A C T A G E O P H Y S I C A P O L O N I C A

Vol. 50, No. 3 2002

PHYSICS OF FLOOD FREQUENCY ANALYSIS. PART I. LINEAR CONVECTIVE DIFFUSION WAVE MODEL

Witold G. STRUPCZEWSKI [1], Stanisław WĘGLARCZYK [2] and Vijay P. SINGH [3]

[1] Institute of Geophysics, Polish Academy of Sciences
ul. Księcia Janusza 64, 01-452 Warszawa, Poland
e-mail: wgs@igf.edu.pl

[2] Institute of Water Engineering and Water Management
Technological University of Cracow, ul. Warszawska 24, 31-155 Kraków, Poland
e-mail: sweglar@smok.wis.pk.edu.pl

[3] Department of Civil and Environmental Engineering
Louisiana State University, Baton Rouge, Louisiana 70803-6405, USA
e-mail: cesing@lsu.edu

Abstract

It is hypothesized that the impulse response of a linearized convective diffusion wave (CD) model is a probability distribution suitable for flood frequency analysis. This flood frequency model has two parameters, which are derived using the methods of moments and maximum likelihood. Also derived are errors in quantiles for these methods of parameter estimation. The distribution shows an equivalency of the two estimation methods with respect to the mean value – an important property in the case of unknown true distribution function. As the coefficient of variation tends to zero (with the mean fixed), the distribution tends to a normal one, similar to the lognormal and gamma distributions.

Key words: flood frequency, probability distribution, convective diffusion model, moments, quantiles, method of moments, maximum likelihood, standard error.

1. INTRODUCTION

Mathematical models of flood frequency analysis (FFA) can be broadly classified into: (1) empirical, (2) phenomenological, and (3) physically-based. An excellent discussion of empirical models is given, e.g., by Stedinger *et al.* (1993), Rao and Hamed (2000). Till today these models continue to be most popular for doing FFA all over the world. Phenomenological models employ a set of probabilistic axioms, which lead to a probabilistic model of one or more flood characteristics. Examples of this type of models are those based on the use of random number of random variables (Todorovic, 1982), the entropy theory (Singh, 1998), and the like. These models received a good deal of attention in the 1970s and the 1980s but did not become popular, partly because of their higher mathematical demands. Physically-based models employ dynamical principles of flood generation. Eagleson (1972) was probably one of the first to employ such a model. Another example is the use of watershed models, as, e.g., the stochastic flood model developed by Schaefer (1998).

Along the lines of physically based models and recognizing that channels are the dominant conduits for transmission of flood waters, it is plausible to develop a model that employs the physics of channel flow routing and in which no explicit consideration is given to the hydrologic processes occurring on the land areas of the watershed. It is well accepted that a good representation of the physics of channel flow is given by the linearized convective diffusion wave approximation. It is then hypothesized that impulse response function (IRF) of such a model can be considered as a probability density function (PDF) for FFA. Although the impulse response of a hydrologic system or the response of an initially relaxed linear deterministic system for the Dirac-δ impulse belongs to the class of purely deterministic functions, it is not difficult to find a stochastic interpretation of the impulse response. If one imagines that the unit volume of the Dirac-δ impulse consists of an infinite number of particles (or drops) then the integral of the impulse response $\int_0^T h(x,t)\,\mathrm{d}t$ determines the probability that a single particle passes the outlet at x during time $(0, T)$, where $h(x,t)$ is the impulse response function at time t and position x. Apart from its stochastic interpretation, it should be noted that the impulse response function fulfills several requirements normally expected of the flood frequency models, namely, (1) semi-infinite lower bounded range with a non-negative value of the bound; (2) positive skewness and the unit integral over the whole range $\int_0^\infty h(x,t)\,\mathrm{d}t$; and (3) uni-modality, which is the property of all single component FF distributions. As an example, the gamma function is used both as the impulse response of a cascade of equal linear reservoirs and the PDF in FFA.

Because of the practice of applying the existing probability distributions in FFA, emphasis in statistical hydrology has been on assessing the accuracy of parameter estimators using the Monte Carlo simulation techniques. As a result, not much attention has been paid to the development of physically based probability distributions taking into account peculiarities of hydrologic phenomena and the attendant statistical reasoning. To that end, this study espouses the use of IRFs as PDFs.

The objective of this two-part paper is to hypothesize the linear convective diffusion wave (CD) model of flow routing as a probability distribution function for flood frequency modeling and to assess its applicability for FFA. This part introduces the CD model and describes the techniques of its use. The second part of the paper is focused on discussion of the validity of this hypothesis on Polish rivers and on the comparison of CD with its competitor – the lognormal distribution.

2. THE LINEAR CONVECTIVE DIFFUSION WAVE MODEL FOR DOWNSTREAM CHANNEL RESPONSE

In one-dimensional flood routing analysis, the prediction of flood characteristics at a downstream section on the basis of the knowledge of flow characteristics at an upstream section is known as the downstream problem. Using the linearization of the Saint-Venant equation, the solution of the upstream boundary problem was derived, e.g., by Deymie (1939), Masse (1939), Dooge and Harley (1967), Dooge *et al.* (1987a, b); a discussion of this problem is presented in Singh (1996). The solution is a linear, physically based model with four parameters dependent on the hydraulic characteristics of the channel reach at the reference level of linearization. However, the complete linear solution is complex in form and is relatively difficult to compute (Singh, 1996). Two simpler forms of the linear channel downstream response are recognized in the hydrologic literature and are designated as linear convective diffusion (CD) model and linear rapid flow (RF) model. These correspond to the limiting flow conditions of the linear channel response, i.e., where the Froude number is equal to zero (Hayami, 1951; Dooge, 1973) and where it is equal to one (Strupczewski and Napiórkowski, 1990).

The linear convective diffusion analogy (CD) model (or the linear convective diffusion model) is based on the solution of the linearized Saint-Venant equation for a semi-infinite, uniform channel with Froude number (F_0) equal to zero and consequently used for small Froude numbers. Its impulse response is given by

$$h(x,t) = \frac{x}{\sqrt{4\pi D t^3}} \exp\left[-\frac{(x-ut)^2}{4Dt}\right], \tag{1}$$

where x is the length of the channel reach, t is the time, u is the convective velocity and D is the hydraulic diffusivity. Both u and D are the functions of channel and flow characteristics at the reference steady state condition. To use CD model for other than

F_0=0 reference flow conditions, a pure lag component has been added (Strupczewski and Dooge, 1996).

The function given by (1) is known in statistics (e.g., Cox and Miller, 1965; p. 221) as the probability density function of the first passage time T for a Wiener process starting at 0 to reach absorbing barrier at the point x, where u is the positive draft and D is the variance of the Wiener process. It was applied by Moore and Clarke (1983) and Moore 1984) as the transfer function of the sediment routing model. The function in (1) is considered as the flood frequency model in this study.

3. CD PROBABILITY DENSITY FUNCTION AND ITS PROPERTIES

Denoting $x/\sqrt{4D} = \alpha$ and $xu/4D = \beta$ and renaming t as x, one gets a two-parameter probability density function of the form:

$$f(x;\alpha,\beta) = \frac{\alpha}{\sqrt{\pi x^3}} \exp\left[-\frac{\left(\alpha - \frac{\beta}{\alpha}x\right)^2}{x}\right], \qquad \alpha,\beta > 0. \tag{2}$$

One can extend (2) to the case of a three-parameter distribution with a lower bound fixed at ε :

$$f(x;\alpha,\beta,\varepsilon) = \frac{\alpha}{\sqrt{\pi(x-\varepsilon)^3}} \exp\left[-\frac{\left(\alpha - \frac{\beta}{\alpha}(x-\varepsilon)\right)^2}{x-\varepsilon}\right], \qquad x > \varepsilon; \qquad \alpha,\beta > 0. \tag{2a}$$

For a length of channel reach approaching zero, the impulse response function of (1) tends to the Dirac-delta function

$$\lim_{x\to 0} h(x,t) = \delta(t). \tag{3}$$

Hence, for the value of the α-parameter approaching zero, the PDF of (2) tends to the Dirac-delta function, which is the limiting case of the normal distribution function:

$$\lim_{\alpha\to 0} h(x;\alpha,\beta) = \delta(x). \tag{4}$$

Tweedie (1957) termed the density function of (2) as an inverse Gaussian PDF, Johnston and Kotz (1970) summarized its properties, and Folks and Chhikara (1978) provided a review of its development.

Cumulants

The cumulants of the two-parameter distribution of (2) are

$$k_1 = \frac{\alpha^2}{\beta}, \tag{5}$$

$$k_r = \frac{1}{2^{r-1}}\{(1)(3)(5)\ \dots(2r-3)\}\,\alpha\left(\frac{\alpha}{\beta}\right)^{2r-1} \quad \text{for } r > 1 \tag{6}$$

and those for the three-parameter distribution are

$$k_1 = \frac{\alpha^2}{\beta} + \varepsilon\,. \tag{5a}$$

Moments

Using the relations between moments and cumulants (Kendall and Stuart, 1969; p. 70) and (5) and (6) the expressions for the first four moments of (2) with $\varepsilon = 0$ are given below:

$$\mu_1' = \frac{\alpha^2}{\beta}, \tag{7}$$

$$\mu_2 = \frac{1}{2}\frac{\alpha^4}{\beta^3}, \tag{8}$$

$$\mu_3 = \frac{3}{4}\frac{\alpha^6}{\beta^5} = 3\frac{\mu_2}{\mu_1'}, \tag{9}$$

$$\mu_4 = k_4 + 3k_2^2 = \frac{3\alpha^8}{8\beta^7}(2\beta+5) = 3\mu_2^2\left(5c_v{}^2+1\right). \tag{10}$$

Dimensionless coefficients

The coefficient of variation is

$$C_v = \frac{1}{\sqrt{2\beta}}. \tag{11}$$

The distribution is positively skewed with the coefficient of skewness given by

$$C_s = \frac{\mu_3}{\mu_2^{3/2}} = \frac{3}{\sqrt{2\beta}} = 3C_v \tag{12}$$

which for the lognormal (LN) distribution is $C_s = 3C_v + C_v^3$ and for the gamma (G) distribution it is $C_s = 2C_v$.

The coefficient of kurtosis (C_k) is

$$C_k = \frac{\mu_4}{\mu_2^2} = 3\left(\frac{5}{2\beta}+1\right) = 2\left(5C_v^2+1\right). \tag{13}$$

The modal value is given by

$$x_{mod} = \frac{3\alpha^2}{4\beta^2}\left(\sqrt{\left(\frac{4}{3}\beta\right)^2+1} - 1\right) \tag{14}$$

and in the dimensionless form

$$\frac{x_{mod}}{\mu_1'} = \frac{3}{2}\left(\sqrt{\frac{4}{9}+C_v^4} - C_v^2\right). \tag{15}$$

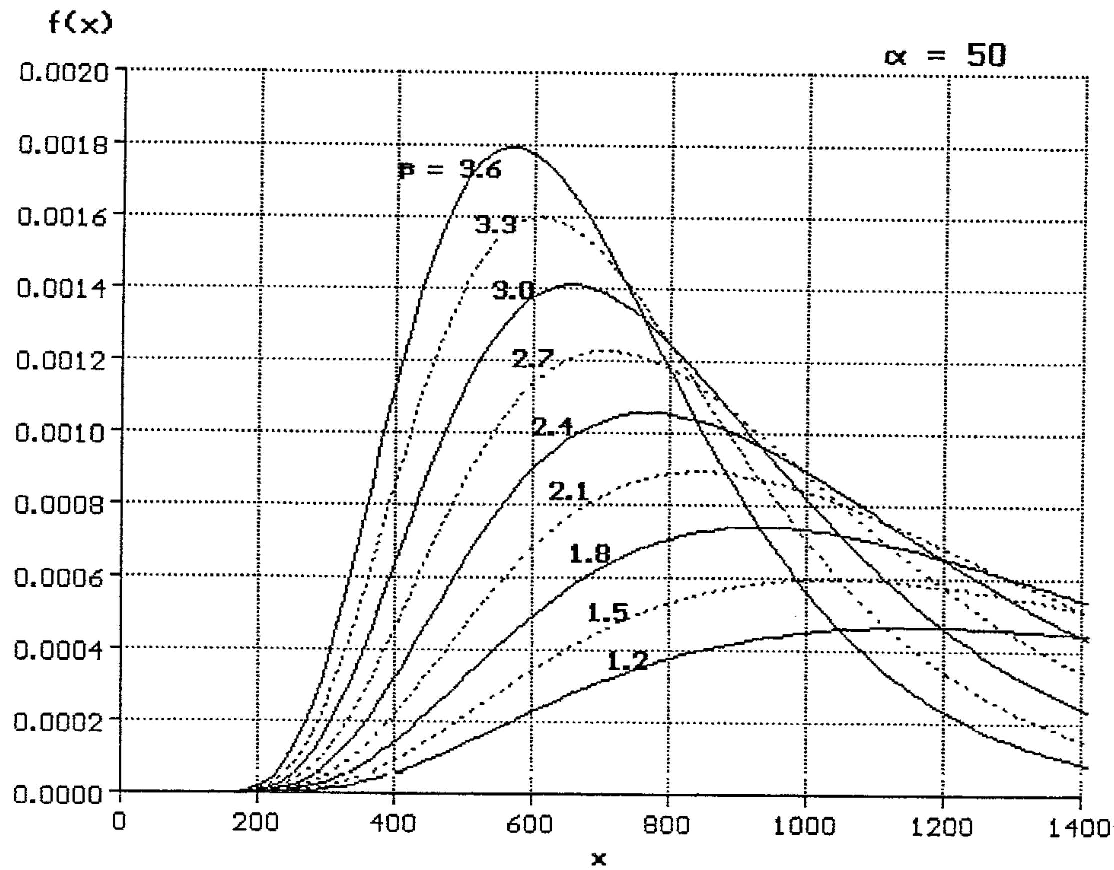

Fig. 1. Typical CD PDFs for α = const. The values $\alpha = 50$ and $\beta = 1.2$ to 3.6 were chosen as covering the range of β found experimentally.

For $C_v \rightarrow 0$ the distribution tends to be symmetric like the two-parameter lognormal (LN2) and the gamma distributions.

Typical graphs of the distribution for some selected values of α and β are presented in Figs. 1 and 2. The values were selected so as to approximately cover the range of α and β obtained in the method of moments (MOM) and maximum likelihood method (MLM) estimation carried out for application of the CD model part of the paper. Although the intensity of changes is greater in Fig. 2, both figures seem to exhibit the same pattern, i.e., for increasing modal value x_{mod} the maximum of $f(x)$ is seen to be decreasing. It should be noted, however, that α is more connected with the absolute scale of x, as the dimension of α^2 is the same as the dimension of x while β, being dimensionless, is responsible for the relative scale of x; see relation (11)–(13).

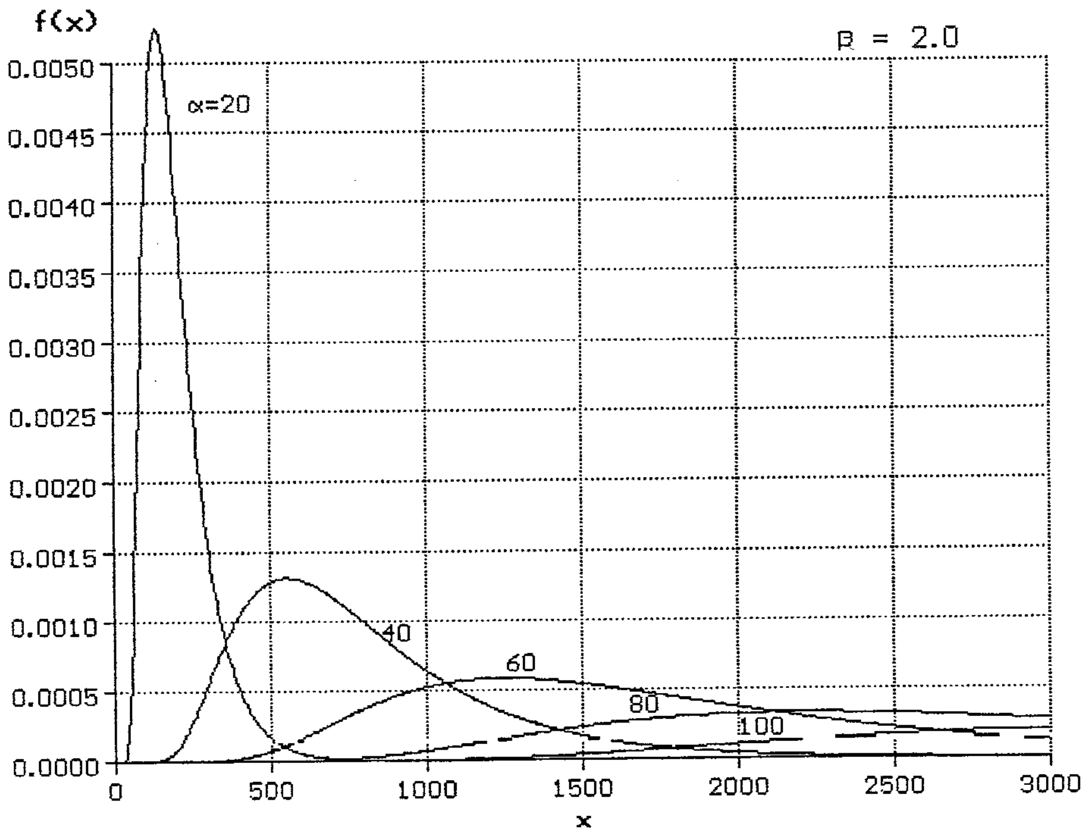

Fig. 2. Typical CD PDFs for β = const. The values $\beta = 2$ and $\alpha = 20...100$ were chosen as covering the range of α found experimentally.

The asymmetry of all graphs in Fig. 2 is the same and all of them are identical if rescaled to a nondimensional coordinate system, e.g., $f(x) \rightarrow f(x)/f(x_{mod})$, $x \rightarrow x/x_{mod}$.

4. PROBABILITY OF EXCEEDANCE

A quantile x_p corresponding to the probability of exceedance p is obtained by integrating (2)

$$x_p = \left(\frac{\alpha}{t_p^{LDA}(\beta)} \right)^2 , \tag{16}$$

where $t_p^{LDA}(\beta)$ is the upper limit of the integral:

$$p = \frac{2}{\sqrt{\pi}} \int_0^{t_p} \exp\left[-\left(z - \frac{\beta}{z} \right)^2 \right] \mathrm{d}z \, . \tag{17}$$

For computational purposes, the integral in (17) can be simplified as follows. The probability of exceedance for the distribution in (2) can be recast as

$$p = \int_{x_p}^{\infty} \frac{\alpha}{\sqrt{\pi x^3}} \exp\left[-\left(\frac{\alpha}{\sqrt{x}} - \frac{\beta}{\alpha}\sqrt{x} \right)^2 \right] \mathrm{d}x \, . \tag{18}$$

Substituting

$$z = \frac{\alpha}{\sqrt{x}} \tag{19}$$

into (18), one obtains

$$p = \frac{2}{\sqrt{\pi}} \int_0^{\alpha/\sqrt{x}} \exp\left[-\left(z - \frac{\beta}{z} \right)^2 \right] \mathrm{d}z \, , \tag{20}$$

which is the same as (17). Therefore, (16) holds.

In order to transform (17) to a more convenient computational form, the integral to be considered is

$$C(t, \beta) = \int_0^{t} \exp\left[-\left(z - \frac{\beta}{z} \right)^2 \right] \mathrm{d}z \, , \tag{21}$$

where

$$t \equiv t_p = \frac{\alpha}{\sqrt{x_p}} \tag{22}$$

is positive.

Substituting

$$y = z - \frac{\beta}{z} \tag{23}$$

into (21) and solving (23) with respect to z we get

$$z = \frac{1}{2}\left(y + \sqrt{y^2 + 4\beta}\right) \tag{24}$$

while

$$dz = \frac{1}{2}\left(1 + \frac{y}{\sqrt{y^2 + 4\beta}}\right)dy . \tag{25}$$

Therefore,

$$\begin{aligned} C(t,\beta) &= \int_{-\infty}^{t-\beta/t} \exp(-y^2)\frac{1}{2}\left(1 + \frac{y}{\sqrt{y^2 + 4\beta}}\right)dy \\ &= \frac{1}{2}\int_{-\infty}^{t-\beta/t} \exp(-y^2)dy + \frac{1}{2}\int_{-\infty}^{t-\beta/t} \exp(-y^2)\frac{y}{\sqrt{y^2 + 4\beta}}dy \\ &= 0.5\left[C_1(t,\beta) + C_2(t,\beta)\right]. \end{aligned} \tag{26}$$

Then, substituting $y = u/\sqrt{2}$ into the first term of (26), we get the value of the first integral $C_1(t, \beta)$

$$C_1(t,\beta) = \int_{-\infty}^{t-\beta/t} \exp(-y^2)dy = \frac{1}{\sqrt{2}}\int_{-\infty}^{\sqrt{2}(t-\beta/t)} \exp(-\frac{1}{2}u^2)du = \sqrt{\pi}\Phi\left[\sqrt{2}\left(t - \beta/t\right)\right], \tag{27}$$

where $\Phi(\cdot)$ is the cumulative probability of the normal distribution $N(0, 1)$.

The function $f(t)$ under the second integral $C_2(t, \beta)$

$$C_2(t,\beta) = \int_{-\infty}^{t-\beta/t} \exp(-y^2)\frac{y}{\sqrt{y^2 + 4\beta}}dy \tag{28}$$

is asymmetric in respect to zero. Hence, for $c < 0$

$$\int_{-\infty}^{c} f(y)\,dy = -\int_{-c}^{\infty} f(y)\,dy \tag{29}$$

and for $c > 0$

$$\int_{-\infty}^{c} f(y)\,dy = -\int_{0}^{\infty} f(y)\,dy + \int_{0}^{c} f(y)\,dy . \tag{30}$$

Therefore, it is enough to compute the integral

$$C_2'(t,\beta) = \int_{0}^{t-\beta/t} \exp(-y^2)\frac{y}{\sqrt{y^2 + 4\beta}}dy \tag{31}$$

along the positive values.

Making substitution into (31)

$$x=\sqrt{y^2+4\beta} \tag{32}$$

we get

$$\mathrm{d}x=\frac{y}{\sqrt{y^2+4\beta}}\,\mathrm{d}y \tag{33}$$

and

$$\begin{aligned} C_2'(t,\beta) &= \int_0^{t-\beta/t} \exp(-y^2)\frac{y}{\sqrt{y^2+4\beta}}\,\mathrm{d}y = \int_{\sqrt{4\beta}}^{t+\beta/t} \exp(-x^2+4\beta)\,\mathrm{d}x \\ &= \frac{\exp(4\beta)}{\sqrt{2}} \int_{\sqrt{8\beta}}^{\sqrt{2}(t+\beta/t)} \exp(-x^2/2)\mathrm{d}x = \sqrt{\pi}\exp(4\beta)\left\{\Phi\left[\sqrt{2}\left(t+\beta/t\right)\right]-\Phi\left(\sqrt{8\beta}\right)\right\}. \end{aligned} \tag{34}$$

Using (34) and taking into account the asymmetry of the function $f(t)$ in (32), it is easy to show that for both negative and positive values of $(t-\beta/t)$ the second integral is expressed by the same equation:

$$C_2(t,\beta)=\int_{-\infty}^{t-\beta/t} \exp(-y^2)\frac{y}{\sqrt{y^2+4\beta}}\,\mathrm{d}y=-\sqrt{\pi}\exp(4\beta)\left[1-\Phi\left(\sqrt{2}\left(t+\beta/t\right)\right)\right]. \tag{35}$$

Finally, we get

$$\begin{aligned} p &= \frac{2}{\sqrt{\pi}}\int_0^{\alpha/\sqrt{x}} \exp\left[-\left(z-\frac{\beta}{z}\right)^2\right]\mathrm{d}z = \frac{2}{\sqrt{\pi}}\left[\frac{1}{2}C_1\left(\alpha/\sqrt{x},\beta\right)+\frac{1}{2}C_2\left(\alpha/\sqrt{x},\beta\right)\right] \\ &= \frac{1}{\sqrt{\pi}}\left[\sqrt{\pi}\,\Phi\left[\sqrt{2}\left(\alpha/\sqrt{x}-\beta\sqrt{x}\right)\right]-\sqrt{\pi}\exp(4\beta)\left\{1-\Phi\left[\sqrt{2}\left(\alpha/\sqrt{x}+\beta\sqrt{x}\right)\right]\right\}\right] \\ &= \Phi\left[\sqrt{2}\left(\alpha/\sqrt{x}-\beta\sqrt{x}\right)\right]-\exp(4\beta)\left\{1-\Phi\left[\sqrt{2}\left(\alpha/\sqrt{x}+\beta\sqrt{x}\right)\right]\right\} \end{aligned} \tag{36}$$

or introducing (22),

$$\begin{aligned} p &= \frac{2}{\sqrt{\pi}}\int_0^{t_p} \exp\left[-\left(z-\frac{\beta}{z}\right)^2\right]\mathrm{d}z \\ &= \Phi\left[\sqrt{2}\left(t_p-\beta/t_p\right)\right]-\exp(4\beta)\left\{1-\Phi\left[\sqrt{2}\left(t_p+\beta/t_p\right)\right]\right\}. \end{aligned} \tag{37}$$

To prove the correctness of derivation one can calculate $-\mathrm{d}p/\mathrm{d}t_p$ getting (2), where $\Phi(\cdot)$ is the cumulative probability of the normal distribution $N(0, 1)$. Some values of t_p for given β and p are listed in Table 1.

Table 1

CD quantile $t_p(\beta)$ for given values of β and probability of exceedance p

	β									
p [%]	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
50	0.8601	1.1149	1.3209	1.4987	1.6572	1.8020	1.9359	2.0610	2.1791	2.2908
40	0.7678	1.0245	1.2313	1.4092	1.5681	1.7129	1.8468	1.9721	2.0900	2.2019
30	0.6787	0.9355	1.1418	1.3194	1.4779	1.6225	1.7562	1.8811	1.9990	2.1108
20	0.5876	0.8414	1.0457	1.2219	1.3793	1.5229	1.6559	1.7804	1.8977	2.0089
10	0.4830	0.7281	0.9273	1.0999	1.2546	1.3961	1.5273	1.6503	1.7663	1.8765
5	0.4136	0.6486	0.8416	1.0100	1.1617	1.3007	1.4300	1.5512	1.6660	1.7749
2	0.3510	0.5723	0.7571	0.9198	1.0673	1.2029	1.3293	1.4483	1.5611	1.6683
1	0.3167	0.5284	0.7073	0.8658	1.0098	1.1430	1.2673	1.3846	1.4958	1.6017
0.5	0.2898	0.4926	0.6658	0.8201	0.9610	1.091*7	1.2141	1.3295	1.4392	1.5439
0.2	0.2618	0.4539	0.6200	0.7693	0.9065	1.0339	1.1535	1.2666	1.3743	1.4773
0.1	0.2446	0.4296	0.5908	0.7365	0.8706	0.9957	1.1134	1.2250	1.3312	1.4331

5. ESTIMATION OF PARAMETERS BY MAXIMUM LIKELIHOOD METHOD

Parameters α and β of the PDF given by (2) can be estimated using the maximum likelihood method (MLM). To that end, the log-likelihood function is

$$\ln L = N \ln \alpha - \frac{N}{2} \ln \pi - \frac{3}{2} \sum_{i=1}^{N} \ln x_i - \sum_{i=1}^{N} \frac{\left(a - \frac{\beta}{\alpha} x_i \right)^2}{x_i}, \tag{38}$$

where N is the sample size and x_i is the i-th sample value.

For equivalency of the MLM and MOM estimators of parameters, the log-likelihood function must be a linear function of the moments (Kendall and Stuart, 1973, p. 12, 26, 67). The term ln L in (38) is a linear function of the first moment only, which is sufficient for the equivalency of the estimators of the mean value. From (38), the maximum likelihood (ML) equations are found to be:

$$\frac{\partial \ln L}{\partial \alpha} = \frac{N}{\alpha} - 2 \sum_{i=1}^{N} \frac{\left(\alpha - \frac{\beta}{\alpha} x_i \right)\left(1 + \frac{\beta}{\alpha^2} x_i \right)}{x_i} = 0, \tag{39}$$

$$\frac{\partial \ln L}{\partial \beta} = 2\sum_{i=1}^{N}\left(1-\frac{\beta}{\alpha^2}x_i\right) = 0. \tag{40}$$

The solution of (39) and (40) yields equations for α and β :

$$\alpha^2 = \frac{1}{2\left[E\left(X^{-1}\right)-\frac{1}{E(X)}\right]}, \tag{41}$$

$$\beta = \frac{\alpha^2}{E(X)} = \frac{1}{2\left[E(X)E\left(X^{-1}\right)-1\right]}. \tag{42}$$

The principle of maximum entropy was also applied to derive parameters α and β (Singh, 1998). It turned out, the estimation equations were equivalent to (41) and (42). Comparing (42) with (7), it is seen that the maximum likelihood (ML)-estimate of the mean is the same as the one of MOM. This favors the contiguity of the values of MOM's and MLM's estimates of the parameters and quantiles of the distribution, which is an important feature for the use of MLM under false distributional assumption (Strupczewski *et al.*, 2002). However, since the statistical characteristics of a sample used in the ML method are $E(X)$ and $E(X^{-1})$, the parameter estimates are not highly sensitive to high values of x in a sample and as such they are robust to outliers. MLM when applied to the CD obtains a good fit of the left part of the CD function to the data. It should be considered, as the true PDF is not known in reality. Hence, to get a good fit of the upper tail to the sample data, MOM is preferred. For comparison, the ML-equations of LN2 contain the two terms related to X, i.e., $E(\ln X)$ and $E(\ln X)^2$ which are more sensitive to the extreme values of X than $E(X^{-1})$.

Accuracy of estimated parameters

Solving (41) and (42) in respect of the mean and the mean of the reciprocals of the variate, one gets

$$E(X) = \frac{\alpha^2}{\beta}, \tag{43}$$

$$E\left(X^{-1}\right) = \frac{1}{\alpha^2}\left(\frac{1}{2}+\beta\right). \tag{44}$$

Taking the second-order derivatives of function $\ln L$ eq. (38), one obtains

$$\frac{\partial^2 \ln L}{\partial \alpha^2} = -\frac{N}{\alpha^2} - 2\sum_{i=1}^{N}\frac{1}{x_i} - 6\frac{\beta^2}{\alpha^4}\sum_{i=1}^{N}x_i \, , \tag{45}$$

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{2}{\alpha^2}\sum_{i=1}^{N}x_i \, , \tag{46}$$

$$\frac{\partial^2 \log L}{\partial \alpha \partial \beta} = 4\frac{\beta}{\alpha^3}\sum_{i=1}^{N}x_i \, . \tag{47}$$

Substituting equations (43) and (44) into (45)–(47), one gets an asymptotic estimate of the second-order derivatives in terms of α and β :

$$\left[-E\left(\frac{\partial^2 \ln L}{\partial \alpha \partial \beta}\right)\right] = \begin{bmatrix} \frac{2N}{\alpha^2}(1+4\beta) & -\frac{4N}{\alpha} \\ -\frac{4N}{\alpha} & \frac{2N}{\beta} \end{bmatrix}, \tag{48}$$

$$\text{Det} = \frac{4N^2}{\alpha^2 \beta}. \tag{49}$$

Therefore, the variance-covariance matrix has the form:

$$\mathbf{M}^{(MLM)}(\alpha,\beta) = \left[r_{\alpha,\beta}\; D(\alpha)\; D(\beta)\right] = \begin{bmatrix} \frac{\alpha^2}{2N} & \frac{\alpha\beta}{N} \\ \frac{\alpha\beta}{N} & \frac{\beta}{2N}(1+4\beta) \end{bmatrix}. \tag{50}$$

In particular, the coefficient of correlation of ML estimators of α and β is

$$r_{\alpha,\beta}^{(MLM)} = \frac{2}{\sqrt{\frac{1}{\beta}+4}} = \frac{1}{\sqrt{1+0.5c_v^{\;2}}}, \tag{51}$$

where r is the coefficient of correlation and c_v is the coefficient of variation. A high value of c_v leads to a small value of r between α and β .

Proceeding in the same manner with the mean (m) and variance (v) as parameters, we get the variance-covariance matrix in the form

$$\mathbf{M}^{(MLM)}(m,v) = \left[r_{m,v}\; D(m)\; D(v)\right] = \begin{bmatrix} \frac{v}{N} & \frac{3v^2}{mN} \\ \frac{3v^2}{mN} & \frac{2v^2}{N}\left(\frac{9v}{2m^2}+1\right) \end{bmatrix} \tag{52}$$

and hence the correlation coefficient of ML estimators of m and v is

$$r_{m,v}^{(MLM)} = \frac{3c_v}{\sqrt{9c_v^{\,2}+2}}. \tag{53}$$

Asymptotic standard error of quantiles

To derive the asymptotic error of quantiles, the logarithmic transformation of (16) is obtained as

$$y_p = \ln x_p = 2\left[\ln\alpha - \ln t_p(\beta)\right]. \tag{54}$$

Therefore,

$$D^2\left(y_p\right) = D^2(\alpha)\left(\frac{\partial y_p}{\partial\alpha}\right)^2 + D^2(\beta)\left(\frac{\partial y_p}{\partial\beta}\right)^2 + 2r_{\alpha,\beta}D(\alpha)D(\beta)\left(\frac{\partial y_p}{\partial\alpha}\right)\left(\frac{\partial y_p}{\partial\beta}\right), \tag{55}$$

where $D(\alpha)$ and $D(\beta)$ are defined in the matrix equation (50) while from (54)

$$\frac{\partial y_p}{\partial\alpha} = \frac{2}{\alpha}, \tag{56}$$

$$\frac{\partial y_p}{\partial\beta} = -\frac{2\partial\ln t_p(\beta)}{\partial\beta} = -A, \tag{57}$$

where

$$A = \frac{2}{t_p(\beta)}\frac{\partial t_p(\beta)}{\partial\beta}. \tag{58}$$

The partial derivative $\partial t_p(\beta)/\partial\beta$ needs numerical evaluation. A handy formula for its computation is developed as follows. To calculate $\partial t_p/\partial\beta$, eq. (37) is rewritten as

$$F[p,t_p(\beta),\beta] = \Phi\left[\sqrt{2}\left(t_p-\beta/t_p\right)\right]-\exp(4\beta)\left\{1-\Phi\left[\sqrt{2}\left(t_p+\beta/t_p\right)\right]\right\}-p. \tag{59}$$

Then, the formula for $\partial t_p/\partial\beta$ can be expressed as the derivative of the implicit function F

$$\frac{\partial t_p(\beta)}{\partial\beta} = -\frac{\partial F/\partial\beta}{\partial F/\partial t_p}. \tag{60}$$

From (59) we have that the numerator is

$$\frac{\partial F}{\partial \beta} = \frac{\partial \Phi(u_1)}{\partial u_1}\left(-\frac{1}{t_p}\right) - 4\exp(4\beta)\left[1-\Phi(u_2)\right] + \exp(4\beta)\frac{\partial \Phi(u_2)}{\partial u_2}\left(\frac{1}{t_p}\right)$$
$$= -4\exp(4\beta)\left[1-\Phi(u_2)\right], \tag{61}$$

where

$$u_1 = \sqrt{2}\left(t_p - \frac{\beta}{t_p}\right) \quad \text{and} \quad u_2 = \sqrt{2}\left(t_p + \frac{\beta}{t_p}\right) \tag{62}$$

and the following equality holds:

$$\frac{\partial \Phi(u_2)}{\partial u_2}\exp(4\beta) = \frac{\partial \Phi(u_1)}{\partial u_1}. \tag{63}$$

The denominator can be also easily calculated:

$$\frac{\partial F}{\partial t_p} = \frac{\partial \Phi(u_1)}{\partial u_1}\sqrt{2}\left(1+\frac{\beta}{t_p^2}\right) + \exp(4\beta)\frac{\partial \Phi(u_2)}{\partial u_2}\sqrt{2}\left(1-\frac{\beta}{t_p^2}\right)$$
$$= \frac{2}{\sqrt{\pi}}\exp\left[-\left(t_p - \frac{\beta}{t_p}\right)^2\right]. \tag{64}$$

Substituting (64) and (61) into (60) we finally get

$$\frac{\partial t_p}{\partial \beta} = 2\sqrt{\pi}\exp(0.5u_2^2)\,\Phi(-u_2), \tag{65}$$

where u_2 is given in (62). Thus,

$$\frac{\partial t_p}{\partial \beta} = 2\sqrt{\pi}\exp\left(t_p + \frac{1}{t_p}\right)^2 \Phi\left[-\sqrt{2}\left(t_p + \frac{1}{t_p}\right)\right]. \tag{66}$$

This relationship is illustrated in Fig. 3.

Substitution of the terms of matrix (50) and eqs. (56)–(57) into eq. (55) yields

$$D^2(y_p) = \frac{1}{N}\xi^{MLM}(\beta, p), \tag{67}$$

where

$$\xi^{MLM}(\beta, p) = 2\beta^2A^2 + \frac{\beta A^2}{2} - 4\beta A + 2. \tag{68}$$

Then,

$$x_p^u = \exp\left[y_p + D(y_p)\right] \quad \text{and} \quad x_p^l = \exp\left[y_p - D(y_p)\right]. \tag{69}$$

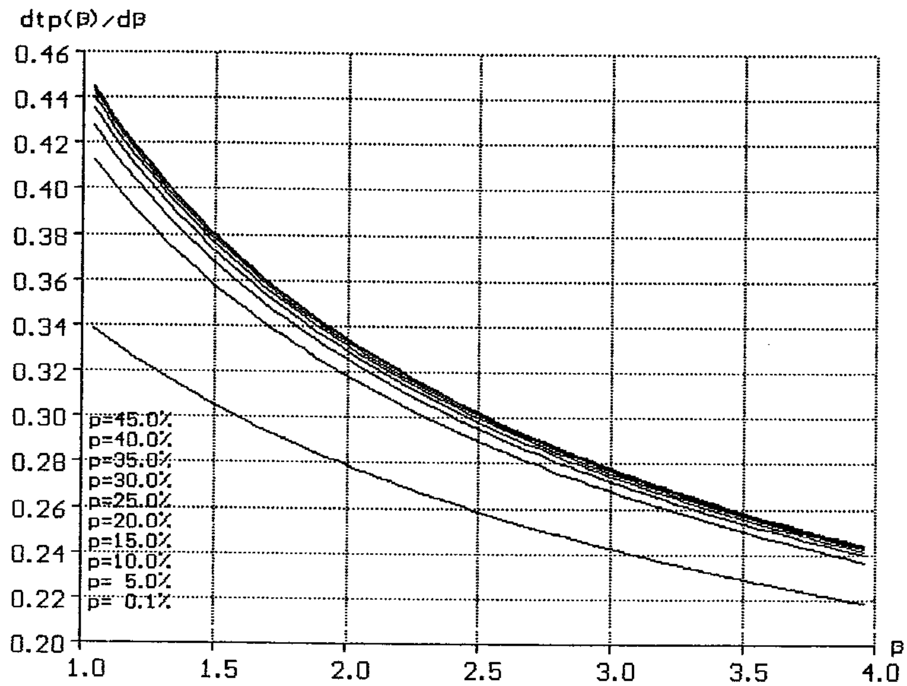

Fig. 3. Graphs of $dt_p(\beta)/d\beta$ as a function of β for 10 chosen values of probability of exceedance p (p increases upwards).

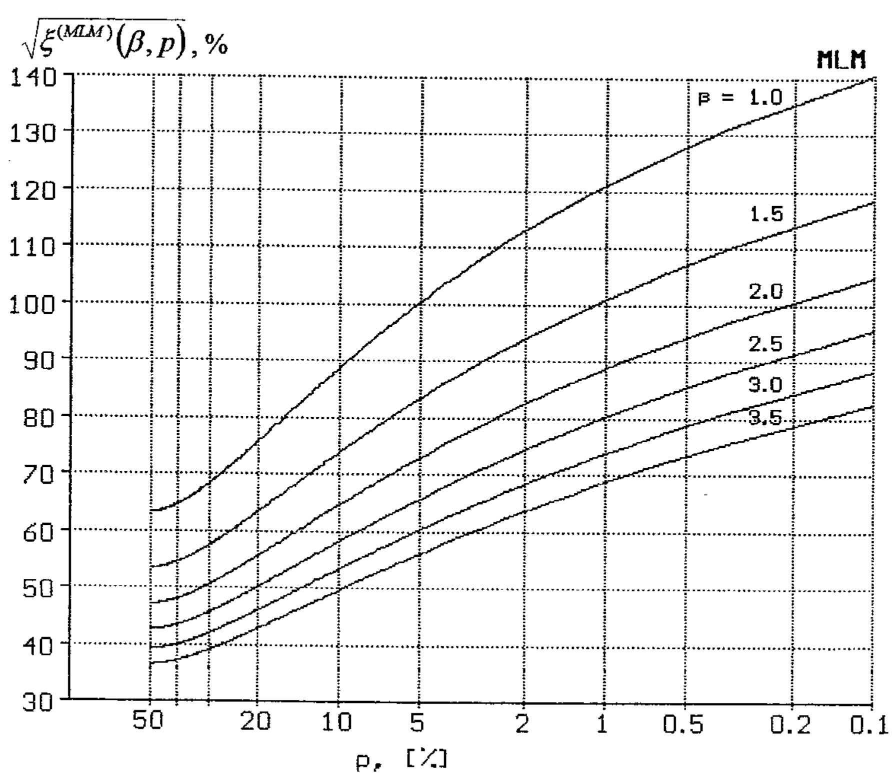

Fig. 4. Graphs of the MLM-estimated quantile relative error (cf., eq. 66) as a function of p for some values of β.

Table 2

MOM and MLM estimated quantile relative error $\sqrt{N}\ D(x_p)/x_p$ for selected values of β and probability of exceedance p

p [%]		β 0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
50	MOM	1.1271	0.7147	0.5692	0.4897	0.4373	0.3994	0.3701	0.3466	0.3272	0.3107
	MLM	0.8170	0.6326	0.5346	0.4714	0.4264	0.3922	0.3652	0.3430	0.3244	0.3086
40	MOM	1.0037	0.6769	0.5548	0.4841	0.4359	0.4001	0.3720	0.3491	0.3301	0.3139
	MLM	0.8375	0.6461	0.5450	0.4802	0.4341	0.3992	0.3715	0.3489	0.3300	0.3138
30	MOM	0.9434	0.6873	0.5758	0.5067	0.4580	0.4211	0.3919	0.3680	0.3480	0.3310
	MLM	0.8944	0.6848	0.5758	0.5065	0.4574	0.4203	0.3910	0.3671	0.3470	0.3300
20	MOM	1.0013	0.7644	0.6432	0.5654	0.5101	0.4682	0.4351	0.4079	0.3853	0.3661
	MLM	1.0013	0.7566	0.6333	0.5557	0.5012	0.4601	0.4278	0.4014	0.3794	0.3607
10	MOM	1.2634	0.9470	0.7851	0.6837	0.6128	0.5598	0.5183	0.4847	0.4568	0.4332
	MLM	1.1935	0.8862	0.7377	0.6458	0.5818	0.5337	0.4960	0.4653	0.4397	0.4180
5	MOM	1.5450	1.1238	0.9206	0.7963	0.7106	0.6472	0.5979	0.5583	0.5253	0.4976
	MLM	1.3631	1.0032	0.8337	0.7297	0.6574	0.6033	0.5607	0.5263	0.4974	0.4730
2	MOM	1.8572	1.3224	1.0748	0.9260	0.8242	0.7495	0.6917	0.6451	0.6067	0.5744
	MLM	1.5425	1.1325	0.9425	0.8263	0.7453	0.6849	0.6374	0.5986	0.5663	0.5388
1	MOM	2.0468	1.4464	1.1731	1.0099	0.8989	0.8172	0.7540	0.7032	0.6614	0.6262
	MLM	1.6500	1.2131	1.0119	0.8889	0.8034	0.7391	0.6885	0.6472	0.6127	0.5833
0.5	MOM	2.2019	1.5518	1.2585	1.0839	0.9650	0.8776	0.8099	0.7557	0.7109	0.6731
	MLM	1.7375	1.2817	1.0724	0.9444	0.8549	0.7876	0.7344	0.6912	0.6549	0.6238
0.2	MOM	2.3677	1.6684	1.3552	1.1685	1.0413	0.9479	0.8757	0.8177	0.7697	0.7291
	MLM	1.8310	1.3577	1.1410	1.0078	0.9145	0.8443	0.7887	0.7433	0.7050	0.6723
0.1	MOM	2.4703	1.7424	1.4178	1.2242	1.0925	0.9954	0.9202	0.8597	0.8097	0.7674
	MLM	1.8888	1.4060	1.1855	1.0498	0.9546	0.8826	0.8255	0.7787	0.7393	0.7054

From (54) and (67) we have that the quantile relative error, $D(x_p)/x_p$, is

$$\frac{D(x_p)}{x_p} = \frac{1}{\sqrt{N}}\sqrt{\xi^{(MLM)}(\beta, p)}\,. \tag{70}$$

The $\sqrt{\xi^{(MLM)}(\beta, p)}$ function versus p is presented in Fig. 4 for selected values of β. The values of this function can also be found in Table 2.

6. PARAMETER ESTIMATION BY THE METHOD OF MOMENTS

Solving eqs. (7) and (8) for parameters α and β, one gets

$$\alpha = \sqrt{\frac{(\mu_1')^3}{2\mu_2}} = \frac{1}{c_v}\sqrt{\frac{\mu_1'}{2}}, \tag{71}$$

$$\beta = \frac{1}{2c_v^2}. \tag{72}$$

Equations (71) and (72) are used in MOM to estimate parameters α and β from sample moments.

Error in quantiles with the mean (m) and variance (v) as parameters

For the asymptotic distribution of moments, the components of the variance -covariance matrix are (e.g., Kaczmarek, 1977, p.168-169):

$$D^2(m) = \frac{v}{N}, \tag{73}$$

$$D^2(v) = \frac{\mu_4 - \mu_2^2}{N}, \tag{74}$$

$$r_{m,v} D(m) D(v) = \frac{\mu_3}{N}. \tag{75}$$

Substitution of (9)–(10) into (74) and (75) gives the variance-covariance matrix

$$\mathbf{M}^{(MOM)}(m,v) = \left[r_{m,v}\, D(m)\, D(v) \right] = \begin{bmatrix} \frac{v}{N} & \frac{3v^2}{Nm} \\ \frac{3v^2}{Nm} & \frac{v^2}{N}\left(15c_v^2 + 2\right) \end{bmatrix}. \tag{76}$$

Hence,

$$r_{m,v}^{(MOM)} = \frac{3c_v}{\sqrt{15c_v^2 + 2}}. \tag{77}$$

The MOM-counterpart of (50) is

$$\mathbf{M}^{(MOM)}(\alpha,\beta)=\left[r_{\alpha,\beta}D(\alpha)\,D(\beta)\right]=\begin{bmatrix}\frac{\alpha^2}{4\beta N}(3+2\beta) & \frac{\alpha}{2N}(3+2\beta)\\ \frac{\alpha}{2N}(3+2\beta) & \frac{\beta}{2N}(7+4\beta)\end{bmatrix}. \tag{78}$$

Hence,

$$r_{\alpha,\beta}^{(MOM)}=\sqrt{\frac{6+4\beta}{7+4\beta}}=\sqrt{\frac{6c_v^2+2}{7c_v^2+2}}. \tag{79}$$

Comparing (53) and (77), one can see that the coefficient of correlation between m and v is greater for MLM than it is for MOM and it equals zero for $c_V = 0$, while for $c_V = 1$ it amounts to 0.90 and 0.73 for MLM and MOM, respectively.

A similar analysis with respect to the coefficient of correlation between the estimators α and β given by (78) and (51) shows that $r_{\alpha,\beta}{}^{(MLM)} \le r_{\alpha,\beta}{}^{(MOM)}$ and they equal one for $c_V = 0$, while for $c_V = 1$ they amount to 0.82 and 0.94 for MLM and MOM, respectively.

Asymptotic standard error of quantiles

Substituting (71) into (54), one gets

$$y_p=\ln x_p=3\ln m-\ln v-\ln 2-2\ln t_p(\beta), \tag{80}$$

where from (72)

$$\beta=\frac{m^2}{2v}. \tag{81}$$

The derivatives of y_p in respect to the mean and variance are

$$\frac{\partial y_p}{\partial m}=\frac{3}{m}-A\frac{\partial\beta}{\partial m}=\frac{3}{m}-A\frac{m}{v}, \tag{82}$$

$$\frac{\partial y_p}{\partial v}=-\frac{1}{v}-A\frac{\partial\beta}{\partial v}=-\frac{1}{v}+A\frac{m^2}{2v^2}, \tag{83}$$

where A is given by (58).

The variance of quantile y_p is

$$D^2(y_p)=D^2(m)\left(\frac{\partial y_p}{\partial m}\right)^2+D^2(v)\left(\frac{\partial y_p}{\partial v}\right)^2+2r_{m,v}D(m)D(v)\frac{\partial y_p}{\partial m}\frac{\partial y_p}{\partial v}. \tag{84}$$

Substituting (82) and (83) into (84) and the respective terms of matrix equation (76), one gets

$$D^2\left(y_p\right) = \frac{\xi^{(MOM)}(\beta, p)}{N}, \tag{85}$$

where

$$\xi^{(MOM)}(\beta, p) = \frac{1}{2\beta}\left[4A\beta^3 + A\beta^2(7A-8) - 4\beta(3A-1)+6\right]. \tag{86}$$

The asymptotic standard error of x_p, i.e., $D(x_p)$, is obtained by substituting (85) into (69). The asymptotic relative standard error, $D(x_p)/x_p$, can be expressed by (70), where $\xi^{(MLM)}(\beta, p)$ is replaced with $\xi^{(MOM)}(\beta, p)$. The $\sqrt{\xi^{(MOM)}(\beta, p)}$ function *versus p* is presented in Fig. 5 for selected values of β. Also, the values of this function are presented in Table 2. The relative efficiency of $\hat{x}_p^{(MOM)}$ in relation to the estimate $\hat{x}_p^{(MLM)}$ is displayed in Fig. 6. For upper quantiles, the efficiency of MOM-estimate decreases with decreasing p and β. For example, if $\beta = 1$ $(c_v = 0.71)$ and p = 0.1%, to achieve from MOM the accuracy of MLM, 54% larger sample would be necessary.

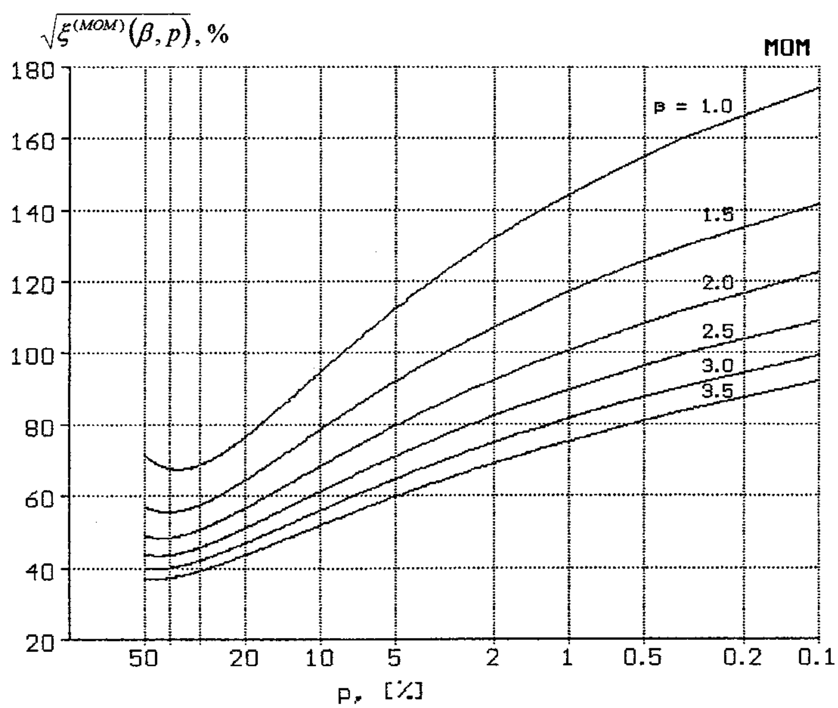

Fig. 5. Graphs of the MOM-estimated quantile relative error (cf., eq. 85) as a function of p for some values of β.

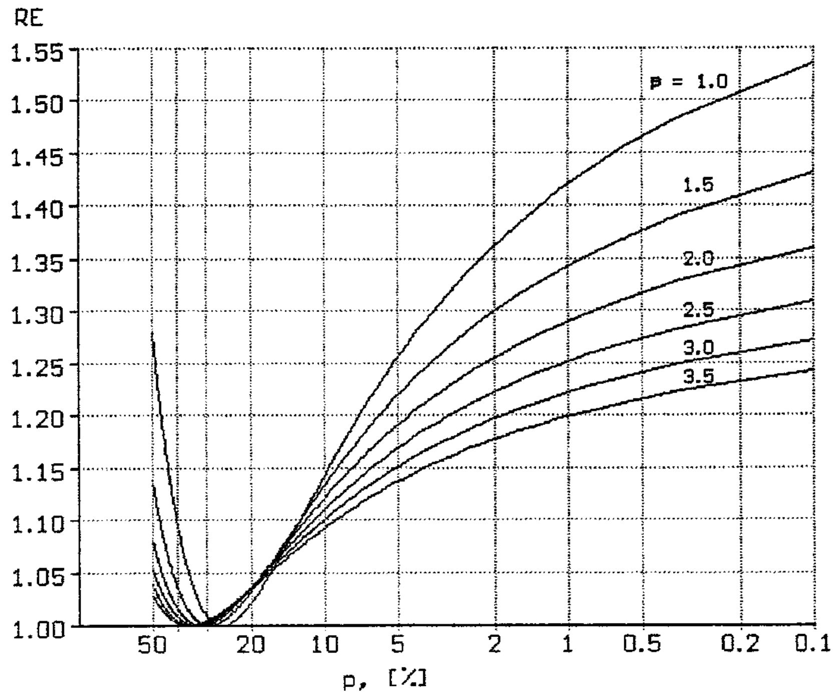

Fig. 6. Relative efficiency (RE) of MOM quantiles, $RE = \xi^{(MOM)}(\beta, p)/\xi^{(MLM)}(\beta, p)$, *vs.* probability of exceedance *p*.

7. CONCLUSIONS

The impulse response of the linearized diffusion wave model is a promising model for flood frequency analysis, being easy from computational point of view and especially attractive if one recognizes that the true distribution function is unknown. Estimates of CD-parameters from the moments and the maximum likelihood method are obtained analytically from simple algebraic equations. The method of the principle of maximum entropy is here equivalent to the maximum likelihood method. The linear moments have not been derived for the CD-distribution yet. The model can be easily extended to the three-parameter form by putting the lower bound as the third parameter.

A c k n o w l e d g m e n t. This study has been partly financed by the Polish Committee for Scientific Research, Grant No. 6 P4D 05617.

References

Cox, D. R., and H.D. Miller, 1965, *The theory of stochastic processes,* Chapman and Hall, London.

Deymie, P., 1939, *Propagation d'une intumescence allongee*, Proc. 5^{th} Int. Cong. Appl. Mech., New York, 537-544.

Dooge, J.C.I., and B.M. Harley, 1967, *Linear routing in uniform open channels*, Proc. Int. Hydrol. Symp. Fort Collins, Co. Sept.1967, Paper No. 8, 1, 57-63.

Dooge, J.C.I., 1973, *Linear theory of hydrologic systems*, Techn. Bull., 1468, Agricultural Research Service, Washington

Dooge, J.C.I., J.J. Napiórkowski and W.G. Strupczewski, 1987a, *The linear downstream response of a generalized uniform channel*, Acta Geophys. Pol. **35**, 277-291.

Dooge, J.C.I., J.J. Napiórkowski and W.G. Strupczewski,1987b, *Properties of the general downstream channel response*, Acta Geophys. Pol. **35**, 405-418.

Eagleson, P.S., 1972, *Dynamics of flood frequency*, Water Resour. Res. **8**, 878-898.

Folks, J.L., and R.S. Chhikara, 1978, *The inverse Gaussian distribution and its statistical application – a review,* J. R. Stat. Soc., Ser. B., **40**, 3, 263-289.

Hayami, S., 1951, *On the propagation of flood waves*, Kyoto Univ. Japan, Disaster Prevention Res. Inst. Bull., 1, 1-16.

Johnston, N.L., and S. Kotz, 1970, *Distribution in Statistics: Continuous Univariate Distributions 1,* Houghton-Mifflin, Boston.

Kaczmarek, Z., 1977, *Statistical Methods in Hydrology and Meteorology,* US Dept. Com., Springfield, Virginia.

Kendall, M.G., and A. Stuart, 1969, *The Advanced Theory of Statistics. Vol. 1: Distribution Theory*, Charles Griffin and Company Ltd., London, p.70.

Kendall, M.G., and A. Stuart, 1973, *The Advanced Theory of Statistics. Vol. 2: Inference and Relationship*, Charles Griffin and Company Ltd., London, pp.12, 26 and 67.

Masse, P., 1939, *Recherches sur la theorie des eaux courantes*, Proc. 5^{th} Int. Cong. Appl. Mech., New York, 545-549.

Moore, R.J., and R.T. Clarke, 1983, *A distributed function approach to modelling basin sediment yield*, J. Hydrol. **65**, 239-257.

Moore, R.J., 1984, *A dynamic model of basin sediment yield*, Water Resour. Res. **20**, 1, 89-103.

Rao, A. R., and K.H. Hamed, 2000, *Flood Frequency Analysis,* CRC Press, Boca Raton, Florida.

Schaefer, M. G., 1998, *General storm stochastic event flood model: Technical support manual*, Technical Report, MGS Eng. Consult. Inc., Olympia, Washington.

Singh, V.P., 1996, *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*, John Wiley and Sons, New York.

Singh, V.P., 1998, *Entropy-Based Parameter Estimation in Hydrology*, Kluwer Academic Publ., Dordrecht.

Stedinger, J.R., M.V. Vogel, and E. Foufoula-Georgiou, 1993, *Frequency analysis of extrema events*. **In**: D.R. Maidment (ed.), "Handbook of Hydrology", 18.1-18.66, McGraw-Hill Inc., New York.

Strupczewski, W.G., and J.J. Napiórkowski, 1990, *Linear flood routing model for rapid flow*, Hydrol. Sc. J. **35**, 1/2, 149-164.

Strupczewski, W.G., and J.C.I. Dooge, 1996, *Relationships between higher cumulants of channel response. II. Accuracy of linear interpolation*, Hydrol. Sc. J. **41**, 1, 61-73.

Strupczewski, W.G., V.P. Singh and S. Węglarczyk, 2002, *Asymptotic bias of estimation methods caused by the assumption of false probability distribution*, J. Hydrol. **258**, 1-4, 122-148.

Todorovic, P., 1982, *Stochastic modeling of floods*. **In**: V.P. Singh (ed.), "Statistical Analysis of Rainfall and Runoff", 597-636, Water Resour. Publ., Littleton, Colorado.

Tweedie, M.C.K., 1957, *Statistical properties of the inverse Gaussian distributions, I*, Ann. Math. Stat. **28**, 362-377.

Received 25 February 2002
Accepted 19 April 2002