Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management (Proceedings of Symposium HS2004 at IUGG2007, Perugia, July 2007). IAHS Publ. 313, 2007.

Can a land surface model simulate runoff with the same accuracy as a hydrological model?

OLGA N. NASONOVA & YEUGENIY M. GUSEV

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

nasonova@aqua.laser.ru

Abstract The ability of the land surface model SWAP (Soil Water–Atmosphere–Plants) to reproduce runoff from the 12 MOPEX (Model Parameter Estimation Experiment) experimental river basins compared to the hydrological Sacramento model (SAC-SMA) was investigated. In previous investigations, the SAC-SMA model (calibrated with 16 model parameters using both manual and automatic calibration techniques) demonstrated much better performance than the SWAP model. In the present study, the behaviour of SWAP was, however, substantially improved by means of advanced model calibration. For the 12 MOPEX basins, the median values of model efficiency and bias for simulated daily streamflow during the calibration period (1960–1979) were found to be 70% and 2.6% respectively for the SAC-SMA model, and 68% and 0.9% respectively for the SWAP model, During the validation period (1980–1998), model efficiency and bias were 65% and 6.3% for SAC-SMA model, and 64% and 3.9% for the SWAP model. It was found that model performance depends greatly on the skill of calibration, and that the LSM SWAP under appropriate calibration can simulate runoff with accuracy comparable to that of the hydrological model.

Key words hydrological Sacramento model; land surface model SWAP; MOPEX river basins; parameter calibration; river runoff

INTRODUCTION

Hydrological models (HMs) treat different parts of the hydrological cycle and primarily aim at maximizing the accuracy of streamflow prediction rather than at representing the underlying physical processes. Land surface models (LSMs) parameterize heat and water exchange processes between the land surface and the atmosphere. Initially, LSMs were designed for coupling with atmospheric models using simple land surface parameterization schemes. More recently they have evolved into complex physically-based models which may be used in a stand-alone mode for simulating different components of the heat and water balance at the land–atmosphere interface as well as different characteristics of the hydrothermal regime. Outputs of LSMs include more than 50 variables, including runoff. In the hydrological community, there is an opinion that LSMs cannot be as successful as hydrological models with respect to runoff simulation because LSMs, being too complicated, suffer from accumulated errors in the larger number of forcing data and model parameters they require relative to the lesser data demands of HMs. Better performance of HMs compared to LSMs (including the land surface model SWAP, Soil Water–Atmosphere–Plants; Gusev & Nasonova, 2000, 2003) was obtained after model calibration within the framework of the MOPEX (Model Parameter Estimation Experiment) project (Duan *et al.*, 2006). In

Copyright © 2007 IAHS Press

this context, the following questions arise: What caused the HMs to perform better than the LSMs? What should be undertaken to make LSMs as successful as HMs? The present work is an attempt to investigate these issues using the LSM SWAP and Sacramento model (SAC-SMA) (Gan & Burges, 2005) streamflow simulations of 12 MOPEX river basins. The SAC-SMA model was one of the best performing hydrological models in the MOPEX project.

MOPEX RIVER BASINS AND DATA

Twelve river basins (with drainage area of 1020 to 4421 km^2) selected within the framework of the Second MOPEX Workshop (Duan *et al.*, 2006) are used in this study. All the basins are located within the southeastern part of the USA and are characterized by a great variety of natural conditions, ranging from arid to humid and from grassland/cropland to forested areas. Snow and frozen ground effects are considered to have a minor impact on the hydrological processes because all basins are located south of 42°N. Full descriptions of the basins can be found in Duan *et al.* (2006).

The data for each of the 12 river basins, which were provided by the Second MOPEX Workshop organizers (Duan *et al.*, 2006), include meteorological forcing data, daily streamflow discharge data at basin outlets for model calibration and validation, as well as basin characteristics data. The forcing data sets for a 39-year period (1960–1998) include basin-averaged hourly near-surface meteorology such as downward shortwave and longwave radiation, air temperature and humidity, atmospheric precipitation, air pressure and wind speed. For HMs, daily climatic potential evaporation was also provided. Data to represent basin characteristics include basin boundary, elevation, spatial distribution of different soil and vegetation classes, soil texture and monthly greenness fraction. Soil and vegetation parameters for each of the 12 basins were derived using available data on the spatial coverage of the USDA soil texture classes and of the University of Maryland vegetation types (Gusev & Nasonova, 2007).

THE LAND SURFACE MODEL SWAP

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 on different scales—from small catchments to continental scale river basins. In the case of a small catchment (up to the order of 10^3–10^4 km^2), a kinematic wave equation is used to simulate runoff at the catchment outlet. In the case of larger river basins, the area is divided into computational grid cells connected by a river network. Runoff is modelled for each cell and then transformed by a river routing model to simulate streamflow at the river basin outlet.

During the last 10 years, different versions of SWAP were validated against various observed hydrothermal characteristics. The validations were performed for "point" experimental sites and for catchments and river basins of different areas (from 10^{-1} to 10^{5} km^2) on a long-term basis and under different natural conditions. The results demonstrated that SWAP is able to reproduce (without calibration) annual and interannual dynamics of the afore mentioned characteristics fairly well, provided that input data of high quality are available. The model structure, its parameterizations and the results of validation are detailed in a number of publications (e.g. Boone *et al.*, 2004; Gusev & Nasonova, 2000, 2003).

THE HYDROLOGIC SAC-SMA MODEL

The Sacramento Soil Moisture Accounting model, SAC-SMA, is a deterministic, conceptually-based rainfall–runoff model with spatially lumped parameters (Burnash *et al.*, 1973). It simulates runoff from a catchment using precipitation and potential evapotranspiration as input data. The simulated runoff is converted into streamflow through a unit hydrograph. The SAC-SMA is well-known in the hydrological community and widely used throughout the world for flood and water supply forecasting, and other operational purposes.

METHODOLOGY

First, it is necessary to understand why HMs simulated streamflow better than LSMs in the MOPEX experiment. The reason seems to be of historical character. HMs were initially aimed at adequate simulation of river runoff and the high accuracy of runoff simulations was mostly reached due to calibration of model parameters using direct streamflow measurements. At the same time, LSMs designed originally for atmospheric science applications (in particular, for coupling with general circulation models), aimed at reproducing vertical energy and water fluxes and up to 50 state variables within each computational grid cell. In doing so, *a priori* estimated model parameters were applied in a coupled model mode at mesoscale or global scale. LSMs can also be used in a stand-alone mode and on smaller scales. In this case, model calibration is feasible but due to the multipurpose destination of LSMs, no simple methodologies exist to calibrate LSM models to reproduce all output variables adequately. This could explain why model calibration is not commonly applied by the land surface modelling community. Inexperience in model calibration among land surface modellers could therefore be one of the main reasons for poor simulation of streamflow by LSMs compared to HMs. Indeed, if an advanced LSM treats the physical mechanisms of heat and water exchange processes adequately, it is reasonable to expect a high accuracy of runoff modelling. Poor runoff simulation may result from a low quality of input data and model parameters. As such, calibration of the most important parameters of LSMs using streamflow observations, in accordance with the traditional hydrological approach, can improve the situation.

To verify this operational hypothesis we investigate different ways of SWAP model calibration and to compare the results of runoff simulations (using different sets of calibrated parameters) with each other, with observations and with the SAC-SMA results. The calibration is performed for each MOPEX-basin. Following the MOPEX strategy, the period 1960–1979 is used for model calibration and the period 1980–1998 is used for validation of simulations from both models. The agreement between simulated and observed daily streamflow for each river basin will be estimated using two goodness-of-fit statistics: the Nash-Sutcliffe coefficient of efficiency (*Eff*) (Nash & Sutcliffe, 1970) and absolute value of bias (*Bias*) calculated as:

$$Eff = 1 - \frac{\sum_{\Omega} (x_{\mathrm{sim}} - x_{\mathrm{obs}})^2}{\sum_{\Omega} (x_{\mathrm{obs}} - \bar{x}_{\mathrm{obs}})^2} \quad \text{and} \quad Bias = \left| \frac{\sum_{\Omega} (x_{\mathrm{sim}} - x_{\mathrm{obs}})}{\sum_{\Omega} x_{obs}} \right| \cdot 100$$

where x_{sim} and x_{obs} are simulated and observed values of a variable x and Ω is a discrete sample set of variable x.

INVESTIGATION OF DIFFERENT WAYS OF SWAP MODEL CALIBRATION

As part of the MOPEX project (Duan *et al*., 2006), SWAP demonstrated poorer streamflow simulations compared to simulation results of HMs. This may be explained by application of inadequate estimates of land surface (soil and vegetation) parameters. According to the MOPEX strategy, modellers were not provided with the land surface parameters and had to estimate the parameters by themselves. Poor *a priori* estimates could be improved by model calibration to measured daily streamflow for each of the 12 MOPEX river basins.

Absence of experience in LSM model calibration often does not allow users to do automatic model calibration. Thus, SWAP was calibrated manually by tuning only one parameter (hydraulic conductivity at saturation, k_0) to minimize mean bias between simulated and measured annual streamflow. Using *a priori* estimated parameters (hereafter this case will be referred to as "SWAP_apriori"), the median *Eff* and the median *Bias* for the 12 basins are 45% and 31% for the calibration period while they are 20 % and 47% for the validation period, respectively (Fig. 1). The calibration of k_0 allowed us to reach better results (referred to as "SWAP_CAL1"), especially for the bias and during the validation period. In this case, the median *Eff* for the 12 basins are 52% and 55%, respectively, for the calibration and validation periods, while the median *Bias* is 11% for both periods (Fig. 1).

These results are compared with the results of streamflow simulations performed by the hydrological model SAC-SMA, which were taken from Gan *et al*. (2007). The SAC-SMA was calibrated using a combination of the global optimization algorithm, the shuffle complex evolution method SCE-UA and manual effort (Gan *et al*., 2007). In doing so, 16 parameters were calibrated: 11 "land parameters", three unitgraph ordinates, the precipitation scaling factor and the potential evapotranspiration adjustment factor. Comparison of the results are shown in Fig. 1. In this figure, the SAC-SMA simulations using *a priori* estimated and calibrated parameters are marked as SAC-SMA_apriori and SAC-SMA_CAL, respectively. As can be seen, SAC-SMA outperformed SWAP in both cases. This means that SAC-SMA modellers were more successful when estimating *a priori* parameters and performing model calibration.

At the next stage, we tried to improve the SWAP model calibration procedure. Instead of manual calibration, an automatic procedure of optimization based on a

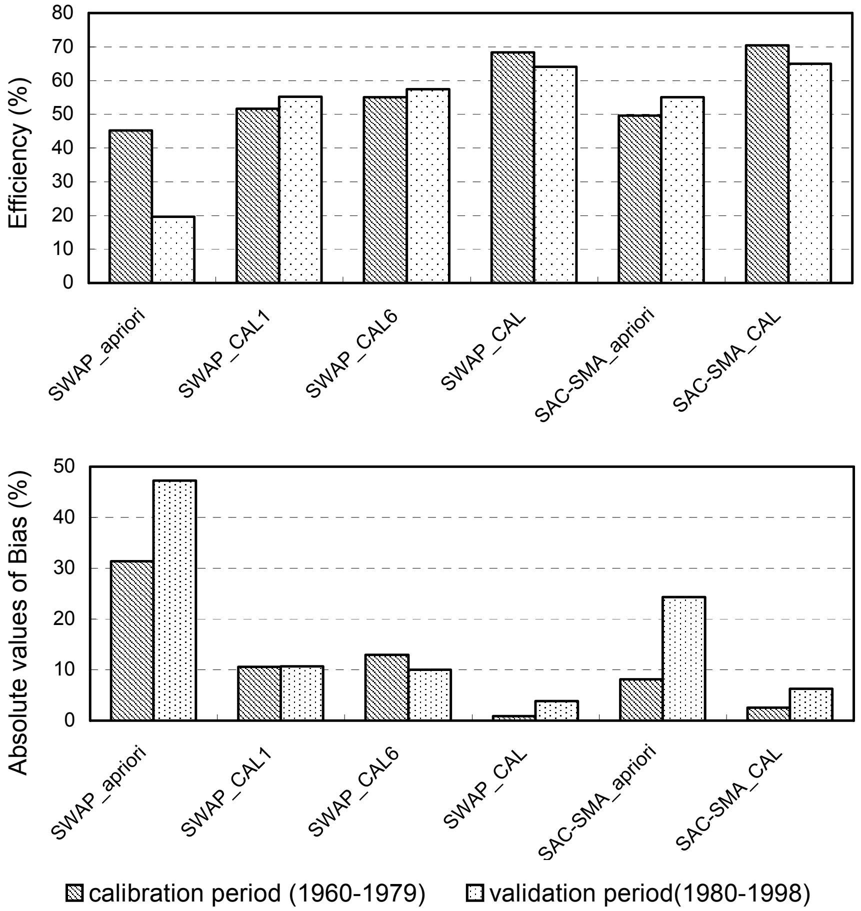

Fig. 1 Median daily Nash-Sutcliffe efficiency and bias for the 12 MOPEX river basins from *a priori* simulation, and calibrated results of SWAP and SAC-SMA models for the calibration and validation periods.

Monte Carlo technique was designed. The objective function used in calibration was changed, and model calibration was performed by minimization of the root mean square deviation, RMSD, between simulated and measured daily streamflow. Application of the automatic procedure allowed us to increase the number of calibrated parameters. Instead of one parameter, six soil parameters (k_0, field capacity W_{fc}, wilting point W_{wp}, soil porosity W_{sat}, soil depth h_{soil}, and root layer depth h_{root}) were calibrated. Improved calibration technique (referred to as "SWAP_CAL6") resulted in slightly better hydrograph simulations: median *Eff* are now 55% and 57%, and median *Bias* are 13% and 10% for the calibration and validation periods respectively. However, SAC-SMA simulations are still better than SWAP simulations (Fig. 1, compare SAC-SMA_CAL and SWAP_CAL6).

Analysis of the behaviour of the two models and their calibration techniques allowed us to conclude that one of the main reasons for better performance of SAC-SMA is the choice of parameters to be calibrated. Following the SAC-SMA calibration strategy, we decided to increase the number of calibrated parameters in the SWAP model. Now the list of calibrated parameters includes several soil parameters (k_0, W_{fc}, W_{wp}, W_{sat}, h_{soil}, h_{root}, soil matric potential at saturation ϕ_0 and the B-parameter in

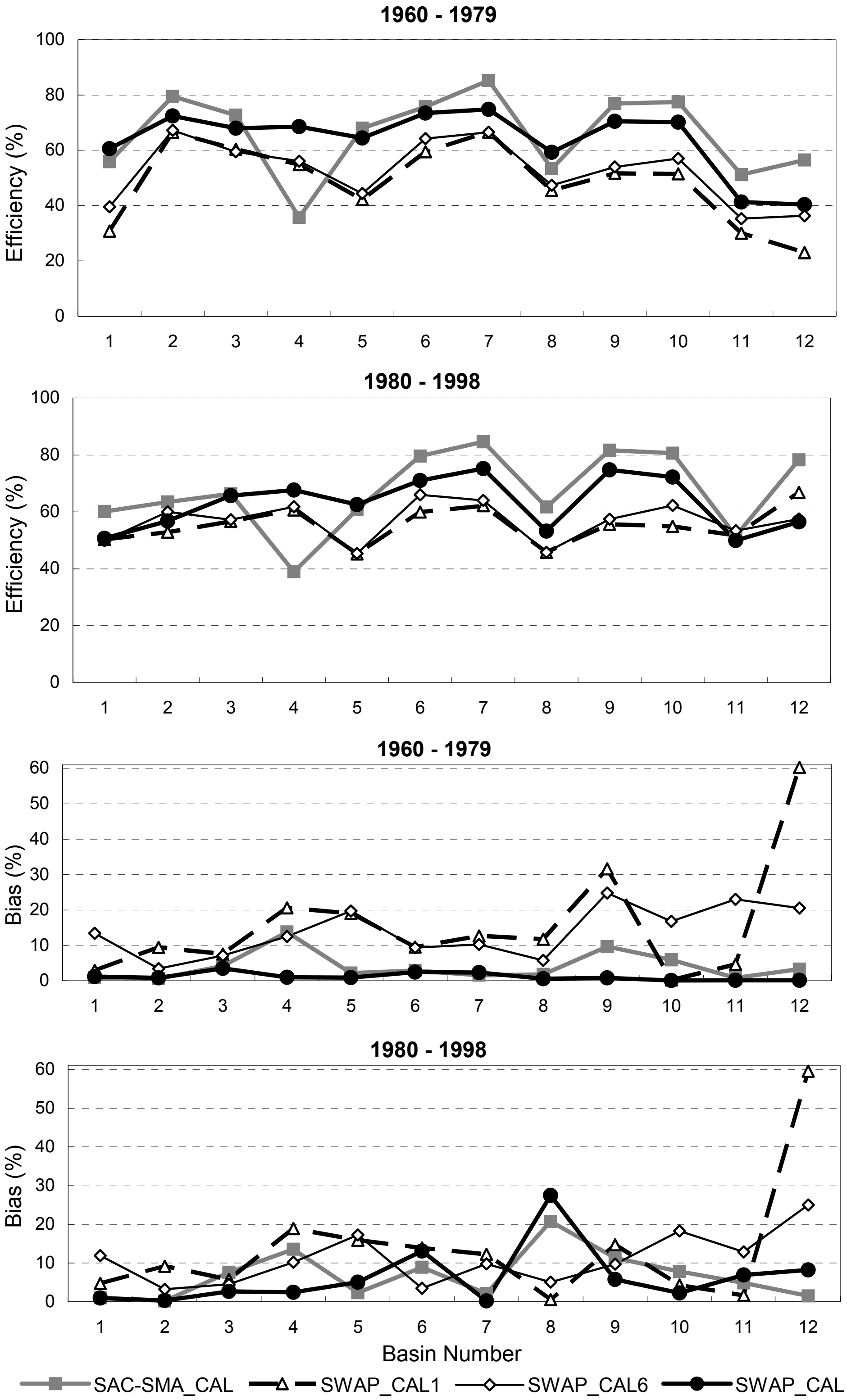

Fig. 2. Daily Nash-Sutcliffe efficiency and bias for each of the 12 MOPEX river basins from calibrated results of SWAP and SAC-SMA models for the calibration and validation periods.

parameterizations by Clapp & Hornberger (1978)) and vegetation parameters (leaf area index LAI, snow-free albedo α_{veg}, zero plane displacement height d_0, roughness length z_0), effective Manning roughness coefficient n and scaling factors for forcing data (rainfall k_{rain}, incoming longwave k_{lw} and shortwave k_{sw} radiation). The total number of calibrated parameters is 16. In the regions where snow processes take place, two more parameters are used: albedo of vegetation with intercepted snow on tree crowns $\alpha_{veg,sn}$ and a scaling factor for snowfall k_{snow}. Since some vegetation parameters (LAI, α_{veg}, d_0, z_0) represent monthly values, one scaling factor for each parameter was calibrated and applied for monthly values. The objective function used in calibration was the same as in the case of six parameters, but bias between simulated and modelled daily streamflow was also taken into account when selecting the optimal set of parameters.

The results of streamflow simulation using the new optimized set of model parameters are shown in Fig. 1 as SWAP_CAL. The new SWAP results are close to those of SAC-SMA. The median *Eff* for simulated daily streamflow is 70% for SAC-SMA and 68% for SWAP for the calibration period and 65% for SAC-SMA and 64% for SWAP for the validation period. The median *Bias* is 2.6% for SAC-SMA and 0.9% for SWAP during the calibration period, and 6.3% for SAC-SMA and 3.9% for SWAP during the validation period.

Figure 2. shows the values of *Eff* and *Bias* for each of the 12 MOPEX basins and for each of the SWAP model calibration techniques compared to SAC-SMA. As can be seen, the last calibration of SWAP resulted in much better performance of SWAP than previous calibrations. For several basins, SWAP managed to out-perform SAC-SMA in terms of *Eff*. As to the bias, SWAP out-performed SAC-SMA for most of basins. It should be noted that the *Eff* of SWAP will be even higher if we do not minimize the absolute value of bias. Further improvement of model calibration may lead to further increase in the accuracy of streamflow simulation.

CONCLUSIONS

1. The model performance in runoff modelling is strongly dependent on the skill of calibration (the technique of calibration, the objective function used, the choice of calibrated parameters). Model calibration appears more important than model type and complexity, and it may be even more important if input data quality is low.
2. It is reasonable to expect that the land surface model SWAP can simulate runoff under appropriate calibration with an accuracy comparable to that of hydrological models.

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

REFERENCES

Boone, A., Habets, F., Noilhan, J., Clark, D., Dirmeyer, P., Fox, S., Gusev, Y., Haddeland, I., Koster, R., Lohmann, D., Mahanama, S., Mitchell, K., Nasonova, O., Niu, G.-Y., Pitman, A., Polcher, J., Shmakin, A. B., Tanaka, K., van den Hurk, B., Verant, S., Verseghy, D., Viterbo, P. & Yang, Z.-L. (2004) The Rhone-aggregation land surface scheme intercomparison project: an overview. *J. Climate* **17**, 187–208.

Burnash, R. J. C., Ferral, R. L. & McGuire, R. A. (1973) A generalized streamflow simulation system – conceptual modeling for digital computers. US Department of Commerce, National Weather Service and State of California, Department of Water Resources, California, USA.

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

Duan, Q., Schaake, J., Andréassian, 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.

Gan, T. Y. & Burges, S. J. (2006) Assessment of soil-based and calibrated parameters of the Sacramento model and parameter transferability. *J. Hydrol.* **320**(1–2), 117–131.

Gan, T. Y., Gusev, Y. M., & Burges, S. J & Nasonova, O. N. (2007) Performance comparison of a complex, physics-based land surface model and a conceptual, lumped-parameter, hydrologic model at the basin-scale. In: *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX* (ed. by V. Andreassian, A. Hall, N. Chahinian & J. Schaake), 196–207. IAHS Publ. 307. IAHS Press, Wallingford, UK.

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

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**(1–4), 162–191.

Gusev, Y. M. & Nasonova, O. N. (2007) Simulating runoff from MOPEX experimental river basins using the land surface model SWAP and different parameter estimation techniques. In: *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX* (ed. by V. Andréassian, A. Hall, N. Chahinian & J. Schaake), 188–195. IAHS Publ. 307, IAHS Press, Wallingford, UK.

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