Hydrological predictability investigation of global data sets for high-latitude river basins

YEUGENIY M. GUSEV, OLGA N. NASONOVA, LARISA Y. DZHOGAN & YEUGENIY E. KOVALEV

Institute of Water Problems, Russian Academy of Sciences, Gubkina St.3, 119991 Moscow, Russia

gusev@aqua.laser.ru

Abstract The aim of the work is to investigate the hydrological predictability of global data sets for the northern river basins using the land surface model SWAP. The input data were taken from the global 1-degree data sets provided within the framework of the Second Global Soil Wetness Project. Application of the global data sets (without calibration of model parameters) for the simulation of the Mezen River runoff, a high-latitude basin, has shown poor results. To improve the results, a stochastic optimization technique was developed for model calibration using the streamflow observations for the period 1986–1990. The period 1991–1995 was used for model validation. The Nash-Sutcliffe efficiency of daily streamflow simulation was 0.80 and 0.82 for the calibration and validation periods, respectively. Analysis of the results showed that global 1-degree data sets may be applied with appropriate calibration for streamflow simulation and prediction in high latitudes.

Key words river runoff; land surface model SWAP; Mezen River; parameter optimization

INTRODUCTION

Currently, high-latitude regions characterized by a long and severe cold season are receiving more and more attention from the hydrometeorological modelling community, because these regions are among the most sensitive to natural and anthropogenic effects and it is necessary to predict the consequences of such effects. At the same time, northern regions have a poor coverage of data that can be used for forcing and estimation of land surface parameters in model simulations. One of the possible ways to provide a model with input data is the application of existing global data sets, which contain meteorological data, land-use information, and soil and vegetation characteristics. Nowadays there are a lot of global data sets, which differ in spatial and temporal resolution, as well as in accuracy and reliability. Nevertheless, this type of information is quite attractive for modellers as it saves them from the quite difficult procedure of model input data preparation. The aim of this study is to assess the possibility of application of global 1-degree data sets for runoff simulation, on a northern river basin using a land surface model (LSM). Herein, the SWAP model (Soil Water–Atmosphere –Plants), developed by the authors during the last 10 years, is used as the LSM and the Mezen River basin, located in the north of the European part of Russia, is selected as the hydrological object.

SWAP MODEL

The land surface model SWAP represents a physically-based model describing the processes of heat and water exchange within a soil–vegetation/snow cover–atmosphere system (SVAS). The model can be applied both for point (or grid cell) simulations of vertical fluxes and state variables of SVAS in atmospheric science applications, and for simulating streamflow at different scales—from small catchments to continental-scale river basins. In the case of a small river basin, a kinematic wave equation is used to simulate runoff at the basin outlet. In the case of a large river basin, the basin area is divided into a number of computational grid cells connected by a river network. Runoff is modelled for each grid cell and then transformed by a river routing model to simulate streamflow at the river basin outlet. Such a transformation may be performed with different approaches. Herein, the simple linear model of water balance formation in the river network is used (Kanae *et al.*, 1995; Oki *et al.*, 1997).

During the last 10 years, different versions of the SWAP model were validated against observations, including characteristics, which relate both to the energy balance or thermal regime of SVAS and to the hydrological cycle or water regime of SVAS. The validation was performed for "point" experimental sites and for catchments and river basins of different area (varying from 10^{-1} to 10^{5} km^2) on a long-term basis and under different natural conditions. The results of validation have shown that SWAP is able to reproduce annual and interannual dynamics of these characteristics fairly well.

The main elements of SWAP and the results of its validation are described in a series of publications (Gusev & Nasonova, 1998, 2002, 2003).

HYDROLOGICAL OBJECT AND ITS SCHEMATIZATION

The Mezen River basin with a 78 000 km^2 drainage area is located in the northeastern part of European Russia, near the taiga and tundra border (64.5–66°N, 44–50°E) (Fig. 1(a)). The Mezen is the longest of the rivers discharging into the White Sea. Its length from head to mouth is 966 km. Its average annual discharge is 27.9 km^3/year. Like most northern rivers, the Mezen is a typical river of the plains. Its basin is 80% covered with boreal forest (with a predominance of coniferous species). Soils in the basin are mostly podsol and gley-podsol. Climate is characterized by a short (~3 month) cool summer and long (6–7 month) cold winter with a stable snow cover.

The Mezen River basin was represented for modelling purposes by ten 1° × 1° computational grid cells (from the head of the river to the Malonisogorskaya gauging station (Fig. 1(b)).

INPUT DATA

Input data (both meteorological forcing data and the land surface parameters) were taken from the global 1-degree data sets, produced within the framework of the international Second Global Wetness Project (GSWP-2) (Dirmeyer *et al.*, 2002; Zhao & Dirmeyer, 2003). The data cover all the land surface of the Earth with the exception of the Antarctic continent. Near-surface meteorological data, provided at a 3-hour time

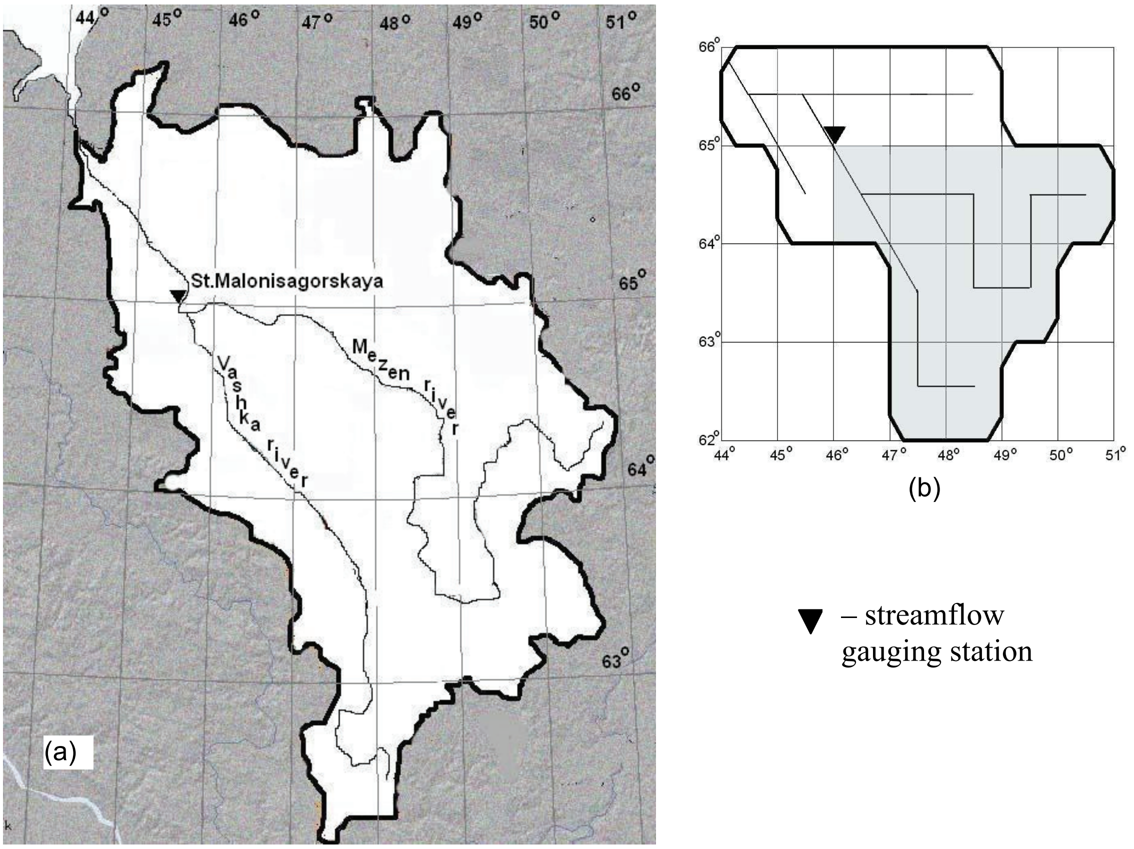

Fig. 1 The Mezen River basin (a), and its schematization for runoff simulations (b).

step for the period from 1 July 1982 to 31 December 1995, include incoming SRB (Surface Radiation Budget) shortwave and longwave radiation; CRU (Climate Research Unit) air temperature and humidity; atmospheric precipitation representing hybrid of NCEP/DOE (National Center for Environmental Prediction/Department of Energy) re-analysis estimates of precipitation with GPCC (Global Precipitation Climatology Center) and GPCP (Global Precipitation Climatology Project) gridded observational precipitation data sets; NCEP/DOE air pressure and wind speed. The land surface parameters (soil and vegetation characteristics) taken from global data sets were previously analysed and checked for consistency (they should be in good agreement with the other parameters and with the nature). In so doing, some corrections were performed. In addition, several SWAP model specific parameters were derived. We will call the values of the land surface parameters obtained in such a manner *a priori* parameters.

SIMULATION OF STREAMFLOW USING *A PRIORI* INFORMATION

The data sets described were applied for modelling the Mezen River runoff at 3-hour time steps for 1982–1995 using the SWAP model. The period from July 1982 to December 1985 was recommended as the spin-up period in the simulation, while the 10-year period 1986–1995 was used as the baseline period (Dirmeyer *et al.*, 2002).

The modelled daily river runoff at the Malonisogorskaya station for 1986–1995 was compared to appropriate observations. The agreement between simulations and observations was evaluated using three statistical criteria: the bias, *Bias*, the coefficient of correlation between the simulated and measured values, *Corr*, and the Nash- Sutcliffe efficiency, *Eff* (Nash & Sutcliffe, 1970). When developing the runoff forecast techniques, it is considered that the accuracy is "good" if $Eff \geq 0.75$ and "satisfactory" if $0.36 < Eff < 0.75$. (Appolov, 1974). For the modelled hydrograph, $Bias = 0.39$ mm/day (or 141 mm/year, i.e. 40% of the measured annual runoff), $Corr = 0.94$ and $Eff = -0.49$, i.e. the quality of simulation is considered as unsatisfactory according to the evaluation scale.

Thus, the analysis of the obtained results has shown that application of *a priori* parameters (from global data sets) for the Mezen River basin, which is typical of the northern rivers, leads to poor streamflow simulation. Therefore, it is reasonable to try to improve this situation using an approach to parameter estimation common within the hydrological modelling community, namely, calibration of a model to historical observations by tuning model parameters (Duan *et al*., 2006). Since river runoff is the only hydrological characteristic measured within the Mezen basin, it is used for such a calibration.

MODEL PARAMETERS OPTIMIZATION (CALIBRATION) PROCEDURE

We developed an automatic calibration procedure, based on the direct search of global optima (corresponding to minimum of root mean square deviation between simulations and observations) by means of stochastic techniques. In so doing we followed some ideas concerning statistical optimization techniques developed, in particular, in Solomatine *et al*. (1999) and Bastidas *et al*. (1999).

First, it is necessary to select model parameters for calibration. Evidently, those parameters whose modification most influences the model outputs (runoff in our case), should be calibrated. Seven land surface parameters were selected: soil hydraulic conductivity at saturation K_0, depth to the upper impermeable soil layer h_0, depth of soil root layer h_r, snow-free albedo of vegetation in the warm period alb_{sum}, snow albedo alb_{sn}, effective Manning roughness coefficient n, effective water speed in river channels u_e (other parameters were taken from the global GSWP-2 data sets). In addition, because meteorological information in global data sets may suffer from systematic errors (Zhao & Dirmeyer, 2003), we also used four scaling factors k_{lp}, k_{sp}, k_{sw} and k_{lw} for the most important meteorological characteristics: rainfall, snowfall, shortwave and longwave radiation, respectively. The values of alb_{sn}, n, u_e, k_{sw}, k_{lw}, k_{lp} and k_{sp} were assumed to be the same for all the basin cells, while the values of K_0, h_0, h_r and alb_{sum} varied from cell to cell resulting in 47 (!) parameters which require calibration. To reduce the number of calibrated parameters and increase their stability, instead of K_0, h_r and alb_{sum} for each grid cell, we calibrated their adjustment factors k_{K0}, k_{hr} and k_{albsum}, which are general for the basin; that allowed us to reduce the number of corresponding parameters from 30 to 3. In addition, we set $h_0 = k_{h0}h_r$ for each cell, where k_{h0} is also an adjustment factor taken as constant for the entire basin (that allowed us to reduce the number of calibrated parameters from 10 to 1). As a result, the total number of calibrated parameters was reduced to 11: seven for the land surface: k_{K0}, k_{hr}, k_{albsum}, k_{h0}, alb_{sn}, n, u_e and four for the forcing data: k_{sw}, k_{lw}, k_{lp} and k_{sp}.

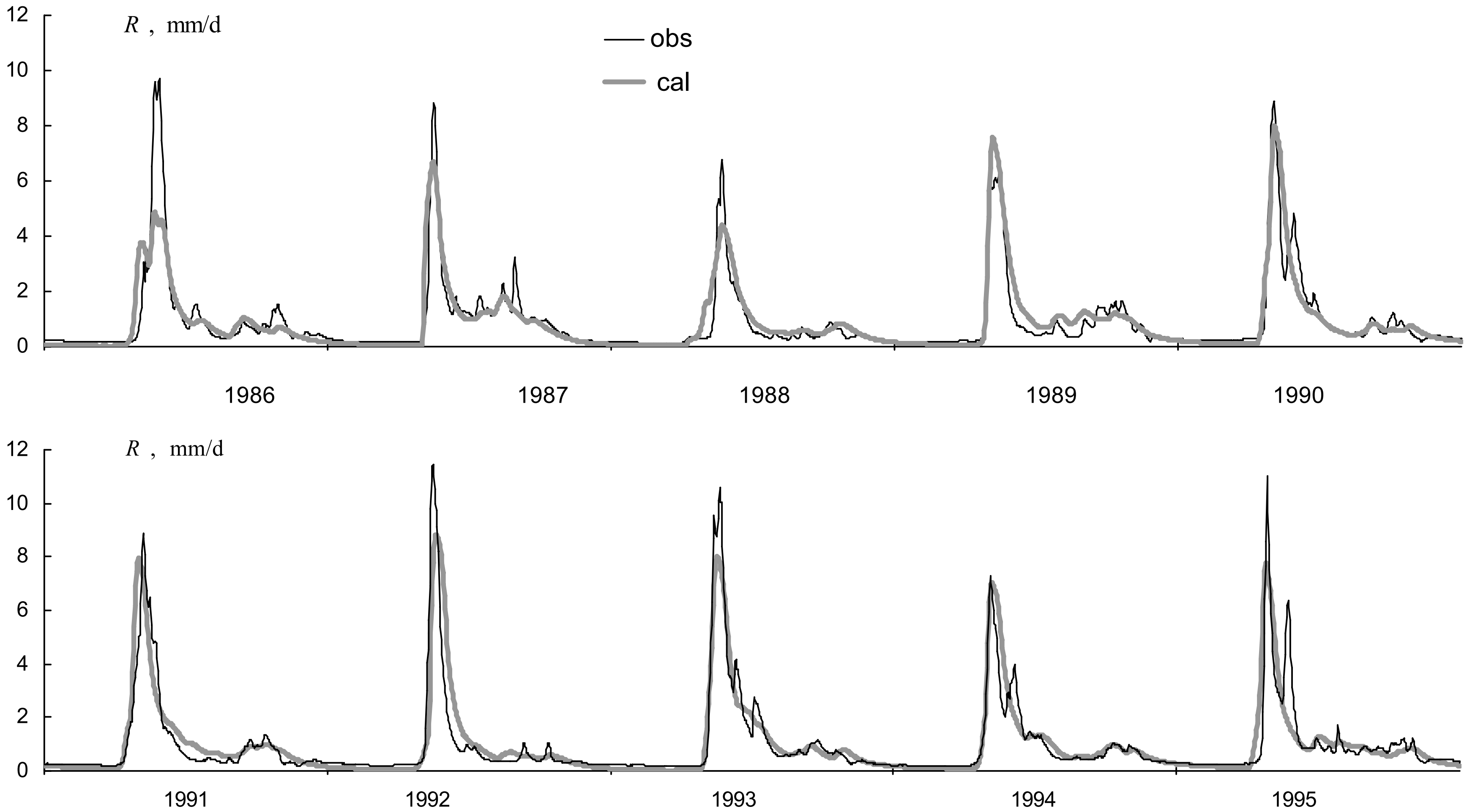

Fig. 2 Comparison of observed (obs) and calculated (cal) with optimal parameters daily runoff for Mezen River at Malonisogorskaya station.

SIMULATION OF STREAMFLOW USING CALIBRATED PARAMETERS

Model calibration was performed for 1986–1990 using daily streamflow measured at the Malonisogorskaya station (Fig. 1). After that, the optimized model parameter values were used for model validation against daily streamflow observed in 1991–1995. The observed streamflow values were taken from the GRDC (Global Runoff Data Centre) database.

The comparison of measured and simulated (with optimal model parameter values) daily streamflow showed a good agreement both for the calibration and validation periods (Fig. 2). Thus, *Bias* = –2 mm/year, *Corr* = 0.89, *Eff* = 0.80 for the calibration period (Fig. 2, upper panel) and *Bias* = 0 mm/year, *Corr* = 0.91, *Eff* = 0.82 for the validation period (Fig. 2, lower panel). As seen, the quality of the river flow modelling can be considered as "good", at least according to the runoff forecast evaluation scale mentioned above.

CONCLUSIONS

1. Application of global 1-degree data sets (without calibration) for simulating runoff over a typical northern river basin using the land surface model SWAP results in poor agreement between simulated and observed hydrographs.
2. The automatic calibration technique developed allows effective optimization of a large number of model parameters.
3. Optimization of model parameters taken from the global data sets significantly improves the quality of streamflow simulation (from "unsatisfactory" to "good").

Thus, the main conclusion from this study is that global 1-degree data sets of near-surface meteorology and land surface parameters along with a land surface model can be used, after appropriate calibration, for streamflow simulations in the northern regions.

Acknowledgements This work was supported by the Russian Foundation for Basic Research (Grant 05-05-64161). We acknowledge the GSWP-2 experiment organizers for providing us with the data to run the model and the GRDC (Global Runoff Data Centre, D-56068 Koblenz, Germany) for daily streamflow observations.

REFERENCES

Appolov, B. A., Kalinin, G. P. & Komarov, V. D. (1974) *Hydrological Forecasts Course*. Gidrometeoizdat, Leningrad, USSR.

Bastidas, L. A., Gupta, H. V., Sorooshian, S., Shuttleworth, W. J. & Yang, Z. L. (1999) Sensitivity analysis of a land surface scheme using multi-criteria methods. *J. Geophys. Res.* **104**(D16), 19481–19490.

Dirmeyer, P., Gao, X. & Oki, T. (2002) The second global soil wetness project. Science and implementation plan. *IGPO Publication Series 37.*

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T. & Wood, E. F. (2006) Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *J. Hydrol.* **320**, 3–17.

Gusev, Ye. M. & Nasonova, O. N. (1998) The land surface parameterization scheme SWAP: description and partial validation. *Global Plan. Change*. **19**(1–4), 63–86.

Gusev, Ye. M. & Nasonova, O. N. (2002) The simulation of heat and water exchange at the land-atmosphere interface for the boreal grassland by the land-surface model SWAP. *Hydrol. Processes* **16**(10), 1893–1919.

Gusev, Ye M. & Nasonova, O. N. (2003) The simulation of heat and water exchange in the boreal spruce forest by the land-surface model SWAP. *J. Hydrol.* **280**(1-4), 162–191.

Kanae, S., Nishio, K., Oki, T. & Musiake, K. (1995) Hydrograph estimations by flow routing modeling from AGCM output in major basins of the world. *Ann. J. Hydraulic Engng* **39**, 97–102.

Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models: 1. A discussion of principles. *J. Hydrol*. **10**(3), 282–290.

Oki, T., Nishimura, T. & Dirmeyer, P. (1997) Validating estimates of land surface parameterizations by annual discharge using Total Runoff Integrating Pathways. *J. Japan Soc. Hydrol. Water Resour*. **10**, 416–425.

Solomatine, D. P., Dibike, Y. B. & Kukuric, N. (1999) Automatic calibration of groundwater models using global optimization techniques. *Hydrol. Sci. J.* **44**(6), 879–894.

Zhao, M. & Dirmeyer, P. A. (2003) Production and analysis of GSWP-2 near-surface meteorology data sets. *COLA Technical Report 159*.