Environmental approaches when determining the calculated hydrological characteristics

A V Zvyagintseva¹, S A Sazonova², O I Marar³ and L V Koval⁴

¹Department of Chemistry and Chemical Technology, Voronezh State Technical University, 84, October 20th Anniversary Street, Voronezh, 394006, Russia
²Department of Technosphere and Fire Safety, Voronezh State Technical University, 84, October 20th Anniversary Street, Voronezh, 394006, Russia
³Department of Modeling and Organization of Social Systems, The Russian Presidential Academy of National Economy and Public Administration, 143, Moskovsky Prospect, Voronezh, 394005, Russia
⁴Department of Philosophy, Sociology and History, Voronezh State Technical University, 84, October 20th Anniversary Street, Voronezh, 394006, Russia

⁵E-mail: zvygincevaav@mail.ru

Abstract. The issues of forecasting dangerous hydrological phenomena in water bodies in the presence of hydrometrically observed data are considered. The analytical distribution functions of annual excess probabilities are applied - sufficiency curves. The features of calculating the empirical annual probability of exceeding hydrological characteristics, variation coefficients and asymmetries for distribution, the scattering of estimates and other distribution parameters are considered. In the case of heterogeneity of the initial data of hydrometric observations, when the series under consideration consists of heterogeneous elements of the hydrological regime, empirical and analytical distribution curves are set separately for each homogeneous totality. Based on the considered data, a system for monitoring and forecasting emergencies of a hydrological nature at water bodies is being constructed.

1. Introduction
The determination of the calculated hydrological characteristics in the presence of hydrometric observations of sufficient duration is carried out by applying the analytical distribution functions of the annual exceedance probabilities - sufficiency curves [1, 2, 3].

The length of the observation period is considered sufficient if the considered period is representative (personable), and the relative mean square error of the calculated value of the studied hydrological characteristics does not exceed 10% for annual and seasonal flows and 20% for the maximum and minimum flows. If the relative mean square errors exceed the indicated limits and the observation period is not representative, it is necessary to bring the hydrological characteristic under consideration to a multi-year period [4, 5, 6].

2. Method
The empirical annual probability of exceeding $P_m,\%$ [7, 8, 9] hydrological characteristics is determined by the formula:
where $m$ is the ordinal number of members of a series of hydrological characteristics arranged in descending order; $n$ is the total number of members of the series [10, 11, 12].

Empirical distribution curves of annual exceedance probabilities are constructed using probability papers. The type of probability paper is selected in accordance with the accepted analytical probability distribution function and the obtained ratio of the asymmetry coefficient $C_s$ to the coefficient of variation $C_v$.

To smooth and extrapolate the empirical curves of the distribution of annual exceedance probabilities, as a rule, three-parameter distributions are used: Kritsky–Menkel distributions for any $C_s / C_v$ ratio, Pearson type III distribution (binomial curve) for $C_s / C_v \geq 2$, log-normal distribution for $C_s \geq (3C_v + C_{v3})$ and other distributions having a range of existence of random variable from zero or positive value [13, 14, 15]. With proper justification, it is allowed to apply two-parameter distributions if the empirical ratio $C_s / C_v$ and the analytical ratio $C_s / C_v$ inherent in a given distribution function are approximately equal. If a number of hydrometric observations are heterogeneous (different conditions for the formation of runoff), truncated and composite probability distribution curves are used [16, 17, 18].

Estimates of the parameters of the analytical distribution curves: the long-term average value, the coefficient of variation $C_v$, and the ratio of the coefficient of asymmetry to the coefficient of variation $C_s / C_v$, are established from the series of observations of the hydrological characteristic under consideration by the method of approximately maximum likelihood and the method of moments [19, 20, 21]. At the initial stages of design, the use of the graphoanalytical method (quantile method) is allowed [22, 23, 24].

The coefficient of variation $C_v$ and the asymmetry coefficient $C_s$ for the three-parameter Kritsky–Menkel gamma distribution should be determined by the approximate maximum likelihood method depending on the statistics $\lambda_2$ and $\lambda_3$ calculated by the formulas:

\[
\lambda_2 = \left( \sum_{i=1}^{n} \log k_i \right) / (n-1); \\
\lambda_3 = \left( \sum_{i=1}^{n} k_i \log k_i \right) / (n-1),
\]

where $k_i$ is the modular coefficient of the considered hydrological characteristics, determined by the formula:

\[
k_i = \frac{Q_i}{\bar{Q}},
\]

here $Q_i$ - annual values of water consumption; $\bar{Q}$ - the arithmetic mean value of water consumption, determined depending on the number of years of hydrometric observations by the formula:

\[
\bar{Q} = \frac{\sum Q_i}{n}.
\]

Coefficients of variation and asymmetry are determined by nomograms based on the obtained statistical data $\lambda_2$ and $\lambda_3$. The coefficients of variation of $C_x$ and asymmetries of $C_v$ are determined by the method of moments according to the formulas:

\[
C_v = (a_1 + a_2 / n) + (a_3 + a_4 / n) \tilde{C}_v + (a_5 + a_6 / n) \tilde{C}_{v3}; \\
C_s = (b_1 + b_2 / n) + (b_3 + b_4 / n) \tilde{C}_s + (b_5 + b_6 / n) \tilde{C}_{s3},
\]

where $\tilde{C}_v$ and $\tilde{C}_s$ are the long-term average values of $C_v$ and $C_s$, respectively.
where \( a_1, ..., a_6; b_1, ..., b_6 \) are the distribution coefficients of Pearson type III and the Kritsky – Menkel distribution; \( \hat{C}_v \) and \( \hat{C}_s \) are respectively mixed estimates of the coefficients of variation and asymmetry, determined by the formulas:

\[
\hat{C}_v = \sqrt{\frac{\sum_{i=1}^{n} (k_i - 1)^2}{n - 1}},
\]

\[
\hat{C}_s = \frac{n \sum_{i=1}^{n} (k_i - 1)^2}{C_v^2 (n - 1)(n - 2)}.
\]

If \( C_v < 0.6 \) and \( C_s < 1.0 \), the coefficients of variation and asymmetry can be determined by formulas (8) and (9) without introducing corrections [25, 26, 27].

The scattering of estimates caused by the limitedness observation data is denoted by \( \varepsilon_{\text{rand}} \), and the scattering due to differences between catchments that were not eliminated during the preparation process, by \( \varepsilon_{\text{geogr}} \), the total variance of the estimate of \( \varepsilon_{\text{total}} \) consists of two components:

\[
\varepsilon_{\text{total}}^2 = \varepsilon_{\text{rand}}^2 + \varepsilon_{\text{geogr}}^2.
\]

The total variance of the estimate of \( \varepsilon_{\text{total}}^2 \) is determined by the formula:

\[
\varepsilon_{\text{total}}^2 = \frac{\sum_{i=1}^{k} (A_i - \bar{A})^2}{k - 1},
\]

where \( i \) is the index (number) of the object. An object is understood either as a catchment basin or a meteorological station; \( k \) is the number of jointly analyzed objects; \( A_i \) is an estimate of the considered parameter of the \( i \)-th object; \( \bar{A} \) - the average rating for all objects [28, 29, 30].

The random component of the scattering of the estimates of \( \varepsilon_{\text{rand}}^2 \) is calculated by averaging the variances of the estimates of these parameters according to the theoretical formulas obtained for individual objects (26) - (28), or according to the results of statistical tests [31, 32, 33].

The geographic component of scattering \( \varepsilon_{\text{geogr}}^2 \) is determined by (10) as the difference between the total and random dispersions. If the estimate \( \varepsilon_{\text{geogr}}^2 \) has a negative sign, then it is taken equal to zero. The dispersion of the result of the joint calculation is determined by the formula:

\[
\varepsilon_{\text{avg}}^2 = \frac{\varepsilon_{\text{rand}}^2 + \varepsilon_{\text{geogr}}^2}{k}.
\]

The correlation between random and geographical components determines the rational composition of objects processed by the method of group assessment. With an increase in the number of collectively analyzed catchments, the random error component decreases [34, 35, 36]. The geographical component should increase due to the involvement of catchments located within a wider geographical area, the conditions, the flow formation of which differ more significantly. Permissible (acceptable) should be considered the number of catchments at which the geographic component does not exceed the random:

\[
\varepsilon_{\text{geogr}} \leq \varepsilon_{\text{rand}}.
\]

The result of the group analysis is estimation of the parameter by the set of own and consolidated observations in the form of the weighted average accuracy of each of the estimates:

\[
A_{\text{comb}} = \frac{A_{\text{own}} \varepsilon_{\text{own}}^2 + A \varepsilon_{\text{cons}}^2}{\varepsilon_{\text{own}}^2 + \varepsilon_{\text{cons}}^2}.
\]
The standard error of such an estimate is calculated by the formula:

\[ \varepsilon'_{\text{comb}} = \sqrt{\varepsilon_{\text{own}}^2 + \varepsilon_{\text{comb}}^2} \]  

(15)

For asymmetry estimates and autocorrelation coefficients, the result of the group analysis is the average of all individual estimates within a homogeneous area [37, 38, 39].

The procedure for performing group analysis (taking into account spatial correlation of observation data) is as follows: parameters of hydrological characteristics distribution used for joint analysis and necessary for calculation of standard errors of parameter \( A \) by formulas (26), (28) are determined for each catchment; cross-row correlation coefficients \( R_{ij}(x) \) are estimated for each pair of catchments; the average value of the parameter is estimated from the sample of values \( A_i \):

\[ \bar{A} = \frac{\sum_{k=1}^{k} A_i}{k} \]  

(16)

and the total dispersion of \( \varepsilon'_{\text{total}} \) is determined by formula (11); the values of the correlation coefficients \( R_{ij}(A) \) between estimates of parameter \( A \) are determined by theoretical dependencies; using samples of volume \( n \), the standard deviation \( \varepsilon_{\text{indep}}(A) \) of the estimates of parameter \( A \) is determined, which characterizes the scattering of estimates for the case of independent samples and is determined by formulas (26), (28) or the results of statistical tests; the standard deviation of the parameter \( \varepsilon_{\text{indep}}(A) \), characterizing independent samples, is adjusted by a value that takes into account the effect of correlation between the combined objects:

\[ \varepsilon_{\text{mod}}(A) = \varepsilon_{\text{indep}}(A) \sqrt{1 - r_{\text{ave}}(A)} \]  

(17)

where \( r_{\text{ave}} \) is the average value of the correlation coefficient between estimates of parameter \( A \) for all \( k \) catchments. The found value of the random component is used to calculate the geographical component according to the formula (10); if condition (13) is fulfilled, then using formulas (14) and (15), the error of the result of the combined calculation, the weighted average accuracy estimate and its standard error are calculated [40].

At the initial stages of design, it is allowed to determine the parameters of the binomial distribution by the graphoanalytical method using the formulas:

\[ S = \frac{Q_5 + Q_{95} - 2Q_{50}}{(Q_5 - Q_{95})}; \]  

(18)

\[ \sigma = \frac{Q_5 - Q_{95}}{(F_5 - F_{95})}; \]  

(19)

\[ \bar{Q} = Q_{50} - F_{50} \sigma, \]  

(20)

where \( Q_5, Q_{50}, Q_{95} \) - values of water flow rates of probability of exceeding 5%, 50%, 95%, respectively, determined by smoothed empirical distribution curve; \( F_5, F_{50}, F_{95} \) are normalized ordinates of binomial distribution curve corresponding to calculated value of curvature coefficient \( S \). Value of asymmetry coefficient \( C_s \) is determined from functional dependence on coefficient \( S \).

The general curve of the distribution of probability of excess is calculated on the basis of the curves established for homogeneous elements in one of two ways:

a) if every year there are observations of all the homogeneous elements of the river’s water regime \( (n_1 = n_2 = n_3 = n) \), the annual probability of exceeding \( P\% \) of the considered hydrological characteristic at any value is determined by the formula:

\[ P = [1 - (1 - P_1)(1 - P_2)(1 - P_3)] \times 100, \]  

(21)

where \( P_1, P_2, P_3 \) - annual probabilities of exceeding homogeneous elements

For two homogeneous hydrological characteristics, formula (21) takes the form:
\[ P = (P_1 + P_2 - P_1 P_2) 100; \]  

(22)

b) if in each year there is only one value of the element of the hydrological characteristic under consideration, the annual probabilities of exceeding for any value are determined by the formula:

\[ P = \frac{n_1 P_1 + n_2 P_2 + n_3 P_3}{n_1 + n_2 + n_3}, \]  

(23)

where \( n_1, n_2, n_3 \) is the number of members of homogeneous elements. For two genetically homogeneous elements, formula (23) takes the form:

\[ P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}. \]  

(24)

If there are zero values in a series of observations of the hydrological characteristic under consideration (for example, minimum water discharge), the annual probability of exceeding is determined by the formula:

\[ P = \frac{n_1 P_1}{n_1 + n_2}. \]  

(25)

The probabilities of exceeding \( P_1, P_2, P_3 \) in formulas (21) and (22) are expressed in fractions of a unit, and in formulas (23) - (25) as a percentage.

Random mean square errors of sample means are determined by an approximate relationship:

\[ \sigma^2_\bar{X} = (\sigma_0^2 / \sqrt{n}) \left( \frac{1 + r}{1 - r} \right), \]  

(26)

which is used when the autocorrelation coefficient between adjacent members of the series \( r \) is less than 0.5. For large autocorrelation coefficients, the formula is used:

\[ \sigma^2_\bar{X} = (\sigma_0^2 / \sqrt{n}) \left[ \frac{1 + \frac{2r}{n(1-r)}}{\left( \frac{1}{1-r} \right) - \frac{n(1-r)}{1-r}} \right]. \]  

(27)

Random mean square errors of the variation coefficients at \( C_s = 2C_v \) are determined by the dependence:

\[ \sigma_{C_v} = \frac{C_v}{n + 4C_v^2} \sqrt{\frac{n(1 + C_v^2)}{2} \left( 1 + \frac{3C_v^2}{1 + r} \right)} \]  

(28)

If there is reliable information about random relative mean square errors of the initial data of hydrometric observations, the estimates of the coefficients of variation and asymmetry are specified by the formulas:

\[ C_v = \sqrt{\left( C_v^2 - \sigma^2_0 \right)/\left( 1 + \sigma^2_0 \right)}; \]  

(29)

\[ C_y = \frac{1}{1 + 3\sigma^2_0} \sqrt{\left( \frac{C_v^2 - \sigma^2_0}{C_v^2 - \sigma^2_0} \left[ \frac{C_v^2 (1 + \sigma^2_0)}{C_v^2 - \sigma^2_0} - C_v^2 - 6\sigma^2_0 \right] \right)} \]  

(30)

Where \( C_v, C_y, \sigma_0 \) respectively, are the coefficients of variation and asymmetry calculated from the observed values; \( \sigma_0 \) - random relative (in fractions of a unit) mean square error of the initial data of hydrometric observations.
The parameters of the distribution curves of hydrological characteristics in the presence of reasonable information about the outstanding values of river flow are determined as follows [41].

When taking into account one outstanding value of the hydrological characteristic that is not included in the continuous n-year series of hydrometric observation data:

a) by the method of approximate maximum likelihood, depending on the statistics $\lambda_2$, and $\lambda_3$, determined by the formulas:

$$\lambda_2 = \frac{1}{N} \left( \frac{\log Q_n}{Q} + \frac{N-1}{n-1} \sum_{i=1}^{n} \frac{\log Q_i}{Q} \right),$$  \hspace{1cm} (31)

$$\lambda_3 = \frac{1}{N} \left( \frac{Q_n \log Q_n}{Q} + \frac{N-1}{n-1} \sum_{i=1}^{n} \frac{Q_i \log Q_i}{Q} \right).$$  \hspace{1cm} (32)

b) by method of moments - according to the formulas:

$$\bar{Q} = \frac{1}{N} \left( Q_n + \frac{N-1}{n} \sum_{i=1}^{n} Q_i \right),$$  \hspace{1cm} (33)

$$C_v = \sqrt{\frac{1}{N} \left( \frac{Q_n}{Q} - 1 \right)^2 + \frac{N-1}{n-1} \sum_{i=1}^{n} \left( \frac{Q_i}{Q} - 1 \right)^2}. \hspace{1cm} (34)$$

When taking into account one outstanding value of the hydrological characteristic included in the n-year series of hydrometric observation data:

a) by the method of approximate maximum likelihood, determined by the formulas, depending on the statistics $\lambda_2$, and $\lambda_3$:

$$\lambda_2 = \frac{1}{N} \left( \log \frac{Q_n}{Q} + \frac{N-1}{n-1} \sum_{i=1}^{n} \log \frac{Q_i}{Q} \right),$$ \hspace{1cm} (35)

$$\lambda_3 = \frac{1}{N} \left( \frac{Q_n \log Q_n}{Q} + \frac{N-1}{n-1} \sum_{i=1}^{n} \frac{Q_i \log Q_i}{Q} \right).$$ \hspace{1cm} (36)

b) the method of moments - according to the formulas:

$$\bar{Q} = \frac{1}{N} \left( Q_n + \frac{N-1}{n} \sum_{i=1}^{n} Q_i \right),$$ \hspace{1cm} (37)

$$C_v = \sqrt{\frac{1}{N} \left( \frac{Q_n}{Q} - 1 \right)^2 + \frac{N-1}{n-1} \sum_{i=1}^{n} \left( \frac{Q_i}{Q} - 1 \right)^2}. \hspace{1cm} (38)$$

In formulas (31) - (38): $\bar{Q}$ - arithmetic mean value calculated taking into account the outstanding value of water flow; n is the number of years of continuous observations; N is the number of years during which the outstanding value of the hydrological characteristic has not been exceeded. The use of formulas (31) - (38) is allowed only if the historical information about the outstanding hydrological value and the number of years it was not exceeded is sufficiently justified. A random $Q_N$ job is not allowed.

Lateral affluent between adjacent transit lines of water objects is determined by one of the following
methods: by summation of expenses taking into account lag time of the water of tributaries, inflowing in the interval between two transit lines; by difference of average water consumption in the lower and upper transit lines of the interval of the river; by method of conditional water balance; by modulus of runoff, determined by the map for part of the area [42].

3. Conclusion
If the data of hydrometric observations are insufficient, the parameters of probability distribution curves of hydrological characteristics, as well as the main elements of the calculated hydrograph, should be brought to a multi-year period with the involvement of observation data of points - analogues.

The need for the given hydrological characteristic arises in cases when the mean square error of the design value of the hydrological characteristic exceeds 10% for annual and seasonal runoff, 20% for maximum and minimum runoffs.

Thus, taking into account the considered data, a system of monitoring and forecasting of emergencies, hydrological on water bodies, is being built.

References
[1] Kurdo A G 1995 Water resources of the Voronezh region: the formation of anthropogenic impact, protection and calculations (Voronezh: Voronezh University Press) 224
[2] Zvyagintseva A V, Neizhmak A N and Rastorguev I P 2013 Convective natural disaster monitoring based on remote sensing data from meteorological spacecraft (Voronezh: Voronezh State Technical University Press) 162
[3] Kotova V P 2008 Protection and actions of the population during natural disasters of a hydrological nature and hydrodynamic accidents (Voronezh: Training and methodological center of the State Emergencies Ministry of the Voronezh region Press)
[4] Zvyagintseva A V, Tenkaeva A S and Mozgovoy N V 2015 Bulletin of the Samara Scientific Center of the Russian Academy of Sciences 17(5) 276-282
[5] Dolzhenkova V V and Zvyagintseva A V 2015 Bulletin of the Samara Scientific Center of the Russian Academy of Sciences 17(6) 70-81
[6] Zvyagintseva A V, Sazonova S A and Kulneva V V 2020 IOP Conference Series: Earth and Environmental Science 459(5) 052047
[7] Zvyagintseva A V and Samofalova A S 2020 AIP Conference Proceedings 2313 060020
[8] Zvyagintseva A V 2020 Bulletin of the Russian Academy of Sciences: Physics 84 9 1097-9
[9] Zvyagintseva A V 2020 IOP Conference Series: Materials Science and Engineering 919 62054
[10] Zvyagintseva A V 2020 International Journal of Hydrogen Energy 45, 46 24991-25001
[11] Zvyagintseva A V and Kravtsova Y G 2007 NATO Security through Science Series A: Chemistry and Biology 661-4
[12] Zvyagintseva A V and Kravtsova Y G 2007 NATO Security through Science Series A: Chemistry and Biology 665-9
[13] Zvyagintseva A V, Kulneva V V and Sazonova S A 2020 IOP Conference Series: Materials Science and Engineering 919(6) 62053
[14] Zvyagintseva A V, Samofalova A S, Sazonova S A and Kulneva V V 2020 Journal of Physics: Conference Series 1679(2) 22076
[15] Kulneva V V, Zvyagintseva A V, Sazonova S A and Akamsina N V 2020 Journal of Physics Conference Series 1679(2) 22077
[16] Zvyagintseva A V, Sazonova S A, Kulneva V V and Panteleev I N 2020 IOP Conference Series: Materials Science and Engineering 919(6) 62034
[17] Zvyagintseva A V, Sazonova S A, Kulneva V V and Panteleev I N 2020 IOP Conference Series: Materials Science and Engineering 919(6) 62036
[18] Zvyagintseva A V, Sazonova S A and Kulneva V V 2020 IOP Conference Series: Materials Science and Engineering 962(4) 042045
[19] Zvyagintseva A V, Sazonova S A and Kulneva V V 2020 IOP Conference Series: Materials
Science and Engineering 962(4) 042066
[20] Zvyagintseva A V, Sazonova S A and Kulneva V V 2020 IOP Conference Series: Materials Science and Engineering 962(4) 042067
[21] Zvyagintseva A V, Sazonova S A, Kulneva V V, Asminin V F and Zyazina T V 2021 IOP Conference Series: Materials Science and Engineering 1047(1) 12192
[22] Kulneva V V, Zvyagintseva A V and Sazonova S A 2021 IOP Conference Series: Earth and Environmental Science 666(2) 022035
[23] Zvyagintseva A V, Sazonova S A and Kulneva V V 2021 IOP Conference Series: Earth and Environmental Science 666(2) 022036
[24] Sysoev D V, Sysoeva A A, Sazonova S A, Zvyagintseva A V and Mozgovoj N V 2021 IOP Conference Series: Materials Science and Engineering 1047(1) 12195
[25] Tereshchenko M A, Bychenok V I and Mozgovoi N V 2009 Thermal Engineering 56 6 522-5
[26] Kuzminov A, Sakharova L, Stryukov M and Zolnikov V K 2020 IOP Conference Series: Earth and Environmental Science 595(1) 012007
[27] Zolnikov V K, Kryukov V P, Achkasov V N and Smerek V A 2011 Modeling of Systems and Processes 1-2 24-6
[28] Zolnikov V K 2012 Modeling of Systems and Processes 1 7-30
[29] Antsiferova V I and Zolnikov V K 2009 Modeling of Systems and Processes 3-4 5-12
[30] Yankov A I, Smerek V A, Kryukov V P and Zolnikov V K 2012 Modeling of Systems and Processes 1 92-5
[31] Novikov A I, Zolnikov V K and Novikova T P 2021 Inventions 7 6(1) 1–11
[32] Sakharova L, Stryukov M and Zolnikov V K 2018 Journal of Physics: Conference Series 973(1) 012045
[33] Belokurov V P, Belokurov S V and Zolnikov V K 2018 Transportation Research Procedia 36 44-9
[34] Smolentseva T E, Sumin V I, Zolnikov V K and Lavlinsky V V 2018 Journal of Physics: Conference Series 973(1) 012045
[35] Belokurov S V, Belokurov V P, Zolnikov V K and Cherkasov O N 2017 Transportation Research Procedia 20 47-52
[36] Vasilyev A V 2020 IOP Conference Series: Earth and Environmental Science 607(1) 012015
[37] Vasilyev A V 2020 IOP Conference Series: Earth and Environmental Science 607(1) 012016
[38] Vasilyev A V 2020 International Conference on Dynamics and Vibroacoustics of Machines, DFM 2020 9243902
[39] Vasilyev A V 2020 International Conference on Dynamics and Vibroacoustics of Machines, DFM 2020 9243900
[40] Vasilyev A V 2020 International Conference on Dynamics and Vibroacoustics of Machines, DFM 2020 9243907
[41] Vasilyev A V 2018 IEEE International Conference 95-7
[42] Vasilyev A V 2018 IEEE International Conference 98-101