ELSEVIER

Journal of Hydrology 270 (2003) 328–351

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Impulse response of the kinematic diffusion model as a probability distribution of hydrologic samples with zero values

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

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

Received 10 December 2001; revised 17 September 2002; accepted 4 October 2002

Abstract

It is hypothesized that the unit impulse response of a linearized kinematic diffusion (KD) model is a probability distribution suitable for frequency analysis of hydrologic samples with zero values. Such samples may include data on monthly precipitation in dry seasons, annual low flow, and annual maximum peak discharge observed in arid and semiarid regions. The hypothesized probability distribution has two parameters, which are estimated using the methods of moments (MOM) and maximum likelihood (MLM). Also estimated are errors in quantiles for MOM and MLM. The distribution shows an equivalency of MOM and MLM with respect to the mean value—an important property for ML-estimation in the case of the unknown true distribution function. The hypothesized KD distribution is tested on 44 discharge data series and compared with the Muskingum-like (M-like) probability distribution function. A comparison of empirical distribution with KD and M-like distributions shows that MOM better reproduces the upper tail of the distribution, while MLM is more robust for higher sample values and more conditioned on the value of the probability of the zero value event. The KD-model is suitable for frequency analysis of short samples with zero values and it is more universal than the M-like model as its modal value cannot be only equaled to zero value but also to any positive value.

Keywords: Probability distribution; Ephemeral streams; Frequency analysis; Hydrologic extremes; Kinematic diffusion; Parameter estimation; Time series

1. Introduction

There exists a multitude of models for flood frequency analysis (FFA). These models can be broadly classified into: (1) empirical, (2) phenomenological, and (3) physically based. An excellent discussion of empirical models is given by Stedinger et al. (1993) and Rao and Hamed (2000), among others. 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 these types 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 great deal of attention in the 1970s

* Corresponding author.
E-mail addresses: wgs@igf.edu.pl (W.G. Strupczewski), sweglar@lajkonik.wis.pk.edu.pl (S. Weglarczyk), cesing@lsu.edu (V.P. Singh).

0022-1694/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0022-1694(02)00309-8

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 one of the first to employ such a model. Despite their appeal, physically based models have yet to become models of choice in hydrologic practice. Another example of such a model is the use of watershed models, as for example, 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 the complete linearized Saint Venant equation and its simplifications provide a reasonable representation of the physics of channel flow. The connection between frequency analysis methods and deterministic methods is not clear yet. Our objective here is to show a connection through techniques of analysis, not through concept. It constituted the subject of this paper.

It is then hypothesized that impulse response function (IRF) of such models 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 is a purely deterministic function, the stochastic interpretation of the impulse response can be explained as follows.

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 $\int_0^T \times h(x,t)\mathrm{d}t$ determines the probability that a single particle passes the outlet at x during time $(0, T)$, where $h(x,t)$ is the IRF at time t and position x. Furthermore, the IRF fulfills several requirements normally expected of the flood frequency models, namely, (1) semi-infinite lower bounded range with a non-negative value at the lower bound; (2) positive skewness and the unit integral over whole range $\int_0^\infty \times h(x,t)\mathrm{d}t = 1$; 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 PDF in FFA.

Because of the practice of applying existing probability distributions in FFA, emphasis in 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.

In the frequency analyses of hydrologic data in arid and semiarid regions, one often encounters data series that contain several zero values while zero is the lower limit of the variability range. Examples of such data may include monthly precipitation in dry seasons, annual low flow in semi-arid and arid regions, and flood damage. Frequency analyses of such data have received relatively less attention. Jennings and Benson (1969), Woo and Wu (1989) and Wang and Singh (1995) among others developed empirical three-parameter models for frequency analysis of hydrologic data containing zero values. No phenomenological or physically based models seem to have been reported for such data in the hydrologic literature. Taking into account that short samples are very common in arid and semiarid regions and that the constraints with respect to the number of parameters are very rigid for hydrologic problems (e.g. Landwehr et al., 1980), even three parameters may be too many for hydrologic samples with zero values. Hence, the possibility of using two-parameter models is explored in this paper.

The main objective of this paper is to hypothesize a linear kinematic diffusion (KD) model of flow routing as a probability distribution function for modeling hydrologic samples with zero values and to evaluate the validity of this hypothesis. The paper is organized as follows. Introducing the objective of study in Section 1, a perspective on frequency modeling of hydrologic samples with zero values is provided in Section 2. The KD and rapid flow (RF) models are briefly revisited in Section 3. The PDF based on the KD model is presented in Section 4. Also presented in this section are the properties of this model-based PDF, the probability of exceedance, estimation of its parameters by MOM and MLM, accuracy of estimated parameters and asymptotic standard error of quantiles. The testing of the model and its comparison

with the Muskingum-like model is presented in Section 5. The paper is concluded in Section 6.

2. Modeling hydrologic samples with zero values

From the viewpoint of the probability theory, the occurrence of zero events can be expressed by placing a nonzero probability mass on a zero value, i.e. $P(X = 0) \neq 0$, where X is the random variable, and P is the probability mass. Therefore, the density functions from which such hydrologic series were drawn would be discontinuous with discontinuity at the zero value having a form:

$$f(x) = \beta\delta(x) + f_c(x;\mathbf{R}) \qquad x \geq 0 \tag{1}$$

where β denotes the probability of the zero event, i.e. $\beta = P(X = 0)$, $f_c(x;\mathbf{R})$ is the continuous function such that $\int_0^\infty f_c(x;\mathbf{R})\mathrm{d}x = 1 - \beta$, $\mathbf{R}$ is the vector of parameters, $\delta(x)$ is the Dirac delta function. Obviously, Eq. (1) violates the basic assumption of continuity made in conventional frequency analyses.

For estimation procedures for a hydrologic sample with zero events, the total probability axiom has been employed (Jennings and Benson, 1969; Woo and Wu, 1989; Wang and Singh, 1995). Then, Eq. (1) takes the form

$$f(x) = \beta\delta(x) + (1 - \beta)f_1(x;\mathbf{g})\cdot 1(x) \qquad \beta \notin \mathbf{g} \tag{1a}$$

where $f_1(x;\mathbf{g})$ is the conditional probability density function (CPDF), i.e. $f_1(x;\mathbf{g}) \equiv f(x|X > 0)$, which is continuous in the range $(0, +\infty)$ with a lower bound of zero value, and $\mathbf{g}$ is the vector of parameters. Wang and Singh (1995) estimated β and the parameters of CPDF separately considering the positive values as a full sample for the purpose. Having estimated $\mathbf{g}$ and β, the conditional distribution can be transformed to the marginal distribution, i.e. to $f(x)$, by Eq. (1a). Among several PDFs with lower bound at zero recognized in FFA (e.g. Rao and Hamed, 2000), the gamma distribution was chosen by Wang and Singh (1995) as an example of CPDF. They applied four estimation methods to this CPDF, i.e. the maximum likelihood method (MLM), method of moments (MOM), probability weighted moments (PWM) and ordinary least squares method.

In order to find the rationale for separately estimating the parameters, consider the likelihood function (L) of a sample drawn from the discontinuous distribution Eq. (1a):

$$L = \beta^{n_1}\cdot(1 - \beta)^{n_2}\prod_{j=1}^{n_2} f_1(x_j;\mathbf{g}) \qquad \beta \notin \mathbf{g} \tag{2}$$

where n_1 and n_2 denote the number of zeros and nonzeros values, respectively. From ML-equations, one can easily find that the ML-estimate of β is

$$\hat{\beta} = \frac{n_1}{n_1 + n_2} \tag{3}$$

i.e. β and $\mathbf{g}$ are estimated by MLM independently. Moreover, it should be noted that the β-estimate of the MOM differs from Eq. (3) and to get it by MOM the whole sample is to be used. It clearly shows that Wang and Singh (1995) applied the mixed method to the marginal distribution (1a) as they have used the estimate given by Eq. (3) in MOM. It reduces by one the number of moments or L-moments used for parameter estimation.

3. Kinematic diffusion and the rapid flow model

Physically based models of flow routing in open channels are based on the St. Venant equation or its simplifications. The complete linearized Saint Venant equation is of hyperbolic type and may be written as

$$a\frac{\partial^2 Q}{\partial x^2} + b\frac{\partial^2 Q}{\partial x\partial t} + c\frac{\partial^2 Q}{\partial t^2} = d\frac{\partial Q}{\partial x} + e\frac{\partial Q}{\partial t} \tag{4}$$

where Q is the perturbation of flow about an initial condition of steady state uniform flow, x is the distance from the upstream boundary, t is the elapsed time, and a, b, c, d and e are the parameters being functions of channel and flow characteristics at the reference steady state condition. Using the linearization of the Saint–Venant equation and assuming a semi-infinite channel, the solution of the upstream boundary problem was derived by Deymie (1939), Masse (1939), Dooge and Harley (1967) and Dooge et al. (1987a,b), among others; a discussion of this problem is presented in Singh (1996). The solution is a linear, physically based model with four parameters depending on 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 (Dooge et al., 1987a,b;

Singh, 1996). A number of models of simplified forms of the complete St. Venant equation have been proposed in the hydrologic literature.

If all three of the second-order terms on the left-hand side of Eq. (4) are neglected, the linear kinematic wave model is obtained. Expressing the second and the third second-order term in terms of the first on the basis of the linear kinematic wave approximation leads to the convective diffusion equation, i.e. the linear convective diffusion analogy model (CD). If the diffusion term is expressed in terms of two other terms using the kinematic wave solution, one gets the RF equation, which is of parabolic-like form (Strupczewski and Napiorkowski, 1990a). Therefore, it filters out the downstream boundary condition. It provides the exact solution for a Froude number equal to one and consequently can be used for river flows with large values of the Froude number. If the alternative approach is taken as one of expressing all the second-order terms as cross-derivatives, one gets the equation representing the diffusion of kinematic waves (Lighthill and Witham, 1955). The linear KD model, being of the parabolic-like form, fits satisfactorily the solution of the complete linearized Saint Venant equation only for small Froude number and slow rising waves (Strupczewski et al., 1989).

For flood routing, i.e. the prediction of flood characteristics at a downstream section on the basis of the knowledge of flow characteristics at an upstream section, two simpler forms of the linear channel downstream response are recognized in the hydrologic literature: (1) linear convective diffusion analogy model (CD), and (2) linear 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 Napiorkowski, 1990a,b).

Although RF and KD models correspond to quite different flow conditions, the structure of their impulse response is similar (Strupczewski et al., 1989; Strupczewski and Napiorkowski, 1986, 1989, 1990a,b). For both models, the impulse response is expressed as:

$$h(x,t) = P_0(\lambda)\delta(t-\Delta) + \sum_{i=1}^{\infty} P_i(\lambda)\cdot h_i\left(\frac{t-\Delta}{\alpha}\right)\cdot 1(t-\Delta) \tag{5}$$

where

$$P_i(\lambda) = \frac{\lambda^i}{i!}\exp(-\lambda) \tag{6}$$

is the Poisson distribution and

$$h_i\left(\frac{t}{\alpha}\right) = \frac{1}{\alpha(i-1)!}\left(\frac{t}{\alpha}\right)^{i-1}\exp\left(-\frac{t}{\alpha}\right) \tag{7}$$

is the gamma distribution and $1(t)$ is the unit step function. Parameters α, λ and Δ are functions of both channel geometry including the longitudinal variable (x) and flow conditions, which differ for the two models. Furthermore, there is no time lag (Δ) in the IRF of the KD model.

Strupczewski and Napiorkowski (1986, 1989) and Strupczewski et al. (1989) showed that for a finite river reach the infinite multiply Muskingum model with physically based parameters is equivalent to the KD model. Similarly, the RF model happens to be identical with the distributed delayed Muskingum model (Strupczewski and Napiorkowski, 1990b). Eq. (5) can be represented as a network of cascades of conceptual elements, namely, linear reservoirs and a linear channel commonly used in hydrology. The upstream boundary condition is delayed in the RF model by a linear channel with time lag, Δ, divided according to a Poisson distribution with mean λ, and then transformed by parallel cascades of equal linear reservoirs (with time constant α) of varying lengths. Note that λ is the average number of reservoirs in a cascade.

Einstein (1942) introduced the function given by Eq. (5) as the mixed deterministic-stochastic model for bed load transport. It has also been applied with $\Delta = 0$ as the probability distribution function of the total rainfall depth on the assumption that storm arrivals follow the Poisson process and storm depths the exponential distribution (Eagleson, 1978). Thus, the function in Eq. (5) is considered to be the flood frequency model in this study.

4. KD probability density function

Since our interest is in frequency estimation for samples lower-bounded at zero and containing zero values, the KD model is more adequate then the RF model, while the RF model can serve to model an

incomplete sample censored by the Δ-value. Therefore equating the delay (Δ) to zero in Eq. (5) and renaming t as x, one gets a two-parameter probability distribution function of the form:

$$f(x) = P(z=0)\cdot\delta(x) + \sum_{i=1}^{\infty} P(z=i)\cdot h_i\left(\frac{x}{\alpha}\right)\cdot 1(x) \quad (5a)$$

where z is the Poisson distributed random variable

$$P(z=i) \equiv P_i(\lambda) = \frac{\lambda^i}{i!}\exp(-\lambda) \quad (6a)$$

and x is the gamma distributed variable

$$h_i\left(\frac{x}{\alpha}\right) = \frac{1}{\alpha(i-1)!}\left(\frac{x}{\alpha}\right)^{i-1}\exp\left(-\frac{x}{\alpha}\right) \quad (7a)$$

and $1(x)$ is the unit step function. Note that Eq. (5a) differs from Eq. (1a) since its second term cannot be expressed as the product of the probability of non-zero value, i.e. $(1 - P_o(\lambda))$, and the CPDF, i.e. $f_1(x;\mathbf{g})$ with $\beta \notin \mathbf{g}$, where $\beta = \exp(-\lambda)$ while $\mathbf{g} = [\alpha, \beta]$ in the KD distribution function. Having in mind the doctrine of parameter parsimony, $\beta \in \mathbf{g}$ better serves the purpose of estimation of upper quantiles than $\beta \notin \mathbf{g}$ in Eq. (1a), which is focused on the estimation of the probability of non-occurrence of event (see Eq. (3)).

The second term of the PDF, i.e.

$$f_c(x) = \sum_{i=1}^{\infty} P_i(\lambda)\cdot h_i\left(\frac{x}{a}\right) \quad (8)$$

can be expressed by the 1st order modified Bessel function of the 1st type, $I_1(\cdot)$ (see Appendix A, Eq. (A1)):

$$f_c(x) = \exp\left(-\lambda - \frac{x}{\alpha}\right)\sqrt{\frac{\lambda}{\alpha x}}I_1\left(2\sqrt{\frac{\lambda x}{\alpha}}\right)\cdot 1(x). \quad (8a)$$

Thus, Eq. (5a) can be expressed as

$$f(x) = P_o(\lambda)\delta(x) + \exp\left(-\lambda - \frac{x}{\alpha}\right)\sqrt{\frac{\lambda}{\alpha x}}I_1 \times\left(2\sqrt{\frac{\lambda x}{\alpha}}\right)\cdot 1(x) \quad (5b)$$

which is the KD-PDF.

4.1. Properties of KD-PDF

It is interesting to look at some useful properties of the KD-PDF.

4.1.1. Modal value

The modal value can be obtained from the solution of $\partial f_c(x)/\partial x = 0$, or equivalently of $\partial f_c(y)/\partial y = 0$, where $y = x/\alpha$. It gives:

$$\frac{I_0(2\sqrt{\lambda y})}{I_1(2\sqrt{\lambda y})} = \sqrt{\frac{y}{\lambda}} + \sqrt{\frac{1}{\lambda y}} \quad (9)$$

for $\lambda \geq 2$, where $I_0(\cdot)$ is the 0th order modified Bessel function of the 1st kind. The modal value (y_{mod}) as the function of λ is shown in Fig. 1. As one can see, the maximum of $f_c(y)$ and consequently of $f_c(x)$ is for $\lambda < 2$ at the origin of y-axis.

4.1.2. Cumulants and moments

The Rth order cumulant of the KD model was expressed by Strupczewski and Napiorkowski (1989) as

$$k_R = R!\alpha^R\lambda. \quad (10)$$

Using the relations between moments and cumulants (Kendall and Stuart, 1969, p. 70) and Eq. (10) the expression for the first four moments of Eq. (5b) are given below:

$$\mu_1' = \alpha\lambda \quad (11)$$

$$\mu_2 = 2\alpha^2\lambda \quad (12)$$

$$\mu_3 = 6\alpha^3\lambda = \frac{3\mu_2^2}{2\mu_1'} \quad (13)$$

$$\mu_4 = k_4 + 3k_2^2 = 12\alpha^4\lambda(2+\lambda) = 3c_v^4(\mu_1')^4(c_v^2+1). \quad (14)$$

4.1.3. Dimensionless coefficients

For Eq. (5b), the coefficient of variation is

$$c_v = \sqrt{\frac{2}{\lambda}}. \quad (15)$$

The coefficient of skewness for the KD is:

$$c_s = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{3}{\sqrt{2\lambda}} = \frac{3}{2}c_v \quad (16)$$

while for the gamma distribution is $c_s = 2c_v$.

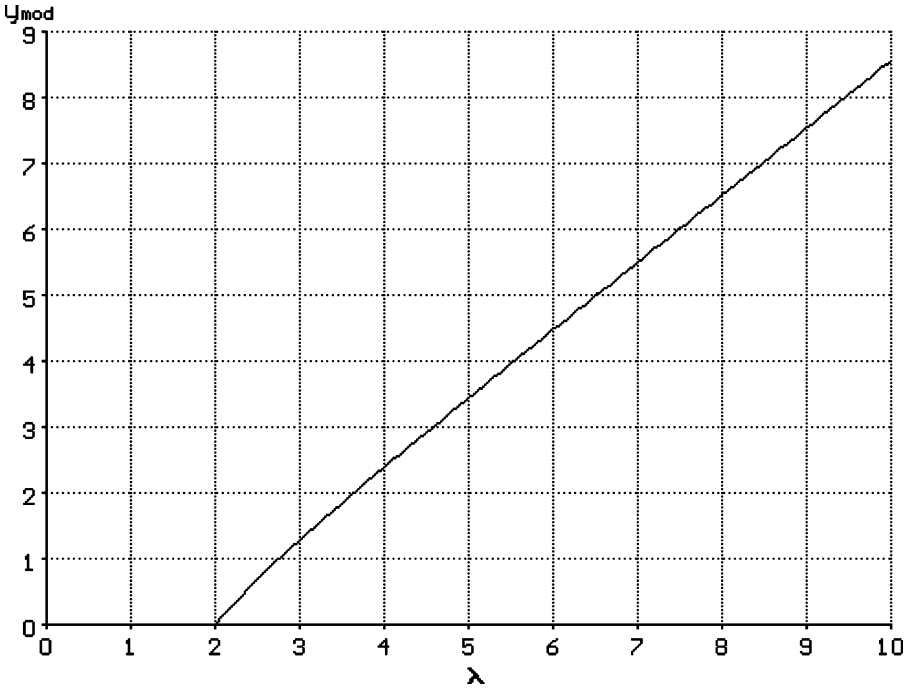


Fig. 1. The modal value ($y_{mod} = x_{mod}/\alpha$) vs the parameter λ.

The coefficient of kurtosis is

$$\text{Kurtosis} = \frac{\mu_4}{(\mu_2)^2} = 3\left(\frac{2}{\lambda} + 1\right) = 3(c_v^2 + 1). \quad (17)$$

4.1.4. Shape of the density function

For $c_v \to 0$ the distribution tends to be symmetric like lognormal (LN), convective diffusion (CD) and gamma distributions for $c_v \to 0$. Typical graphs of the distribution for some selected values of λ vs $y = x/\alpha$ are presented in Fig. 2. Note that α is the scale parameter. For increasing λ-value the maximum of $f(x)$ decreases and shifts along the y-axis. The value of λ defines from Eq. (6a) the probability of no-occurrence of the event, i.e. $P_o(\lambda)$.

4.1.5. Probability of exceedance

The quantile corresponding to the probability of exceedance p, x_p, is obtained by integrating Eq. (5b) as

$$x_p = \alpha \cdot t_p(\lambda) \quad (18)$$

where $t_p(\lambda)$ is the lower limit of the integral:

$$p = \int_{t_p}^{\infty} \exp(-\lambda - y)\sqrt{\frac{\lambda}{y}} I_1(2\sqrt{\lambda y})\mathrm{d}y \quad (19)$$

and p must be less than $1 - e^{-\lambda}$, otherwise $t_p = 0$. Some values of t_p for given λ and p are listed in Table 1.

4.2. Estimation of parameters by the maximum likelihood method

Let the sample contain n_1 zeros and n_2 positive values. The likelihood function L of the density function Eq. (5b) can be written as follows:

$$L = (e^{-\lambda})^{n_1} \prod_{j=1}^{n_2} \left(e^{-\lambda - x_j/\alpha} \sqrt{\frac{\lambda}{\alpha x_j}} I_1\left(2\sqrt{\frac{\lambda x_j}{\alpha}}\right)\right). \quad (20)$$

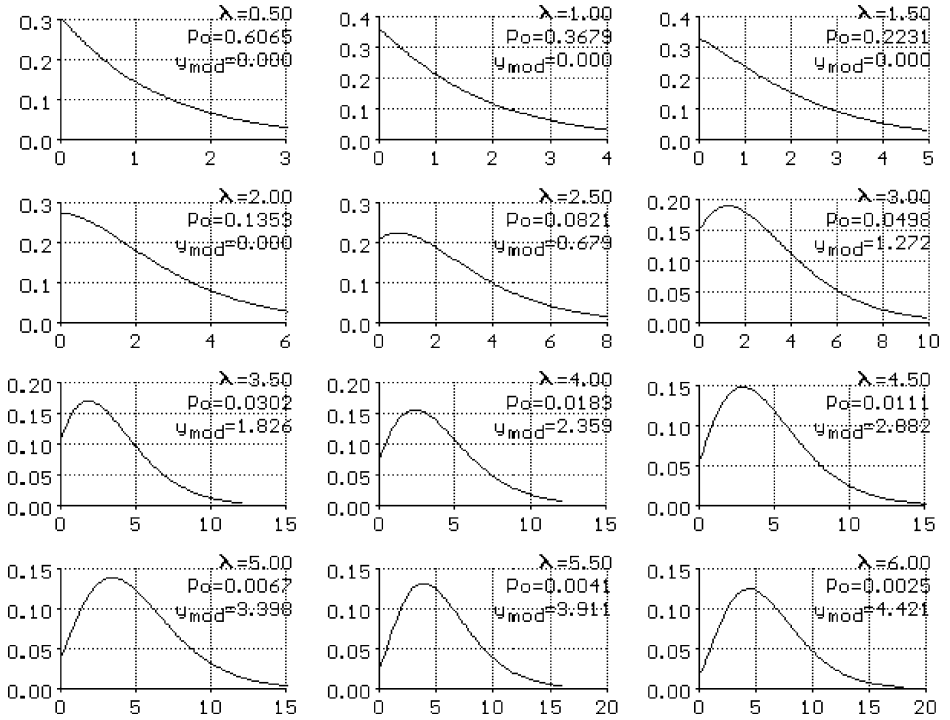


Fig. 2. Graphs of $f_c(x)$ (Eq. (8a)) for $\alpha = 1$ and some values of λ. P_0 denotes the probability of zero-value. $P_0 = e^{-\lambda}$ (see Eq. (6)).

So the log-likelihood function $\ln L$ will be:

$$\ln L = -\lambda n - \frac{1}{\alpha}\sum_{j=1}^{n_2} x_j + \frac{n_2}{2}\ln\lambda - \frac{n_2}{2}\ln\alpha - \frac{1}{2} \times \sum_{j=1}^{n_2}\ln x_j + \sum_{j=1}^{n_2}\ln I_1\left(2\sqrt{\frac{\lambda x_j}{\alpha}}\right). \quad (21)$$

Denoting for convenience

$$z_j = 2\sqrt{\frac{\lambda x_j}{\alpha}}. \quad (22)$$

Table 1
KD quantile $t_p(\lambda)$ for given values of λ and probability of exceedance p

p (%)	λ											
	0.5	1	1.5	2	3	4	5	6	7	8	9	10
50	0.000	0.396	0.950	1.469	2.483	3.487	4.490	5.492	6.493	7.494	8.495	9.496
40	0.000	0.764	1.394	1.980	3.111	4.215	5.305	6.385	7.458	8.524	9.588	10.65
30	0.350	1.226	1.939	2.599	3.857	5.065	6.248	7.410	8.559	9.698	10.83	11.95
20	0.871	1.861	2.673	3.420	4.823	6.156	7.446	8.706	9.945	11.17	12.37	13.57
10	1.749	2.906	3.860	4.728	6.333	7.836	9.275	10.67	12.03	13.37	14.68	15.98
5	2.616	3.918	4.989	5.956	7.729	9.371	10.93	12.44	13.90	15.33	16.74	18.12
2	3.750	5.216	6.420	7.500	9.460	1.259	12.96	14.59	16.17	17.71	19.21	20.69
1	4.597	6.178	7.468	8.622	10.71	12.61	14.41	16.12	17.77	19.38	20.95	22.49
0.5	5.439	7.122	8.494	9.716	11.91	13.91	15.79	17.58	19.31	20.98	22.61	24.21
0.2	6.542	8.352	9.820	11.13	13.46	15.58	17.56	19.44	21.25	23.00	24.71	26.38
0.1	7.372	9.269	10.81	12.17	14.60	16.80	18.85	20.80	22.67	24.47	26.23	27.95

One obtains the MLM equations:

$$\frac{\partial \ln L}{\partial \lambda} = -n + \frac{n_2}{2\lambda} + \frac{1}{2\lambda}\sum_{j=1}^{n_2} z_j \frac{\partial \ln I_1(z)}{\partial z}\bigg|_{z=z_j} = 0 \tag{23}$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{1}{\alpha^2}\sum_{j=1}^{n_2} x_j - \frac{n_2}{2\alpha} - \frac{1}{2\alpha}\sum_{j=1}^{n_2} \times z_j \frac{\partial \ln I_1(z)}{\partial z}\bigg|_{z=z_j} = 0. \tag{24}$$

Denoting

$$B = \sum_{j=1}^{n_2} B_j = \sum_{j=1}^{n_2} z_j \frac{\partial \ln I_1(z)}{\partial z}\bigg|_{z=z_j} \tag{25}$$

one obtains after a little manipulation:

$$-2n\lambda + n_2 + B = 0 \tag{26}$$

$$\frac{2}{\alpha}\sum_{j=1}^{n_2} x_j - n_2 - B = 0. \tag{27}$$

Adding Eqs. (26) and (27), one can eliminate B:

$$\lambda\alpha = \frac{1}{n}\sum_{j=1}^{n_2} x_j = \bar{x} \tag{28}$$

which means that this MLM equation is the same as its MOM equivalent given by Eq. (11).

The second MLM equation is different and its solution requires some numerical manipulation. Using Eq. (A4), one gets B as

$$B = \sum_{j=1}^{n_2} z_j \frac{I_0(z_j)}{I_1(z_j)} - n_2. \tag{25a}$$

Then, substituting it into Eq. (26) we have

$$\sum_{j=1}^{n_2} \frac{z_j I_0(z_j)}{I_1(z_j)} - 2n\lambda = 0 \tag{29}$$

where

$$z_j = 2\sqrt{\frac{\lambda x_j}{\alpha}} = 2\lambda\sqrt{\frac{x_j}{\bar{x}}}. \tag{22a}$$

For an iterative procedure of solution one can get an initial value by combining Eqs. (3) and (6a) for $i = 0$:

$$\hat{\lambda} = -\ln(n_1/n) \tag{30}$$

or use the MOM estimate of λ from Eq. (47).

4.2.1. Accuracy of estimated parameters

To get the asymptotic variance of quantiles the asymptotic variance–covariance matrix of the KD-parameters should be derived as the inverse of the expected information matrix (e.g. Kendall and Stuart, 1973):

$$[\mathrm{cov}(\lambda, \alpha)] = \left[-E\left(\frac{\partial^2 \ln L}{\partial\lambda\partial\alpha}\right)\right]^{-1} \tag{31}$$

where E represents the expected values.

Taking the second-order derivatives of $\ln L$ function given by Eq. (21), one obtains

$$\left[-\frac{\partial^2 \ln L}{\partial\lambda\partial\alpha}\right] = \begin{bmatrix} \frac{n}{4\lambda^2}(A+2\lambda) & -\frac{n}{4\alpha\lambda}(A-2\lambda) \\ -\frac{n}{4\alpha\lambda}(A-2\lambda) & \frac{n}{4\alpha^2}(A+2\lambda) \end{bmatrix} \tag{32}$$

where

$$A = \frac{n_2}{n} - \frac{D}{n} \tag{33}$$

while

$$D = \sum_{j=1}^{n_2} D_j = \sum_{j=1}^{n_2} z_j^2 \frac{\partial^2 \ln I_1(z)}{\partial z^2}\bigg|_{z=z_j}. \tag{34}$$

The inverse of the symmetric matrix Eq. (32) is:

$$\left[-\frac{\partial^2 \ln L}{\partial\lambda\partial\alpha}\right]^{-1} = \begin{bmatrix} \frac{(A+2\lambda)\lambda}{2An} & \frac{(A-2\lambda)\alpha}{2An} \\ \frac{(A-2\lambda)\alpha}{2An} & \frac{(A+2\lambda)\alpha^2}{2A\lambda n} \end{bmatrix}. \tag{35}$$

The asymptotic value of A, i.e. $E(A)$, has been derived in Appendix B. It is given by Eq. (B5) and shown in Fig. B1(b) for various values of λ. Substituting it into matrix (35), one can get the asymptotic variance–covariance matrix of

the KD-parameters, Eq. (31) as:

$$[\mathrm{cov}(\lambda,\alpha)] = E\begin{bmatrix} \dfrac{(A+2\lambda)\lambda}{2An} & \dfrac{(A-2\lambda)\alpha}{2An} \\ \dfrac{(A-2\lambda)\alpha}{2An} & \dfrac{(A+2\lambda)\alpha^2}{2A\lambda n} \end{bmatrix}. \quad (36)$$

In particular, the asymptotic coefficient of correlation of ML estimators of λ and α is:

$$r_{\lambda,\alpha}^{(\mathrm{MLM})} = \frac{A-2\lambda}{A+2\lambda}. \quad (37)$$

4.2.2. Asymptotic standard errors of quantiles

To derive the asymptotic error of quantiles defined by Eq. (18), the logarithmic transformation of Eq. (18) is obtained as

$$y_p = \ln x_p = \ln\alpha + \ln t_p(\lambda) \quad (38)$$

then

$$D^2(y_p) = D^2(\lambda)\left(\frac{\partial y_p}{\partial\lambda}\right)_m^2 + D^2(\alpha) \times\left(\frac{\partial y_p}{\partial\alpha}\right)_m^2 + 2r_{\lambda,\alpha}D(\lambda)D(\alpha) \times\left(\frac{\partial y_p}{\partial\lambda}\right)_m\left(\frac{\partial y_p}{\partial\alpha}\right)_m \quad (39)$$

where $D(\lambda)$ and $D(\alpha)$ are elements of the covariance matrix (36) while from Eq. (38)

$$\frac{\partial y_p}{\partial\lambda} \equiv G = \frac{\partial \ln t_p(\lambda)}{\partial\lambda} \quad (40)$$

$$\frac{\partial y_p}{\partial\alpha} = \frac{1}{\alpha} \quad (41)$$

and the index m indicates that the estimators of λ and α should be replace by their mean values in the partial derivatives.

The function (40) is derived in Appendix B, given by Eq. (B8) and displayed in Fig. B2 for various values of p and λ. Substitution of the terms of matrix equations (36), (40) and (41) into Eq. (39) yields

$$D^2(y_p) = \frac{1}{n}\xi^{(\mathrm{MLM})}(\lambda,p) \quad (42)$$

where

$$\xi^{(\mathrm{MLM})}(\lambda,p) = \frac{1}{2A}\left[\frac{A}{\lambda}(\lambda\cdot G+1)^2 + 2(\lambda\cdot G-1)^2\right] \quad (43)$$

and A is the function of λ defined by Eq. (B5). Then,

$$x_p^{\mathrm{u}} = \exp[y_p + D(y_p)] \text{ and } x_p^{\mathrm{l}} = \exp[y_p - D(y_p)]. \quad (44)$$

From Eqs. (38) and (42), one obtains the quantile relative error, $D(x_p)/x_p$, is

$$\frac{D(x_p)}{x_p} = \frac{1}{\sqrt{n}}\sqrt{\xi^{(\mathrm{MLM})}(\lambda,p)}. \quad (45)$$

The $\sqrt{\xi^{(\mathrm{MLM})}(\lambda,p)}$ function vs p is presented in Fig. 3 for selected values of λ. The values of this function can also be found in Table 2.

4.3. Parameter estimation by the method of moments

Solving Eqs. (11) and (12) for parameters α and λ, one gets

$$\alpha = \frac{\mu_2}{2\mu_1'} = \frac{\mu_1'}{2}c_{\mathrm{v}}^2 \quad (46)$$

$$\lambda = \frac{2(\mu_1')^2}{\mu_2} = \frac{2}{c_{\mathrm{v}}^2}. \quad (47)$$

Eqs. (46) and (47) are used in MOM to estimate parameters α and λ from sample moments. Note the equivalency with MLM-estimate of the mean given by Eq. (28).

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

The variances and covariances of the moments m and v are given among others by Kendall and Stuart (1969), Kaczmarek (1977) and Kite (1988). Substituting into them the KD-moments given by Eqs. (12)–(14), one can get the variance–covariance matrix in the form

$$[\mathrm{cov}(m,v)] = \begin{bmatrix} \dfrac{\mu_2}{n} & \dfrac{\mu_3}{n} \\ \dfrac{\mu_3}{n} & \dfrac{\mu_4-\mu_2^2}{n} \end{bmatrix} = \begin{bmatrix} \dfrac{v}{n} & \dfrac{3v^2}{2mn} \\ \dfrac{3v^2}{2mn} & \dfrac{v^2}{n}\left(\dfrac{3v}{m^2}+2\right) \end{bmatrix}. \quad (48)$$

Hence, the coefficient of correlation of MOM

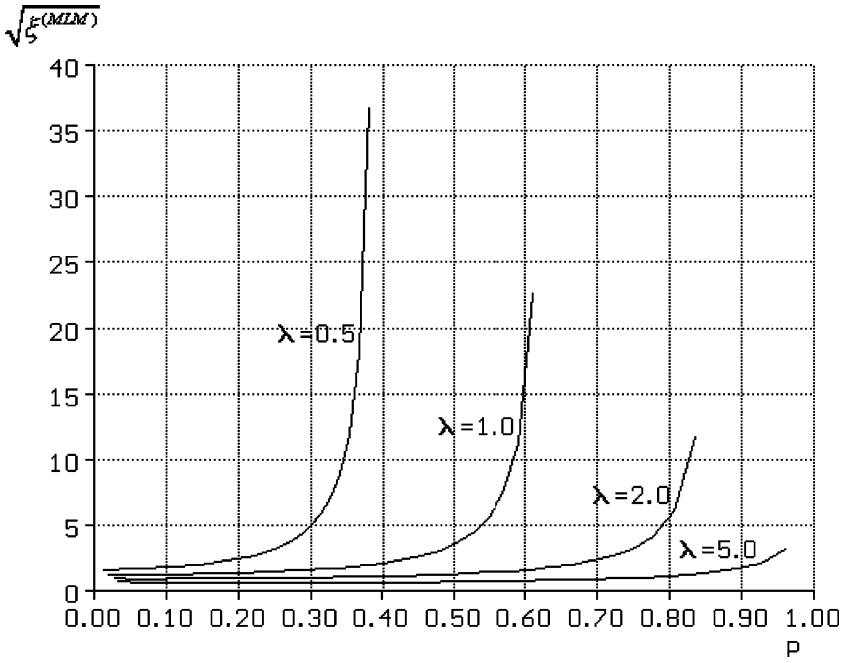


Fig. 3. Graphs of the MLM-estimated quantile relative error $\sqrt{nD(x_p)}/x_p = \sqrt{\xi^{(\mathrm{MLM})}(\lambda, p)}$ (c.f. Eq. (43)) as function of p for some values of λ.

Table 2
MOM and MLM estimated quantile relative error $\sqrt{nD(x_p)}/x_p = \sqrt{\xi(\lambda, p)}$ for selected values of λ and probability of exceedance p

p (%)		λ											
		0.5	1	1.5	2	3	4	5	6	7	8	9	10
50	MOM	–	4.673	1.929	1.363	0.968	0.795	0.692	0.621	0.568	0.527	0.494	0.466
	MLM	–	3.566	1.700	1.271	0.939	0.782	0.685	0.617	0.565	0.525	0.492	0.465
40	MOM	–	2.406	1.419	1.116	0.855	0.724	0.640	0.581	0.536	0.500	0.471	0.446
	MLM	–	2.116	1.362	1.097	0.851	0.723	0.640	0.581	0.536	0.500	0.471	0.446
30	MOM	6.221	1.642	1.188	0.997	0.801	0.692	0.620	0.567	0.526	0.493	0.466	0.443
	MLM	4.874	1.603	1.186	0.997	0.800	0.690	0.618	0.565	0.524	0.491	0.464	0.441
20	MOM	2.531	1.371	1.102	0.958	0.792	0.694	0.626	0.576	0.537	0.506	0.479	0.457
	MLM	2.431	1.367	1.087	0.940	0.774	0.679	0.615	0.567	0.529	0.499	0.473	0.451
10	MOM	1.851	1.326	1.110	0.981	0.825	0.730	0.663	0.614	0.575	0.543	0.516	0.493
	MLM	1.811	1.247	1.033	0.912	0.772	0.690	0.633	0.590	0.555	0.527	0.502	0.481
5	MOM	1.835	1.363	1.151	1.022	0.866	0.771	0.704	0.654	0.614	0.581	0.554	0.530
	MLM	1.673	1.210	1.019	0.909	0.784	0.710	0.658	0.618	0.585	0.557	0.533	0.512
2	MOM	1.893	1.417	1.203	1.073	0.916	0.819	0.752	0.701	0.660	0.627	0.598	0.574
	MLM	1.612	1.193	1.017	0.916	0.803	0.737	0.690	0.652	0.621	0.594	0.571	0.550
1	MOM	1.934	1.452	1.236	1.106	0.948	0.851	0.783	0.731	0.690	0.656	0.628	0.603
	MLM	1.593	1.189	1.020	0.924	0.817	0.755	0.711	0.675	0.645	0.619	0.595	0.575
0.5	MOM	1.969	1.482	1.265	1.135	0.976	0.879	0.810	0.758	0.717	0.683	0.654	0.629
	MLM	1.583	1.188	1.023	0.931	0.830	0.772	0.730	0.696	0.666	0.641	0.618	0.598
0.2	MOM	2.007	1.515	1.297	1.167	1.008	0.911	0.842	0.789	0.748	0.713	0.684	0.659
	MLM	1.575	1.189	1.029	0.941	0.846	0.792	0.753	0.720	0.692	0.666	0.644	0.624
0.1	MOM	2.031	1.536	1.318	1.188	1.029	0.932	0.863	0.810	0.768	0.734	0.704	0.679
	MLM	1.572	1.190	1.033	0.948	0.857	0.805	0.768	0.736	0.708	0.684	0.662	0.642

estimators of m and v is

$$r_{m,v}^{(\mathrm{MOM})} = \frac{3c_v}{2\sqrt{3c_v^2+2}}. \tag{49}$$

A high value of c_v leads to a high value of r between m and v. For $c_v = 1$, the coefficient of correlation equals 0.67.

Substituting Eqs. (46) and (47) into Eq. (42), one gets

$$y_p = \ln x_p = \ln v - \ln m + \ln t_p(\lambda) - \ln 2 \tag{50}$$

where

$$\lambda = \frac{2m^2}{v}. \tag{47a}$$

The derivatives of y_p with respect to the mean and variance are:

$$\frac{\partial y_p}{\partial m} = -\frac{1}{m} + G\frac{\partial \lambda}{\partial m} = -\frac{1}{m} + \frac{4m}{v}G \tag{51}$$

$$\frac{\partial y_p}{\partial v} = \frac{1}{v} + G\frac{\partial \lambda}{\partial v} = \frac{1}{v} - \frac{2m^2}{v^2}G \tag{52}$$

where G is given by Eq. (B8).

The variance of quantile y_p equals:

$$D^2(y_p) = D^2(m)\left(\frac{\partial y_p}{\partial m}\right)_m^2 + D^2(v) \times \left(\frac{\partial y_p}{\partial v}\right)_m^2 + 2r_{m,v}D(m)D(v) \times \left(\frac{\partial y_p}{\partial m}\right)_m\left(\frac{\partial y_p}{\partial v}\right)_m \tag{53}$$

where index m indicates that the partial derivatives are evaluated at $(m, v) = (E(\hat{m}), E(\hat{v}))$.

Substituting Eqs. (51) and (52) into Eq. (53) and the respective terms of matrix of Eq. (48), one gets

$$D^2(y_p) = \frac{\xi^{(\mathrm{MOM})}(\lambda, p)}{n} \tag{54}$$

where

$$\xi^{(\mathrm{MOM})}(\lambda, p) = \frac{2}{\lambda}[(\lambda+1)(\lambda G - 1)^2 + \lambda G] \tag{55}$$

where $G = G(\lambda, p)$ is given by Eq. (B8).

The asymptotic relative standard error, $D(x_p)/x_p$, can be expressed by Eq. (45) where $\xi^{(\mathrm{MLM})}(\lambda, p)$ is replaced with $\xi^{(\mathrm{MOM})}(\lambda, p)$. The $\sqrt{\xi^{(\mathrm{MOM})}(\lambda, p)}$ function vs p is presented in Fig. 4 for selected values of λ. Also, the values of this function are presented in Table 2. The relative efficiency (RE) of MOM is shown in Fig. 5.

5. Application of KD model

Annual peak flow discharge data for 44 flow gauging stations located in arid regions of the United States were selected from the USGS data bank to evaluate the performance of the KD-model. Some basic characteristics of the data and the river regime, such as drainage area, length of records (n), number of zero values (n_1), average peak flow [$E(x)$], coefficient of variation (c_v), and skewness coefficient (c_s) are given for each gauging station in Table 3. The skewness of coefficient was plotted against the coefficient of variation as shown in Fig. 8. This scatter diagram indicates a positive correlation of the two moment characteristics. Obviously, the values of the both characteristics for ephemeral streams are much greater than for perennial rivers.

5.1. Comparison of parameter estimates by MOM and MLM

Using MOM equations (46) and (47) and MLM equations (28) and (29), the values of parameters λ and α for the KD-distribution were calculated as given in Table 4. With the exception of Series No. 9, MOM resulted in lower values of λ and greater values of α than did MLM. Consequently, the MOM gave greater values of the upper quantiles estimates than did MLM as shown in Column 8 by the ratio of the respective estimates of $x_{1\%}$. Recognizing the differences between the MOM-estimates (Columns 4 and 5) and MLM-estimates (Columns 6 and 7), and noting from Eq. (6a) that

$$P(x = 0) = \exp(-\lambda) \tag{56}$$

one can deduce that the empirical probability of zero values (Column 3) was better reproduced by MLM than was by MOM. In fact, MLM was more robust for higher sample values than was MOM and more conditioned on the value of $P(x = 0) = n_1/n$. It is worth remembering that for a probability distribution of the form in Eq. (1a) MLM exactly

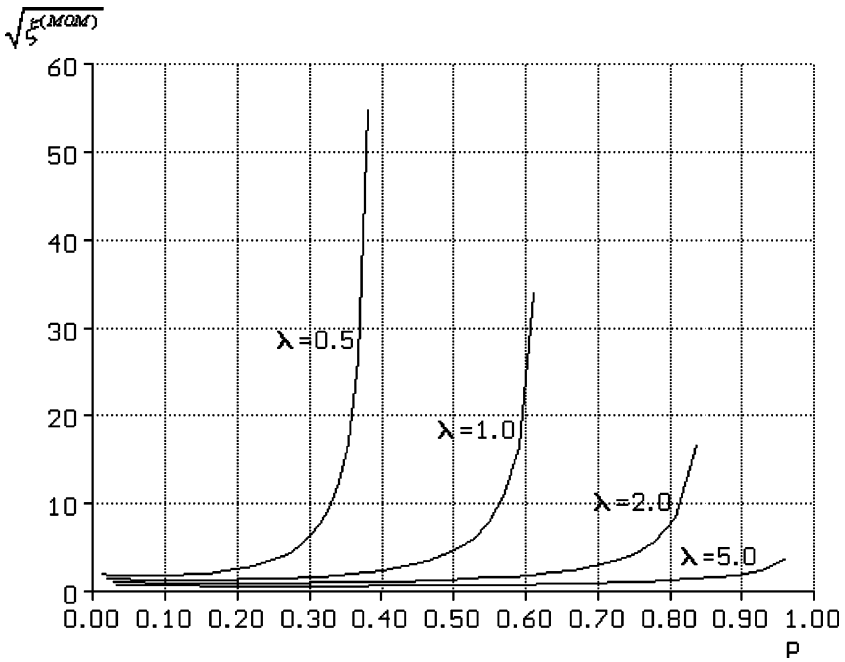


Fig. 4. Graphs of the MOM-estimated quantile relative error $\sqrt{n}D(x_p)/x_p = \sqrt{\xi^{(\mathrm{MOM})}(\lambda, p)}$ (c.f. Eq. (55) as function of p for some values of λ.

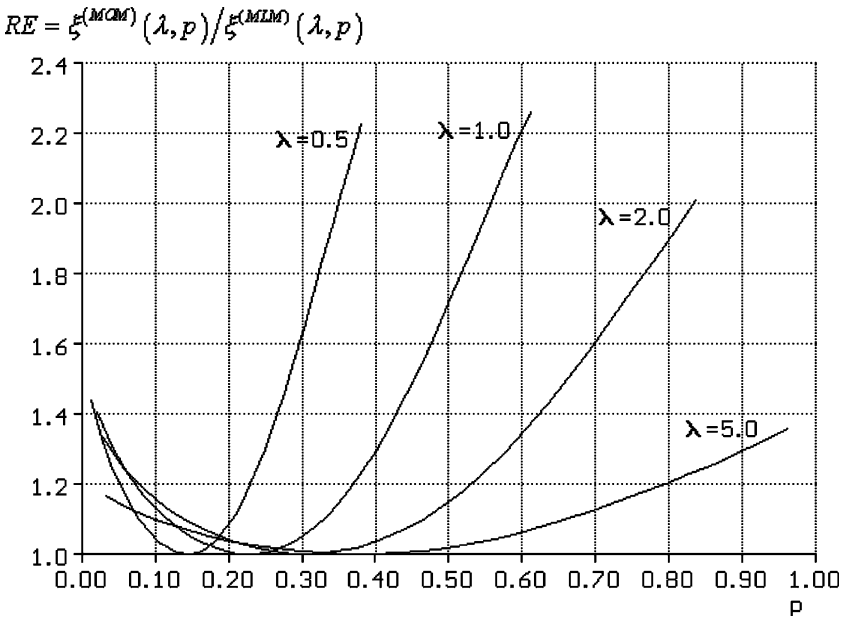


Fig. 5. Relative efficiency RE $= \xi^{(\mathrm{MOM})}(\lambda, p)/\xi^{(\mathrm{MLM})}(\lambda, p)$ of MOM quantiles, vs probability of exceedance p.

Table 3
Relevant information on data used in study

Series No.	USGS code	Location of Gauge, State	River or Creek	Drainage area (km^2)	Record length n (yr)	No. of zeros n_1	$E(x)$ (m^3/s)	c_v	c_s
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	11063680	nr San Bernardino, CA	Devil cyn Ck	14.2	79	2	7.118	2.376	4.511
2	11028500	nr Ramona, CA	Santa Maria Ck	149.0	61	4	28.99	2.257	4.283
3	11042000	Oceanside, CA	San Luis Rey Rv	1441	70	18	116.5	3.008	5.951
4	11012500	nr Campo, CA	Campo Ck	219.8	64	7	5.110	1.690	2.280
5	11012000	nr Dulzura, CA	Cottonwood Ck abv Tecate	801.8	64	5	24.58	2.568	3.750
6	10259300	Indio, CA	Whitewater Rv	2775.0	32	4	58.69	1.742	1.993
7	10259200	nr Palm Desert, CA	Deep Ck	79.1	39	1	23.74	1.717	3.065
8	10258500	nr Palm Springs, CA	Palm Cyn Ck	240.8	66	2	30.54	1.352	1.860
9	10264675	Edwards Afb, CA	Rogers Lk trib	4.5	11	2	0.127	0.866	0.53343
10	11042631	Temecula, CA	Pechanga Ck	35.7	11	1	10.78	2.307	2.718
11	11046250	San Onofre, CA	San Onofre Ck	109.1	23	8	17.00	1.415	1.260
12	08404000	Blv Avalon Dam, NM	Pecos R	46,761	48	9	149.1	2.021	3.088
13	09520500	Dome, AZ	Gila Rv	149,622	96	10	367.3	2.231	3.782
14	09519800	Painted Rock Dam, AZ	Gila Rv	131,672	40	2	59.11	2.487	4.826
15	09513860	nr Phoenix, AZ	Skunk Ck	167.1	39	2	71.85	1.220	1.363
16	09484500	Tucson, AZ	Tanque Verde Ck	566.4	31	1	132.9	1.119	1.888
17	09502000	Stewart Mountain, AZ	Salt Rv	16,118	65	1	190.7	1.789	3.209
18	09517000	nr Arlington, AZ	Hassayampa Rv	3802.0	37	1	139.0	1.463	3.161
19	13132500	nr Arco, ID	Big Lost Rv	3647	51	5	17.60	1.009	1.000
20	06713000	Cherry Creek Lake, CO	Cherry Ck	995.8	50	8	11.07	1.061	1.374
21	07126140	nr Tyrone, CO	Van Bremer Arroyo	341.4	17	3	3.819	1.161	0.9372
22	07126480	At mouth nr Timpas, CO	Bent Canyon Ck	145.4	14	1	12.50	1.854	2.022
23	06846500	Cedar Bluffs, KS	Beaver Ck	4185	56	1	21.90	1.711	3.605
24	06846000	Ludell, Ks	Beaver Ck	3649	45	1	20.31	1.170	2.105
25	06866900	nr Wakeeney, KS	Saline Rv	1800	31	1	72.08	1.416	1.658
26	06870300	nr Gypsum, KS	Gypsum Ck	310.4	40	1	75.06	0.9211	1.505
27	06871800	Kirwin, KS	Nf Solomon R v	3536	70	1	87.11	2.530	4.596
28	06873200	Bl Webster Re, KS	Sf Solomon Rv	2974	44	3	10.32	1.249	2.280
29	07139000	Garden City	Arkansas Rv	70,016	78	1	179.4	2.534	6.051
30	07139500	Dodge City, KS	Arkansas Rv	79,143	59	3	123.8	2.640	5.316
31	07140850	nr Burdett, KS	Pawnee Rv	2822	19	2	21.70	1.235	2.617
32	07141200	Rozel, KS	Pawnee Rv	5556	76	1	86.79	0.9869	2.113
33	07141780	Nekoma, KS	Walnut Ck	3083	31	1	1494	0.9270	1.496
34	07155590	nr Elkhart, KS	Cimarron Rv	7498	29	2	75.15	1.590	3.094
35	08128400	Tankersley, TX	Middle Concho Rv	5390	40	1	109.0	1.052	1.138
36	08129300	Tankersley, TX	Spring Ck	1098	40	1	136.4	2.840	4.680
37	08134000	nr Carlsbad, TX	N Concho Rv	3274	76	1	330.0	1.600	2.450
38	08155300	Austin, TX	Barton Ck at loop 360	300.0	26	1	191.5	1.252	2.253

Table 3 (*continued*)

Series No.	USGS code	Location of Gauge, State	River or Creek	Drainage area (km^2)	Record length n (yr)	No. of zeros n_1	$E(x)$ (m^3/s)	c_v	c_s
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
39	08180500	nr Riomedina, TX	Medina Rv	1681	23	1	95.10	2.118	2.516
40	08185000	Selma, TX	Cibolo C	708.7	55	3	347.8	1.664	2.318
41	08190500	Nr Brackettville, TX	W Nueces Rv	1795	56	2	933.2	2.542	4.706
42	08197500	nr Uvalde, TX	Frio Rv	1632	49	9	621.6	1.286	1.673
43	08200700	nr Hondo, TX	Hondo Ck	385.4	40	6	366.6	1.312	1.843
44	08202700	nr D'Hanis, TX	Seco Ck	434.5	41	8	305.0	1.295	1.486

reproduces the empirical probability of the zero event and the weighting factor β is then determined by the ratio n_1/n (Eq. (3)). It is the only one of six terms in $\log L$ function of KD in Eq. (21), which is a function of x and at the same time of both parameters.

5.2. Comparison of distribution functions

The KD cumulative distribution functions (CDF's) were computed and are presented for four sample series in Fig. 6a–d. Also presented are the empirical CDFs defined as $F_{\text{emp}}(x_{(i)}) = i/(n+1)$, where $x_{(i)}$ is the ith element of the sample arranged in ascending order. The respective MOM and MLM-estimated PDFs are shown in Fig. 7. In nine out of 44 series, i.e. the samples No. 11, 19, 20, 21, 26, 33, 35, 42 and 44, both methods of parameter estimation were in satisfactory agreement. For other 35 series, there are noticeable discrepancies between the two methods.

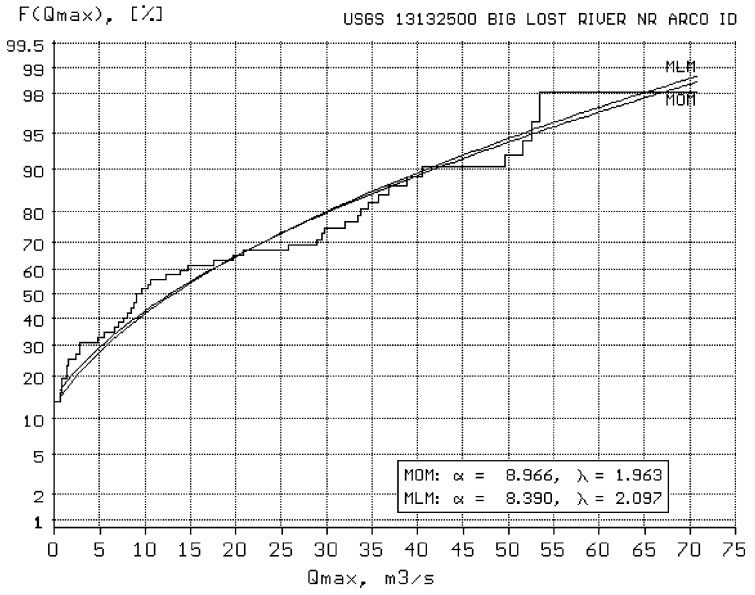
Obviously, the estimate of the two methods (Table 4) should asymptotically converge if the chosen PDF is the true one. When dealing with a finite sample, the discrepancy between the estimates of the two methods might be caused both by the incorrect choice of the distribution function and by the sampling error. In practice, the assumed model always differs from the true one. Therefore, no model can properly reproduce the data set in its entire interval of variability. This fact should be taken into account when selecting the estimation method. In FFA, the practical interest is in the upper quantiles estimation. Dealing with the two-parameter lower bounded model and recalling the sampling properties of the first two moments (Curetan, 1968; Wallis et al., 1974; Cunnane, 1989), it seems reasonable to base here on the MOM-estimates rather than on MLM.

If the KD model does not satisfactorily fit the data, then two possibilities may be considered: (1) replacement by another two-parameter model, which may be the only reasonable solution in case of a short sample, or (2) extension of the model to a three-parameter form, i.e. to the 3-KD, or replacement by another three-parameter model.

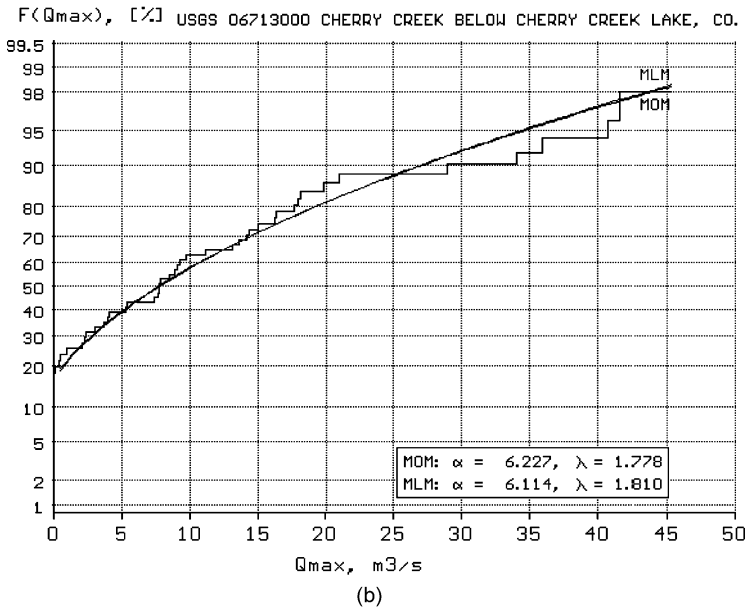
Table 4
Fitting KD and M-like models to the data

No.	USGS No.	n_1/n	Kinematic difussion (KD)					Muskingum-like (M-like)				
			MOM		MLM		$\hat{x}_{1\%}^{(MOM)}/\hat{x}_{1\%}^{(MLM)}$	MOM		MLM		$\hat{x}_{1\%}^{(MOM)}/\hat{x}_{1\%}^{(MLM)}$
			$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$		$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	11063680	0.0253	0.354	20.09	1.723	4.13	2.44	0.6990	23.65	0.0253	7.303	2.41
2	11028500	0.0656	0.392	73.84	1.559	18.59	2.18	0.6719	88.34	0.0656	31.02	2.19
3	11042000	0.2571	0.221	527.2	1.070	108.9	2.53	0.8010	585.5	0.2571	156.9	2.59
4	11012500	0.1094	0.700	7.29	1.553	3.29	1.54	0.4813	9.851	0.1094	5.737	1.51
5	11012000	0.0781	0.303	81.04	1.424	17.25	2.43	0.7367	93.34	0.0781	26.66	2.53
6	10259300	0.1250	0.659	89.02	1.478	39.71	1.56	0.5042	118.4	0.1250	67.07	1.54
7	10259200	0.0256	0.679	34.98	1.882	12.61	1.73	0.4934	46.86	0.0256	24.37	1.65
8	10258500	0.0303	1.094	27.92	2.023	15.1	1.37	0.2929	43.19	0.0303	31.50	1.28
9	10264675	0.1818	2.665	0.05	1.870	0.07	0.84	−.1426	0.1115	0.1818	0.1558	*
10	11042631	0.0909	0.376	28.66	1.423	7.57	2.13	0.6836	34.05	0.0909	11.85	2.20
11	11046250	0.3478	0.999	17.02	1.032	16.47	1.02	0.3339	25.53	0.3478	26.07	0.98
12	08404000	0.1875	0.490	304.2	1.298	114.9	1.73	0.6065	378.8	0.1875	183.5	1.72
13	09520500	0.1042	0.402	914.1	1.443	254.6	2.06	0.6654	1098	0.1042	410.0	2.09
14	09519800	0.0500	0.323	182.8	1.577	37.48	2.46	0.7217	212.4	0.0500	62.22	2.49
15	09513860	0.0513	1.344	53.45	2.056	34.94	1.24	0.1961	89.38	0.0513	75.74	1.14
16	09484500	0.0323	1.597	83.16	2.372	56.01	1.21	0.1119	149.6	0.0323	137.3	1.07
17	09502000	0.0154	0.625	305.0	2.006	95.06	1.87	0.5238	400.4	0.0154	193.7	1.74
18	09517000	0.0270	0.934	148.8	2.154	64.51	1.54	0.3633	218.3	0.0270	142.8	1.39
19	13132500	0.0980	1.963	8.97	2.097	8.39	1.03	0.0094	17.77	0.0980	19.51	0.93
20	06713000	0.1600	1.778	6.23	1.810	6.11	1.01	0.0589	11.76	0.1600	13.18	0.92
21	07126140	0.1765	1.368	2.65	1.595	2.27	1.08	0.1478	4.481	0.1766	4.637	1.02
22	07126480	0.0714	0.572	20.99	1.612	7.45	1.77	0.5491	27.73	0.0714	13.46	1.75
23	06846500	0.0179	0.699	31.89	2.082	10.7	1.79	0.4910	43.02	0.0179	22.30	1.63
24	06846000	0.0222	1.508	14.52	2.460	8.9	1.27	0.1555	24.06	0.0222	20.78	1.11
25	06866900	0.0333	1.046	71.14	1.975	37.69	1.39	0.3347	108.3	0.0323	74.48	1.30
26	06870300	0.0270	2.810	28.86	3.292	24.63	1.07	−.0820	69.37	0.0250	76.98	*
27	06871800	0.0143	0.312	278.9	1.591	54.76	2.53	0.7299	322.5	0.0143	88.38	2.62
28	06873200	0.0682	1.281	8.06	2.172	4.75	1.30	0.2191	13.22	0.0682	11.08	1.15
29	07139000	0.0128	0.311	576.0	1.708	105.0	2.62	0.7305	665.8	0.0128	181.8	2.63
30	07139500	0.0509	0.287	431.3	1.542	80.27	2.62	0.7491	493.3	0.0509	130.4	2.68
31	07140850	0.1053	1.311	16.56	2.010	10.79	1.24	0.2083	27.41	0.1053	24.25	1.10
32	07141200	0.0132	2.053	42.27	3.115	27.86	1.21	−0.0131	85.67	0.0132	87.95	*
33	07141780	0.0323	2.327	18.18	2.903	14.57	1.11	−0.0756	1389	0.0323	1544	*
34	07155590	0.0690	0.791	94.99	1.810	41.51	1.56	0.4332	132.6	0.0690	80.72	1.46
35	08128400	0.0250	1.807	60.27	2.364	46.07	1.14	0.0506	114.8	0.0250	111.7	1.02
36	08129300	0.0250	0.248	549.8	1.511	90.22	2.83	0.7793	618.0	0.0250	139.9	2.98
37	08134000	0.0132	0.781	422.4	1.954	168.9	1.63	0.4382	587.5	0.0132	334.4	1.54
38	08155300	0.0385	1.276	150.1	2.180	87.85	1.31	0.2211	245.9	0.0385	199.2	1.18
39	08180500	0.0435	0.446	213.4	1.539	61.79	2.00	0.6355	260.9	0.0435	99.42	2.07
40	08185000	0.0546	0.722	481.4	1.728	201.2	1.60	0.4694	655.4	0.0546	367.8	1.56
41	08190500	0.0357	0.309	3015	1.588	587.7	2.54	0.7320	3482	0.0357	967.8	2.59
42	08197500	0.1837	1.209	514.3	1.562	398.0	1.14	0.2467	8251	0.1837	761.4	1.06
43	08200700	0.1500	1.161	315.7	1.675	218.9	1.21	0.2654	499.0	0.1500	431.3	1.12
44	08202700	.1951	1.192	255.8	1.493	204.3	1.13	.2531	408.4	.1951	379.0	1.06

PDFs used to model the annual peak flows of perennial rivers have a semi-infinite lower bounded range with a non-negative value at the lower bound. For ephemeral streams, there is a non-zero probability mass on a zero value in a data series. Consequently, one can expect a non-zero density at the lower bound of the continuous part of the function in Eq. (1). Low values of the λ-parameter (Table 4, Columns 4 and 6)

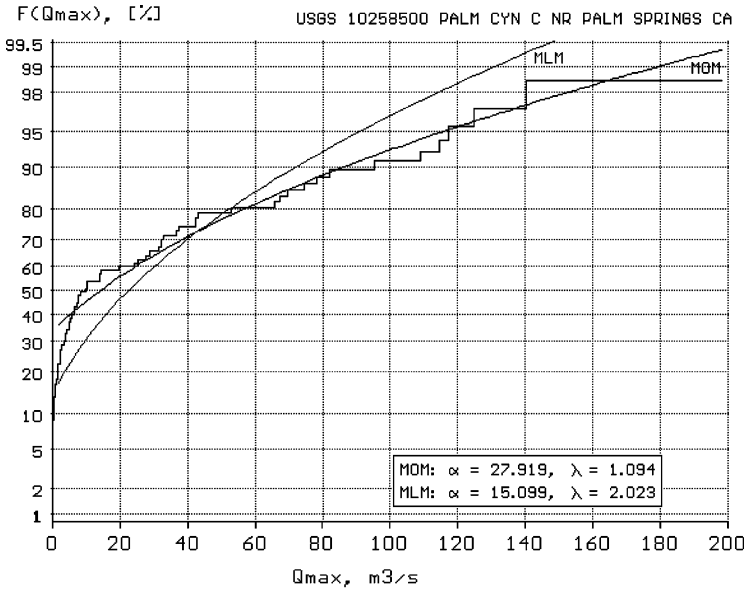


(a)

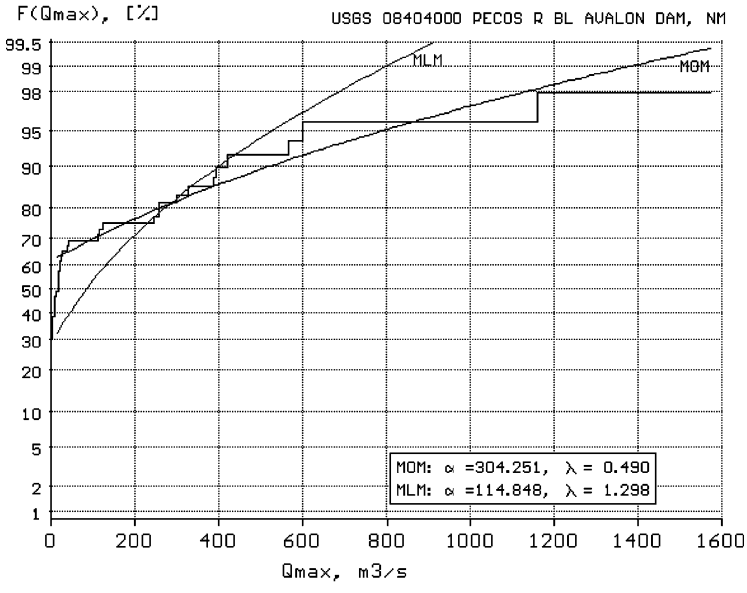


(b)

Fig. 6. Empirical and two theoretical (estimated with MOM and MLM) KD cumulative distribution functions (CDFs). (a) The Big Lost River, Arco, ID; (b) Cherry Creek, Cherry Creek Lake, CO; (c) The Palm Cyn River, Palm Springs, CA; (d) Pecos River, Avalon Dam, NM.



(c)



(d)

Fig. 6 (*continued*)

seem to validate the above hypothesis. Furthermore, a majority of the data series used shows the value of $\lambda < 2$. According to Eq. (9), the maximum of the function in Eq. (8) is for $\lambda \leq 2$ at the zero lower bound.

5.3. Muskingum-like distribution and its comparison with KD model

Having defined the general properties for the function in Eq. (1), one can acknowledge that one

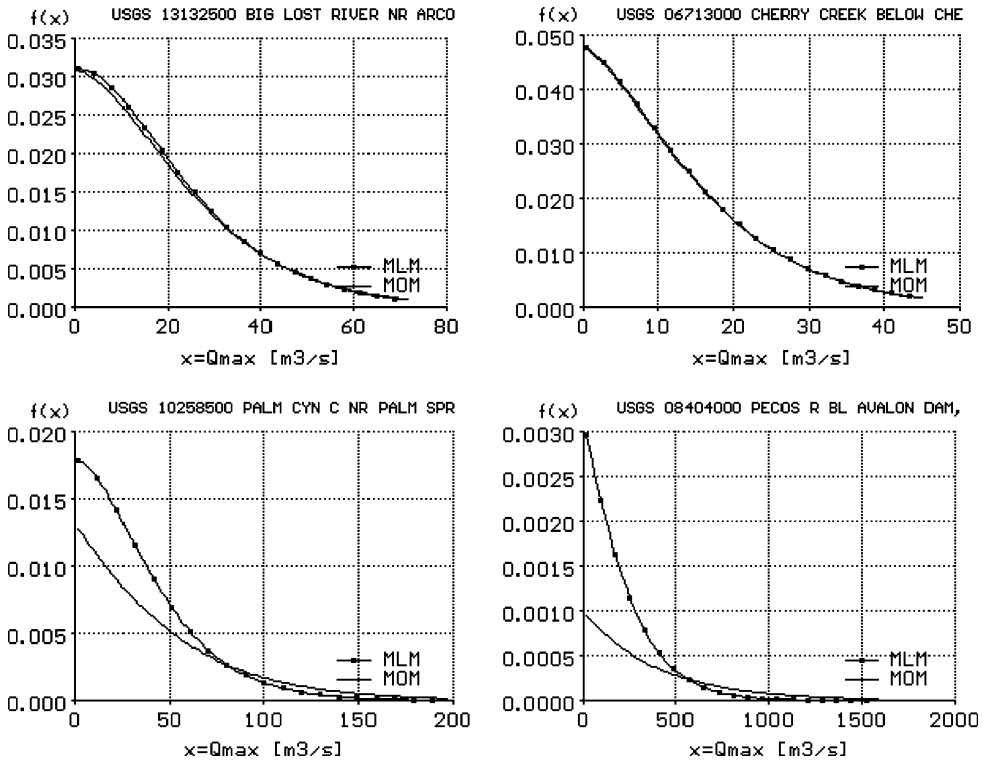


Fig. 7. Two theoretical KD PDFs for the four selected flow gauging stations estimated with MOM and MLM.

plausible alternative within the family of two-parameter distributions for discontinuous samples with zero values may be the Muskingum-like probability distribution function (M-like). Therefore, once again the unit impulse response of the Muskingum flood routing model

$$h(t) = -\frac{\alpha}{1-\alpha}\delta(t) + \frac{1}{K(1-\alpha)^2} \exp\left[-\frac{t}{K(1-\alpha)}\right] \tag{57}$$

was selected as the probability distribution for modeling hydrologic samples with zero values. It has been shown by Strupczewski and Kundzewicz (1980) and Dooge et al. (1982) that modeling flood routing along a short reach of a lowland river might result in negative values of α. Then,

$$0 \le \left(-\frac{\alpha}{1-\alpha}\right) \le 1 \tag{58}$$

where K is the average travel time along a reach, and α is a parameter. Denoting $-\alpha/1-a=\beta$ and $K/1-\beta=\gamma$ and mapping from time t to x, one gets a two-parameter probability distribution function:

$$f(x) = \beta\delta(x) + \frac{(1-\beta)}{\gamma}\exp(-x/\gamma). \tag{59}$$

The PDF given by Eq. (59) is a weighted sum of two functions: a delta function and an exponential function. It is interesting to note that in this function parameter β is a weighting factor and parameter $K=(1-\beta)\gamma$ becomes the average of X. Thus, the original expressions of the weighting factor and the average travel time are modified somewhat under mapping, but the conceptual meaning of the modified expressions remains more or less intact.

The parameter estimation for Eq. (59) is given in Appendix C. Using MOM equations (C7) and

(C8) and ML equations (C5) and (C6), the values of parameters β and γ for the M-like distribution were calculated as shown in Table 4. It is interesting to note a similarity to the KD pattern in the relation between MOM (Columns 9 and 10) and MLM estimated values (Columns 11 and 12). MOM for the M-like model resulted in greater values of β and greater values of γ than did MLM in 36 of 44 series. It led to greater values of upper quantiles of MOM than of MLM as shown in Column 13 by the ratio of estimates of $x_{1\%}$. For the M-like model, MOM (Table 4, Columns 9 and 10) and MLM (Table 4, Columns 11 and 12) parameter values were in satisfactory agreement for only five out of 44 data series, i.e. for samples No. 11, 19, 21, 42 and 44. For these five samples, the same behavior was observed in case of the KD model as well.

Inverting MOM-relation (C7):

$$c_v = \sqrt{\frac{1+\beta}{1-\beta}} \tag{60}$$

and recalling the admissible range of the weighting factor $\beta \in (0,1)$, it is seen that the coefficient of variation is limited to the range $c_v \in (1,\infty)$. Note that the c_v range of the KD model is $(0,\infty)$ for $\lambda \geq 0$ (Eq. (15)). Therefore, the c_v interval (0,1) is not covered by the M-like model. In fact, the c_v-values of four of the 44 series (Table 3, Column 9) are less than unity and an attempt to apply MOM resulted in a negative probability, i.e. $\hat{\beta} = P(x=0) < 0$ for these series (Table 4, Series marked by ' * ' in Column 13). According to Eq. (15) $c_v < 1$ corresponds to $\lambda > 2$, i.e. to the positive modal value.

For the M-like distribution, the coefficient of skewness is related to the coefficient of variation as

$$c_s = \frac{3}{2}c_v + \frac{1}{2c_v^3} \tag{61}$$

which is shown in Fig. 8. The scatter band of (c_v, c_s) points indicates small preference of the KD over the M-like distribution for the given data. A comparison of Eqs. (16) and (61) shows that for $c_v > 2$ (or for $\lambda < 1/2$) the difference in c_s values is less than 1/16. Then, both models are deemed equally good if measured by the moment-ratio criterion. In fact, there are 14 of 44 data series with $c_v > 2$ (Table 3, Column 9).

The basic problem of selecting from a set of competing models one that best describes the data lies in the appropriate choice of the selection criterion. A flexible approach to model discrimination is

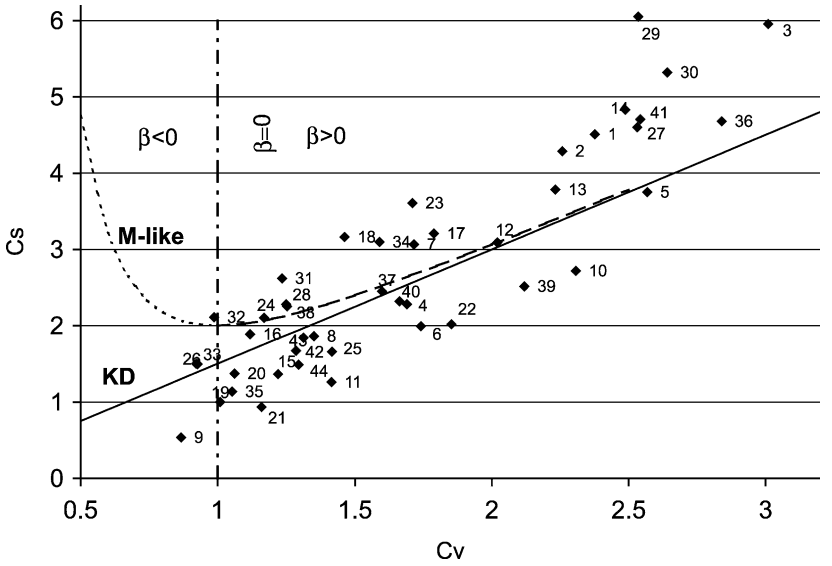


Fig. 8. Relationship between coefficient of variation and coefficient of skewness for KD and M-like probability distributions and for the empirical data.

the Akaike information criterion (Akaike, 1974), which builds on the log L function, makes allowance for different numbers of model parameters. In practical situations, the assumed models are always different from reality. Therefore, selecting a criterion applicable for FFA, it seems reasonable to choose one, which is sensitive to a fit of a model in the upper part of a sample. If so, it automatically eliminates the criterion based on the likelihood function, as its value highly depends on the range of the main probability mass.

The ratio of $x_{1\%}$ for MOM and MLM estimates is used here for the purpose. In fact, it can serve not only for the selection of a better-fitted model but also for the assessment of the goodness of fit of the upper tail at the same time. Obviously, in the case of the true distribution function the $x_{1\%}$ ratio should be close to one. Arbitrarily taking the admissible interval as $0.5 \leq x_{1\%}^{(\mathrm{MOM})}/x_{1\%}^{(\mathrm{MLM})} \leq 2$ one can find in Table 4, that as many as 12 samples out of 44 and 13 out of 40 are above the upper limit for the KD and M-like distributions, respectively. Neither sample was beyond the lower limit for both distributions. Therefore, the criterion indicates a slight superiority of KD over M-like model, confirming the result of the moment-ratio comparison (Fig. 8). Narrowing the interval down to $\pm 25\%$ of the difference, one finds 14 samples within the interval both for KD and M-like distributions. There are one and three samples, which produce the value of the ratio less than the unit for KD and M-like models, respectively. Since the same samples fall to the same range of the ratio for both assumed distributions, no considerable improvement can be expected by replacing one model by the other. Both the high range of the $x_{1\%}$ ratio variability and the asymmetry of the $x_{1\%}$ ratio values around the unit value point out a need to enlarge a set of competitive models. Due to a lack of candidates within the two-parameter models family, the three-parameter options should be considered, provided a sample size legitimates it.

6. Conclusions

The following conclusions are drawn from this study:

(a) The impulse response of the linear KD model is a promising model for frequency analysis of hydrologic samples with zero values. It is easy from a computational point, particularly if the MOM is applied.
(b) Being a two-parameter model, the KD distribution is especially attractive in case of short samples, which is common in arid and semiarid regions, especially in developing countries.
(c) With its two parameters, the KD model can reproduce a variety of probability density forms encountered in nature (Fig. 2).
(d) For the KD distribution, the MLM estimate of the mean is equivalent to that of the MOM, i.e. it is arithmetic mean, which is especially attractive, one recognizes, when applying the ML-method, that the true distribution function is unknown.
(e) A comparison of both methods (Fig. 6a–d) shows that MOM better reproduces the upper tail of the distribution, while MLM is more robust for higher sample values than MOM and more conditioned on the value of $P(X = 0)$.
(f) The KD distribution represents flood frequency characteristics of arid zones quite well, as seen from the use of the USGS data.
(g) The M-like model can be considered as an alternative of the KD model in frequency analysis of samples with zero values.
(h) For larger samples, an extension of the models to three-parameter forms is advisable as shown by the discrepancies between MOM and MLM estimates.

Acknowledgements

This study was supported in part by Polish Committee for Scientific Research, Grant No. 6 P 4D 056 17, ‘Revision of applicability of the parametric methods for estimation of statistical characteristics of floods’. This support is gratefully acknowledged. The authors are thankful to A. Kozlowski for his help in retrieval and preliminary processing of the USGS data. Two anonymous reviewers made constructive and comprehensive comments on the manuscript.

The authors gratefully express their gratitude to these reviewers for their thoughtful and thorough review.

Appendix A. The modified Bessel function

The nth order modified Bessel function of the 1st kind (e.g. Korn and Korn, 1961, Sec. 21.8):

$$I_n(z) = \sum_{i=1}^{\infty} \frac{1}{(i-1)!(n+i-1)!} \left(\frac{z}{2}\right)^{2(i-1)+n} \tag{A1}$$

as the solution to the differential equation

$$z^2 \frac{\partial^2 y}{\partial z^2} + z\frac{\partial y}{\partial z} - (z^2 + n^2)y = 0. \tag{A2}$$

The $I_n(z)$ is sometimes known as hyperbolic Bessel function. It fulfills the recursive equation:

$$I_{n+1}(z) = I_{n-1}(z) - \frac{2n}{z} I_n(z)$$

$$= 2\frac{\mathrm{d}}{\mathrm{d}z} I_n(z) - I_{n-1}(z) \tag{A3}$$

which for $n = 1$ gives

$$\frac{\partial}{\partial z} I_1(z) = I_0(z) - \frac{I_1(z)}{z}. \tag{A4}$$

Appendix B. Asymptotic variance of the KD-quantiles estimates got by MLM

To get the asymptotic variance–covariance matrix (31), the asymptotic value of A in matrix (35) should be derived. Its sampling value is given by Eq. (33) and consists of the two terms. The expected value of its first term can be found from Eq. (6a) as

$$E\left(\frac{n_2}{n}\right) \approx 1 - P_{i=0}(\lambda) = 1 - \exp(-\lambda) \tag{B1}$$

while the second term

$$ED(Z) \approx \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n_2} D_j = \int_0^{\infty} D(z) f(z) \mathrm{d}z$$

$$= \int_0^{\infty} D(z) f_c(z;\lambda) \mathrm{d}z \tag{B2}$$

or substituting $D(z)$ from Eq. (34) into Eq. (B2) and the continuous term of the PDF got from Eqs. (8a) and (22a), i.e.

$$f_c(z;\lambda) = \exp\left(-\lambda - \frac{z^2}{4\lambda}\right) I_1(z) \tag{B3}$$

one gets

$$ED(z) \approx \int_0^{\infty} z^2 (\ln I_1(z))'' I_1(z) \times \exp(-\lambda - z^2/4\lambda)\mathrm{d}z \tag{B4}$$

which was integrated numerically (Fig. B1(a)).

Hence, substituting Eqs. (B1) and (B4) into Eq. (33), one gets the required asymptotic value of A as shown in Fig. B1(b).\eqalign{

$$EA(z) \approx 1 - \exp(-\lambda) - \int_0^{\infty} z^2 (\ln I_1(z))'' I_1(z) \times \exp(-\lambda - z^2/4\lambda)\mathrm{d}z. \tag{B5}$$

Substituting it into Eq. (36), the asymptotic variance–covariance matrix is obtained.

To get the asymptotic variance of quantiles y_p defined by Eq. (39), the derivative (40) should be obtained. Note that the function

$$t_p = \varphi(\lambda, p) \tag{B6}$$

does not exist in an explicit form. Its derivative with respect to λ can be obtained by the method of implicit partial differentiation. Using Eq. (19), one gets

$$\frac{\partial t_p(\lambda)}{\partial \lambda} = \sqrt{\frac{t_p}{\lambda}} \frac{I_0(2\sqrt{\lambda t_p})}{I_1(2\sqrt{\lambda t_p})} \tag{B7}$$

and therefore,

$$G = \frac{\partial \ln t_p(\lambda)}{\partial \lambda} = \frac{1}{t_p(\lambda)} \frac{\partial t_p(\lambda)}{\partial \lambda}$$

$$= \frac{1}{\sqrt{t_p \lambda}} \frac{I_0(2\sqrt{\lambda t_p})}{I_1(2\sqrt{\lambda t_p})}. \tag{B8}$$

Fig. B2 displays the function G for various values of p and λ.

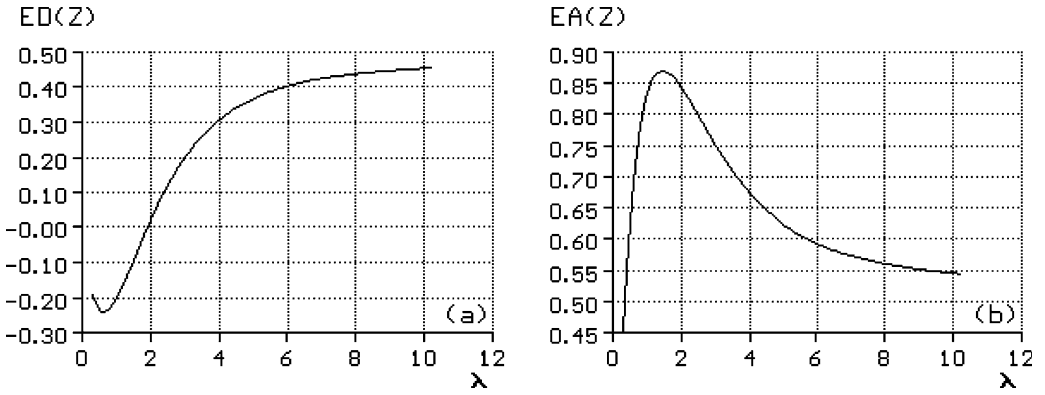


Fig. B1. Asymptotic values of the conditional D and A vs the λ parameter.

Appendix C. Parameter estimation of the M-like probability distribution

MLM. As before, let the sample contains n_1 zeros and n_2 positive values. We determine the likelihood function of Eq. (59) in the form:

$$L = (\beta)^{n_1} \prod_{j=1}^{n_2} \frac{(1-\beta)}{\gamma} \exp\left(-\frac{x_j}{\gamma}\right). \tag{C1}$$

So the log-likelihood function $\ln L$ will be:

$$\ln L = n_1 \ln \beta + n_2 \ln(1-\beta) - n_2 \ln \gamma - \frac{1}{\gamma} \sum_{j=1}^{n_2} x_j.$$

We get the MLM equations:

$$\frac{\partial \ln L}{\partial \beta} = \frac{n_1}{\beta} - \frac{n_2}{1-\beta} = 0 \tag{C3}$$

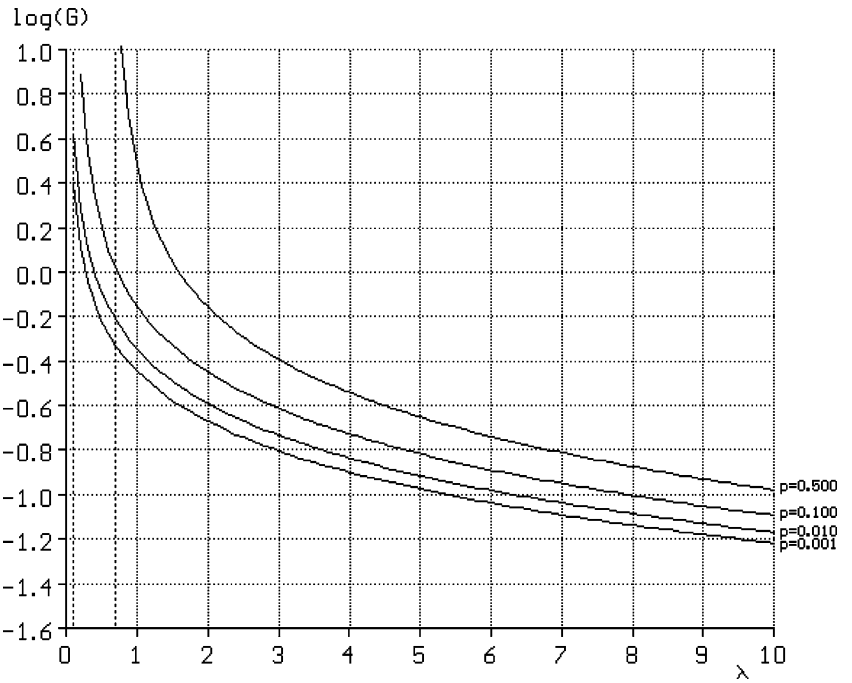


Fig. B2. Function G vs the λ parameter in log scale. Vertical dotted lines for $p = 0.1$ and 0.5 define the left bound of intervals where G exists.

$$\frac{\partial \ln L}{\partial \gamma} = -\frac{n_2}{\gamma} - \frac{1}{\gamma^2}\sum_{j=1}^{n_2} x_j = 0 \tag{C4}$$

and therefore

$$\hat{\beta} = \frac{n_1}{n_1 + n_2} = \frac{n_1}{n} \tag{C5}$$

$$\hat{\gamma} = \frac{1}{n_2}\sum_{j=1}^{n_2} x_j. \tag{C6}$$

MOM. Matching the first two theoretical and sampling moments we get

$$\hat{\beta} = \frac{c_v^2 - 1}{c_v^2 + 1} \tag{C7}$$

$$\hat{\gamma} = \frac{\mu_1'}{1 - \hat{\beta}}. \tag{C8}$$

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19 (16), 716–722.

Curetan, E.E., 1968. Unbiased estimation of the standard deviation. Am. Stat. 22 (1), 22.

Cunnane, C., 1989. Statistical distributions for flood frequency analysis. WMO—Operational Hydrology Report No. 3, WMO-No. 718, Geneva, 1989.

Deymie, P., 1939. Propagation d'une intumescence allongee (Problem aval) [Propagation of an elongated intumescence]. Proceedings of the Fifth International Congress on Appllied Mechanics, New York (in French), pp. 537–544.

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

Dooge, J.C.I., Harley, B.M., 1967. Linear routing in uniform open channels. Proceedings of International Hydrology Symposium, Fort Collins, CO, September 1967, vol. 1, Paper No. 8, pp. 57–63.

Dooge, J.C.I., Strupczewski, W.G., Napiorkowski, J.J., 1982. Hydrodynamic derivation of storage parameters of the Muskingum model. J. Hydrol. 54, 371–387.

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

Dooge, J.C.I., Napiorkowski, J.J., Strupczewski, W.G., 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.

Eagleson, P.S., 1978. Climate, soil and vegetation. 2. The distribution of annual precipitation derived from observed storm sequences. Water Resour. Res. 5, 713–721.

Einstein, H.A., 1942. Formulas for the transportation of bed load. Trans. ASCE, Paper No. 2140, 561–597.

Hayami, S., 1951. On the propagation of flood waves. Kyoto Univ. Jpn Disaster Prevent. Res. Inst. Bull. 1, 1–16.

Jennings, M.E., Benson, M.A., 1969. Frequency curve for annual flood series with some zero events or incomplete data. Water Resour. Res. 5 (1), 276–280.

Kaczmarek, Z., 1977. Statistical Methods in Hydrology and Meteorology, US Department of Commerce, Springfield, VA.

Kendall, M.G., Stuart, A., 1969. The Advanced Theory of Statistics. V. 1. Distribution Theory, Charles Griffin and Company Ltd, London.

Kendall, M.G., Stuart, A., 1973. The Advanced Theory of Statistics, Inference and Relationship, vol. 2. Charles Griffin and Company Ltd, London, Ch. 24.

Kite, G.W., 1988. Frequency and Risk Analyses in Hydrology, Water Resource Publications, Littleton, CO.

Korn, G.A., Korn, T.M., 1961. Mathematics Handbook for Scientists and Engineers, 21.8-6, McGraw-Hill, New York.

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

Lighthill, M.H., Witham, G.B., 1955. On kinematic waves. I. Flood movements in long rivers. Proc. R. Soc., Lond. Ser. A 229, 281–316.

Masse, P., 1939. Recherches sur la theorie des eaux courantes [Researches on the theory of water currents]. Proceedings of the Fifth International Congress on Applied Mechanics, New York (in French), pp. 545–549.

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

Schaefer, M.G., 1998. General storm stochastic event flood model: technical support manual. Technical Report, MGS Engineering Consultants Inc., Olympia, Washington.

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

Singh, V.P., 1998. Entropy-based Parameter Estimation in Hydrology, Kluwer, Dordrecht.

Stedinger, J.R., Vogel, M.V., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidment, D.R., (Ed.), Handbook of Hydrology, McGraw-Hill, New York, Ch. 18.

Strupczewski, W.G., Kundzewicz, Z.W., 1980. Muskingum method revisited. J. Hydrol. 48, 327–342.

Strupczewski, W.G., Napiorkowski, J.J., 1986. Asymptotic behaviour of physically based multiple Muskingum model. In: Shen, H.W., Obeysekera, J.T.B., Yevjevich, V., Decoursey, D.G. (Eds.), Proceedings of the Fourth International Hydrological Symposium, Engneering Research Center, Colorado State University, Fort Collins, CO, pp. 372–381.

Strupczewski, W.G., Napiorkowski, J.J., 1989. Properties of the distributed Muskingum model. Acta Geophys. Pol., V.XXXVII 3–4, 299–314.

Strupczewski, W.G., Napiorkowski, J.J., 1990a. Linear flood routing model for rapid flow. Hydrol. Sci. J. 35 (1/2), 49–64.

Strupczewski, W.G., Napiorkowski, J.J., 1990b. What is the distributed delayed Muskingum model? Hydrol. Sci. J. 35 (1/2), 65–78.

Strupczewski, W.G., Napiorkowski, J.J., Dooge, J.C.I., 1989. The distributed Muskingum model. J. Hydrol. 111, 235–257.

Todorovic, P., 1982. Stochastic modeling of floods. In: Singh, V.P., (Ed.), Statistical Analysis of Rainfall and Runoff, Water Resource Publications, Littleton, CO, pp. 597–636.

Wang, S.X., Singh, V.P., 1995. Frequency estimation for hydrological samples with zero value. J. Water Resour. Plan. Mgmt, ASCE 121 (1), 98–108.

Wallis, J.R., Matalas, N.C., Slack, J.R., 1974. Just a moment!. Water Resour. Res. 13 (1), 159–182.

Woo, M.K., Wu, K., 1989. Fitting annual floods with zero flows. Can. Water Resour. J. 14 (2), 10–16.