



Good-quality Long-term Forecast of Spring-summer Flood Runoff for Mountain Rivers

Yuri B. Kirsta¹ · Olga V. Lovtskaya¹

Received: 17 June 2020 / Accepted: 10 December 2020/

Published online: 26 January 2021

© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract

The universal simulation model was developed with the use of system-analytical modeling to ensure a long-term forecast of mountain river runoff during spring-summer floods. Prediction quality of this SAM-model is characterized by Nash-Sutcliffe efficiency of 0.68–0.88 and is very high for long-term flood forecasts, including ones for inundations and mountain reservoirs filling in spring. The model was tested on the example of 34 medium and small rivers (1630 values of runoff observations for 1951–2016) located in the Altai-Sayan mountain country (2,000,000 km²). Its input factors include monthly precipitation, monthly mean air temperature, GIS data on landscape structure and orography of river basins. Meteorological factors are calculated as percentage of their “in situ” long-term mean values averaged for the whole study area. This helps to explain and quantify the influence of autumn-winter-spring soaking, freezing and thawing of mountain landscape soils on spring-summer flood. We apply a simple novel method to evaluate model sensitivity to variations in environmental factors expressed in terms of their contribution to variance of the observed flood runoff. It turns out that sensitivity of the latter decreases in the following sequence of factors: autumn precipitation, landscape structure of river basins, winter precipitation, winter air temperature, landscape altitude. The developed SAM-model provides a three-month lead-time estimate of runoff in a high water period with the threefold less variance as compared to forecasts based on the observed long-term mean values.

Keywords System-analytical modeling · Flood · Forecast · Mountains · Altai · Sayan

1 Introduction

Solving the problems of efficient use of river runoff under climate warming and increased number of extreme weather events requires the development of adequate models of flood

✉ Yuri B. Kirsta
kirsta@iwep.ru

¹ Institute for Water and Environmental Problems SB RAS, Barnaul, Russian Federation

forecasting. Such models enable to assess flood risks (including catastrophic ones), take measures in advance to prevent adverse consequences of inundation, and provide sustainable water supply to population (Mohammad-Azari et al. 2020; Parisouj et al. 2020). The models should take into account changes in meteorological, geomorphological and other factors over time and space (Conrad 2019; Brinkerhoff et al. 2020). As soon as the most important factors are identified, a quantitative forecast of river hydrological regime is made (Ahmadi et al. 2019; Feng et al. 2020). This task becomes much more complicated for mountain territories due to their complex orography, varied spatial-temporal structure of meteorological fields, diverse landscapes, and a sparse network of weather stations (Roessler et al. 2014). We have found the solution and developed the universal simulation model for a long-term forecasting of river runoff during spring-summer floods. Such a prediction is extremely important for flood control (Kundzewicz et al. 2019), planning of hydropower reservoir filling (Zhang et al. 2019), and management of water resources in the mountain areas (Tullos et al. 2016).

Most modern flood forecasting models are designed for a specific river and provide a short, medium- or long-term forecast with a lead-time from a few days up to a season. They use statistical methods for processing meteorological and hydrological data, differential equations of hydrological processes, artificial neural networks, genetic programming, GIS technologies, or different techniques taken together (Corripio and López-Moreno 2017; Mosavi et al. 2018; Musselman et al. 2018; Tabari 2019; Wu et al. 2020). GIS is used in accounting for spatial data distribution, visualization of calculation results, modeling of individual processes, etc. Statistical methods alone do not provide the accuracy required in water management; that is why differential equations of hydrodynamics and mathematical physics supplemented by simulation equations are applied to describe hydrological processes in river basins (Wijayarathne and Coulibaly 2020). Differential equations require the detailed spatially distributed information on precipitation, air temperature, evaporation, underground aquifers, morphometry, slopes, properties of water-saturated soils and fractured rocks, indicators of flow resistance on slopes and in river channels, etc. Because of the lack of such an information for mountain areas, the application of differential equations prevents an increase in the lead-time and good quality of forecasts and, therefore, becomes impractical. Since it is hardly feasible to account for impacts of soaking, freezing and thawing of mountain soils on flood, the accuracy of all predictive models leaves much to be desired.

Our method of system-analytical modeling (**SAM**) (Kirsta and Puzanov 2020) allows to avoid the problems mentioned above and to create the simulation balance SAM-models not based on differential equations of mathematical physics but distinguished by adequate description of real hydrological and physical processes occurred in the mountains. Following Beven (2002), adequacy is a conformity of the model to physical principles and laws complemented by appropriate assumptions. In this paper, the established dependences of flood runoff on environmental factors characterize dozens of different catchments and landscapes that makes them universal.

2 System-analytical Modeling

SAM enables to identify and quantitatively describe major processes and their relations with environmental factors in natural systems. SAM is used to search for and quantify relationships of flood runoff with meteorological factors, morphometry, and landscape structure of river basins. As the Tardy expert-analytical method (Tardy et al. 2004), SAM provides the selection of information on these relationships from experimental data series.

Main SAM steps are as follows (Kirsta and Puzanov 2020; Kirsta 2020):

1. Preparation of temporally and spatially homogeneous consistent data samples for a simulated process (flood runoff) and influencing factors.
2. Selection of averaging periods for primary data and model steps with regard for features and inertia of the process.
3. Estimation of admissible number of model parameters, which should be an order of magnitude less than the amount of process observation data.
4. Identification and verification of different variants of simulation balance equations (not differential, but with piecewise-linear dependencies on factors) by solving the inverse mathematical problem based on observation data. Choosing the model equations with the least quadratic discrepancy.
5. Assessment of adequacy of the model and its sensitivity to input factor variations.
6. Estimation of factors' contributions to model residual variance.

The simulation balance SAM-model for long-term forecasting spring-summer floods is constructed as a system of algebraic equations. Functional relationships between river runoff and environmental factors are found through hydrology-based selection and adjustment of equations to provide minimum model discrepancy (the sum of squared residuals) between calculated and observed values of runoff. Equation parameters are identified by solving the inverse mathematical problem using optimization methods of the MATLAB software package via substituting runoff observation data into equations. At the same time, the tested version of the model is evaluated for the resulting quadratic discrepancy. The SAM-model is considered to be developed when an equation system provides the least discrepancy.

In SAM, the form of sought dependencies on environmental factors is not fixed in advance, unlike differential equations when this form is fixed by their choice. To support this feature, we apply the universal match function H defined as (Kirsta and Puzanov 2020):

$$H(X1, X2, Y1, Y2, Z1, Z2, X) = \begin{cases} Y1 + Z1(X - X1), & \text{if } X < X1 \\ \frac{Y2 - Y1}{X2 - X1}(X - X1) + Y1, & \text{if } \begin{cases} X1 \leq X < X2 \\ X1 \neq X2 \end{cases} \\ Y2 + Z2(X - X2), & \text{if } X \geq X2 \end{cases}, \quad (1)$$

where $X1, X2, Y1, Y2, Z1, Z2$ are parameters; X is any model variable. Function H is a continuous piecewise-linear function consisting of three arbitrary linear fragments. It is used to approximate different dependencies between variables by changing parameters' values.

Any mathematical models (including predictive ones) need verification. To do that, we propose a criterion for assessing adequacy of any models or calculation methods through the comparison of observed and calculated data series (Kirsta 2011; Kirsta and Puzanov 2020):

$$A = \frac{1}{\sqrt{2}} \frac{S_{\text{dif}}}{S_{\text{obs}}}, \quad (2)$$

where A is the criterion of model adequacy; S_{dif} is the standard (RMS) deviation of the difference between calculated and observed data patterns (standard deviation of model residuals); S_{obs} is the standard deviation of the observed pattern; $1/\sqrt{2}$ is the introduced multiplier. According to (2), A is actually a model error normalized to the standard deviation of observation data.

Criterion A in (2) is similar to commonly accepted indicators of model quality RSR (RMSE-observation Standard deviation Ratio, where RMSE is Root Mean Square Error

(Moriassi et al. 2007; Koch and Cherie 2013)) and NSE (Nash-Sutcliffe model Efficiency (Nash and Sutcliffe 1970; Koch and Cherie 2013)). A is related to RSR and NSE through dependencies $RSR=A\sqrt{2}$ and $NSE=1-RSR^2=1-2A^2$. As compared to RSR and NSE, the range of criterion A applications is wider and additionally involves assessment of statistical forecasting when $NSE < 0$.

Mathematical models have another important characteristic, i.e. sensitivity to variations in environmental factors (Iooss and Lemaître 2015; Song et al. 2015; Wang et al. 2020). We have proposed a simple method for quantifying the model sensitivity due to the use of adequacy criterion A (Kirsta 2020):

$$FS = (A')^2 - (A)^2 = \frac{(S'_{dif})^2 - (S_{dif})^2}{2(S_{obs})^2} = \frac{2(S_{fac})^2}{2(S_{obs})^2} = \frac{(S_{fac})^2}{(S_{obs})^2}, \quad (3)$$

where FS is the model sensitivity to natural variations of the selected input factor; A is the criterion (2); A' is the value A obtained from (2) by using randomly mixed values of input factor instead of initially ordered. In this case, the randomly mixed observed data pattern, obviously, has a former statistical distribution and variance. Here, $(S_{dif})^2$ is the variance for the difference between calculated and observed data patterns of the model output variable (river runoff); $(S'_{dif})^2$ is the same variance obtained by substituting randomly mixed values of the selected input factor in the model; $(S_{fac})^2$ is the contribution of natural variations of input factor to variance of the output variable (river runoff); $(S_{obs})^2$ is the variance of the observed output variable used for FS normalization.

In Eq. (3), the variance determined by errors in input factor observations is present in both $(S'_{dif})^2$ and $(S_{dif})^2$. Therefore, it does not affect FS because of its mutual subtraction in the numerator of Eq. (3). Hence, FS evaluates the model sensitivity *directly to natural variations of input factor*, excluding any errors in its observations. Obviously, FS also indicates the relative importance of environmental factors. Similar to A , FS can be expressed as a function of RSR or NSE. Taking into account the equality of $RSR=A\sqrt{2}$, we have $FS = [(RSR')^2 - (RSR)^2]/2$. Similarly, we get $FS = (NSE - NSE')/2$. As A' in (3), RSR' and NSE' are equal to RSR and NSE obtained via using randomly mixed values of the selected input factor instead of initially ordered ones.

3 Source Materials

The study territory of the Altai-Sayan mountain country (50–56° N and 82–90° E) represents a part of the world watershed located between the humid zone of the Arctic Ocean and the arid drainless area of Central Asia. It covers more than 2,000,000 km² and serves as the catchment for Siberian large rivers Ob, Irtysh and Yenisei distinguished by a complex hydrographic network. Generally, mountain ridges are of 2000–2500 m a.s.l., while in Altai they reach 3500–4500 m a.s.l. The river flow regime depends mainly on snow melting in springtime and precipitation amount during summer and autumn. The share of snow melting in the annual flow exceeds 50%.

The climate of the Altai-Sayan mountain country is sharply continental. According to meteorological observations, long-term average temperature is −16 °C in January and +18 °C in July. Long-term average precipitation varies from 20 mm in February to 84 mm in July. More detailed climate characteristics can be found in (Kirsta 2011; Kokorin 2011). Climatic diversity of the country contributes to that of its landscapes: glacial-nival, tundra,

alpine and subalpine meadows, forest, steppe and semi-desert ones (Gvozdetskiy 1968; KIRSTA and Puzanov 2020). For instance, some northern slopes at altitudes over 3000 m get 1200–2500 mm of precipitation per year, the middle parts of the slopes – up to 600 mm, and the foot – about 200 mm (Gvozdetskiy and Mikhailov 1987).

SAM was based on river runoff observations in the Altai-Sayan mountain country made by the Hydrometeorological service of the USSR and Russia in 1951–2016. For the study, 34 medium and small rivers with defined basin boundaries (i.e. Anuy, Berd, Charysh, Inya, Katun, Tom, et al.) were selected (KIRSTA and Puzanov 2020). A large number of simultaneously analyzed basins (from 177 to 21,000 km² in size) allowed leveling their specific features and determining general (universal) patterns of hydrological processes for the entire mountain territory.

Taking into consideration annual dynamics of streamflow, we specified four hydrological periods/seasons: winter low water (XII–III months), spring-summer flood (IV–VI), summer low water (VII–VIII) and autumn low water with possible flood in case of heavy rains (IX–XI). The data on daily runoff observations at each of 34 river gauges were averaged over spring-summer flood period for each year of observations. During flooding, runoff was about 140 m³/s with standard deviation of streamflow time series of 24% on average for all basins. Note that this runoff was much higher compared to other periods (10, 50, 30 m³/s for the first, third and fourth seasons, respectively).

We made typification of landscapes of the Altai-Sayan mountain country reflecting the conditions of hydrological and hydrochemical runoff formation including altitudinal-belt and structural-layering heterogeneity of the territory (KIRSTA and Puzanov 2020). To account for the landscape structure of river basins and for spatial separation of different types of a hydrological regime, we used maps “Landscapes of Altai” (Chernykh and SamoiloVA 2011) and “Landscape map of Altai Krai” (Tsimbaley 2011) of 1:500,000 scale. A total of 12 landscapes and one extra (13th) small aquatic landscape were selected. Each landscape was considered as an independent typological group of geosystems with its own hydrological characteristics and a certain contribution to the overall flow from the river basin.

The elevation and areas of the selected landscapes were calculated for all 34 river basins. The mentioned above maps as well as topographic ones of 1:200,000 scale served as a basic cartographic material processed in ArcGIS 10.2 to create digital versions of maps. We calculated the morphometry of landscape structure of each river basin (namely, the area and mean altitude of landscapes) by means of TIN models constructed by ArcGIS 3D Analyst.

For most analyzed river basins, regular meteorological observations are absent. We have shown that long-term dynamics of monthly mean air temperature and monthly precipitation is the same throughout the Altai-Sayan mountain country, if these factors are expressed in percent/fraction relative to their long-term in situ average for corresponding month (KIRSTA 2011). For calculations of dynamics of such relative/normalized factors, we used the data from 11 reference weather stations for 1951–2016 years. Four stations (with WMO indexes 29,822, 29,915, 29,923, 29,939) were located in the plains adjacent to the Mountain Altai, six (36,038, 36,045, 36,055, 36,064, 36,229, 36,442) – in the Mountain Altai, and one (29,849) – in the Kuznetsk intermountain basin at the altitudes of 125–2601 m a.s.l.

First, we calculated monthly and interannual dynamics of normalized values of air monthly mean temperature and monthly precipitation for each weather station separately. For the cold period of the year, we expressed the normalized temperature sets of 11 stations as percentage of long-term average January in situ temperature, and for the warm period – of July one. July provided the normalized precipitation sets for all months of the year. Then, using the average for 11 stations, we

obtained spatially averaged monthly dynamics of normalized meteorological factors. This dynamics adequately characterized the real meteorological situation in any part of the Altai-Sayan mountain country, regardless of coordinates and elevation. The absence of dependence on changing elevation was particularly important for the description of dynamics of normalized monthly mean temperature and monthly precipitation over the entire area of each basin, where elevation difference reached 2000 m. Moreover, this dynamic was uniform for all river basins. Thus, $66 \times 12 = 792$ normalized values were calculated both for temperature and precipitation for the years 1951–2016. To present monthly and interannual dynamics of temperature and precipitation in degrees Celsius and millimeters for any site, it is sufficient to have their long-term in situ average for January and July.

The adequacy of normalized meteorological characteristics was assessed by criterion (2). Via comparison of their calculated (for reference stations) and observed (for other non-reference stations) multi-year monthly series with criterion A , we got $A_T=0.39$ for air temperature and $A_P=0.62$ for precipitation for all months of the year on average (Kirsta 2011). It is expected to use values of A in evaluating the SAM-model adequacy.

Since normalized air temperature and precipitation were among input factors of the SAM-model, streamflow values should be normalized as well. For this purpose, the observed values were divided by their long-term average for each basin. Thus, we passed from streamflow measurements in m^3/s to dimensionless units of normalized runoff. Such a normalization made it possible to create suitable for SAM single homogeneous sample of runoff data on all 34 river basins. In a similar manner, landscape areas (km^2) in each basin were converted into percent/fraction via division by the basin area. Elevation of landscapes remained unchanged.

Overall, the created database to execute SAM of flood runoff includes the following characteristics for 34 river basins: hydrological (1630 normalized values of river runoff for spring-summer flood period), meteorological ($792 + 792 = 1584$ normalized monthly values of air temperature and precipitation) and landscape ($160 + 160 = 320$ values of area and altitude) ones.

4 Development of the SAM-model for Flood Forecasting

Just as in regression analysis, to use SAM successfully, the array of actual data on the process dynamics should greatly exceed the number of SAM-model equation parameters (Kirsta and Puzanov 2020). There are 1630 values of river runoff. Theoretically, it allows us to enter about 160 parameters into the model. However, the number of river basins comes to 34 resulting in 3–4 parameters allowed, if we describe the runoff dependence on any of basin characteristics (e.g. basin area or elevation of landscapes). The same applies to the dependence of runoff on temperature and precipitation if we use, for example, a 33-year observation period to identify the model parameters. The limitation to 3–4 parameters implies the presence of two linear fragments in piecewise-linear function H in (1). This requirement is implemented according to (1) as dependence $H(X_1, X_2, 1, 1, Z_1, Z_2, X)$.

The forecast of average river runoff during spring-summer flood (IV–VI months) was made for each year (1984–2016) using the data on air temperature and precipitation for preceding IX–III months. The parameters of the developed predictive SAM-model were identified by solving the inverse mathematical problem due to optimization methods. For their calculation, we simultaneously used actual normalized flood flows of all 34 rivers over a 33-year moving identification period preceding the forecast years. The defined parameters were applied to forecast runoff only in the next three years. For subsequent years, their values were calculated again using further identification period with a three-year shift.

During SAM, various sets of equations describing runoff formation influenced by environmental factors during high water periods were tested. As a result, the SAM-model for flood forecasting with 39 parameters and the least quadratic discrepancy between predicted and actual flood flows was constructed. Its general/universal (for all river basins) simulation balance equation has the form:

$$Q^i = H(c_1, c_2, 1, 1, c_3, c_4, P_1) \left\{ \sum_k a_k S_k^i P_1 H(c_9, c_{10}, 1, 1, c_{11}, c_{12}, h_k^i) + \sum_k b_k S_k^i P_2 H(c_5, c_6, 1, 1, c_7, c_8, T_2) H(c_9, c_{10}, 1, 1, c_{11}, c_{12}, h_k^i) \right\} + d \tag{4}$$

where Q^i – the predicted average normalized runoff during spring-summer flood period (IV–VI months) for outlet of basin i , $i = 1–34$; the first and second sums in the right part of (4) correspond to contributions of recent autumn period (IX–XI) and current winter one (XII–III), respectively; a_k, b_k – the parameters characterizing k -th landscape contributions to river runoff in relevant period, $k = 1–13$; S_k^i – the relative area of k -th landscape in the basin i ; h_k^i – the landscape elevation, m a.s.l.; P_1, P_2 – the mean deviations of normalized monthly precipitation from its long-term average monthly values in recent autumn and current winter periods, respectively; T_2 – the analogous mean deviation but for normalized monthly air temperature in current winter period; H – the piecewise-linear function (1); $c_{1–4}, c_{5–8}, c_{9–12}$ – the parameters describing the influence of autumn precipitation P_1 and winter temperature T_2 on runoff volume during a flood period as well as landscape elevation h_k^i on precipitation amount; d – the constant fraction of normalized runoff ($d \leq 1$) equal for all river basins.

The right part of Eq. (4) summarizes the contributions of each landscape to Q^i of river basin i . In the first summand of (4), the contribution of k -th landscape is formed due to autumn precipitation P_1 in the preceding year and depends on hydrological features of this landscape (parameter a_k), its area S_k^i , and landscape elevation h_k^i (which affects precipitation amount). In the second summand of (4), the contribution of k -th landscape is formed by winter precipitation P_2 and depends on landscape features (parameter b_k), area S_k^i , temperature T_2 (responsible for evaporation from the snow surface), and elevation h_k^i . Multiplier $H(c_1, c_2, 1, 1, c_3, c_4, P_1)$ accounts for moisture exchange between soils and snow cover in winter, which depends on autumn precipitation P_1 (Tanasienko et al. 1999).

To identify parameters through solving the inverse problem, we used the runoff observation data for 1951–2013. We replaced the left part of Eq. (4) with normalized values of runoff observed in the proper 33-year moving identification period. These values (200–1000) made up a system of the same number of equations solved by means of MATLAB optimization methods to calculate parameters as unknown variables. Through solving this system of equations, the quadratic mean of identification residuals was also determined; it characterized quadratic discrepancy between calculated and observed river flows in the 33-year identification periods. River runoff data, which gave significant residuals, were excluded from calculations. Such a random disagreement may occur in a particular year due to (a) considerable difference between value of a meteorological factor averaged for 11 reference weather stations and its actual fluctuations in some river basins, (b) technical errors in data recording. Significant residuals were detected by comparison of calculated and observed runoff according to the known “three-sigma rule” for normal statistical distribution. For SAM, deviation of calculated from experimental/observed values was close to normal as in adequate mathematical models. Any runoff observation beyond three residual standard deviations from the calculated runoff was considered to be an error. Up to 4% of the data were excluded in this way. Obviously, these data could be reliable and correspond to real fluctuations of precipitation or

air temperature in the basin. In fact, few errors exclusion did not affect SAM. From the above it follows that reliability of performed forecasts is $100 - 4 = 96\%$ (commonly accepted reliability of statistical estimates makes up 95%).

The quality of the SAM-model (4) for forecasting spring-summer flood runoff was verified for 1984–2016 using adequacy criterion A in (2). This period was divided into short three-year intervals; for each we made individual forecasts using (4) with parameters found for the preceding 33-year moving identification period. When calculating criterion A , standard deviation S_{diff} characterizing discrepancy between predicted and actual normalized river runoff was used. For 1984–2016, 381 normalized values of observed runoff were available for all 34 river basins. It is obvious that such samples of predicted and actual runoff are sufficient for reliable statistical estimates. According to the calculations, adequacy A of model (4) is equal to 0.65. Note that the same value of A was obtained for winter runoff forecasting; it was based on meteorological data of preceding summer and autumn seasons (Kirsta and Puzanov 2020). The resulting value $A = 0.65$ is less than the threshold value of 0.71 (consistent with $NSE = 1 - 2A^2 = 0$). The latter corresponds to a simple predictive model based on the use of statistical average of a characteristic. Thus, preceding meteorological conditions considered in (4) really affect the volume of spring-summer flood runoff in river basins of the Altai-Sayan mountain country. It should be noted that we used the data from weather stations located outside the river basins. The possibility of more accurate flood forecast based on in situ meteorological observations is discussed below.

5 Discussion of Simulation Results

Most predictive models of river floods are developed due to the data obtained for a single river basin. In such cases, to characterize the impact of runoff from each basin landscape as a separate natural hydrological system is extremely difficult. One value of elevation (or any other characteristic) is assigned to each landscape. To construct the function of landscape runoff dependence on this characteristic using only one plot point is impossible. In addition, one can not exclude individual characteristics of the analyzed basin from the model, which limit its application in another territory. In our case, parameters of model (4) characterize 34 basins concurrently analyzed in SAM; they essentially differ in orography, land cover and climate. In other words, the developed SAM-model is universal and can be used in forecasting river floods throughout the Altai-Sayan mountain country. It is worth noting that predicted runoff from each of 34 river basins is normalized to its long-term average value. Multiplying runoff by this value, we can convert it to m^3/s . Using Eq. (4), it is also easy to forecast runoff from individual landscapes in each basin.

5.1 Runoff Dependence on Meteorological Factors

Figure 1 shows a typical dependence of spring-summer flood runoff Q on input meteorological factors of model (4) obtained by SAM. Proportionality of runoff to amount of winter precipitation P_2 as well as rather complex dependence of runoff on autumn precipitation P_1 and winter temperature T_2 are marked.

Let us consider runoff Q behavior at small amount of winter precipitation $P_2 < 0$. A drastic decrease in autumn precipitation $P_1 < -0.3$ (Fig. 1a) or low winter temperature $T_2 < 0$ (Fig. 1b)

bring to an increase in runoff Q . This is due to growing depth of soil freezing (up to 2–3 meters (Tanasienko et al. 1999)) in winter because of a thin snow cover ($P_2 < 0$). When snow melts in spring, deep freezing facilitates an ice layer formation in the upper soil layers and prevents meltwater infiltration into soils (Tanasienko et al. 1999). As a result, water flows downslope into rivers, thus intensifying flooding. With autumn precipitation increase $P_1 > -0.3$ (Fig. 1a), the ice layer is formed already in the transitional autumn–winter period and occupies more and more soil areas. In winter, it enlarges due to upward diffusion of water vapor from deep soil layers. In spring, it again prevents meltwater infiltration and increases Q . In turn, with winter temperature rise $T_2 > 0$ (Fig. 1b), the increase in runoff Q can be explained by two reasons: (a) more intensive melting of snow cover at all altitudes due to slightly frozen soils, (b) greater moistening of soils because of their less drying in winter caused by a reduced temperature gradient between them and snow cover/air.

With a large amount of winter precipitation ($P_2 > 0$), decrease in winter temperature $T_2 < 0$ brings to decrease in Q (Fig. 1b). This is because of temperature fall in snow cover and soil resulting in delayed snowmelt in spring and partial flood transfer to summer. With temperature rise $T_2 > 0$, decrease in Q is explained by (a) slower freezing of soils that causes moisture transition to the lower soil layers (Nikolaev and Skachkov 2012) and then to winter runoff, (b) increase in evaporation from the snow surface.

5.2 Sensitivity of the SAM-model to Factors' Variations

Another important characteristic of model (4), namely, its sensitivity FS to variations in environmental factors, makes it possible to quantify the significance of these factors in forecasting spring-summer floods. Since FS in (3) is expressed as a fraction of variance of observed values of output variable (S_{obs})², FS can be expressed as percentage via multiplying it by 100. Flood sensitivity assessment with the use of series of forecasted river runoff Q^i in (4) and observed ones is given in Table 1.

According to Table 1, sensitivity of flood forecasting model (4) to variations in environmental factors goes down in the following order: both autumn and winter precipitation P_1 and P_2 , autumn precipitation P_1 , hydrological characteristics of landscapes a_k and b_k , winter precipitation P_2 , air temperature T_2 , and elevation of landscapes h_k^i . As expected, the SAM-

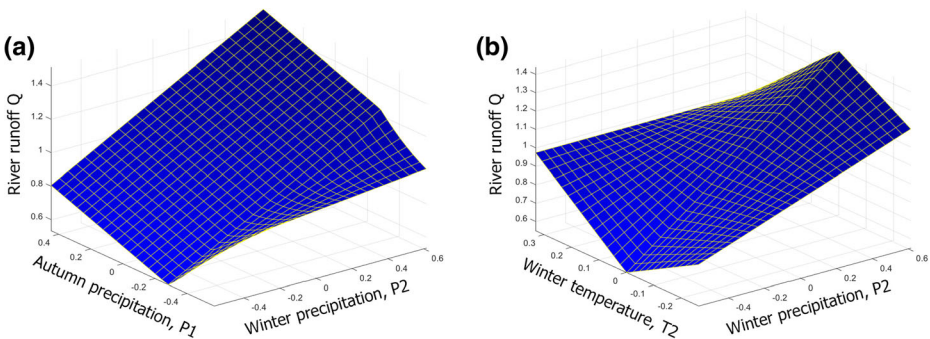


Fig. 1 Dependence of Katun river runoff during spring-summer flood on air temperature and precipitation for the preceding seasons (we show the deviations of normalized meteorological characteristics from their long-term averages – see notation for Eq. (4)); **a** – runoff Q (IV–VI months) as a function of precipitation P_1 (IX–XI) and P_2 (XII–III); **b** – runoff Q as a function of temperature T_2 (XII–III) and precipitation P_2

model is most sensitive to two factors, i.e. precipitation and hydrological features of landscapes. It shows the least sensitivity to landscape elevation, regardless of significant dependence of precipitation on elevation (Gvozdetkiy and Mikhailov 1987). This independently confirms adequacy of the description of meteorological fields in the Altai-Sayan mountain country using normalized values of monthly precipitation and temperature.

A special emphasis should be put on the small value of flood sensitivity $FS_T = 1.1\%$ to variations in winter temperature T_2 (Table 1). The value of T_2 is calculated as deviation of normalized monthly mean air temperature from its long-term value for four months (XII–III) on average. Therefore, interannual variations of T_2 are small and, consequently, their impact on flood runoff is insignificant too. In rare years, T_2 deviations can be great and have essential effect on runoff (Fig. 1b).

5.3 Assessing the Accuracy of Flood Prediction for a Single Basin

When spring-summer flood forecasts are made for a single river basin, its landscape structure, elevation and area of landscapes remain constant and, therefore, they can be excluded from the predictive SAM-model (4). Thus, Eq. (4) takes a simplified form:

$$Q = H(c_1, c_2, 1, 1, c_3, c_4, P_1)\{aP_1 + bP_2H(c_5, c_6, 1, 1, c_7, c_8, T_2)\} + d, \quad (5)$$

where the same notation as in (4) but for a single river basin is used. The values of parameters c_{1-4} and c_{5-8} are also not changed. Three parameters a , b , and d in (5) should be SAM calculated anew from the data on flood observations in the analyzed basin for the preceding 33 years. Further, these parameters are used when making flood forecasts in the following three years. After that, they are calculated again for the next preceding 33-year identification period shifted forward by three years.

Assume that we deal with a single typical river basin and use in situ meteorological observations. For this reason, forecast adequacy (A_0) for the simplified SAM-model (5) should differ from that for model (4) (A in Table 1). Indeed, in (5) there is no influence of (a) landscape hydrological features of 34 different basins, which increase discrepancy between forecasted and observed runoff as well as (b) errors in values of precipitation and air

Table 1 Sensitivity of flood forecasting model (4) to variations in environmental factors

Characteristic	Value
Adequacy A of model (4) according to Eq. (2)	0.65
Standard deviation ¹ S_{obs} of actual river runoff Q	0.30
Sensitivity ² FS_P to combined variations of autumn and winter precipitation (P_1 and P_2)	34
Sensitivity FS_{P_1} to autumn precipitation P_1	21
Sensitivity FS_L to variations in landscape structure of river basins ³	16
Sensitivity FS_{P_2} to winter precipitation P_2	10
Sensitivity FS_T to winter air temperature T_2	1.1
Sensitivity FS_h to landscape elevation h_k^i	0.2

¹ calculated as mean standard deviation of normalized observed runoff in 34 river basins. At the same time, it corresponds to mean standard deviation in fractions (or as percentage when multiplied by 100%) of non-normalized observed runoff

² estimated by (3) and expressed in percent of variance (S_{obs}^2)

³ calculated via joint random mixing the values of landscape hydrological characteristics (a_k and b_k in (4)) among 34 basins, $k = 1-12$

temperature arising from their spatial averaging over the Altai-Sayan mountain country. Obviously, A_0 is more important than A , since a flood forecast is almost always needed for a specific river if there is a threat of extreme flooding, overflowing a hydropower reservoir, or breakdown in normal water supply of local population.

Sensitivity of the SAM-model (4) to variations in environmental factors (Table 1) makes it possible to assess adequacy A_0 for spring-summer flood forecasts based on in situ meteorological observations. Normalized monthly mean air temperature and normalized monthly precipitation are the changing input factors of model (4). They are the same for the entire territory of the Altai-Sayan mountain country because of their spatial averaging characterized by adequacy criterion (2) for each month of the year (Kirsta 2011). Let us consider what makes up discrepancy variance $(S_{dif})^2$ for river runoff predicted by means of (4).

In mathematical models of complex natural systems, variance of model residuals is composed of components determined by variability of input factors, observation errors and equation errors (Mirkin and Rozenberg 1978; Kirsta 2020). Given such summation and Eqs. (2), (3), we can distinguish the components of residual variance $(S_{dif})^2$ for the SAM-model (4), which are absent in residual variance for the simplified SAM-model (5). Excluding them from $(S_{dif})^2$ for (4), we find the remaining component that characterizes A_0 . Using normalization of (2) and (3) to S_{obs} and $(S_{obs})^2$, we get:

$$2A^2 = (S_{dif})^2 / (S_{obs})^2 \approx FS_L + FS_P \times 2A_P^2 + FS_T \times 2A_T^2 + 2A_0^2, \tag{6a}$$

where A – the adequacy (2) of model (4); $(S_{dif})^2$ – the variance of residuals between predicted and observed stream flows; $(S_{obs})^2$ – the variance of observed stream flows; FS_L – the contribution of variations of landscape hydrological characteristics a_k , and b_k (Table 1); FS_P , FS_T – the contributions of precipitation and air temperature variations, respectively (see (3) and Table 1); A_P , A_T – the adequacy (2) of relevant meteorological factors; $2A_P^2$, $2A_T^2$ – the shares in FS_P , FS_T , formed by errors of P and T spatial averaging; A_0 – the desired variance component that characterizes adequacy of simplified model (5) for a single river basin (see above). Substituting in (6a) the values in fractions of a unit $A = 0.65$, $FS_L = 0.16$, $FS_P = 0.34$, $FS_T = 0.011$ (Table 1) as well as $A_P = 0.73$ and $A_T = 0.32$ obtained by averaging their monthly values for autumn-winter and winter seasons (Kirsta 2011), we easily identify A_0 :

$$2 \times (0.65)^2 \approx 0.16 + 0.34 \times 2(0.73)^2 + 0.011 \times 2(0.32)^2 + 2A_0^2 \text{ or } A_0 \approx 0.40. \tag{6b}$$

Let us consider a typical river basin, for which model adequacy A_0 should remain the same as in Eqs. (6a), (6b). This is because samples of normalized observed runoff for a typical basin and for 34 basins have the same statistical characteristics S_{obs} and differ only in their volumes. Therefore, we can use $A_0 \approx 0.40$ from (6b) to evaluate the quality of model (5) for an individual basin. Calculating A_0 in Eqs. (6a), (6b), we excluded the terms characterizing relatively small influence of (a) landscape elevation h_k^i on normalized precipitation, (b) errors in determining h_k^i , S_k^i , and (c) errors in evaluating parameters c_{9-12} in model (4). This means that a real value of A_0 is less than obtained 0.40, and $A_0 < 0.40$ is final assessment of adequacy of the predictive SAM-model (5) based on meteorological observations in the basin.

Consider the extent to which the SAM-model (5) reduces residual variance $(S_{dif})^2$ of predicted flood runoff in comparison with variance $(S_{obs})^2$ for a trivial forecast based on the long-term average of observed runoff. Given Eq. (2) and $A_0 < 0.40$, we obtain threefold reduction in residual variance:

$$A_0^2 = (S_{\text{dif}})^2 / 2(S_{\text{obs}})^2 < (0.40)^2 \text{ or } (S_{\text{dif}})^2$$

It is also easy to calculate another characteristic of predictive model quality, i.e. the Nash-Sutcliffe efficiency NSE. The flood forecast by model (5) in a river basin with in situ meteorological observations is characterized by $NSE_0 = 1 - 2A_0^2$ (see Section 2). Substitution of $A_0 < 0.40$ gives $NSE_0 > 0.68$. Values $NSE_0 > 0.68$ correspond to a good ($0.65 < NSE \leq 0.75$) quality of “descriptive” hydrological models (Koch and Cherie 2013) and are incredibly high for models of long-term forecast of flood runoff. In our case, river runoff forecasting is made for the entire period of spring-summer high water (IV–VI months), i.e. for 3 months ahead.

The average share of current precipitation in spring-summer runoff from 34 river basins makes up 0.3 (Kirsta and Puzanov 2020) and proves the possibility of achievement of high accuracy in flood forecasting. With regard for the standard deviation of precipitation amount (28%) (Kirsta 2011), we get theoretically best accuracy of flood forecasting models: $2 \times 0.3 \times 28 \approx 17\%$. The latter characterizes the confidence interval of the forecast $\pm 2 \times \text{sigma}$ (i.e., $\pm 2 \times$ “forecast standard deviation”) with 95% reliability.

Using the standard deviation $S_{\text{obs}} \approx 24\%$ of the observed flood runoff, we can also estimate theoretically best value NSE_{best} for forecasting river flows during spring-summer floods:

$$NSE_{\text{best}} \approx 1 - (0.3 \times 28 / S_{\text{obs}})^2 \approx 1 - (0.3 \times 28 / 24)^2 \approx 0.88.$$

Thus, for the rivers of the Altai-Sayan mountain country, the values of NSE_0 and $RSR_0 = \sqrt{1 - NSE_0}$ are within 0.68–0.88 and 0.35–0.57, respectively.

6 Conclusion

The proposed system-analytical modeling (SAM) of complex natural systems is an efficient method for developing adequate simulation models of river runoff. SAM ensures quantitative assessments of their accuracy and sensitivity to variations in environmental factors as well as evaluation of factors contributions to the variance of discrepancy between calculated and observed values of output model variable. Based on these estimates, individual components of this variance can be identified and analyzed separately. Such an analysis provides more objective and complete description of the mathematical model quality as compared to index RSR (the ratio of RMSE to standard deviation of observation data) or Nash-Sutcliffe efficiency NSE.

The runoff from 34 river basins of the Altai-Sayan mountain country was analyzed using SAM. The universal predictive SAM-model (4)–(5) was developed to forecast spring-summer flood runoff for 3 months ahead. The forecasts are based on relative (normalized for long-term average values) monthly precipitation and monthly mean air temperature for recent autumn and winter periods. Relative values of these factors more adequately reflect their complex distribution over the mountain country and are employed to improve the accuracy of hydrological calculations. In the course of SAM, the influence of autumn-winter-spring freezing and thawing of soils on spring-summer flood was quantified due to the data on autumn-winter-spring precipitation, air temperature and observed flood runoff.

Sensitivity of the SAM-model (4)–(5) to environmental factors decreases in their order: autumn precipitation, landscape structure of river basins, winter precipitation, winter air temperature, landscape elevation. The predicted flood runoff is normalized similar to meteorological data and can be converted to m^3/s via multiplying by its long-term average value.

The simplified SAM-model (5) has a threefold reduction in variance of forecast error as compared to that of trivial prediction based on the long-term average of observed runoff. The quality of forecasts is characterized by the Nash-Sutcliffe efficiency of 0.68–0.88. These values are excellent for long-term forecast hydrological models. Such a quality can be achieved using precipitation and air temperature observed directly in the river basin under study. The versatility of the model makes it applicable to all rivers of the Altai-Sayan mountain country, and after additional identification (specifying the values of some parameters) – to rivers in other mountain territories. It can be successfully applied in real-life water resources management under changing climate conditions, e.g. regulation of mountain hydropower reservoir filling or advance preparation for extreme flooding in spring.

Acknowledgements The work was carried out within the framework of the Research Program of the Institute for Water and Environmental Problems SB RAS (Project 0383-2019-0005).

Author Contributions Y. B. Kirsta designed the study and conducted data analysis; O. V. Lovtskaya participated in the design of the study and GIS application. Both authors read and approved the final manuscript.

Data Availability Not applicable.

Compliance with Ethical Standards

Ethical Approval Not applicable.

Consent to Participate Not applicable.

Consent to Publish Not applicable.

Conflict of Interest The authors declare that they have no competing interests.

References

- Ahmadi M, Moeini A, Ahmadi H et al (2019) Comparison of the performance of SWAT, IHACRES and artificial neural networks models in rainfall-runoff simulation (case study: Kan watershed, Iran). *Phys Chem Earth Parts A/B/C* 111:65–77. <https://doi.org/10.1016/j.pce.2019.05.002>
- Beven K (2002) Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrol Processes* 16:189–206. <https://doi.org/10.1002/hyp.343>
- Brinkerhoff CB, Gleason CJ, Feng D, Lin P (2020) Constraining remote river discharge estimation using reach-scale geomorphology. *Water Resour Res* 56(11):e2020WR027949. <https://doi.org/10.1029/2020WR027949>
- Chemykh DV, SamoiloVA GS (2011) Landscapes of Altai (Republic of Altai and Altai Krai). In: The 1:500 000 Map. FGUP Novosibirskaya Kartograficheskaya Fabrika, Novosibirsk (in Russian)
- Conrad CP (2019) Seasonal precipitation influences streamflow vulnerability to the 2015 drought in the western United States. *J Hydrometeorol* 20(7):1261–1274. <https://doi.org/10.1175/JHM-D-18-0121.1>
- Corripio JG, López-Moreno JI (2017) Analysis and predictability of the hydrological response of mountain catchments to heavy rain on snow events: a case study in the Spanish Pyrenees. *Hydrology* 4(2):20. <https://doi.org/10.3390/hydrology4020020>
- Feng Z, Niu W, Tang Z et al (2020) Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization. *J Hydrol* 583:124627. <https://doi.org/10.1016/j.jhydrol.2020.124627>
- Gvozdet'skiy NA (1968) Physical-geographical zoning of the USSR. Characteristics of regional units. MSU Publishing house, Moscow (in Russian)

- Gvozdetkiy NA, Mikhailov NI (1987) Physical geography of the USSR. The Asian part. Mysl' Publishing House, Moscow (in Russian)
- Iooss B, Lemaître P (2015) A review on global sensitivity analysis methods. In: Meloni C, Dellino G (eds) Uncertainty management in simulation-optimization of complex systems: algorithms and applications. Springer, Boston, pp 101–122. https://doi.org/10.1007/978-1-4899-7547-8_5
- Kirsta YB (2011) Spatial generalization of climatic characteristics in mountain areas. World Sci Cult Educ 3: 330–337. <http://amko.ru/index.php/english/journals/> Accessed 1 June 2020
- Kirsta YB (2020) System-analytical modelling: 2. Assessment of runoff model sensitivity to environmental factor variations. EJMSA 8(3):67–77. <https://doi.org/10.32523/2306-6172-2020-8-3-67-77>
- Kirsta YB, Puzanov AV (2020) System-analytical modelling: 1. Development of regional models for mountain river runoff. EJMSA 8(2):69–85. <https://doi.org/10.32523/2306-6172-2020-8-2-69-85>
- Koch M, Cherie N (2013) SWAT-modeling of the impact of future climate change on the hydrology and the water resources in the upper blue Nile river basin, Ethiopia. In: Proceedings of the 6th International Conference on Water Resources and Environment Research. Koblenz, Germany, pp 428–523
- Kokorin AO (2011) Assessment report: climate change and its impact on ecosystems, population and economy of the Russian part of the Altai-Sayan Ecoregion. WWF-Russia, Moscow. https://wwf.ru/upload/iblock/7c5/assessment_climate_altai_eng_.pdf. Accessed 1 June 2020
- Kundzewicz ZW, Su BD, Wang YJ et al (2019) Flood risk and its reduction in China. Adv Water Resour 130: 37–45. <https://doi.org/10.1016/j.advwatres.2019.05.020>
- Mirkin BM, Rozenberg GS (1978) Phytosociology: Principles and methods. Nauka, Moscow (in Russian)
- Mohammad-Azari S, Bozorg-Haddad O, Loaiciga HA (2020) State-of-art of genetic programming applications in water-resources systems analysis. Environ Monit Assess 192:73. <https://doi.org/10.1007/s10661-019-8040-9>
- Moriasi DN, Arnold JG, Van Liew MW et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulation. Trans ASABE 50(3):85–90. <https://doi.org/10.13031/2013.23153>
- Mosavi A, Ozturk P, Chau K (2018) Flood prediction using machine learning models: Literature review. Water 10(11):1536. <https://doi.org/10.3390/w10111536>
- Musselman KN, Lehner F, Ikeda K et al (2018) Projected increases and shifts in rain-on-snow flood risk over western North America. Nat Clim Chang 8(9):808–812. <https://doi.org/10.1038/s41558-018-0236-4>
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I – A discussion of principles. J Hydrol 10:282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Nikolaev AN, Skachkov YB (2012) Snow cover and permafrost soil temperature influence on the radial growth of trees in Central Yakutia. J Siberian Federal Univ Biol 5(1):43–51 (in Russian). <http://elib.sfu-kras.ru/handle/2311/3009>. Accessed 1 June 2020
- Parisouj P, Mohebzadeh H, Lee T (2020) Employing machine learning algorithms for streamflow prediction: a case study of four river basins with different climatic zones in the United States. Water Resour Manag 34: 4113–4131. <https://doi.org/10.1007/s11269-020-02659-5>
- Roessler O, Froidevaux P, Boerst U et al (2014) Retrospective analysis of a nonforecasted rain-on-snow flood in the Alps – a matter of model limitations or unpredictable nature? Hydrol Earth Syst Sci 18:2265–2285. <https://doi.org/10.5194/hess-18-2265-2014>
- Song X, Zhang J, Zhan C et al (2015) Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. J Hydrol 523:739–757. <https://doi.org/10.1016/j.jhydrol.2015.02.013>
- Tabari H (2019) Statistical analysis and stochastic modelling of hydrological extremes. Water 11:1861. <https://doi.org/10.3390/w11091861>
- Tanasienko AA, Putilin AF, Artamonova VS (1999) Environmental aspects of erosion processes: An Analytical review. SPSTL SB RAS & ISSA SB RAS, Novosibirsk (Series “Ecology” No. 55) (in Russian). <http://www.spsl.nsc.ru/win/lc3/tanasi.pdf>. Accessed 1 June 2020
- Tardy Y, Bustillo V, Boeglin JL (2004) Geochemistry applied to the watershed survey: hydrograph separation, erosion and soil dynamics: A case study: the basin of the Niger River, Africa. Appl Geochem 19:469–518. <https://doi.org/10.1016/j.apgeochem.2003.07.003>
- Tsimbaley Y (2011) Landscape map of Altai Krai: Maps (with using the materials from Purdik LN, Bulatov VI, Kovanova AA, and with the participation of Chernykh DV, Smirnov SB, Vinokurova OM). Library materials of IWEP SB RAS, Barnaul (in Russian)
- Tullos D, Byron E, Galloway G et al (2016) Review of challenges of and practices for sustainable management of mountain flood hazards. Nat Hazards 83(3):1763–1797. <https://doi.org/10.1007/s11069-016-2400-3>
- Wang F, Huang GH, Fan Y, Li YP (2020) Robust subsampling ANOVA methods for sensitivity analysis of water resource and environmental models. Water Resour Manag 34:3199–3217. <https://doi.org/10.1007/s11269-020-02608-2>

- Wijayarathne DB, Coulibaly P (2020) Identification of hydrological models for operational flood forecasting in St. John's, Newfoundland, Canada. *J Hydrol Regional Stud* 27:100646. <https://doi.org/10.1016/j.ejrh.2019.100646>
- Wu WY, Emerton R, Duan QY et al (2020) Ensemble flood forecasting: Current status and future opportunities. *Wiley Interdiscip Rev Water* 7(3):e1432. <https://doi.org/10.1002/wat2.1432>
- Zhang X, Peng Y, Xu W, Wang B (2019) An optimal operation model for hydropower stations considering inflow forecasts with different lead-times. *Water Resour Manage* 33(1):173–188. <https://doi.org/10.1007/s11269-018-2095-1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.