Modeling of rare rain floods on the example of middle rivers of Angara basin

A Yu Belyakova, Ya M Ivanyo and S A Petrova*

Department of informatics and mathematical modeling, FSBEI HE Irkutsk SAU, Irkutsk district, Irkutsk region, Russia

*E-mail: sofia.registration@mail.ru

Abstract. The paper presents the results of creating mathematical models for predicting the maximum water discharge of rare rain floods and an algorithm for the probabilistic assessment of such hydrological phenomena. The developed models are implemented to describe the catastrophic flood observed on the river Iya at Tulun. The prognostic and actual values of the maximum water levels are compared, and the likelihood of a rare hydrological event with a dispersion interval for given significance levels is obtained. The results are important for warning the public about the dangers and designing protective structures.

1. Introduction

In [1], it is proposed to simulate future situations by evaluating the adequacy of the models used based on a comparison of model parameters with previous data. At the same time, the authors of the article associate scenario of events with climate warming. According to [2], an analysis of the risks of floods caused by liquid precipitation in seven regions of the world is given. It is estimated that the number of local hydrological events may increase. The authors of [3], when describing the floods on the Indigirka River observed in the spring and summer of 2017, note that their recurrence will decrease in the future.

Climate variability also applies to Eastern Siberia, so it is possible to form more often than usual rare climatic events, which include rain floods. Analysis of rare rain floods on the rivers of the Angara and Lena basins with catchment areas of 8420-24700 km2 shows that approximately every 12 years, one high flood with the highest maximum water discharge during the observation period is formed on one of the left tributaries of the Angara and in the upper Lena. Of particular note is the catastrophic rain flood on the river Iya, formed at the end of June 2019, resulting in human casualties and damages, amounting to tens of billions of rubles. In addition, high floods were observed on the rivers Oka, Uda and Belaya.

Obviously, the probabilistic assessment of such a catastrophic phenomenon is important for determining a set of measures to minimize future damage and prevent loss of life. The forecasting of such phenomena is associated with the operational actions of decision makers. At the same time, the calculated values of the levels and maximum water discharge are necessary for the design of water management systems, roads, transmission lines and other objects.

The aim of the study is to create models for predicting the rare maximum water flow rates and levels from previous precipitation and temperatures recorded at remote observation sites from the hydrological target under consideration, as well as a stochastic estimate of the probability of not exceeding the maximum flow.
To achieve the goal, the following tasks were solved:

- development of forecasting models for the rare maximum water discharge caused by rain floods in the middle rivers of the Angarsk and Lena basins;
- a statistical assessment of the rare maximum water discharge in the summer; on the example of a rain flood formed on the river Iya in 2019;
- the implementation of forecasting models and determining the value and probability of occurrence of a rare maximum water flow rate on the example of the river Iya at Tulun.

2. Materials and methods

The article [4] mentions the development of forecasting the level in the river and groundwater with a lead time of 1-3 months. In [5–8 et al.], Results are presented on modeling the maximum runoff of rain floods as applied to rivers of different regions. In addition to using the above sources, this article continues the authors’ own research [9]. During its preparation, the materials of the hydrometeorological stations of the Irkutsk region on daily precipitation and temperatures in the warm period during the formation of rare rain floods were collected. The data on long-term maximum water discharge on the Lena (Kachug), Irkut (Irkutsk), Kitoi (Angarsk), Belaya (Michelevka), Iya (Tulun), Biryusa (Biryusinsk), Uda (Ukar) rivers were used.

Since the work evaluated the recurrence of rare hydrological phenomena, together with the data of hydrometeorological observations, historical-archival evidence systematized in was considered [10].

To solve the problems of forecasting rare floods, a correlation and regression analysis was used, as a result of which the dependences of the maximum water discharge on the previous values of precipitation and temperatures recorded in remote areas from the studied point and in the observation point were built.

When assessing the probability of the formation of the largest hydrological event, the methods of constructing the laws of distribution and scattering estimation, as well as the Monte Carlo method, were applied.

The paper proposes an algorithm for assessing the rare hydrological phenomenon of 2019 based on the relationship of the levels and maximum water discharge for events corresponding to and exceeding critical water levels.

3. Results and discussion

When forecasting high maximum water flow rates, two tasks are solved:

- the construction of mathematical models according to which predictive values can be obtained;
- an assessment of the significance and probability of the formation or repeatability of a hydrological phenomenon.

Consider the second problem on the example of rain flood on the river Iya, which led to disastrous consequences. Table 1 shows the regression equations characterizing the relationship between the modular coefficients of long-term maximum water discharge $k_Q = Q_i / Q$ and modular coefficients of water levels $k_H = H_i / \bar{H}$. In these formulas, $Q_i, H_i$ are the annual highest water discharges and levels, and $\bar{Q}, \bar{H}$ are the average values of long-term series of maximum water discharges and levels of rain floods.

Table 1 shows the calculation results for two options for determining the highest water discharge of rain flood on the basis of hydrological events and annual phenomena. In the second case (second row), annual levels and maximum water discharges corresponding to the highest values in the summer period are considered. In the second case, from a long-term series of water levels and discharges, events are distinguished by which the dependence $Q = f(H)$ is built.
To transfer the phenomenon into an event, a criterion $H \geq H_{cr}$ was used, where $H_{cr}$ is the level of transition of the value to the event - the river overflows from the banks and flooding of the floodplain zones. For two calculation options, the determination coefficient $R^2$ is high. Moreover, the parabolic regression equations according to the Fisher F-criterion are significant, as are the coefficients for unknowns according Student's t-test.

Based on the obtained models, an urgent maximum water discharge was determined for the catastrophic flood of the Iya River at Tulun. According to the actual highest water level, it amounted to 6330 and 6380 m$^3$/s.

A series of events is more consistent with a homogeneous sample than the sequence of all annual maximum water discharges, including values not exceeding the critical level. Therefore, it is proposed to use the first value of the given options, 6330 m$^3$/s, as the maximum water discharge for the observation period.

**Table 1.** Assessment of the urgent maximum water discharge observed on the river Iya at Tulun in June 2019 according to 1932-2019.

| Parameter   | Regression equation | $\bar{Q}$, m$^3$/s | $Q$, m$^3$/s | $R^2$ |
|-------------|---------------------|---------------------|--------------|-------|
| Events      | $k_Q = 1.681-4.2907k_H + 3.55708k_H^2$ | 1265 | 6330 | 0.91 |
| Phenomenon  | $k_Q = 2.301 - 5.22k_H + 3.837k_H^2$ | 1088 | 6380 | 0.87 |

Having received the maximum water discharge, it is possible to evaluate the probability of occurrence of the found value. To achieve this goal, a three-parameter power gamma distribution law was used [11]. Its probability density has the form

$$p = \left[ \frac{\Gamma(\gamma+b)}{\Gamma(\gamma)} \right]^{\gamma/b} \frac{1}{\Gamma(\gamma)b\bar{Q}} \left( \frac{Q}{\bar{Q}} \right)^{\gamma/b-1} \times \exp \left\{ - \left[ \frac{Q}{\bar{Q}} \frac{\Gamma(\gamma+b)}{\Gamma(\gamma)} \right]^{1/b} \right\},$$

(1)

where $\bar{Q}$ is the average value of the distribution; $\gamma$, $b$ are distribution parameters associated with the coefficient of variation $c_v$ and asymmetry $c_s$; $\Gamma(\gamma)$, $\Gamma(\gamma+b)$ are gamma function with corresponding arguments.

The use of a three-parameter power gamma probability distribution to describe the empirical values of the maximum water discharge of rain flood observed on the river Iya at Tulun, due to:

- development and the possibility of widespread use of analytical expression (1);
- the applicability of formula (1) for series with a high asymmetry parameter;
- recommendations of regulatory documents on the use of distribution [12, 13].

In figure 1 shows the function of a three-parameter power gamma distribution constructed from empirical data (1) for two variants. In the first case, the statistical parameters of a number of maximum water discharge of rain floods were estimated using the moment method without taking into account the bias (2). In the second case (3), The method of A.V. Rozhdestvensky was used to eliminate the negative bias [14].
The results of determining the probabilities of the occurrence of a rare hydrological event are shown in table 2. According to the results obtained, the frequency of occurrence of such a hydrological phenomenon was 424 years (the probability of occurrence is \( p = 0.00236 \)).

Meanwhile, fluctuations in the probability of occurrence of an event for significance levels \( \alpha = 0.05 \) and 0.95, obtained according to a power law [11]

\[
p_{\alpha} = 1 - (1 - \alpha)^{1/n}
\]

are significant (table 2). In the formula (2), \( n \) is the number of sample values equal to 90.

To this we add the value of the empirical probability of excess based on historical and archival evidence collected and systematized in [10]. In the XX and XXI centuries the highest maximum water discharge is the value of the end of June 2019. It should be noted that in the XIX century three high rain floods were observed on the river Iya (1820, 1851 and 1870). In other words, the period of not exceeding the 2019 flood \( N \) can correspond to 200, 169 and 150 years. Based on the analysis of extreme hydrological phenomena, a situation similar to 2019 was observed in 1870. It was noted in the Irkutsk Provincial Gazette [15] that, near the village of Tulun, water flooded the road at a distance of more than 5 versts. Near the village, the water level rose by more than the height of the telegraph pole. Bridges, houses, a mill, a wine warehouse, outbuildings and other buildings were destroyed and demolished. During the flood, “up to 170 pieces of various small and large cattle” died.

**Figure 1.** Functions of a three-parameter power gamma of the probability distribution of the maximum annual water discharge of rain floods with biased (2) and bias (3) momentary statistical estimates.
Thus, if we assume that in 1870 a flood similar to 2019 was observed, then the empirical value of the probability of occurrence of a hydrological phenomenon will be \( p_r = 0.00662 \).

Let us return to the second task - forecasting rare hydrological events. To solve it, a factor model is proposed that characterizes the relationship between the maximum water discharge or level \( y_i \) and the previous values of daily precipitation and air temperatures according to the observation points located near and in the site of the investigated river. The general view of the model can be written as follows:

\[
y_i = a_0 + a_1 t_{k-1}^{op} + a_2 x_{k-1}^{op} + \ldots + a_{k-1} t_{k-1}^{op} + a_k x_{k-1}^{op},
\]

(3)

\[
y_i = a_0 + a_1 t_k^{op} + a_2 x_k^{op} + \ldots + a_{k-1} t_k^{op} + a_k x_k^{op} + a_{k+1} t_{k+1}^{op} + a_{k+2} x_{k+1}^{op} + \ldots + a_{2k} t_{2k}^{op} + a_{2k+1} x_{2k+1}^{op},
\]

(4)

where \( t_{k-1} \), \( x_{k-1} \), \( t_k \), \( x_k \), are the average daily air temperatures and the sum of daily precipitation for the previous period \( t_k \) and \( t_k \), corresponding to the lead time. Linear and nonlinear expressions (3) and (4) allow prediction based on a plurality of observation points \( (op_1, op_2) \). In particular, \( t_{k-1} \), \( x_{k-1} \) are the average daily air temperatures and the sum of daily precipitation for the previous period \( t_k \) and \( t_{k+1} \) for another observation point.

The given regression equations were used in [9] to predict rare hydrological phenomena observed on the rivers of different basins of the Irkutsk region. The developed models for the catchments of the left tributaries of the Angara allow us to obtain satisfactory forecasts with different lead times, which is associated with the peculiarities of the rivers and territories of their feeding. Regression equations, as a rule, are of a linear type. They are characterized by the significance of the expressions and coefficients of the variables, as well as adequacy. Moreover, the anticipation of the manifestation of the maximum water discharge is from 1 to 4 days. The simulation results show that the prognostic values of the hydrological parameter are lower than the actual values.

To predict the maximum water level of rain flood \( (H_i) \) on the river Iya, which took place in June 2019, obtained a factor model of the form

\[
H_i = 2034.83 - 57.22 t_{i-4}^{Tulun} + 1.41 x_{i-4}^{Tulun}.
\]

(5)

According to (5), the expression takes into account the data of the observation point at the site of the flood. The regression equation (5) is significant in accordance with the Fisher F-criterion. Moreover, the coefficients for unknowns are also significant according to Students t-test.

In addition to data on Tulun, you can use information from the meteorological parameters of other observation points. As a result of the studies, a dependence was constructed taking into account the influence of previous average daily air temperatures and the amounts of daily precipitation in Tulun and Nizhneudinsk on the peak water levels of the catastrophic flood on the river Iya at Tulun at the end of June 2019:

\[
H_i = 1679.57 - 31.66 t_{i-4}^{Tulun} + 2.26 x_{i-4}^{Tulun} - 7.18 y_{i-4}^{Nizhneudinsk} + 4.1 y_{i-4}^{Nizhneudinsk}.
\]

(6)

Like the previous expression, the regression equation (6) is significant, as are the coefficients for unknowns. The calculations showed that the discrepancy between the analytical and actual peak water levels for expression (5) was 1%, and for dependence (6) it corresponded to 2%. The regression equation allows predicting the values of the highest maximum water level with a lead time of 4 days.

**Table 2.** Estimation of the likelihood of a catastrophic maximum water discharge on the river Iya at Tulun in June 2019.

| Q, m³/s | c₁ | c₂/c₁ | p | Repeatability, years | a = 0.05 | a = 0.95 |
|--------|----|-------|---|----------------------|--------|--------|
| 6330   | 0.74 | 6.0   | 0.00218 | 459 | 0.00057 | 0.0327 |

| Q, m³/s | c₁ | c₂/c₁ | p | Repeatability, years | a = 0.05 | a = 0.95 |
|--------|----|-------|---|----------------------|--------|--------|
| 6330   | 0.76 | 6.0   | 0.00236 | 424 | 0.00057 | 0.0327 |
For comparison, we present the prognostic expression obtained on the basis of data on the second most significant rain flood on the river Iya, formed in 1984:

\[ Q = 3186 - 121.011_{Tulun}^{i} + 226.6x_{Tulun}^{i} \]  

(7)

According to the result, the lead time of the forecast was two days for air temperature and nine for daily precipitation. In this case, the discrepancy between the actual and forecasted value of the maximum discharge was 11%.

It should be noted the advantages and disadvantages of the proposed algorithm for the probabilistic assessment of a rare phenomenon and prognostic value.

In the first problem, for a more accurate estimate of the maximum water discharge, truncated probability distribution laws [16] or the synthesis of different functions describing the corresponding distribution sections [17, 18] can be used. In this case, one should bear in mind the different accuracy of the methods for estimating the statistical parameters of series [17, 19, 20], as well as the features of the description of event flows [21].

In this work, an attempt was made to use a truncated distribution, but the results were worse than the given probabilistic estimates based on a three-parameter power gamma distribution.

As for the factor prognostic dependences, for their use a network of observations is required that characterizes air flows of different directions. This is possible thanks to the use of special sensors. They can significantly supplement the information of the hydrometeorological network and partially solve the problem of estimating the parameters of heat and humidification for various directions of movement of air masses. On the other hand, the use of different observation points around the considered hydrological alignment of the river allows us to identify the corresponding scenario.

The proposed factorial prognostic model underestimates the value of the maximum water discharge relative to the actual value. Therefore, it is necessary to introduce a correction factor for increasing the forecast value. For the territory under consideration, the empirical coefficient fluctuates at the level of 8-11%.

Obtaining a satisfactory forecast of the maximum water discharge of rain floods based on the proposed models with a given lead time allows timely warning of the population about the danger of flooding and reducing risks.

4. Conclusions
An algorithm for predicting rare hydrometeorological events and its probabilistic assessment using the example of rare rain floods in the middle rivers of the Angarsk and Lena basins is proposed.

The developed algorithm is implemented to determine the catastrophic maximum water discharge of rain flood observed on the river Iya at Tulun at the end of June 2019.

The advantages and disadvantages of the forecasting algorithm and the probabilistic assessment of the maximum water discharge of rare rain flood are given.

Acknowledgment
Some results were obtained with funded by RFBR according to the research project no. 19-07-00322

References
[1] Hayhoe K et al. 2007 Past and future changes in climate and hydrological indicators in the US Northeast Climate Dynamics 28 381-407
[2] Kundzewicz Z W et al. 2014 Flood risk and climate change: global and regional perspectives Hydrological Sciences Journal 59(1) 1-28
[3] Shunsuke T et al. 2020 An extreme flood caused by a heavy snowfall over the Indigirka River basin in Northeastern Siberia Hydrological Processes 34(3) 522-37
[4] Van Lanen J H A et al. 2016 Hydrology needed to manage droughts: the 2015 European case Hydrological Processes 30(17) 3097-104
[5] Arefieva E V and Bolgov M V 2018 Features of forecasting natural floods in order to reduce the
risk of emergencies on the example of the Krasnodar Territory Civil Security Technologies 15 4(58) 40-7

[6] Befani N F 1977 Forecasting Rain Floods Based on Territories Common Dependencies (Leningrad: Gidrometeoizdat) p 182

[7] Bolgov M V and Korobkina E A 2013 Reconstruction of the rain flood on the Adagum River based on mathematical models of runoff formation Russian Water Management: Problems Technologies Management 3 87-102

[8] Zhou Y-L and Guo S-L 2014 Risk analysis for flood control operation of seasonal flood-limited water level incorporating inflow forecasting error Hydrological Sciences Journal 59(5) 1006-19

[9] Ivanyo Ya M and Petrova S A 2015 Optimization Models of Agricultural Production in Solving the Problems of Assessing Natural and Technological Risks (Irkutsk: Publishing house Irkutsk SAU) p 180

[10] Ivanyo Ya M 1997 Extreme Natural Phenomena of the Historical Past in the Territory of the Irkutsk Region (Irkutsk: Publishing house Irkutsk university) p 96

[11] Kritsky S N and Menkel M F 1981 Hydrological Foundations of River Flow Management (Moscow: Nauka) p 270

[12] SNiP 1985 “Determination of calculated hydrological characteristics” 2.01.14-83 (Moscow: Stroyizdat) p 36

[13] SP “Determination of the main calculated hydrological characteristics” 33-101-2003

[14] Rozhdestvenskiy A V 1977 Evaluation of the Accuracy of the Distribution Curves of Hydrological Characteristics (Leningrad: Gidrometeoizdat) p 270

[15] Floods in the Irkutsk Province that Occurred During 1870 (Irkutsk Provincial Gazette 29 July 1872) p 3-4

[16] Blokhinov E G 1974 The Probability Distribution of the Values of River Flow (Moscow: Publishing house of Science) p 169

[17] Van Gelder P 2004 Statistical estimation methods in hydrological engineering Analysis and stochastic Modeling of extreme Runoff in Euroasian Rivers under Conditions of Climate Chang, Proc. Int. Scientific Seminar (Irkutsk: Publishing House in the Institute of the Geography SB RAS) p 11-57

[18] Bolgov M V et al. 2016 The method of composition of distributions to estimate the maximum water discharge of the Amur River taking into account the effect of flow regulation Water Resources 43(3) 265-76

[19] Grinevich G A et al. 1972 Compositional Modeling of Hydrographs (Moscow: Nauka) p 181

[20] Bezak N et al. 2014 Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis Hydrological Sciences Journal 59(5) 959-77

[21] Belyakova A Yu and Ivanyo Ya M 2009 Probabilistic Models of Extreme Hydrological Phenomena in the Problems of Agricultural Production Optimization (Irkutsk: Publishing house Irkutsk SAU) p 151