Hydrology and Earth System Sciences Discussions

An opportunity of application of excess factor in hydrology

V. Kovalenko, E. Gaidukova, and A. Kachalova

Russian State Hydrometeorological University (RSHU), Saint-Petersburg, Russia

Correspondence to: E. Gaidukova (oderiut@mail.ru)

HESSD

9, 13635–13649, 2012

An opportunity of application of excess factor in hydrology

V. Kovalenko et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

|◀ | ▶|
◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

In last few years in hydrology an interest to excess factor has appeared as a reaction to unsuccessful attempts to simulate and predict evolving hydrological processes, which attributive property is statistical instability. The article shows, that the latter has a place at strong relative multiplicative noises of probabilistic stochastic model of a river flow formation, phenomenological display of which are "the thick tails" and polymodality, for which the excess factor "answers", by being ignored by a modern hydrology in connection to the large error of its calculation because of insufficient duration of lines of observation over a flow. However, it is found out, that the duration of observation of several decades practically stabilizes variability of the excess factor, the error of which definition appears commensurable with an error of other calculated characteristics used in engineering hydrology.

1 Motivation of researches

At the end of the 80s in USSR in several organizations the researches of hydrological consequences of antropogeneous change of climate (basically have used the climatic scenarios of Budiko, 1974) were carried out. In them active participation was taken by faculty of hydrophysics and hydroforecasts of the Russian state hydrometeorological university (Kovalenko, 1993). For a basis of the technique a stochastic model of flow formation was taken, nucleus of which served the ordinary differential equation of the first order:

$$\mathrm{d}Q/\mathrm{d}t = -(1/k\tau)Q + \dot{X}/\tau, \tag{1}$$

where Q – sliding average consumption of water (modules or layers) in closing target of watershed; k – factor of a flow, to which the different sense can be given depending on a kind of a long-term flow (it can be accepted both constant and dependent from the factors of a spreading surface (population of pool, waterlogging level, degree of

urbanization, etc.) or from the hydrometeorological factors – intensity of precipitation $\dot{X}$ and temperature of air T °C); τ – relaxation time of watershed appropriate to radius of correlation, usually one year.

After transformations and introduction into Eq. (1) white noises we shall receive the stochastic differential equation (model of the linear forming filter):

$$dQ = [-(\bar{c} - \tilde{c})Q + \bar{N} + \tilde{N}]dt, \tag{2}$$

where $c = 1/k\tau = \bar{c} + \tilde{c}$; $N = \dot{X}/\tau = \bar{N} + \tilde{N}$ (where $\bar{c}$, $\bar{N}$ – mathematical expectations; $\tilde{c}$, $\tilde{N}$ – correlated with each other white noise with intensities $G_{\tilde{c}}$, $G_{\tilde{N}}$ and mutual intensity $G_{\tilde{c}\tilde{N}}$).

In a science the procedure of replacement (Eq. 2) statistically by equation Fokkera-Planka-Kolmogorova (FPK), equivalent to it, describing Markov's evolution of density of probability $p(Q, t)$ is known:

$$\frac{\partial p(Q,t)}{\partial t} = -\frac{\partial}{\partial Q}[A(Q,t)p(Q,t)] + 0.5\frac{\partial^2}{\partial^2 Q}[B(Q,t)p(Q,t)], \tag{3}$$

where A and B – drift and diffusion coefficients, determined by physical and statistical parameters which are included in Eq. (2):

$$A(Q,t) = -(\bar{c} - 0.5G_{\tilde{c}})Q - 0.5G_{\tilde{c}\tilde{N}} + \bar{N},$$

$$B(Q,t) = G_{\tilde{c}}Q^2 - 2G_{\tilde{c}\tilde{N}}Q + G_{\tilde{N}}.$$

As in a practical hydrology because of short series of observations are limited to the several moments of probabilistic distributions, it is meaningful to approximate (Eq. 3) the system of the differential equations for the initial moments m_i (procedure of similar approximation in a science is known):

$$\begin{aligned}
dm_1/dt &= -(\bar{c} - 0.5\,G_{\tilde{c}})\,m_1 + \bar{N} - 0.5\,G_{\tilde{c}\tilde{N}}; \\
dm_2/dt &= -2\,(\bar{c} - 0.5\,G_{\tilde{c}})\,m_2 + 2\bar{N}m_1 - 3\,G_{\tilde{c}\tilde{N}}m_1 + G_{\tilde{N}}; \\
dm_3/dt &= -3\,(\bar{c} - 1.5\,G_{\tilde{c}})\,m_3 + 3\bar{N}m_2 - 7.5\,G_{\tilde{c}\tilde{N}}m_2 + 3G_{\tilde{N}}m_1; \\
dm_4/dt &= -4\,(\bar{c} - 2\,G_{\tilde{c}})\,m_4 + 4\bar{N}m_3 - 4\cdot3.5\,G_{\tilde{c}\tilde{N}}m_3 + 6G_{\tilde{N}}m_2.
\end{aligned} \tag{4}$$

This system of equations is enough to define all calculated hydrological characteristics: mean $\bar{Q} = m_1$; factors of variation $C_v = f(m_1, m_2)$ and asymmetry $C_s = f(m_1, m_2, m_3)$, and also the excess factor $\text{Eh} = f(m_1, m_2, m_3, m_4)$, which are given by:

$$C_v = \sqrt{\frac{\sum_{i=1}^{n} \left(\frac{Q_i}{\bar{Q}} - 1\right)^2}{n}}, \tag{5}$$

$$C_s = \frac{\sum_{i=1}^{n} \left(\frac{Q_i}{\bar{Q}} - 1\right)^3}{nC_v^3}, \tag{6}$$

$$\text{Eh} = \frac{\mu_4}{\sigma^4} - 3, \tag{7}$$

where $\bar{Q}$ – average annual consumption of water; Q_i – consumption of water for the i-th year; n – the number of terms of consumption of water (time series); $\mu_4 = \frac{\sum (Q_i - \bar{Q})^4}{n}$ – the central point of the fourth order; σ – standard deviation.

The real application of system (Eq. 4) in tasks of modeling and forecasting has resulted in understanding of limits of its applicability in connection with a problem of loss of the decision stability, which can take place not only at the adverse climatic scenarios, but also for a hydrological mode, existing for last decades. System (Eq. 4) (we shall designate a vector, components of which are m_1, m_2, m_3, m_4, through $\boldsymbol{m}$) steadily describes process of flow formation under condition of $\text{div}\dot{\boldsymbol{m}}_i = \sum \mathrm{d}\dot{m}_i/\mathrm{d}m_i < 0$ ($\bar{c} > 2G_{\bar{c}}$). This compressibility condition of phased space of four initial moments provides existence of K. Pearson's family of curve density of probability in a stationary mode. In certain sense, this family could be considered as an attractor of decisions (in extended and metaphorical interpretation of this term), the type each of which (decisions) is defined by a ratio between parameters of the equations of system (Eq. 4). In process of

increase of numerical meaning of criterion $\beta = G_{\tilde{c}}/\bar{c}$ (relative intensity of multiplicative noises), the stability is lost by, consistently, fourth ($\beta > 0.5$), third ($\beta > 0.67$) and second ($\beta > 1$) moments. Earlier (Kovalenko and Haustov, 1998) parameter β was mapped (for the existing mode of flow) and has appeared, that almost a half of the territory of Russia (southern regions mostly) gets in a zone of instability on factor of asymmetry C_s (on C_v naturally less).

This circumstance (incapacity, under certain conditions, of the model of formation of a long-term flow based on the linear forming filter) has caused numerous attempts to leave from a deadlock situation, which has resulted, at the end, in creation of a new scientific direction "Partially infinite hydrology" (Kovalenko, 2007). In it instability is considered as an attribute of development (evolution, complication) of the system, which can be described, by expanding phase space, in which the unstable task is immersed. The necessary number of phase variables is defined by methods of fractal diagnostics (Kovalenko, 2007), but their representation could not be predicted, particularly, by rational methods, it is necessary to involve intuition, experience of a hydrologist and similar unformalizable concepts (from here is the name, "partially infinity", i.e., "partially indeterminate"). Practically (coming back to the stated method) it means, that for the steady statistical description of flow formation, it is necessary to operate not one-dimensional ($p(Q)$, but, for example, two-dimensional density of probability ($p(Q, E)$, where E — evapotranspiration).

For a practical hydrologist, use of rather complex stochastic models of formation of a flow is uncommon. Such term as "the relative intensity of multiplicative noise of model" cuts hearing and causes a frank irritation. Therefore it is desirable to put into use a clearer and more accessible, for calculation on actual series of observations, measurement, the applicant for which is the excess factor.

2 Basic parameter in genetic model of formation of a long-term flow, responsible for excess

The marked above fact of instability of the initial moments (also of derived hydrological characteristics) causes two questions. (1) Why in engineering hydrology no one never lifted a question on instability? If such term was used, the instability of statistical estimations of some derived characteristic because of limitation of series of observations, and not in any way physical instability, as meant in our case. (2) Why the excess factor is thrown out from consideration, though it essentially influences the form of the curve density of probability and security determining reliability of projected hydraulic-engineering objects?

It is simpler to answer the second question. In hydrology the conventional formulas circulate, which allow an error of definition of the derived hydrological characteristics depending on length of the series of observations (Directory of applied statistics, 1989; Klibashev and Goroshkov, 1970; Standard 8.508-84, 2002) to be calculated:

$$\sigma_{(\bar{Q})_n} = C_{\mathrm{v}}/\sqrt{n}; \tag{8}$$

$$\sigma_{(C_{\mathrm{v}})_n} = (\sqrt{1 + C_{\mathrm{v}}^2})/\sqrt{2n}; \tag{9}$$

$$\sigma_{(C_{\mathrm{s}})_n} = \sqrt{\frac{6(n-1)}{(n+1)(n+3)}}; \tag{10}$$

$$\sigma_{(\mathrm{Eh})_n} = \sqrt{\frac{24n(n-2)(n-3)}{(n+5)(n+3)(n-1)^2}}, \tag{11}$$

where n – number of years of observations.

As seen from Table 1, the error, in all cases, decreases with growth of n, but for the excess factor it makes 46 % even at length of a series of 100 yr.

It is natural, that at such error hydrologists, in general, dismiss from consideration this characteristic (given below data put under doubt such indifference to the excess factor).

The answer to the first question requires small digression in methodology of science. In nature there is no instability as such. Only our representations about of that or other natural phenomenon can feature instability. These representations in mathematical sciences associate with stability of decision models, which are these representations. In our case such model is the system of Eq. (4). As soon as it has appeared in hydrology (i.e., the representation has appeared, how density of probability of the flow characteristic in closing target of watershed is generically formed), so almost at once there was a condition of stability of this representation. At construction of distributions on actual series, such representations are not required as there is nothing to be physically unstable.

To understand which parameters in the model answer for excess (at least answer for it in greater degree), we will address to statistically stationary variant of FPK model – Pearson's equation

$$\frac{\mathrm{d}p}{\mathrm{d}Q} = \frac{Q - a}{b_0 + b_1 Q + b_2 Q^2} p,$$ (12)

decision of which is the family of curves $p(Q)$, used by engineering hydrology. Originally this equation has not been related to hydrology; it was offered for approximation of empirical distributions without any connection with the FPK equation. Coefficients a, b_0, b_1, b_2 till now are considered in hydrology as "fitting", that the analytical curve $p(Q)$ best met empirical distribution.

The situation cardinally changes when Eq. (12) is approached (to its conclusion) on the part of the FPK equation: coefficients communicate by formulae with parameters $\bar{c}$, $\bar{N}$, $G_{\tilde{c}}$, $G_{\tilde{N}}$, $G_{\tilde{c}\tilde{N}}$ and get physical (instead of "fitting") sense. These formulae are shown below:

$$a = \frac{G_{\tilde{c}\bar{N}} + 2\bar{N}}{2\bar{c} + G_{\tilde{c}}}; \quad b_0 = \frac{-G_{\bar{N}}}{2\bar{c} + G_{\tilde{c}}}; \quad b_1 = \frac{2G_{\tilde{c}\bar{N}}}{2\bar{c} + G_{\tilde{c}}}; \quad b_2 = \frac{-G_{\tilde{c}}}{2\bar{c} + G_{\tilde{c}}}. \tag{13}$$

They open an opportunity to consider the Eq. (12) with the physically reasonable factors (Eq. 13) as genetic model of flow formation, and its decision $p(Q)$ – as theoretical distribution, to which will be coordinated (or will not be coordinated) the empirical dis-
5 tribution. Most importantly – is possible, changing the factors of flow formation (climatic or spreading surface of pool) through parameters $\bar{c}$ and $\bar{N}$, i.e., through precipitation $\dot{X}$ and values influencing factor of the flow k to estimate sensitivity to similar changes of the form of density curve of probability.

Now (in a context of given article) it is necessary to find out what parameter (or their
10 combination) renders influence on excess factor Eh. There are two "hints": (1) from the previous text follows, that the instability causes criterion $\beta = G_{\tilde{c}}/\bar{c}$; (2) standard distributions, used in a practical hydrology, does not require knowledge of the fourth moment; thereof in the Eq. (12) $b_2 = 0$. Thus it is obvious, that the excess is influenced, first of all, by complex parameter β, containing $G_{\tilde{c}}$ and $\bar{c}$. Before integration of the
15 differential Eq. (12) could be begun, it shall be reminded, that at zero meaning of excess factor Eh the distribution of $p(Q)$ becomes normal (at zero asymmetry), at Eh > 0 – peak-like, at Eh < 0 – plateau (or with a "pothole").

In Fig. 1a, distributions of $p(Q)$ are shown, equations, received by integration of Eq. (12), at setting of factors a, b_0, b_1, b_2 with the help of Eq. (13).
20 From this figure, tails of distributions rise (thus increasing probability of accidents) at reduction of $\bar{c}$ and increase of $G_{\tilde{c}}$; however criterion of stability is served not by every one of them separately, but the relation $G_{\tilde{c}} / \bar{c}$.

3 Estimation of error of definition of the excess factor on actual series of observations

From the previous statement follows, that: (1) existing estimations of error of definition of the excess factor are considered so large, that it is generally dismissed from consideration; (2) influences of intensity of internal noise $G_{\tilde{c}}$ on the form of the curve $p(Q)$ are noticeable, which is especially important concerning tails that determine charges of small security levels, bringing to accidents. The degree of this influence is a not speculative estimation, but is a direct consequence of formation model of flow, bringing to family of Pearson's curves, everywhere used in practical hydrology. Attempt therefore will be made below to modernize a technique of estimation of a definition error of the excess factor with purpose of its improvement. That will open opportunities of its practical use in hydrology. It will be executed on actual series of observations, which have rather large duration and are located in different regions of the world (see Table 2). The account was reduced to calculation of the derived characteristics in process of lengthening of series by 10 yr (last 30 yr – by 1 yr). In several decades the curve dependences of the derived characteristics from increasing length of the series were stabilized, by testing small fluctuations (example for Eh see Fig. 2). For these stabilized sites (for uniformity the interval of 30 yr long was taken) the meanings of variation factors describing an error of received estimations of four considered characteristics were calculated. To a line of the excess factors appropriate to this site, it is possible to apply the formula of estimation of the calculation error of the norm $\overline{Eh}$: $\sigma_{\overline{(Eh)}_n} = |C_v|/\sqrt{30}$ (Table 2). From the table it could be seen, that the error varies in limits from 1.1 % up to 30.9 % at average meaning 8.0 %. It is quite acceptable to application Eh in engineering calculations. From the practical point of view it will result in an opportunity to make more reasonable estimations of the characteristics of emissions of casual process, which result in catastrophic floods.

4   Conclusions

As a result of work performed, established:

1. Basic parameter in genetic model of flow formation, answering for the excess factor, is the intensity of internal white noise ($G_{\tilde{c}}$) of watershed. With its increase the tails of probabilistic distributions aggravate (making more probable hydrological accidents), and at the large intensity of relative multiplicative noises of the equation decisions for the initial moments become unstable, which takes density distributions of probability over the class of K. Pearson's distributions.

2. It appears durations of series of observations of a flow in some decades in most cases is sufficient for rather reliable definition of the excess factor (its variation factor after stabilization of fluctuations of the last makes an average of 0.44).

3. The occurrence of partially infinite hydrology was connected to theoretical and experimental revealing of conditions, at which senior moments of probabilistic distributions used in engineering hydrology, appear unstable, making unreasonable application of distributions of Pearson's III types and its modifications offered by S. N. Kritskii and M. F. Menkel (BC 435-72, 1972). Identifier of these conditions is the relative intensity of multiplicative noises which are included in genetic formation model of density of probability of a long-term flow. It is shown in this article that practically its acceptable analogue could be the excess factor, which meanwhile is ignored as unreliably determined parameter.

References

BC 435-72: Guidelines for determination of design hydrological characteristics, Leningrad, Gidrometeoizdat, 17 pp., 1972.

Budiko, M. I.: Climate change, Leningrad, Gidrometeoizdat, 280 pp., 1974.

5 Directory of applied statistics: T1: Moscow: Finance and statistics, Translated from English by: Lloid, A., Lederman, U., and Turin, U. N., 510 pp., 1989.

Klibashev, K. P. and Goroshkov, I. F.: Hydrological accounts, Leningrad, Gidrometeoizdat, 460 pp., 1970.

Kovalenko, V. V.: Modeling of hydrological processes, St.-Petersburg, Gidrometeoizdat, 256 pp.,

10 1993.

Kovalenko, V. V.: Partially infinitive hydrology, St.-Petersburg, RSHU, 230 pp., 2007.

Kovalenko, V. V. and Haustov, V. A.: Criterion of steady development of hydrological processes and mapping of zones of expected anomalies of parameters of an annual flow of the rivers of CIS at antropogeneous change of a climate, Meteorology and hydrology, 12, 96–102, 1998.

15 Standard 8.508-84: Metrological characteristic of measurement means and of accuracy characteristic of automation means of GSP, General methods of estimation and control/State system of maintenance of measurement unity, – Moscow: IPK Publishing House of the Standards, 39 pp., 2002.

Table 1. Relative average quadratic errors (%) of calculation of the derived hydrological characteristics.

The settlement characteristic	Number of years of observations					
	10	20	30	40	50	100
m_1	12.6	8.9	7.6	6.3	5.8	4.0
C_v	24	17	14	12	11	7,6
C_s (at $C_s = 2C_v$)	61	49	41	36	33	24
Eh	92	84	75	68	62	46

Table 2. Estimation of errors of calculation of the excess factors on the stabilized sites of their dependences from duration of observation.

| River, station	The total length of the series	The basin area, km^2	Error $\sigma_{\overline{(Eh)}_{30}}$
Susquehanna, Harrisburg, USA	93	62 419	23.0
Massena, Massena, USA	123	764 600	8.0
Elbe, Decin, Czechoslovakia	134	51 104	2.9
Vuoksi, Imatra, Finland	138	61 275	2.2
Loire, Blois, France	117	38 240	4.2
Loire, Montjean, France	117	110 000	1.1
Danube, Drobeta-Turnu, Romania	149	576 232	2.7
Vannern, Vanersborg, Sweden	178	46 830	4.4
Niagara, Queenston, Canada	125	686 000	5.7
Columbia, Columbia, USA	109	613 830	30.9
Unzha, Makarov, Russia	96	18 500	4.0
Tikhvinka, Goreluha,Russia	99	3910	7.3
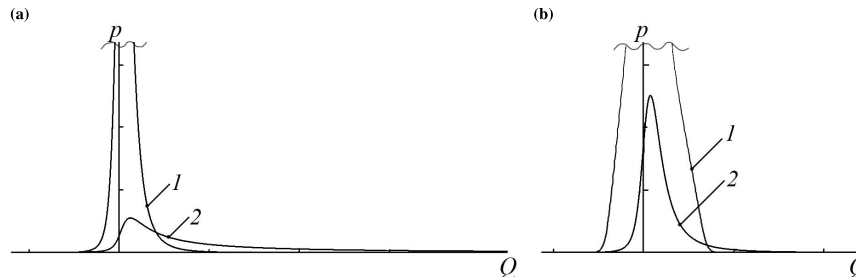
Fig. 1. Influence of parameters of genetic model of formation of probabilistic mode of a flow on form of density curve of probability: **(a)** at different meanings of parameter $\bar{c}$ ($\bar{c}_1 > \bar{c}_2$); **(b)** at different meanings of intensity of noise $G_{\tilde{c}}$ ($G_{\tilde{c}_2} > G_{\tilde{c}_1}$) (in the given inequalities indexes 1 and 2 correspond to curves in **a** and **b**).
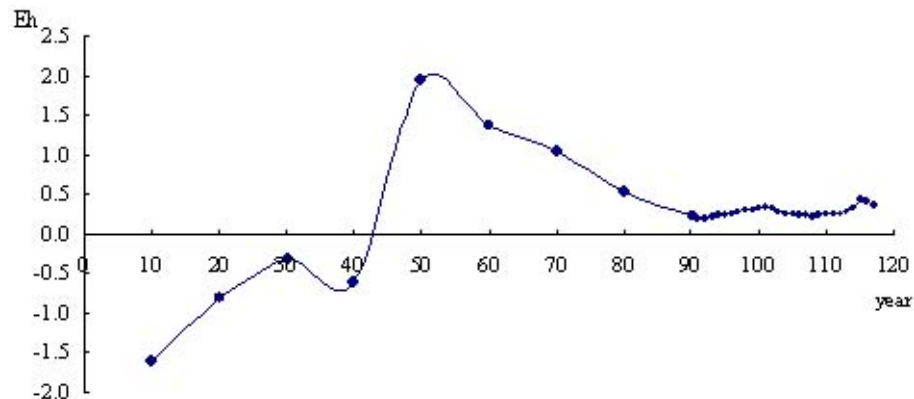
Fig. 2. Dependence of the excess factor on number of the members of the series used in calculations (R. Loire – St. Blois, France).