A universal approach to runoff processes modelling: coping with hydrological predictions in data-scarce regions

O. M. SEMENOVA[1] & T. A. VINOGRADOVA[2]

1 *Department of Experimental Hydrology and Mathematical Modelling of Hydrological Processes, State Hydrological Institute, 23 2-ya liniya VO, 199053 St Petersburg, Russia*
omakarieva@gmail.com

2 *Department of Geography and Geoecology, St Petersburg State University, 31/33 10-ya liniya VO, 199178 St Petersburg, Russia*

Abstract This paper discusses the features which a hydrological model should possess to be successfully applied in the task of hydrological predictions in poorly gauged regions. The Deterministic Modelling Hydrological System developed on the basis of the principle of universality is described as an example of such a model. The results of the simulations conducted across the data scarce basins of eastern Siberia are presented.

Key words principle of universality; Deterministic Modelling Hydrological System, DMHS; data scarce regions; eastern Siberia

INTRODUCTION

Physically-based distributed modelling can be considered to be a core component of modern hydrological science. The ability of the hydrological community to identify and master the principal problems of hydrological modelling is crucial for progress to be made in the applied task of hydrological prediction; the issue of its further development is currently under intense consideration by many authors, for example, Beven (2006).

Data scarce regions present a particular challenge for hydrological modelling. Models developed for specific river basins and/or for describing specific groups of runoff formation processes become useless in the absence of long-term or detailed observations. The reason is that they have incomplete physical validity and therefore require intensive parameter calibration for any new application.

What are the approaches to which the hydrological modelling has to move from calibration-based ones? The essential fundamentals of physics suggest that the process of runoff formation must be the same in any point of space, implying the possibility of simulating runoff formation processes in any basin within the framework of a single (or universal) methodological approach, its mathematical realization and unified informational support.

The principle of universality is the base of the Deterministic Hydrological Modelling System (DHMS) (often called “Hydrograph”). This model is being developed at the State Hydrological Institute (St Petersburg, Russia) by Prof. Yu. B. Vinogradov. Its successful application to basins of different scales in a variety of geographical zones without change of model structure, algorithms and based on a unified set of model parameters, has suggested the possibility of a general approach in hydrological modelling (Vinogradov, 1988; Vinogradov & Vinogradova, 2009).

This paper briefly presents the main properties and basic concepts of the DHMS. Several basins of eastern Siberia are used to test the model under conditions of considerable lack of any kind of data.

THE DMHS

The principle of universality

What can be the measure of model universality? And how does it relate to the adequate presentation of the natural processes? We claim that only the successful results of multifold testing of a model over basins of any type, regardless of their scale and landscape/climate characteristics, can lead us in the right direction to answer these questions.

The idea of universality requires the fulfilment of several important concepts which are strongly related to the approach of model parsimony.

The universal model should use only existing and commonly available meteorological forcing data, which should be available for any territory, even if they are very limited.

The appropriate characteristics describing runoff conditions (in other words, parameters of the model) should be general for any basin and at the same time take into account their unique properties. Reflecting the objective properties of watersheds they obviously would have very clear physical meaning. The possibility for *a priori* estimation of the model parameters on the basis of some general idea about the features of the basin, and with the use of any indirect information or expert evaluation, is essential. The model is valuable if the parameters can be systematized, generalized and normalized in relation to calibrated parameters of models of other types.

General description of the DMHS

The DHMS is a distributed model of runoff formation processes. It describes all components of the land hydrological cycle, including: precipitation and its interception; snow accumulation and melt; evaporation from snow, soil and vegetation cover; surface flow and infiltration; soil water dynamics and flow; heat dynamics and phase change in soil layers; underground flow formation, slope and channel flow transformation; and flow discharge. It is designed to be applied in any geographical area of the Earth.

The model forcing data consist of the standard and most simple meteorological information from observational networks that can be obtained even for data scarce regions; these include daily values of air temperature, moisture deficit, and precipitation. The performance of the model is improved by some approaches which are used for estimation of the effective characteristics of the input data. For example, effective temperature differs from its common analogue by an additive constant that is computed as a function of given latitude, elevation, terrain slope and orientation toward direct solar radiation. The specific definition accounts for climatic thermal gradients for temperature and for precipitation interpolation from the precipitation ratio to its annual total that is independently estimated.

The various outputs of the model are the continuous runoff hydrographs at the outlet, from any part of the basin or a specified landscape; the distributed state variables, reflecting water and heat dynamics in soil layers and snow cover; spatial and temporal distribution of water balance elements including precipitation; evaporation from snow, soil and vegetation cover; surface, subsurface and underground runoff.

The parameters and characteristics of the model are horizontally distributed as a system of representative points and runoff formation complexes in space, and vertically, deep into the soil column and layers of underground runoff. Most parameters have strong physical meaning and are assessed *a priori*. The set of parameters describing one landscape can be used both for small and large basins without change of values.

The spatial-computational schematization of the basin

In the framework of the DMHS, the basin is represented by a system of ordered points which are situated within the basin territory. A regular hexagonal mesh is used (the distance between neighbouring points is equal). Each calculation point, or so-called “representative point” (RP) corresponds to its own area, which is considered to be homogenous in all characteristics such as absolute altitude, orientation, inclination and others. The RP quantity depends nonlinearly on basin area, its orographic structure, density and evenness of meteorological station locations.

The regular system of RPs is combined with the scheme of runoff formation complexes (RFC). The RFC can be identified with any kind of landscape; it is the part of a river basin where the process of runoff formation is assumed to be uniform. While allocating RFCs, the principles of homogeneity of soil, vegetable and landscape characteristics of the basin surface are followed. By means of these criteria, the areas with considerably different relief and elevation are separated into different RFCs. The information about most of the model parameters is related and systematized to the RFC; their values remain fixed within its range and change step-wise at its borders.

The distribution of the RFCs within each RP area is defined. The exact location of the specific RFC in an RP is not important. Runoff variables are calculated for each RFC and summed up according to the relative fractions of area within the RP hexagon; then they are summed up for the whole basin according to the RP's lag time.

The weather forcing data is interpolated from meteorological stations into RPs, and all RFCs within one RP get the same input data. Further, the forcing data (for example, precipitation) are modified according to the landscape qualities.

The example of basin schematization for one of the studied basins is shown in Fig. 1.

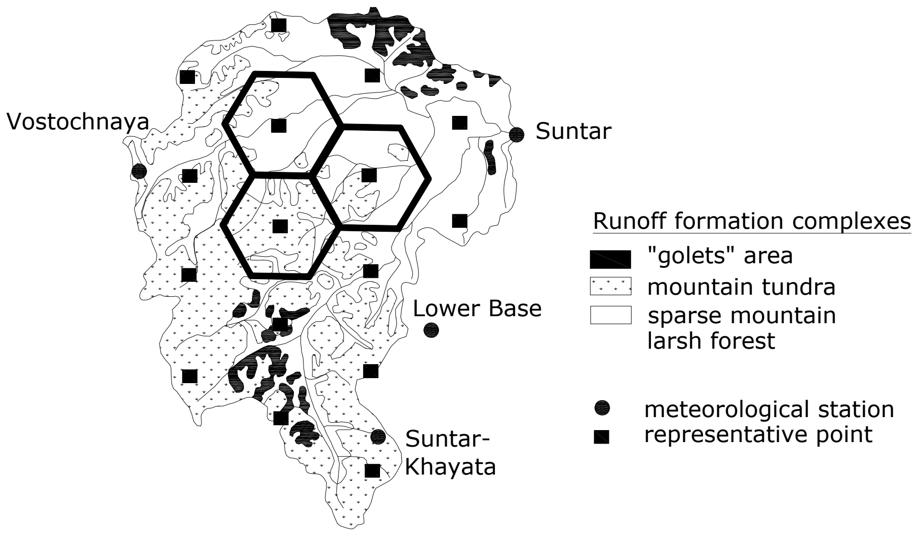


Fig. 1 The spatial-computational schematization. The Suntar River basin at the Sakharynya mouth (basin area: 7680 km^2).

The concept of runoff elements

The approach used for describing water dynamics within the river basin greatly determines the structure of runoff formation model. The use of partial differential equations, such as equations of St Venant or kinematic wave for surface and channel flow, and the Boussinesq equation for underground waters is prevalent in modern deterministic hydrological models. For the numerical solution of the differential equations of movement and continuity, it is required to approximate the basin surface structure by a set of finite elements, usually taking into consideration relief and landscape type. The assumption of the presentation of water movement as a thin continuous water layer is very approximate and not in correspondence with natural process of water motion over the real surface and underground slopes. The incomplete physical validity of these models is aggravated by the lack of reliable information about the real conditions of water movement, requirement of large amount of information about inclinations, morphology, "roughness" coefficients for solving these equations. Here the calibration procedure is necessary in which the parameters are optimized to fit the observations. Because the parameters are evaluated not individually, but in complexes, they often lose any physical meaning and can not be reliably transferred to the data scarce basins.

The DMHS uses an integral approach for describing water movement within the basin – the concept of runoff elements (Vinogradov, 1988, 2003b). A runoff element is a part of surface or underground elementary slope or watershed limited by micro-divides which is oriented with its open part towards the slope non-channel or underground drainage system. The basin is constituted by a set of elementary slopes or watersheds, which, in turn, consists of a system of runoff elements which can be surface, subsurface and underground ones. The size of a runoff element depends on inclination; the underground runoff elements are larger than the surface ones.

The types and features of runoff elements determine the transformation character of runoff formation hydrographs at the point of origin of water flux into the channel system.

According to Vinogradov (1988, 2003b), the equation for water flux, Q, from all runoff elements of the given level to the channel system is:

$$Q = \frac{S+b}{1+[(S-Q_0)/(Q_0+b)]\exp[-a\Delta t(S+b)]} - b \tag{1}$$

Here, Q_0 is the initial value of runoff Q, and S is the runoff formation intensity (m^3 s^{-1}); Δt is the computation time interval (s) during which S is constant; $a = a^* \times F^{-1}$ and $b = b^* \times F$, where a^* and b^* are normalized hydraulic coefficients with units m^{-1} and m s^{-1}; and F is the basin area (m^2).

The total water flux to the channel system is described by an equation system such as equation (1) when values S, Q, a and b are different for the multitude of surface, soil and underground runoff elements of different levels that form a river basin.

The derivation of equation (1), the hierarchical system of runoff elements and related description of flow types, the ranges of a^* and b^* parameters values are worked out in Vinogradov (1988, 2003b). He marks out surface, subsurface types of runoff and 15 underground layers corresponding to rapid ground, ground, upper, deep and historical underground types of runoff. Each runoff type (and layer of runoff elements) is characterized by the specific values of residence time $\tau^* = 1/(a^* b^*)$, outflow q and water storage H. It is assumed that the outflow rate decreases and water storage increases with depth in groundwater aquifers.

The concept of runoff elements allows carrying out simulations for basins of any size; since the value of basin area is introduced into the calculation scheme.

The parameters of the DMHS

The parameters of the DMHS can be divided into five groups according to landscape components: soil column (unsaturated zone), vegetation cover, slope surface, underground runoff and climate parameters.

The problem of heterogeneity of earth surface characteristics (and thus of model parameters) is considered to be one of the fundamental challenges of hydrological science. In particular, it refers to the physical properties of soils. However, consideration must be given to the fact that the variation of soil properties depends on the areal extent used for estimating or measuring their values.

The DMHS parameters of the soil (unsaturated zone) for different soil depths are density; porosity; maximum water holding capacity; infiltration coefficient; specific heat capacity and conductivity; the index of ice content influence at infiltration; the contribution ratio to evaporation; hydraulic parameter of soil runoff elements. They should represent the values of the soil strata typical (or representative) for the specific RFC. For Russian basins such values are obtained from published agricultural-hydrometeorological surveys, or estimated on the basis of observations at water-balance stations (a highly instrumented small watershed intended for long-term collection of observations) or experimental watersheds. Our experience of runoff process simulations indicates that such data can be systematized for different landscapes and are stable enough. In seeking the appropriate range of values that do not have much variation, it is not necessary to search for the exact values of these parameters. In contrast to calibrated and optimized parameters they can be extrapolated to ungauged basins for any forecasting and research task.

Parameters of the vegetation cover include: four phenological dates; maximum and minimum values of seasonal shadow fraction by vegetation cover; interception water capacities; landscape albedos; the coefficients of potential evaporation; and the coefficient of evaporation from the interception storage during the maximum development of vegetation cover. This information can be found in special literature related to geobotanical, agricultural and climate research.

The parameters of the slope surface are: maximum and minimum values of the snow redistribution coefficient; spatial variation coefficient of SWE in snow cover; spatial variation of infiltration capacity of upper soil layer; maximum ponding fraction; maximum surface depression storage; and hydraulic parameter of surface runoff elements. The parameters describing snow characteristics are assessed against snow survey data. The spatial variation of the infiltration capacity of the upper soil layer can be calibrated against runoff observations on small watersheds or water-balance stations. The properties describing the surface storage process are obtained from the literature.

The parameters of the underground system of runoff elements are: the hydraulic parameter and redistribution values. The hydraulic parameter is usually assumed to be constant; the values of redistribution of water volume among modelling groundwater layers are assigned on the basis of observed hydrograph analysis within the concept of runoff elements. For runoff hydrographs of similar type these values are closely related.

Climate parameters, if not obtainable, are estimated using the forcing meteorological data. Here the serious problem of interpolation of precipitation in data poor regions (mainly mountainous areas) is to be mentioned. While implementing the DHMS, we use the approach of normalizing daily precipitation layer by its mean annual value. The assessment of mean annual values for each representative point depending on its altitude and location is a special problem to be solved.

THE STUDY OBJECTS

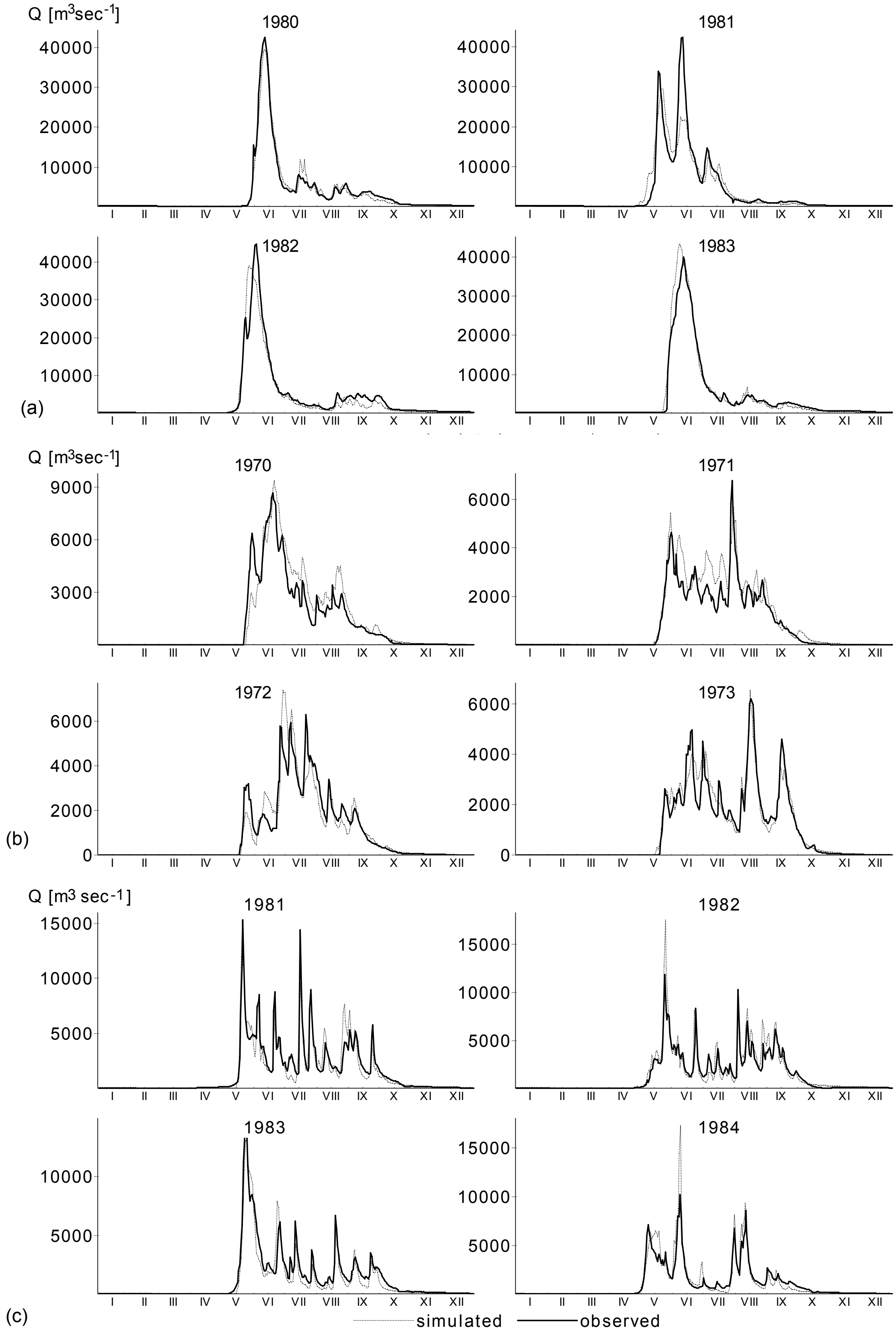
To illustrate the described principles, four basins within the territory of eastern Siberia have been chosen for the simulation of runoff formation processes. The basins are of different sizes and represent different landscape characteristics. The basin selection is summarized in Table 1.

All study basins are situated in the zone of continuous permafrost; they have mountainous relief, and the climate is characterized as severe continental. The main landscape type is taiga dominated by larch. The rivers have mixed snow melting and rainfall supply.

Well-defined maximum discharge in June, due to snowmelt, is typical only for the Nizhnaya Tunguska; the other basins are subject to intensive summer rains causing flows comparable to and greater than spring snowmelt.

Table 1 Description of watersheds used in the study.

No.	River; outlet	Basin area (km^2)	Average elevation (m)	Mean annual runoff (mm) / discharge ($m^3 s^{-1}$)	Number of RP	Number of meteorological stations (including those situated inside the basin)	Location
1	Nizhnaya Tunguska; Bolshoy Porog	418 000	500	245 / 3400	51	19 (8)	The right tributary of the Yenisey River
2	Yana; Dgangky	216 000	830	120 / 965	45	15 (8)	The river of the Laptev Sea basin
3	Uchur; Chul'bu	108 000	1100	380 / 1300	48	9(3)	The tributary of the Aldan River (Lena River basin)
4	Suntar; Sakharynya River Mouth	7 680	1500	176 / 43	16	3 (1)	The Headstream Of The Indigirka River
5	Detrin; Vakhanka River Mouth	5 630	920	324 / 58	15	6 (2)	The headstream of the Kolyma River



T

Fig. 2 Simulated and observed hydrographs for: (a) the Nizhnaya Tunguska River at Bolshoy Porog (basin area: 418 000 km^2), 1980–1983; (b) the Yana River at Dgangky (basin area: 216 000 km^2), 1970–1973; and (c) the Uchur River at Chul'bu (basin area: 108 000 km^2), 1981–1984.

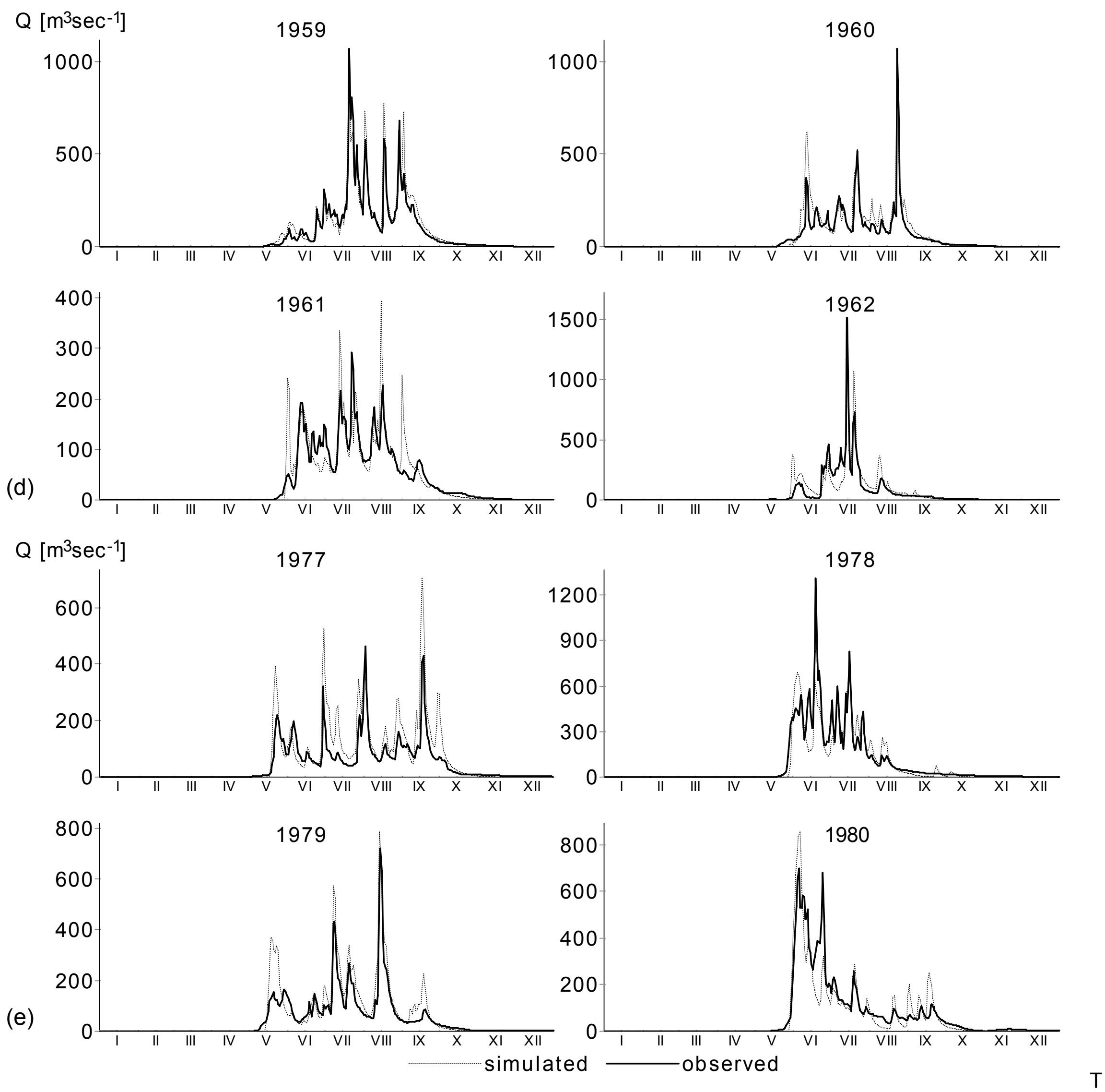


Fig. 2 *continued* Simulated and observed hydrographs for: (d) the Suntar River at the Sakharynya River mouth (basin area: 7680 km^2), 1959–1962; and (e) the Detrin River at the Vakhanka River mouth (basin area: 5630 km^2), 1977–1980.

RESULTS

The study involved runoff modelling for four watersheds with 24-hour calculation interval for the different periods (from 7 to 19 years). The comparison of observed and simulated hydrographs is shown in Fig. 2(a)–(e).

In general, the model simulations capture the shape and depletion curves of observed hydrographs during both the snowmelt and rainy periods; the discrepancy between observed and simulated volumes of flood peaks can be related to the uncertainties of precipitation data.

Table 2 presents the statistical characteristics evaluating the model performance of runoff simulations. They include: mean annual observed runoff, H_{obs}, and calculated runoff, H_{calc} (mm); the Nash-Sutcliffe efficiency, *Ef*, describing the quality of simulated runoff compared to observed data for daily and annual values; and the relative error, *Er*, between simulated and observed daily and mean annual runoff.

The *Ef* and *Er* are calculated as:

$$Ef = 1 - \frac{\sum_{i=1}^{n}(Q_{\text{calc}}^{i} - Q_{\text{obs}}^{i})^2}{\sum_{i=1}^{n}(Q_{\text{calc}}^{i} - \overline{Q_{\text{obs}}})^2} \quad (2)$$

$$Er(\%) = \frac{\sum_{i=1}^{n}\frac{\left|Q_{\text{calc}}^{i} - Q_{\text{obs}}^{i}\right|}{Q_{\text{calc}}^{i}}}{n} \times 100 \quad (3)$$

where Q_{calc}^{i} and Q_{obs}^{i} are the calculated and observed runoff at day (or year) *i* for daily (or annual) values; $\overline{Q_{\text{obs}}}$ is either: the observed annual average in the case of daily flows, or the long-term average in a case of annual flow values; and *n* is the number of the days in the year (or the number of years).

Table 2 Statistical characteristics of flow simulations.

Basin	Period	H_{obs}	H_{calc}	*Ef* Daily	*Ef* Annual	*Er, %* Daily High flow	*Er, %* Daily Low flow	*Er, %* Daily Average	*Er, %* Annual
Nizhnyaa Tunguska River at Bolshoy Porog	1978–1984	242	234	0.91	0.98	28	35	32	4
Yana River at Dgangky	1966–1984	120	130	0.81	0.80	40	88	72	10
Uchut at Chul'bu	1977–1984	380	380	0.81	0.95	32	36	35	8
Suntar River at the Saharynya River mouth	1957–1964	176	181	0.71	0.91	38	76	63	10
Detrin River at the Vakhanka River mouth	1977–1984	324	354	0.75	0.94	40	79	66	12

The relative error, *Er*, for daily values is presented in three variants: for the periods of high and low flow and the average for the year. This follows the seasonal cycle of runoff which indicates low flows during October–April and high flows during May–September. Therefore, by the high-flow period we imply the warm part of the year with 90–98% of annual runoff; depending on the basin it starts in different 10-day periods of May and finishes in September or October. For periods of river freeze-up (observed value equals zero), *Er* was not calculated.

For all basins, the calculated Nash-Sutcliffe efficiency, *Ef*, exceeds 0.70 and for three of them it exceeds 0.80 for daily values; for annual values it exceeds 0.80. Such values of *Ef* are usually considered to be rather high and indicate good model performance. But the main weakness of this criterion is the overestimation of simulation efficiency for the periods of high flow and small contribution of low flow periods to its estimation; it is insensitive to systematic under- or overestimations of simulated flow.

The annual *Er* varies within the range 4–12%, showing maximum values for smaller basins. This may be due in part to the higher dependence of small basins on precipitation input from very few, or even a single, meteorological station(s).

For the basins of Nizhnyaa Tunguska and Uchur rivers, the average *Er* for daily values is around 35%, with little variation through the year. The average *Er* for the other three basins (nos 2, 4 and 5 in Table 1) exceeds 60% (even amounting to 72% for the Yana River). This is caused by high values of *Er* (76–88%) during the low-flow period; it can hardly be seen from Fig. 2, but for these specific basins the model systematically overestimates the baseflow.

The reason for such high discrepancies is that while the rivers 2, 4 and 5 are subject to complete freezing during the cold period when the observed runoff values equal zero, the calculation algorithm for runoff elements does not reflect these processes. Some modification of the calculation scheme should be done for the conditions of very low winter temperatures when the outflow from the deep groundwater reservoirs is hampered by ice.

DISCUSSION

Estimation of the credibility and acceptability of the simulation results is the natural stage of any model application. It is clear that the greater the quantity and variety of verified simulations, the more likely it is that the concepts underlying the model are not defective.

The possibility of getting similar results, complying with the observational data, by the use of different models (the problems of non-uniqueness and equifinality: Beven, 2001) relates to the active use of parameter calibration and ignoring of the principle of universality in the methodology of model development. Numerous calculations for the basins situated in different climate and landscape zones can impressively improve the models' value.

The task of choosing of appropriate evaluation criteria is important. According to (Vinogradov, 2003a), the relative error at every time step of the calculations, its annual distribution and its variation over the simulation period, can objectively affect the model efficiency. It is significant that the value of the relative error if averaged should be taken in its absolute value without mutual compensation of negative and positive deviations.

REFERENCES

Beven, K. (2001) How far can we go in distributed hydrological modelling? *Hydrol. Earth System Sci.* **5**(1), 1–12.

Beven, K. (2006) Searching for the Holy Grail of scientific hydrology: Qt=H(S,R,Δt)A as closure. *Hydrol. Earth System Sci.* **10**, 609–618.

Vinogradov, Yu. B. (1988) *Mathematical Modelling of Runoff Formation. A Critical Analysis.* Gidrometeoizdat, Leningrad (in Russian).

Vinogradov, Yu. B. (2003a) River runoff modelling. In: *Hydrological Cycle* (ed. by I. A. Shiklomanov), *Encyclopedia of Life Support Systems (EOLSS).* Developed under the auspices of UNESCO. EOLSS Publishers, Oxford, UK (http://www.eolss.net).

Vinogradov, Yu. B. (2003b) Runoff generation and storage in watershed. In: *Hydrological Cycle* (ed. by I. A. Shiklomanov), *Encyclopedia of Life Support Systems (EOLSS).* Developed under the auspices of UNESCO, EOLSS Publishers, Oxford, UK (http://www.eolss.net).

Vinogradov, Yu. B. & Vinogradova, T. A. (2008) *Current Problems in Hydrology.* Academia Publishers, Moscow, Russia (in Russian).

Vinogradov, Yu. B. & Vinogradova, T. A. (2009) *Mathematical Modelling in Hydrology.* Academia Publishers, Moscow, Russia (in Russian) (in press).