FLOOD RISK ASSESSMENT IN A POORLY GAUGED BASIN

Alexander Gelfan

1Water Problems Institute, Russian Academy of Sciences, Moscow, Russia.

5.1 ABSTRACT
An approach for flood risk assessment in a poorly gauged basin has been proposed and tested for the Sosna River basin in European Russia. The approach involves searching a data-rich small proxy basin which is hydrologically similar to the poorly gauged study basin, developing a physically based model of flood generation in the proxy basin, and transferring the developed model with adjustments to the study basin. In this case study, the adjustment was carried out through the model calibration against snow and soil freezing survey data in the study basin; streamflow data were not used for the calibration. Long-term artificial time series of daily weather variables were Monte Carlo simulated and input to the hydrological model to generate a corresponding series of snowmelt flood hydrographs in the study basin. Frequency distributions of flood characteristics (volume and peak discharge) were derived from the long-term series of the modelled hydrographs. The approach allows the derivation of frequency distribution of flood volume without utilizing any streamflow observations in the study basin; however, in order to obtain reliable frequency distribution of flood peak discharge, several years of streamflow observations should be used for the additional calibration of the model. The proposed approach is targeted for hydrological engineering practice and considered as a suitable alternative to the traditional methods of flood risk assessment in ungauged or poorly gauged basins.

5.2 RÉSUMÉ
Une approche d’évaluation des risques d’inondation dans un bassin fluvial avec un nombre insuffisant de limnimètres a été proposée et essayée pour la rivière de Sosna située dans la partie européenne de Russie. Cette approche
comprend une recherche d’un petit bassin témoin riche en données qui est semblable du point de vue hydrologique au bassin exploré mal calibré, un développement d’une modèle physiquement justifiée de la formation des inondations dans le bassin témoin ainsi qu’un transfert de la modèle développée avec les ajustements dans le bassin exploré. Dans cette étude de cas l’ajustement a été fait par le moyen d’étalonnage du modèle contre les données d’un sondage sur la couverture de neige et congélation du sol dans le bassin exploré. L’étalonnage n’a pas utilisé les données sur le débit d’eau fluvial. Monte Carlo simulation a été effectuée pour obtenir des séries de temps artificielles à long terme des variables météorologiques quotidiens qui ont été entrées dans la modélisation hydrologique afin de générer des séries correspondantes d’hydrographes d’inondation de la fonte des neiges dans le bassin exploré. Distributions de fréquence des caractéristiques d’inondation (volume et débit de pointe) ont été dérivées de séries à long terme des hydrographes modelés. Cette approche permet de dériver les distributions de fréquence du volume de l’inondation sans utiliser aucunes observations sur le débit de l’eau fluvial dans le bassin exploré. Pourtant, pour obtenir les données fiables sur la distribution de fréquence du débit de pointe de l’inondation, plusieurs années d’observations sur le débit de l’eau fluvial doivent être utilisées pour avancer l’étalonnage du modèle. L’approche proposée est axée sur la pratique d’ingénieur hydrologique et est considérée une alternative convenable à méthodes traditionnelles d’évaluation des risques d’inondation dans les bassins manquants ou sans limnimètres.

5.3 INTRODUCTION

Planning and design of water resources systems, flood management, and protection are fundamentally dependent on reliable estimates of flood risk. Most countries use a set of empirical methods for flood risk assessment, and among these methods the flood frequency analysis (FFA) is the most commonly used one for over a century. The standard at-site FFA is based on acquisition of data of flood extremes, computation of observed probabilities of occurrence, fitting of the appropriate probability distribution to the observed probabilities with use of an appropriate parameter estimation technique, and, finally, estimation of flood quantiles of the desired probabilities. The fundamental weakness of the FFA is widely known (see, for example, discussions in Klemeš, 1986; Singh and Strupczewsky, 2002) and arises, first of all, from lack of available streamflow data for the
overwhelming majority of river basins. This is the case especially when the interest is in the assessments of extreme floods with return periods of hundreds of years, i.e. much longer than the period of flood observations. The data deficit results in increasing uncertainty of assessments for the desired extreme floods that can negate the practical value of these assessments. Moreover, in many countries, the gauging network has been reduced for the past few decades. Presently, the number of stream gauges in Russia, for instance, is about 70% of the density of the hydrological network that existed during the 1980s. The density of the hydrological network is about an order of magnitude less than the minimum density recommended by WMO (1994), and one can assume that the difference between the number of gauged basins and ungauged basins is of the same order. Regional statistical analyses of flood frequency can to some extent compensate for the lack of temporal data, but an additional, spatial dimension is introduced (Bobée and Rasmussen, 1995) that leads to increasing uncertainty of the desired estimates.

Attempts at improving the FFA so far have been focused on the statistical aspects, such as improvement of the parameter estimation techniques, seeking probability distribution for improving goodness of fit, etc. Lack of knowledge about the form of the parent distribution limits these attempts; however, even if one assumes that the distribution form is known, the paucity of the available observations per se leads to the unreliable results of the extrapolation above the maximum observed flood. As an example, Figure 5.1 presents the annual maximum peak discharges of the Seim River (centre of European Russia) observed during 61 years (beginning from 1928) and fitted by a three-parameter gamma-distribution curve. Note that in the range of the available observations, this distribution is indistinguishable from the other distributions typically used for FFA of annual maximum discharge, e.g. Log-Pearson type III distribution. Another gamma-distribution curve is fitted to a 56-year sample obtained by exclusion of the first five years of discharge (from 1928 to 1932) from the original, 61-year sample. The legitimacy of the fitted curves cannot be doubted, but there is an obvious difference between these curves when extrapolated to extreme floods. The exceedance probability of the maximum observed discharge (2230 m³/s), estimated by the 61-year sample, is 0.012 (83-year return period). If, for some reason, the observations began 5 years later (in 1933 instead of 1928), the corresponding exceedance probability of the same discharge would be four times less (0.003, i.e. 333-year return
period). This presents a major difficulty in applying such statistical techniques with limited time series for flood prediction purposes in ungauged basins.

An opportunity for refinement of the extreme flood frequency assessment is associated with the inclusion of deterministic, physical information in addition to statistical information extracted solely from the runoff observation series. As Klemeš (1993) states: “if more light is to be shed on the probabilities of hydrological extremes, then it will have to come from more information on the physics of the phenomena involved, not from more mathematics.” Such additional information may be both \textit{a posteriori}; empirical information about factors affecting flood generation (e.g., meteorological factors, watershed conditions), and \textit{a priori} information, reflecting accumulated knowledge on flood generation physics. In other words, lacking homogeneous runoff data for the standard FFA, the data deficit may be partly compensated by deterministic information on physical processes and stochastic information on better defined forcing variables, e.g. meteorological variables.

\textbf{Figure 5.1} Gamma-distribution curves fitted to 61-year (solid line; solid circles) series of observations at the Seim River beginning from 1928 and 56-years (dashed line; open circles) series beginning from 1933 (explanations are in the text).
Development of such a model which is based on the deterministic description of the hydrological processes and takes into account available stochastic information on input meteorological variables is the subject of the dynamic-stochastic approach to flood risk assessment (Kuchment and Gelfan, 1991). The resulting dynamic-stochastic model integrates two components: a physically based deterministic model of runoff generation and stochastic models of the meteorological variables which are the inputs into the deterministic model. The integration of dynamic and stochastic approaches opens an opportunity for assessment of magnitude/frequency of extreme floods in the basins where series of streamflow data are too short and/or statistically non-homogeneous due to anthropogenic pressure on environment and climate change (“virtually ungauged basins” (He et al., 2011)), in other words, in the basins where the standard FFA is ineffective.

Eagleson (1972) was the first who proposed a dynamic-stochastic model based on the physically based description of hydrological processes of rainfall flood generation; he derived a distribution function of flood peak discharge through integration of joint probability distribution of rainfall intensity and duration over the domain determined from the analytic solutions of the kinematic wave equation. Eagleson’s (1972) approach to flood frequency estimation has been used and extended by Carlson and Fox (1976), Chan and Bras (1979), Hebson and Wood (1982), Diaz-Granados et al. (1984), Bras et al. (1985), Blöschl and Sivapalan (1997), and others.

Bras et al. (1985) compared abilities of the models of Eagleson (1972), Hebson and Wood (1982), and Diaz-Granados et al. (1984) to derive flood frequency distribution for ungauged basins. Five river basins located in the different physiographic and climatic conditions with catchment areas from 100 to 1000 km² were selected and it was assumed that no streamflow data were available, so that the parameters of the rainfall-runoff components of the models were assigned a priori. The return periods of flood peak discharges derived by each of the models for the five basins were compared with the return periods estimated from the available data of observations. The comparison has shown that none of these models agreed well with the observations. Bras et al. (1985) concluded that performance of these models could be significantly improved if some observation data could be used for calibration of the rainfall-runoff models.
Blöschl and Sivapalan (1997) used the derived distribution approach in order to test the “index flood” concept underlying the standard procedure of determination of homogeneous geographical regions in regional FFA often used for ungauged basins. The authors interpreted data from hundreds of catchments in Austria and showed that the coefficient of variation of peak discharge depends on the basin area, which contradicts the principal “index flood” assumption of independence of variation of peak discharge.

An alternative, numerical technique for deriving flood frequency is a Monte Carlo simulation-based dynamic-stochastic modelling allowing one to combine a sophisticated model of runoff generation with a stochastic weather generator. A physically based model of rainfall flood was first combined with the Monte Carlo continuous simulation of precipitation and air humidity series by Kuchment et al. (1983); they showed for small basins that extreme flood frequency numerically derived by the use of the available short series of streamflow data for the model calibration is more reliable than flood frequency obtained by the standard statistical analysis of that series. Recently, the numerical dynamic-stochastic approach has been applied, for example, by Franchini et al. (1996), Blazkova and Beven (1997), Hashemi et al. (2000), Sivapalan et al. (2005), Fiorentino et al. (2007), and Haberlandt et al. (2008) who used different deterministic models (TOPMODEL, ARNO, HEC-HMS and others) for derivation of rainfall flood frequency.

Considerably fewer authors have applied the numerical dynamic-stochastic approach to derive the frequency of extreme snowmelt floods (Kuchment and Gelfan, 2002; Blazkova and Beven, 2004; Gelfan, 2010), and there are no publications regarding snowmelt flood frequency assessment for ungauged or poorly gauged basins. This is rather surprising, given (1) the dominant role of snowmelt in the flow regime of rivers over vast cold regions and the associated high cost of snowmelt flood events for the economy of cold-region countries (e.g. in Russia more than 65% of the disastrous floods are of snowmelt origin) and (2) the sparse gauge network in most cold regions.

The objective of this study was to develop an approach for assessment of snowmelt flood risk in the basins where no streamflow data are available or data are too scarce for application of the FFA but for which long-term meteorological data are assumed to be available. The developed approach consists of the following steps:
1. Select a data-rich small proxy-basin which is hydrologically similar to the ungauged study basin; the criteria of similarity proposed by Kuchment and Gelfan (2009) are used for the selection (details about the criteria are presented in the next section).

2. Develop a physically based model of snowmelt flood generation for the proxy-basin and “transpose” the model to the study basin without use of the streamflow data in the latter basin.

3. Construct a stochastic weather generator using long-term meteorological observations in the study basin.

4. Assess flood risk in the study basin on the basis of the dynamic-stochastic approach combining the “transposed” model with the weather generator.

In the next sections, the proposed approach is demonstrated by the example of the Sosna River basin as the study basin.

5.4 STUDY BASIN AND PROXY-BASIN: BRIEF DESCRIPTION AND CRITERIA OF SIMILARITY

The Sosna River basin is located in the centre of European Russia, draining west into the Don River. The study area of approximately 16 300 km² (up to the outlet at Elets town, 52°37'N, 38°28'E) is situated at the steppe-forest physiographic zone (Figure 5.2). The basin terrain is a smooth plain. Soils in the area belong to a steppe type of soil formation, being mainly represented by common chernozem (black soils) and podzol with a texture varying from heavy loam to clay. About 80% of the basin area is farmed; of the remainder, pastures take up about 10%, ravines and gullies occupy 8%, and forest about 2% of the basin area. Annual air temperature is +5.9 °C, the mean air temperature in the coldest month (January) is -7.0 °C and +19.5 °C in the warmest month (July). Annual precipitation is 475 mm, about 29% of which falls as snow. The maximum snow water equivalent (SWE) varies significantly from year to year (mean SWE is 69 mm with a maximum observed value of 180 mm and minimum of 17 mm). The mean date of the beginning of snowmelt is March 27. During the snowmelt period, which averages 26 days, from 39% to 73% of annual runoff is generated (55% on average). The snowmelt runoff coefficient varies from year to year over a wide range: from 0.21 to 0.88. Mean snowmelt runoff is 72 mm; the mean peak discharge of snowmelt floods is 1783 m³s⁻¹ which is much greater than
the highest observed peak discharge of rainfall flood, 388 m$^3$s$^{-1}$. The highest peak discharge from a snowmelt flood was 4950 m$^3$s$^{-1}$ on April 5, 1970. The Sosna River basin will be considered hereafter as an ungauged basin, a so-called “pseudo-ungauged basin” (He et al., 2011).

Kuchment and Gelfan (2009) suggest that small experimental basins, particularly those included in the network of the Russian water balance stations (WBS), can be considered as good proxy-basin candidates. The network of WBS was created in different physiographic zones over the former USSR in the 1930s-1950s; more than 20 WBS existed at the end of the 1980s. Presently, the number of Russian WBS sites is shrinking significantly, but they are still a source of the unique long-term detailed data including measurements of streamflow, meteorological and snow characteristics, soil properties, hydrothermal regime of vadose zone, evaporation, groundwater, etc. (Kane and Yang, 2004).

A number of hydrological similarity criteria have been proposed in the literature (e.g. Wagener et al., 2007). Beginning from the simplest criterion of spatial proximity, the Yasenok Creek experimental basin located within the territory of the Nizhnedevitskaya WBS (51°33'N, 38°22'E), near the south-eastern boundary of the Sosna River catchment (Figure 5.2) was considered as the initial choice of the proxy-basin for the Sosna River basin.

The Yasenok Creek catchment (22 km$^2$) is located in the upper part of the Devitsa River basin draining east into the Don River. Relief is flat and the dominant soils are chernozems with some podzol. The bottom water-bearing horizon of 25-30 m depth is the main aquifer, which is drained only by main watercourses. The vegetation cover of the station belongs to a band of steppes rich in herbs with oak forests. Forests cover 4% of the catchment. The main part of the Yasenok Creek basin is occupied by arable lands. The mean annual temperature is 5.8°C; the mean annual precipitation is 507 mm. The maximum SWE varies considerably from year to year, from 34 to 124 mm. Location of the measurement gauges within the catchment area is shown in Figure 5.2.

Comparing the above descriptions one can see that the Sosna River and the Yasenok Creek basins have similar catchment attributes, such as topographic characteristics, soil and vegetation type, etc. This likeness allows one to expect that the catchments behave in a hydrologically similar manner.
Figure 5.2 Location of the study basin (Sosna River) and the proxy-basin (Yasenok Creek).
Measures of hydrological similarity (such as aridity index, topographic wetness index, runoff coefficient, bifurcation ratio, etc.) differ in terms of the processes they aim to represent (Blöschl, 2005) and a reasonable choice of the measures depends on the understanding of the dominant runoff generation mechanism in both gauged and ungauged catchments. In other words, one can talk about the similarity of the prevailing features of runoff generation rather than similarity of hydrological systems as a whole. Kuchment and Gelfand (2009) found that the dimensionless indices derived from the Richards’ equation work well as similarity measures for the arid steppe region where the infiltration excess mechanism of runoff generation is dominated. Four dimensionless similarity indices proposed by Kuchment and Gelfand (2009) were used here as the criteria of hydrological similarity: (1) the Peclet number, which is the ratio of the rates of moisture transfer by gravitational filtration and capillary diffusion; (2) the index of maximum soil capacity, which is the ratio of the infiltration to water-bearing capacities of soil; (3) the gravitational filtration efficiency, which is the ratio of the saturated hydraulic conductivity to the mean precipitation rate; and (4) the capillary filtration efficiency which is the ratio of the mean rate of capillary filtration to the mean precipitation rate. Hydraulic properties of soils needed for calculation of the above indices were adopted from the soil survey data (Department of Hydrometeorological Service for the Central-Chernozem Regions, 1975) as well as the data of the experiments made in the Hydrophysical Lab. of the State Hydrological Institute (SHI) published in (Kalyuzhny et al., 1988). For the Sosna and Yasenok catchments, the following values of the similarity indices were obtained: the Peclet numbers are 0.43 and 0.38, the free soil capacity criteria are 1.59 and 1.78, the gravitational filtration efficiencies are 81.0 and 66.7, and the capillary filtration efficiencies are 277.9 and 215.1, respectively. Taking into account a large spatial variability of the hydraulic soil properties, the differences between the similarity indices for the catchments are insignificant. Thus, the closeness of the indices was interpreted as similarity of the processes of runoff generation in the basins, and the data-rich Yasenok basin was assigned as the proxy-basin for the Sosna River basin.

5.5 MODEL OF SNOWMELT FLOOD GENERATION

Description of the model and its development for the proxy-basin

The model of snowmelt flood generation used in this study presents a modification of the model version reported in Gelfan (2010). The model
describes processes of snow accumulation and melt, water and heat transfer in a soil during its freezing and thawing, infiltration into frozen and unfrozen soil, detention of meltwater by basin storage, and overland and channel flow. Below the main equations are shown; algorithms of their solution as well as relationships for the parameters are presented in Gelfan (2010).

Dynamics of snow depth, water/ice content of snow are calculated by the equations:

\[ \frac{dH_s}{dt} = \rho_w \left[ X_s \rho_0^{-1} - (M + E_s)(\rho_i I_s)^{-1} \right] - V \quad (1) \]

\[ \frac{d}{dt} (\rho_i I_s H_s) = \rho_w(X_s - M - E_s) + F_i \quad (2) \]

\[ \frac{d}{dt} (\rho_w \theta_s H_s) = \rho_w(X_i + M - E_i - R_s) - F_i \quad (3) \]

where \( H_s \) is the snow depth; \( I_s \) and \( \theta_s \) are the volumetric content of ice and liquid water, respectively; \( X_s \) and \( X_i \) are the snowfall rate and the rainfall rate, respectively; \( M \) is the melt rate; \( E_i \) and \( E_s \) are the evaporation and sublimation rates, respectively; \( F_i \) is the rate of refreezing of meltwater in snow; \( R_s \) is the meltwater outflow from snowpack calculated taking into account the maximum liquid water-retention capacity; \( V \) is the snowpack compression rate.

Water and heat transfer in a soil during the processes of soil freezing, thawing and infiltration of water are described by the following equations:

\[ \frac{\partial W}{\partial t} = \frac{\partial}{\partial z} \left( D \frac{\partial \theta}{\partial z} + D_I \frac{\partial I}{\partial z} - K \right) \quad (4) \]

\[ c_T \frac{\partial T}{\partial t} - \rho_w LH \frac{\partial W}{\partial t} = \frac{\partial}{\partial z} \left( \lambda \frac{\partial T}{\partial z} \right) + \rho_w c_w \left( D \frac{\partial \theta}{\partial z} + D_I \frac{\partial I}{\partial z} - K \right) \frac{\partial T}{\partial z} \quad (5) \]

where \( W, \theta \) and \( I \) are the total water content, liquid water content, and ice content of soil, respectively; \( K = K(\theta, I) \) is the hydraulic conductivity of soil; \( T \) is the temperature of soil; \( \lambda = \lambda(\theta, I) \) is the thermal conductivity; \( D \) and \( D_I \) are the diffusivities under the constant values of \( I \) and \( \theta \), respectively; \( c_T = c_T(\theta, I) \) is the heat capacity of soil; \( LH \) is the latent heat of ice fusion.
The hydraulic parameters of Equations (4) - (5) for a frozen soil are calculated from the relationships (Gelfan, 2010) derived from van Genuchten’s (1980) formulae for an unfrozen soil.

Cumulative detention, $\text{DET}_\Sigma$, of meltwater by surface depressions is calculated by the formula assuming exponential distribution of the storage capacity:

$$\text{DET}_\Sigma = \text{DET}_0 \left[ 1 - \exp\left( \frac{-R_\Sigma}{\text{DET}_0} \right) \right]$$  \hfill (6)

where $\text{DET}_0$ is the mean value of the free storage capacity before the beginning of melt; $R_\Sigma$ is the cumulative snowmelt outflow from snowpack.

The rate of evaporation, $E$, from an unfrozen, snow-free soil is calculated as:

$$E = K_E \, d_a \, S_1$$  \hfill (7)

where $S_1$ is the relative saturation of the upper soil layer; $d_a$ is the air humidity deficit; $K_E$ is an empirical coefficient.

Runoff excess over the rectangular reaches is calculated taking into account the variability of snow water equivalent before spring melt and saturated hydraulic conductivity of soil. The same scheme was used and described in detail by Kuchment and Gelfan (2002) for simulating sub-grid variability within the finite-elements.

To simulate overland flow over the Yasenok Creek catchment, its area was schematized as a series of 18 rectangular reaches located along the main channel. Overland flow along each of the schematized reaches was described by the length-integrated kinematic wave equation written as:

$$L \frac{dh}{dt} = RL - q_l$$  \hfill (8)

$$h = \frac{m}{m + 1} \left( \frac{q_l n_l}{i_l^{0.5}} \right)^{\frac{1}{m}}$$  \hfill (9)

where $h$ is the average flow depth; $L$ is the length of the reach; $R$ is the rate of water inflow per unit length of the reach; $q_l$ is the lateral overland inflow rate per unit length of the channel; $i_l$ is the slope of the overland flow; $n_l$ is the Manning’s roughness coefficient for slope; $m$ is equal to $5/3$. 
Equations (8) and (9) were used instead of the quasi two-dimensional description of the overland flow used previously Gelfan (2010) in order to make simulations more computationally fast. Overland flow is the main mechanism of water inflow to the Yasenok Creek; subsurface contribution is negligible, so it is not considered in the model. To simulate channel flow, the one-dimensional kinematic wave equation is applied.

The procedure used for the assessment of the soil parameters of Equations (4) - (5) was described in (Gelfan, 2006) and illustrated, partially, by the measurements in the Nizhnedevitskaya WBS; only the essence of the parameterization procedure utilized in that paper is shown below.

Parameters of the van Genuchten’s formulae for the hydraulic soil properties, as well as of the formulae for the thermal conductivity and the heat capacity of soil were calculated from their dependences on the measurable soil characteristics (bulk density, porosity, field capacity, and wilting point). Saturated hydraulic conductivity $K_s$ and the coefficient $K_E$ (Eq. 7) were adjusted through calibration against the measured soil moisture profiles over 5 warm seasons, as well as measurements of soil evaporation. In addition, the coefficient $K_E$ was refined with the use of the evaporation measurements for the same seasons. The parameters of the snow model (1) - (3) were adjusted through calibration against the available snow measurements at the NWBS for 5 cold seasons. To test an ability of the model (4) - (5) to reproduce the hydrothermal regime of a frozen soil during the melt season, the measurements of the infiltration-excess overland flow, which were carried out at bounded rectangular 100 m$^2$ plots representing sections of the watershed slope, were used. Runoff excess simulated for four spring melt periods was compared satisfactorily with the observations. In this study, snow and soil parameters were taken the same as found in Gelfan (2006). The remaining 3 parameters ($DET_0$, $n_l$, and the Manning’s roughness coefficient for channel, $n_r$) were adjusted through calibration against observed discharge in the Yasenok Creek for the period 1970-1974. In Figure 5.3, the simulated hydrographs of snowmelt floods are compared with those observed. Figure 5.3 shows that the model better reproduced large floods than the small ones; however, generally the obtained results can be interpreted as satisfactory: Nash and Sutcliffe’s (1970) efficiency criterion $E_{\text{discharge}} = 0.69$. The complete list of the model parameters is presented in Table 5.1.
The next step is to transfer the model developed for the proxy-basin (Yasenok Creek) to the study basin (Sosna River) considered as an ungauged basin. The transferring procedure is described below.

**Transferring the developed model to the study basin**

In order to apply the developed model to the study basin, the latter was schematized as a series of 91 rectangular reaches located along the main channel and along the main tributaries and completely covered the catchment area. Each reach is characterized by the set of the topography parameters for simulation of overland flow by Equations (8) - (9). Hydraulic constants of soil (porosity, density, field capacity, and wilting point) were adopted from the catalog (Department of Hydrometeorological Service for the Central-Chernozem Regions, 1975) containing data of measurements at the agrometeorological stations located in the study basin.

*Figure 5.3* Comparison of observed (bold line) and simulated hydrographs of snowmelt floods in the Yasenok Creek.
Parameters of van Genuchten’s formulae were calculated from these soil constants as before. The evaporation coefficient $K_E$, as well as Manning’s coefficients of roughness were assigned the same as obtained for the proxy-basin. The other 4 parameters listed in Table 5.1 (the saturated hydraulic conductivity, $K_s$, two snow parameters, namely the coefficients of melt, $\beta$ and snow evaporation, $K_{E^*}$, and the mean value of the free storage capacity, $DET_0$) were assumed to be refined in comparison with their values obtained for the proxy-basin.

### Table 5.1 Parameters of the model of runoff generation in the Yasenok Creek basin.

<table>
<thead>
<tr>
<th>Physical meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hydraulic and thermal parameters of soil</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type of soil</td>
</tr>
<tr>
<td>Volumetric porosity</td>
<td>0.510</td>
</tr>
<tr>
<td>Bulk density, kg m$^{-3}$</td>
<td>1260</td>
</tr>
<tr>
<td>Volumetric field capacity</td>
<td>0.289</td>
</tr>
<tr>
<td>Volumetric wilting point</td>
<td>0.126</td>
</tr>
<tr>
<td>Parameters of the formulae of van Genuchten</td>
<td>0.088/1.219</td>
</tr>
<tr>
<td>$\alpha$, cm$^{-1}/n$</td>
<td>Depends on soil temperature and moisture content (Gelfan, 2006)</td>
</tr>
<tr>
<td><strong>Specific heat capacity of ground matrix, J kg$^{-1}$°C$^{-1}$</strong></td>
<td>Depends on soil temperature and moisture content (Gelfan, 2006)</td>
</tr>
<tr>
<td><strong>Thermal conductivity, J m$^{-1}$ s$^{-1}$°C$^{-1}$</strong></td>
<td>Depends on soil temperature and moisture content (Gelfan, 2006)</td>
</tr>
<tr>
<td><strong>Saturated hydraulic conductivity, m/s</strong></td>
<td>0.3×10$^{-5}$</td>
</tr>
<tr>
<td><strong>Snow parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Density of fresh fallen snow, kg m$^{-3}$</td>
<td>Depends on air temperature (Pomeroy, Hedstrem, 1998)</td>
</tr>
<tr>
<td>Coefficient of snow evaporation rate, m hPa$^{-1}$ s$^{-1}$</td>
<td>2.9×10$^{-8}$</td>
</tr>
<tr>
<td>Coefficient of melt rate, m$^4$ °C$^{-1}$ kg$^{-1}$ s$^{-1}$</td>
<td>1.8×10$^{-10}$</td>
</tr>
<tr>
<td><strong>Soil evaporation and water detention parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Coefficient of soil evaporation rate formula, m hPa$^{-1}$</td>
<td>4.4×10$^{-8}$</td>
</tr>
<tr>
<td>Mean value of the free storage capacity, m</td>
<td>0.008</td>
</tr>
<tr>
<td><strong>Roughness coefficients</strong></td>
<td></td>
</tr>
<tr>
<td>Manning’s coefficient of roughness for the river channels, s m$^{-1/3}$</td>
<td>0.04</td>
</tr>
<tr>
<td>Manning’s coefficient of roughness for the slope surface, s m$^{-1/3}$</td>
<td>0.15</td>
</tr>
</tbody>
</table>
The procedure of the parameter refinement was designed as if no streamflow
data measurements are available in the Sosna River basin. The data sets used
for the refinement were obtained from the following sources: the
hydrometeorological archive of the World Data Centre in Obninsk, Russia
(http://www.meteo.ru/data_b/); the published materials of the field experiments
of SHI in the 1960s-1970s (Vershinina et al., 1985); and catalogue (Morduhy-
Boltovsky and Zubchenko, 1971). The archive, materials, and catalogue
summarize a vast amount of information on the Don River basin (where both
the Yasenok and Sosna basins are located), its physiographic peculiarities and
hydrological behavior, regionalized values of the hydrological characteristics,
etc. The data used included: (1) maximum SWE and maximum soil freezing
depth measured at four meteorological stations before the beginning of the melt
seasons for 25 years (1952-1976); (2) snow and soil freezing depths measured
once per 10 days at the Livny meteorological station for 17 years (1965-1981);
and (3) regionalized value of the long-term mean of snowmelt runoff obtained
by interpolation from the large-scale runoff maps presented in Morduhy-
Boltovsky and Zubchenko (1971).

Initially, the snow model (Equations (1) - (3)) was calibrated against the
maximum SWE and snow depth data. Then the model of heat and moisture
transfer (Equations (4) - (5)) was calibrated against the freezing depth data;
the calibrated snow model was utilized to assign the upper boundary
conditions for Equations (4) - (5). In both cases, daily meteorological data
(air temperature and humidity, precipitation) measured at 4 meteorological
stations located within the study area were used for 30 years (1952-1981) as
the inputs into the models. The manual calibration procedure was used and
a search of the optimal parameter values was carried out within the intervals
specified a priori on the basis of the simulations in the proxy-basin.

In addition, to specify the intervals for $K_s$, the data of the field infiltration
experiments (Nazarov, 1970) made in the steppe-forest zone of European
Russia were utilized. As a result of the calibration, three aforementioned
parameters ($K_s$, $\beta$, and $K_\text{f}^*$) were refined in comparison with ones obtained
for the proxy-basin. As an example, Figure 5.4 shows the maximum values of
SWE and freezing depth measured at the Livny station versus the
corresponding ones calculated under the following values of the refined
parameters: $K_s$ (podzol) = $0.8 \times 10^{-5}$ m s$^{-1}$; $K_s$ (chernozem) = $2.9 \times 10^{-5}$ m s$^{-1}$;
$\beta = 0.1 \times 10^{-9}$ m$^4$ °C$^{-1}$ kg$^{-1}$ s$^{-1}$; $K_\text{f}^*$ = $4.5 \times 10^{-8}$ m hPa$^{-1}$ s$^{-1}$. Nash and Sutcliffe’s
(1970) efficiency of simulations of the maximum SWE and the maximum
Freezing depth equals $E_{SWE} = 0.92$ and $E_{FD} = 0.80$, respectively. The listed values of the parameters were assumed as the final ones for the Sosna River basin.

Long-term (climatic) mean snowmelt runoff volume averaged over the Don River basin (where the Sosna basin is located) is 75 mm (Morduhy-Boltovsky and Zubchenko, 1971). This information was used for calibration of the parameter $DET_0$ that is one of two key-parameters controlling runoff losses during a melt period ($K_s$ is the second one). Snowmelt runoff volume for 30 years (1952-1981) was calculated by the model and the value of $DET_0$ was adjusted under the unchanged, listed above values of other parameters. Mean 30-year snowmelt runoff of 75 mm was calculated under $DET_0 = 0.014$ m.

Figure 5.4  Calculated vs. observed values of maximum SWE (a) and soil freezing depth (b) (Sosna River basin, meteorological station Livny, 1952-1976).
The weather generator (WG) is a set of stochastic models that use existing weather records to produce long series of synthetic daily weather variables, for which statistical properties are expected to be similar to those of the actual data. The WG used includes stochastic models of daily precipitation, air temperature, and the air humidity deficit. To represent the tendency of wet or dry weather spells to persist, the widely used two-state, first-order Markov chain was applied. Daily precipitation amount was considered as a gamma distributed random variable with different parameters for the cold season and the warm season. For the dry spell, the average air humidity deficit is considered as a lognormal variable; for the wet spell, the air humidity deficit was set equal to zero. In order to simulate the daily air temperature occurrences, the method of fragments (Srikanthan and McMahon, 1985) was applied.

### Table 5.2

Stochastic weather generator parameters estimated by the long-term meteorological observations in the Sosna River basin (standard deviations of the estimations are shown in brackets).

<table>
<thead>
<tr>
<th>Model of Daily Precipitation</th>
<th>Period of simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>May-October</td>
</tr>
<tr>
<td>Probability of dry day after dry day</td>
<td>0.70 (0.08)</td>
</tr>
<tr>
<td>Probability of wet day after wet day</td>
<td>0.54 (0.07)</td>
</tr>
<tr>
<td>Mean daily precipitation amount, mm</td>
<td>4.10 (1.08)</td>
</tr>
<tr>
<td>Coefficient of variation of daily precipitation amount</td>
<td>1.31 (0.35)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model of daily air temperature</th>
<th>November-April</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean seasonal temperature, °C</td>
<td>-3.55 (0.28)</td>
</tr>
<tr>
<td>Standard deviation of mean seasonal temperature, °C</td>
<td>1.89 (0.16)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model of daily air humidity deficit</th>
<th>May-October</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean air humidity deficit for the dry spell, mb</td>
<td>7.482 (0.89)</td>
</tr>
<tr>
<td>Standard deviation of mean air humidity deficit, mb</td>
<td>2.31 (0.33)</td>
</tr>
</tbody>
</table>
Time series of daily precipitation, air temperature, and humidity deficit observed in the Livny meteorological station located at the Sosna catchment for 54 years from 1949 to 2005 (3 years containing long periods of missed data were removed) were utilized for estimating the parameters of the developed stochastic models on mean areal basis. Parameters of the precipitation model were estimated by the methods presented by Katz (1977). Parameters of the air temperature and humidity models are estimated by the method of moments. The complete list of the weather generator parameters is presented in Table 5.2.

The stochastic models were comprehensively tested through their ability to reproduce the main features of meteorological processes at the Sosna River basin. For testing, only those characteristics of the observed and simulated time series which are neither the parameters of the models nor a single-valued function of the parameters were compared. The following characteristics of the observed and simulated time series of precipitation were compared: histograms of wet and dry spell durations, autocorrelation functions of precipitation occurrence and daily precipitation series, variance of precipitation sum for 30 and 365 successive days, and distribution of maximum daily precipitation for 30 and 365 successive days. For the model of air temperature we tested how

![Figure 5.5](image_url)  
**Figure 5.5** Frequency histograms of the observed (gray columns) and calculated (striped columns) characteristics of precipitation: (a) - duration of a wet-day sequence; (b) - duration of a dry-day sequence; (c) - annual maximum of daily precipitation; (d) - daily precipitation.
the model reproduces mean value and variance of air temperature for 30 successive days and autocorrelation function of temperature time series. Histograms of mean air humidity deficit for dry spell intervals of different duration were compared for testing the model of air humidity deficit. Some results demonstrating comparison between statistical properties of the observed and simulated precipitation series are shown in Figure 5.5.

5.7 ASSESSING PROBABILITY DISTRIBUTION OF SNOWMELT FLOOD CHARACTERISTICS FOR THE PSEUDO-UNGAUGED BASIN

The dynamic-stochastic model consisting of the runoff generation model and the stochastic weather generator described in sections 5.4, and 5.5, respectively, is applied to the assessment of flood frequency in the pseudo-ungauged Sosna River basin.

Five thousand weather scenarios were Monte Carlo generated (hereafter, the weather scenario is determined as the 1-year, from May 1 to April 30, time series of daily generated meteorological data) (Gelfan, 2010). The weather scenario was used as an input into the hydrological model to simulate a single snowmelt flood hydrograph, i.e. each simulated hydrograph was in accordance with the respective weather scenario. Thus, the series of 5000 hydrographs of snowmelt flood were simulated at first by the model parameterized without using streamflow data measurements. The assessed exceedance probabilities of flood volume and peak discharge were calculated from the series and are shown in Figure 5.6a, b, respectively.

In order to estimate the accuracy of the assessments we need “to remember” the available runoff data in the Sosna River basin which we ignored so far. So, the exceedance probabilities assessed from the artificial hydrograph series are compared in Figure 5.6a, b with the probabilities of the available 61-year series of the observed flood volume (1927-1989) and the 49-year series of the observed peak discharges (1936-1989). Statistical characteristics of the simulated and the observed series are compared in Table 5.3.

Flood volume statistics are satisfactorily reproduced by the dynamic-stochastic model as it follows from Figure 5.6a and Table 5.3 and, importantly, this result was obtained without use of the streamflow data measurements. At the same time, the model overestimates both mean peak discharge and its year-to-year variation.
Inaccuracy in reproduction of the flood peak discharge statistics is caused by the errors in the roughness parameters, which were transposed from the proxy-basin as are, without any refinement through the local calibration. Let us assume now that we have a few observed values of annual peak discharge in the study basin and use these data for adjustment of Manning’s roughness.
coefficients for channel and overland flow. The coefficients were adjusted through the kinematic wave model calibration against annual flood peak discharge data for the period of 1965-1974. The adjusted coefficients turned to be equal to 0.17 s m$^{-1/3}$ for overland flow and 0.07 s m$^{-1/3}$ for channel flow.

<table>
<thead>
<tr>
<th>Table 5.3</th>
<th>Statistical characteristics of the measured and calculated flood characteristics of the Sosna River.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Flood Volume, mm</td>
<td></td>
</tr>
<tr>
<td>Observation Data (1927-1989)</td>
<td>72.5</td>
</tr>
<tr>
<td>Simulation Data</td>
<td>73.7</td>
</tr>
<tr>
<td>Flood Peak Discharge, m$^3$/s</td>
<td></td>
</tr>
<tr>
<td>Observation Data (1936-1989)</td>
<td>1727</td>
</tr>
<tr>
<td>Simulation Data</td>
<td>1991</td>
</tr>
</tbody>
</table>

Figure 5.7 Simulated vs. observed flood peak discharge at the Sosna River basin: open circles – roughness parameters are transferred from the proxy-basin; solid circles – roughness parameters are refined through calibration against 10 observed values of annual peak discharge in the Sosna River basin.
Figure 5.7 shows flood peak discharge calculated before and after the calibration procedure versus the observed discharges. The refinement of the parameter values in comparison with ones obtained for the proxy-basin has lead to removing positive bias in the peak discharge simulations (Figure 5.7). Five thousand generated weather scenarios were used once again as the inputs into the model with the refined parameters of roughness. The exceedance probabilities of flood peak discharge assessed from the simulated 5000-year hydrograph series are compared in Figure 5.8 with the corresponding probabilities obtained from 49-year observations.

Comparison of Figure 5.8 with Figure 5.6b suggests that use of even a relatively short series of the hydrograph observations to calibrate the parameters improved the simulation results regarding estimates of both mean peak discharge and its year-to-year variation. Mean flood peak discharge was found to be 1767 m$^3$s$^{-1}$, while the coefficient of variation equals 0.64. Comparing these values with those in Table 5.3, the calibration error of the mean value was reduced from 15% to 2% and error of coefficient of variation was reduced from 13% to 6%.

**Figure 5.8** Exceedance probabilities of the observed peak flood discharge (open circles) and the probabilities of the discharge simulated by the calibrated model (solid circles).
5.8 CONCLUSION

Hydrological models are widely recognized as the main tool for prediction of streamflow time series from meteorological data and are used for a huge range of scientific and engineering applications. Applicability of models is dramatically reduced when the basin in question is ungauged, i.e. there are no past streamflow observations, so the model parameters can not be adjusted through calibration against streamflow data. In this case, as well as in the cases of poorly gauged basins and virtually ungauged ones (when the available observation series are inhomogeneous because of the changes that occurred), data-based models become inapplicable. Kuchment and Gelfan (2009) argued that the physical foundation of the model, particularly, \textit{a priori} information, reflecting accumulated knowledge on runoff generation mechanisms in the basin under consideration, can compensate, to some extent, for the lack of homogeneous streamflow observation data. Kuchment and Gelfan (2009) suggested a methodology of assessing the parameters of the physically based model from both the observations in the hydrologically similar proxy-basin and the observations in the poorly gauged study basin. They then concluded that 3-4 years of streamflow observations in the poorly gauged basin are enough for obtaining stable results of hydrograph simulation by the model used in the study.

The methodological approach developed by Kuchment and Gelfan (2009) was extended here and applied for the classical problem of the engineering design – flood risk assessment in a poorly gauged basin; this is among the most crucial problems of putting the PUB-decade achievements into the practice.

The essence of the approach suggested in this paper is the following. In order to assess extreme snowmelt flood magnitude and frequency in a study basin where streamflow data are assumed to be unavailable, a data-rich small proxy-basin which is hydrologically similar to the study basin is selected. A physically based model of snowmelt flood generation was developed for the proxy-basin and then “transposed” to the study basin. Thereafter, some key parameters of the model were refined by the use of the snow and soil freezing survey data available for the study basin. At last, the deterministic hydrological model was linked to the developed stochastic weather generator and forced by the long series of the artificial meteorological data simulated by this generator. As a result of the applied
dynamic-stochastic approach, multi-year hydrograph series were calculated and the exceedance probability curves for flood volume and peak discharge were constructed for the study basin without using any streamflow data. It was found that flood volume statistics (mean and coefficient of variation) were satisfactorily reproduced by the applied approach but the corresponding statistics of flood peak discharge were overestimated. In order to remove the detected bias, we had to adjust flow roughness parameters of the model through its calibration against annual peak discharge values registered for the 10-year period of observations.

After calibration of the model, both the calculated mean value and variance of the annual flood peak discharge turned much closer to the corresponding values obtained from the long-term observations: error of mean value reduced from 15% to 2%, error of coefficient of variation reduced from 13% to 6%. One can consider this approach as a suitable alternative to the traditional engineering methods of flood risk assessment in the ungauged or poorly gauged basins.

5.9 ACKNOWLEDGEMENTS

I am sincerely grateful to Professor John Pomeroy for many helpful discussions on PUB problems and their peculiarities for cold region hydrology, and I thank him for inviting me to present this paper for the PUB monograph. I wish to express my appreciation to my teacher, Professor Lev Kuchment: some ideas of this paper are the continuation of our cooperative works.