0.5\sigma^2 = \ln\frac{\alpha_1}{\sqrt{1 + c_v^2}} \tag{16}$$

where μ and σ^2 are the mean and the variance of $\ln X$.

For the gamma distribution, the first moments are

$$\alpha_1 = \frac{\lambda}{\alpha} \tag{17}$$

$$\mu_2 = \frac{\lambda}{\alpha^2} \tag{18}$$

and

$$c_v^2 = \frac{1}{\lambda} \tag{19}$$

Hence,

$$\alpha = \frac{\alpha_1}{\mu_2} = \frac{1}{\alpha_1 c_v^2} \tag{20}$$

$$\lambda = \frac{\alpha_1^2}{\mu_2} = \frac{1}{c_v^2} \tag{21}$$

The expression for quantile x_p of exceedance probability p is denoted as

$$x_p = g(p;\alpha_1,\mu_2) \tag{22}$$

which for the lognormal distribution takes the form

$$x_p = \exp\left(\mu + \sigma t_p^N\right) = \exp\left(-\sigma^2/2 + \sigma t_p^N\right)\alpha_1 \quad (23)$$

where t_p^N is the quantile of the order p of $N(0,1)$, σ is defined by Eq. (15), and p is probability of exceedence, while for the gamma distribution

$$x_p = \frac{t_p^\Gamma(\lambda)}{\alpha} = \frac{t_p^\Gamma(\lambda)}{\lambda}\alpha_1 \quad (24)$$

where t_p^Γ is the lower limit of the integral:

$$p = \frac{1}{\Gamma(\lambda)}\int_{t_p^\Gamma}^{\infty} t^{\lambda-1}\,\mathrm{e}^{-t}\,\mathrm{d}t \quad (25)$$

To use parameter estimation methods for fitting one distribution by another, an infinite sample is considered.

6. Maximum likelihood as approximation method

6.1. Finite sample size

For the ML estimation, it is usually more convenient to work with the log-likelihood function

$$\ln L = \sum_{i=1}^{N} \ln\, \phi_j(x_i; \alpha_1, \mu_2) \quad (26)$$

than with the likelihood function L. By differentiating $\ln L$ with respect to each of the parameters separately and equating to zero, one obtains as many equations as the number of parameters. Obviously, multiplication of Eq. (26) by any number does not affect the solution. Choosing $(1/N)$ as a multiplier

$$\Lambda = \frac{1}{N}\ln L = \frac{1}{N}\sum_{i=1}^{N} \ln\, \varphi_j(x_i; \alpha_1, \mu_2) \quad (27)$$

emphasizes the well-known fact that the ML-estimation method involves sample averages. It is in contrast with the POME method, which involves population expectations. Consequently, the sample average of the maximum likelihood function, M, is introduced

$$M = \max_{\alpha_1,\mu_2} \Lambda = \max_{\alpha_1,\mu_2}\left[\frac{1}{N}\ln L\right] \quad (28)$$

Therefore, one seeks a solution of

$$\frac{\partial \Lambda}{\partial \alpha_1^{(j)}} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\varphi_j\left(x_i; \alpha_1^{(j)}, \mu_2^{(j)}\right)} \frac{\partial \varphi_j\left(x_i; \alpha_1^{(j)}, \mu_2^{(j)}\right)}{\partial \alpha_1^{(j)}} = 0 \quad (29)$$

$$\frac{\partial \Lambda}{\partial \mu_2^{(j)}} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\varphi_j\left(x_i; \alpha_1^{(j)}, \mu_2^{(j)}\right)} \frac{\partial \varphi_j\left(x_i; \alpha_1^{(j)}, \mu_2^{(j)}\right)}{\partial \mu_2^{(j)}} = 0 \quad (30)$$

for which $(\Lambda)'' < 0$.

6.2. Infinite sample size

Going to the asymptotic case, i.e. when N tends to infinity, the two cases are distinguished:

1. The hypothesized distribution is the right choice $(H = T)$:

$$M_\infty(H = T) = \lim_{N\to\infty} M(H = T) \quad (31)$$

2. The hypothesized distribution is the false one ($H = F$) while the true PDF is known:

$$M_\infty(H = F|T) = \lim_{N\to\infty} M(H = F|T) \quad (32)$$

While the first case is the subject of interest in the classical statistical theory, the second case reflects the reality with respect to the hypothetical distribution, which serves as an approximation of the true unknown distribution. Every function of APDFS can stand either for the true or false distribution in our study. Doing so, we hope to be able to compare the robustness of various estimation methods with respect to the statistics of interest for the wrong choice of a distribution function.

6.3. Known distribution function (case 1): $H = T$

The hypothesis is $H = T = \varphi_k(x_i; \alpha_1, \mu_2)$, $k \in$ APDFS.

The asymptotic average of the $\ln L$ function (27) is

$$\Lambda_\infty = \lim_{N\to\infty} \frac{1}{N}\sum \ln\, \varphi_k = E[\ln\, \varphi_k] \quad (33)$$

The MLM-conditions then become:

$$\frac{\partial \Lambda_\infty(H=T)}{\partial \alpha_1} = \lim_{N\to\infty} \frac{1}{N} \sum \frac{\partial \ln \varphi_k}{\partial \alpha_1} = E\left[\frac{\partial \ln \varphi_k}{\partial \alpha_1}\right] = \int_0^\infty \frac{\partial \varphi_k}{\partial \alpha_1} \mathrm{d}x = 0 \quad (34)$$

$$\frac{\partial \Lambda_\infty(H=T)}{\partial \mu_2} = \lim_{N\to\infty} \frac{1}{N} \sum \frac{\partial \ln \varphi_k}{\partial \mu_2} = E\left[\frac{\partial \ln \varphi_k}{\partial \mu_2}\right] = \int_0^\infty \frac{\partial \varphi_k}{\partial \mu_2} = 0 \quad (35)$$

Since both the MOM and MLM estimates of parameters are asymptotically unbiased, there is convergence of MOM and MLM estimates of any statistics Z. Therefore, the solution of Eqs. (34) and (35) is for all k:

$$\alpha_1(x) = \int_0^{+\infty} x\varphi_k(x; \alpha_1, \mu_2)\mathrm{d}x \quad (36)$$

$$\mu_2(x) = \int_0^{+\infty} (x-\alpha_1)^2 \varphi_k(x; \alpha_1, \mu_2)\mathrm{d}x \quad (37)$$

and the asymptotic average of the ln L function (27) is

$$\Lambda_\infty(H=T) = \int_0^{+\infty} \varphi_k \log \varphi_k \, \mathrm{d}x \quad (38)$$

For illustration, consider the case of the gamma distribution. The MLM Eqs. (34) and (35) expressed in terms of the original parameters for this distribution take the form:

$$\frac{\partial \Lambda_\infty(H=T=\Gamma)}{\partial \alpha} = \frac{\lambda}{\alpha} - \int x\varphi^\Gamma \, \mathrm{d}x = 0 \quad (39)$$

$$\frac{\partial \Lambda_\infty(H=T=\Gamma)}{\partial \lambda} = \ln \alpha - \psi(\lambda) + \int \ln x\varphi^\Gamma \, \mathrm{d}x = 0 \quad (40)$$

or substituting logarithmically transformed Eq. (39) into Eq. (40), we get

$$\ln \lambda - \psi(\lambda) = \ln \int x\varphi^\Gamma \, \mathrm{d}x - \int \ln x\varphi^\Gamma \, \mathrm{d}x \quad (41)$$

where $\psi(\lambda) = \mathrm{d} \ln \Gamma(\lambda)/\mathrm{d}\lambda$ is the digamma function, while Eqs. (36) and (37) take forms, respectively,

$$\alpha_1 = \frac{\lambda}{\alpha} \quad (42)$$

$$\mu_2 = \frac{\lambda}{\alpha^2} = \frac{1}{\lambda}\alpha_1^2 \quad (43)$$

The asymptotic average ln L function (38) is

$$\Lambda_\infty(H=T=\Gamma) = \lambda \ln \alpha - \ln \Gamma(\lambda) + (\lambda - 1)\int_0^\infty \ln x\varphi^\Gamma(x)\mathrm{d}x - \alpha \int_0^\infty x\varphi^\Gamma(x)\mathrm{d}x \quad (44)$$

Substituting Eqs. (42) and (39) into Eq. (40) in Eq. (44), we get $M_\infty = \max \Lambda_\infty$

$$M_\infty(H=T=\Gamma) = -\lambda + \ln \lambda + (\lambda-1)\psi(\lambda) - \ln \Gamma(\lambda) - \ln \alpha_1^\Gamma \quad (45)$$

where λ is the function (21) of c_v^Γ.

6.4. Wrong choice of distribution function (case 2): $H \neq T$

The hypothesis to be dealt with is:

$$H = \varphi_j\left(x; \alpha_1^{(j)}, \mu_2^{(j)}\right) \text{ and } T = \varphi_k\left(x; \alpha_1^{(k)}, \mu_2^{(k)}\right)$$
$$j \neq k; \quad j,k \in \text{APDFS} \quad (46)$$

The asymptotic Λ value is

$$\Lambda_\infty(H=\varphi_j, T=\varphi_k) = \int_0^\infty \varphi_k \log \varphi_j \, \mathrm{d}x \quad (47)$$

with parameters derived from the MLM conditions:

$$\frac{\partial \Lambda_\infty(H \neq T)}{\partial \alpha_1^{(j)}} = E\left[\frac{\partial \ln \varphi_j}{\partial \alpha_1^{(j)}}\right] = \int_0^{+\infty} \frac{1}{\varphi_j} \frac{\partial \varphi_j}{\partial \alpha_1^{(j)}} \varphi_k \, \mathrm{d}x = 0 \quad (48)$$

$$\frac{\partial \Lambda_\infty(H \neq T)}{\partial \mu_2^{(j)}} = E\left[\frac{\partial \ln \varphi_j}{\partial \mu_2^{(j)}}\right] = \int_0^{+\infty} \frac{1}{\varphi_j} \frac{\partial \varphi_j}{\partial \mu_2^{(j)}} \varphi_k \, \mathrm{d}x = 0 \quad (49)$$

Solving them we get $\alpha_1^{(j)}$ and $\mu_2^{(j)}$ for given $\alpha_1^{(k)}$ and $\mu_2^{(k)}$:

$$\left(\alpha_1^{(k)}, \mu_2^{(k)}\right) \Rightarrow \left(\alpha_1^{(j)}, \mu_2^{(j)}\right) \tag{50}$$

Substituting both the exact value, i.e. true, and that got by the ML-approximation into Eq. (3) defines the asymptotic bias of statistic Z. For the moments, we have

$$B(\mu_1') = \alpha_1^{(j)} - \alpha_1^{(k)} \tag{51}$$

$$B(\mu_2) = \mu_2^{(j)} - \mu_2^{(k)} \tag{52}$$

Similarly, the asymptotic bias of quantile can be obtained from Eq. (3):

$$B(x_p) = x_p^{(j)} - x_p^{(k)} \tag{53}$$

In general, the roots of Eqs. (48) and (49) may differ from the moments of the true distribution. It happens when MLM and MOM are not equivalent for the assumed distribution (H). Note that if $H =$ Normal distribution then for any T-distribution

$$\alpha_1^{(j)} = \int x\varphi_k \, \mathrm{d}x = \alpha_1^{(k)} \tag{54}$$

and

$$\mu_2^{(j)} = \int \left(x - \alpha_1^{(k)}\right)^2 \varphi_k \, \mathrm{d}x = \mu_2^{(k)} \tag{55}$$

that is, MLM of the normal distribution produces unbiased estimators of the first two moments independently of the form of the true distribution and the asymptotic value of M, i.e. M_∞ does not depend on the true distribution form. The fact that for the normal distribution the method gives asymptotically unbiased estimators of the mean and standard deviation belongs to the fundamental statistical statements. In general, if the H-distribution belongs to the exponential type of distributions (Kendall and Stuart, 1973, vol. 2, pp. 12, 26, 67), then the MLM and MOM are equivalent. The Pearson distributions do not belong to this type except as approximations which are sufficient to preserve for $H =$ Gamma unbiased MLM-estimator of the first moment.

Let us exchange the places of the two distributions, assuming φ_j to be the true distribution and φ_k to be the false one. In general, the MLM as the approximation method is irreversible, i.e. the solution of the inverse problem to the one posed by Eqs. (48) and (49) for input data taken from the solution of Eqs. (48) and (49) differ from the input used in the previous problem, i.e. if

$$\left(\alpha_1^{(k)} = a, \mu_2^{(k)} = b\right) \Rightarrow \left(\alpha_1^{(j)} = c, \mu_2^{(j)} = d\right) \tag{56}$$

then

$$\left(\alpha_1^{(j)} = c, \mu_2^{(j)} = d\right) \not\Rightarrow \left(\alpha_1^{(k)} = a, \mu_2^{(k)} = b\right) \tag{57}$$

Therefore, in general,

$$B(z(H = \varphi_k | T = \varphi_j)) \neq -B(z(H = \varphi_j | T = \varphi_k)) \tag{58}$$

which is the distinguishing property of the MLM-approximation.

To illustrate, consider that the chosen distribution is lognormal while the real distribution is gamma ($F = \mathrm{LN}$, $T = \Gamma$). As in the previous example, it is more convenient to present derivations using the original parameters of both distributions, i.e. (μ, σ) and (λ, α), and convert them into moments than to operate directly on the moments given by Eqs. (48) and (49):

$$\frac{\partial \Lambda_\infty}{\partial \mu} = \frac{1}{\sigma^2}\left(\mu - \int_0^\infty \ln x \varphi^\Gamma \, \mathrm{d}x\right) = 0 \tag{59}$$

$$\frac{\partial \Lambda_\infty}{\partial \sigma} = \frac{1}{\sigma^3}\left[\sigma^2 - \int_0^\infty (\ln x - \mu)^2 \varphi^\Gamma \, \mathrm{d}x\right] = 0 \tag{60}$$

It is clearly seen from Eqs. (59) and (60) as well as from Eq. (A8) in Appendix A that the approximate distribution must have the same range as the true distribution or its domain must cover one of the true distribution.

Substitution of $\int_0^\infty \ln x \varphi^\Gamma \, \mathrm{d}x$ from Eq. (40) for the gamma distribution into Eq. (59), gives

$$\mu = \psi(\lambda) - \ln \alpha \tag{61}$$

The variance of the logarithm of the gamma-distributed variable equals (Strupczewski, 1999):

$$\int_0^\infty (\ln x - \mu)^2 \varphi^\Gamma \, \mathrm{d}x = \psi'(\lambda) \tag{62}$$

Therefore,

$$\sigma^2 = \psi'(\lambda) \tag{63}$$

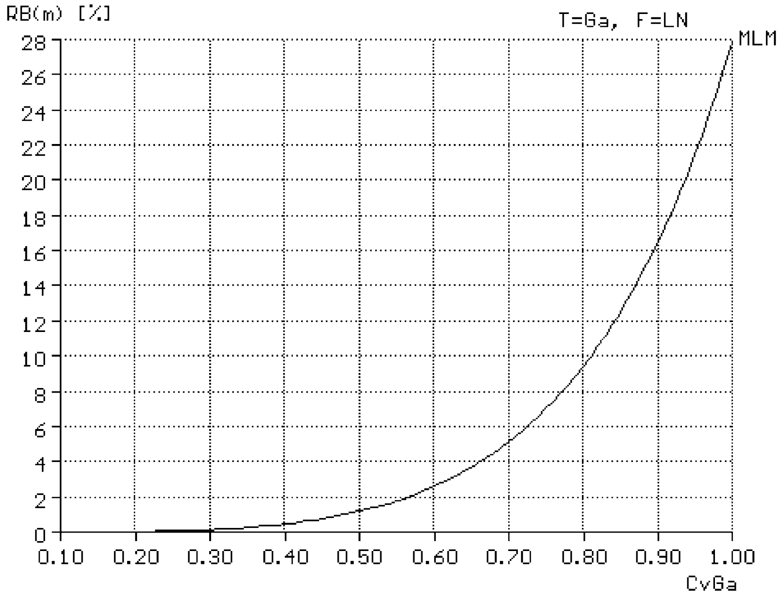

Fig. 1. Asymptotic relative bias of ML-estimator of mean if lognormal distribution is applied to gamma distributed data.

and

$$M_\infty(H = \mathrm{LN}, T = \mathrm{Gamma}) = -\ln\sqrt{2\pi} - \frac{1}{2} - \frac{1}{2}\ln\psi'(\lambda) - \psi(\lambda) + \ln\lambda - \ln\alpha_1^\Gamma \tag{64}$$

where λ is a function of c_v^Γ given by Eq. (21).

6.4.1. Bias of mean

Substituting Eqs. (61) and (63) into Eq. (11), we get

$$\alpha_1^{\mathrm{LN}} = \frac{1}{\alpha}\exp\left[\psi(\lambda) + \frac{\psi'(\lambda)}{2}\right] = \frac{1}{\lambda}\exp\left[\psi(\lambda) + \frac{\psi'(\lambda)}{2}\right]\alpha_1^\Gamma \tag{65}$$

Therefore, the asymptotic bias of the mean is

$$B(\alpha_1) = \alpha_1^{\mathrm{LN}} - \alpha_1^\Gamma = \left\{\frac{1}{\lambda}\exp\left[\psi(\lambda) + \frac{\psi'(\lambda)}{2}\right] - 1\right\}\alpha_1^\Gamma \tag{66}$$

where λ is the function (21) of c_v^Γ. Its relative value is displayed in Fig. 1 as the function of c_v^Γ. The relative bias grows with increasing value of the coefficient of variation, approaching zero for c_v tending to zero and being over 10% for c_v greater than 0.82. In the opposite case (Appendix A), i.e. when ($H = \Gamma$; $T = \mathrm{LN}$), the mean is unbiased (Eq. (A7)).

6.4.2. Bias of variance

Substituting Eqs. (61), (63) and (43) into Eq. (12), we obtain

$$\mu_2^{\mathrm{LN}} = \frac{1}{\lambda}[\exp(\psi'(\lambda)) - 1]\exp[2\psi(\lambda) + \psi'(\lambda)]\mu_2^\Gamma \tag{67}$$

where λ is a function of the coefficient of variation c_v^Γ defined by Eq. (21).

Therefore, the asymptotic bias of the variance equals

$$B(\mu_2) = \mu_2^{\mathrm{LN}} - \mu_2^\Gamma = \left\{\frac{1}{\lambda}[\exp(\psi'(\lambda)) - 1]\exp[2\psi(\lambda) + \psi'(\lambda)] - 1\right\}\mu_2^\Gamma \tag{68}$$

Its relative value is shown in Fig. 2. The relative

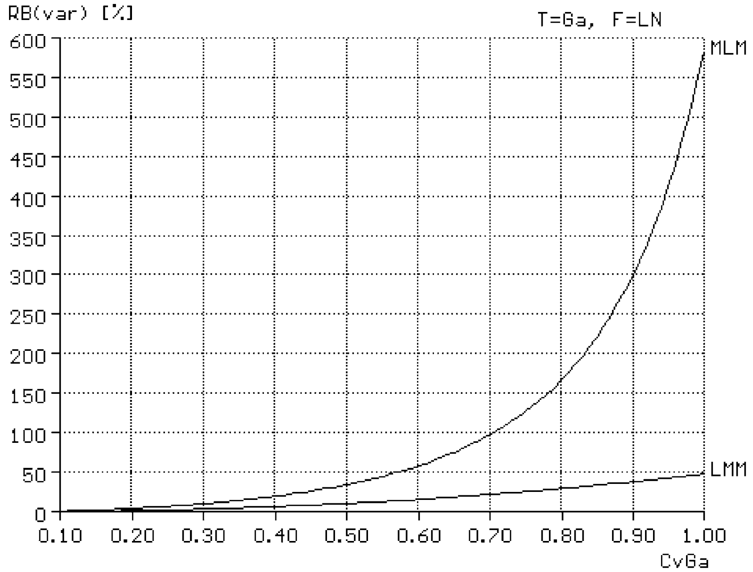

Fig. 2. Asymptotic relative bias of ML-estimator of variance if lognormal distribution is applied to gamma distributed data.

bias is an increasing function of the coefficient of variation c_v^Γ and it reaches over 100% for $c_v^\Gamma > 0.7$. It is approximately two order higher than one for the mean. The relation of the opposite case, i.e. ($H = \Gamma$; $T = \mathrm{LN}$), is given by Eq. (A14) in Appendix A and displayed in Fig. 3. The bias is of the opposite sign, i.e. negative, and its value is one order lower than the previously determined value, which reflects the fact that the MLM estimate of the mean for the gamma distribution remains unbiased irrespective of what the true distribution function is. For c_v tending to zero, the relative bias also approaches zero, as both distributions tend to normal distribution.

6.4.3. Bias of quantiles

Substituting Eqs. (23) and (24) into Eq. (53), we get

$$B(x_p) = x_p^{\mathrm{LN}} - x_p^\Gamma = \exp\left(\mu + \sigma t_p^N\right) - \frac{t_p^\Gamma(\lambda)}{\lambda}\alpha_1^\Gamma \quad (69)$$

Expressing the parameters of the LN distribution by those of the gamma distribution using Eqs. (61) and (63), the result is

$$B(x_p) = \frac{1}{\lambda}\left\{\exp\left[\psi(\lambda) + \sqrt{\psi'(\lambda)}t_p^N\right] - t_p^\Gamma(\lambda)\right\}\alpha_1^\Gamma \quad (70)$$

and the relative bias

$$\mathrm{RB}(x_p) = \frac{\exp\left[\psi(\lambda) + \sqrt{\psi'(\lambda)}t_p^N\right]}{t_p^\Gamma(\lambda)} - 1 \quad (71)$$

where λ is the reciprocal of the squared variation coefficient given by Eq. (21).

Therefore, the RB of the quantile is a function of the coefficient of variation of the gamma-distributed variable (c_v^Γ) and the quantile order (p) as shown in Fig. 4(a) and (b). Its value grows rapidly with increasing value of the coefficient of variation and with the probability of exceedance, and it exceeds 100% for $p < 2\%$ and $c_v = 1$. The relation for the opposite case relation, i.e. ($H = \Gamma$; $T = \mathrm{LN}$), is given by Eq. (A13) and displayed in Fig. 5(a) and (b). For $c_v = 0.2$, the MLM-quantile bias curve is almost a mirror reflection of that from Fig. 4(a) being a dozen percent smaller in the relative absolute values, while for $c_v = 1.0$ the growth of the MLM-quantile bias with c_v is in negative values not so rapid as all the values of RB lie in the (-45%, -18%) interval for $p < 2\%$. The differences between the MLM-quantile in Figs. 4(a) and (b), and 5(a) and (b) reflect the fact that the MLM estimate of the mean for the gamma distribution remains unbiased irrespective of what the true

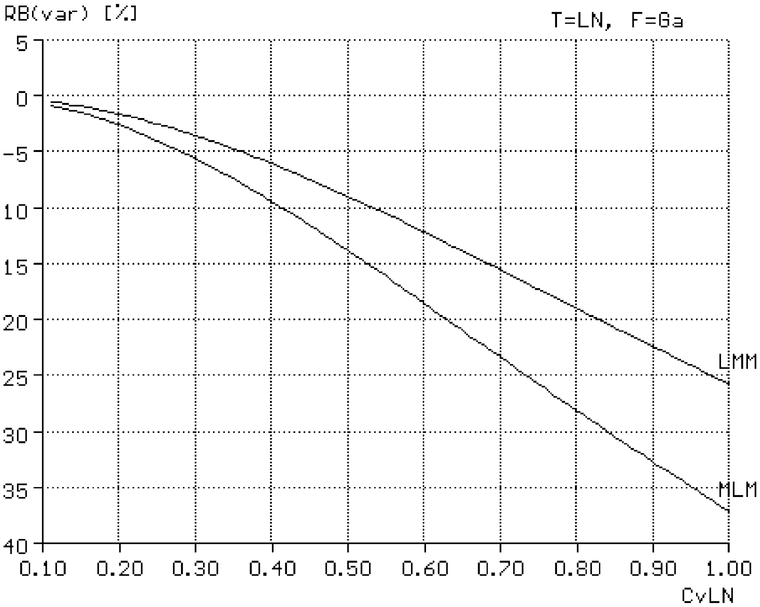

Fig. 3. Asymptotic relative bias of ML and LM methods' estimator of variance if gamma distribution is applied to lognormal distributed data.

distribution function is. That is, the MLM-estimate of the upper tail quantiles is more resistant with respect to the wrong choice of the distribution function if the gamma distribution is taken as the hypothetical distribution than in the opposite case. Fortunately, there are some other PDFs having this convenient property of equivalency of MLM and MOM estimators in respect to the mean. One of them is described by Strupczewski et al. (2001).

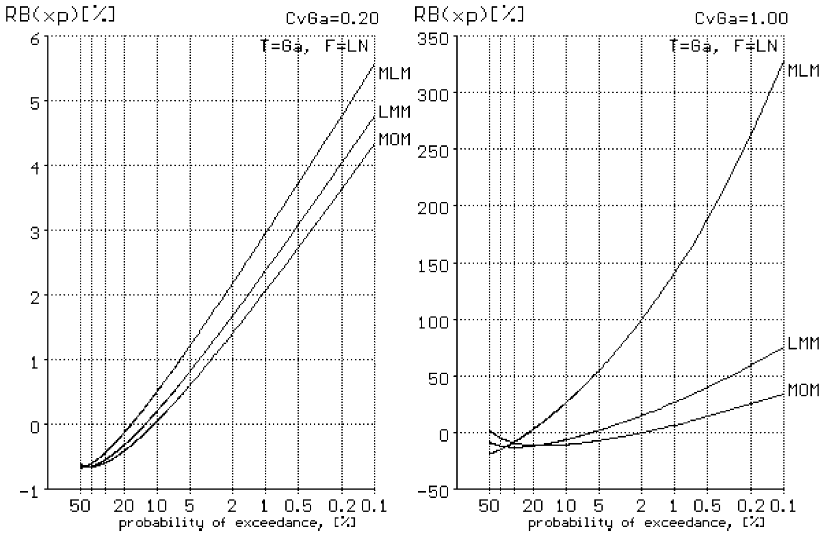

Fig. 4. Asymptotic relative bias of MLM, LMM and MOM estimates of quantiles if lognormal distribution is applied to gamma distributed data versus probability of exceedance: (a) for coefficient of variation $c_{vGa} = 0.2$; (b) for coefficient of variation $c_{vGa} = 1.0$.

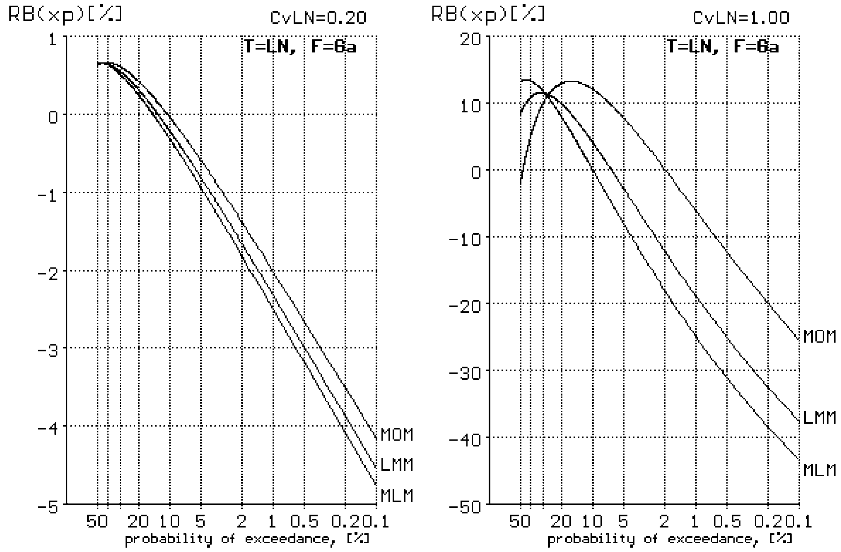

Fig. 5. Asymptotic relative bias of MLM, LMM and MOM estimates of quantiles if gamma distribution is applied to lognormal distributed data versus probability of exceedance: (a) for coefficient of variation $c_{vLN} = 0.2$; (b) for coefficient of variation $c_{vLN} = 1.0$.

6.4.4. *Influence of wrong distribution assumption on the value of the M_∞ function*

Since the ML-ratio test is frequently used for choosing the best flood frequency model, it is interesting to examine the sensitivity of the maximum likelihood function, M_∞, to the wrong model choice as outlined by Eqs. (5)–(7). Substituting Eqs. (45) and (64) into Eq. (7), we get

$$\Delta M_\infty(H = \mathrm{LN}; T = \Gamma)$$

$$= \lambda[\psi(\lambda) - 1] - \ln \Gamma(\lambda) + \frac{1}{2}[\ln \psi'(\lambda) + \ln(2\pi) + 1] \tag{72}$$

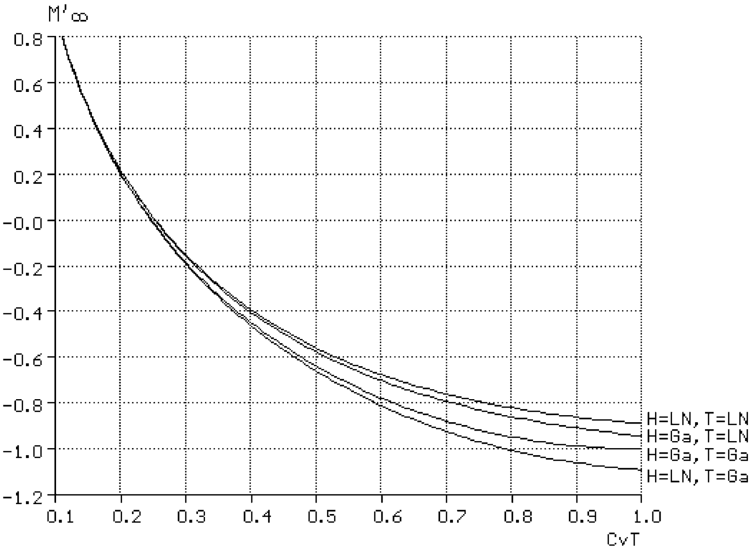

Fig. 6. Asymptotic value of ML as a function of the coefficient of variation for various combinations of (*True*; *Hypothetical*) distributions.

where λ is a function of the coefficient of variation c_v^Γ defined by Eq. (21).

To get rid of the first moment about the origin (α_1) in Eqs. (45) and (64), it is convenient to compare the functions $M'_\infty = [M_\infty(\cdot;\cdot) + \ln \alpha_1]$. Their graphs are shown in Fig. 6 both for $T = \Gamma$ and $T = \mathrm{LN}$, while the respective Eqs. (A4) and (A13) are given in Appendix A. It should be remarked that for small c_v, RB of quantiles is small and the difference between asymptotic ML values is small, which makes the ML ratio weak for identifying the true PDF. In other words, the standard deviation of each ML statistics will be greater than the ML difference even for large samples.

7. Method of moments

The approximation by the MOM is widely used in hydrological linear system theory, where two impulse response functions of known form are fitted by the so-called moment matching technique. The system response for a polynomial signal is usually taken as justification and for evaluation of the accuracy of the approximation.

Distributions, which have finite number of lower moments in common, are, in a sense, approximations of one another. Some mathematical support for this so-called principle of moments is given by Kendall and Stuart (1969, vol. 1, Section 3.34, p. 87). They approximate both functions by the finite series of powers and use the principle of least squares to determine their coefficients. If two distributions have moments up to order s equal, they must have the same least-squares approximation of the first s coefficient of the polynomial expansion.

As both distributions, i.e. *True* and *False*, are in our case two-parameter distributions, one can match their first two moments:

$$\int x\varphi_k(x)\mathrm{d}x = \int x\varphi_j(x)\mathrm{d}x = \alpha_1 \tag{73}$$

$$\int x^2\varphi_k(x)\mathrm{d}x = \int x^2\varphi_j(x)\mathrm{d}x = \alpha_2 \tag{74}$$

Therefore, the moments of higher than the second order may be biased only. There are no constraints with respect to the range of both distributions (Kendall and Stuart, 1969, vol. 1, Section 3.34, p. 87) but the overlapping range enabling fit of the first moment. The asymptotic bias of any statistics Z caused by the false choice of the distribution, i.e. φ_j taken instead of φ_k, determines the asymptotic bias of the opposite case, i.e. taking φ_k instead of φ_j :

$$B(z(H = \varphi_k|T = \varphi_j)) = -B(z(H = \varphi_j|T = \varphi_k)) \tag{75}$$

To illustrate, consider an example. Substituting Eqs. (17) and (11) into Eq. (73), we get

$$\frac{\lambda}{\alpha} = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{76}$$

and matching the coefficients of variation given by Eqs. (19) and (13)

$$\frac{1}{\lambda} = \exp(\sigma^2) - 1 \tag{77}$$

Obviously, the first two moments are asymptotically unbiased. Note that, if the indirect MOM for approximation of Γ by LN is applied, moments of all orders of the original variable X will be biased but Eq. (75) still holds. In any case, the bias should be related to the true value of moments, i.e. to α_1^Γ and μ_2^Γ in this case.

The asymptotic bias of the quantile approximation (53) and its relative value can be determined by Eq. (69) with all parameters derived by MOM (note the difference from the MLM approach: μ and σ are derived from Eqs. (15) and (16) but not from Eqs. (61) and (63)). After substituting Eqs. (15), (16) and (21) into it, we get the asymptotic bias (4)

$$\mathrm{RB}(x_p) = \frac{\exp\left[t_p^N\sqrt{\ln\left[1 + c_v^2\right]}\right]}{c_v^2\sqrt{1 + c_v^2}\,t_p^\Gamma\left(c_v^{-2}\right)} - 1 \tag{78}$$

Therefore, the relative bias of the quantile approximated by MOM is a function of the coefficient of variation c_v and the probability of exceedance p as in the case of the MLM approximation. It is displayed in Fig. 4(a) and (b).One can see that bias arising from the ML-method and caused by the wrong distribution assumption is for $p < 15\%$ greater than one due to the MOM. The difference between them grows rapidly with the increasing value of the coefficient of variation and decreasing probability of exceedance, exceeding 300% for $c_v = 1$ and $p = 0.1\%$.

The RB caused by taking the gamma distribution function instead of lognormal and MOM as the approximation method is given by Eq. (A15) and shown in Fig. 5(a) and (b). It is a function of the coefficient of variation c_v^{LN} and the probability of exceedance p. It has a smaller absolute value than either the one of the ML-method for p less than a few percent or the one of the previous case (Fig. 4(a) and (b)). The difference, $RB^{MLM}(x_p; c_v) - RB^{MOM}(x_p; c_v)$, reaches 20% for $c_v = 1$ and $p = 0.1\%$. That is, having a choice between the lognormal and the gamma, the latter should be selected if the MLM is to be applied.

8. Method of L-moments

As in MOM, the matching of the lower L-moments of both distributions permits approximation of one distribution by another. Let, once again, $H = \varphi_j(x; \alpha_1^{(j)}, \mu_2^{(j)})$ and $T = \varphi_k(x; \alpha_1^{(k)}, \mu_2^{(k)})$, $j \neq k$; $j, k \in$ APDFS. The first two L-moments of the true distribution expressed in terms of the moments are:

$$\lambda_1 = \alpha_1^{(k)} \tag{79}$$

$$\lambda_2 = \alpha_1^{(k)} \phi_k\left(c_v^{(k)}\right) \tag{80}$$

Similar relations can be written for the H distribution:

$$\lambda_1 = \alpha_1^{(j)} \tag{81}$$

$$\lambda_2 = \alpha_1^{(j)} \phi_j\left(c_v^{(j)}\right) \tag{82}$$

Matching the L-moments yields

$$\alpha_1^{(j)} = \alpha_1^{(k)} \tag{83}$$

$$c_v^{(j)} = \phi_j^{-1}\left[\phi_k\left(c_v^{(k)}\right)\right] \tag{84}$$

Therefore, the mean is unbiased, while the variance and quantiles are biased.

Furthermore, as in MOM, we have $(\alpha_1^{(k)} = a, \mu_2^{(k)} = b) \Leftrightarrow (\alpha_1^{(j)} = c, \mu_2^{(j)} = d)$, where $a = c$ for the L-moment matching, and Eq. (75) holds. It is noted that the bias should be related to the true values of moments.

Consider an example. As before, the chosen distribution is lognormal while the real distribution is gamma ($F = \text{LN}, T = \Gamma$). The first L-moment is identical with the first moment about the origin. Therefore, the match of the first L-moment is equivalent to the match of the first moment about the origin. Hence, the first moment α_1^{LN} obtained from the L-moment technique remains unbiased. Substituting Eqs. (11) and (17) into Eq. (73), we obtain Eq. (76). Doing so, with the L-coefficient of variation

$$\tau^{LN} = \tau^{\Gamma} \tag{85}$$

where (e.g. Hosking and Wallis, 1997)

$$\tau^{LN} = 2\left[\frac{1}{2} - p\left(\frac{\sigma}{\sqrt{2}}\right)\right] \tag{86}$$

where

$$p(w) = \int_w^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz \tag{87}$$

is the probability of exceedance for $N(0, 1)$, and

$$\tau^{\Gamma} = \frac{\Gamma(\lambda + 0.5)}{\sqrt{\pi}\,\Gamma(\lambda + 1)} \tag{88}$$

we get the relation between λ and σ, which differs from both the ML-relation (63) and the MOM-relation (77). Expressing λ and σ by c_v^{Γ} and c_v^{LN}, respectively, we get the relationship between the coefficients of variation of both distributions arising from the L-moment method:

$$\frac{\Gamma\left(\left(c_v^{\Gamma}\right)^{-2} + 0.5\right)}{\sqrt{\pi}\,\Gamma\left(\left(c_v^{\Gamma}\right)^{-2} + 1\right)} = 2\left[0.5 - p\left(\frac{\sqrt{\ln\left[1 + (c_v^{LN})^2\right]}}{\sqrt{2}}\right)\right] \tag{89}$$

i.e. $c_v^{\Gamma} \Leftrightarrow c_v^{LN}$, which needs numerical solution.

To find the RB of variance, it is convenient to reformulate Eq. (52) as:

$$B(\mu_2) = \left[\left(c_v^{LN}\right)^2 - \left(c_v^{\Gamma}\right)^2\right]\alpha_1^2 \tag{90}$$

Then, the asymptotic relative bias (4) takes the form:

$$RB(\mu_2) = \left(\frac{c_v^{LN}}{c_v^{\Gamma}}\right)^2 - 1 \tag{91}$$

The value of RB versus the c_v^{Γ} together with the respective relation for MLM-approximation is displayed in Fig. 2. RB(μ_2) for LMM increases slowly and near linearly with c_v^Γ, reaching about 50% for $c_v^\Gamma = 1$, while for MLM its increase is rapid, reaching almost 600% for $c_v^\Gamma = 1$. Fig. 3 shows the relation for the case ($H = \Gamma$; $T = \text{LN}$). Now both the biases are negative and the difference between them is much smaller, slightly exceeding 10% for $c_v^\Gamma = 1$.

The bias of quantiles is defined by Eq. (69), which, taking into account the equality of the first moments, $\alpha_1^{\text{LN}} = \alpha_1^\Gamma$, takes the form:

$$B(x_p) = x_p^{\text{LN}} - x_p^\Gamma = \left[\exp\left(\sigma t_p^N - 0.5\sigma^2\right) - \frac{t_p^\Gamma(\lambda)}{\lambda}\right]\alpha_1 \tag{92}$$

Introducing the direct moments of both distributions into Eq. (92), one gets the RB as

$$\text{RB}(x_p) = \frac{\exp\left[t_p^N\sqrt{\ln\left(1 + (c_v^{\text{LN}})^2\right)}\right]}{t_p^\Gamma\left((c_v^\Gamma)^{-2}\right)(c_v^\Gamma)^2\sqrt{1 + (c_v^{\text{LN}})^2}} - 1 \tag{93}$$

In Eqs. (90), (91) and (93), the coefficient of variation of the lognormally distributed variable, c_v^{LN}, is to be computed from Eq. (89) for given c_v^Γ. Therefore, the RB of the quantile is a function of the coefficient of variation of the Γ-distributed variable, c_v^Γ, and the quantile order p. The result is displayed in Fig. 4(a) and (b) together with the results of MLM and MOM approximations. For $c_v^\Gamma = 0.2$, all three RB(x_p) functions are close to each other (the maximum difference is about 1.3%) with LMM and MOM being closer, while for $c_v^\Gamma = 1.0$ the bias for MLM grows rapidly being for $p = 0.1\%$ one order higher than the other two which are still close.

For the case ($H = \Gamma$; $T = \text{LN}$) (Fig. 5(a) and (b)), this image is in a way mirrored for $c_v^\Gamma = 0.2$: all three bias functions run close to each other. However, the LMM bias is now close to that for MLM. Unlike Fig. 4, the pattern from Fig. 5(a) occurs also in Fig. 5(b) (for $c_v^\Gamma = 1.0$). The closeness of all three biases is not so tight as in Fig. 5(a) but they all increase (in absolute values) in a very similar way.

9. Least squares approximation

The approximation of the T-distribution by the H-distribution, i.e. the determination of $\alpha_1^{(j)}$ and $\mu_2^{(j)}$ from $\alpha_1^{(k)}$ and $\mu_2^{(k)}$, can be made at least in three different ways:

1. *In the variable domain*. Estimation of parameters from a sample of size N is done by minimizing the sum of squares of differences between the estimated and observed values:

$$\min_{\alpha_1,\mu_2} \frac{1}{N}\sum_{m=1}^{N}[\hat{x}(p_m) - x(p_m)]^2 \tag{94}$$

where $x(p_m)$ is the mth largest value in the sample, p_m is its exceedance probability, and $\hat{x}(p_m)$ is the estimated value. The continuous counterpart of Eq. (94) is

$$\min_{\alpha_1^{(j)},\mu_2^{(j)}} \int_{p=0}^{1}\left[x\left(p;\alpha_1^{(j)},\mu_2^{(j)}\right) - x\left(p;\alpha_1^{(k)},\mu_2^{(k)}\right)\right]^2\text{d}p \tag{95}$$

2. *In the probability domain*.

$$\min_{\alpha_1,\mu_2} \frac{1}{N}\sum_{m=1}^{N}[\hat{p}(x_m) - p(x_m)]^2 \tag{96}$$

where $\hat{p}(x_m)$ and $p(x_m)$ are the theoretical and empirical probabilities of the mth largest value, respectively. The continuous counterpart of Eq. (96) is

$$\min_{\alpha_1^{(j)},\mu_2^{(j)}} \int_{x=0}^{+\infty}\left[p_j\left(x;\alpha_1^{(j)},\mu_2^{(j)}\right) - p_k\left(x;\alpha_1^{(k)},\mu_2^{(k)}\right)\right]^2\text{d}x \tag{97}$$

3. *In the density domain*.

$$\min_{\alpha_1^{(j)},\mu_2^{(j)}} \int_{0}^{+\infty}\left[\varphi_j\left(x;\alpha_1^{(j)},\mu_2^{(j)}\right) - \varphi_k\left(x;\alpha_1^{(k)},\mu_2^{(k)}\right)\right]^2\text{d}x \tag{98}$$

which, in fact, leads to MOM (Kendall and Stuart, 1969, vol. 1, Section 3.34, p. 87).

"It is known that a function which is continuous in a

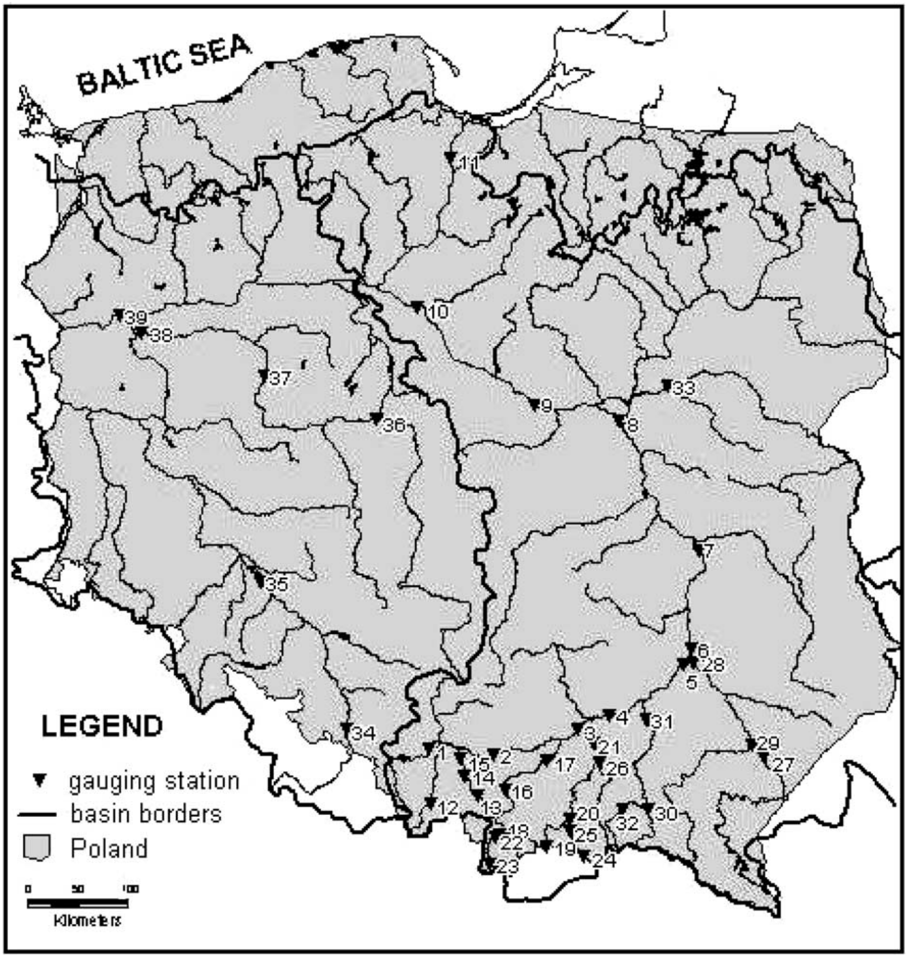

Fig. 7. Location of hydrological stations.

finite range a–b can be represented in that range by a uniformly convergent series of polynomials in x, say $\sum_{n=0}^{\infty} P_n(x)$ where $P_n(x)$ is of degree n. Suppose we wish to represent such a function approximately by a finite series of powers:

$$\varphi(x) \approx \sum_{n=0}^{s} a_n x^n \tag{99}$$

The coefficients a_n may be determined by the principle of least squares, which makes

$$\int_a^b (\varphi - \sum a_n x^n)^2 \, dx \tag{100}$$

a minimum. Differentiating by a_i, we have

$$2\int_a^b (\varphi - \sum a_n x^n) x^i \, dx = 0$$

or

$$\int_a^b \varphi x^i \, dx = \alpha_i = \int_a^b \sum a_n x^{n+i} \, dx \tag{101}$$

If two distributions, φ_j and φ_k, have equal moments of up to order s, then they must have the same least squares approximation, for coefficients a_n are determined by the moments by virtue of Eq. (101). A similar line of approach may be adopted when the range is infinite, the distributions in such cases being, under certain general conditions,

Table 3
Pertinent characteristics of annual flood data of 39 gauging stations in Poland

Basin/river	Gauging station		Drainage area (10^3 km^2)	Average peak flow (m^3/s)	Coefficient of variation (c_v)	Skewness coefficient (c_s)	Winter floods contribution (%)
	No.	Name					
Vistula	1	Jawiszowice	0.971	149.4	0.637	1.258	25.7
	2	Tyniec	7.520	719.1	0.600	1.573	40.0
	3	Jagodniki	12.060	1126.4	0.551	1.298	44.3
	4	Szczucin	23.900	1906.0	0.589	1.214	44.3
	5	Sandomierz	31.850	2485.4	0.533	0.827	51.4
	6	Zawichost	50.730	3281.4	0.460	0.882	52.9
	7	Pulawy	57.260	3003.6	0.436	0.778	57.1
	8	Warsaw	84.540	2998.0	0.394	0.659	62.9
	9	Kepa	169.000	3937.0	0.355	0.796	71.4
	10	Torun	181.000	3916.7	0.368	1.094	74.3
	11	Tczew	194.400	3962.6	0.404	1.294	77.1
Vistula/Sola	12	Zywiec	0.785	301.4	0.729	1.819	25.7
Vistula/Skawa	13	Sucha	0.468	153.5	0.795	1.533	30.0
	14	Wadowice	0.835	256.8	0.730	1.286	32.9
Vistula/Skawa/Wieprzowka	15	Rudze	0.154	53.0	0.762	0.816	45.7
Vistula/Raba	16	Stroza	0.644	219.0	0.787	1.373	31.4
	17	Proszowki	1.470	459.8	0.739	1.135	32.9
Vistula/Dunajec	18	Kowaniec	0.681	250.9	0.748	2.388	42.9
	19	Kroscienko	1.580	458.7	0.793	2.223	20.0
	20	Nowy Sacz	4.340	933.6	0.750	1.498	31.4
	21	Zabno	6.740	1161.0	0.734	1.537	32.9
Vistula/Dunajec/Czarny Dunajec	22	Nowy Targ	0.432	172.2	0.837	2.131	32.9
Vistula/Dunajec/Bialy Dunajec	23	Zakopane	0.058	37.9	0.885	2.213	27.1
Vistula/Dunajec/Poprad	24	Muszyna	1.510	228.1	0.779	2.409	55.7
	25	Stary Sacz	2.070	319.0	0.655	1.729	48.6
Vistula/Dunajec/Biala	26	Koszyce W.	0.957	267.4	0.724	1.210	40.0
Vistula/San	27	Jaroslaw	7.040	794.1	0.580	0.982	64.3
	28	Radomysl	16.800	985.2	0.480	2.174	65.7
Vistula/San/Wislok	29	Tryncza	3.520	240.4	0.694	3.958	68.6
Vistula/Wisloka	30	Zolkow	0.581	167.2	0.812	1.953	44.3
	31	Mielec	3.690	545.0	0.565	2.156	52.9
Vistula/Wisloka/Ropa	32	Kleczany	0.482	114.7	0.715	1.358	32.9
Vistula/Bug	33	Wyszkow	39.100	667.3	0.563	1.721	95.7
Oder	34	Miedonia	6.74	557.8	0.512	1.374	25.0
	35	Trestno	20.40	1341.9	0.766	1.781	45.6
Oder/Warta	36	Konin	13.400	255.4	0.626	2.417	84.3
	37	Poznan	25.900	378.4	0.645	2.221	88.6
	38	Skwierzyna	32.100	416.1	0.597	1.626	94.3
	39	Gorzow	52.400	545.8	0.475	1.360	94.3

capable of representation by a series of terms such as $e^{-x^2} P_n(x)$. The same conclusion is reached." Therefore, the solution of Eq. (98) is

$$\alpha_1^{(j)} = \alpha_1^{(k)} \tag{102}$$

and

$$\mu_2^{(j)} = \mu_2^{(k)} \tag{103}$$

10. Empirical testing

10.1. Annual flood data

Thirty-nine uninterrupted annual peak flow series from Polish territory covering the period 1921–1990 were analysed. They are from drainage basins ranging in area from 100 to 194,000 km^2, as shown in Fig. 7. A

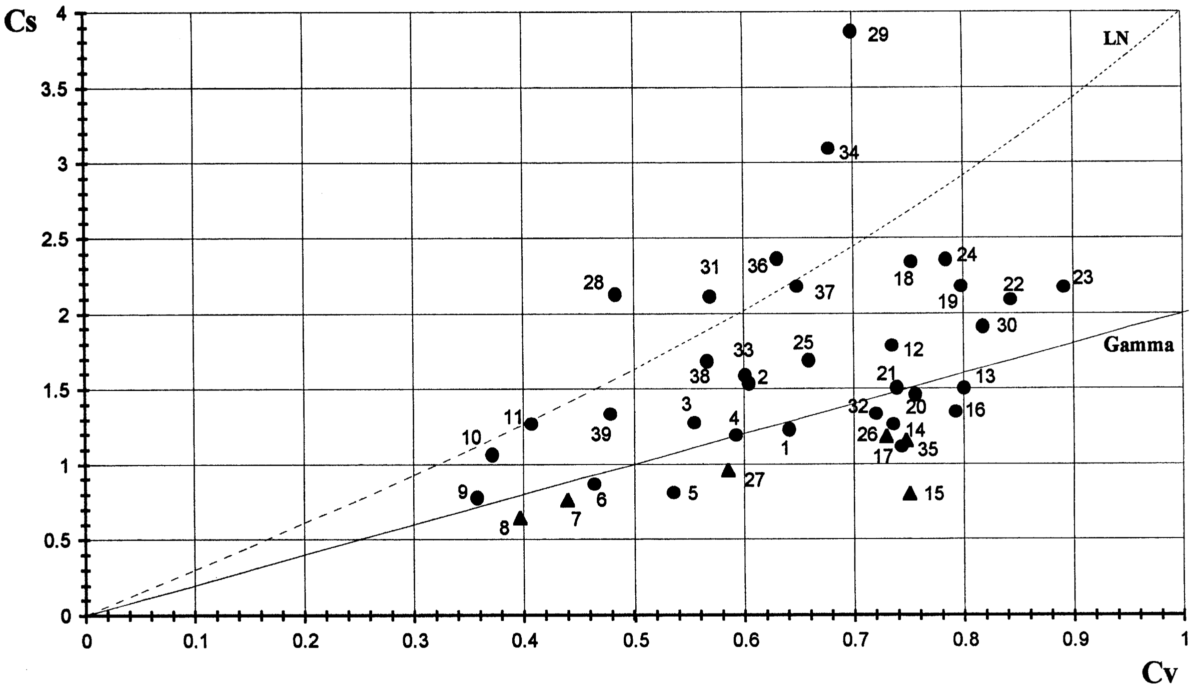

Fig. 8. Relation of sampling variation and skewness coefficients.

majority of the basins are from the south of Poland, which is a mountainous part of the country. Some pertinent characteristics of the data are given in Table 3 and Fig. 8. These data were selected on the basis of the length, completeness and homogeneity of records. Each of the 39 data sets was non-parametrically tested (Mitosek and Strupczewski, 1996) for homogeneity and independence basing on WMO (1988) guidelines. For stationarity of the mean and variance, the Mann test was used (Mann, 1945), while for detection of abrupt changes Lombard's test (Lombard, 1988) and Pettitt's test (Pettitt, 1979) were applied. Independence of elements in the series was tested by the runs test (Fisz, 1963). In each case the sample was found homogenous and independent at 5% significant level.

10.2. Fitting of PDFs

Seven two-parameter distribution functions, namely, normal, lognormal, gamma, Gumbel (extreme value type I), Weibull, log-Gumbel and log-logistic, were fitted by the ML method to the data. The criterion of the maximum log-likelihood value was used for choosing the best model choice. From the above competing models, lognormal was selected in 32 cases out of 39, gamma in six cases, Gumbel in one case, and the remaining four were not identified as the best model even in one case.

Finally, it was decided to form a set of APDFS by lognormal and gamma distributions. In Fig. 8 and in all subsequent figures, the circles and triangles mark the sample where LN and gamma were selected as the best model, respectively.

10.3. Relative difference of moments

Since neither the true PDF nor the true value of moments, i.e. values of the general population, are known, the MOM estimate of moments were regarded as the best approximation of the moments. Taking additionally into account a limited size of the random samples (see Eq. (1)), we will use the term 'relative difference' for the empirical values instead of 'asymptotic relative bias' giving credit

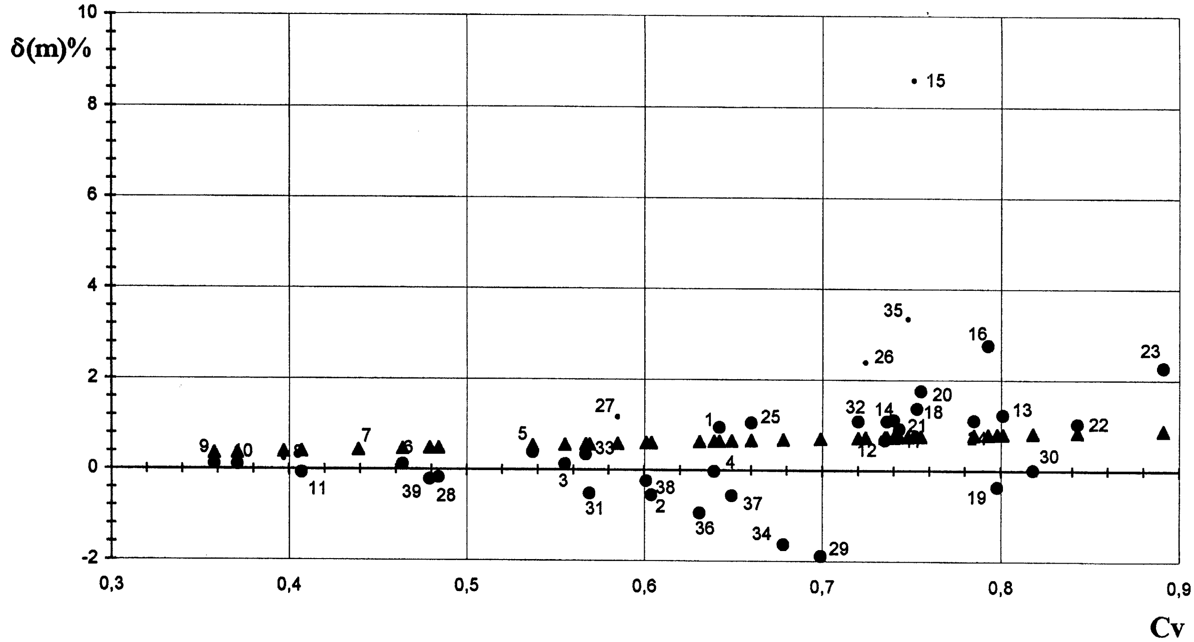

Fig. 9. Relative difference of the ML-mean (Eq. (104)) assuming LN distribution vs. MOM-coefficient of variation.

to the 70-year long sample as big enough not to mask the asymptotic properties of the bias. Therefore,

$$\delta(m) = \frac{m^{(\mathrm{MLM})} - m^{(\mathrm{MOM})}}{m^{(\mathrm{MOM})}} \tag{104}$$

and

$$\delta(\mathrm{var}) = \frac{\mathrm{var}^{(\mathrm{MLM})} - \mathrm{var}^{(\mathrm{MOM})}}{\mathrm{var}^{(\mathrm{MOM})}} \tag{105}$$

correspond to Eq. (4) for the mean and variance, respectively. The scattering points in Fig. 9 of the relative differences of the mean show a strong similarity to Fig. 1 with respect to either the sign of the differences, their range of variability or the relation to the coefficient of variation (c_v). A similar correspondence is observed with respect to the variance, i.e. Fig. 10 compared with Figs. 2 and 11 with Fig. 3. The magnitudes of differences δ(var) of Figs. 10 and 11 differ considerably; this finding is in conformity with the asymptotic differences RB(var) shown in Figs. 2 and 3. Furthermore, a growing tendency with the coefficient of variation value is observed both for lognormal and gamma distributions, which correspond to the increasing functions RB(var) and (A14) as displayed in Figs. 2 and 3.

10.4. Relative difference of quantiles

Since the true PDF is not known, no estimation method is bias-free. Therefore, the average of quantile estimates of the two distributions, i.e. of LN and Γ, obtained by the same method is taken in the denominator of Eq. (4) instead of the true value:

$$\delta(x_p) = \frac{x_p^{(\mathrm{LN})} - x_p^{(\Gamma)}}{0.5\left(x_p^{(\mathrm{LN})} + x_p^{(\Gamma)}\right)} \tag{106}$$

A comparison of results of MLM and MOM presented in Fig. 12 for probability of exceedance $p = 0.1\%$ and in Fig. 13 for $p = 1\%$ shows a greater relative difference for the MLM quantiles than for the MOM quantiles. That is, for small probability of exceedance, the sensitivity of MOM-estimate of quantiles to the distribution assumption is smaller than it is for the MLM-estimates. This finding is in conformity with the

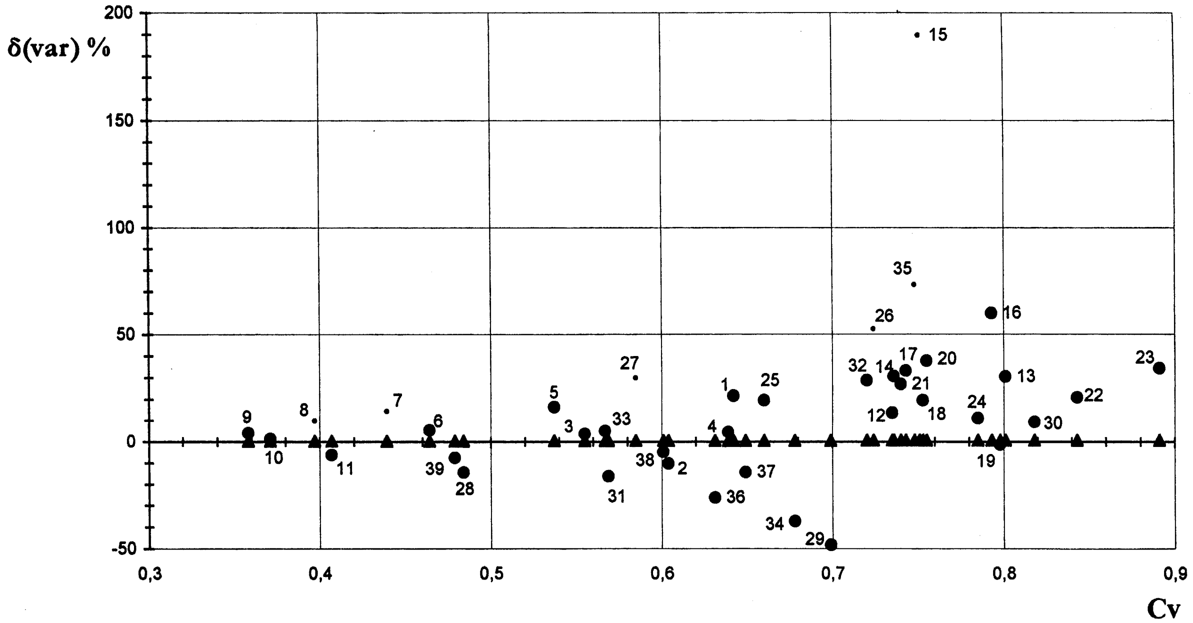

Fig. 10. Relative difference of the ML-variance (Eq. (105)) vs. MOM-coefficient of variation for assumed LN distribution.

theoretical findings on the two distributions obtained for the asymptotic case and presented in Figs. 5 and 6.

11. Conclusions

An analytical method for evaluation of the resistance of the estimates of moments and quantiles by various estimation methods with respect to the distribution choice has been proposed and illustrated using two two-parameter distribution functions. Both the theoretical and empirical findings using the gamma and lognormal distributions show that the bias caused by the wrong distribution choice cannot be disregarded in evaluation of the efficiency of estimation methods in FFA. The RB of the MLM-estimate of moments can be considerable and grows rapidly with increasing value of the coefficient of variation, while the MOM estimates of the two first moments are asymptotically bias free. Since the lognormal and gamma distributions converge with each other and with the normal distribution for the coefficient of variation tending to zero, the RB of the MLM-estimate of moments tends to zero for $c_v \rightarrow 0$.

Similarly, the MOM estimate of the quantiles of upper tails is more resistant to distribution choice than is the MLM estimate. The bias of LMM estimates lies between these two. A comparison of the statistics involved in the Λ-function of other two-parameter distributions of the range $(0, \infty)$ might exhibit similar results for them.

Since MLM used as the approximation method is irreversible, the asymptotic bias of the MLM-estimate of any statistical characteristic (see Eq. (58)) is not asymmetric as is for the MOM and LMM. The MLM-estimate of the upper tail quantiles is more resistant with respect to the wrong choice of the distribution function if the gamma distribution is taken as the hypothetical distribution than in the opposite case. It comes from the fact that the MLM estimate of the mean for the gamma distribution remains unbiased irrespective of what the true distribution function is. That is, having a choice between the lognormal and the gamma, the latter should be selected if the MLM is to be applied. In fact, there are some other PDFs having this convenient property of equivalency of MLM and MOM estimators in respect to the mean. Analysis of the loss of the asymptotic ML value due to a wrong distribution choice points out the weakness of the ML-ratio for identifying the true PDF.

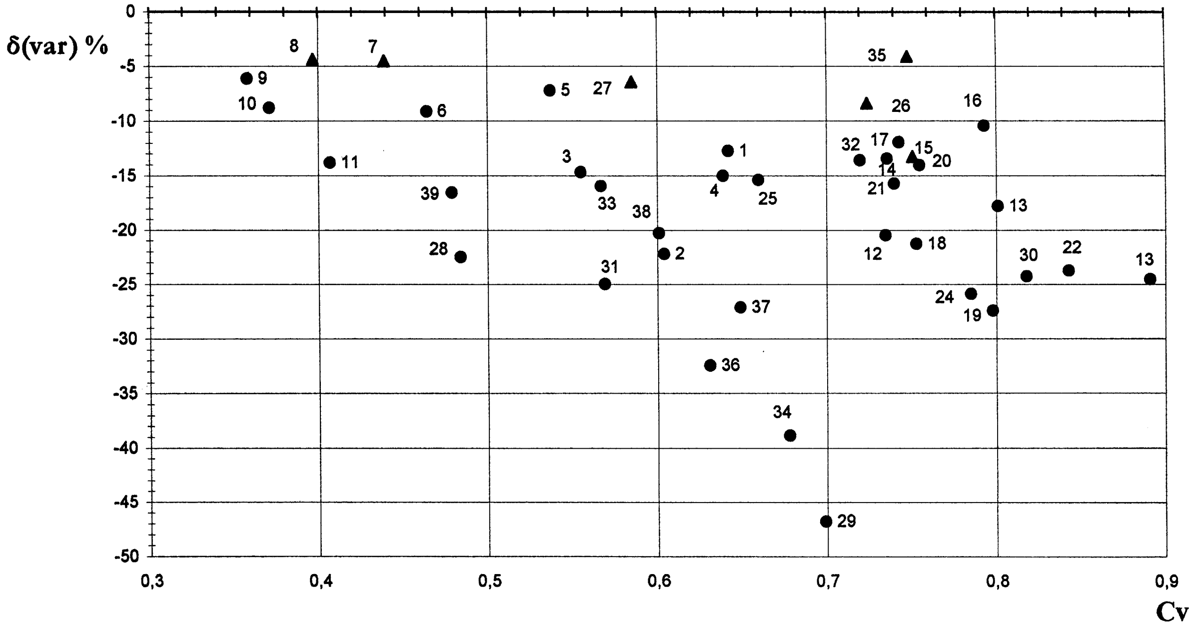

Fig. 11. Relative difference of the ML-variance (Eq. (106)) vs. MOM-coefficient of variation for assumed gamma distribution.

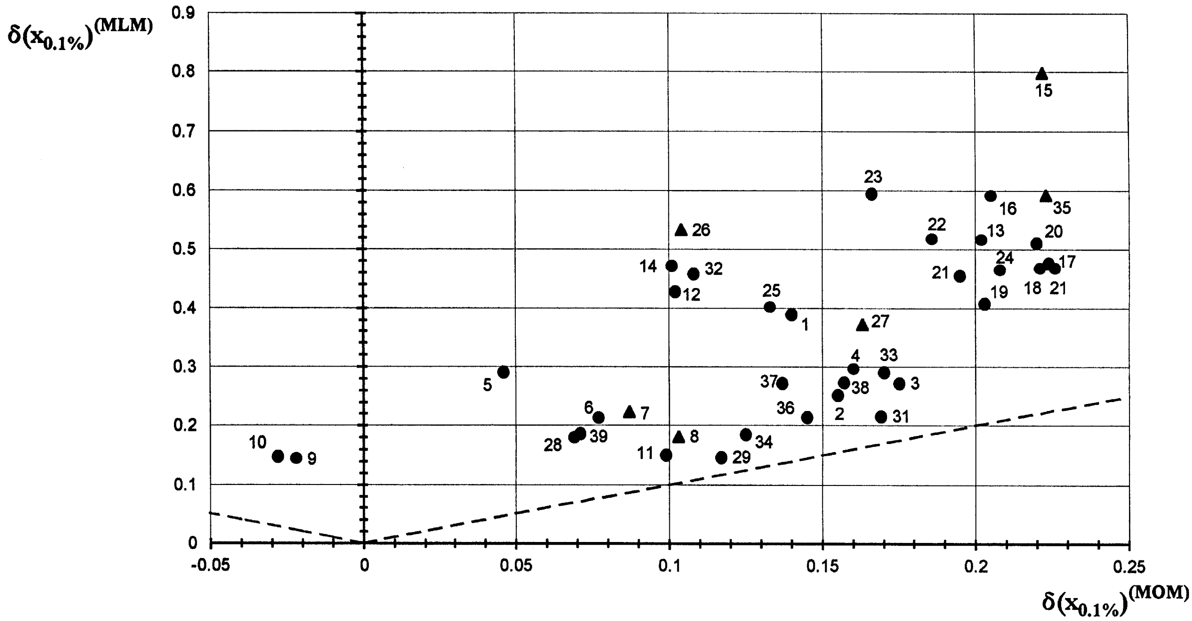

Fig. 12. Relative differences of MOM and MLM estimates of the quantile of order $p = 0.1\%$.

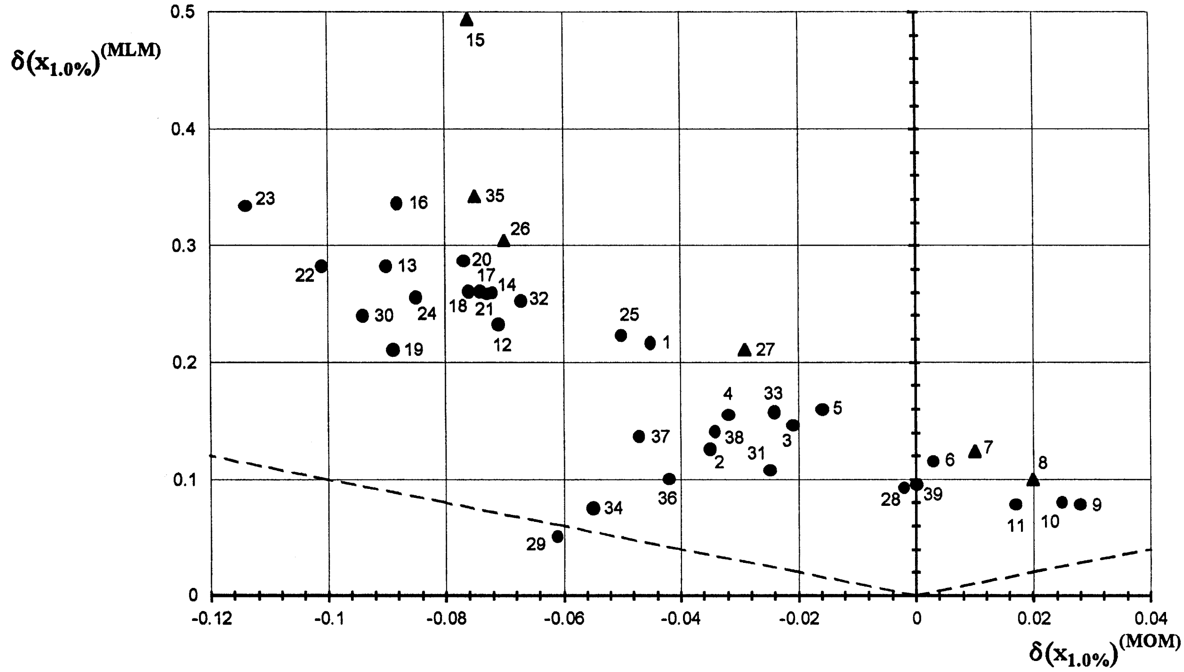

Fig. 13. Relative differences of MOM and MLM estimates of quantile of order $p = 1\%$.

The results of analysis performed on 39 70-year long annual peak flow series of Polish rivers are in agreement with theoretical findings and provide, therefore, an empirical evidence for the necessity to include bias in evaluation of the efficiency of PDF estimation methods. When dealing with random samples of hydrological size, one should take into account both terms of MSE (given by Eq. (1)) in numerical simulation experiments.

Acknowledgements

This work was supported by the Polish State Committee for Scientific Research under the grant KBN No. 6 PO 4D 056 17 entitled 'Revision of Applicability of the Parametric Methods for Estimating Statistical Characteristics of Floods'. The authors are grateful to Meteorology and Water Resources Institute in Warsaw for providing with flood data and to A. Kozlowski (MSc) for data processing of Section 10 of the paper. Two anonymous reviewers made critical but constructive and comprehensive comments on the manuscript. Incorporation of these comments significantly improved the text and the authors gratefully express their gratitude to these reviewers for their thoughtful and thorough review.

Appendix A. MLM approximation of lognormal distribution by gamma distribution

For lognormal distribution (9), the ML-asymptotic estimates of the two parameters are equal to the first moment about the origin and the second central moment of logarithmically transformed variable, respectively:

$$\mu = \int_0^\infty \ln x \varphi^{\mathrm{LN}}(x)\mathrm{d}x \tag{A1}$$

and

$$\sigma^2 = \int_0^\infty (\ln x - \mu)^2 \varphi^{\mathrm{LN}}(x)\mathrm{d}x \tag{A2}$$

The asymptotic maximum log-likelihood function (6) is

$$M_\infty(H = \text{LN}, T = \text{LN}) = -\ln\sqrt{2\pi} - \frac{1}{2} - \mu - \ln\sigma \quad \text{(A3)}$$

Substituting Eq. (16)

$$M_\infty(H = \text{LN}, T = \text{LN}) = -\ln\sqrt{2\pi} - \frac{1}{2} - \ln\sigma + \frac{\sigma^2}{2} - \ln\alpha_1^{\text{LN}} \quad \text{(A4)}$$

where σ^2 is determined by Eq. (15) with c_v^{LN} instead of c_v.

The asymptotic Λ-function for ($H = \Gamma$, $T = \text{LN}$):

$$\Lambda_\infty(H = \Gamma, T = \text{LN}) = \lambda\ln\alpha - \ln\Gamma(\lambda) + (\lambda - 1)\int_0^\infty \ln x\varphi^{\text{LN}}(x)\text{d}x - \alpha\int_0^\infty x\varphi^{\text{LN}}(x)\text{d}x \quad \text{(A5)}$$

MLM-equations:

$$\frac{\partial\Lambda_\infty(H = \Gamma, T = \text{LN})}{\partial\alpha} = \frac{\lambda}{\alpha} - \int x\varphi^{\text{LN}}\,\text{d}x = 0 \quad \text{(A6)}$$

Substituting Eq. (17), we get

$$\alpha_1^\Gamma = \alpha_1^{\text{LN}} \quad \text{(A7)}$$

This means that the mean is not biased for any T-distribution if the gamma distribution is assumed instead of another distribution, say lognormal distribution: $B(\alpha_1) = 0$. For $T = \text{LN}$, we have α_1 given by Eq. (11).

The second MLM-condition reads:

$$\frac{\partial\Lambda_\infty(H = \Gamma, T = \text{LN})}{\partial\lambda} = \ln\alpha - \psi(\lambda) + \int \ln x\varphi^{\text{LN}}\,\text{d}x = 0 \quad \text{(A8)}$$

Substituting logarithmically transformed Eq. (A6), we get

$$\ln\lambda - \psi(\lambda) = \ln\int x\varphi^{\text{LN}}\,\text{d}x - \int \ln x\varphi^{\text{LN}}\,\text{d}x \quad \text{(A9)}$$

Furthermore, we can replace the first term of the RHS of Eq. (A9) by logarithmically transformed Eq. (11) and the second term by Eq. (A1), getting:

$$\ln\lambda - \psi(\lambda) = \frac{\sigma^2}{2} \quad \text{(A10)}$$

Introducing the direct moments of the lognormal distribution as

$$\ln\lambda - \psi(\lambda) = \frac{1}{2}\ln\left[1 + \left(c_v^{\text{LN}}\right)^2\right] \quad \text{(A11)}$$

and the moments of the gamma distribution as

$$-2\ln c_v^\Gamma - \psi\left(\left(c_v^\Gamma\right)^{-2}\right) = \frac{1}{2}\ln\left[1 + \left(c_v^{\text{LN}}\right)^2\right] \quad \text{(A12)}$$

which is the relationship between the ML-approximated coefficient of variation of the gamma distribution, c_v^Γ, and its true value for the LN distribution, c_v^{LN}.

A.1. Asymptotic ML function, M_∞

Substituting the MLM-solutions given by Eqs. (A6), (A7), (A8) and (A9) into Eq. (A5), we get Eq. (7) in the form:

$$M_\infty(H = \Gamma, T = \text{LN}) = \lambda(\ln\lambda - 1) - \ln\Gamma(\lambda) - (\lambda - 1)\frac{\sigma^2}{2} - \ln\alpha_1^{\text{LN}} \quad \text{(A13)}$$

Then, by Eqs. (A9) and (15) both λ and σ^2 can be expressed by the respective function of c_v^{LN} only. To get rid of the first moment about the origin (α_1) in Eqs. (A4) and (A13), it is convenient to compare the functions $M'_\infty(H = \cdot, T = \cdot) = M_\infty(H = \cdot, T = \cdot) + \ln\alpha_1$. The graphs of both functions vs. c_v^{LN} are shown in Fig. 6.

A.1.1. Bias of variance

Substituting Eqs. (17), (18) and (12) into Eq. (4) gives the RB of variance:

$$\text{RB}(\mu_2) = \frac{1}{\lambda[\exp(\sigma^2) - 1]} - 1 \quad \text{(A14)}$$

where in the ML method λ is the function (A9) of σ, or, using the direct moment, the function (A11) of $c_v^{\text{LN}}(X)$. Therefore, the RB of variance given by Eq. (A14) can be determined from the coefficient of variation (c_v^{LN}) only (Fig. 3).

A.1.2. Bias of quantiles

Substituting Eqs. (24) and (23) into Eq. (52) and expressing α_1 by Eq. (11), we get the relative asymptotic bias (4) of quantile as

$$\mathrm{RB}(x_p) = \frac{t_p^{\Gamma}(\lambda)}{\lambda}\exp\left(0.5\sigma^2 - \sigma t_p^N\right) - 1 \qquad \text{(A15)}$$

where λ is determined by Eq. (A9) and σ is the function (13) of the coefficient of variation. Noting Eq. (15), we see that the relative bias of quantile is a function of the coefficient of variation of the *True*-distribution, $c_{\mathrm{v}}^{\mathrm{LN}}$, and the probability of exceedance, p (Fig. 5(a) and (b)).

References

Bobee, B., 1973. Sample error of *T*-year events computed by fitting a Pearson type 3 distribution. Water Resources Research 9 (5), 1264–1270.

Bras, R.L., Moghamian, M.S., McLaughlin, D.B., 1985. Estimation of flood frequency: a comparison of physically-based procedures. US–China Bilateral Symposium on the Analysis of Extra-ordinary Flood Events.

Buishand, T.A., 1990. Bias and variance of quantile estimates from a partial duration series. Journal of Hydrology 120, 35–49.

Chong, S.K., Moore, S.M., 1983. Flood frequency analysis for small watersheds in southern Illinois. Water Resources Bulletin 19 (2), 277–282.

Condie, R., 1977. The log-Pearson type 3 distribution: the *T*-year event and its asymptotic standard error by maximum likelihood theory. Water Resources Research 3 (6), 987–991.

Condie, R., 1986. Flood samples from a three-parameter population with historic information: the asymptotic standard error of estimate of the *T*-year flood. Journal of Hydrology 85, 139–150.

Cunnane, C., 1985. Factors affecting choice of distribution for flood studies. Hydrological Sciences Journal 30 (1), 25–36.

Douglas, J.R., Clarke, R.T., Newton, S.G., 1976. The use of likelihood functions to fit conceptual models with more than one dependent variable. Journal of Hydrology 29, 181–198.

Fisz, M., 1963. Probability Theory and Mathematical Statistics. Wiley, New York chapter 11.

Greewood, 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 Resources Research 15 (5), 1049–1054.

Greis, N.P., 1983. Flood frequency analysis: a review of 1979–1982. Reviews of Geophysics and Space Physics 21 (3), 699–706.

Gupta, V.L., 1970. Selection of frequency distribution models. Water Resources Research 6 (4), 1193–1198.

Hoshi, K., Burges, S.J., 1981. Sampling properties of parameter estimates for the log-Pearson type 3 distribution, using moments in real space. Journal of Hydrology 53, 305–316.

Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of Royal Statistical Society, Series B 5 (1), 105–124.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Flood Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, New York 224 pp..

Kendall, M.G., Stuart, A., 1969. The Advanced Theory of Statistics. Distribution Theory, Charles Griffin & Company, London. Vol. 1.

Kendall, M.G., Stuart, A., 1973. The Advanced Theory of Statistics. Inference and Relationship, vol. 2. Charles Griffin & Company, London.

Klemes, V., 2000a. Tall tales about tails of hydrological distributions: I. Journal of Hydrologic Engineering 5 (3), 227–231.

Klemes, V., 2000b. Tall tales about tails of hydrological distributions: II. Journal of Hydrologic Engineering 5 (3), 232–239.

Kuczera, G., 1982. On the relationship between the reliability of parameter estimates and hydrologic time series data used in calibration. Water Resources Research 18 (1), 146–154.

Landwehr, J.M., Matalas, N.C., Wallis, J.R., 1979. Probability-weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. Water Resources Research 15, 1055–1064.

Landwehr, J.M., Matalas, N.C., Wallis, J.R., 1980. Quantile estimation with more or less floodlike distribution. Water Resources Research 16 (3), 547–555.

Lombard, F., 1988. Detecting change points by Fourier analysis. Technometrics 30 (3), 305–310.

Mann, H.B., 1945. Nonparametric tests against trend. Econometircs 13, 245–259.

Mitosek, H.T., Strupczewski, W.G., 1996. Can series of maximal annual flow discharges be treated as realizations of stationary process? Acta Geophysica Polonica XLIV (1), 61–77.

Mockus, V., 1960. Selecting a flood frequency method. Transactions of ASAE 3, 48–54.

Mutua, F.M., 1994. The use of the Akaike information criterion in the identification of an optimum flood frequency model. Hydrological Science Journal 39-3, 235–244.

Nash, J.E., 1959. Systematic determination of unit hydrograph parameters. Journal of Geophysical Research 64 (1), 111–115.

Pettitt, A.N., 1979. A nonparametric approach to the change-point problem. Applied Statistics 28 (2), 126–135.

Phien, H.N., Hsu, L.C., 1985. Variance of the *T*-year event in the log-Pearson type 3 distribution. Journal of Hydrology 77, 141–158.

Pilon, P.J., Adamowski, K., 1993. Asymptotic variance of flood quantile in log-Pearson type III distribution with historical information. Journal of Hydrology 143, 491–503.

Rao, A.R., Hamed, K.H., 2000. Flood Frequency Analysis. CRC Press, Boca Raton 350 pp..

Singh, V.P., 1988. Hydrologic Systems. Rainfall-Runoff Modeling, vol. 1. Prentice-Hall, Englewood Cliffs, NJ.

Singh, V.P., 1998. Entropy-Based Parameter Estimation in Hydrology. Kluwer Academic Publishers, Boston 365 pp.

Singh, V.P., Adrian, D.D., 2000. Information theory for design of transportation systems. Technical Report, Louisiana Transportation Research Center, Baton Rouge, LA.

Singh, V.P., Rajagopal, A.K., 1986. A new method of parameter estimation from hydrologic frequency analysis. Hydrological Science and Technology 2 (3), 33–40.

Snyder, W.M., 1972. Fitting of distribution functions by nonlinear least squares. Water Resources Research 8 (4), 1423–1432.

Spence, E.S., 1973. Theoretical frequency distributions for the analysis of plains streamflow. Canadian Journal of Earth Science 10, 130–139.

Stedinger, J.R., Tasker, G.D., 1985. Regional hydrologic analysis: 1. Ordinary, weighted, and generalized least squares compared. Water Resources Research 21 (9), 1421–1432.

Stevens, E.W., 1992. Uncertainty of extreme flood estimates incorporating historical flood data. Water Resources Bulletin 28 (6), 1057–1068.

Strupczewski, W.G., 1999. A farewell to the ML method in flood frequency analysis. In: Singh, V.P., et al. (Ed.), Hydrologic Modeling. Proceedings of WEESHE, Seoul, WRP, LLC, Colorado, pp. 291–306.

Strupczewski, W.G., Singh, V.P., Weglarczyk, S., 2001. Impulse response of linear diffusion analogy model as a flood frequency probability density function. Hydrological Science Journal 46 (5), 761–780.

Turkman, K.F., 1985. The choice of extremal models by Akaike's information criterion. Journal of Hydrology 82, 307–315.

Wang, S.X., Singh, V.P., 1994. Sampling variance of a *T*-year flood estimated by curve fitting. Stochastic Hydrology and Hydraulics 8, 253–258.

WMO, 1988. Analyzing long time series of hydrological data with respect to climate variability. Project description, WCAP Report No. 3. WMO/TD—No. 224.