Soil–Vegetation–Atmosphere Transfer Schemes and Large-Scale Hydrological Models (Proceedings of a symposium held during the Sixth IAHS Scientific Assembly at Maastricht, The Netherlands, July 2001). IAHS Publ. no. 270, 2001.

Investigating the ability of a land surface model to simulate runoff on a large river basin scale

YEUGENIY M. GUSEV & OLGA N. NASONOVA

Water Problems Institute, Russian Academy of Sciences, Gubkina St. 3, Moscow 117971, Russia

e-mail: gusev@aqua.laser.ru

Abstract The aim of the present work is to investigate the ability of the physically-based land surface model SWAP-2 (Soil–Water–Atmosphere–Plants) to simulate runoff on a large river basin scale, and to reveal the potentials for model improvement. The model treats heterogeneity of a large basin explicitly, i.e. it divides the basin into a number of computational units provided with deterministic effective values for land surface parameters and atmospheric forcings. The SWAP-2 model was validated against naturalized streamflow from 15 drainage catchments located within the Red-Arkansas River basin for the period 1979–1988. The accuracy of the model was found to be close to the estimated maximum accuracy under the chosen discretization of drainage basins and prescribed effective input data. The SWAP-2 model does not incorporate significant systematic error into the results and can operate at a regional scale with satisfactory accuracy under appropriate discretization of a basin without calibration of its land surface parameters.

Key words land surface model; runoff; Red-Arkansas River basin; regional-scale modelling

INTRODUCTION

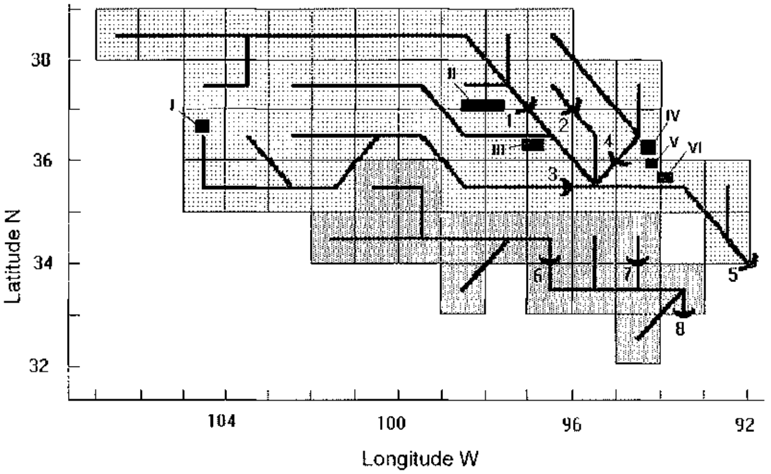
The present paper is based on our previous investigations carried out within the framework of the Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase 2(c) experiment (Wood *et al.*, 1998; Lohmann *et al.*, 1998). The primary goal of the PILPS Phase 2(c) experiment was to evaluate the ability of sixteen land surface models (LSMs) intended for coupling with atmospheric models to reproduce heat and water exchange between the land surface and the atmosphere at various spatial scales varying from small catchments (of the order of 10^2–10^3 km^2) to a continental-scale river basin—the Red-Arkansas River basin (of the order of 10^6 km^2) characterized by different natural conditions. Large catchments and river basins were divided into a number of computational units ($1° \times 1°$ grid cells) consistent with the grid-scale of atmospheric models and connected by a stream network. Streamflow from different catchments and river basins simulated by sixteen LSMs were intercompared and validated against observations (Wood *et al.*, 1998; Lohmann *et al.*, 1998).

The first step in our investigations was to participate in the PILPS Phase 2(c) experiment with our LSM SWAP-2 (soil–water–atmosphere–plants) model, and the next step was the analysis of the results obtained to achieve a better understanding of the behaviour of the model, to identify shortcomings, and to find the ways and limits for improvement of the model performance. The current paper represents results obtained during the second step.

MODEL AND DESIGN

The SWAP-2 model is a one-dimensional physically-based model describing the interactions between the land surface and the atmosphere and being oriented to the coupling with atmospheric and hydrological models. It is based on a system of physical–mathematical equations for the surface energy balance, the water balances of the canopy, the soil column and snow cover, and for the heat and water transfer within a soil–vegetation/snow cover–atmosphere system. Being connected with the routing scheme which transforms the model generated runoff into river streamflow, SWAP-2 treats the following processes: interception of precipitation by a canopy, evaporation of intercepted precipitation, partitioning non-intercepted precipitation into infiltration and surface runoff, transpiration, soil evaporation, dynamics of the frozen and liquid water in a soil column, drainage, snow cover formation (snow accumulation including its solid and liquid fractions, snow evaporation, and snowmelt), formation of the surface energy balance, soil freezing, soil thawing, transformation of grid-generated runoff into streamflow at the outlet of a grid cell, and into river network streamflow. A detailed description of the latest version of SWAP-2 is given in Gusev & Nasonova (2000a).

The model was provided with three groups of data by the PILPS Phase 2(c) organizers: (a) atmospheric forcing data, (b) soil, vegetation and snow parameters, and (c) validation data. All these data sets are detailed in Wood *et al.* (1998). Using these data, SWAP-2 was run (RUN1) for the period 1979–1988 with a 3-h time step to simulate different components of the heat and water balances for 15 drainage basins (Fig. 1, Table 1). In so doing, the data of observed (to be more exact, "naturalized")



Legend:
A: Arkansas River basin
B: Red River basin
I: Canadian River
II: Chikaskia River
III: Black Bear Creek
IV: Illinois River
V: Lee Creek
VI: Mulberry River

Streamflow evaluation sites:
1: Ralston
2: Lenapah
3: Eufaula
4: Ft Gibson
5: Little Rock (outlet of the Arkansas River basin)
6: Texoma
7: Millwood
8: Shreveport (outlet of the Red River basin)

Fig. 1 Schematization of the Red-Arkansas River basin.

Table 1 Catchments under study, their size and the number of constituting computational units.

Catchments	No. of computational units	Drainage area (km^2)
1. Canadian River catchment	1	593
2. Mulberry River catchment	1	966
3. Lee Creek catchment	1	1 103
4. Black Bear Creek catchment	1	1 491
5. Illinois River catchment	1	2 483
6. Chikaskia River catchment	1	4 814
7. Lenapah sub-catchment	1	9 424
8. Millwood sub-catchment	1	10 668
9. Ft Gibson sub-catchment	4	32 360
10. Texoma sub-catchment	9	87 492
11. Eufaula sub-catchment	13	123 074
12. Ralston sub-catchment	14	141 056
13. Red River basin (outlet at Shreveport)	16	156 978
14. Arkansas River basin (outlet at Little Rock)	45	409 273
15. Red-Arkansas River basin	61	566 251

streamflow represent the main source of information for a solution of the above problems related to understanding the quality of the model. In addition, SWAP-2 was run (RUN2) with aggregated land surface parameters (averaged over 61 grid cells, land surface parameters were applied for each grid cell) to study the influence of spatial heterogeneity of the land surface on the accuracy of streamflow calculation from the whole basin.

ANALYSIS OF RESULTS

The most informative statistical characteristic to estimate the quality of the model, in our opinion, is the root mean squared (RMS) deviation $\sigma_{R,d}$ between simulated R_{cal} and observed R_{obs} daily streamflow, because it is sensitive both to a systematic error in simulations and to discrepancies in the dynamics of R_{cal} and R_{obs}. Hence, here most emphasis will be on $\sigma_{R,d}$.

Since drainage basins consist of different numbers, N, of grid cells (Table 1), data on $\sigma_{R,d}(N)$ estimated for each basin should be transformed to $\sigma_{R,d}(N_0)$ (where $N_0 = 61$ is the number of grid cells composing the Red-Arkansas River basin) to obtain statistics of $\sigma_{R,d}$. In particular, it can be shown (Gusev & Nasonova, 2000b) that, given $\sigma_{R,d}(N)$ for each of the 15 drainage basins, $\sigma_{R,d}(N_0)$ is calculated by the relationship:

$$\sigma_{R,d}(N_0) = \sqrt{\frac{N}{N_0}}\sigma_{R,d}(N) \qquad (1)$$

Then, the values of $\sigma_{R,d}(N_0)$ derived from simulated and observed streamflow from each basin are averaged to obtain their mean value $\overline{\sigma}_{R,d}(N_0)$, which characterizes a quality of the model's performance. The statistical characteristics of the deviation of $\sigma_{R,d}(N_0)$ from $\overline{\sigma}_{R,d}(N_0)$ can be estimated on the basis of the statistics of χ^2 distribution of $\sigma_{R,d}(N_0)$.

To answer the question, whether it is possible to improve the performance of the model, and to what extent, it is necessary to estimate the lower limit of error $\sigma_{\mathrm{lim}}(N_0)$ to which $\overline{\sigma}_{R,d}(N_0)$ can tend. The lower limit is connected with the uncertainty of the effective parameters and forcing data used for model simulations with a coarse (1° × 1°) spatial resolution. For the first approximation, $\sigma_{\mathrm{lim}}(N_0)$ was estimated by an expression which is analogous to equation (1) (Gusev & Nasonova, 2000b):

$$\sigma_{\mathrm{lim}}(N_0) = \frac{1}{\sqrt{N_0}}\sigma_{\mathrm{lim}}(1) \tag{2}$$

where $\sigma_{\mathrm{lim}}(1)$ can be interpreted as a maximum accuracy of the determination of streamflow for one computational grid cell. The value of $\sigma_{\mathrm{lim}}(1)$ was calculated on the basis of the consideration of streamflow differences for three small catchments (the Mulberry River, Lee Creek, and the Illinois River) located within one computational grid cell. Although these three catchments are situated within a common cell, the observed streamflow and precipitation data for these catchments are sufficiently different. This means that the observed data from small catchments allows us to estimate effective parameters for a grid cell only with an error $\sigma_{\mathrm{lim}}(1)$ which can be estimated as:

$$\sigma_{\mathrm{lim}}(1) = \sigma_{1\text{-}3} \tag{3}$$

where $\sigma_{1\text{-}3}$ is the RMS deviation between streamflow measured at the three catchments and its mean value.

The suggested technique was applied to evaluate the performance of SWAP and to compare the two model runs (RUN1 and RUN2, mentioned above). As may be seen in Fig. 2, explicit accounting for spatial heterogeneity of the land surface by application of different deterministic effective land surface parameters in each grid cell does not practically influence streamflow simulations. The reason may be connected, in particular, with the fact that the effective parameters provided for each grid cell are not

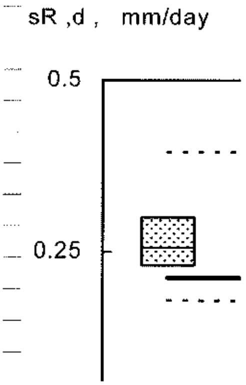


Fig. 2 Root mean squared errors of simulated by SWAP-2 daily streamflow for the Red-Arkansas River basin for 1979–1988 (bars) and their lower limits (lines) estimated using observed streamflow. The errors of simulated streamflow are given by the average value (horizontal solid line within a bar), and the upper and lower quartile (boxed region). The lower limits of errors are also given by the average value (horizontal solid line) and the upper and lower quartile (dashed lines).

optimal, and that is why their replacement by coarser values (averaged over 61 grid cells) cannot sufficiently spoil the results. The accuracy of streamflow simulation in both cases is close to the maximum accuracy under the given schematization of drainage basins (e.g. the error of modelled daily streamflow is 0.25–0.26 mm day^{-1}, as compared to the estimated minimum error of 0.21 mm day^{-1}).

It is interesting to apply the suggested technique for intercomparison of the performance of the 16 LSMs participating in the PILPS Phase 2(c) experiment. Here, the intercomparison of the models can be made using only a rather integrated characteristic: deviation between simulated and observed mean annual streamflow published for seven drainage basins in Lohmann *et al.* (1998). Unfortunately, more detailed data were not available to us. The estimated mean calculation error for each of 16 models (including SWAP-2) and the lower limit of error (both with the upper and lower quartile) are presented in Fig. 3. Since, in the PILPS Phase 2(c) experiment, all LSMs have used the same forcing data, the same routing models and the same values for most parameters, the scatter in the simulation errors can be explained by the differences in parameterizations and in the structure of the LSMs, and, to a lesser extent, by uncertainties in those parameters which were calibrated by each modelling group. As may be seen in Fig. 3, there are several models (including SWAP-2) whose accuracy is very close to the maximum accuracy under the given schematization of drainage basins. However, final conclusions about the quality of models can be made only after consideration of the statistical characteristics of runoff hydrographs.

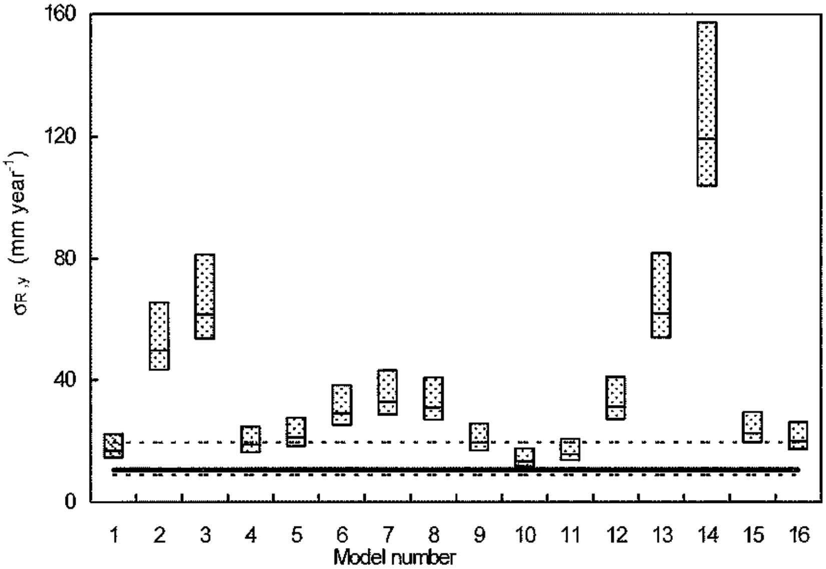


Fig. 3 Root mean squared errors of annual streamflow $\sigma_{R,y}$ simulated by 16 LSMs for the Red-Arkansas River basin for 1980–1986 (bars) and their lower limits (lines) estimated using observed streamflow. The errors of simulated streamflow are given by the average value (horizontal solid line within a bar), and the upper and lower quartile (boxed region). The lower limits of errors are also given by the average value (horizontal solid line) and the upper and lower quartile (dashed lines). The number 1 corresponds to the model SWAP-2.

Analysis of SWAP-2-simulated streamflow from different drainage basins has shown that the larger the basin, the better agreement with observations. Thus, for the whole Red-Arkansas River basin the coefficient of correlation for daily values is equal to 0.71, the relative deviation ε_Q between modelled and naturalized mean annual discharge (1979–1988) is nearly 3%. At the same time, ε_Q and $\sigma_{R,d}$ increase with a decrease in basin size; to be more exact, in the number of computational grid cells, N, comprising the basin (Fig. 4) that is in a good agreement with equation (1). The calculated values of $\sigma^2_{R,d}$ follow the law derived from equation (1):

$$\sigma^2_{R,d}(N) = \sigma^2_{R,d}(1)/N \tag{4}$$

where $\sigma^2_{R,d}(1)$ was estimated by averaging $\sigma^2_{R,d}$ values for eight unit catchments. Since equation (1) was derived under the assumption that errors in different grid cells represent independent random variables (i.e. the systematic error component is absent), a good agreement between empirical and theoretical estimation of $\sigma^2_{R,d}$ means that the SWAP-2 model does not incorporate significant systematic error into the results. As such, the model can reproduce the dynamics of streamflow on a regional scale with satisfactory accuracy under appropriate discretization of the area. In so doing, errors in the modelled streamflow for different computational units resulted from random errors in estimation of effective input data and parameters compensate each other to a large extent.

As to unit basins which use "point" inputs, errors in the calculations may be significant. In such cases, more accurate parameter estimation, or adjustment of their values by means of calibration, is needed. However, as discussed above, in using this approach there is a limit to increasing the accuracy of simulations due to the uncertainties in effective parameters.

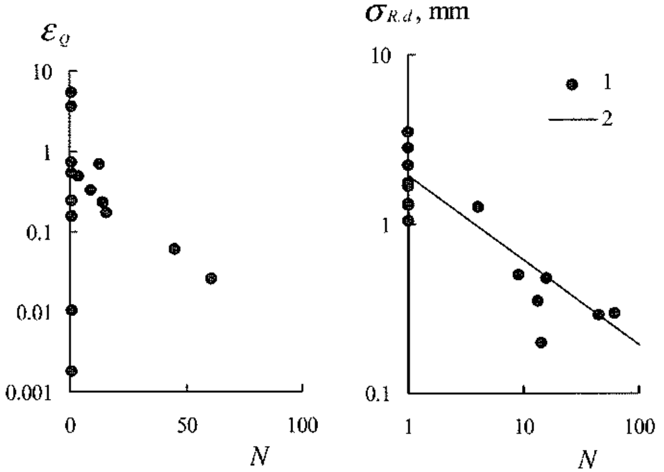


Fig. 4 Relative deviations between mean annual (1979–1988) modelled and naturalized discharge, ε_Q, and RMS deviations between daily values of modelled and naturalized streamflow, $\sigma_{R,d}$, for different drainage basins *vs* the number of grid cells, N, composing a basin. 1: empirical results; 2: theoretical curve according to equation (4).

CONCLUSION

Analysis of the SWAP-2 model validation results shows that it simulates runoff from large drainage basins without significant systematic error, and the accuracy of simulations is close to the expected maximum accuracy under the given schematization of basins and prescribed effective inputs. Better performance of the model can be achieved only in the case of better quality of data and finer spatial resolution.

Acknowledgements The work was supported by the Russian Foundation for Basic Researches. We acknowledge PILPS Phase 2(c) experiment organizers (Eric F. Wood, Dennis P. Lettenmaier and Xu Liang), Xu Liang for providing us with the data for this study and Dag Lohmann for the river routing model.

REFERENCES

Gusev, Ye. M. & Nasonova, O. N. (2000a) Modeling of land surface heat and water exchange on a regional scale. *Wat. Resour.* 27(1), 27–40.

Gusev, Ye. M. & Nasonova, O. N. (2000b) An experience of modelling heat and water exchange at the land surface on a large river basin scale. *J. Hydrol.* 233(1–4), 1–18.

Lohmann, D., Lettenmaier, D. P., Liang, X., Wood, E. F., Boone, A., Chang, S., Chen, F., Dai, Y., Desborough, C., Dickinson, R. E., Duan, Q., Ek, M., Gusev, Y. M., Habets, F., Irannejad, P., Koster, R., Mitchell, K. E., Nasonova, O. N., Noilhan, J., Schaake, J., Schlosser, A., Shao, Y., Shmakin, A. B., Verseghy, D., Warrach, K., Wetzel, P., Xue, Y., Yang, Z.-L. & Zeng, Q. (1998) The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase 2(c) Red-Arkansas River basin experiment: 3. Spatial and temporal analysis of water fluxes. *Global Planet. Change* 19(1–4), 161–179.

Wood, E. F., Lettenmaier, D. P., Liang, X., Lohmann, D., Boone, A., Chang, S., Chen, F., Dai, Y., Dickinson, R. E., Duan, Q., Ek, M., Gusev, Y. M., Habets, F., Irannejad, P., Koster, R., Mitchell, K. E., Nasonova, O. N., Noilhan, J., Schaake, J., Schlosser, A., Shao, Y., Shmakin, A. B., Verseghy, D., Warrach, K., Wetzel, P., Xue, Y., Yang, Z.-L. & Zeng, Q. (1998) The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase 2(c) Red-Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. *Global Planet. Change* 19(1–4), 115–135.