Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment–MOPEX. IAHS Publ. 307, 2006.

Application of global 1-degree data sets to simulate runoff from MOPEX experimental river basins

OLGA N. NASONOVA & YEUGENIY M. GUSEV

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

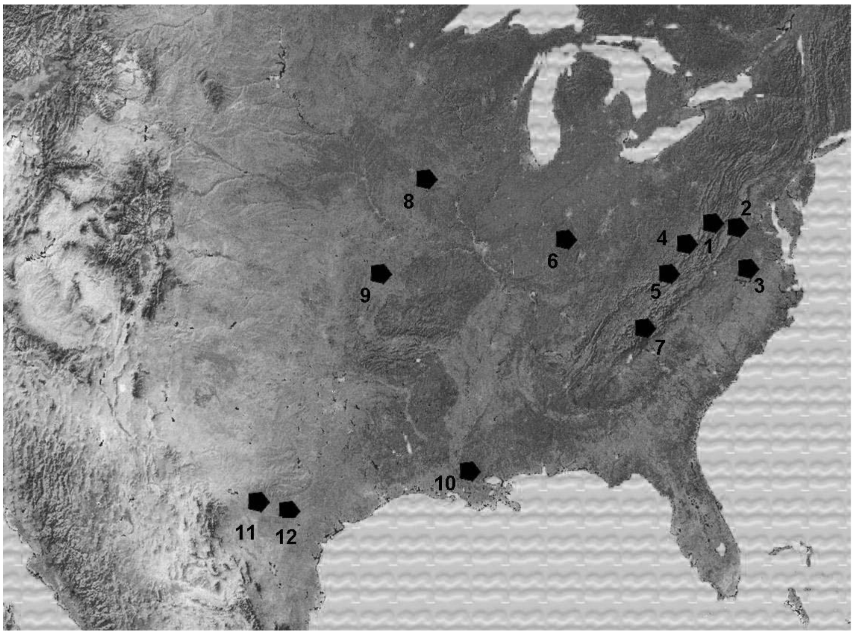
nasonova@aqua.laser.ru

Abstract Global 1-degree data sets provided within the framework of the Second Global Wetness Project (GSWP-2) were applied for the simulation of river flow from 12 MOPEX (Model Parameter Estimation Experiment) experimental river basins (with an area of the order of 10^3 km^2) to reveal the applicability of global forcing data and land surface parameters for regional runoff predictions. The simulations were performed at 3-h time steps for a 10-year period (1986–1995) using the SWAP land surface model (Soil Water–Atmosphere–Plants). The results of simulations were compared with the observed streamflow and with analogous simulations based on the regional data set. The comparison allowed us to reveal the influence of uncertainties in the forcing data and the land surface parameters on runoff prediction.

Key words global data; GSWP-2; land surface modelling; MOPEX; river runoff; uncertainty analysis

INTRODUCTION

At the moment there are a lot of global data sets containing hydrometeorological data, land use information, and soil and vegetation characteristics with 1-degree spatial resolution. Global data sets are widely used for global and macroscale runoff simulations (e.g. Oki *et al*., 1995; Jayawardena & Mahanama, 2001), while their resolution is supposed to be crude for regional and local applications. In the latter case the downscaling procedure is usually applied to span the gap between large-scale forcing data generated by general circulation models (GCMs) and regional or local hydrological simulations (e.g. Salathé, 2003). Downscaling techniques are generally divided into statistical and dynamical, and spatial and temporal ones; for a review see e.g. Wilby & Wigley (1997). Each of them has its own advantages and disadvantages and is more or less costly. Also, different techniques result in different spatial/temporal patterns of downscaled data that may necessitate the investigation of a number of different downscaling techniques before a suitable methodology is identified. All these circumstances have led us to the following questions. Is it possible to simulate runoff at regional scale (for the basins, comparable in size with a 1-degree grid cell) without using the spatial downscaling technique? What will be the reliability of such simulations? Uncertainties in what type of data will be crucial for the final results? The present work is aimed at the investigation of these issues.



Basin	Area (km^2)
1	3810
2	2116
3	4134
4	2372
5	1020
6	4421
7	2448
8	1484
9	3015
10	3315
11	3406
12	2170

Fig. 1 Location of the 12 MOPEX river basins.

DESCRIPTION OF STUDY SITES AND DATA SETS

Twelve river basins (with an area ranging between 1020 and 4421 km^2), which were selected within the framework of the Second MOPEX (Model Parameter Estimation Experiment) Workshop (Duan *et al.*, 2006), were used for the given investigation (Fig. 1). The basins are located within the southeastern part of the United States and characterized by a wide range of hydrological and climatic conditions varying from desert conditions to very wet ones. The dominant vegetation types represent deciduous broadleaf (basins: 1, 2, 4, 5, 9), evergreen needleleaf (basin 10) and mixed (basins: 3, 7) forests, as well as cropland (6, 8) and grassland (basins: 11, 12).

Regional and global data sets were used for model simulations. Both data sets include near-surface meteorological (forcing) data and land surface parameters (soil and vegetation characteristics).

The regional data set was described in Gusev & Nasonova (this issue). The forcing data were provided by the Second MOPEX Workshop organizers. The values of model parameters were partly provided by the organizers, and partly derived by the authors (see the set of model parameters referred to as "CAL1-MOP" in Gusev & Nasonova (this issue)).

The applied global data represent 1-degree data sets (including forcing data and land surface parameters) produced within the framework of the Second Global Wetness Project (GSWP-2) (Dirmeyer *et al.*, 2002; Zhao & Dirmeyer, 2003). The global forcing data are based on re-analyses and gridded observational data used in ISLSCP (the International Satellite Land-Surface Climatology Project) Initiative II. ISLSCP-II global database includes two versions of meteorological data representing the products of NCEP/DOE (the National Center for Environmental Prediction—Department of Energy) and ERA-40 (ECMWF Re-analysis-40, European Centre for Medium-Range Weather Forecasts) re-analysis and meteorological data sets based on observations. The former contain systematic errors, the latter are characterized by low temporal resolution (as a rule, one month) and cannot be used directly for modelling. That is why corrections to the systematic biases in the re-analysis fields were made by hybridization of the 3-hourly re-analysis data with global observation-based gridded data sets within the framework of GSWP-2 project. As a result, several alternative data sets with different interpretations of global meteorological fields were suggested (http://www.iges.org/gswp/):

B0 Baseline data set (*SRB downward shortwave and longwave radiation*, CRU near-surface air temperature and humidity, ECOR surface pressure, NCEP wind speed, *hybrid NCEP/DOE precipitation* with GPCC and GPCP precipitation data sets).
M1 All *NCEP/DOE forcing data* (no hybridization with observational data).
M2 All *ECMWF forcing data* (no hybridization with observational data).

All the other global forcing data sets are the same as B0, but either precipitation or radiation is replaced by an alternative one:

PE *Hybrid ERA-40 precipitation* (instead of NCEP/DOE).
P1 *ERA-40 precipitation without hybridization* (pure re-analysis).
P2 *Hybrid NCEP/DOE precipitation, but without the relaxation to GPCP* (satellite-estimated) precipitation.
P3 *Hybrid precipitation as in P2, but without correction for gauge undercatch*.
P4 *NCEP/DOE precipitation without hybridization* (pure re-analysis).
R1 *NCEP/DOE downward shortwave and longwave radiation* (instead of SRB).
R2 *ERA-40 downward shortwave and longwave radiation* (instead of SRB).
R3 *ISCCP downward shortwave and longwave radiation* (instead of SRB).

Following the GSWP-2 strategy (Dirmeyer *et al*., 2002), we used the described data sets to perform different sensitivity experiments.

To reveal the influence of model parameters on the runoff simulations, in addition to MOPEX and GSWP-2 parameter data sets, we produced two alternative global data sets of soil parameters on the basis of their relationships with the clay and sand contents in a soil. The relationships were derived from generalized tables published in Clapp & Hornberger (1978) and Dunne & Willmott (1996) for one soil data set and in Cosby *et al*. (1984) for another.

RESULTS

The described data sets were applied for the simulations of river runoff for the 12 basins at 3-hour time steps for a 10-year period (1986–1995) using the SWAP land surface model (Soil Water–Atmosphere–Plants) (Gusev & Nasonova, 2003). Some results are shown in Fig. 2, where annual observed runoff from each basin and for each year is compared with runoff simulated using the MOPEX regional data and 11 global forcing data sets. Different statistics for validation of the simulated annual runoff against observations (the ratio between modelled and observed runoff (*rat*), the Nash-Sutcliffe efficiency (*eff*) and the coefficient of correlation (*corr*)) are also given in the panels. Analysis of different chains of experiments allows us to reveal how uncertainties in forcing data influence the final results. Thus, following the sequences

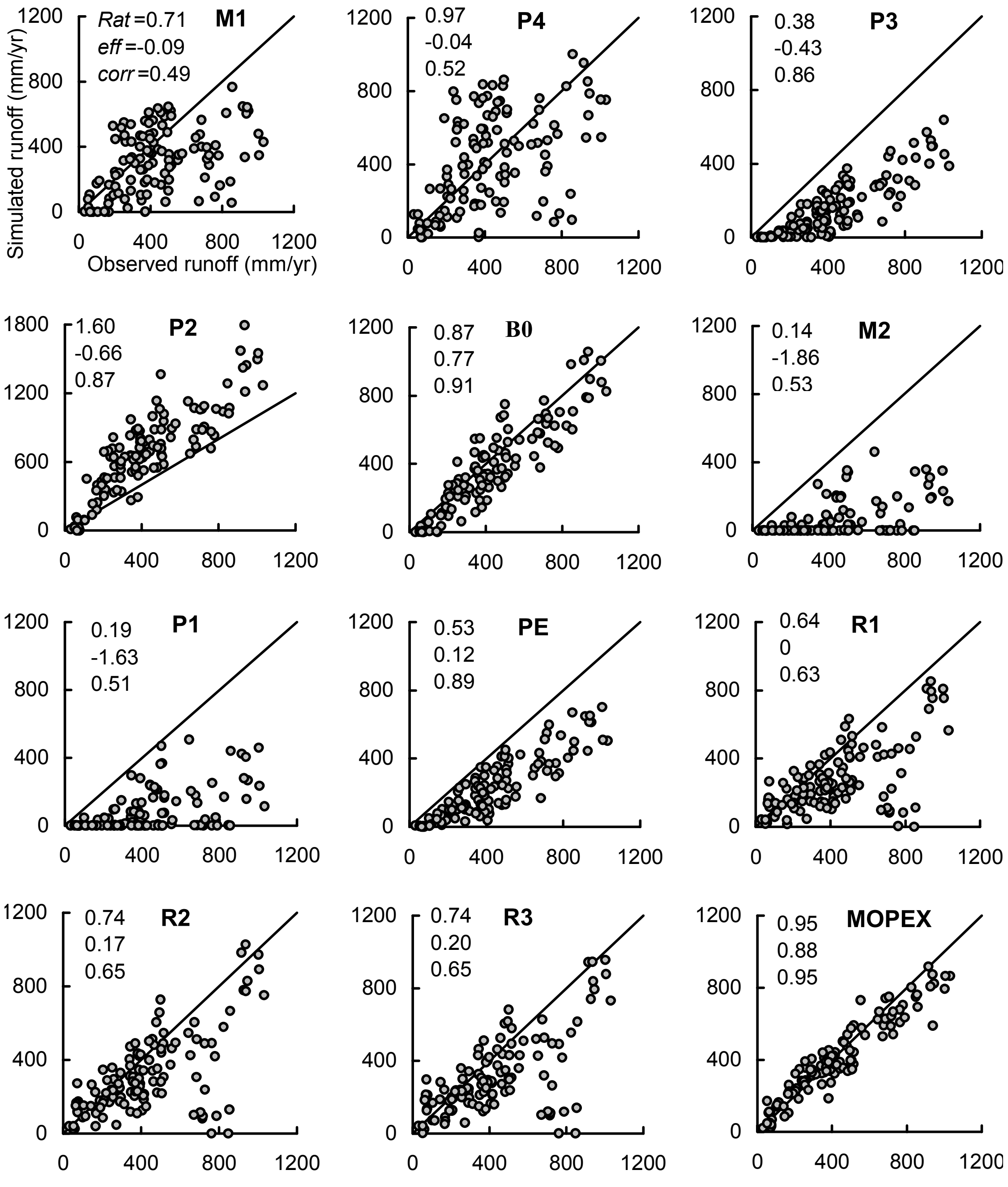


Fig. 2 Comparison of simulated and observed annual streamflow for each of the 12 basins over the 1986-1995 period. All designations are placed on the first panel.

of experiments M1→P4→P3→P2→B0 and M2→P1→PE, we move step by step from pure re-analysis NCEP/DOE (M1) or ERA-40 (M2) products to fully hybridized ones. The best result among the experiments with global data is in the case of B0 experiment: *rat* = 0.87, *eff* = 0.77 and *corr* = 0.91. For this experiment the simulated monthly and annual runoff from each basin was validated against observed runoff. The results of validation are shown in Table 1, which also contains, for comparison, the results from the MOPEX experiment simulation. As seen from the Table, for some

Table 1 Comparison of modelled and observed runoff from the twelve basins on monthly (mm month^{-1}) and annual (mm year^{-1}) basis for 1986–1995. The numerator refers to the B0-experiment results, the denominator represents the MOPEX-experiment results.

Basin	Mean observed	Mean simulated	Bias	RMSD	Nash-Sutcliffe efficiency	Correlation
Monthly statistics						
1	25	32.4/26.8	7.5/1.8	26.9/20.7	0.01/0.51	0.82/0.77
2	32.9	26.4/33.1	–6.6/0.2	22.3/19.8	0.57/0.66	0.86/0.83
3	29.7	30.4/27.5	0.8/–2.1	27.3/17.2	–0.07/0.57	0.77/0.77
4	59.6	43.6/59.1	–16.0/–0.54	34.5/33.6	0.5/0.53	0.81/0.74
5	34.5	38.7/31.5	4.3/–2.9	26.5/24.1	0.45/0.55	0.84/0.75
6	31.0	28.8/34.1	–2.2/3.1	25.8/13.3	0.07/0.75	0.74/0.88
7	65.5	58.6/59.3	–7.0/–6.2	39.6/19.8	–0.33/0.67	0.82/0.84
8	24.2	18.8/24.4	–5.4/0.2	22.5/19.7	0.58/0.68	0.78/0.83
9	33.1	24.7/28.2	–8.5/–4.9	33.3/34.0	0.46/0.44	0.71/0.68
10	60.1	54.9/53.1	–5.1/–7.0	36.7/34.7	0.62/0.66	0.84/0.85
11	12.1	4.4/10.6	–7.7/–1.5	15.8/9.2	0.37/0.79	0.73/0.89
12	16.4	6.6/12.7	–9.8/–3.7	18.8/13.3	0.33/0.67	0.76/0.83
Mean	35.3	30.7/33.4	–4.6/–2.0	27.5/21.6	0.30/0.62	0.79/0.80
Annual statistics						
1	299.6	389.3/321.6	89.8/22.0	122.1/39.2	–1.17/0.78	0.92/0.92
2	395.3	316.6/397.8	–78.7/2.5	110.1/71.3	0.21/0.67	0.88/0.83
3	356.1	365.4/330.3	9.3/–25.8	79.7/63.7	0.44/0.64	0.85/0.84
4	715.2	522.9/708.7	–192.3/–6.5	206.9/103.6	–0.30/0.67	0.93/0.87
5	413.6	464.9/378.6	51.3/–35.1	112.4/81.7	–0.23/0.35	0.86/0.70
6	371.5	345.6/408.7	–25.9/37.2	66.2/57.3	0.76/0.82	0.89/0.95
7	786.5	702.6/711.8	–83.8/–74.6	145.5/123.5	0.60/0.71	0.93/0.93
8	290.1	225.5/292.9	–64.6/2.7	93.8/133.2	0.87/0.75	0.97/0.95
9	397.3	295.9/338.5	–101.4/–58.7	123.0/122.0	0.52/0.53	0.97/0.85
10	720.9	659.1/637.0	–61.8/–83.9	123.6/101.3	–0.06/0.29	0.76/0.89
11	145.2	53.1/127.3	–92.1/–17.9	107.5/45.9	0.24/0.86	0.91/0.94
12	196.7	79.2/152.4	–117.5/–44.3	133.6/90.8	0.21/0.64	0.91/0.90
Mean	424.0	368.3/400.5	–55.6/–23.5	118.7/86.1	0.17/0.64	0.90/0.88

RMSD; root mean square deviation.

basins the values of the coefficient of correlation in the B0 experiment are even higher than in the MOPEX experiment, while the values of bias and RMSD are greater in B0. No wonder that the efficiency of simulations in B0, as a rule, is lower, than in the MOPEX run.

All experiments with different precipitation data sets are generalized in Fig. 3(a),(b) where annual precipitation values (averaged over 10 years and 12 basins) are sorted in increasing order. As seen, they vary among the experiments greatly: from 924 (experiment P1) to 1732 (experiment P2) mm year^{-1}. The appropriate values of modelled runoff (Fig. 3(a)) and runoff ratio (Fig. 3(b)) are within the ranges 80–680 mm year^{-1} (i.e. modelled runoff may differ by an order of magnitude in dependence of precipitation) and 0.09–0.39, respectively. In the case of the lowest precipitation, annual runoff is underestimated by 343 mm year^{-1} (or 81%). In the case of the largest precipitation, annual runoff is overestimated by 256 mm year^{-1} (or 60%).

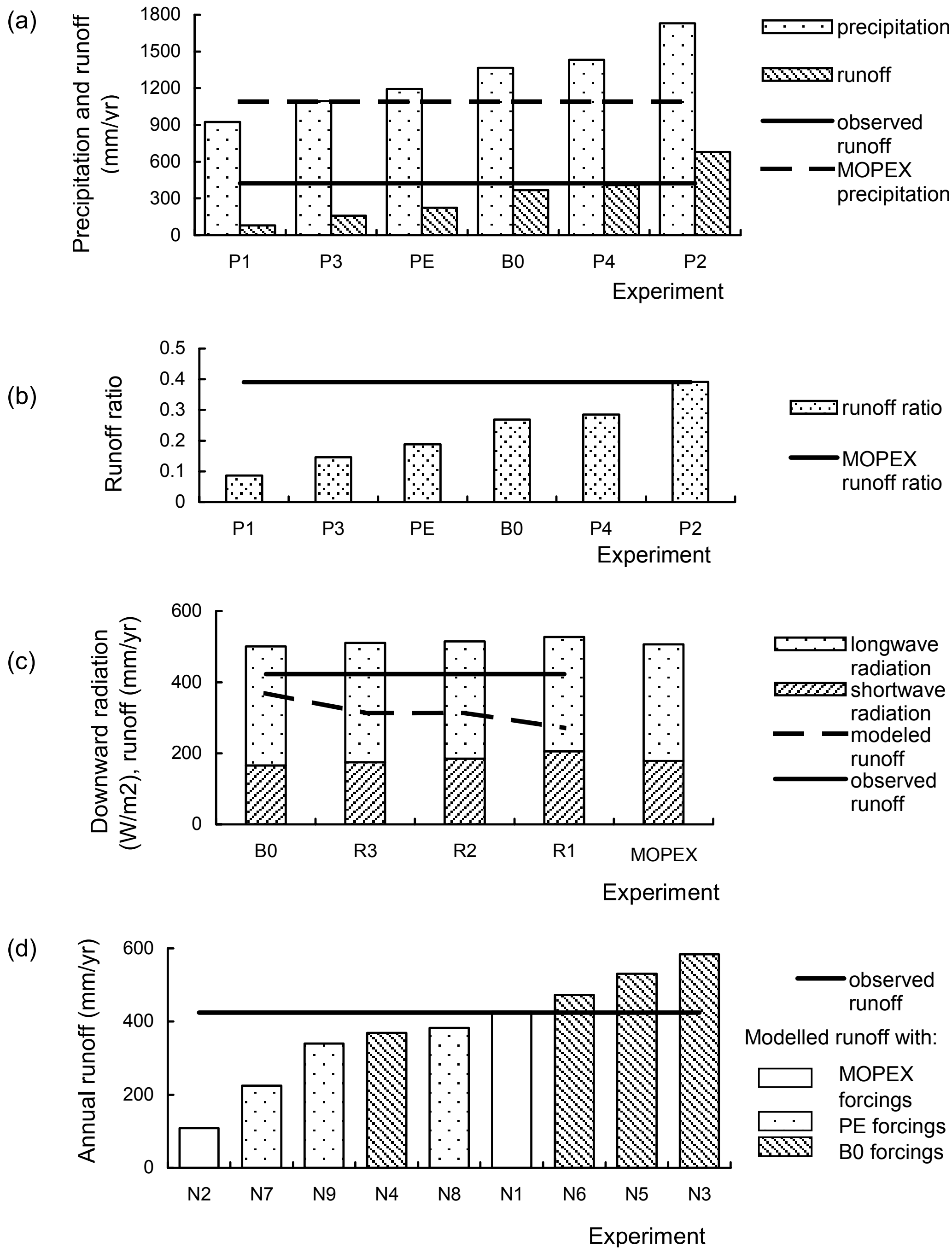


Fig. 3 Generalized results from the sensitivity experiments with different precipitation (a, b), downward radiation (c), and soil parameter data sets (d). The annual values of precipitation, runoff and radiation averaged over 10-year period (1986–1995) and over 12 basins are shown.

Four experiments with downward radiation $R\downarrow$ are generalized in Fig. 3(c). Mean annual $R\downarrow$ with partitioning between shortwave and longwave radiation is shown, being also sorted in increasing order, MOPEX radiation is given for comparison. The largest difference, equalled to 27 W m^2 (≈5%), is between SRB and NCEP/DOE radiation. This corresponds to a difference in simulated annual runoff equal to 98 mm $year^{-1}$ (≈27%). Compared to the observed runoff, the simulated annual runoff is underestimated in all these experiments (by 13% for the lowest *R*↓ and by 36% for the highest *R*↓).

Different combinations of forcing data (MOPEX forcings and fully hybridized global forcing data sets: B0 and PE) and land surface parameters (from MOPEX and GSWP-2, as well as two soil data sets derived by us) allowed us to perform the following sensitivity experiments:

N1 MOPEX forcing data and MOPEX land surface parameters.
N2 MOPEX forcing data and GSWP-2 land surface parameters.
N3 B0 forcing data and MOPEX land surface parameters.
N4 B0 forcing data and GSWP-2 land surface parameters.
N5 B0 forcing data and global soil parameters derived from Clapp & Hornberger (1978) and Dunne & Willmott (1996), the other land surface parameters are from GSWP-2.
N6 B0 forcing data and global soil parameters derived from Cosby *et al*. (1984), the other land surface parameters are from GSWP-2.
N7 PE forcing data and GSWP-2 parameters.
N8 PE forcing data and global soil parameters derived from Clapp & Hornberger (1978) and Dunne & Willmott (1996), the other land surface parameters are from GSWP-2.
N9 PE forcing data and global soil parameters derived from Cosby *et al*. (1984), the other land surface parameters are from GSWP-2.

The results from these experiments with different sets of parameters are presented in Fig. 3(d). The simulated annual runoff is sorted in increasing order and compared to the observed one. As seen, the impact of uncertainties in parameters on the simulated runoff is of the same order of magnitude as the impact of uncertainties in precipitation. The simulated annual runoff, averaged over 10 years and over all basins, varies from 109 to 584 mm $year^{-1}$. In experiment N2, being the lowest runoff is underestimated by 315 mm $year^{-1}$ (74%), while in experiment N3 it is overestimated by 160 mm $year^{-1}$ (38%). Combination of B0 forcings with soil parameters from GSWP-2 (N4 experiment) and PE forcings with soil parameters derived by using relationships reported by Clapp-Hornberger and Dunne-Willmott (N8 experiment) show the best results among simulations with different global 1-degree parameter data sets. The statistics for the simulated annual runoff in both cases are nearly the same: for PE, *rat* = 0.90, *eff* = 0.77 and *corr* = 0.90; for B0, they are 0.87, 0.77 and 0.91, respectively. Further improvement of these results may be achieved by means of model calibration by tuning some model parameters (e.g. the hydraulic conductivity at saturation) (Gusev & Nasonova, this issue).

The results obtained have shown that uncertainties in the two primary factors that force hydrological processes—precipitation and incoming radiation, as well as in the land surface parameters may cause large differences in simulated runoff. That is why global data sets should be tested for a wide range of river basins located all over the world (because data quality may vary spatially) before application on ungauged basins. It should also be noted that more attention should be paid to the development of alternative global data sets with the land surface (especially soil) parameters since their impact on runoff simulations is comparable to that of the main forcing factors.

Acknowledgements This work was supported by the Russian Foundation for Basic Research (Grant 05-05-64161). We acknowledge the MOPEX and GSWP-2 experiment organizers for providing us with the data to run the model.

REFERENCES

Clapp, R. B. & Hornberger, G. M. (1978) Empirical equations for some soil hydraulic properties. *Water Resour. Res*. **14**, 601–604.

Cosby, B. J., Hornberger, G. M., Clapp, R. B. & Ginn, T. R. (1984) A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res*. **20**(6), 682–690.

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.

Dunne, K. A. & Willmott, C. J. (1996) Global distribution of plant-extractable water capacity of soil. *Int. J. Climatol.* **16**, 841–859.

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

Gusev, Ye. M. & Nasonova, O. N. (2006) Simulating runoff from MOPEX experimental river basins using the SWAP land surface model and different parameter estimation techniques. In: *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment–MOPEX* (ed. by V. Andréssian, A. Hall, N. Chahinian & J. Schaake), 188–195. IAHS Publ. 307. IAHS Press, Wallingford, UK (this issue).

Jayawardena, A. W. & Mahanama, S. P. P. (2001) Daily river discharge prediction using GCM generated atmospheric data. In: *Soil-Vegetation–Atmosphere Transfer Schemes and Large-Scale Hydrological Models* (ed. by A. J. Dolman, A. J. Hall, M. L. Kavvas, T. Oki & J. W. Pomeroy), 159–166. IAHS Publ. 270. IAHS Press, Wallingford, UK.

Oki, T., Musiake, K., Matsuyama, H. & Masuda, K. (1995) Global atmospheric water balance and runoff from large river basins. *Hydrol. Processes* **9**, 655–678.

Salathé, E. P. (2003) Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *Int. J. Climatol*. **23**, 887–901.

Wilby, R. L. & Wigley, T. M. L. (1997) Downscaling general circulation model output: a review of methods and limitations. *Progr. Phys. Geogr.* **21**, 530–548.

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