Method of Prediction of the Stream Flows in Poorly Gauged and Ungauged Basins

Vadim Kuzmin\textsuperscript{1,3}, Inna Pivovarova\textsuperscript{2*}, Kirill Shemanaev\textsuperscript{3}, Daria Sokolova\textsuperscript{3}, Artur Batyrov\textsuperscript{3}, Ngoc Anh Tran\textsuperscript{4}, DinhKha Dang\textsuperscript{4}

\textsuperscript{1} State Hydrological Institute, 23, 2 Line, Vassilievsky Ostrov, St. Petersburg, 199053, Russia
\textsuperscript{2} Saint-Petersburg Mining University, 2, 21 Line, Vassilievsky Ostrov, St. Petersburg, 199026, Russia
\textsuperscript{3} Russian State Hydrometeorological University, Malookhtinskiy 98, St. Petersburg 195196, Russia
\textsuperscript{4} VNU University of Science, Vietnam National University (VNU), 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

* Corresponding author’s e-mail: i_pivovarova@mail.ru

ABSTRACT

The underlying principles and content of new technology for automated hydrological predictions of poorly gauged and ungauged basins were reviewed herein. Basin classification principles depending on the availability and spatial-temporal discreteness of the observations of meteorological and hydrological variables were proposed. The prediction procedure for large river systems insufficiently covered by hydrometeorological survey was outlined. The prediction methodology for Sông Sê San river basin, which is the left tributary of Mekong river, was tested. The possible options for preliminary calibration and validation of MLCM3 predictive model (Multi-Layer Conceptual Model, 3\textsuperscript{rd} generation), developed within the framework of set task, were described. The software, implementing the streamflow prediction method for ungauged and poorly gauged special basins of large rivers tributaries, was tested.

Keywords: hydrological prediction, basins, mathematic model, parametrization, ungauged Basin, hydrometeorological survey.

INTRODUCTION

Prediction of stream flows in poorly gauged and ungauged basins is an exceptional challenge for the international hydrological community. That is exactly why the international Association of Hydrological Sciences (IAHS) held a special decade in 2003-2012 dedicated to solving this problem (Decade on Predictions in Ungauged Basins (PUB)).

Despite great research achievements, there is no community consensus about a universal framework to deal with the problem of continuous streamflow simulations in ungauged basins [Hrachowitz et al. 2013; Parajka et al. 2013]. The main PUB science focus – a reduction of predictive uncertainty – is still alive across the globe: there are neither only right land cover and meteorological forcing datasets [Nasonova et al. 2011; Essou et al. 2016], nor true physically-based model working with the same efficiency under different geographical conditions [Arsenault and Brissette 2016; Goswami et al. 2006; Duan et al. 2006], nor the best regionalization technique for model parameters transfer [Razavi and Coulibaly 2013]. Despite these difficulties in hydrological theory development, the most state-of-art technique for research in PUB, still is within the framework of the coherent bundle “data – model – regionalization technique – prediction”.

There are many different types of hydrological models implemented for a variety of scientific and practical purposes, including (i) water resources assessment, (ii) flood forecasting, (iii) runoff calculations, (iv) climate impact uncertainties and assessment, etc. Nevertheless, modern hydrological model development, testing, and further implementation face the same limitations
as five decades ago [Smith et al. 2013; Paniconi and Putti 2015; Hrachowitz et al., 2013]. A lot of physically-based [Stromqvist et al. 2012; Semenova et al. 2015], conceptual [Winsemius et al. 2009; Arsenault and Brissette 2014; Razavi and Coulibaly 2016; Oudin et al. 2008; Merz and Blosch 2004], and data-driven models [Yang et al. 2008; Besaw et al. 2010] are used for continuous streamflow predictions in ungauged basins, but only some of them take into consideration the issues devoted to equifinality of model parameters and model robustness.

The streamflow prediction in ungauged basins is a policy-relevant science theme emerging as a major challenge to the hydrology. The observation network of hydrologic data is in decline around the world, yet the data is needed for more efficient water resources management, flood forecasting, and policy development based on water quality and quantity modeling [Sivapalan et al., 2003]. Advanced methods in hydrology are required to learn how to best use the hydrologic information that is available for streamflow prediction [Bandaragoda 2008].

The definition of an ungauged basin is “one with inadequate records (in terms of both data quantity and quality) of hydrological observations to enable computation of hydrological variables of interest (both water quantity and quality) at the appropriate spatial and temporal scales, and to the accuracy acceptable for practical applications” [Sivapalan et al., 2003]. Precipitation and runoff are generally the variables of interest in rainfall-runoff modeling, but if the applications are broader, they may involve erosion rates, sediment and nutrient concentrations, or stream temperature. From this perspective, all drainage basins are ‘ungauged’ to some degree and research towards understanding the application of advanced technologies to ungauged basins is applicable to all basins.

MATERIALS AND METHODS

Classification of basins by exploration degree

In general, basins are considered to be ungauged or poorly gauged, when hydrometeorological and other studies whereon do not allow the application of standard prognostic methods, techniques and technologies.

For streamflow prediction in gauged basins (in relation to meteorology) ground observations data is used as an “input” for the hydrological model; in poorly gauged basins – the output data of numerical weather models that have passed the procedure of assimilation according to available data of ground observations, from ungauged – output data of numerical weather models, which have passed the procedure of assimilation according to ground observations in neighbouring watersheds [Kuzmin 2009].

Figure 1 illustrates the possible combinations of a meteorological study (characterized by availability of ground meteorological observations points) and a hydrological study (characterized by availability of hydrometrical stations).

In the case when the number of weather station and posts at basins allows for precise evalu-

Figure 1. Explanation of «Basins exploration» term
ation of precipitation— in the context of the tasks, then basin is considered to be sufficiently explored in relation to meteorology (these are basins A, B, C and D in Figure 1), if the weather observation are not performed, then basin is considered to be ungauged, but if the weather observations are not sufficient, and require additional information, the basin is considered to be poorly gauged. The same applies for the hydrological study.

It is obvious that basins can be divided into 9 types based on their hydrological and meteorological study (Table 1).

The basins (a) with a sufficient number of observations points and (b) having the hydrometric observations stations (e.g., basins A and D in Figure 1) are most preferable in the context of forecasting, since simulation of the stream flow in such basins shall be in accordance with the standard procedure:

- take one or the other predictive model, enabling to calculate the stream flow according to precipitations and evaporation data considering the parameters of basin;
- calibrate the model – identify its parameters that provide the best (in the context of objective function) matching of the simulated and actual flow;
- perform verification of parameters to verify the accuracy of stream flow simulation using identified parameters of independent material.
- if the verification confirmed the effectiveness of calibration and reliability of the identified parameters then the prediction of stream flow according to continuously incoming data with ground observations of precipitations shall be performed.

Generally, this is the approach to be applied for small and moderate-sized basins located in flood-prone areas of technically developed countries. It should be noted that the basins of Russian Federation are considered to be small, provided that their area is less than 2000 km², moderate-sized – ranging from 2,000 to 50,000 km², and large – with an area of over 50,000 km² [Doganovsky 2012]. Principle of geographic and/or climatic zoning shall be used as criterion.

Other countries have adopted different threshold values, generally determined in accordance with specific features of the tasks to be performed [Pivovarova and Kuzmin 2017]. In the predictive context, it makes sense to focus on the key predicting elements of the stream flow, but not on the basin area and its belonging to one or several geographical and/or climatic zones, for example:

- small-sized basins – zero “memory” (response time) basins: the initial flow stream in the main-stream station does no matter; prediction of stream flow according to actual precipitations has extremely slight (1-2 hours), zero or negative lead time; therefore, for realistic predictions with the lead time up to several hours, the predicted precipitations survey data produced by weather radar shall be used, and for the forecasts with greater lead time or flood risk, the out-coming data of synoptic models shall be used;
- moderate-sized basins – minor nonzero “memory” basins: to predict the stream flow in main-stream station, the stream flow at initial time and the actual precipitation shall be known; stream flow prediction lead time usually does not exceed 1-2 days. In case the lead time increases (even with slight discrepancies!) the data of ground-based surveys shall be replaced with the data produced by the synoptic models;
- large basins – significant response time basins: prediction is usually based on calculating the distribution of water masses from private basins; stream flow prediction lead-time may exceed 10 days.

Table 1. Classification of basins by hydro-meteorological exploration degree (in the context of availability, insufficient or absence of ground-based hydrometeorological studies)

<table>
<thead>
<tr>
<th>Exploration degree</th>
<th>In hydrological relation</th>
<th>gauged</th>
<th>poorly gauged</th>
<th>ungauged</th>
</tr>
</thead>
<tbody>
<tr>
<td>In meteorological relation</td>
<td>The data of ground-based observations</td>
<td>yes</td>
<td>not sufficient</td>
<td>no</td>
</tr>
<tr>
<td>gauged</td>
<td>yes</td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>poorly gauged</td>
<td>not sufficient</td>
<td>IV</td>
<td>V</td>
<td>VI</td>
</tr>
<tr>
<td>ungauged</td>
<td>no</td>
<td>VII</td>
<td>VIII</td>
<td>IX</td>
</tr>
</tbody>
</table>
Prediction technique of large river flows, hydrometeorological surveys data is not sufficient

In the course of implementation of a joint Russian-Vietnamese research project “Development of methodological bases and technologies Water resources management of river systems insufficiently covered by hydrometeorological surveys (example of the Mekong river basin)” the method of river flow prediction, insufficiently covered by hydrometeorological observations was developed. This method is focused on stream flow prediction not only for the Mekong river and any other major rivers, and implying that:

- large ones shall be understood to mean a basin with significant response time without any quantitative restrictions on the area;
- large basin consists of private basins that differ in terms of meteorological and hydrological exploration (Table 1);
- at least main stream station flow of large basin is known;
- access to the output data (precipitations) of weather models is available [WRF (Weather Research and Forecasting Model) / www.wrf-model.org, HIRLAM (High Resolution Limited Area Model) / http://hirlam.org/ etc.].

Prediction procedure, depending on basin classification

Subject to fulfillment of the above-mentioned conditions, the stream flow shall be predicted in the following order:
- large basin under review shall be divided into private basins, the dimensions of which generally depend on the spatial resolution of weather models ‘outcome’, computing resources, software, geoinformation basis, etc. Particularly, if any GIS technology is used for delineation of private basins, then SW and geoinformation basis will determine limitation of size from below [Pivovarova 2016]. The abundance of small private basins, the precipitations of which can only be assessed using weather models, and stream flow is unknown, increases the uncertainty of predicted stream flow in main-stream station; therefore, moderate size of a private basin is recommended to be taken as one of the objective variables (based on minimizing the objective function in main-stream station);
- allocated private basins shall be formally classified in accordance with Table 1;
- the prediction of stream flows from private basins shall be performed according to conceptual hydrological models developed by Russian State hydrometeorological University MLCM3 (Multi-Layer Conceptual Model3) [Sokolova et al. 2018]. However, it should be noted that any hydrological model acceptable for the predictor can be used. Exploration class influences the choice of input data for the prediction, determination of calibration parameters and model validation;
- water masses displacement along the tributaries to each main-stream station shall be calculated. For this purpose, both simple methods (such as kinematic wave model, Muskingum-Cung method and their analogues, which parameters shall be optimized by minimizing the objective function MSOF in main-stream station), and more complex software can be applied (such as HEC - RAS [http://www.hec.usace.army.mil/software/hec-ras/hecras-demo.html] etc. [https://www.fema.gov/hydrologic-models-meeting-minimum-requirement-national-flood-insurance-program] considering the transformation of stream flow on certain sections of principal river and its tributaries.

It should be noted that taking into account the transformation of the flood wave is extremely important in the lower part of the Mekong river, and especially in its Delta, where a significant effect of variable backwater from the South China sea is present [Annual Mekong Flood Report 2012; Manual for Training Trainers in Integrated Water Resources Management in the Mekong Basin 2014]. It can also be important for large rivers, where the stream flow of major tributaries is relatively large, which may affect the displacement of water masses along the main channel (for example, the Oka, the Rhine, the Danube, the Amazon, etc.). In the meantime, if precipitation occurs locally, and the flood formed in a limited area at considerable distance from river mouth, simple computational methods of water masses displacement, not considering the spreading of the waves can be sufficient, which significantly simplifies the prediction procedure.
Prediction of stream flow in the presence of hydrometrical observations for certain private basins

If the stream flow of certain tributaries of these large rivers is known (for example, the stream flow of Emchu, Mun, the Tonle Sap Jiechi and Sông Sê San, which are the main tributaries of the Mekong), it is advisable to use a more flexible calibration procedure of hydrological models.

Regardless of the meteorological study degree of such tributaries basins (relating to classes IV or VII) and, hence, the procedures of predicting their stream flows, the available hydrological data can be used to identify the adjustment factors for precipitation and evaporation. If the basin is homogeneous in terms of topography and vegetation, then their distribution across these large basin can be determined according to the obtained adjustment factors (e.g. by the means of map development).

If the basin is not homogeneous, such a procedure can also be applied; however, it should be noted that the uncertainty introduced by inhomogeneity of certain basin characteristics can be significantly more substantial than the gain of objective function reduction. Therefore, the use of this procedure should be combined with the analysis of model parameters stability and robustness of response surface of the models applied.

RESULTS AND DISCUSSION

The prediction method was tested for the Sông Sê San river basin, which is the left tributary of the Mekong river, and flowing through Vietnam (Kon tum province and Gia Lai) and Cambodia (Ratanakiri and Stung Treng provinces). The basin area of Sông Sê San river is 17 000 km², of which 11 000 km² is in Vietnam.

The data from hydrological points- Đác Môt, Kon Tum and Kon Plong was used for the forecasting experiments. These three stations form a certain spatial triangle inside the river basin and hydrological posts in peaks (Table 2). River basin belongs to the poorly explored class (V).

Prediction was made for several hydrological models SAC-SMA and HBV MLCM3:
• the well-known Sacramento Soil Moisture Accounting Model (SAC-SMA) [Burnash et al.1973, Reed et al. 2005] has successfully been implemented in the United States for flash flood forecasting for four decades. This model is very popular among hydrologists. The original version of the Sacramento model includes 16 optimized parameters, the values of which can be estimated by observing the soil moisture and entering data on soil type; however, in most cases, it is possible to replace the parameters with 5-7 constants without any quality loss modeling;
• the HydrologiskaByrånsVattenbalansavdelning model (HBV) was developed by Bergström at the Swedish Meteorological and Hydrological Institute for modeling and analysis of river stream flow. The HBV model is a conceptual model that converts precipitation, temperature and atmospheric potential natural water loss into either snow melting, or into runoff or inflow into the reservoir. The HBV model can be considered as a model with half-distributed parameters: the basin area is divided into private basins, and the altitudinal zoning method is also applied [Johansson 1997]
• MLCM3 hydrological model is a model of “precipitation–stream flow” type with a flexible structure and a high level of conceptualization. From a technical point of view, while performing calibration of the model, it can be easily narrowed down to relatively simple models of basin or channel flow, and to more complex hydrogeological models, taking into account the hydraulic properties of basin soils [Sokolova and Kuzmin 2017].

The precipitation data obtained at the “output” of the weather model WRF (Weather Research and Forecasting Model) were used as input parameters for all models. The timeliness of issued predictions is 5–7 days. The processed data on stream flow encompassed the period from 01.01.2013 to 30.12.2016. MLCM3 calibration was performed by the SLS Stepwise Line Search

<table>
<thead>
<tr>
<th>Table 2. Stations g/m used for observations on Sông Sê San river</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station</td>
</tr>
<tr>
<td>Đác Môt</td>
</tr>
<tr>
<td>Kon Tum</td>
</tr>
<tr>
<td>Kon Plong</td>
</tr>
</tbody>
</table>
– step-wise linear search is a type of automatic optimization) and Nelder–Mid (simplex method, which is based on using nonlinear optimization techniques applied for double-differentiated and unimodal tasks). Since MLCM3 software tools allow the user to choosing the type of objective function for calibration and validation, three options were studied: mean-root square error, a criterion of S/σ and multiscale objective function MSOF. The criterion of Nash-Sutcliffe was used for efficiency evaluation.

The results are provided in Table 3 and illustrated in Figures 2 and 3. Their analysis allows for drawing the following conclusions:

1. The optimal value of objective function MSOF was shown by MLCM3 model. Apparently, this method is the most efficient in predicting stream flow from the moderate-sized basins (ranging from 1000 to 10 000 km$^2$), since in this case the dependence of stream flow prediction accuracy on spatiotemporal uncertainties of simulated precipitation is decreased.

Table 3. Results of numerical experiments (effectiveness study for different models)

<table>
<thead>
<tr>
<th>No.</th>
<th>Stream flow Code</th>
<th>Optimal value of objective function MSOF (during the period of validation, best-case scenario)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MLCM3</td>
</tr>
<tr>
<td>1</td>
<td>Kon Tum</td>
<td>13.98</td>
</tr>
<tr>
<td>2</td>
<td>Đắc Môt</td>
<td>7.50</td>
</tr>
<tr>
<td>3</td>
<td>Kon Plong</td>
<td>4.81</td>
</tr>
</tbody>
</table>

Figure 2. Example of numerical tests on Đắc Môt using Nelder-Mid method (the test result)

Figure 3. Example of numerical tests on Đắc Môt using SLS (test result)
2. The calibration using Nelder-Mid method showed the best result compared to SLS. The SLS method is more suitable for the cases when the a priori (initial) point in quasilocall optimization is set correctly, and the basin area is well-covered by the hydro-meteorological observations.

3. The use of weather hydrodynamic model “output” as an “input” for hydrological model allowed to improve the prediction timeliness from 1–3 to 5–7 days.

CONCLUSION

Improving the timeliness of short-term predictions of river stream flow on retention of acceptable accuracy is one of the most important and difficult challenges in hydrological science. Under the conditions of lack of basin hydrometeorological study, the task becomes much more complicated and the statistical prediction methods lose their effectiveness; therefore, the only method of qualitative hydrological prediction is mathematical simulation. Prediction effectiveness depends on the correct compilation of input data – identification of model parameters – mathematical algorithm and software.

The article shows the development stage of modern technologies for water resources management of large river basins, consisting of private basins with varying degrees of meteorological and hydrological explorarion. The core task of the project is the development of scientific-technical base to ensure the on-line monitoring of long-term and short-term hydrological risks in the absence or insufficiency of surface meteorological and hydrological observations.

Acknowledgements

The presented work is funded by the Ministry of Education and Science of the Russian Federation (Project ID No. RFMEFI58316X0059; the code of the application 2016-14-585-0005-002) within the framework of the Federal Target Program.

REFERENCES


12. High Resolution Limited Area Model [Electronic resource], http://hirlam.org/


16. Manual for Training Trainers in Integrated Wa-


